The Cost of Air Pollution for Workers and Firms

Marion Leroutier[*], Hélène Ollivier[†]

February 15, 2023

Abstract

Poor air quality negatively affects workers' health and cognitive functions, but we know little about the countrywide consequences for firms. In this paper, we estimate the causal effects of fine particulate matter ($PM_{2.5}$) exposure on workers' absenteeism and firms' monthly sales using unique employer-employee data and granular measures of air pollution in France from 2009 to 2015. We exploit variation in air pollution induced by changes in monthly wind directions at the postcode level. We find that a 10% increase in monthly $PM_{2.5}$ exposure increases worker absenteeism in the same month by 1% and reduces sales in manufacturing, construction, and professional services, with different lags. Sales losses are several orders of magnitude larger than what we would expect if workers' absenteeism was the only factor affecting firms' performance. This suggests a potentially large effect of pollution on the productivity of non-absent workers. We estimate that reducing air pollution in France in line with the World Health Organization's guidelines would have avoided sales losses worth around 6 billion euros every year (0.3% of the French GDP).

Keywords: Cost of air pollution, Absenteeism, Firm performance
JEL codes: Q53, I1, J22

[*]Stockholm School of Economics, Misum. Email: marion.leroutier@hhs.se.
[†]Paris School of Economics, CNRS. Email: helene.ollivier@psemail.eu.
We are grateful to Tatyana Deryugina, Geoffrey Barrows, Abhijeet Singh, Antoine Dechezleprêtre, Maiting Zhuang, Jonathan Lehne, Paul Dutronc-Postel, Sampreet Goraya and Martina Bjorkman-Nyqvist for valuable feedback, as well as attendees at the 9th IZA Workshop on Environment, Health and Labor Markets, the 13th Toulouse Conference on the Economics of Energy and Climate and the 27th EAERE annual conference, and seminar participants at SSE, RWI Berlin, and OARES. We thank Augustin Colette and Elsa Real from INERIS for sharing the air pollution data for France, Camille Regaert from IRDES for answering our questions on the Hygie data, and Tom Verrier for outstanding research assistance. Authorization to use the Hygie data was granted from CNIL in September 2019 (decision DR-2019-286). The results contain modified Copernicus Climate Change Service information from 2022. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. Funding from the Mistra foundation and from a subsidy managed by the French National Research Agency under the ANR-17-EURE-001 programme are gratefully acknowledged.

1 Introduction

It is widely acknowledged that air pollution has detrimental effects on human health.[1] Air pollution exposure causes higher emergency admissions and mortality (Schlenker and Walker, 2016; Deryugina et al., 2019), higher medical expenditures (Barwick et al., 2018), and a higher number of work loss days (Holub et al., 2021). Cognitive functions and intellectual performance may also be impaired (Aguilar-Gomez et al., 2022). These large health costs directly affect the utility of many individuals and are sufficient to justify public intervention. Yet, there might be even wider economic implications to air pollution shocks if workers' productivity decreases as a result of pollution-induced health effects. While several papers have examined how pollution affects workers and firms in specific settings, often based on a handful of production sites, there is limited evidence on the economy-wide costs of air pollution for firms.

In this paper, we assess quantitatively to what extent air pollution health costs on workers translate into economic costs for their employers. We estimate countrywide effects of air pollution on workers' sick leave and on firms' sales using confidential employer-employee data from France. We focus on exposure to fine particulate matter pollution ($PM_{2.5}$) since it can penetrate deep into the respiratory tract and enter the brain, with particularly detrimental health effects.[2] It can also easily penetrate indoors, thus affecting most workers while at work. We assemble a unique dataset which combines detailed data on sick leave episodes for a representative sample of 400,000 French private sector employees with monthly sales data of the firms that employ them, as well as granular measures of air pollution and weather conditions at their workplace between 2009 and 2015. Two key challenges with identifying the causal effects of pollution exposure on countrywide work absenteeism and firms' sales is that air pollution is often a co-product of production, and individual exposure to pollution exposure is always measured with noise.[3] To circumvent these challenges, our analysis leverages variation in air pollution induced by changes in monthly wind directions at the postcode level.

The identifying assumption of our instrumental variable (IV) approach is that, after flexibly controlling for postcode and month-by-year fixed effects and other weather variables, changes in a postcode's monthly wind direction are unrelated to changes in the postcode's work absenteeism except through their influence on air pollution. The benefit of our approach is that it does not

[1] Exposure to fine particulate matter ($PM_{2.5}$), for instance, is associated with approximately 4.2 million premature deaths every year globally (WHO, 2014a). Even in Europe, where air pollution has been regulated for several decades, an annual 307,000 premature deaths are attributed to $PM_{2.5}$ pollution (European Environment Agency, 2020).

[2] The 2.5 subscript in $PM_{2.5}$ means that these particles have a size lower than 2.5 μm.

[3] In an ideal setting, pollution exposure would be measured by multiplying pollution levels from each location where an individual spend some time by the number of hours spent in each location. In this paper, we proxy pollution exposure by pollution levels measured at the postcode of the workplace, where workers spend most of their waking hours.

require identifying the sources of pollution in each area. Instead, we allow wind directions to influence pollution differently across 100km-by-100km grid cell areas covering all metropolitan France. Our analysis thus employs a similar strategy to Deryugina et al. (2019) and Anderson (2020), except that Deryugina et al. (2019)'s instrument is made of binary variables indicating the dominant wind direction at the daily level whereas we use continuous variables measuring the share of hours in each month where the wind blows from each direction (North, East, and West, with South winds being the excluded category).[4] We further address the additional challenge coming from multi-establishments firms when estimating the effects of air pollution on monthly sales. Many firms own several establishments that are not located in the same postcode, rendering the previous IV approach inapplicable. We therefore exploit another IV approach by computing a weighted average of predicted pollution exposure at the firm-month level.

Our study has three main results. First, we estimate that a 1 microgram per cubic meter (µg/m^3) increase in monthly $PM_{2.5}$ exposure causes 0.15 additional sickness leave episodes per 1,000 workers within the month of exposure, which corresponds to a 0.6 percent increase relative to the mean. It also increases the number of sick days by 2.7 per 1,000 workers and the associated sickness leave spending by €86.5 per 1,000 workers. Our estimate implies that a 10 percent increase in monthly $PM_{2.5}$ increases by 1 percent the number of sickness leaves per 1,000 workers in the same month. Our results are robust to various specifications and placebo tests, including studying the relationship between pollution and work absenteeism at the weekly level or at the individual level and excluding months with air quality alerts. We also find that the effect of pollution on work absenteeism is larger than average for relatively low-wage workers, and varies in magnitude across sectors of activity.

Second, we find that firm-level $PM_{2.5}$ exposure has heterogeneous effects on sales depending on the economic sector. We estimate that a 1 µg/m^3 increase in firm-level pollution exposure in month t decreases manufacturing sales by 0.24 percent and construction sales by 0.44 percent in the following month. This is consistent with the existence of a lag between the timing of production – when workers are directly affected by air pollution – and the timing of recorded sales for these sectors. By contrast, we detect a decrease in sales by 0.55 percent in the contemporaneous month of exposure in the professional services sector.[5] These results also hold for the subsample of single-establishment firms, for which pollution exposure is measured in only one location. We perform a range of robustness checks, including discarding 2009 – when France was hit by the Great

[4]We use wind directions as instruments for air pollution rather than infrequently occurring events, such as thermal inversions, because Bagilet and Zabrocki (2022) show that an IV strategy with low frequency events as instruments may lead to inflated estimates due to low statistical power when estimating acute health effects.

[5]Professional services refer to business-to-business specialized services and include two types of firms: about two-third of the firms provide high-skilled services such as legal and financial advice, engineering and consulting; about one-third of the firms provide low-skilled services such as cleaning, security, or are temporary work agencies.

Recession, discarding months with air quality alerts, including ozone, and two-way clustering standard errors, which corroborate our results for the construction and professional services sector whereas the effects for manufacturing are less precisely estimated. We fail to detect a significant effect of pollution on sales in other services sectors (including retail and restaurants, business-to-consumer services and other business-to-business services).

We present a conceptual framework that describes how workers and firms can be impacted by pollution shocks. The framework highlights three channels through which air pollution can influence sales: i) a decrease in labor supply, as measured by an increase in work absenteeism in our context; ii) a decrease in productivity among non-absent workers, which can result from non-absent workers suffering from mild health symptoms or reduced cognitive capacities, as well as from the potential disruptions in production value chains when their co-workers take sick leaves; and iii) a decrease in demand.[6] Finding a negative effect of air pollution only for firms in the manufacturing, construction and professional services sectors indicate the potentially detrimental cumulative effects of ambient pollution and work emissions in some sectors, the role of complementarities among workers in production value chains, and sectoral differences in the ability of firms' managers to temporarily replace sick workers. In professional services, where a fraction of the workers is high-skilled, non-absent workers may mostly suffer from the negative effects of pollution on their cognitive abilities.

With our previous estimates, we can evaluate the cost of pollution-induced work loss days valued at the marginal product of labor, a proxy for the sales loss imputable to the first channel. This cost is several orders of magnitude smaller than the total sales loss in manufacturing, construction and professional services sectors. Thus, the productivity and demand channels must play an important role in reducing firms' sales. While we cannot rule them out entirely, we argue that demand responses to pollution shocks should be limited in our context because awareness of air pollution levels is low in our study period and we fail to detect a significant decrease in sales in business-to-consumers services.[7] Our results thus reflect the importance of the productivity channel among non-absent workers.

Third, we quantify the benefits associated with reducing pollution levels so as to meet the World Health Organization (WHO)'s recommendations of not exceeding 15 µg/m^3 for daily $PM_{2.5}$ exposure, a threshold exceeded 37% of the days in our data. Based on our estimates, bringing pollution levels down to 15 µg/m^3 every day would have avoided annually 2 million of sick days,

[6] In our study, we define the labor supply response in terms of whether or not a worker calls sick to work, which, in France, requires a medical certificate signed by a general practitioner. If a worker chooses to go to work while shortening his number of hours per day, we consider that the channel is lower productivity.

[7] There is no air quality alert system in place for $PM_{2.5}$ in France. Even in the most polluted city of France, Paris, air pollution alerts for PM_{10} – which involve recommendations from the health authorities targeting the most vulnerable individuals – were issued on 4% of the days in our study period. More severe alerts involving restrictions in car traffic were issued on 0.7% of the days only.

corresponding to 1% of total sickness leave spending, and between €6 billion and €11 billion of foregone sales (between 0.3 and 0.5% of French GDP) on average between 2009 and 2015. The lower bound excludes manufacturing sales losses, for which the point estimate is less precise, and the upper bound includes them. While there is no readily available estimate for the cost of meeting the WHO threshold, we compare our estimated benefits with existing estimates for a scenario reducing particulate emissions by 16% (lower bound) and for a scenario reducing them by 33% (upper bound) and evaluate that they represent between 80% and 150% of the cost. While we only consider benefits from short-term reductions in exposure and ignore population-wide morbidity and mortality benefits, our results already provide evidence of large benefits from tightening regulatory standards, even at the firm level.

To the best of our knowledge, this paper provides the first countrywide estimates of the effect of air pollution on firms' performance and their workers in a high-income country. Previous literature has examined how pollution affects workers, in terms of productivity (Graff Zivin and Neidell, 2012; Chang et al., 2016; Lichter et al., 2017; Meyer and Pagel, 2017; He et al., 2019; Chang et al., 2019; Adhvaryu et al., 2022) and decision-making (Meyer and Pagel, 2017; Dong et al., 2019), based on specific settings where workers are paid by the hour or productivity is easy to observe. Other studies have examined pollution effects on labour supply (Aragón et al., 2017) or firms' performance (Fu et al., 2021) using representative data at the country level in the context of developing countries. We expect pollution to affect workers' health, labor supply and productivity differently in high-income countries, where the levels and saliency of pollution are lower, the sectoral composition of the economy is different, and workers often benefit from institutionalized sickness leave. Average pollution levels in France are similar to those in Europe and slightly above those in the US.[8] Thus, we expect our results to have external validity for other high-income countries.

A closely related study by Holub et al. (2021) estimates the effects of PM_{10} on sickness leaves in Spain and derive an economic cost associated with pollution-induced work loss days by multiplying the number of work loss days with workers' daily wage. We differ from this study in the type of pollutant, in the choice of instrument – adopting wind directions instead of episodes of Sahara wind –, but more importantly in the combination of employer-employee data that allows us to estimate the cost of pollution in terms of foregone sales, which we find to be much larger than the cost related to work absenteeism only. Furthermore, another related study by Borgschulte et al. (2022) finds that air pollution shocks induced by wild-fire smoke in the US reduce per capita earnings in the medium run, using county-level data. Combining worker- and firm-level data, our

[8] In 2015, population-weighted $PM_{2.5}$ exposure was 13 µg/m^3 in France, 8 µg/m^3 in the US, 11 µg/m^{3}in Spain and the UK, 13 µg/m^3 in Germany, and 17 µg/m^3 in Italy. Source: https://www.who.int/data/gho/data/themes/air-pollution/modelled-exposure-of-pm-air-pollution-exposure.

study sheds light on two possible mechanisms at play: first, workers may not be fully (or at all) compensated for the income loss associated with taking a sickness leave; second, lower sales in some sectors may reduce the demand for labor.

Our paper contributes to the literature that evaluates the cost of air pollution based on micro-level data. This literature has focused almost exclusively on the health costs to individuals (Deryugina et al., 2019; Barwick et al., 2018; Mink, 2022). For instance, Mink (2022) estimates that reducing nitrogen dioxide concentrations by 27% would save at least €5.2 billion in healthcare costs in France. Our results suggest that considering only the health effects from pollution while ignoring its effects on firms may widely underestimate the economic costs associated with pollution shocks. This corroborates the findings in Dechezleprêtre et al. (2019) that a 10% increase in annual $PM_{2.5}$ decreases real GDP by 0.8%, based on GDP data from European regions. By using micro data, our study highlights the heterogeneous effects across economic sectors and provide suggestive evidence of the channels underlying the output loss.

More broadly, our paper adds to the literature on the effects of environmental shocks on workers and firms. Previous literature has examined the effects of high temperatures on workers' productivity (Somanathan et al., 2021), labor supply (Graff Zivin and Neidell, 2014), firms' sales (Addoum et al., 2020), and work accidents (Park et al., 2021) in the context of climate change. In contrast to our findings using pollution shocks, Addoum et al. (2020) fails to detect a significant impact of temperature shocks on US firms' economic performance. This difference may arise from the fact that temperature shocks are salient, which may trigger a private adaptation response which consists in using air conditioning at the workplace, whereas air pollution shocks are largely unnoticed in our context and adaptation is limited.

The rest of the paper is organized as follows. Section 2 provides a brief background on the effect of fine particulate matter on health and presents a conceptual framework that encompasses the potential channels through which pollution can affect firms' sales. Section 3 describes our data. Section 4 describes our empirical strategy. Section 5 presents the results, and section 6 some robustness checks. Section 7 discusses the magnitudes and implications of the effects on workers' absenteeism and firms'sales, and section 8 concludes.

2 Background and Conceptual Framework

2.1 Effects of Particulate Matter on Health and Productivity in the French Context

Particulate matter with a diameter below 2.5 micrometers ($PM_{2.5}$) enters the lungs and can pass into the bloodstream, resulting in significant health problems such as increased mortality and cardiovascular diseases (World Health Organization, 2016; European Environment Agency, 2020).[9] A large literature has shown the negative effects of short- and long-term exposure to $PM_{2.5}$ on human health, even at low levels of exposure. For instance, Deryugina et al. (2019) found that, in the US, a 1 µg/m^3 increase in $PM_{2.5}$ exposure for one day causes 0.69 additional deaths per million elderly individuals over the three following days. $PM_{2.5}$ also readily penetrates indoors (Chang et al., 2016; Krebs et al., 2021), thereby being likely to affect individuals in their working environment. Exposure to fine particulate matter can temporarily affect cognitive functions: mounting toxicological evidence suggests that it can enter the brain and increase neuro-inflammation and oxidative stress in the central nervous system. Furthermore, $PM_{2.5}$ can travel far (hundreds of kilometres) and remain in the atmosphere for a long period of time (US EPA, 2018).

As in many high-income countries, air quality in France is regulated via command and control taking the form of maximum concentration thresholds defined at the European level. Depending on the pollutant, the thresholds are defined at the annual and/or 24-hour level.[10] For $PM_{2.5}$, the annual threshold is 25 µg/m^3 and there is no threshold for daily exposure. By contrast, the recent recommendations from the WHO set the threshold at 15 µg/m^3 for *daily* exposure for $PM_{2.5}$.[11]

The recent literature has put forward different ways through which air pollution can affect workers' productivity and firms' performance. In the context of developing countries or in settings where workers are paid by the hour, several studies find that pollution reduces workers' productivity primarily through a decrease in output per hour (Graff Zivin and Neidell, 2012; Chang et al., 2016; Adhvaryu et al., 2022; Chang et al., 2019; He et al., 2019). Other papers find that air pollution reduces labor supply, both in the short run (Hanna and Oliva, 2015; Aragón et al., 2017; Holub et al., 2021) or in the medium run (Borgschulte et al., 2022). By reducing non-absent workers' productivity or by reducing labor supply, air pollution will likely also reduce firms' output and sales. Fu et al. (2021) shows that air pollution decreases annual firm-level productivity for a large representative sample of Chinese manufacturing firms. However, in few cases where firms' response could be explored, studies find that firms can dampen the productivity loss from their most affected employees by reallocating tasks among all employees (Adhvaryu et al., 2022), by hiring new employees (Fu et al., 2021), or by asking unaffected workers to work longer hours.

[9] $PM_{2.5}$ is related to other air pollutants. In particular, it is by definition included in PM_{10} concentration levels, but it is deadlier because smaller-sized particles penetrate deeper into the respiratory system. $PM_{2.5}$ can be either directly emitted as "primary" particles, for which the main contributors are the residential and tertiary sector (52%), transportation (20%), manufacturing (18%) and agriculture (11%) (CITEPA, 2021) or formed in the atmosphere as "secondary" particles from the chemical reactions of gaseous pollutants, including SO_2 and NO_2.

[10] The legal thresholds are defined in European Union legislation and transposed into French law. The French government must comply with these thresholds or risks incurring sanctions: in 2020, France has been referred to the Court of Justice of the European Union for exceeding the daily thresholds for particulate matter PM_{10} (European Commission, 2020).

[11] See the 2021 recommendations from the World Health Organization at https://apps.who.int/iris/handle/10665/345329.

We expect that labour market institutions, industry, and the saliency of pollution shocks influence how workers and firms respond to these shocks. In particular, workers may benefit from different levels of job protection across countries, sectors, and firms, which will lead to different abilities to take sick leaves when being ill. In France, private sector employees are entitled to relatively generous sickness allowances, under some conditions. Sickness allowances consist of three parts, all conditioned on providing a medical certificate and having worked at least 150 hours in the past three months. First, workers receive publicly funded benefits from the fourth day of a sickness leave episode (hereafter, SLE), which amount to roughly 50% of their gross daily wage, with a cap of 1.8 times the daily equivalent of the minimum wage (i.e., €43 per day in 2015). Second, they receive an allowance from mandatory employer-funded funds from the eighth day of leave, which amounts to 40% of their gross daily wage initially and then decreases to 16% after 30 to 90 days, and is paid for a maximum of 60 to 180 days, depending on the workers' seniority in the firm. Third, they receive an optional employer-funded allowance that is negotiated in collective agreements and generally covers the difference between the gross daily wage and the publicly-funded plus mandatory employer-funded benefits. According to survey evidence (Pollak, 2015), two-thirds of private sector employees receive this optional allowance and are granted a 100 percent replacement rate from the first day of leave.

