Master Project - Barcelona GSE

Does Air Pollution Exacerbate Covid-19 Symptoms? Evidence from France

Hubert Massoni[1]

James Newland

Mattia Laudi

June, 2020

Abstract

For patients infected by Covid-19, underlying health conditions are often cited as a source of increased vulnerability, of which exposure to high levels of air pollution has proven to be an exacerbating cause. We investigate the effect of long-term pollution exposure on Covid-19 mortality, admissions to hospitals and admissions to intensive care units in France. Using cross-sectional count data at the local level, we fit mixed effect negative binomial models with the three Covid-19 measures as dependent variables and atmospheric $PM_{2.5}$ concentration ($\mu g/m^3$) as an explanatory variable, while adjusting for a large set of potential confounders. We find that a one-unit increase in $PM_{2.5}$ concentration raised on average the mortality rate by 22%, the admission to ICU rate by 11% and the admission to hospital rate by 14% (rates with respect to population). These results are robust to a large set of sensitivity analyses. As a novel contribution, we estimate tangible marginal costs of pollution, and suggest that an marginal increase in pollution resulted on average in 61 deaths and created a 1 million euro surcharge in intensive care treatments over the investigated period (March 19^{th} - May 25^{th}).

[1]All authors are affiliated to the Barcelona Graduate School of Economics. Contact details: hubert.massoni@barcelonagse.eu.

Authors thank Larbi Alaoui, Gianmarco León and Sergio Pirla Lopez for their insightful guidance and comments throughout the preparation of this study.

Contents

I. Introduction

The Covid-19 pandemic, to date having taken the lives of over 376,000 people around the world, is the most severe public health crisis since the Spanish flu of 1918. Controlling the transmission of the deadly virus has required an all-out response from policymakers, resulting in nationwide 'lockdowns' and the closure of all non-essential sectors of economic activity. France has been heavily affected, with its regions suffering to varying degrees of affliction. *Départements*, sub-regional administrative areas, have faced drastic differences in human and healthcare costs as a result of the pandemic, to the point where some local healthcare systems have threatened to break down. Therefore, understanding the drivers of more acute Covid-19 prognoses across affected areas is crucial in assessing and controlling the epidemiological crisis.

Severe cases of Covid-19 are characterised by symptoms such as respiratory distress or other vital organ failure, which require the use of a ventilator or the admission to intensive care units (ICU) (Zu et al. 2020; CDC 2020). So far, noted sources of individual vulnerability include age ($\geq$60), sex (male), ethnicity (Black, Asian), obesity, diabetes, immunocompromise, smoking, severe heart conditions, chronic lung disease and asthma (BMJ 2020; WHO 2020; Brandt, Beck, and Mersha 2020). Within the sources of increased vulnerability, an underlying factor of many is prolonged exposure to fine particulate matter, or $PM_{2.5}$, defined as atmospheric pollution containing minute particles smaller than 2.5 micrometres in diameter. $PM_{2.5}$ pollution can contain harmful solids which lead to serious health conditions, such as respiratory and cardiovascular illnesses, which increase vulnerability to infectious diseases and elevates mortality risk. In this paper, we provide evidence that long-term exposure to $PM_{2.5}$ worsens the symptoms and prognosis of Covid-19 patients.

Our central hypothesis of adverse health effects of long-term exposure to air pollution, is grounded in a plethora of existing biomedical and epidemiological literature.

In a review of European epidemiological studies, Pelucchi et al. (2009) highlight that overall mortality is exacerbated by long-term exposure to $PM_{2.5}$ and other PM measures. Excesses in pollution-associ,ated mortality are thought to be driven by increased risk for an array of diseases, such as cardiovascular, ischemic heart, and respiratory diseases (Dockery et al. 1993; Pope III et al. 2002; Henschel and Chan 2013; Beelen, Raaschou-Nielsen, et al. 2014). Perhaps most compellingly, Cui et al. (2003) found that rates of mortality during the SARS virus outbreak - another closely related coronavirus - of 2002 were positively associated with exposure to air pollution in China. There seems to be a wide consensus on the adverse effects of air pollution on human health, and vulnerability to chronic and infectious diseases.

As the pandemic has unfolded, many researchers have observed a *correlation* between air pollution and local 'hotspots' of Covid-19 deaths. Ogen (2020), finds significant correlations between NO2 concentrations and abnormal spikes in death counts. Setti et al. (2020) provide evidence that areas of increased PM_{10} particulate matter saw higher rates of infections.[2] Conticini, Frediani, and Caro (2020) conclude that higher levels of $PM_{2.5}$ and PM10 correlate strongly with higher mortality rates. All three studies examine the case of Italy, specifically Lombardy, which was the first European region to experience the outbreak of the virus.[3]

In a recent epidemiological study, X. Wu et al. (2020) quantify the effect of exposure to air pollution on the number of Covid-19 deaths in US counties, and establish air pollution as a key environmental risk factor. They find that a one-unit increase in $PM_{2.5}$ ($\mu g/m^3$) is associated with an 8% increase in the Covid-19 death rate, controlling for a wide range of socioeconomic and behavioural factors.

[2] Where PM_{10} is particulate matter smaller than 10 micrometers.

[3] Although running parallel to our discussion, the present study does not address the phenomenon of reductions in global air pollution following enforced lockdowns and diminishing economic activity, as in Dutheil, Baker, and Navel (2020).

We perform a similar analysis, but in a setting which favours the empirical strategy, hopefully allowing for cleaner causal inference. We present our results in the language and context of health and environmental economics. As a novel contribution, we extend the analysis to other measures of severity of Covid-19 symptoms, namely admissions to intensive care units and admissions to hospitals. We use the latter two results to provide tangible estimates of the direct health and economic costs of the pandemic associated with pollution. This paper is - to our knowledge - the only empirical investigation into the *causal* effect of long-term exposure to $PM_{2.5}$ air pollution on Covid-19 deaths and hospitalisations in Europe.

Ultimately, we find that a one-unit increase in long-term $PM_{2.5}$ concentrations is associated with a 22% increase in the Covid-19 mortality rate (deaths over population), an 11% increase in admissions to ICU and a 14% increase in admissions to hospitals, across a sample of 96 départements in France. In real terms, these 'marginal costs of pollution' represented on average 61 excess deaths and roughly 1 million euros in excess intensive care costs since March 19^{th}. We adjust estimates for the marginal mortality effect by accounting for under-reporting in Covid-19 death counts.

The following section introduces the case of France in the Covid-19 crisis. Section 3 presents the data collection and the rationale behind the inclusion of confounding factors. Section 4 describes our main empirical strategy and robustness checks. We present our results in section 5. Section 6 comprises a discussion of the results and their potential limitations, with an extended discussion on under-reporting of deaths, and finally an estimate of the healthcare costs of pollution. Section 7 concludes. All tables and figures are included in the appendix.

II. Background: France facing the pandemic

We choose to work within the borders of metropolitan France to avoid differences in reporting methods and other qualitative inconsistencies which could not reasonably be controlled for. Conticini, Frediani, and Caro (2020) find that differences in reporting and count methods account for large differences in mortality rates, so we hope to minimise these. As we write the present work, data collection is imperfect and the degree of accuracy varies greatly between countries.

As of the 25^{th} May 2020, 182,942 Covid-19 cases were confirmed, although inconsistencies in the testing and reporting methods shed doubts on the accuracy of these numbers, which might be largely under-estimated. France reported 28,432 Covid-19-related deaths, from which 18,405 were reported in hospitals. North-eastern France and the greater Paris area were the first areas affected by the virus, and are consequentially the worst affected (Figure 1). Hospitals in these broad regions have constantly reported financial and material shortages.

As of March 17^{th}, 2020, a nation-wide 'lockdown' has been implemented across all départements to counter the surge in Covid-19 cases. In contrast to the varied state-level approaches taken in the United States, the lockdown was uniform across the country, where stay-at-home orders and restriction of movement were strictly enforced. Only key sectors remained active, such as agriculture, construction and healthcare.