While developing countries such as India or China face very high air pollution levels that may render pollution shocks more visible to managers and firms, the moderate levels of pollution in high-income countries such as France have ambiguous effects on the severity of economic consequences for firms. On the one hand, few workers may suffer severe health symptoms or impaired productivity under such moderate pollution levels. This would suggest that firms would experience small decline in their productivity and output. On the other hand, pollution shocks being less salient implies that managers are less able to respond appropriately to mitigate the reduction in output.

2.2 Conceptual Model

We illustrate how pollution shocks might affect workers and firms in a stylized model that connects workers' exposure to air pollutants with firms' productivity and sales. We model the production function (in logs) for firm f in industry i at time t as:

$$q_{fit} = F_{it}(k_{fit}, l_{fit}) + \omega_{fit}, \tag{1}$$

where q denotes the quantity of output; k and l denote capital and labor, respectively, and ω is a persistent Hicks-neutral productivity shock that is known to the firm when making its period t decisions. Capital is assumed to be chosen a period ahead in $t-1$ whereas labor is chosen flexibly in period t. The observed quantity of output is given by:

$$y_{fit} = q_{fit} + \mu Z_{fit} + \epsilon_{fit}, \tag{2}$$

where the discrepancy between the originally planned output quantity q and the actual final quantity y can arise from either pollution-induced productivity shocks μZ_{fit} or from other ex-post shocks to output ϵ_{fit} capturing measurement error or demand shocks. Z denotes the concentration of air pollution measured at the firm location. Both ex-post shocks are not known to the firm when making its period t decisions and cannot be anticipated. Assuming that both labor productivity and labor supply are affected by air pollution and that consumers may potentially respond to pollution shocks by staying home and reducing their consumption, pollution-induced ex-post shocks can be decomposed in the following way:

$$\mu Z_{fit} = \frac{\partial F_{it}}{\partial l_{fit}} \frac{\partial l_{fit}}{\partial Z_{fit}}\bigg|_{l_{fit}=l^*_{fit}} + \frac{\partial F_{it}}{\partial \sigma^L_{fit}} \frac{\partial \sigma^L_{fit}}{\partial Z_{fit}}\bigg|_{l_{fit}=\tilde{l}_{fit}} + \zeta_{fit}, \tag{3}$$

where the first term in the right hand side of equation (3) captures the loss in output associated with reduced labor force if observed labor $\tilde{l}_{fit}$ is lower than originally hired labor l^*_{fit} assuming that the marginal product of labor $\sigma^L_{fit} \equiv \frac{\partial F_{it}}{\partial l_{fit}}$ remains unaffected; the second term captures the loss in output associated with reduced marginal product of labor given labor force $\tilde{l}_{fit}$, and the third term captures pollution-induced ex-post demand shocks.

We assume an exogenous first-order Markov process for persistent productivity shock: $\omega_{fit} = g_0 + g_1\omega_{fit-1} + \eta_{fit}$, where η_{fit} denotes an innovation to productivity that is by assumption uncorrelated with any lagged choice variables of the firm. Combining (2) with the Markov process for ω yields the following equation:

$$y_{fit} = F_{it}(k_{fit}, l_{fit}) + \mu Z_{fit} + g_0 + g_1\omega_{fit-1} + \xi_{fit}, \tag{4}$$

where the error term ξ_{fit} combines both ϵ_{fit} and η_{fit}. Assuming orthogonality between pollution shocks and the original input choices of the firm and ω_{fit-1}, we can directly estimate the firm-level pollution-induced ex-post shock by regressing observed sales on pollution levels and controls.[12]

The focus of this paper is twofold. First, we assess the workers' response to pollution shocks through the sickness-induced labor supply channel, $\frac{\partial l_{fit}}{\partial Z_{fit}}\Big|_{l_{fit}=l^*_{fit}}$. We explore this channel by focusing on missed work days due to sickness. Second, we directly estimate μZ_{fit}, the total response of sales to air pollution shocks.

[12] Anticipating our empirical analysis, we abstract from any price response to pollution shocks either because of perfect competition or because we expect that firms do not adjust their prices at the monthly level.

3 Data

We combine nationwide gridded reanalysis pollution and weather data, a representative panel dataset of French private sector employees affiliated to France's universal sickness-leave insurance, and value added tax records for the universe of French firms, over the period spanning 2009 to 2015. This section describes the data sources, cleaning steps and key variables used in the analysis.

3.1 Pollution Data

We use air pollution data from the French National Institute for Industrial Environment and Risks (INERIS), which provides gridded reanalysis historical pollution data for metropolitan France (Real et al., 2021). The dataset combines background measurements of air quality from monitoring stations with modelling from the chemistry-transport model CHIMERE, using a kriging method (Real et al., 2021). It contains hourly concentrations of $PM_{2.5}$, PM_{10}, NO_2, and O_3 with a spatial resolution of approximately 4 km x 4 km for the period 2000-2018. We refer to the grid scale of this data as delimited by "Chimere grid cells". We aggregate pollution data at the monthly level, focusing on $PM_{2.5}$, for the 33,252 Chimere grid cells located in metropolitan France. We limit the analysis to the 2009-2015 period, the coverage of the absenteeism data.

Gridded reanalysis pollution data are better suited to capture the average pollution exposure of local residents than pollution-monitor readings. Indeed, monitors are sparse and sometimes strategically placed to capture locally produced emissions (e.g., from a highway).[13] As a result, monitor readings may not be informative of the average pollution exposure in a given grid cell. By contrast, reanalysis data combine these monitor readings with a chemistry-transport model that takes account of all sources of pollution to give a measure of average exposure.

The average $PM_{2.5}$ exposure of French workers, based on the postcode of their workplace, is 15.4 µg/m^3 over the study period. Figure B.4 shows the spatial distribution of annual exposure at different points in time whereas Figure 1 shows the distribution of average monthly exposure. Pollution has been decreasing over the period, a trend also observed in the US and in the rest of Europe (Champalaune, 2020; Currie et al., 2020; Sicard et al., 2021). While the annual EU standard of 25 µg/m^3 is rarely exceeded, the WHO recommended threshold of 15 µg/m^3 for daily exposure is exceeded for 37% of the sample of worker-days.

[13] The network of $PM_{2.5}$ background monitoring stations is particularly sparse in France. Over the study period, there are between 62 and 105 stations for this pollutant, to be compared with between 173 and 251 for PM_{10}, between 318 and 385 for ozone, and between 282 and 337 for NO_2. The reanalysis data take into account the correlation between $PM_{2.5}$ and PM_{10} using a co-kriging method to exploit the higher density of PM_{10} monitoring stations for estimating $PM_{2.5}$.

3.2 Weather Data

We use gridded reanalysis weather data derived from the ERA5 dataset of the Copernicus Climate Change Service (C3S).[14] We obtain hourly precipitations, surface temperatures, wind direction, and wind speed at the 0.25° x 0.25° resolution, which corresponds to grid cells of around 28 km x 28 km. We aggregate weather data at the monthly level for the 1,156 Copernicus grid cells in metropolitan France. We take the monthly averages of daily maximum temperatures and hourly wind speeds, and we sum hourly precipitation over each month. For wind direction, we compute for each month the share of hours when the wind blows from each of four directions: North (below 45° or above 315°), East (between 45° and 135°), South (between 135° and 225°) and West (between 225° and 315°).

3.3 Sickness Leave Episodes

We obtain data on sickness leave episodes (SLE) from the Hygie dataset, which is a representative sample of private sector employees born between 1935 and 1989 and affiliated to France's universal sickness-leave insurance.[15] For each worker, we know the exact start date and duration of each SLE that occurred during the period 2009-2015, the associated state-funded sickness benefits, and characteristics such as age, gender, annual wage, contract type, and annual medical expenditures. We restrict our dataset to employees that we can match to their exact workplace via an establishment-level identifier denoted SIRET used by all French administrations (see Appendix A for more details). This restriction allows us to precisely geolocate the workplace so as to allocate pollution exposure and to link employees information from the Hygie panel to other establishment- and firm-level datasets.[16] We have three primary measures of absenteeism: an indicator for an individual starting a SLE in a given month, a count of sick days associated with SLEs that started in a given month, and the total sickness leave spending associated with SLEs that started in a given month. In the main analysis, we only consider SLEs lasting less than three months, which represent 93% of SLEs.[17]

[14] We acknowledge using the ERA5 dataset (Hersbach et al., 2018) downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. See https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

[15] The Hygie dataset combines administrative data on health from the organization managing the public health insurance (CNAM) with administrative data on employees' careers from the organization managing the public pension system (CNAV).

[16] French administrations have created a unique identifier for each establishment composed of 14 digits, of which the first 9 digits correspond to the unique identifier of the firm that owns the establishment (denoted SIREN).

[17] In our data, the average sickness leave episode lasts 29 days whereas the median duration is only 9 days. Figure A.2 shows the small proportion of SLEs that last more than 3 months and their strong influence on the average number of sick days. We therefore focus on SLEs lasting less than 3 months to avoid that our results for the number of sick days are driven by long-term illnesses.

The resulting workers' sample is a monthly unbalanced panel of around 450,000 individuals working in roughly 430,000 private sector establishments over the period 2009-2015. Employees are 55 percent male and 40 years old on average; they earn an average annual gross wage of €25,910, and 75 percent of them are full time employed (see Table A.1). On average, roughly 22 workers per 1,000 start a sickness leave episode in a given month. Comparing our data with the exhaustive matched employer-employee data called *DADS-Postes*, we validate the representativeness of our sample of private sector employees.

We geolocate workers based on their establishment's postcode. We link this dataset to the corresponding pollution and weather information by identifying the Chimere and Copernicus grid cells which includes each postcode's centroid. As a result, our measurement of pollution exposure rests on exposure at the workplace.[18] To overcome limited computational power associated with our access to confidential datasets, we aggregate the data and run the analysis at the postcode level (with roughly 6,000 postcodes in France).[19] In a robustness test, we take advantage of the individual panel dimension and estimate a model with individual-fixed effects on a 10% subsample. Table 1 shows descriptive statistics at the postcode level.

3.4 Firm-Level Data

We observe monthly sales for almost the universe of French firms from monthly VAT records. We restrict the sample to firms that declare their VAT every single month in a year and employ workers identified in the Hygie dataset (see Appendix A for details).[20] We define six sectors of activity based on the sectoral classification available at the establishment level: manufacturing, construction, retail and restaurants, other business-to-consumer services, professional services, and Information and Communications Technologies (ICT) and other business-to-business services. The final sample includes 182,929 firms. Table 1 shows that firms employ on average 65 workers. 17% of firms belong to the manufacturing sector, 14% to the construction sector, 21% to retail and restaurants, 11% to other business-to-consumer services, 16% to professional services and 21% to ICT and other business to business services. Average monthly sales amount to €1,311,465 whereas median monthly sales amount to €141,676.

[18] Individual exposure depends on the location of residence, the location of work, transportation between the two, as well as the location of leisure activities. Based on the 2015 population census, we note that 27% of employees actually live and work in the same postcode. Additionally, the median commuting distance was only 9.2 kilometres in 2017 (INSEE, 2021). Comparing the distributions of pollution exposure at the workplace and at the place of residence for the population of French workers using exhaustive matched employer-employee data (*DADS-Postes*), we find that the two distributions almost overlap.

[19] Access to the data is obtained through the CASD (Secure Data Access Center), which provides the service of making confidential data sets available to researchers using a secured server, for which there are constraints on the size devoted to each project, thereby computational constraints.

[20] Firms with monthly VAT declarations represent 66% of French firms, but 91% of total sales (France Stratégie and Inspection générale des Finances, 2021).

To link firm-level sales data to pollution and weather data, we address a challenge regarding multi-establishment firms. Indeed, 37% of the firms – representing 75% of total sales– own more than one establishment in the VAT sample of 2015. Whereas we simply attribute pollution and weather exposure to single-establishment firms based on their postcode as before, for multi-establishment firms we build a measure of weighted exposure to pollution and weather characteristics, where the weights are the annual number of workers in each plant owned by the firm. To compute these weights, we exploit exhaustive matched employer-employee data called *DADS-Postes*, which records the number and location of establishments owned by a firm and the number of workers in each of these establishments.

4 Empirical Strategy

4.1 Air Pollution Effects on Sickness Leaves

Fixed-effects specification. We model the relationship between short-run exposure to fine particulate matter and workers' sickness leave outcomes with the following equation:

$$Y_{g,t} = \alpha + \beta PM_{2.5g,t} + W_{g,t}\gamma + h_{g\subset d,t}\delta + \nu_g + \theta_t + \epsilon_{g,t}, \tag{5}$$

where the dependent variable $Y_{g,t}$ is the sickness leave outcome measured in month-of-sample t in postcode g. $PM_{2.5g,t}$ measure monthly pollution exposure at the postcode level, and the parameter of interest is β. Postcode fixed effects ν_g isolate monthly variation in pollution exposure at the postcode level, absorbing any time-invariant area-specific characteristic at a fine geographical scale. In a robustness test, we examine the sensitivity of our results to the inclusion of postcode-year fixed effects to isolate monthly variation in pollution exposure within a year assuming that area-specific characteristics may vary over the years. Controlling for month-of-sample fixed effects θ_t captures time-specific shocks influencing both pollution and absenteeism nationwide.

Weather conditions can jointly affect pollution and health. For example, high levels of $PM_{2.5}$ are more frequent in winter when low temperatures can also affect health. For this reason, we generate indicators for monthly averages of daily maximum temperatures, wind speed and precipitation in each postcode.[21] We then generate a set of indicators for all possible interactions of these weather controls and include it in all our regressions as $W_{g,t}$. We also include the count of days in month t associated with school holidays and the monthly average of flu cases per week per 100,000 individuals in the *departement* d where postcode g is located as variable $h_{g\subset d,t}$ (where the

[21] Monthly average of daily maximum temperatures falls into 8 potential bins. The bins span 5°C each, except for the first bin including all negative temperatures, and for the eighth bin including all temperatures above 30°C. For wind speed and precipitation, we compute indicators for each quintile of these variables.

departement is an administrative and jurisdictional unit).[22] Many workers go on vacation during school breaks in France, and they would likely not take a sickness leave if they fall sick during their vacation. At the same time, economic activity and pollution are typically lower during these holiday periods. Flu epidemics spread at the local level and influence both the number of sickness leaves and potentially the economic activity in the region. We weight all estimates by the number of workers present in each postcode in a given year.

Equation (5) assumes a linear relationship between pollution concentration and work absenteeism. We assess the plausibility of this assumption by looking at a residualized binned scatter plot of the two variables. Figure 3 reports the binned scatterplot of the residualized number of SLEs for each bin (of equal size) of residualized monthly average $PM_{2.5}$ concentrations, controlling for the fixed effects and control variables from equation (5). The scatter plot suggests that the effect of $PM_{2.5}$ on sickness leave episodes is approximately linear in the monthly average concentration of particulate matter.

Instrumental variable approach using local changes in wind direction. Despite the use of high-dimensional fixed effects, OLS estimates of equation (5) are prone to bias because exposure to $PM_{2.5}$ is likely to be measured with noise. Indeed, measuring pollution exposure based on the workplace location fails to capture exposure to pollution at the other places (e.g., location of residence and leisure activities) where workers spend time during the month. Assuming that the measurement error is classical – mean zero and i.i.d – , this gives rise to an attenuation bias, which can be exacerbated by the use of fixed effects (Griliches and Hausman, 1986). Another potential source of bias pertains to unobserved local shocks that may influence pollution concentration while also affecting workers' absenteeism (e.g., road work). Finally, there may be a simultaneity bias if high levels of absenteeism lead to lower economic activities and lower commuting, thereby decreasing local pollution levels.

To address these remaining potential biases, we rely on an instrumental variable approach exploiting month-to-month variation in wind direction at the postcode level, in the spirit of Deryugina et al. (2019) and Anderson (2020). We instrument monthly pollution in a postcode g with the share of hours in a month where wind blows from each of three directions (excluding South winds). We allow wind directions to influence pollution differently in different areas. This flexible approach acknowledges that a given wind direction might not affect pollution in the same way in all postcodes in France. In our main specification, we define these areas as 100km-by-100km grid

[22] In metropolitan France, the median size of a *departement* is 5 880 km^2, which is equivalent to 3.5 times the size of a median US country. Beside the July-August school break, which is the same for all schools in France, nationally the two-week school breaks in the Fall, Winter, and Spring are staggered by region. The data on flu cases is publicly available on the following website that records the prevalence of different diseases and epidemics in France: https://www.sentiweb.fr/france/fr/?page=table&maladie=25.

cells (see Figure B.7).[23] Imposing that all postcodes within a 100km-by-100km grid cell see their pollution exposure respond to wind direction in similar ways reduces the influence of pollution variation coming from local sources and rather capture pollution variation coming from nonlocal sources and transported by the wind (Deryugina et al., 2019). In a robustness check, we check that our results are similar with smaller grid cells of 50 km by 50 km.

The specification of our first stage is:

$$PM_{2.5g,t} = \alpha + \sum_{j=2}^{4}\sum_{k=1}^{K} \beta_{jk} \text{WIND}_{j,k,t} \mathbb{1}(g \subset k) + W_{g,t}\gamma + h_{g\subset d,t}\delta + \nu_g + \theta_t + \epsilon_{g,t}. \tag{6}$$

The excluded instruments are $\text{WIND}_{j,k,t}\mathbb{1}(g \subset k)$, where $\text{WIND}_{j,k,t}$ identifies the share of hours in month t where the wind blows from direction j, with $j = 1$ being from the South (omitted category), $j = 2$ from the West, $j = 3$ from the East, and $j = 4$ from the North. The variable $\mathbb{1}(g \subset k)$ is an indicator for postcode g to be included in 100km-by-100km area k, with K being the total number of such areas. The other variables are defined as in equation (5). The parameters of interest are the β_{jk}s. For a given wind direction j and 100km x 100km grid cell area k, β_{jk} captures the effect of a marginal increase in wind blowing from j on pollution in postcode g located in area k (relative to South winds).

For the identification of the β_{jk}s, we rely on within-postcode variation in wind direction, where wind direction is observed at the Copernicus grid cell level. Therefore, we cluster standard errors at the Copernicus grid cell level. Figure B.6 shows that there is substantial within-postcode variation in monthly wind directions for three postcodes located in different parts of France (Paris, Marseille, center of France).

4.2 Air Pollution Effects on Firms' Sales

To examine the total response of sales to variation in air quality, we amend our previous empirical strategy in four ways. First, since we observe sales only at the firm level and not at the establishment level, we calculate firms' exposure to pollution as a weighted average of pollution exposure at each establishment owned by the same firm, where the weights correspond to the relative number of workers in each establishment in each year. Since establishments are not located in the same area, we can no longer rely on wind direction to instrument firms' pollution exposure. Instead, we adopt another instrumental variable approach where we instrument firms' observed pollution exposure by their predicted pollution exposure according to equation (6) averaged across

[23] We use a publicly available gridded dataset splitting metropolitan France into 94 such areas. See https://geo.data.gouv.fr/fr/datasets/d97be74ddacf262d81737f783da1b5264deb0adb. On average, there are 61 postcodes and 14 Copernicus grid cells in each 100km-by-100km area.

establishments.