France, moreover, displays high spatial variability in levels of air pollution. The industrial north and north-eastern regions, as well as the greater Paris area, display the highest concentrations of atmospheric $PM_{2.5}$, as illustrated by Figure 1. Visual inspection already indicates that regions with higher pollution exposure have been relatively more affected by the pandemic.

III. Data

Our study investigates the relationship between long-term $PM_{2.5}$ concentrations and Covid-19 related deaths, hospitalisations and admissions to ICU. These three indicators inform on the impact of pollution on different degrees of severity of Covid-19 symptoms. Data is gathered at the départemental level – an administrative sub-regional division of metropolitan France. This creates 96 individual observations in our sample, representing the entire population of metropolitan France (including Corsica). Table 1 provides summary statistics on dependent and independent variables.

i. Covid-19 data

We obtained data on confirmed Covid-19 cases from Santé Publique France (National Agency of Public Health), provided on a daily basis since March 19^{th} 2020. Data is reported by hospitals and frequently updated for reporting errors. We collect the cumulative Covid-related number of deaths, admissions to intensive care units and admissions to hospitals up to and including the 25^{th} of May 2020 at the départemental level. It should be noted that the data does not account for the total number of Covid-related deaths in France, since Santé Publique does not report disaggregated data on deaths in nursing or private homes. We comment on this issue in the discussion section.

ii. Pollution data

We obtained temporally averaged annual $PM_{2.5}$ concentrations from the NASA Socioeconomic Data and Applications Center. This dataset combines aerosol optical depth (AOD) retrievals from multiple satellite instruments. The GEOS-Chem chemical transport model is used to account for weather factors, then data is gridded at the 0.01x0.01-degree resolution, implying one km^2 pixels (Donkelaar et al. 2019). For context, the median land area of a département is 5,965 km^2. These annual values are averaged across the period 2000-2016, then départemental averages are calculated by computing mean concentrations for all latitude/longitude combinations in a département. We define these averages as the long-term exposure to pollution of each département's inhabitants.

iii. Confounding factors

To better isolate the role of long-term pollution exposure in the prevalence of strong Covid-19 symptoms, we control for 18 confounding factors, distinguished in 6 classes (a detailed list of sources can be found in Table 7):

(i) Socio-economic characteristics: median income, poverty rate and share of disposable income from wealth. (ii) Demographic characteristics: the share of people above 65 years of age and the male share of the population. (iii) Healthcare capacity: the number of doctors per 100,000 inhabitants and the number of hospital beds used for Covid-19 patients.[4] (iv) Health and behavioural characteristics capturing existing health risks and sources of vulnerability: smoking, obesity and diabetes rates. (v) Weather factors affecting pollution levels and Covid-19 spread: average rain fall (mm/year) and average temperature (°C) over the 2016-2018 period.[5] (vi) Factors influencing the spread of the virus: Covid-19 infections are naturally the most important predictor of Covid-19 fatalities. Without reliable test data, we must rely on additional factors that accelerate or slow the propagation of the virus. Close contact is a precursor of its contagion so we include population density, proportion of overcrowded residential apartments and the share of the labour force who continue to work in non-confined sectors. We include inter-départemental migration around the implementation date of the lockdown (*circa* 17$^{\text{th}}$ of March), proxied by inter-département cell phone movements, as data shows an overall migration out of densely populated areas into less populated areas. If not taken into account, these migration movements may bias our estimates, given that the popu-

[4]We exclude beds used for cardiology and neuro-vascular care given that only reanimation and general medicine beds are used to treat Covid-19 cases in France.

[5]Weather factors are found to influence the transmission of Covid-19 (Ma et al. 2020).

lation exposed to long term pollution in cities would migrate and contract the virus in rural areas, which are arguably less polluted. For completeness, we include the number of positive tests in cities as a measure of the state of the epidemic in each département - although biased given different urbanisation characteristics. Finally, we compute the number of days before the 50th case admitted to hospital as an additional measure of the state of the epidemic in each départements. Available data does not allow us to compute the exact day of the 1st case, hence the use of this specification - which will be the basis for additional robustness checks.

IV. Empirical strategy

Our main goal is to estimate the impact of long-term exposure to pollution on the mortality rate, the admission to ICU rate and the admission to hospital rate. We defined long-term pollution exposure as the annual average of estimated concentrations of $PM_{2.5}$ over the 2000-2016 period. Mortality, ICU and hospital rates are the total count of each variable, divided by the département's population (per 100,000).

We fit a cross-sectional mixed effect negative binomial model (MENB) with random intercept using $PM_{2.5}$ as the explanatory variable and the Covid-related rates as the dependent variables. We control for 18 possible confounding factors, and run 24 secondary regressions and sensitivity analyses per dependent variable, to examine the robustness of our model to various specifications.

i. Model

The nature of the dependent variables, namely count data, is a call for caution in causal inference, as it usually follows a non-normal distribution. A preferred choice for explaining variability in counts is a distributional assumption belonging to the Poisson log-linear family (Booth et al. 2003). In addition, the Covid-19 data presents strong signs of over-dispersion, illustrated in Table 2. Whilst a Poisson distribution has its mean equal to its variance, the data displays variances greater than means. This feature of the data is found across départements with both low and high levels of pollution. We instead assume a negative binomial distribution, which is more appropriate to deal with overdispersed data (Booth et al. 2003).[6]

Another concern stemming from the data is the potential spatial correlation of errors. Hospitals and healthcare services in France are mostly regulated and administered at the regional level. The response to the pandemic and the institutional capacities to manage patients are likely to share similar characteristics across départements belonging to the same region. Furthermore, air pollution is extremely likely to show autocorrelation across neighbouring départements (Gautam, Teraiya, and Patra 2018). A typical solution would involve adding spatial lags, or allowing for spatial dependence in the residual term. However, the negative binomial model cannot have an endogenous spatial lag, as the distribution is not stable. Subsequently when modelling Poisson-distributed count data, different techniques are required (Mohebbi, Wolfe, and Jolley 2011). We therefore follow one of the recommended methods of Arora and Brown (1977) and include a random intercept at the regional level (13 groups) to capture regional heterogeneity in administrative response to the pandemic. This is also a reasonable approximation of spatial correlation in pollution and in the spread of Covid-19.

We, therefore, fit a mixed effect negative binomial model (MENB), defined as follows:

$$\log(y_i) = \mathbf{z}_i^T \mathbf{b}_i + \mathbf{x}_i^T \beta + \log(population_i) + r_i \tag{1}$$

where $y_i \sim NB\left(y_i | \mu_i, \theta\right)$ and i is the département index, $\mathbf{b}_i$ is the random intercept vector, where $\mathbf{b}_i \sim N_K(0, \Psi)$. $\mathbf{x}_i$ is a vector of explanatory variables and β is estimated by maximum likelihood. Region identifiers $\mathbf{z}_i$ introduce spatial dependence of counts

[6]The negative binomial distribution is defined as: $y_i \sim NB\left(y_i | \mu_i, \theta\right) = \frac{\Gamma(y_i+\theta)}{\Gamma(\theta) y_i!} \cdot \left(\frac{\theta}{\mu_i+\theta}\right)^{\theta} \cdot \left(\frac{\mu_i}{\mu_i+\theta}\right)^{y_i}$

and r_i is a Pearson residual. The model is adjusted for 18 covariates and includes a population size (per 100,000) exposure.

Under this specification, dependent variables are transformed to rates: dependent variable over population (per 100,000). Our empirical strategy hence informs about the mortality, admission to ICU and to hospital rates. Estimates from the MENB models - exponentiated MENB coefficients - are interpreted as Incidence Rate Ratios (IRR): a percentage increase in the dependent variable rate resulting from one-unit increase in an independent variable.

ii. Robustness checks

The risk of over-fitting, the use of count data and limited sample size raise concerns over the robustness of the findings to changes in specifications, modelling choices and assumptions (notably distributional). We first run 7 secondary analyses for each of our three dependent variables to confirm the magnitude and significance of the estimate of interest when removing potential confounders - so that our main results are not biased by over-controlling. These include the removal of different covariate groups: healthcare capacities, tests performed in cities, pre-existing health conditions, weather conditions and Covid-19 spread channels. A last secondary analysis removes the hot spots of the epidemic in France (the 5 highest death counts), to confirm that the results are not driven by outliers (mainly, the region Ile-de-France, including Paris) which suffer from a very high number of Covid-19 cases. This particular specification notably alleviates possible bias stemming from the way we defined the start of the epidemic. Given that we defined the variable as the day when a département reaches 50 cases admitted to hospital, several départements were attributed a similar start date - the very first day included in the data. Removing hot spots palliates possible bias embedded in our measure of the start of the epidemic by representing cross-départemental differences more accurately in the epidemic outbreak.

Second, we perform a set of sensitivity analyses with alternative specifications and assumptions. These include (i) MENB with clustered standard errors at the regional level, (ii) $PM_{2.5}$ estimation as quintiles, (iii) density of population as quintiles, (iv) more disaggregated and progressive age categories, (v) including the natural logarithm of population as a covariate. In addition, to cross-check the results of the MENB models under different distributional assumptions, we estimate the effect of the long-term pollution exposure on the natural logarithm of the Covid-19 rates by OLS. We use spatially correlated standard errors up to a 200km radius, and perform comparable secondary and sensitivity analyses. Although more standard in causal inference analyses of this kind in economics, OLS is not our favoured specification given the possible bias when dealing with count data (Cameron and Trivedi 2003). Our OLS model is defined as follows:

$$y_i = \beta_0 + \mathbf{x}_i^T \beta_1 + \epsilon_i \tag{2}$$

where y_i is our dependent variable and $\mathbf{x}_i$ is a vector of explanatory variables, as listed before. Also, $\epsilon_i \sim N(0, \sigma^2 + \hat{C}(k))$, $\hat{C}(k)$ estimates spatial covariances at distance k.

V. Results

Main analyses

Of the départements in our sample, all 96 had reported at least one death by the time of our analysis. We analyse cumulative death, admission to ICU and admission to hospital tolls up to May 25^{th}, 2020, expressed in respective rates per départemental population size (per 100,000).[7]

[7] It is to be noted that the likelihood ratio test of negative binomial vis-à-vis Poisson regression and the likelihood ratio test of negative binomial with random intercept vis-à-vis without random intercept are both significant. This indicates respectively over-dispersion and intra-regional correlation of errors and confirms the model specifications.

Table 3 reports the results of our main specification. Regarding the mortality rate, we find an estimated IRR of 1.22 on $PM_{2.5}$ (95% CI: [1.10; 1.36]), statistically highly significant at the 0.1% level. Regarding the admission to ICU rate, we find an estimated IRR of 1.11 on $PM_{2.5}$ (95% CI: [1.01; 1.22]), statistically significant at the 5% level. Regarding the hospitalisation rate, we find an estimated IRR of 1.14 (95% CI: [1.05; 1.25]), statistically significant at the 1% level. That is, our results suggest that a one-unit increase in long-term average $PM_{2.5}$ (μg/m^3) is associated with an average increase of 22% in the Covid-19 death rate, 11% in the admission to ICU rate and 14% in the admission to hospital rate in France.

Meanwhile, we find that the share of the population aged above 60 years, the male share of the population and rates of diabetes are positive determinants of all three dependent variables (significant at the 5% level or less). These results are consistent with previous research into risk factors for Covid-19 patients, particularly with findings regarding age and sex differences in mortality rates (BMJ 2020; WHO 2020). We also find significant positive coefficients for median income, but sensitivity analyses cast doubts on causality (see discussion section). Finally, we find that a strong predictor of all three measures is the start date of the epidemic, which highlights that areas which were first hit by the pandemic are worse affected. This is intuitive, the longer the virus is allowed to spread, the greater the number of infections and fatalities in a population.

Secondary and sensitivity analyses

Importantly, these findings are consistent with results from 24 secondary and sensitivity analyses (per dependent variable, all reported in Table 4 and 5). Figure 2 displays the robustness of our coefficient of interest and respective confidence intervals to secondary analyses, in which we separately omitted groups of covariates or observations with higher mortality rates. Secondary analyses confirm the magnitude of the effect of long-term pollution on mortality, admission to ICU and admission to hospital rate, with an IRR ranging between 1.17-1.31 for the mortality rate, 1.04-1.16 for the admission to

ICU rate and 1.09-1.21 for the admission to hospital rate across 8 specifications.

We report in Table 4 and 5 the IRR and confidence intervals obtained from the sensitivity analyses, including a set of alternative OLS specifications with spatially correlated standard errors, up to a radius of 200km. Sensitivity analyses are consistent with our main findings, and OLS estimates are notably extremely similar to MENB estimates, with the primary specification obtaining an e^{β} of 1.22.

VI. Discussion

i. Determining causality

Our results and their robustness to different specifications show that there is a clear association between $PM_{2.5}$ concentrations and Covid-19 outcomes. We hypothesize, based on scientific empirical work such as Pelucchi et al. (2009) and Lim et al. (2012), that long-term exposure to air pollution damages respiratory health, making a patient more vulnerable to Covid-19 and therefore more likely to exhibit more acute symptoms of the disease. This channel of influence is supported by a wide scientific consensus, which, combined with the robustness of our results, allows us to infer a causal relationship between long-term exposure to air pollution and Covid-19 outcomes in France, and likely elsewhere.

Furthermore, our findings are consistent in magnitude with Wu et. al. (2020) and Cui et al. (2003) regarding respectively Covid-19 patients in the US, and SARS patients in China, although we find a larger mortality IRR of 1.22 than findings of 1.08 by Wu et. al., 2020. We believe France offers a cleaner setting to the study, due to its uniform lockdown and broad exposure to the virus. Our model also echoes recent results from the BMJ (2020) and WHO (2020), finding statistically significant effects of demographic factors such as age and sex, and of underlying health conditions (diabetes). We also note that areas which recorded 50 cases earlier in the year see higher mortality rates and more severe symptoms.

However, we find that log median household income is extremely positively associated with mortality, ICU admission and hospital admission rates (IRR of 2.91, 15.69, 17.18 respectively) and the latter two coefficients are statistically significant at the 5% level. Whilst this appears troubling, this does not necessarily imply that wealthier individuals are likely to suffer more greatly. Instead, wealthier départements are centres of economic activity, host many more visitors and returning citizens from abroad, and are densely populated. These dynamics accelerate the spread of the virus and therefore its cumulative impact. In fact, when Paris - the economic center and primary recipient of foreign visitors in the country - is excluded from the sample, we observe a negative coefficient on log income. The result is not robust to any of our sensitivity analyses, either.

There may exist alternative channels of influence which can explain our result. One such theory is that air pollution increases airborne transmission of the virus, whereby atmospheric particulate matter is able to 'carry' the virus and aid its propagation. In their own evaluation of the negative impacts of air pollution on the Covid-19 pandemic, Setti et al. (2020) remark that rates of infection were also far higher in areas of higher pollution. We support the view that present-day air pollution may aid the transmission of the virus. However, we stress that our study is unrelated to their hypothesis, given that we use historic pollution data which precedes the pandemic.

We must also consider the possibility of external factors, which may create a 'spurious regression'. In controlling for 18 potential confounders, we hope to exhaust the list of possible influences. Our estimates are broadly in line with findings in the USA and from previous pandemics, so we hope that we have succeeded in integrating all potential confounding factors.