Second, we use firm-by-year fixed effects instead of postcode fixed effects. This isolates variation in pollution exposure around the mean exposure of a firm at the annual level, thereby absorbing any time-invariant firm characteristics while also controlling for annual shocks jointly affecting exposure to pollution and sales. Such shocks include any productivity shock or any change in the number or location of establishments belonging to a firm, which we only observe at the annual level.

Third, we run our firm-level analysis sector by sector to account for the differing lag between production and sales across sectors. As pollution shocks potentially impair workers' health and productivity, we expect that only shocks occurring at the time of production will affect sales. The lag between production and sales may differ across sectors. For instance, in manufacturing, production may take a few months for some products, and products are sold whenever they meet a demand. In the construction sector, clients sign for a project, pay some percentage of the bill in advance, but only pay the full amount at the end of the project. Similarly, professional and business-to-business services may face similar lags between the time of production that starts after signing a contract and the payment by the client at the end of the contract. By contrast, the retail and restaurant sectors are characterized by simultaneity between the service provided to consumers and the payment. Additionally, the rules defining the business month when the firm must declare sales and the VAT to the tax administration are different for goods vs for services, as well as for domestic sales vs for exports.[24] In the absence of data on the exact lag between the time of production or service delivery and the time of the sales declaration, we explore dynamic effects separately for each sector. Finally, we use month-of-sample-by-industry fixed effects instead of month-of-sample fixed effects to capture time-varying shocks that are common across firms in the same industry, where 38 industries are defined according to the European nomenclature (A38).

The main specification is as follows:

$$Y_{i,t+l} = \alpha + \sum_{n=0}^{l} \beta_n \overline{PM_{2.5i,t+n}} + \sum_{n=0}^{l} \overline{W_{i,t+n}}\gamma_n + \sum_{n=0}^{l} \overline{h_{i,t+n}}\delta_n + \nu_{i,y} + \theta_{c,t} + \epsilon_{i,t}, \tag{7}$$

where $Y_{i,t}$ refers to the inverse hyperbolic sine transformation of the sales of firm i in month-of-sample t. This specification resembles a log-linear model while better accommodating the 2.5% of observations with zero monthly sales (Burbidge et al., 1988).[25] Variables $\overline{PM_{2.5i,t}}$, $\overline{W_{i,t}}$ and

[24] Specifically, the VAT on the sales of domestic goods has to be declared in the month where the good is delivered to the buyer; the VAT on the sales of domestic services has to be declared when the service is paid for; the VAT on exported goods and services has to be paid one month after the delivery. For a service that runs over several months or years, the VAT on the sales of the year are allowed to be declared at the end of the year. See https://entreprendre.service-public.fr/vosdroits/F31412.

[25] The log-linear model is the most widely used approach in the literature linking environmental shocks and sales

$\overline{h_{i,t}}$ correspond to the weighted averages of pollution exposure, weather controls, flu incidence and holidays across establishments j belonging to firm i, where the weights are the number of workers in establishment j at the end of the calendar year t. $\theta_{c,t}$ are month-of-sample-by-industry fixed effects and $\nu_{i,y}$ are firm-year fixed effects. In a robustness test, we control for firm fixed effects only to absorb time-invariant firms' characteristics.

The coefficients of interest are β_n, with $n \in \{0, 1, ..., l\}$. If we expect no lag between production and sales, $l = 0$ and equation (7) captures the contemporaneous effect of air pollution on sales. When we expect a one-month lag between production and sales, we impose $l = 1$ and thus also control for the leads of $\overline{PM_{2.5i,t}}$, $\overline{W_{i,t}}$ and $\overline{h_{i,t}}$ to avoid capturing the impact of contemporaneous pollution or weather conditions. Similarly, when we expect a two-month lag, we consider $l = 2$ and control for two leads. Future air pollution shocks, by contrast, should have no effect on current sales. Hence, we provide a placebo check by studying the effect of pollution at time t on sales at time $t-1$, while also controlling for one lag of $\overline{W_{i,t}}$ and $\overline{h_{i,t}}$.

Similar endogeneity issues arise at the worker level and at the firm level. The risk of reverse causality is accentuated at the firm level: any increase in sales is likely to increase pollution as a by-product of higher production. Measurement error in the allocation of pollution exposure still threatens identification at the firm level when the effects of pollution on sales are channelled through workers' labor force and productivity. We therefore also rely on an instrumental variable approach and proceed in two steps. First, we run equation (6) at the postcode level and save the vector of estimated $\widehat{\beta_{jk}}$. For each postcode, we compute the predicted pollution exposure as $\widehat{PM_{2.5g}}_{,t} = \sum_{k=1}^{K} \sum_{j=2}^{4} \widehat{\beta_{jk}} \text{WIND}_{j,k,t} \mathbb{1}(g \subset k)$. Weather controls and other variables beside wind directions from equation (6) are not used to build the instrument in this specification. Otherwise, the exclusion restriction assumption would not hold.

Second, we compute the weighted average of predicted pollution exposure across the establishments belonging to a firm. We obtain a firm-level predicted pollution measure, $\overline{\widehat{PM_{2.5i,t}}}$, which we use as an instrument for $\overline{PM_{2.5i,t}}$ in equation (7).[26] We cluster the standard errors at the firm level, the scale at which the instrument varies. As a robustness checks, we cluster standard errors two-way at the firm and year-month level. We explore the robustness of our results by comparing estimates using predicted pollution exposure as an instrument and using the same IV strategy as the one described in section 4.1 for the subsample of single-plant firms. For these firms, pollution and weather exposure is simply the one observed at the postcode of their unique establishment –

(see, for example, Dechezleprêtre et al. 2019; Fu et al. 2021; Addoum et al. 2020). After an inverse hyperbolic sine transformation of the outcome variable, the estimated coefficient can be interpreted as a semi-elasticity when the outcome variable takes large values (Bellemare and Wichman, 2020) – which is our case.

[26]In OLS models, inference using predicted regressos should be corrected for first-stage sampling variance. When the predicted regressor is used as an instrumental variable, like we do here, the standard errors of the 2SLS regression are unbiased under a set of weak assumptions (Wooldridge, 2010). Predicted regressors have similarly been used as instruments in Schlenker and Walker (2016) and Dahl and Lochner (2012).

similar to the analysis on workers' absenteeism.

4.3 First Stage Results: Wind Direction and Air Pollution

Local variations in wind direction are a strong predictor for local $PM_{2.5}$ concentrations. Figure 2 shows a map of the estimated $\widehat{\beta_{jk}}$ coefficients obtained with an OLS regression of equation (6) and their associated t-stat. Compared to South winds, winds blowing from the West (from the Atlantic ocean) significantly decrease pollution levels in the vast majority of postcodes. By contrast, winds blowing from the East increase pollution in the North-West, and decrease pollution in the South, which reflects the importance of South winds bringing dust from the Sahara, thereby increasing pollution, for the Southern region.[27] We show in the results section that the Kleibergen-Paap rk Wald F statistics of the two-stage least square estimation is well above critical values, suggesting that we do not have a weak instrument problem.

Given that each wind direction can affect pollution differently depending on the location of the postcode, we cannot perform the reduced form analysis linking the instrument directly to our outcomes. Yet, Figure 2 reveals that Eastern winds tend to increase pollution and Western winds tend to decrease pollution in areas concentrating economic activities (North, East, and Rhone region in the South-East). To provide intuition, Figure 4 shows the relationship between residualized West and East winds and residualized absenteeism in a binned scatter plot. On average for the whole country, more hours under West (East) wind decrease (increase) the number of sick leaves.

The main threat to the exclusion restriction assumption in our context is that other pollutants that also affect health outcomes co-vary with wind direction. Of the four other regulated air pollutants (SO_2, NO_2, PM_{10} and ozone), SO_2 and NO_2 are primary pollutants that convert to particulate matter within two to three days. By aggregating pollution concentration at the monthly level, we cannot estimate their effect on health independently. Additionally, PM_{10} is highly correlated with $PM_{2.5}$ (Pearson correlation coefficient: ρ=0.93) and actually includes $PM_{2.5}$. As a result, caution should be taken in interpreting our causal estimates as reflecting only the effect of $PM_{2.5}$. By contrast, ozone is a pollutant that is typically anti-correlated with other pollutants due to how it is formed in the atmosphere: ozone results from the chemical reaction between solar radiation, nitrogen oxide and volatile organic compound (Nasa Earth Observatory, 2003). In our data, the Pearson correlation coefficient between monthly $PM_{2.5}$ and ozone is -0.3. Figures 1 and B.5 illustrate this anti-correlation by showing the reverse seasonality of ozone vs $PM_{2.5}$ and NO_2 concentrations. In section 6, we show that when we instrument ozone with wind direction, we fail to detect a significant effect on absenteeism, suggesting that $PM_{2.5}$ is the main driver of the effect.

[27] In contrast to Spain (Holub et al., 2021), France does not report these dust events associated with Sahara winds.

5 Results

5.1 Effects on Workers' Sickness Leaves

Main results. Table 2 reports the main OLS and IV estimates of the effect of $PM_{2.5}$ on sickness leave outcomes. Panel a reports the estimates based on equation (5) using postcode fixed effects whereas panel b reports estimates of the same equation using postcode-by-year fixed effects. Column (1) of panel a shows that a one unit (1 $\mu g/m^3$) increase in monthly average $PM_{2.5}$ concentrations is associated with a significant increase of 0.07 workers taking a sickness leave per 1,000 workers. Column (2) of panel a shows that the IV estimate is about twice as high, with an increase of 0.15 workers taking a sickness leave per 1,000 workers. The effect corresponds to a 0.6 percent increase given the baseline average of 22 per 1,000 workers.

The IV estimate is also larger than the OLS estimate for the two other outcomes considered, which we interpret as evidence of an attenuation bias due to classical measurement error in the OLS estimate, potentially combined with a downward simultaneity bias. The Kleibergen-Paap F-statistic is 815, suggesting that weak instrument is not an issue in our setting. Hence, we base our interpretation of the results on the IV estimates in the rest of the paper. Column (4) of panel A shows that an increase by 1 $\mu g/m^3$ in monthly $PM_{2.5}$ concentrations increases the number of sick days associated with SLEs starting the same month by 2.7 days per 1,000 workers (about a 0.8 percent increase relative to the mean for SLEs that last less than 3 months). Column (6) shows that the associated increase in sickness leave spending amounts to €86.5 per 1,000 workers (about a 1 percent increase relative to the mean). Panel B shows that using postcode-year fixed effects yields similar effects to using postcode fixed effects.

One way to assess the magnitude of our results is to compare our IV estimates to prior studies of the effect of air pollution on labor supply. Consider our estimate on the number of work loss days (for sickness spells shorter than 3 months), which resembles the outcomes used in previous studies: a 10% increase in monthly $PM_{2.5}$ yields an elasticity of labor supply to pollution shocks of –0.12. Aragón et al. (2017) shows that a 10% increase in $PM_{2.5}$ exposure in the previous week reduces weekly hours worked by 2% for working adults in Peru, which implies an elasticity of labor supply to pollution shocks of –0.20. Holub et al. (2021) examines a representative sample of Spanish workers and shows that a 10% reduction in weekly PM_{10} concentration in the main urban areas in Spain reduces sickness-related absenteeism by 0.8% of the mean, implying an elasticity of the labor supply of –0.08. While these studies and ours differ in the type of pollutant, the time horizon, the IV strategy and the source of data, it is interesting that the elasticity of 0.12 of our study aligns with these other estimated elasticities.

Dynamic specification. Figure 5 shows dynamic specification coefficients that reflect the effect of future exposure to $PM_{2.5}$ on current outcomes as placebo checks (in event months $t-1$ and $t-2$) and the contemporaneous and lagged effects of previous months' exposure to $PM_{2.5}$ (in event months t, $t+1$, and $t+2$). Panel a in Figure 5 shows significant effects of air pollution on work absenteeism not only in the month of exposure, but also in the following month, with a smaller magnitude. This indicates that some pollution-related health symptoms take time to set in. The small non-statistically significant estimates of the effects of pollution in event month $t+2$ suggest that pollution-related health shocks only last two months on average and allow us to rule out a displacement effect.[28] We fail to detect an impact of pollution at t on absenteeism in the previous periods. As a result, changes in pollution induced by wind direction can be interpreted as quasi-random. Panel b of Figure 5 reveals that the number of sick days per 1,000 workers significantly increase in the month of exposure, while the effect in the following month is smaller in magnitude and not statistically significant at the 5% level. This suggests that the effects of pollution are more severe (as evaluated by the number of missed days from work) in the contemporaneous month than in the following month.

Heterogeneous effects by wage, industry, and firm size. We report IV estimates of the effect of $PM_{2.5}$ on sickness leave episodes by wage, industry, and firm size. Workers with different working status and wages may face different incentives to take a sick leave from work. For instance, low-wage employees in precarious contracts may be less willing to take a sickness leave if it endangers their job security. However, they may also be exposed to higher environmental risks (Hsiang et al., 2019). By contrast, high-wage employees may have a lower replacement rate given the cap on publicly funded sickness leave benefits and may also feel that their skills are less substitutable, inducing them to take sick leave less often (Hensvik and Rosenqvist, 2019). Only considering full-time employees, we break down annual wages in 2009 by deciles and run the 2SLS estimation outlined in equations (5) and (6) on 10 postcode-level datasets that only include workers from a given wage decile in 2009.

Figure B.8 reports estimated effects on sickness leaves separately by wage decile. We find that workers with relatively low wages (D2, D3, and D4) experience the largest absolute effect of air pollution on work absenteeism, with roughly a doubling of the positive effect observed overall. These estimates are also all statistically significant at the 5% level. By contrast, workers with the lowest annual wage (D1), as well as workers with above-median annual wages (especially, D5, D6, and D9), have lower absolute and non-significant effects of air pollution on absenteeism. Interestingly, workers in higher income deciles are exposed to higher pollution concentrations on

[28] If pollution only caused already vulnerable workers to fall sick earlier than they would have under a counterfactual with no pollution shock, we would see a decrease in absenteeism in month $t+2$.

average, as indicated by the average level reported below each decile in Figure B.8. In France, high-wage jobs are concentrated in large and highly polluted urban areas, such as Paris. As a result, the observed heterogeneity in the pollution-sickness leave relationship cannot be driven by different exposures to air pollution, but rather by different vulnerabilities and different incentives to take sick leave.

Firms differ in the work conditions they offer to their employees, labor contracts, as well as moral pressure relative to absenteeism, which influence the ability of workers to take sick leave. Figure B.9 reports estimated effects by economic sector. We run the regression at the establishment level instead of the postcode level and use establishment fixed effects. The magnitudes and relative effects are largest in the retail and restaurant sector and in the other business-to-consumer services sector, followed by the manufacturing sector. The fact that point estimates are often imprecise within a sector suggests substantial within-sector heterogeneity. In the professional services sector, workers rarely take any sick leave and the effect of pollution on absenteeism is not statistically different from zero. There are two explanations corresponding to the two categories of workers in that sector. First, for workers from temporary work agencies, job security is low and they may not be willing to enter sickness leave even if they feel sick. Second, for high-skilled professional workers, there may be a composition effect, since the marginal effect of pollution on absenteeism is lower for high-wage workers in Figure B.8. Finally, Figure B.10 reports estimated effects by firm status and size, also running the regression at the establishment level. We do not find substantial differences in the effect when comparing single and multi-establishment firms, nor when comparing firms with sales above and below the median.

5.2 Effects on Firms' Sales

Main results. Due to substantial heterogeneity across sectors in the lag between production and reported sales, we examine the effect of pollution by sector. Figure 6 reports the dynamic effects of air pollution on firms' sales for the six different sectors. Both OLS and IV estimates are reported for the whole sample, and we additionally report the IV estimates for the subsample of single-establishment firms, for which pollution exposure only depends on one location.

Panel a in Figure 6 shows that the OLS estimate is positive and statistically significant at t but decreases at $t+1$ and $t+2$ for manufacturing firms. In contrast, the IV estimates reveal a significant and negative effect of air pollution on sales in the two months following exposure: a one-unit increase of pollution at t decreases sales by 0.24% in $t+1$ and by 0.29% in $t+2$. The discrepancy between the OLS and the IV estimates is probably reflecting an upward bias due to reverse causality: within a firm-year and controlling for country-wide time-varying shocks, months with a greater economic activity bring more sales but also more emissions, and thus, pollution concentration. The patterns for all manufacturing firms are similar to the patterns for single-establishment firms only. Although single-establishment firms may differ from multi-establishment firms in many dimensions (e.g., in size, productivity, capacity to export, capacity to innovate), finding similar results gives some credibility to our empirical strategy where multi-establishment firms' exposure to pollution is measured by a weighted average of establishments' exposure.

Panel b in Figure 6 also reveals significant negative effects of air pollution on sales for construction firms in months t, $t+1$ and $t+2$. A one-unit increase of pollution at t decreases sales by 0.30% in t, by 0.44% in $t+1$ and by 0.58% in $t+2$. By contrast, the OLS point estimates are higher, consistent with an upward bias. The magnitudes of the IV estimates are larger in the construction sector than in manufacturing, which is consistent with construction workers being more exposed to air pollution since they work outside and cumulate ambient air pollution with construction emissions (especially, dust, ultra fine particulate matter, and nitrogen dioxide). From Figure B.9, we previously found that construction workers tend to have a low labor force response to air pollution shocks. As a result, we can infer that even though construction workers do not often take sickness leaves, they experience lower productivity while at work when air pollution is high.

The IV estimates of panel c suggest that air pollution exposure in month t reduces sales at t for firms supplying professional services, while the effects on sales at $t-1$, $t+1$, and $t+2$ are negligible, and the OLS estimates are close to 0 at all periods. The patterns are similar for single-establishment firms and for all firms. The magnitude of the effect in the contemporaneous month is large: a one-unit increase in pollution at t reduces sales by 0.55%. While we failed to detect an effect of pollution on sickness leave in that sector (see Figure B.9), the large impact on sales could be explained by the fact that workers still suffer from pollution-induced physical symptoms while at work, thereby experiencing a decrease in their productivity. They may also adjust their labor force at the margin by leaving their job early during these episodes. Interestingly, there is no significant rebound in sales in the following months $t+1$ and $t+2$. This indicates that workers are not be able to postpone some of their tasks to smooth the effects of pollution shocks over time.

Panel d in Figure 6 reveals that pollution shocks do not reduce sales in the ICT and other business-to-business services sector at t and $t+1$, but it slightly decreases sales at $t+2$ for multi-establishment firms. This is compatible with the payment occurring a few months after the service is provided to the client firm. The decrease is not statistically significant for single-establishment firms.

Services to consumers record their sales at the time of the payment, which often corresponds to the time when the service is provided. This suggests there will be no lag between the time of pollution exposure and sales effects. Panels e and f in Figure 6 shows that the IV point estimate is negative at all time periods considered, including at $t-1$, which renders the interpretation of the results as a causal effect of pollution at t difficult. This correlation may be driven by some specific seasonality patterns in these industries, which we fail to capture with the firm x year fixed effects.