Given the robustness of our results, their concordance with other existing empirical findings, and the strength of our hypothesised channel of influence, we believe there is strong reason to infer a causal relationship. Our study is not without limitations, however, which future research should hope to evade.

ii. Limitations

Observational issues

Whilst we maintain that limiting the scope of our study to continental France improves the precision of our results, as doing so controls for potential inconsistencies which may occur in a cross-border study, it comes at the cost of a small sample size (maximum of 96 possible observations). Estimates from the MENB are asymptotically unbiased, so greater sample size will improve the accuracy of results (Davis and R. Wu 2009). Secondly, when using spatial data such as air pollution, results can vary according to the scale of aggregation (Crouse et al. 2019). A quick robustness check is to change the level of aggregation of data and re-estimate the model. However, in our case, we can only increase the level of aggregation to the regional level, which permits 12 observations, unlikely to have the statistical power to validate our results. Low observation count also prevents us from performing quintile analysis of our variable of interest to investigate the hypothesis of linear relationship between pollution and Covid-19 outcomes, as each quintile of our sample contains only about 20 observations.

Exposure Model

Using satellite-measured $PM_{2.5}$ data is an imprecise estimate for actual exposure to ambient air $PM_{2.5}$. One method is the use of exposure models, which are unavailable for $PM_{2.5}$ in Europe and tend to focus on NO_2 instead (Beelen, Hoek, et al. 2013; Vienneau, Hoogh, and Briggs 2009). Ideally, we would use ground-level monitoring stations as our data source, but unfortunately, such data does not properly cover the territory.

Data issues

The reliability of the hospital and infection data is a possible cause for concern. Infection data only represents tests performed in cities, while we will show the official death count is likely to vastly underestimate the true figures (see extended discussion). Secondly, our model does not include ethnicity data which is proven to greatly affect mortality rates (Brandt, Beck, and Mersha 2020).[8] However, the authors argue the 'ethnicity' effect is captured in large part by socioeconomic factors such as income, lower access to healthcare capacities and employment in non-confined economic sectors, which we are able to control for.

Hence, we acknowledge that our results could be more informative at a larger scale, and that a low number of observations may be a limiting factor to the success of our investigation. A larger scale study could improve the reliability and external validity of our estimates. Unfortunately, this is not possible until data regarding Covid-19 deaths and infections are more reliable and consistent across different countries. The weakness of Covid-19-related data is a consequence of its novelty. Once such data is consistent and accurate, future work should replicate our methods in neighbouring European countries, and further afield, in order to develop a comprehensive understanding for the area.

iii. Extended discussion on estimating pollution impacts

In this section we attempt to provide estimations of the Covid-19-related human and financial costs associated with long-term pollution exposure. Our findings so far inform on the marginal Covid-19 effects of pollution, but we seek to understand how these translate to tangible costs over the investigated period.

Death reporting

The rapid spread of the pandemic has revealed national statistic agencies to be ill-

[8] France does not allow collection of data based on race or ethnicity, even at the aggregated level.

equipped to deal with accurately tracking case and death counts. Before April 4th, Santé Publique France had been reporting solely deaths occurring in hospitals, without taking into account deaths in pension and private homes. Therefore, any attempt to quantify the effect of pollution on Covid-19 deaths, beyond marginal effects, is likely to be severely biased downward.

We follow Chopin (2020) and propose a simple procedure to calculate the true number of Covid-19-related deaths. We take advantage of data provided by INSEE during the pandemic, which reports the départemental total death counts from *all causes* from early 2018 to today (Figure 3). We calculate excess deaths by subtracting the average 2018-2019 weekly death counts from the 2020 weekly death count in each French département. We regress excess deaths reported from the 19th of March to the 11th of May, 2020 on Covid-19 deaths reported in hospitals. Weekly data are more stable than daily data and less affected by delays in the statistical reporting procedures.[9] We use panel data and estimate the following model:

$$ex\_d_{it} = \beta_0 + \beta_1 hosp\_d_{it} + \varepsilon_{it}$$

Where ex_d_{it} represents excess deaths in département i at time t, $hosp\_d_{it}$ is number of deaths in hospitals in département i at time t and ε_{it} is the error term. We use clustered standard errors at the regional level to account for differences in reporting quality (Table 6).

We find a positive relationship between the number of hospital deaths and excess deaths (Figure 4). The model yields a coefficient of 1.69 (95% CI: [1.62, 1.75]), statistically highly significant at 0.1% level, implying that any death count reported by hospitals should be augmented by 62-75% to take into account under-reporting of Covid-19 deaths. Although overly simple, we hope this method informs on the magnitude of

[9]INSEE reports weekend deaths on the following Monday.

the discrepancy between deaths reported by hospitals and the estimated total Covid-19 deaths. An ideal study of long-term exposure to pollution would be based on more accurate mortality data.

Nevertheless, using this estimate to adjust for under-reporting, we attempt to assess how many additional Covid-19 deaths can be attributed to higher air pollution since the outbreak of the pandemic. We do so by multiplying the average Covid-19-related mortality rate in hospitals across France by our estimated marginal increase in the death rate (22%), multiplying by the average département population (per 100,000) and then multiplying by the coefficient estimated in the above regression (1.69) - full calculations are in the Appendix.

We obtain that a one-unit increase in long-term $PM_{2.5}$ level ($\mu g/m^3$) resulted on average in 61 additional Covid-19 deaths per département since March 19^{th}.[10] Repeating this method for ICU and hospital admissions, we predict marginal increases of 15 and 131 respective admissions owing to this one-unit increase and over the investigated period.[11] For context, a standard deviation of pollution concentrations equates to a 2.4 $\mu g/m^3$ change in $PM_{2.5}$.

Cost estimation

In order to emphasize the economic costs of air pollution whilst Covid-19 is in the spotlight, we attempt to provide a conservative estimate of the cost of ICU responses – crucial defence mechanisms against pandemics of respiratory diseases. Assistance publique-Hôpitaux de Paris estimates the daily cost for ICU treatment to be €4,658 per day (AP-HP 2019). Furthermore, Grasselli et al. (2020) estimate the average stay of patients in ICU to be 14 days. We hence estimate the Covid-19-cost of air pollution

[10] With a lower bound and an upper bound of 27 and 103, respectively. For context, in-hospital Covid-19 deaths average 192 per département.

[11] With lower and upper bounds of (1.4, 31) for ICU and (47, 234) for hospital admissions.

by multiplying the daily cost by the average number of days spent in an ICU, and multiplying again by the estimate of the increased number of ICU admissions attributable to one additional unit of $PM_{2.5}$ - see calculations in Appendix. We obtain that the marginal cost of a one-unit increase in $PM_{2.5}$ has been €1,007,895, purely in terms of intensive care treatments and since March 19^{th}.[12]

This method is crude. However, the provided estimates are likely a lower bound to the true current cost of Covid-19 on the healthcare system. We had hoped to estimate the costs of treatment for less-severe hospitalisations, but providing a central estimate is made difficult by extreme variability in treatment costs, since duration of stay and severity of symptoms differ significantly from patient to patient.

That said, our base estimate should alert policymakers and highlight air pollution as a costly surcharge associated with the pandemic. Air pollution is a well-documented negative externality which here has direct social costs (both in human and financial terms). If overlooked as a risk factor, or under-estimated due to Covid-19 outcomes reporting errors, air pollution can lead to misallocation of resources within France's public health system and failures in the optimisation of the response to the pandemic. In turn, this may lead to a greater number of deaths and severe cases, especially when considering a gradual re-opening of the economy following the lockdown.

VII. Conclusion

The present study is a strong indication that air pollution is a crucial environmental factor in mortality risks and vulnerability to Covid-19. The health risks associated with air pollution are well documented, but with Covid-19 in the spotlight we hope to increase awareness of the threat caused by pollution, not only through direct increased

[12] With lower and upper bounds of €90,438 and €1,989,635, respectively. The lower bound is driven by a small lower bound of the IRR for ICU admissions.

health risks, but also through external factors, such as pandemics.

We show the aggravating effect of long-term pollution exposure to three levels of severity of Covid-19 symptoms in France: admission to hospitals for acute Covid-19 cases, admission to intensive care units for the most severe vital organ failures, and fatalities (all expressed per 100,000 inhabitants). Using cross-sectional data at the départemental level, we fit mixed effect negative binomial models with the three Covid-19 measures as dependent variables and the average level of atmospheric concentration of $PM_{2.5}$ ($\mu g/m^3$) as an explanatory variable. We adjust for a set of 18 potential confounders to isolate the role of pollution in the spread of the Covid-19 disease across départements. We find that a one-unit increase in average $PM_{2.5}$ levels increases on average the mortality rate by 22%, the admission to ICU rate by 11% and the admission to hospital rate by 14%. These results are robust to a set of 24 secondary and sensitivity analyses per dependent variable, confirming the consistency of the findings across a wide range of specifications.

We further provide numerical - and hence more tangible - estimates of the marginal costs of pollution since March 19^{th}. Adjusting for under-reporting of Covid-19 deaths, we estimate that long-term exposure to pollution marginally resulted in an average 61 deaths across French départements. Moreover, based on average daily costs of intensive care treatments, we estimate that pollution induced an average 1 million euros in costs borne by hospitals treating severe symptoms of Covid-19. These figures strongly suggest that areas with greater air pollution faced substantially higher casualties and costs in hospital services, and raise concerns about misallocation of resources to the healthcare system in more polluted areas.

Our paper provides precise estimates and a reproducible model for future work, but is limited by the novelty of the phenomenon at the center of the study. Our empirical investigation is restricted to the scope of France alone due to cross-border inconsistencies in Covid-19 data collection and reporting. Once Covid-19 data reporting is complete and consistent, we hope future studies will examine the effects of air pollution at a greater scale, or in greater detail. On the other hand, more disaggregated data - at the individual or hospital level - would allow more precise estimates and a better understanding of key factors of Covid-19 health risks and would also allow the use of surface-measured air pollution. Measured pollution data is available for France, but is inherently biased when aggregated at the départemental level, due to lack of territorial coverage.

If precise data tracking periodic Covid-19 deaths becomes available for a wider geographic region, we specifically recommend a MENB panel regression incorporating a PCFE for spatially correlated errors. This will produce the most accurate estimates.

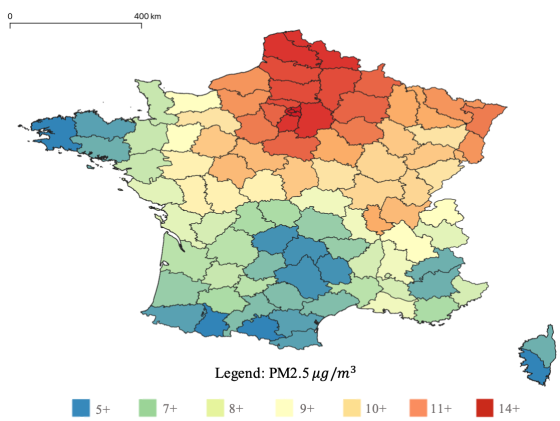
Going forward, more accurate and granular data should motivate future research to uncover the exact financial costs attributable to air pollution during the pandemic. Precise estimation of costs of Covid-19 treatments and equipment (e.g. basic protective equipment for personnel or resuscitation equipment), should feature in a more accurate cost analysis. Hospital responses should be thoroughly analysed to understand the true cost of treatments across all units.

It is crucial that the healthcare costs of pollution are globally recognised so that future policy decisions take them into account. Ultimately, this paper stresses that failure to manage and improve ambient air quality in the long run only magnifies future burdens on healthcare resources, and cause more damage to human life. During a global pandemic, the costs of permitting further air pollution appears ever more salient.

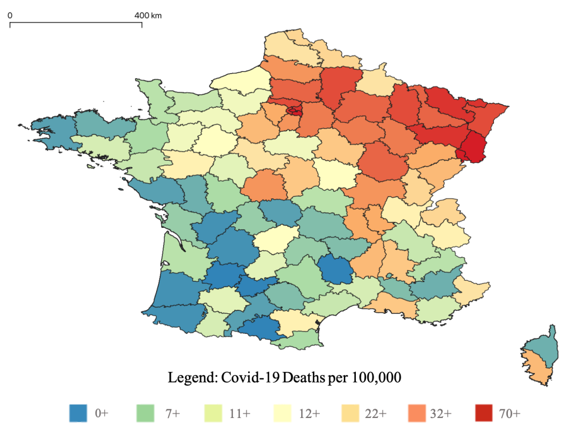
VIII. Appendix

Figure 1: Départemental-level Covid-19 rates and pollution concentration

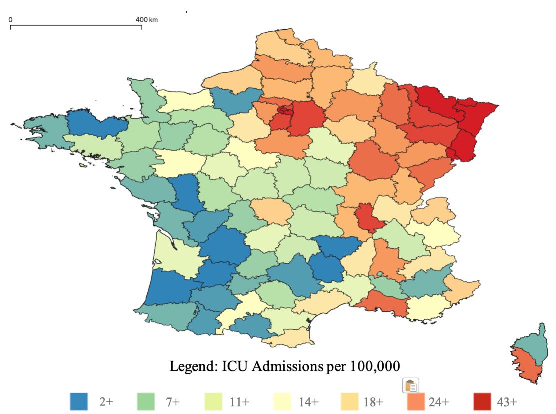
(a) Atmospheric $PM_{2.5}$ Concentrations



(b) Covid-19 Deaths



(c) ICU Admissions



(d) Hospital Admissions

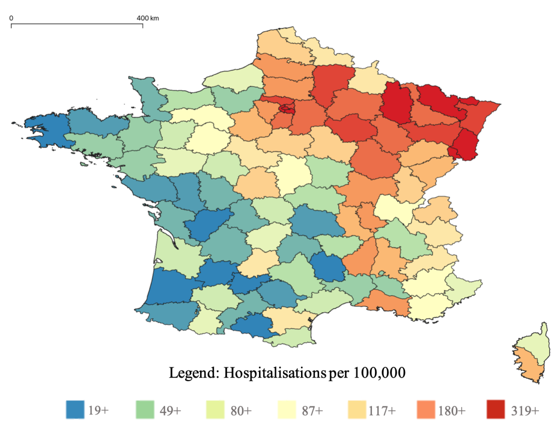


Table 1: Summary Statistics

	Mean	Std. Dev.	Min.	Max.
Covid-19 Counts				
Deaths	191.7	286.2	1	1691
Admissions to intensive care unit	179.6	291.0	2	2056
Admissions to hospital	1105.3	1575.2	24	8871
Covid-19 Rates				
Mortality rate (per 100,000)	24.2	23.3	1.31	124.2
Admission to intensive care unit rate (per 100,000)	20.5	16.7	2.62	95.7
Admission to hospital rate (per 100,000)	138.5	111.1	20.6	515.3
Pollution				
Average $PM_{2.5}$ ($\mu g/m^3$)	9.91	2.40	5.95	16.1
Demographic Characteristics				
Population (100,000)	6.76	5.20	0.76	25.9
% $\geq$ 60 years of age	29.6	4.83	16.7	39.3
% Men	48.5	0.49	47.0	49.6
Socio-economic Characteristics				
Median household income (€)	20780.4	1616.7	17310	27400
% In poverty	8.13	1.97	5	17.5
% Wealth in disposable income	9.75	1.93	5.90	19.6
Healthcare Capacity				
Doctors (per 100,000)	305.0	89.3	167	858
Hospital beds (per 100,000)	17.9	7.75	4.33	38.0
Covid-19 Spread Channels				
Tests (in cities, per 100,000)	520.2	453.3	0	2623.7
Population density (person/m^2)	565.8	2425.1	14.8	20459.7
% Inter-dep. migration	0.21	6.15	-38	7
% Over-crowed flat	5.61	4.92	1.80	30.9
% Non-confined sectors	10.3	2.91	1.50	19.3
Days before 50 cases in hospitals	8.15	7.05	1	40
Weather Factors				
Average rain fall (mm/year)	1167.4	214.3	795.5	1772.5
Average temperature (C)	11.9	1.63	7.25	16.1
Health Factors				
% Smoking	27.7	2.94	21.3	32.2
% Diabetic	4.39	0.67	2.84	6.94
% Obese	5.63	4.53	0.50	18.5
Observations	96			

Table 2: Overdispersion in Covid-19 counts

	Mortality		ICU		Hospital	
	Mean	Variance	Mean	Variance	Mean	Variance
Average $PM_{2.5}$ ($\mu g/m^3$)						
1st half	63.2	7153.8	76.0	11391.0	76.0	11391.0
2nd half	320.2	124734.6	283.2	137800.2	283.2	137800.2
Total	191.7	81933.4	179.6	84659.3	179.6	84659.3
Observations	96		96		96	

Note: Death, admission to ICU and admission to hospital counts, reported by hospitals (19/03 to 25/05). Data is presented by 1s & 2nd halfs of average $PM_{2.5}$ levels (threshold at 9.91 $\mu g/m$).