Heterogeneous effects by firm size by sector. We investigate heterogeneous effects by firm size within each sector, based on the total annual sales of the firm in 2012. Tables C.2 and C.3 show that the effect on manufacturing firms' sales at $t+1$ is significant only for smaller firms whereas pollution affects larger and smaller construction firms in similar ways at $t+1$. The effect of pollution on firms in the professional services sector is only significant for larger firms at t. We detect a positive effect of pollution on sales for large firms in the retail and restaurant sector, in contrast to a negative effect for large firms in other business to consumer services. Again, a causal interpretation of these results is problematic because we also find a negative effect of pollution at t on sales at $t-1$ for these two sectors.

6 Robustness checks

6.1 Worker-Level Robustness Checks

Time and unit dimensions. To validate the evidence of a causal effect of $PM_{2.5}$ concentrations on sickness leave episodes, we explore the unit and time dimensions of analysis. Panel A of Table 3 reports the effect of $PM_{2.5}$ on worker-level sickness leave episodes using a 10% sub-sample where the data have not been aggregated at the postcode level, so that individual fixed effects are used instead of postcode fixed effects. We thus control for time-invariant individual characteristics that may influence the decision to call sick from work. Magnitudes differ from the main results due to sampling variation, but we still find significant effects of $PM_{2.5}$ on the likelihood to start a sickness leave episode, the number of days spent in sickness leave, and the total spending at the individual level. Panel B of Table 3 reports the effect of $PM_{2.5}$ on the usual outcomes when the data is aggregated at the weekly level instead of the monthly level. Since $PM_{2.5}$ affects the human body within hours of exposure, we should be able to detect an impact of weekly pollution exposure on weekly absenteeism. Magnitudes are lower than in the main results at the monthly level, with two possible explanations. First, increasing weekly concentrations by 1 $\mu g/m^3$ entails a smaller pollution shock than increasing monthly concentrations by the same amount. Second, the outcome only captures pollution-induced sickness leave on the contemporaneous week of the shock, while the monthly specification captures also lagged responses within the month.

Other pollutants. One concern with interpreting our estimates as the causal effects of $PM_{2.5}$ is that other pollutants – ozone in particular – may also be influenced by wind directions while being omitted from our main specification. Given our specification of monthly concentration in $PM_{2.5}$, as mentioned in section 4.3 we are not able to disentangle the independent effects of SO_2, NO_2 and PM_{10} from those of $PM_{2.5}$. By contrast, ozone (O_3) is anti-correlated with $PM_{2.5}$ and may have independent effects that we can capture. In column (2) of Table 4, we instrument ozone with wind direction instead of $PM_{2.5}$. Although the first stage is still strong, we fail to detect any impact on sickness leave incidence. This could be due to a lack of causal effect of ozone on work absenteeism, potentially because high ozone levels occur mostly during summer vacation period (see figure B.5b), when French workers are less likely to be at work and thus take a sickness leave. But this lack of effect may also reflect a downward bias from omitting $PM_{2.5}$, which is negatively correlated with ozone and positively correlated with absenteeism.

To mitigate the risk of violating the exclusion restriction, we also examine the relationship between the French air quality index, a synthetic index based on the daily concentrations of different pollutants (SO_2, NO_2, PM_{10}, $PM_{2.5}$ and ozone), ranging from 1 (best air quality) to 6 (worst air quality). We build the index using daily data for four pollutants available in the CHIMERE dataset and take the monthly average in each postcode.[29] Column (3) shows that a one-unit increase in the air quality index increases the number of workers starting a sickness leave that week by 2.9 workers per 1,000 workers. Although the magnitude of the effect is not comparable with our main results, the positive and statistically significant relationship between pollution and absenteeism holds when pollution is measured with this synthetic air quality index.

Other robustness checks. We report a series of additional robustness checks in Table 5. Column (1) reports our baseline estimate of the effect of $PM_{2.5}$ on the number of SLEs per 1,000 workers. In Column (2), weather controls take the form of three continuous variables – monthly averages of daily maximum temperatures and daily wind speeds, and total rainfall – rather than a set of indicators and interactions, as in the main specification. The results are similar to the baseline estimates.

To be able to interpret our IV estimates as local average treatment effects (LATE), the monotonicity assumption must hold. In our setting, the monotonicity assumption implies that each 90° wind direction should either increase pollution for all postcodes located in a given 100 x 100 km area, or decrease it for all of them. Like Deryugina et al. (2019), we note that this assumption could be violated if the wind direction-pollution relationship varies across postcodes or over the

[29] Following the method used by the regional air quality agencies, we first create a sub-index ranging from 1 to 6 for each pollutant based on official thesholds; then we allocate the maximum value of all sub-indices to the air quality index that day. In the data, $PM_{2.5}$ takes the maximum value of all sub-indices 70% of the days from October to March, while ozone takes the maximum value 80% of the days From April to September.

course of the year within a given area. To test whether the monotonicity assumption is likely to hold in our setting, we check whether the point estimate varies when we allow the first stage coefficients to vary over smaller geographic areas and over the course of the year. Column (3) uses 50km-by-50km areas instead of 100km-by-100km areas, and Column (4) allows the interaction between wind direction and local areas to vary by quarter. We find that the results are similar to the one from the baseline specification.

Another potential concern is that some SLEs starting at the end of December might have been censored or reported as starting in January by the social security administration. To the extent that this reporting error holds for the entire country, its effect should be captured by month-by-year fixed effects. We however re-run the main analysis excluding December and January (column (5)) to verify that the results are not driven by this potential measurement error. We confirm that our results are robust to these restrictions.

Last, we check that our results are not driven by a behavioural response to air quality alerts. Air quality alerts do not exist for $PM_{2.5}$ in France but are issued for PM_{10}. Owing to the high correlation between PM_{10} and $PM_{2.5}$, we use the regulatory thresholds for the issuance of PM_{10} alerts.[30] For each postcode-month we build a variable corresponding to the number of days where a PM_{10} air quality alert was issued, and we re-run the IV regression after excluding all the postcode-months with at least one day of air quality alerts. From column (6), we see that the point estimate slightly increases after excluding these observations, but remains similar to the main estimate.

6.2 Firm-Level Robustness Checks

Since we only find significant effects of air pollution on the sales of firms in the manufacturing, construction, and professional services sectors using our main specification, we also focus on these sectors for our robustness checks.

Comparing empirical strategies for single-establishment firms. By definition, single-establishment firms have a unique geographic location and their workers' pollution exposure is simply the pollution exposure at their unique establishment. This allows us to use an empirical strategy akin to the one adopted for work absenteeism: the first stage equation corresponds to (6), except that we examine effects at $t+1$ for manufacturing and construction controlling for one lead of the instrument and control variables, and at t for professional services. Figure 7a shows that point estimates are similar across specifications, especially for construction and professional services

[30] Two levels of alerts exist: level 1 provides information on air pollution levels and advises vulnerable individuals to avoid physical activities outside ; level 2 adds strict enforcement measures such as driving restrictions (see https://www.airparif.asso.fr/procedure-dinformation-et-dalerte for more information). Until November 2014, level 1 was triggered when daily average PM_{10} exceeded 80 µg/m^3 and level 2 when it exceeded 125 µg/m^3. From November 2014 onwards, the thresholds were lowered to 50 µg/m^3 for level 1 and 80 µg/m^3 for level 2.

sectors. This robustness check validates that using predicted pollution as an instrument instead of wind direction yield relatively similar results. For the manufacturing sector, the point estimate has a lower magnitude with the wind direction instrument, rendering the result not statistically different from zero at conventional significance levels. This small difference in magnitude might be related to the sample of single-establishment firms in manufacturing or might suggest that a wind direction instrument could yield slightly more conservative results.

Polynomial Distributed Lag. In contrast to the 2SLS regression on workers' absenteeism, we can explore dynamic effects of pollution on firms' sales over 6 months using a distributed lag specification. This allows us to explore the timing of the effects of pollution on sales assuming relatively long lags between production and sales. To reduce the noise due to the serial correlation in wind direction and pollution exposure over time, we use a polynomial distributed lag (PDL) (Schwartz, 2000; He et al., 2019), where we impose a smooth polynomial function on the lag structure to discipline the coefficients. Assuming a cubic polynomial functional form, we examine in a single regression the effects of pollution at t, $t-1$,... up to $t-5$ on sales at t by sector. It implies that we impose the following relationships on the coefficients β_l, for $l \in \{0, .., 5\}$: $\beta_l = \sum_{k=0}^{3} \gamma_k l^k$, resulting in $\beta_0 = \gamma_0$, $\beta_1 = \gamma_0 + \gamma_1 + \gamma_2 + \gamma_3$, and $\beta_2 = \gamma_0 + 2\gamma_1 + 4\gamma_2 + 8\gamma_3$ for the first parameters. Using these relationships, we rewrite the regression equation as a function of γ_ks and estimate by OLS and by 2SLS the coefficients γ_1, γ_2, and γ_3. Combining these point estimates and associated standard errors, we recover the point estimates β_ls and associated standard errors by sector. We report in Figure 7 the OLS and IV estimates for β_0 with label t, β_1 with label $t+1$, up to β_5 with label $t+5$. The OLS and IV results from t to $t+2$ are similar to our main results described above. Focusing on the IV estimates, adding more time periods reveal that the negative effect of pollution on sales is the largest at $t+3$ in manufacturing, at $t+1$ or $t+2$ in construction, and at t for professional services. The effect of pollution on sales slowly fades out over time until it reaches zero at $t+5$.

Other robustness checks. Table 6 shows that the effects of pollution shocks at t on sales at $t+1$ for manufacturing and construction and sales at t for professional services obtained with our main specification are robust to various checks. First, we exclude 2009 from the sample since it was the year where an economic crisis strongly affected manufacturing and construction. Second, we use continuous weather variables instead of indicators and interactions, and column (3) shows that the magnitudes of the effects are slightly larger when doing so. Third, we use firm fixed effects instead of firm-year fixed effects, which cannot absorb annual productivity shocks and changes in the number of plants over time, and also find in column (4) that the magnitudes are higher. Fourth, in column (5) we discard the months with at least one day of PM_{10} alerts and find that the magnitude barely changes. Fifth, beside the correlation of the error terms across time within a firm, there could also be a correlation of the error terms across firms within a time period. In column (6) we cluster the standard errors both at the firm and month-by-year level to account for this second type of correlation. We find that two-way clustered errors are larger, resulting in a non statistically significant point estimate for manufacturing sales but robust findings in construction and professional services sector. Table 7 reproduces the same robustness test as in Table 4 and tests for the effect of ozone only and of the multi-pollutant air quality index on sales. From column (2), instrumented ozone has no significant effect on sales across the three sectors considered. From column (3), a one-unit increase in the air quality index has a negative but non-statistically significant effect on manufacturing sales, and a negative and significant effect on sales in construction and professional services.

Spillover effects among competing firms A potential concern would arise if firms that are non-exposed in a given month and compete with firms that are exposed increase sales due to the decrease in productivity among their exposed competitors. When competition occurs within a relatively local market, such as for firms in the retail and restaurant sector or in the construction sector, spillovers are by construction excluded because firms in other cities cannot compete for the same customers. When competition occurs at the national or international level, as for manufacturing firms for instance, the temporary nature of the labor force and productivity responses to pollution shocks at the monthly level (as evidenced by Figures 5 and 7) make changes in pricing, and therefore within-industry spillovers, unlikely. Our data does not allow us to test for a price effect of air pollution shocks. Yet, past evidence suggests that manufacturing firms adjust their prices every six months whereas business-to-business services firms adjust their prices every twelve months (Gautier, 2007). These time intervals are therefore longer than the impacts of air pollution shocks.

7 Discussion

7.1 Pollution-Induced Costs on Workers and Firms

In this section, we compare the magnitude of the total response of sales to air pollution shocks and the labor supply channel measured by work absenteeism by sector.

To provide back-of-the-envelope estimates of the magnitude of the labor supply channel, we need to multiply the number of missed work days by the average marginal product of labor in each sector. From the statistical agency INSEE, we obtain average marginal products of labor equal to

€257 per day in manufacturing and to €189 per day in construction.[31] Combining our estimate of 2.75 work loss days per 1,000 workers with an average of 92 workers per manufacturing firm implies 0.25 work loss days for the average manufacturing firm for a one-unit increase in $PM_{2.5}$, which results in a monthly cost of €65. This small cost contrasts with the total pollution-induced monthly loss in sales of roughly €5,488, given the average monthly sales of €2,386,088 in manufacturing. Work absenteeism valued at the marginal product of labor thus represents only 1.2% of total sales losses. Similarly, for construction, we calculate a labor supply cost of roughly €18 per month, which represents only 0.7% of the pollution-induced monthly loss in sales of roughly €2,440. Our results suggest that the increase in absenteeism due to pollution is not the main driver of the total response of sales.

We can offer three potential explanations for why the cost associated to work absenteeism represents such a small share of the total sales losses, and also for why the sales losses are negligible in some sectors. First, we expect that firms can compensate for the temporary decrease in labor supply with an increase in the productivity of non-absent workers, for example by readjusting working hours or changing the allocation of tasks across workers, or even by hiring temporary workers. Such strategies could explain why sales in the retail and restaurant sector and in other business-to-consumer services sector seem unaffected by pollution shocks. In sectors where workers are complementary along a production value chain, such as manufacturing, such strategies may not be available.

Second, some sectors probably face stronger negative effects of pollution on non-absent workers' productivity. While we are unable to quantify this productivity channel, we can formulate hypotheses on the types of workers most likely to be affected by it: first, workers who breathe a more polluted air while working, either because they are working outdoors or because they also breath specific work emissions, such as construction and manufacturing workers; second, high-skilled workers, who may be particularly vulnerable to the effects of pollution on their cognitive skills. Third, some sectors may face a strong decrease in demand due to pollution. While we cannot quantify this channel either, we would expect demand responses to be stronger in the business-to-consumer non-tradable services, where sick consumers may avoid grocery shopping or going to a restaurant. By contrast, customers in professional services or manufacturing sectors are less likely to live in the vicinity of sellers, which implies that they probably will not be exposed to the same pollution shocks. We interpret the lack of significant decrease in sales in the business-to-consumer services as suggestive evidence for a limited role of the demand channel.

Lastly, firms within a sector may differ widely in the lag between the time of production and the time when sales are recorded depending on the time necessary for production, on their contracts, or on their exporting behaviors. This generates measurement error. As a result, our inability to

[31] See INSEE webpage: https://www.insee.fr/fr/statistiques/4255787?sommaire=4256020.

detect any significant effect of air pollution on sales in some sectors (for instance, for business-to-business services) may be related to a wider within-sector heterogeneity in the lag between production and sales.

7.2 Benefits from Meeting the Daily $PM_{2.5}$ Threshold from WHO

To illustrate the economic significance of our results, we provide back-of-the-envelope calculation of the benefits of meeting the daily $PM_{2.5}$ WHO target in terms of avoided sickness leave spending and avoided lost sales. Over our 7-year study period, the 15µg/m^3 threshold is exceeded for 37% of the worker-days. Bringing each day above the threshold to 15µg/m^3 would decrease monthly average pollution exposure from 15.4 to 11.5µg/m^3, a 25% decrease compared to the levels observed over 2009-2015.

Avoided sickness leave spending. To quantify the monetary benefits in terms of avoided sickness days, we compute, for each postcode-month, (i) the decrease in $PM_{2.5}$ that is required to meet the 15 µg/m^3 threshold; (ii) the associated number of avoided sickness days from Table 2 column (4)'s estimates; (iii) the associated avoided sickness leave benefit spending from Table 2 column (6)'s estimates. For each postcode-month, we multiply the estimated marginal effect of a one-unit change in pollution on absenteeism by the required decrease to reach the WHO recommendations. Scaling the coefficients obtained at population size, we obtain that meeting the WHO thresholds would have saved an annual 1.9 million sick days and avoided €66 million of publicly funded sickness leave benefits.[32] This represents 3% of the social security spending on episodes shorter than 3 months and 1% of the total spending.[33] Additionally, using data on the average duration of SLEs and the schedule for employer-funded sickness leave benefits, we estimate that it would have avoided €65 million in employer-funded sickness leave spending.[34] In total, the avoided sickness leave spending amounts to €130 million, shared equally between the employer-funded and state-funded benefits.

Our analysis is restricted to private sector employees, so any avoided sickness leave spending for public sector employees is excluded from the calculation. The analysis does not include other

[32] Our sample has on average 344,052 individuals per year, 1.84% of the total population of private sector employees, which amounts to 18,730,000 workers in 2015. (See INSEE: https://www.insee.fr/fr/statistiques/2496914).

[33] Total public spending for private employees' sickness leaves amounted to €7,091 million each year over the period (DREES, 2020).

[34] Among SLEs lasting less than three months, the average duration is 16 days. The avoided 1.9 million of sick days per year correspond to $1,932,965/16 = 120,810$ avoided sickness leave episodes. Given the average daily wage of $25,865/365 = 71$€, the benefit of avoiding one 16-day SLE for private sector employers is approximately $(16-7)\times 0.4\times 71+(2/3)[3\times 71+4\times 0.5\times 71+(16-7)\times 0.1\times 71] =$€534. The avoided spending is $534\times 120,810 = 64,512,540$.

benefits from reduced air pollution for workers, such as avoided healthcare costs, avoided disutility from being sick, and additional income losses associated to a low replacement rate.

Additional sales. We combine the estimates from Figure 6 with the decrease in pollution exposure at the firm level implied by meeting the WHO thresholds. To be conservative, we focus on effects at period $t + 1$ for manufacturing and construction, although sales also decrease at $t + 2$. For professional services, we focus on effects at period t. We obtain that annual sales would be higher by €1.9 billion for the construction sector, by €4.2 billion for the professional services sector, and by €5.3 billion for the manufacturing sector – although the estimates for the manufacturing sector are less robust than for construction and professional services. In total, meting the WHO targets would have increased annual sales by €6.1 billion in an average year between 2009 and 2015 excluding manufacturing – or 0.3% of the French GDP in 2015 –, and €11.3 billion if we include manufacturing – or 0.5% of the French GDP in 2015.

Comparison of benefits to the costs of meeting the WHO thresholds. To our knowledge, there is no estimate of the cost of bringing $PM_{2.5}$ concentrations down to the daily WHO thresholds for France. Drawing on Dechezleprêtre et al. (2019), we rely on the estimates of the cost of reducing $PM_{2.5}$ *emissions* – not concentrations – obtained in a report published by the European Commission.[35] The cost estimates are country-specific and available for different scenarios with varying levels of emission reductions. Absent a scenario reducing emissions by 25% (as implied by the respect of the WHO threshold), we consider the two closest scenarios for France: the scenario reducing emissions by 16 (option 6C) would cost €375 million annually whereas the scenario reducing emissions by 33% (option 6D) would cost €7,675 million annually. Our estimate of the economic benefits of meeting the WHO targets therefore largely exceeds the lower bound, and represents 80% of the upper bound if we exclude effects on the manufacturing sector and 150% if we include them.

7.3 Policy implications

Our results have several policy implications. First, introducing a daily threshold of 15µg/m^3 in France could generate substantial economic benefits, in addition to the benefits from avoided pollution-induced healthcare consumption which we do not quantify. These benefits may partly compensate for the costs of implementing such regulations. If we consider that manufacturing firms will have to invest in cleaner technologies to abate their air emissions, we could compare these investment costs with the increased profits from enhanced worker productivity. Given that France

[35]See part 3, page 43 of the following report: https://ec.europa.eu/environment/archives/air/pdf/Impact_assessment_en.pdf.

is similar to the average EU member state in terms of pollution levels and composition of the economy, we could use our results to inform the debate about a future tightening of the regulatory standards at the European Commission to bring them closer to the WHO recommendations.[36] Our analysis sheds light on the large economic benefits it could bring for a lof of firms.

Second, cost-benefit assessments of policies improving air quality typically do not take into account the positive effects that cleaner air could have on firms' economic performance, beyond reducing work absenteeism. For example, in a flagship report, the OECD mentions the importance of including productivity losses in air pollution cost assessments (OECD, 2016). Their recommendation is to use the average market wage rate as a proxy for the marginal cost of labour and multiply it by the number of work loss days due to pollution. Our results show that the direct effect of pollution on sales goes well beyond this monetary value of work loss days in several sectors. Thus, this approach – also used by the WHO (WHO, 2014b) – will tend to underestimate the economic cost from air pollution.

8 Conclusion

In this paper, we show that an increase in workers' exposure to particulate matter causes an increase in sickness-related absenteeism. Separately, we identify a negative effect on sales in three sectors: construction, manufacturing, and professional services. We find that the economic cost of pollution associated with these firm-level sales losses exceeds by far the monetary value of pollution-induced absenteeism valued at the marginal product of labor.

Our analysis has several implications for research and policy. First, our analysis suggests that the productivity channel plays an important role in the transmission of pollution shocks to firms' sales. Based on the sectors where the sales losses are the largest, we laid out potential explanations of the importance of this channel: difficulty to substitute sick workers, complementarities in the production value chain, cumulative effects of work emissions, and cognitive impacts of pollution. Developing research designs to better understand the underlying mechanisms for each affected sector could be a promising path for future research.

Second, there is a large literature in economic geography and urban economics that relates high density with a high productivity, one of the benefits of agglomeration (Combes et al., 2012; Ahlfeldt and Pietrostefani, 2019). Recent work separately shows that high density also causes high levels of air pollution (Carozzi and Roth, 2019). Our work suggests that pollution levels may be an important omitted variable in the estimation of agglomeration effects. This omitted variable is expected to bias the effect of density on productivity downward, given the positive relationship

[36] see https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12677-Qualite-de-lair-revision-de-la-reglementation-de-lUE_fr

between density and pollution and the negative relationship between pollution and productivity. Revisiting estimates of agglomeration effects on productivity net of pollution effects would be an interesting avenue for urban and environmental economists.

Third, ex-ante cost-benefit analyses of environmental regulation that do not account for the negative effect of pollution on firms' performance will significantly underestimate the net benefits of these regulations. As the European Commission is currently considering to change regulatory standards to bring them closer to the WHO recommendations, it seems all the more important to properly quantify the costs and benefits of doing so. In our analysis, we provide an estimate of the benefits of bringing daily exposure to $PM_{2.5}$ down to the WHO recommendations that represents between 80% and 150% of an available cost estimate. Adding health benefits for the entire population to our estimates that depend exclusively on work loss days and sales losses, the benefits will significantly exceed the costs.

References

Addoum, J. M., D. T. Ng, and A. Ortiz-Bobea (2020). Temperature Shocks and Establishment Sales. *The Review of Financial Studies 33*(3), 1331–1366.

Adhvaryu, A., N. Kala, and A. Nyshadham (2022). Management and shocks to worker productivity. *Journal of Political Economy 130*(1).

Aguilar-Gomez, S., H. Dwyer, J. S. Graff Zivin, and M. J. Neidell (2022). This is Air: The "Non-Health" Effects of Air Pollution. Working Paper 29848, National Bureau of Economic Research. Series: Working Paper Series.

Ahlfeldt, G. M. and E. Pietrostefani (2019). The economic effects of density: A synthesis. *Journal of Urban Economics 111*, 93–107.

Anderson, M. L. (2020). As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality. *Journal of the European Economic Association 18*(4), 1886–1927.

Aragón, F. M., J. J. Miranda, and P. Oliva (2017). Particulate matter and labor supply: The role of caregiving and non-linearities. *Journal of Environmental Economics and Management 86*, 295–309.

Bagilet, V. and L. Zabrocki (2022). Why acute health effects of air pollution could be inflated. *Mimeo*.

Barwick, P. J., S. Li, D. Rao, and N. B. Zahur (2018). The Healthcare Cost of Air Pollution: Evidence from the World's Largest Payment Network. Technical Report 24688, National Bureau of Economic Research, Inc. Publication Title: NBER Working Papers.

Bellemare, M. F. and C. J. Wichman (2020). Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics 82*(1), 50–61. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12325.

Borgschulte, M., D. Molitor, and E. Zou (2022). Air Pollution and the Labor Market: Evidence from Wildfire Smoke.

Burbidge, J. B., L. Magee, and A. L. Robb (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. *Journal of the American Statistical Association 83*(401), 123–127. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1988.10478575.

Carozzi, F. and S. Roth (2019). Dirty density: air quality and the density of American cities. Technical Report dp1635, Centre for Economic Performance, LSE. Publication Title: CEP Discussion Papers.

Champalaune, P. (2020). Inequality in Exposure to Air Pollution in France: Measurement and Impact of a City-Level Public Policy. pp. 67.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016). Particulate Pollution and the Productivity of Pear Packers. *American Economic Journal: Economic Policy 8*(3), 141–169.

Chang, T. Y., J. G. Zivin, T. Gross, and M. Neidell (2019). The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China. *American Economic Journal: Applied Economics 11*(1), 151–172.

CITEPA (2021). Secten – le rapport de référence sur les émissions de gaz à effet de serre et de polluants atmosphériques en France.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica 80*(6), 2543–2594.

Currie, J., J. Voorheis, and R. Walker (2020). What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality.

Dahl, G. B. and L. Lochner (2012, May). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review 102*(5), 1927–1956.

Dechezleprêtre, A., N. Rivers, and B. Stadler (2019). The economic cost of air pollution: Evidence from Europe. Technical report, OECD, Paris.

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

Dong, R., R. Fisman, Y. Wang, and N. Xu (2019). Air Pollution, Affect, and Forecasting Bias: Evidence from Chinese Financial Analysts. *Journal of Financial Economics 139*(3).

DREES (2020). Les dépenses de santé en 2019 - Edition 2020 - Les indemnités journalières.

European Commission (2020). Air quality: Commission decides to refer France to the Court.

European Environment Agency (2020). Air pollution: how it affects our health.

France Stratégie and Inspection générale des Finances (2021). Comité de suivi et d'évaluation des mesures de soutien financier aux entreprises confrontées à l'épidémie de covid-19 - Rapport final. Technical report.

Fu, S., V. B. Viard, and P. Zhang (2021). Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China. *The Economic Journal 131*(640), 3241–3273.

Gautier, E. (2007). La dynamique des changements de prix à la production: une analyse à partir des relevés de prix à la production. *Economie et Statistique 407*, 3–26.

Graff Zivin, J. and M. Neidell (2012). The Impact of Pollution on Worker Productivity. *American Economic Review 102*(7), 3652–3673.

Graff Zivin, J. and M. Neidell (2014). Temperature and the Allocation of Time: Implications for Climate Change. *Journal of Labor Economics 32*(1), 1–26. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Griliches, Z. and J. A. Hausman (1986, February). Errors in variables in panel data. *Journal of Econometrics 31*(1), 93–118.

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

He, J., H. Liu, and A. Salvo (2019). Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China. *American Economic Journal: Applied Economics 11*(1), 173–201.

Hensvik, L. and O. Rosenqvist (2019). Keeping the Production Line Running Internal Substitution and Employee Absence. *Journal of Human Resources 54*(1), 200–224. Publisher: University of Wisconsin Press.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut (2018). Era5 hourly data on single levels from 1959 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Accessed on 20-JUL-2022).*

Holub, F., L. Hospido, and U. Wagner (2021). Urban Air Pollution and Sick Leaves: Evidence from Social Security Data.

Hsiang, S., P. Oliva, and R. Walker (2019). The Distribution of Environmental Damages. *Review of Environmental Economics and Policy 13*(1), 83–103. Publisher: The University of Chicago Press.

INSEE (2021). Déplacements domicile-travail – Même sur de très courts trajets, l'usage de la voiture reste majoritaire - Insee Flash Provence-Alpes-Côte d'Azur - 70.

Krebs, B., J. Burney, J. G. Zivin, and M. Neidell (2021). Using Crowd-Sourced Data to Assess the Temporal and Spatial Relationship between Indoor and Outdoor Particulate Matter. *Environmental Science & Technology 55*(9), 6107–6115. Publisher: American Chemical Society.

Lichter, A., N. Pestel, and E. Sommer (2017). Productivity effects of air pollution: Evidence from professional soccer. *Labour Economics 48*, 54–66.

Meyer, S. and M. Pagel (2017). Fresh Air Eases Work – The Effect of Air Quality on Individual Investor Activity. Technical Report 24048, National Bureau of Economic Research, Inc. Publication Title: NBER Working Papers.

Mink, J. (2022). Putting a price tag on air pollution: the social healthcare costs of air pollution in France.

Nasa Earth Observatory (2003). Chemistry in the Sunlight. Publisher: NASA Earth Observatory.

OECD (2016). Social Costs of Morbidity Impacts of Air Pollution. OECD Environment Working Papers 99. Series: OECD Environment Working Papers Volume: 99.

Park, R. J., N. Pankratz, and A. P. Behrer (2021, July). Temperature, Workplace Safety, and Labor Market Inequality. Technical Report 14560, Institute of Labor Economics (IZA). Publication Title: IZA Discussion Papers.

Pinna, M. (2022). Binned scatterplots with marginal histograms: binscatterhist. *The Stata Journal 22*(2), 430–445. Publisher: SAGE Publications.

Pollak, C. (2015). L'effet du délai de carence sur le recours aux arrêts maladie des salariés du secteur privé.

Real, E., F. Couvidat, A. Ung, L. Malherbe, B. Raux, and A. Colette (2021). Historical reconstruction of background air pollution over france for 2000–2015. *Earth System Science Data preprint, in review.*(https://doi.org/10.5194/essd-2021-182).

Schlenker, W. and W. R. Walker (2016). Airports, Air Pollution, and Contemporaneous Health. *Review of Economic Studies 83*(2), 768–809.

Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology (Cambridge, Mass.) 11*(3), 320–326.

Sicard, P., E. Agathokleous, A. De Marco, E. Paoletti, and V. Calatayud (2021). Urban population exposure to air pollution in Europe over the last decades. *Environmental Sciences Europe 33*(1), 28.

Somanathan, E., R. Somanathan, A. Sudarshan, and M. Tewari (2021). The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing. *Journal of Political Economy 129*(6), 1797–1827. Publisher: The University of Chicago Press Chicago, IL.

US EPA (2018). EPA Report on the Environment - Particulate Matter Emissions. Technical report.

WHO (2014a). 7 million premature deaths annually linked to air pollution.

WHO (2014b). *WHO Expert Meeting: Methods and tools for assessing the health risks of air pollution at local, national and international level.* WHO Regional Office for Europe.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data.* The MIT Press.

World Health Organization (2016). WHO Expert Consultation: Available evidence for the future update of the WHO Global Air Quality Guidelines (AQGs). pp. 50.

9 Figures

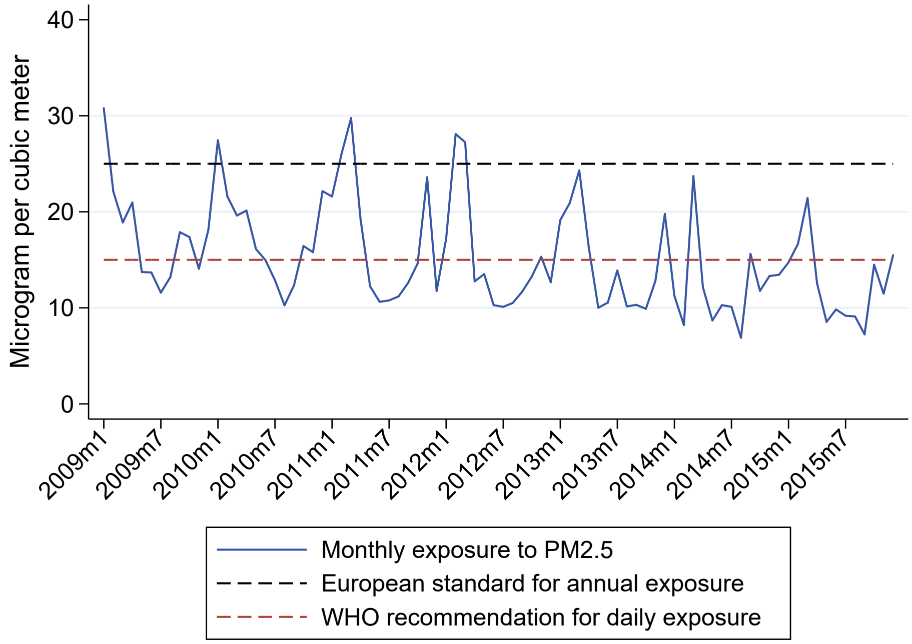
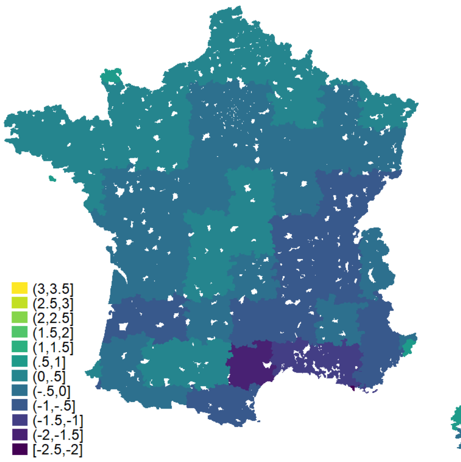
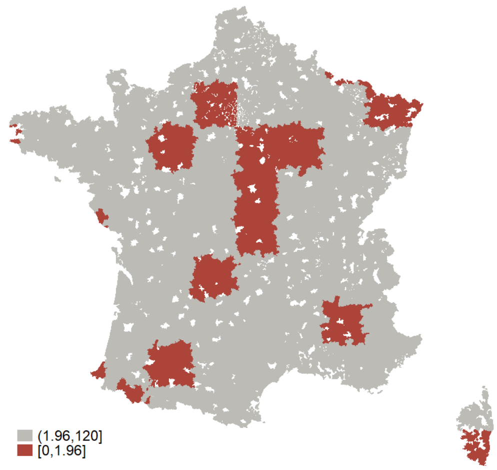


Figure 1: Average monthly exposure to $PM_{2.5}$ (µg/m^3)

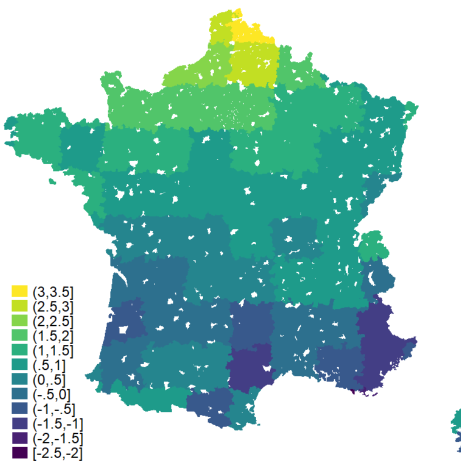
Notes: Figure presents the monthly average of workers' exposure to $PM_{2.5}$ measured at workers' postcodes. The sample of workers is the one used for the analysis of pollution effects on sickness leaves described in section 3.3 (unbalanced panel, N≈450,000). The European standard for annual exposure is 25µg/m^3 while the WHO's recommendation for daily exposure is 15 µg/m^3.



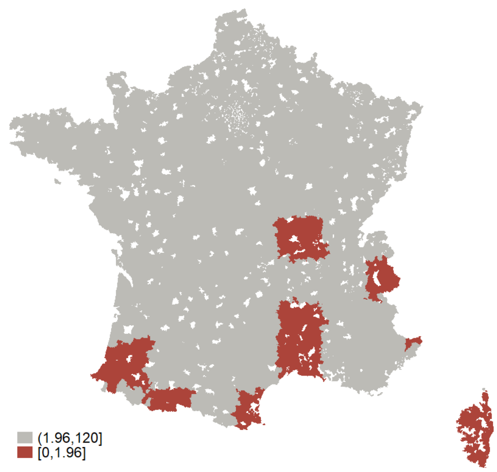
(a) West winds estimates $\widehat{\beta_{West,k}}$



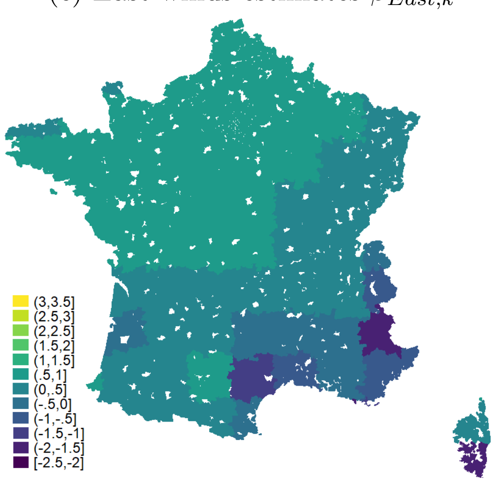
(b) Associated t-stat for $\widehat{\beta_{West,k}}$



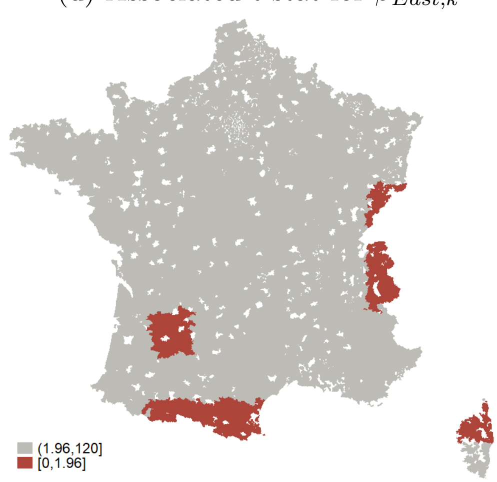
(c) East winds estimates $\widehat{\beta_{East,k}}$



(d) Associated t-stat for $\widehat{\beta_{East,k}}$



(e) North winds estimates $\widehat{\beta_{North,k}}$



(f) Associated t-stat for $\widehat{\beta_{North,k}}$

Figure 2: First stage results - point estimates and t-stat

Notes: Figure presents the estimated coefficients $\widehat{\beta_{jk}}$ and associated t-stat from equation (6) for wind direction $j = 2$ for West in panels a and b, $j = 3$ for East in panels c and d, and $j = 4$ for North in panels e and f, respectively. $\widehat{\beta_{jk}}$ express the average increase in monthly $PM_{2.5}$ concentration in the postcodes located in a given 100km x 100km grid k in month t when wind blows 10 percentage point more from direction j, compared to blowing from the South in month t. In panels b, d, f, the red cells are those where the estimated coefficient is not statistically significantly different from zero at the 5% level. The grey outline shows the postcode boundaries.