Table 3: Main Analyses

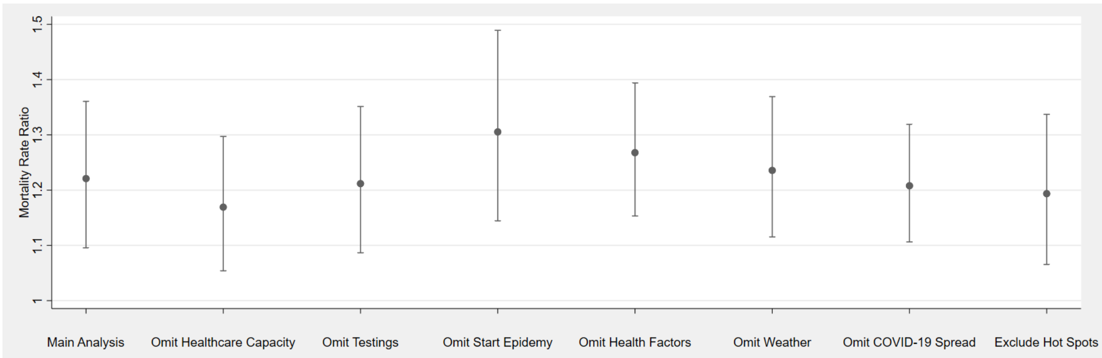
Variables	*Mixed Effect Negative Binomial Regressions* Mortality		ICU		Hospital	
Average $PM_{2.5}$ (μg/m^3)	1.22***	(0.07)	1.11*	(0.05)	1.14**	(0.05)
% $\geq$ 60 years of age	1.12***	(0.02)	1.05*	(0.02)	1.09***	(0.02)
% Men	1.87***	(0.27)	1.45**	(0.18)	1.85***	(0.22)
Doctors (per 100,000)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Hospital beds (per 100,000)	1.01	(0.01)	1.03***	(0.01)	1.02*	(0.01)
Tests (in cities, per 100,000)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Population density (person/km^2)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
% Inter-dep. migration	1.01	(0.01)	1.00	(0.01)	1.00	(0.01)
% Over-crowed flat	1.03	(0.02)	1.00	(0.02)	1.01	(0.02)
% Non-confined sectors	0.99	(0.03)	1.03	(0.03)	1.03	(0.03)
Days before 50 cases in hospitals	0.94***	(0.01)	0.97***	(0.01)	0.95***	(0.01)
log(Median household income)	2.91	(4.83)	15.69*	(21.80)	17.18*	(23.32)
% In poverty	0.96	(0.06)	1.06	(0.06)	1.03	(0.06)
% Wealth in disposable income	1.03	(0.03)	1.01	(0.02)	1.01	(0.02)
Average rain fall (mm/year)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Average temperature (C)	0.95	(0.05)	1.02	(0.04)	0.99	(0.04)
% Smoking	0.99	(0.03)	1.00	(0.03)	1.01	(0.03)
% Diabetic	1.39*	(0.22)	1.36*	(0.19)	1.42**	(0.18)
% Obese	1.02	(0.02)	1.02	(0.02)	1.03	(0.02)
log(α)	0.11***	(0.02)	0.08***	(0.02)	0.08***	(0.01)
Random intercept (region)	1.03	(0.02)	1.03	(0.02)	1.02	(0.02)
Exposure	Population		Population		Population	
Observations	92		92		92	

Exponentiated coefficients (IRR); Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

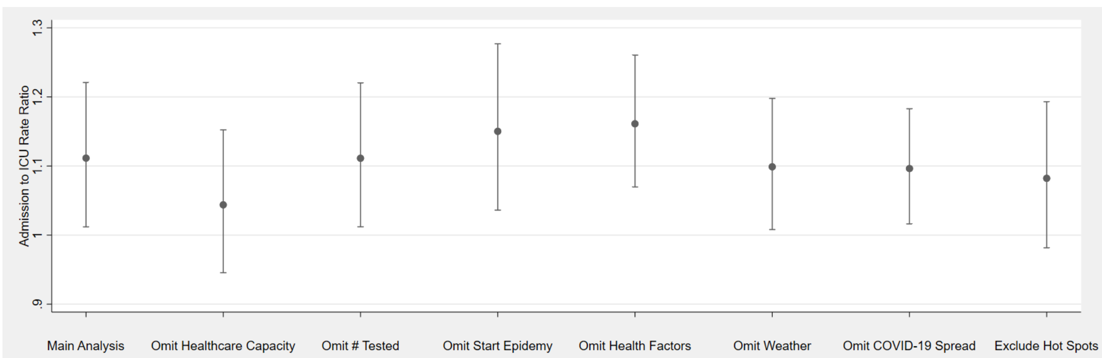
Note: Incidence Rate Ratios (e^{β}) and standard errors for all variables on the mortality rate, the admission to ICU rate and the admission to hospital rate. The IRR can be interpreted as a percent increase in the dependent rates resulting from a one-unit increase in the explanatory variable. Negative binomial models include a random intercept at the regional level (12 groups). Log(α) is a coefficient of dispersion of the data with respect to a Poisson distribution.

Figure 2: MENB Secondary analyses, $PM_{2.5}$ estimates and 95% CI

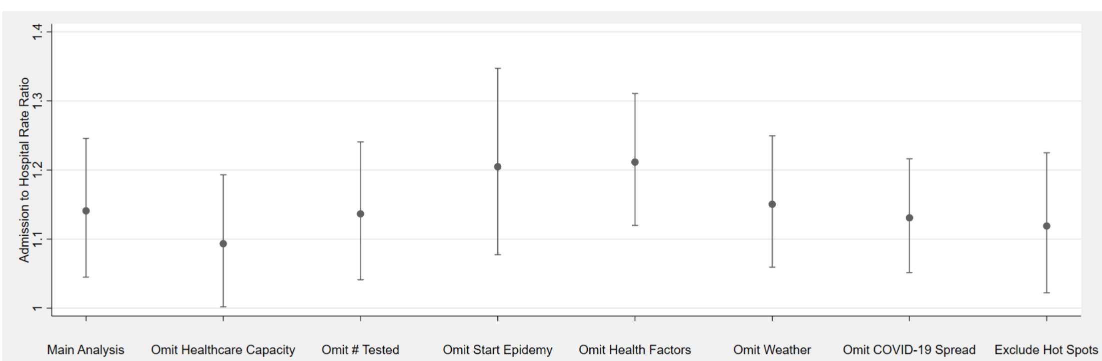
(a) Mortality



(b) Admission to ICU



(c) Admission to hospitals



Note: Incidence Rate Ratios and 95% confidence intervals for Mixed Effect Negative Binomial models (main analyses). The panels display the exponentiated estimates of average $PM_{2.5}$ levels on (a) mortality rate, (b) admission to ICU rate, (c) admission to hospital rate. The IRR can be interpreted as a percent increase in the rates (a), (b) or (c) resulting from a one-unit increase in $\mu g/m^3$ of $PM_{2.5}$.