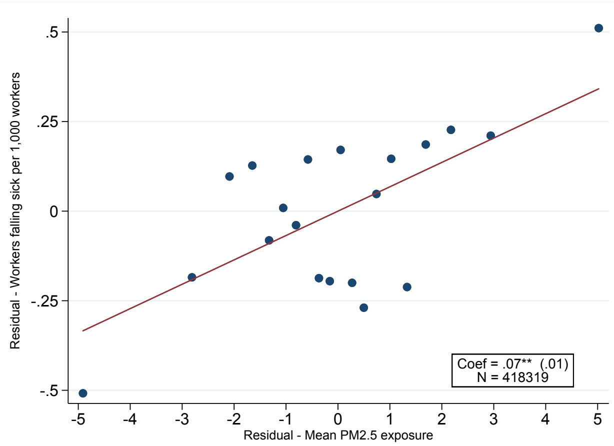
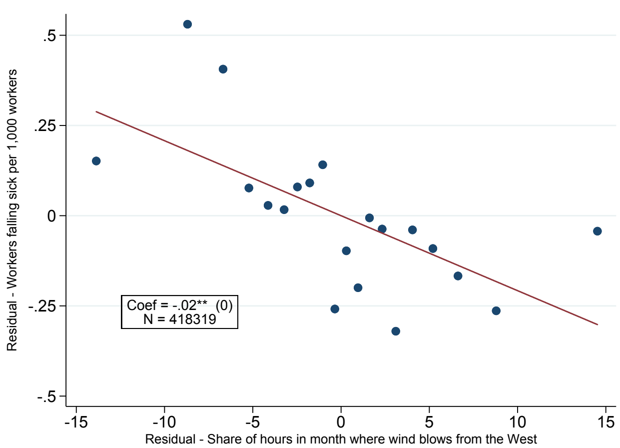
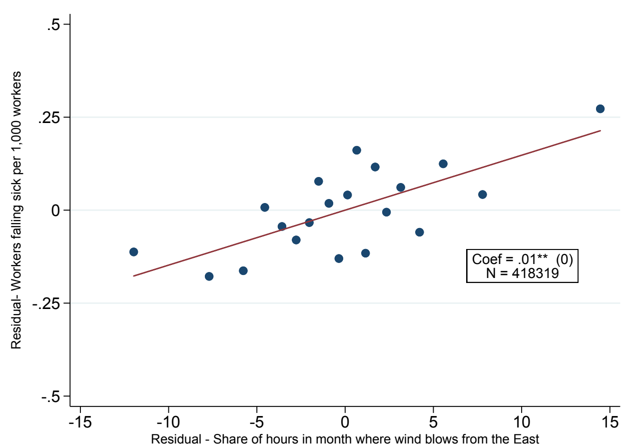


Figure 3: Binned scatter plot of work absenteeism and particulate matter pollution

Notes: The data is split into equal-sized bins based on the value of the x-variable – monthly $PM_{2.5}$ concentrations. Each data point shows the mean residual of that bin for the x-axis and y-axis variables after controlling for postcode and month-by-year fixed effects, as well as all weather variables (and their interactions) and flu incidence and holidays. Observations are weighted by the number of workers in each postcode. The graph is done using the stata command *binscatterhist* by Pinna (2022).



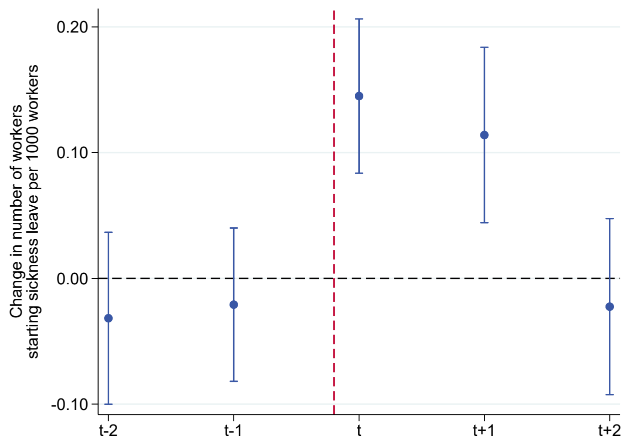
(a) Wind from West



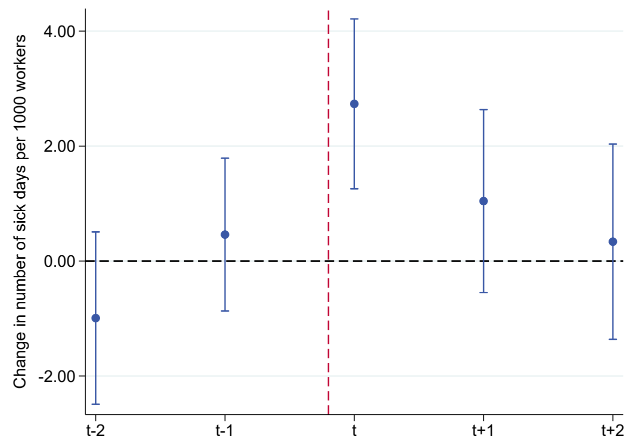
(b) Wind from East

Figure 4: Binned scatter plot of work absenteeism and monthly wind direction

Notes: The data is split into equal-sized bins based on the value of the x-variable – the share of hours in a month where wind blows from wind direction Est (Figure a) and West (Figure b). Each data point shows the mean residual of that bin for the x-axis and y-axis variables after controlling for flu incidence and the number of holiday days in the *departement*; for a flexible weather controls including interactions between categories of average daily maximum temperatures, quintiles of average daily wind speed and quintiles of total daily precipitations; grid cell level fixed effects and month by year fixed effects. Observations are weighted by the number of workers in each postcode. The graphs are done using the stata command *binscatterhist* by Pinna (2022).



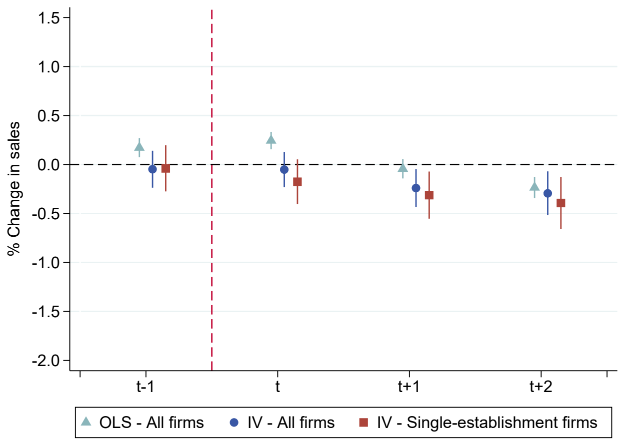
(a) Effects on work absenteeism



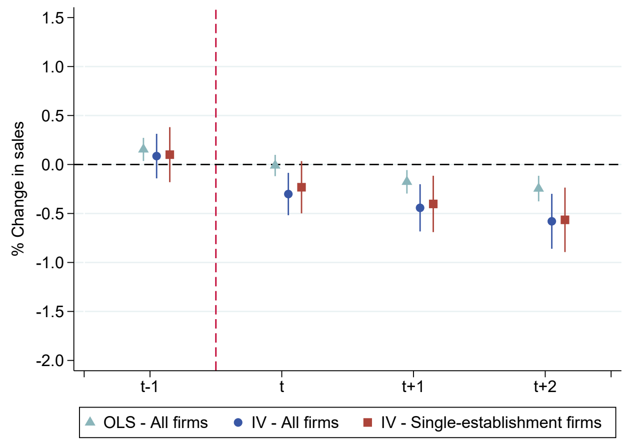
(b) Effects on the number of sick days

Figure 5: Dynamic effects of $PM_{2.5}$ on work absenteeism and on the number of sick days per 1,000 workers

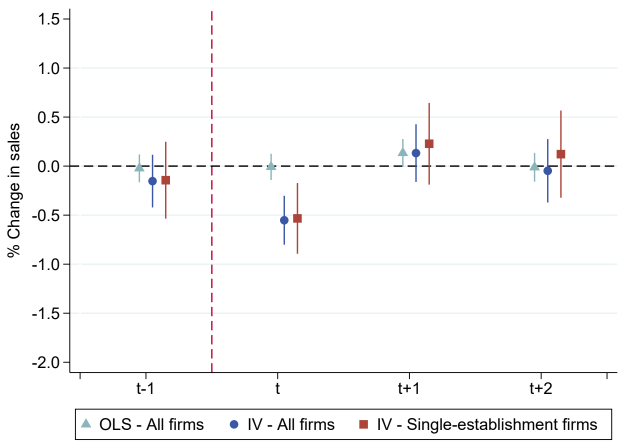
Notes: Figure shows the point estimates and 95% confidence intervals for the effect of $PM_{2.5}$ measured at t on sickness leave at $t-2$, $t-1$, t, $t+1$ and $t+2$. All regressions include month-by-year fixed effects, postcode fixed effects, weather controls, and holidays and flu controls at t. For the effects on sickness leave at $t+1$ and $t+2$, controls for weather, holiday and flu and the instrument at the relevant leads are added. For the effects on sickness leave at $t-11$ and $t-2$, controls for weather, holiday and flu at the relevant lags are added. Observations are weighted by the number of workers in each postcode. The confidence intervals are based on standard errors clustered at the Copernicus grid cell level.



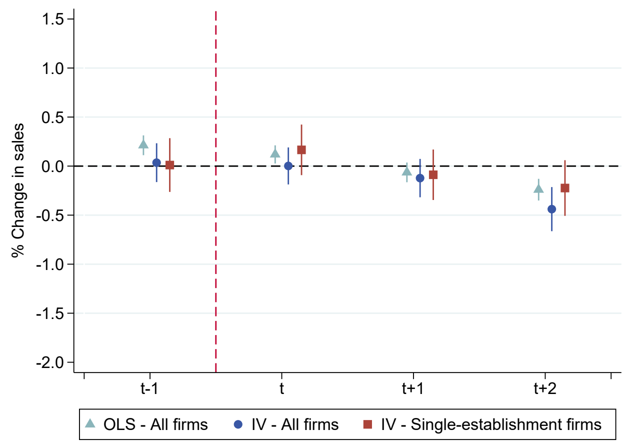
(a) Manufacturing



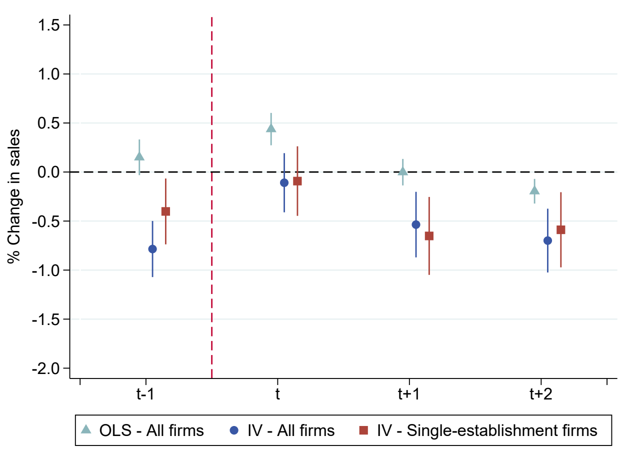
(b) Construction



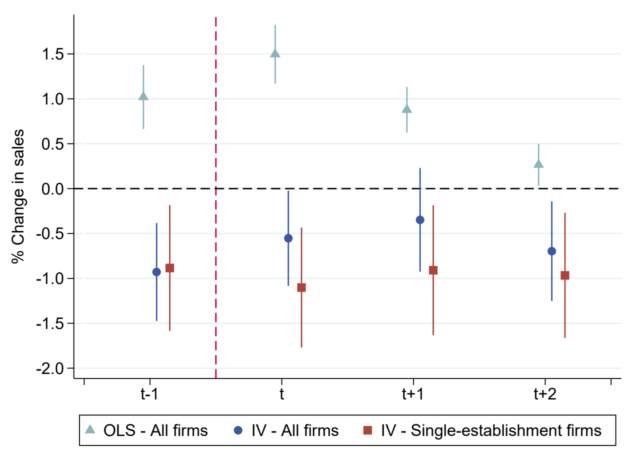
(c) Professional services



(d) ICT and other B-to-B services



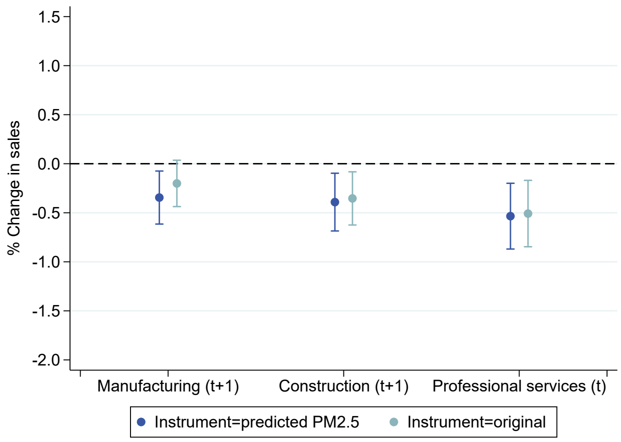
(e) Retail and restaurants



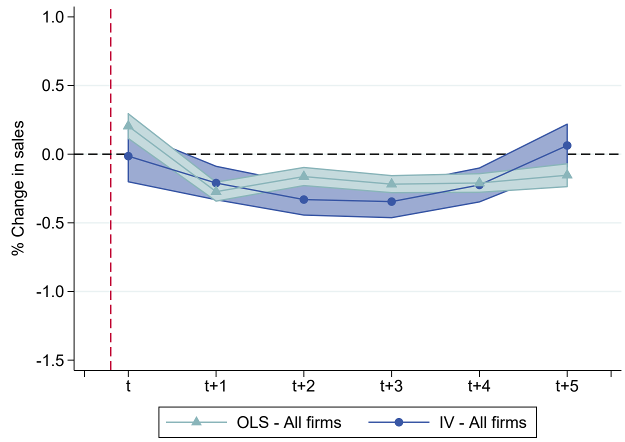
(f) Other business-to-consumer services

Figure 6: Dynamic effects of $PM_{2.5}$ on firms' sales, by sector

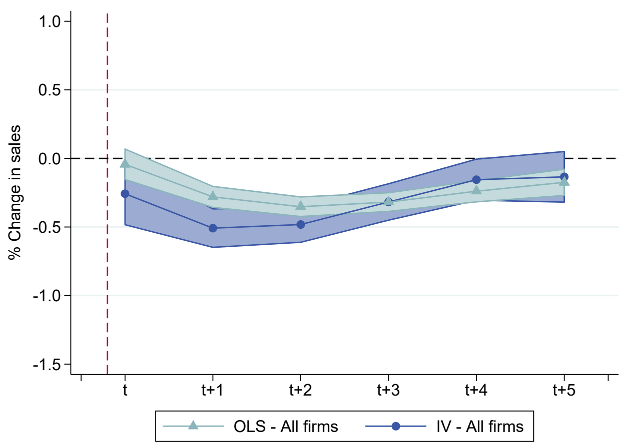
Notes: Figure shows the OLS and IV point estimates and 95% confidence intervals from equation (7) for the effect of $PM_{2.5}$ on t on firms' sales by sector at $t-1$, t, $t+1$ and $t+2$. All regressions include month-by-year-by-industry fixed effects, firm-by-year fixed effects, weather controls, and holidays and flu controls. For the leads, controls for weather, holiday and flu at the relevant leads and lags are added, and both PM25 at t (for which the coefficient is reported) and $PM_{2.5}$ for the relevant leads are instrumented. For the placebo test at $t-1$, controls for weather, holiday and flu at $t-1$ are added. The confidence intervals are based on standard errors clustered at the firm level.



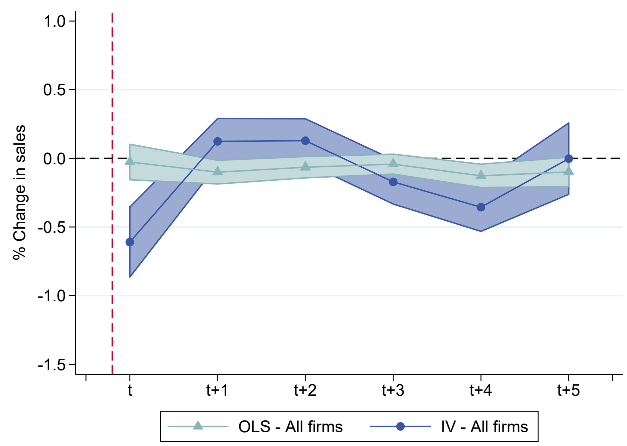
(a) Comparing the predicted $PM_{2.5}$ instrument and the wind direction instrument for single-plant firms



(b) PDL - Manufacturing



(c) PDL - Construction



(d) PDL - Professional services

Figure 7: Robustness checks for the effects of $PM_{2.5}$ on firms' sales, by sector

Notes: Panel a) compares, for the subsample of single-establishment firms, two point estimates and their confidence intervals for the effect of pollution at t on manufacturing and construction sales at $t+1$, and sales in professional services at t: the point estimate obtained with the predicted pollution instrument (main result), and the point estimate obtained with the wind direction instrument. Panels b), c) and d) show the OLS and IV point estimates and 95% confidence intervals for the effect of contemporaneous $PM_{2.5}$ (t) and five lags of $PM_{2.5}$ on firms' sales at t using a polynomial distributed lag (PDL) model with a cubic polynomial specification. All regressions include month-by-year-by-industry fixed effects, firm-by-year fixed effects, weather controls, and holidays and flu controls. The confidence intervals are based on standard errors clustered at the firm level.

10 Tables

Table 1: Summary Statistics, 2009-2015

	Mean	Sd	Count
Panel a: Postcode-level summary statistics for Workers			
Age	40.43	2.17	393,756
Annual wage (euros €)	25,909.84	6,569.49	393,756
Annual total medical expenditures (€)	462.51	132.18	393,756
Annual total out-of-the-pocket expenditures (€)	143.18	39.29	393,756
Monthly exposure to $PM_{2.5}$ (µg/m^3)	15.37	6.32	393,756
Nb workers falling sick each month, per 1,000 workers	23.85	20.42	393,756
incl: for <93 days	22.13	19.48	393,756
Nb of associated sick days per 1,000 workers	758.77	1,605.93	393,756
incl: for <93 days	354.94	448.23	393,756
Sickness leave spending per 1,000 workers (€)	21,477.44	51,624.77	393,756
incl: for <93 days	9,099.00	12,769.65	393,756
Panel b: Firm-level summary statistics			
Share of single-establishment firms	0.63	0.48	10,991,760
Number of workers	65.06	888.24	10,991,760
Share in: Manufacturing	0.17	0.37	10,991,760
Construction	0.14	0.35	10,991,760
Retail and Restaurants	0.21	0.41	10,991,760
Other business to consumer services	0.11	0.31	10,991,760
Professional services	0.16	0.36	10,991,760
ICT and other business to business services	0.21	0.42	10,991,760
Monthly exposure to $PM_{2.5}$ (µg/m^3)	15.22	6.23	10,991,760
Monthly sales (€)	1,311,465.43	18,385,016.90	10,483,016
Monthly sales (€) in: Manufacturing	2,386,087.68	27,111,797	1,790,571
Construction	532,055.97	3,140,304	1,464,654
Retail and Restaurants	946,355.82	16,002,400	2,216,708
Other business to consumer services	545,468.55	13,594,881	1,106,004
Professional services	754,394.65	10,944,304	1,646,831
ICT and other business to business services	2,067,678.96	23,498,297	2,258,248

Table 2: Pollution and Sick Leave Episodes

	Nb. of workers falling sick		Nb. of sick days		Sickness leave spending	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel a: Postcode Fixed Effects						
$PM_{2.5}$	0.0659***	0.146***	0.811**	2.734***	26.54***	86.54***
	(0.0206)	(0.0313)	(0.339)	(0.754)	(9.779)	(22.46)
Kleibergen Paap F-statistic		815		815		815
N	382,187	382,187	382,187	382,187	382,187	382,187
R-squared	0.23	0.23	0.10	0.10	0.08	0.08
Panel b: Postcode-Year Fixed Effects						
$PM_{2.5}$	0.0674***	0.167***	0.676**	2.594***	20.87**	78.23***
	(0.0207)	(0.0316)	(0.337)	(0.749)	(9.725)	(22.47)
Kleibergen Paap F-statistic		637		637		637
N	382,187	382,187	382,187	382,187	382,187	382,187
R-squared	0.30	0.30	0.18	0.18	0.16	0.16
Dep. var. mean	22	22	355	355	9,099	9,099

Notes: Table reports OLS and IV estimates from equation (5) for the effect of $PM_{2.5}$ on the number of workers taking a sick leave in a postcode (columns 1 and 2), on the number of sick days associated with this leave (columns 3 and 4), and on the sickness leave spending (columns 5 and 6). All regressions include month-by-year fixed effects, weather controls, and holidays and flu controls. Panel a also includes postcode fixed effects, while Panel b includes postcode-by-year fixed effects. Observations are weighted by the number of workers in each postcode. Standard errors in parentheses are clustered at the Copernicus grid cell level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 3: Pollution and Sickness Leave Episodes, Varying Unit and Time Dimensions

	Nb. of workers falling sick		Nb. of sick days		Sickness leave spending	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel a: 10% Sample and Individual Fixed Effects						
$PM_{2.5}$	0.106**	0.256***	0.912	5.024**	19.96	159.0**
	(0.0439)	(0.0912)	(1.037)	(2.217)	(28.98)	(66.89)
N	2,531,172	2,531,172	2,531,172	2,531,172	2,531,172	2,531,172
R-squared	0.0525	0.0525	0.0301	0.0301	0.0257	0.0257
Panel b: Weekly Outcomes and Pollution						
$PM_{2.5}$	0.00459***	0.0132***	0.0509	0.391***	1.448	12.86***
	(0.00160)	(0.00376)	(0.0306)	(0.0776)	(0.904)	(2.233)
Dependant variable mean	5.2	5.2	85	85	2,079	2,079
N	1,711,430	1,711,430	1,711,430	1,711,430	1,711,430	1,711,430
R-squared	0.0645	0.0644	0.0226	0.0225	0.0178	0.0177

Notes: Table reports OLS and IV estimates from equation (5) for the effect of $PM_{2.5}$ on the number of workers taking a sick leave in a postcode (columns 1 and 2), on the number of sick days associated with this leave (columns 3 and 4), and on the sickness leave spending (columns 5 and 6). Panel a reports estimates from a 10% sample of Hygie using worker fixed effects instead of postcode fixed effects. Panel b reports estimates when outcomes and all controls are defined at the weekly level and using quarter-by-year fixed effects instead of month-by-year fixed effects.Observations are weighted by the number of workers in each postcode. Standard errors in parentheses are clustered at the Copernicus grid cell level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 4: Pollution and Sickness Leave Episodes, the exclusion restriction and the role of Ozone

	(1) Baseline	(2) O3 only	(3) AQI index
$PM_{2.5}$ exposure	0.145***		
	(0.0313)		
Ozone exposure		0.0342**	
		(0.0171)	
Air quality index (higher: worse air quality)			2.922***
			(0.495)
N	393,755	393,755	393,755
R-squared	0.2299	0.2300	0.2300

Notes: Table reports IV estimates from equation (5) for the effect of $PM_{2.5}$ on the number of workers taking a sick leave in a postcode using the baseline specification (column 1), for the effect of ozone only (column 2), and considering the effect of a one-unit increase in the French air quality index ranging from 1 (best air quality) to 6 (worst air quality), instead of a one-unit increase in $PM_{2.5}$ (column 3). All regressions include month-by-year fixed effects, weather controls, holidays and flu controls, and postcode fixed effects. Observations are weighted by the number of workers in each postcode. Standard errors in parentheses are clustered at the Copernicus grid cell level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 5: Pollution and Sickness Leave Episodes, Various Robustness Checks

	(1) Baseline	(2) Cont. weather	(3) 50km grid	(4) Quarter -specific Interactions	(5) No dec & jan	(6) No month with PM10 alerts
$PM_{2.5}$ exposure	0.145***	0.193***	0.138***	0.161***	0.166***	0.180***
	(0.0313)	(0.0398)	(0.0307)	(0.0316)	(0.0349)	(0.0361)
Holiday and flu controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	No	Yes	Yes	Yes	Yes
Month by Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Postcode FE	Yes	Yes	Yes	Yes	Yes	Yes
N	393,755	393,756	393,755	393,755	328,130	384,276
R-squared	0.2299	0.2287	0.2300	0.2299	0.1603	0.2186

Notes: Table reports IV estimates from equation (5) for the effect of $PM_{2.5}$ on the number of SLEs for 1,000 workers using the baseline specification (column 1), using three weather controls for average daily maximum temperature, total monthly precipitation, and average wind speed, instead of the interaction of categorical weather variables (column 2), using smaller grid cell areas of 50 km x 50 km (column 3), allowing the influence of wind direction to vary not only by grid cell area but also by quarter (column 4), removing January and December (column 5), and removing the postcode-months where at least one day has PM_{10} levels corresponding to an air quality alert level 1. All regressions include month-by-year fixed effects, weather controls, holidays and flu controls, and postcode fixed effects. Observations are weighted by the number of workers in each postcode. Standard errors in parentheses are clustered at the Copernicus grid cell level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 6: Pollution and Sales, Various Robustness Checks

	(1) Baseline	(2) No 2009	(3) Cont. weather	(4) Firm FE	(5) No month with PM10 alerts	(6) Two-way clustering
Panel a: Manufacturing – Effect at $t+1$						
$PM_{2.5}$ exposure	-0.00240**	-0.00245**	-0.00335***	-0.00426**	-0.00259**	-0.00240
	(0.000989)	(0.00107)	(0.00102)	(0.00193)	(0.00110)	(0.00162)
N	1,726,477	1,472,433	1,726,477	1,728,519	1,706,923	1,726,477
R-squared	0.8268	0.8285	0.8267	0.7192	0.8262	0.8262
Panel b: Construction – Effect at $t+1$						
$PM_{2.5}$ exposure	-0.00442***	-0.00490***	-0.00540***	-0.00809***	-0.00498***	-0.00442**
	(0.00123)	(0.00133)	(0.00121)	(0.00180)	(0.00138)	(0.00189)
N	1,409,093	1,205,874	1,409,093	1,410,660	1,346,239	1,409,093
R-squared	0.7565	0.7533	0.7563	0.6119	0.7562	0.7565
Panel c: Professional services – Effect at t						
$PM_{2.5}$ exposure	-0.00552***	-0.00520***	-0.00527***	-0.00849***	-0.00612***	-0.00552***
	(0.00127)	(0.00132)	(0.00132)	(0.00250)	(0.00147)	(0.00159)
N	1,636,769	1,414,198	1,636,771	1,637,481	1,636,769	1,525,976
R-squared	0.7184	0.7173	0.7173	0.5623	0.7186	0.7186
Weather controls	Yes	Yes	No	Yes	Yes	Yes
Holidays and Flu	Yes	Yes	Yes	Yes	Yes	Yes
Month-Year-Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Year FE	Yes	Yes	Yes	No	Yes	Yes
Firm FE	No	No	No	Yes	No	No

Notes: Table reports IV estimates from equation (7) for the effect of $PM_{2.5}$ on firms' sales in manufacturing (panel a), construction (panel b), and professional services (panel c). Column (1) reports the baseline result, column (2) uses a sample that excludes 2009, column (3) uses continuous weather variables as controls, column (4) uses firm fixed effects instead of firm-year fixed effects, column (5) discards months with at least one day of PM_{10} alert, and column (6) reports standard errors in parentheses clustered at the firm and month x year level instead of the firm level. Standard errors in parentheses are clustered at the firm level, except in column (6). *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table 7: Pollution and Sales, the exclusion restriction and the role of Ozone

	(1)	(2)	(4)
	Baseline	O3 only	AQI index
Panel a: Manufacturing – Effect at $t+1$			
$PM_{2.5}$ exposure	-0.00240**		
	(0.000989)		
Ozone exposure		-0.00055	
		(0.00089)	
Air quality index (higher: worse air quality)			-0.0131
			(0.0164)
N	1,726,477	1,726,477	1,706,923
R-squared	0.8268	0.8267	0.8267
Panel b: Construction – Effect at $t+1$			
$PM_{2.5}$ exposure	-0.00442***		
	(0.00123)		
Ozone exposure		0.00069	
		(0.00107)	
AQI index			-0.0422**
			(0.0189)
N	1,409,093	1,409,093	1,393,402
R-squared	0.7565	0.7562	0.7562
Panel c: Professional services – Effect at t			
$PM_{2.5}$ exposure	-0.00552***		
	(0.00127)		
Ozone exposure		0.00135	
		(0.000866)	
AQI index			-0.0603***
			(0.0204)
N	1,636,769	1,636,769	1,622,131
R-squared	0.7184	0.7183	0.7182

Notes: Table reports IV estimates from equation (7) for the effect of $PM_{2.5}$ on firms' sales in manufacturing (panel a), construction (panel b), and professional services (panel c). Column (1) reports the baseline result, column (2) reports results when ozone is instrumented and column (3) reports the estimated effect of an increase by one unit of the multi-pollutant air quality index, ranging from 1 (best air quality) to 6 (worst air quality). Standard errors in parentheses are clustered at the firm level, except in column (5). *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Appendix

A Data Appendix

A.1 Sickness Leave Episodes

We obtain data on sickness leave episodes (SLE) from the Hygie dataset, which follows roughly 900,000 employees during the period 2009-2015. To build our sample of analysis, we make three restrictions. First, we only keep individuals to whom we are able to assign a place of work based on the establishment's unique identifier. This makes us discard individuals with no employment history declared between 2009 and 2015, who represent 25% of the sample. Although we cannot check the exact reason for missing information, these individuals are probably retired, unemployed or out of the labour force over the whole period. Two-thirds of them should be retired in 2009 given their age. We also discard individuals for whom we do not have an establishment identifier despite the fact that they did work and contribute to the pension system over the 2009-2015 period, who represent 6% of the sample. Two third of these individuals have zero employers declared over the period. They may have switched to the public sector or to the agricultural sector or started their own business, or they may work in the domestic care sector, where there is no establishment-level identifier (since they are employed by private individuals).

Second, we discard individuals whose establishment identifier corresponds to a public institution such as hospital or schools, because we want to focus the analysis on private sector employees. Some individuals working in these institutions have a private sector type of contract and are thus eligible to enter the Hygie sample.

Third, we discard a few individuals who did not work enough to contribute to the public pension system for any of the years included in the period. Each year, these individuals worked less than 150 equivalent hours valued at the minimum wage per year, which is the minimum to contribute to public pension. With such a low labour supply, they are unlikely to experience sickness leave episodes.

We assign each worker to the postcode of her workplace (there are around 6,000 postcodes in France). Figure A.1 shows the geographic distribution of the employees' workplaces in 2009, which is consistent with the distribution of the French population across the territory.

We use the exhaustive matched employer-employee data (DADS-Postes) to compare the characteristics of our sample of workers to the characteristics of the whole population of private sector employees. Applying the same restrictions as in the Hygie dataset,[37] we find that those workers

[37] Namely, we keep private sector employees born between 1935 and 1989, less those older than 71 who should be retired. Note that in the matched employer-employee data, a worker having two different employers appears twice. We aggregate wage information at the worker level, summing up the wages she receives from different employers.

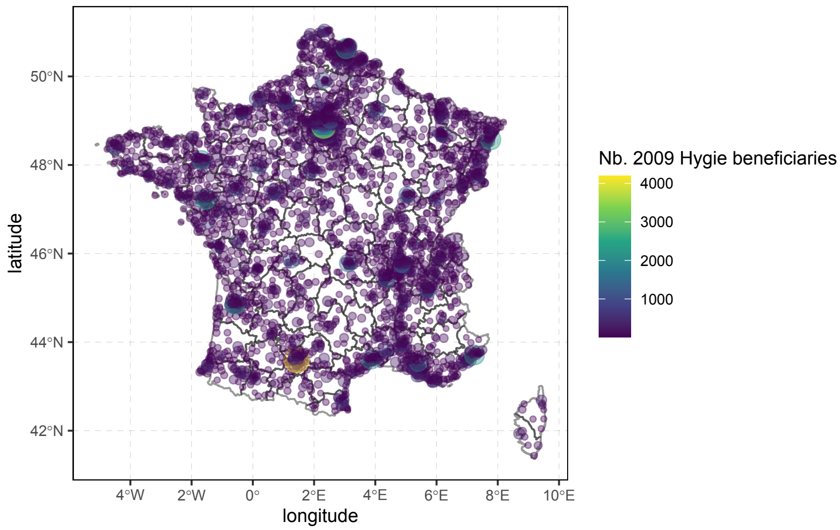


Figure A.1: Location of workers from the Hygie dataset based on the workplace postcode, in 2009

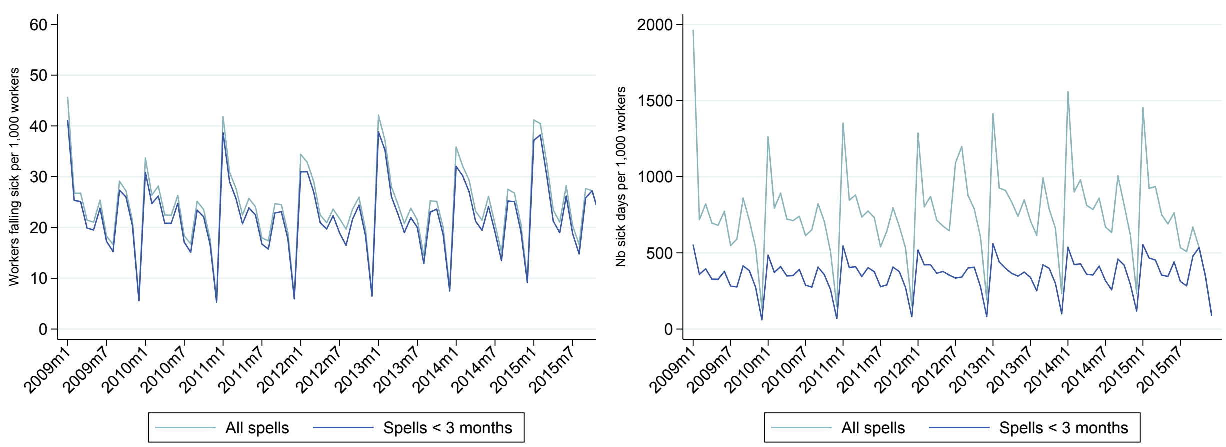
representing the population from which our sample is drawn are 55% male, 41 on average, and earn an average annual gross wage of €26,204. Thus, the average individual in our final worker sample – as shown in Table A.1 – is very close to the average private sector employee.

In our sample, 21 percent of employees take at least one sickness leave episode within a year. By comparison, a national survey on Working Conditions estimated that 28 percent of private sector employees in France took at least one sick leave during 2013.[38]

[38]Source: https://www.fonction-publique.gouv.fr/files/files/statistiques/rapports_annuels/2015/RA2015_dossier_1.pdf

Table A.1: Summary statistics at the worker (top) and sickness leave episode (bottom) level, 2009-2015

	Mean	Sd	Median	Count
Share men	0.55	0.50		2,395,035
Age	40.43	10.98		2,395,035
Annual wage	25,907.65	26,176.07		2,395,035
Share full-time employed	0.75	0.43		2,395,035
Works in a single-establishment firm	0.38	0.49		2,395,035
Works in: Manufacturing	0.17	0.37		2,395,035
Construction	0.07	0.26		2,395,035
Retail and Restaurants	0.12	0.33		2,395,035
Other business to consumer services	0.07	0.25		2,395,035
Professional services	0.17	0.37		2,395,035
Financial services	0.04	0.19		2,395,035
ICT and other business to business services	0.15	0.36		2,395,035
Health, education and charitable organizations	0.14	0.35		2,395,035
Other	0.07	0.25		2,395,035
Sick at least once	0.21	0.41		2,395,035
Nb. sickness leave episodes in a year	0.29	0.68		2,395,035
SLE duration (days)	29	69	9	753,522
SLE benefits (publicly-funded component) (€)	808	2,291	183	753,522



(a) Number of workers falling sick (b) Number of sick days

Figure A.2: Number of workers falling sick and number of sick days per 1,000 workers

Notes: Figure presents the average number of workers falling sick and average number of sick days per 1,000 workers over time. While the spells larger than 3 months represent a small proportion of total spells, their tend to strongly increase the average number of sick days.

A.2 Firm-Level Sales

We compute firms' monthly sales by adding up different components included in the VAT records, following the methodology of France Stratégie and Inspection générale des Finances (2021). In the raw data, total sales are broken down into different components based on two main criteria that determine VAT liability: the location of the buyer (whether in France, in another EU country, or in a non EU-country) and whether the buyer is herself liable to VAT. In addition, the sales value of goods and services subject to specific tax rules is reported separately.[39] Our measure of sales includes both domestic sales and exports to EU and non-EU countries. The French tax administration imposes monthly declarations to firms with annual sales above €818,000 for the manufacturing sector and the hospitality industry and to those with annual sales above €247,000 for the other sectors. Firms below this threshold are allowed to fill declarations on a quarterly basis.

We discard the entire firm-year series for firms not reporting sales each month within a year. However, we make one exception for zero sales records in July since it is a relatively common pattern in the data. A large number of French firms close for vacation during some weeks in August, the month where the July VAT declaration is expected since the VAT declaration corresponding to the business month t is typically made on month $t + 1$. French tax authorities allow firms to report their July sales together with the August sales.[40] We indeed observe in the data that when the sales are 0 in July, the sales for August are frequently twice as high as those in June or September. We re-allocate sales for July and August by splitting August sales in two. After implementing this correction for zero sales in July followed by high sales in August, we discard around 3% of firm-year observations with at least one zero sales value.

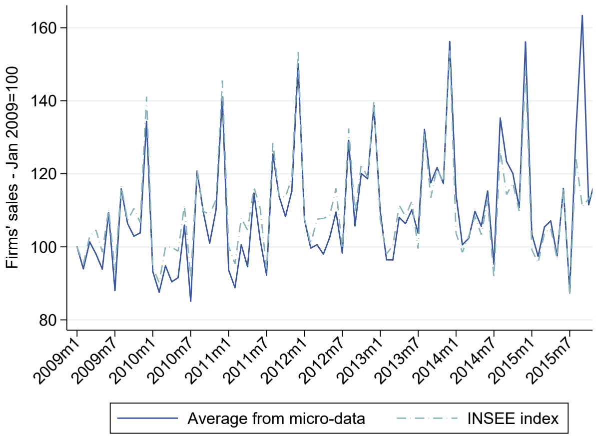
We determine sectors of activity based on the sectoral classification available at the establishment level and we use the mode of sector categories across establishments for multi-establishment firms. Following a methodology developed by the statistical agency INSEE (https://www.insee.fr/fr/statistiques/1372801?sommaire=1372813), we define 6 sectors: manufacturing, construction, retail and restaurants, other business-to-consumer services (including, hospitality industry, passenger transport, real estate, travel agencies, recreational services, repairing services), professional services, and ICT and other business-to-business services. We discard firms belonging to the financial services sector and the health, education and charitable sectors, which are often not-for-profit.