Table 4: Sensitivity analyses, $PM_{2.5}$ estimates

Dependent variable: **Mortality rate**

	MENB		OLS	
Specifications	e^{β}	95% CI	e^{β}	95% CI
Main analysis	1.22^{***}	[1.10 1.36]	1.22^{**}	[1.07 1.39]
Omit healthcare capacity	1.17^{**}	[1.05 1.30]	1.17^{*}	[1.04 1.31]
Omit # tested	1.21^{***}	[1.09 1.35]	1.21^{**}	[1.07 1.37]
Omit start epidemy	1.31^{***}	[1.14 1.49]	1.39^{***}	[1.19 1.62]
Omit health factors	1.27^{***}	[1.15 1.39]	1.29^{***}	[1.19 1.40]
Omit weather	1.24^{***}	[1.12 1.37]	1.25^{***}	[1.15 1.36]
Omit Covid-19 spread	1.21^{***}	[1.11 1.32]	1.21^{***}	[1.08 1.34]
Exclude hot spots	1.20^{**}	[1.07 1.34]	1.20^{*}	[1.04 1.39]
$PM_{2.5}$ as quintile	1.22^{*}	[1.05 1.42]	1.25^{*}	[1.01 1.54]
Density as quintile	1.20^{***}	[1.09 1.31]	1.19^{**}	[1.06 1.34]
Disaggregated age	1.22^{***}	[1.09 1.36]	1.22^{**}	[1.05 1.41]
Log(population) as covariate	1.18^{***}	[1.07 1.30]	-	-
Clustered S.E. (region)	1.21^{*}	[1.04 1.41]	1.22^{*}	[1.01 1.47]

Dependent variable: **Admission to ICU rate**

	MENB		OLS	
Specifications	e^{β}	95% CI	e^{β}	95% CI
Main analysis	1.11^{*}	[1.01 1.22]	1.11^{*}	[1.01 1.23]
Omit healthcare capacity	1.04	[0.95 1.15]	1.05	[0.95 1.16]
Omit # tested	1.11^{*}	[1.01 1.22]	1.11^{*}	[1.01 1.23]
Omit start epidemy	1.15^{**}	[1.04 1.28]	1.19^{**}	[1.06 1.33]
Omit health factors	1.16^{***}	[1.07 1.26]	1.18^{***}	[1.10 1.26]
Omit weather	1.09^{*}	[1.01 1.20]	1.11^{**}	[1.04 1.19]
Omit Covid-19 spread	1.07^{*}	[1.02 1.18]	1.09^{*}	[1.01 1.18]
Exclude hot spots	1.08	[0.98 1.19]	1.09	[0.98 1.21]
$PM_{2.5}$ as quintile	1.09	[0.96 1.24]	1.13	[0.96 1.33]
Density as quintile	1.11^{**}	[1.03 1.20]	1.11^{**}	[1.03 1.19]
Disaggregated age	1.12^{*}	[1.02 1.23]	1.12^{*}	[1.01 1.24]
Log(population) as covariate	1.09^{*}	[1.00 1.21]	-	-
Clustered S.E. (region)	1.09^{*}	[1.00 1.20]	1.11^{*}	[1.00 1.24]

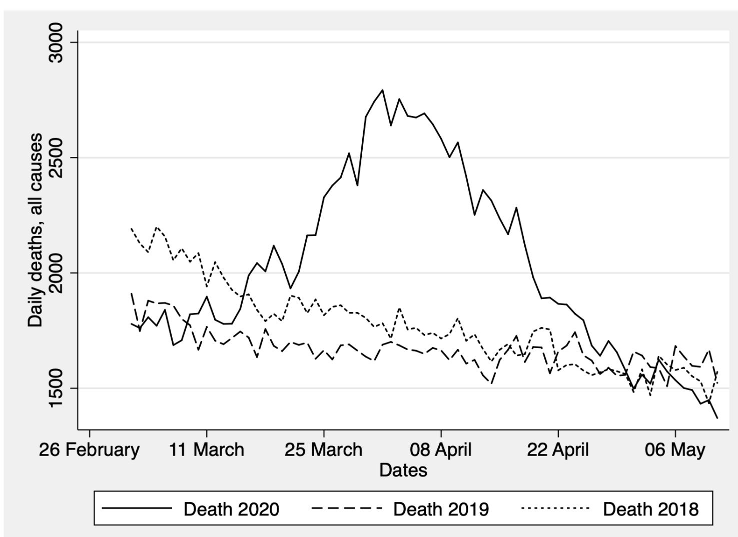
Table 5: Sensitivity analyses, $PM_{2.5}$ estimates (continued)

Dependent variable: **Admission to hospital rate**				
	MENB		OLS	
Specifications	e^{β}	95% CI	e^{β}	95% CI
Main analysis	1.14**	[1.05 1.25]	1.14*	[1.03 1.26]
Omit healthcare capacity	1.09*	[1.00 1.19]	1.09	[0.99 1.20]
Omit # tested	1.14**	[1.04 1.24]	1.13*	[1.02 1.25]
Omit start epidemy	1.21**	[1.08 1.35]	1.25***	[1.11 1.41]
Omit health factors	1.21***	[1.12 1.31]	1.23***	[1.16 1.31]
Omit weather	1.15***	[1.06 1.25]	1.16***	[1.11 1.22]
Omit Covid-19 spread	1.13***	[1.05 1.22]	1.13**	[1.03 1.23]
Exclude hot spots	1.12*	[1.02 1.23]	1.12*	[1.00 1.24]
$PM_{2.5}$ as quintile	1.18**	[1.04 1.33]	1.19*	[1.02 1.39]
Density as quintile	1.13***	[1.05 1.22]	1.13**	[1.03 1.23]
Disaggregated age	1.13**	[1.03 1.23]	1.13*	[1.01 1.26]
Log(population) as covariate	1.12**	[1.03 1.21]	-	-
Clustered S.E. (region)	1.14*	[1.02 1.27]	1.14	[0.99 1.31]

Exponentiated coefficients; 95% confidence intervals in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

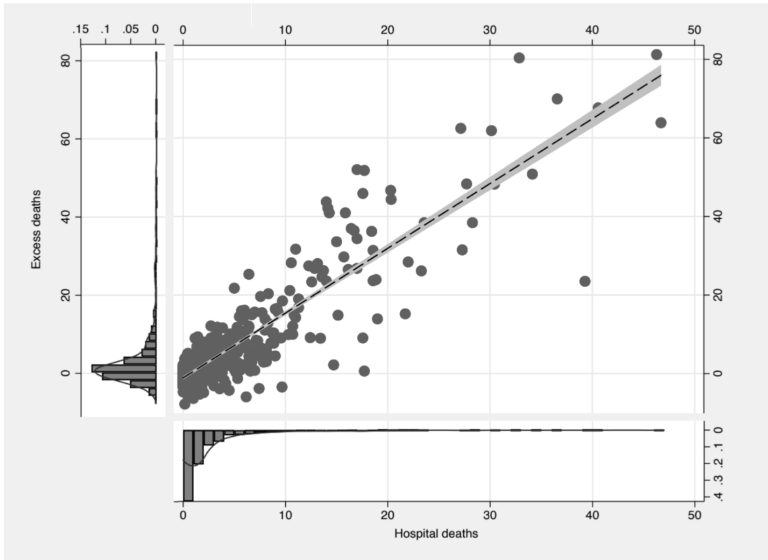
Note: Exponentiated estimates of average $PM_{2.5}$ ($\mu g/m^3$) levels (to be interpreted as Indicence Rate Ratios for MENB). Negative binomial regressions (MENB) include random intercept at the regional level, OLS regressions use Conley S.E. with radius 200km (except when "clustered" is specified).

Figure 3: Daily deaths trends



Note: Data obtained from INSEE (from March 1st to May 11th), which has been reporting the number of deaths per day per département on a weekly basis. In addition, for comparison, they have also been reporting the same data for years 2018 and 2019.