We check the quality of the reported data in two different ways. First, for a few large French

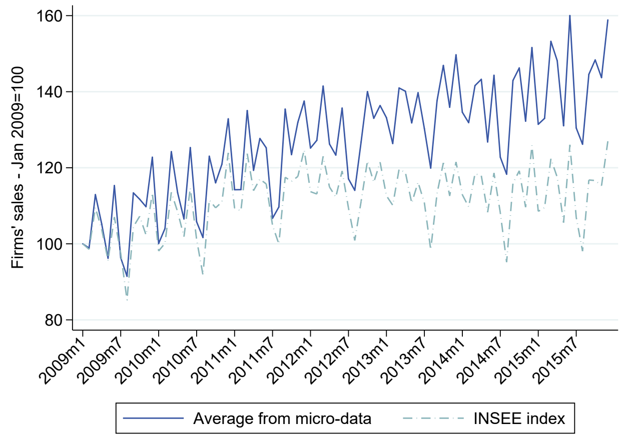
[39] For instance, the sales of natural gas and electricity is subject to a specific VAT rule in the French tax code, so they have their own subcomponent in the VAT records. See https://www.impots.gouv.fr/sites/default/files/formulaires/3310-ca3-sd/2022/3310-ca3-sd_3947.pdf

[40] See https://www.impots.gouv.fr/professionnel/questions/comment-declarer-ma-tva-en-periode-de-conges-pa

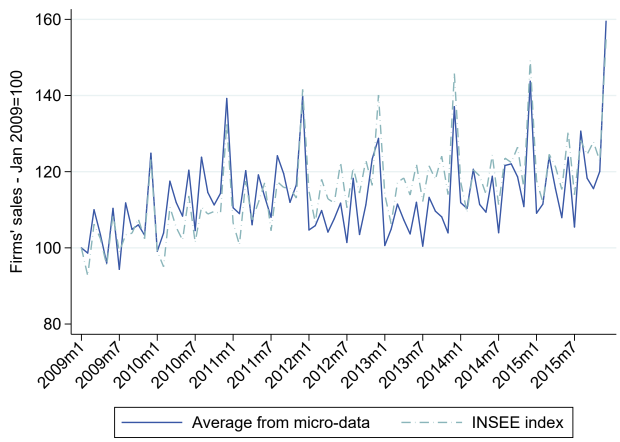
companies for which annual financial reports are publicly available, we manually check that the sum of monthly sales of a given year is close to the official annual sales value. Second, we compare the time series of monthly sales value aggregated by economic sector to the data published at the industry level by the French statistical institute, using the same source. Figure A.3c shows the time series of monthly sales in construction (A.3a), manufacturing (A.3b) and all services (A.3c) as constructed from the VAT micro-data compared with the INSEE index. Differences may arise between our sales value and the statistical agency's because of different choices in data cleaning or the subcomponents entering the sales variable, but the correlation between the two series are above 0.9 for the three broad sectors.



(a) Construction



(b) Manufacturing



(c) Services

Figure A.3: Average firms' nominal sales in construction, manufacturing and service sector, 2009=100

Notes: Figure presents the average nominal sales from our VAT micro-data in blue for construction, manufacturing, and services and the INSEE sales index in dashed green, using January 2009 as the reference point. We exclude several service industries (trade - sector G in NACE classification, banking - sector K and health - sector Q) to compare with the INSEE index which also excludes these industries.

B Additional Figures

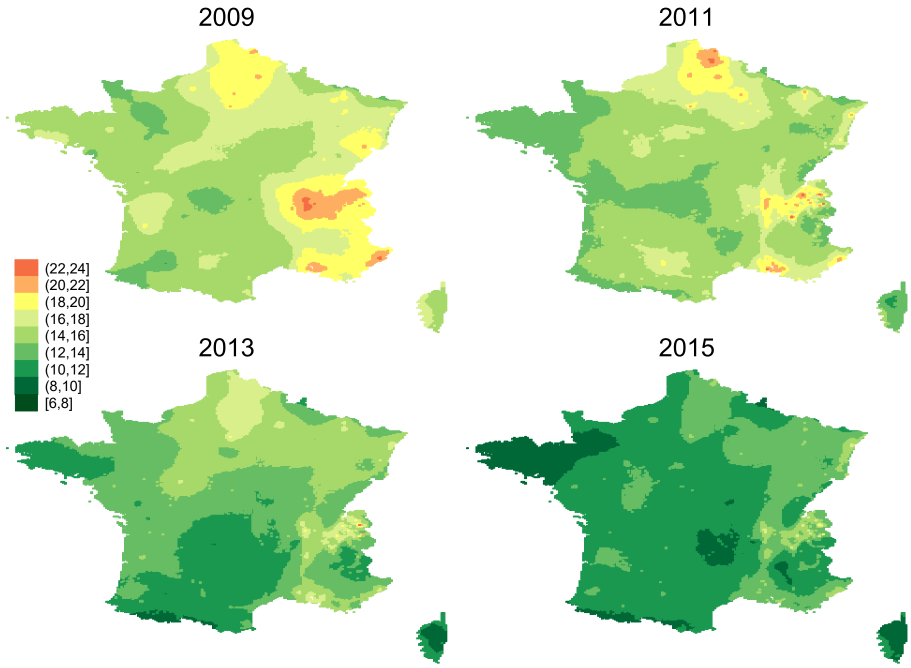
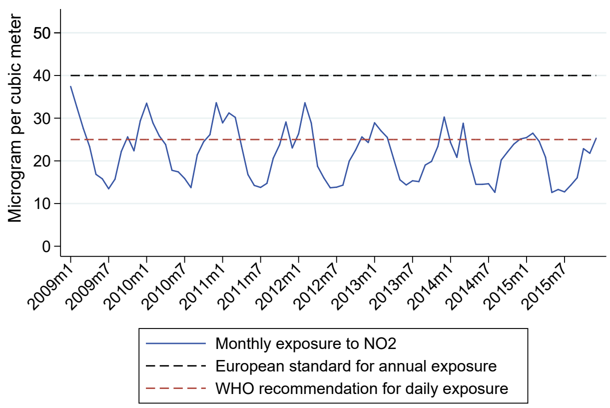
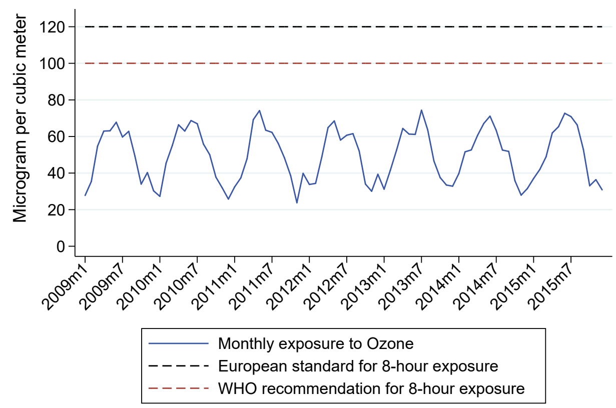


Figure B.4: Average annual concentrations of $PM_{2.5}$ (µg/m^3)

Notes: Figure shows the average annual concentration of $PM_{2.5}$ measured at the 4km x 4 km grid cell level using the reanalysis CHIMERE data, for selected years. There are 33,252 Chimere grid cells in metropolitan France.



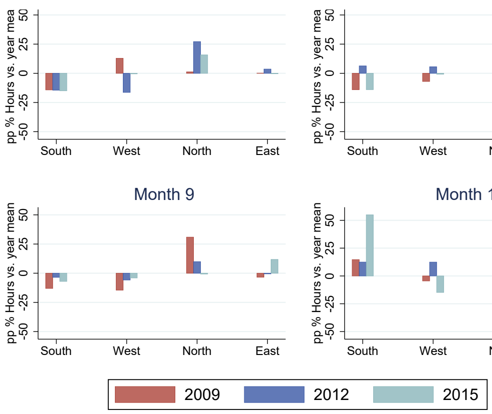
(a) NO_2



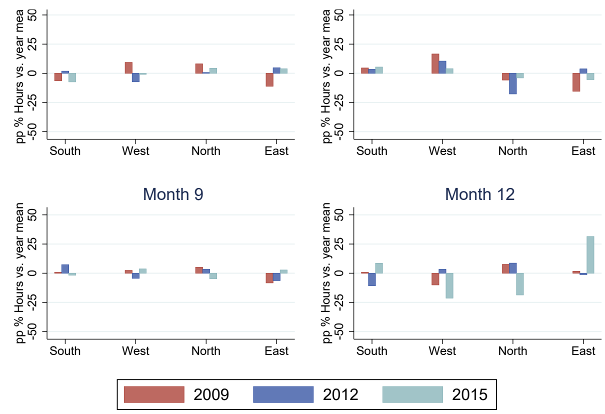
(b) Ozone

Figure B.5: Average monthly exposure to other pollutants

Notes: Figure presents the monthly average of workers' exposure to $PM_{2.5}$ measured at workers' postcodes. The sample of workers is the one used for the analysis of pollution effects on sickness leaves described in section 3.3 (unbalanced panel, N≈450,000). For NO_2, the European standard for annual exposure is 40µg/m^3 while the WHO's recommendation for daily exposure is 25µg/m^3. For ozone, the European standard for 8-hour exposure is 120µg/m^3 while the WHO's recommendation for 8-hour exposure is 100µg/m^3. Exposure in each postcode is weighted by the number of workers working in that postcode.



(a) Paris



(b) Marseille

(c) Center of France

Figure B.6: Variation in monthly wind direction within local areas

Notes: Figure shows for each local area (Paris/Marseille/countryside in the center of France) the share of hours in a month in which the wind blows from a given direction, demeaned by the average for the year, for four different months (Month 3=March, Month 6=June, Month 9=September, Month 12=December and three different years (2009, 2012, 2015).

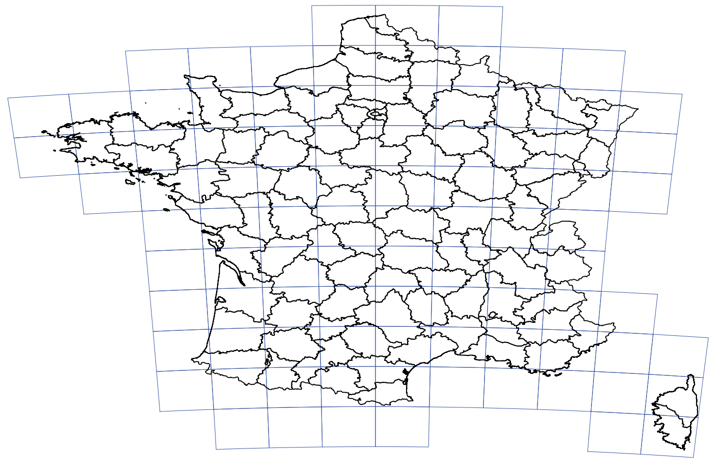


Figure B.7: 100 x 100 km grid and French *departements*

Notes: The figure shows the 100 x 100 km areas k used in equation (6 in blue, and the borders of French *departements* in black

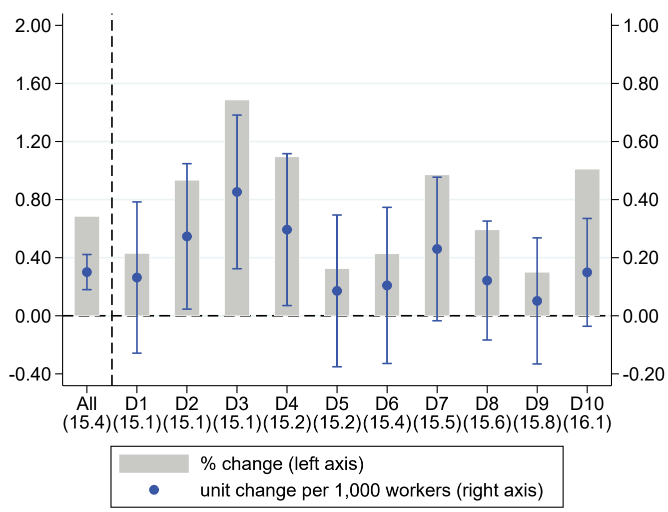


Figure B.8: Effect of pollution on sickness leaves by wage level

Notes: Figure shows the point estimates and 95% confidence intervals for the effect of $PM_{2.5}$ on sickness leave by wage decile. We run equations (5) and (6) for 10 postcode-level datasets that only include workers from a given wage decile in 2009. Number of observations varies between 274,067 (D1) and 172,391 (D10). On the x-axis, we report the average monthly $PM_{2.5}$ exposure below the relevant subsample. Observations are weighted by the number of workers in each postcode. The confidence intervals are based on standard errors clustered at the Copernicus grid cell level.

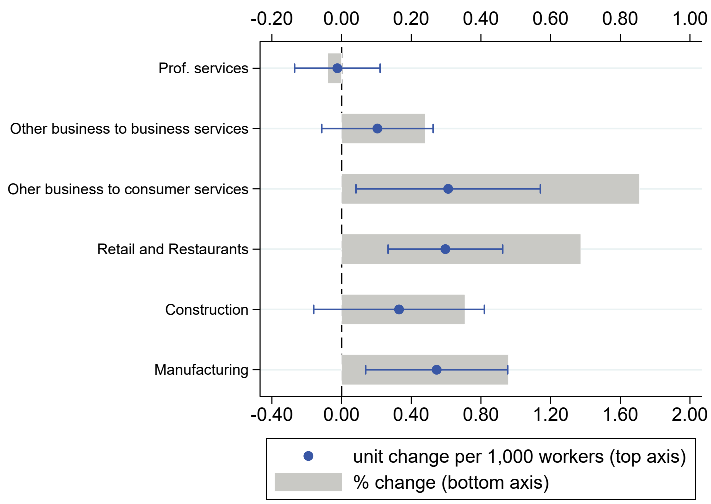


Figure B.9: Effect of pollution on sickness leaves by sector

Notes: Figure shows the point estimates and 95% confidence intervals for the effect of $PM_{2.5}$ on sickness leave by sector. We run equations (5) and (6) with individual establishment-months as the unit of observation instead of postcode-months. Establishment fixed effects replace postcode fixed effects. Absenteeism is measured as the number of workers starting a sickness leave on a given month in that establishment, per 1,000 workers. The regressions are weighted by the number of workers for whom we observe absenteeism in each establishment. The regressions are run separately for 6 subsamples of establishments, that only include establishments part of firms from a given sector. Number of observations varies between 1,413,286 (construction) and 2,805,238 (retail). The confidence intervals are based on standard errors clustered at the Copernicus grid cell level.

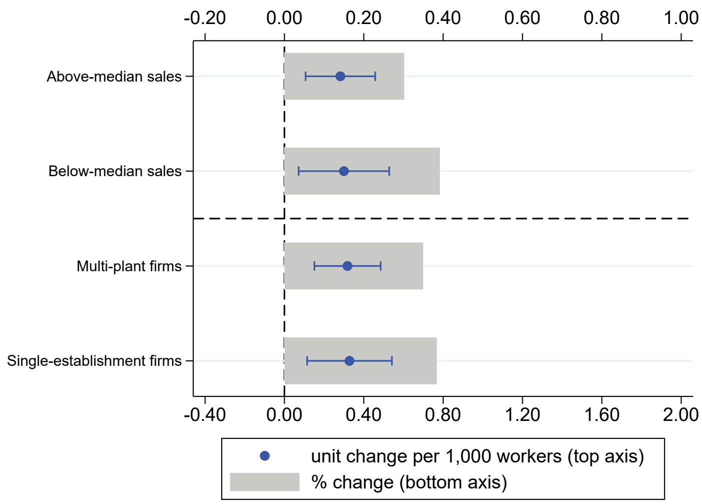


Figure B.10: Effect of pollution on sickness leaves by firms' size and status

Notes: Figure shows the point estimates and 95% confidence intervals for the effect of $PM_{2.5}$ on sickness leave by firm type. We run equations (5) and (6) with individual establishment-months as the unit of observation instead of postcode-months. Establishment fixed effects replace postcode fixed effects. Absenteeism is measured as the number of workers starting a sickness leave on a given month in that establishment, per 1,000 workers. The regressions are weighted by the number of workers for whom we observe absenteeism in each establishment. The regressions are run separately for 2 subsamples of establishments for each heterogeneity dimension: establishments part of firms with above-median vs. below-median sales (above), and establishments part of multi-plant vs. single-establishment firms (below). Number of observations varies between 3,293,614 (establishments part of firms with below median sales in 2012) and 5,855,816 (establishments part of multi-plant firms). The confidence intervals are based on standard errors clustered at the Copernicus grid cell level.

C Additional Tables

Table C.2: Effect of pollution on sales by firm size (1/2)

	All firms (t)		Manufacturing ($t+1$)		Construction ($t+1$)		Professional services (t)	
	Below median	Above median	Below median	Above median	Below median	Above median	Below median	Above median
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$PM_{2.5}$ exposure	-0.00095	-0.00121*	-0.00326**	-0.00108	-0.00567***	-0.00412***	-0.00286	-0.00348*
	(0.00101)	(0.00068)	(0.00154)	(0.00125)	(0.00203)	(0.00149)	(0.00176)	(0.00194)
N	4,534,726	4,861,799	780,574	824,185	599,272	659,826	697,766	739,016
R-squared	0.69	0.67	0.71	0.73	0.70	0.62	0.69	0.59

Notes: Table reports IV estimates from equations (7) (columns (1)-(2) and (7)-(8)) and (**??**) (columns (3)-(6)) for the effect of $PM_{2.5}$ on firms' sales by sector. Regressions are run separately for each sector and each sector is split in two subsamples: one for firms with below-median sales in 2012 and one for firms with above-median sales in 2012. All regressions include month-by-year fixed effects, weather controls, and holidays and flu controls. Standard errors in parentheses are clustered at the Firm level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.

Table C.3: Effect of pollution on sales by firm size (2/2)

	ICT and other BtoB services (t)		Retail and Restaurants		Other consumer services	
	Below median	Above median	Below median	Above median	Below median	Above median
	(1)	(2)	(3)	(4)	(5)	(6)
$PM_{2.5}$ exposure	-0.00033	-0.00033	-0.00199	0.00248*	0.00483	-0.0128***
	(0.00141)	(0.00137)	(0.00283)	(0.00142)	(0.00159)	(0.00170)
N	997,506	1,041,220	926,620	1,020,672	461,249	523,473
R-squared	0.70	0.66	0.71	0.69	0.69	0.54

Notes: Table reports IV estimates from equation (7) for the effect of $PM_{2.5}$ on firms' sales by sector. Regressions are run separately for each sector and each sector is split in two subsamples: one for firms with below-median sales in 2012 and one for firms with above-median sales in 2012. All regressions include firm-year and month-by-year fixed effects, weather controls, and holidays and flu controls. Standard errors in parentheses are clustered at the firm level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$.