Figure 4: Excess deaths and hospital deaths



Note: This figure shows pairs of data points for *excess deaths* and *hospital deaths* with respective densities, from March 19th to May 11th. The upward slope indicates correlation between hospital deaths and non-reported Covid-19 deaths. The data has been obtained from INSEE, as in Figure 3.

Table 6: Mortality Under-reporting

	Excess deaths		
	Coeff.	Std. Err.	95% CI
Hospital Deaths	1.69***	0.036	[1.62 1.76]
Constant	-1.22***	0.17	[-1.55 -0.90]
Observations	96		
Periods	8		

Clustered standard errors reported (13 regions), 95% confidence intervals in brackets
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Estimates computed using pooled OLS with département level data for the cross-sectional dimension and weekly data for the time series dimension. We cluster at regional level to account for intra-regional correlation within broad administrative boundaries.

Calculations of marginal increases for deaths, ICU and hospital admissions.

$$y_c = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{\text{counts}_i}{\text{population}_i} \cdot 100,000 \right) \cdot (IRR_c - 1) \cdot \left(\frac{1}{N} \sum_{i=1}^{N} \text{population}_i \right) \cdot 100,000^{-1} \cdot \gamma_c$$

With c = counts = $\{1, 2, 3\}$ and 1 = deaths, 2 = ICU admissions, 3 = hospital admissions. Furthermore, counts_i represents the number of units by département, β is the IRR coefficient from our MENB regressions and $\gamma = 1.69$ if $c = 1$ and $\gamma = 1$ if $c = 2, 3$ (γ is the under-reporting ratio calculated from our regression).

Table 7: Data and sources

Category	Variables	Source
Socio-economic characteristics	Median income (€, 2017), poverty rate (2017), share of disposable income from wealth (2016).	INSEE
Covid-19 outcomes	(i) Covid-19-related deaths, admissions to ICU and admissions to hospital counts (March 19th to May 25th, 2020); (ii) Total deaths, all causes (2018-2020, March 1st to May 11th)	(i) Santé Publique; (ii) INSEE
$PM_{2.5}$ (μg/m^3)	$0.01° \times 0.01°$ grid resolution $PM_{2.5}$ prediction (averaged across grid cells in each département, 2000–2016).	SEDAC
Demographics	Population (2020), age (2020) and male share of the population (2020).	INSEE
Healthcare Quality	Number of doctors, number of hospital beds (per 100,000 inhabitants, 2018).	DREES
Health Factors	(i) Diabete (2012) and smoking (2017) prevalence; (ii) obesity rate (regional level, 2017).	(i) Santé Publique; (ii) ObEpi-Roche
Weather Factors	Yearly average rain fall (mm/year, 2016-2018) and average temperature (°C, 2016-2018)	Météo France
Spread Factors	(i) Tests (in cities, March 19th to May 25th, 2020), days before 50 cases in hospitals; (ii) Population density (person/m^2, 2020), share of inter-départemental migration (March 16th-17th 2020), share of overcrowded flats (2016), share on non-confined sectors (employed population, 2018)	(i) Santé Publique; (ii) INSEE

References

Arora, Swarnjit S. and Murray Brown (1977). “Alternative Approaches to Spatial Autocorrelation: An Improvement Over Current Practice”. In: *International Regional Science Review* 2.1, pp. 67–78.

Beelen, Rob, Gerard Hoek, et al. (2013). “Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project”. In: *Atmospheric Environment* 72, pp. 10–23.

Beelen, Rob, Ole Raaschou-Nielsen, et al. (2014). “Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project”. In: *The Lancet* 383, pp. 785–795.

BMJ (2020). “Covid-19: risk factors for severe disease and death”. In: *BMJ* 368. Jordan, Rachel E and Adab, Peymane and Cheng, K K.

Booth, James G et al. (2003). “Negative binomial loglinear mixed models”. In: *Statistical Modelling* 3.3, pp. 179–191.

Brandt, Eric B., Andrew F. Beck, and Tesfaye B. Mersha (May 2020). “Air pollution, racial disparities, and COVID-19 mortality.” In: *Journal of Allergy and Clinical Immunology* 51.

Cameron, Colin A. and Pravin K. Trivedi (2003). *A Companion to Theoretical Econometrics*. Ed. by Badi H. Baltagi. 2nd. Blackwell Publishing, pp. 331–348.

CDC (Mar. 2020). “Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States”. In: *Morbidity and Mortality Weekly Report*.

Chopin, Nicolas (Apr. 2020). *A quick, preliminary study of COVID death under-reporting in France*. https://statisfaction.wordpress.com/2020/04/12/a-quick-preliminary-study-of-covid-death-under-reporting-in-france/.

Conticini, Edoardo, Bruno Frediani, and Dario Caro (2020). “Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy?” In: *Environmental Pollution* 261, p. 114465.

Crouse, Dan et al. (Nov. 2019). "Evaluating the Sensitivity of PM2.5-Mortality Associations to the Spatial and Temporal Scale of Exposure Assessment". In: *Epidemiology* 31, p. 1.

Cui, Yan et al. (Dec. 2003). "Air pollution and case fatality of SARS in the People's Republic of China: An ecologic study". In: *Environmental health : a global access science source* 2, p. 15.

Davis, Richard A. and Rongning Wu (July 2009). "A negative binomial model for time series of counts". In: *Biometrika* 96.3, pp. 735–749.

Dockery, D.W et al. (Dec. 1993). "An association between air pollution and mortality in six U.S. cities". In: *N Engl J Med* 329.

Donkelaar, Aaron van et al. (2019). "Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors". In: *Environmental Science & Technology* 53.5.

Dutheil, Frédéric, Julien S. Baker, and Valentin Navel (2020). "COVID-19 as a factor influencing air pollution?" In: *Environmental Pollution* 263, p. 114466.

Gautam, Sneha, Jaydev Teraiya, and Aditya Kuma Patra (2018). "Spatial statistics, spatial correlation and spatial graph theory in air pollution". In: *Environmental Technology Innovation* 11, pp. 384–389.

Grasselli, Giacomo et al. (Apr. 2020). "Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy". In: *JAMA* 323(16), pp. 1574–1581.

Henschel, Susann and Gabrielle Chan (2013). *Health risks of air pollution in Europe - HRAPIE project : New emerging risks to health from air pollution - results from the survey of experts.*

AP-HP (2019). *Liste des tarifs applicables à l'AP-HP*. http://pitiesalpetriere.aphp.fr/wp-content/blogs.dir/58/files/2019/03/Affiche-tarifs-mars2019-2.pdf. Source: Assistance publique-Hôpitaux de Paris.

Lim, Stephen S et al. (2012). “A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010”. In: *The Lancet* 380.

Ma, Yueling et al. (2020). “Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China”. In: *Science of The Total Environment* 724, p. 138226.

Mohebbi, Mohammadreza, Rory Wolfe, and Damien Jolley (2011). “A Poisson regression approach for modelling spatial autocorrelation between geographically referenced observations”. In: *BMC Medical Research Methodology* 11.Art. No: 133, pp. 1–11.

Ogen, Yaron (2020). “Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality”. In: *Science of The Total Environment* 726.

Pelucchi, Claudio et al. (2009). “Long-term particulate matter exposure and mortality: a review of European epidemiological studies”. In: *BMC public health* 9.

Pope III, C. Arden et al. (Mar. 2002). “Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution”. In: *JAMA* 287.9, pp. 1132–1141.

Setti, Leonardo et al. (2020). *The Potential role of Particulate Matter in the Spreading of COVID-19 in Northern Italy: First Evidence-based Research Hypotheses.*

Vienneau, D., K. de Hoogh, and D. Briggs (2009). “A GIS-based method for modelling air pollution exposures across Europe”. In: *Science of The Total Environment* 408.2, pp. 255–266.

WHO (Mar. 2020). “Coronavirus disease 2019 (COVID-19) Situation Report”. In: 51.

Wu, Xiao et al. (2020). *Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study.*

Zu, ZY et al. (Feb. 2020). “Coronavirus Disease 2019 (COVID-19): A Perspective from China”. In: *Radiology* 200490.