Inequality in exposure to air pollution in France: bringing pollutant cocktails into the picture

Camille Salesse

Inequality in exposure to air pollution in France: bringing pollutant cocktails into the picture

Camille Salesse

CEE-M Working Paper 2022-11

Inequality in exposure to air pollution in France: bringing pollutant cocktails into the picture

Camille Salesse*

November 30, 2022

Abstract

I estimate the relationship between income, the number of days of exposure to the four main air pollutants and the proportion of "cocktail days" with French municipal data over the period 2012-2018. I find contrasting results between rural and urban areas. The most affluent urban municipalities have on average a lower number of pollution days compared to the poorest urban municipalities. In urban areas, the pollution days are composed of an equal proportion of cocktail days between the poorest and the most affluent municipalities. On the other hand, in the rural areas the better-off municipalities have on average a higher number of days of pollution, composed of more toxic mixtures, compared to the poorer municipalities. I also show that the pollution levels and the difference in the number of pollution days between the better-off and poorer municipalities are higher in urban areas.

Keywords: air pollution, cocktail, inequality, environmental justice

JEL Codes: D63, Q53, I14

*CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France. E-mail address: camille.salesse@umontpellier.fr

1 Introduction

Exposure to air pollution has negative effects on health, productivity, cognitive performance and educational outcomes[1]. The overexposure of a part of the population to pollution thus reduces its earning potential, challenges the principle of equal opportunity and increases health inequalities. Quantifying environmental inequalities is important for determining optimal environmental policies that take into account these equity concerns. In this paper, I have two main objectives. First, to quantify the difference in air pollution exposure days between the richest and poorest municipalities. Second, to assess inequalities in exposure to air pollutant cocktails.

This study focuses on the four main pollutants identified by the World Health Organization (WHO) and regulated in France: fine particulate matter (PM2.5), PM10, nitrogen dioxide (NO2) and ozone (O3). The health effects of these pollutants are widely documented (Deryugina et al., 2019; Knittel, Miller, and Sanders, 2016; Anderson, 2020; Moretti and Neidell, 2011) and each of them has a recommended concentration threshold defined by the WHO. I evaluate the exposure to these pollutants because of their different origins and spatial dispersion patterns, some of which are more urban and others rural (Figure 1). In addition, exposure to these different pollutants can accumulate and form an even more toxic mixture[2]. For these reasons, it is necessary to assess inequalities in exposure to the different main pollutants, taking into account their simultaneous presence when measuring exposure.

I use French daily concentration data for these pollutants as well as socio-economic data at the municipal level over the period 2012-2018. I exploit the daily availability of our pollution data to construct a new exposure indicator measuring the number of days in the year for which the concentration thresholds defined by the WHO for these pollutants are exceeded. Exposure to pollutant cocktails is measured by the share of cocktail days[3] in the total number of days for which at least one pollutant is above the thresholds. I use these indicators to examine the relationship between municipal income[4], the number of pollution days and the share of cocktail days among these pollution days. I examine these relationships in rural and

[1]See Aguilar-Gomez et al. (2022) for a literature review. See Anderson (2020) and Deryugina et al. (2019) for a causal analysis of the effect of air pollution on mortality, Carneiro, Cole, and Strobl (2021) and Ebenstein, Lavy, and Roth (2016) for a recent analysis of the consequences of pollution on educational outcomes, Graff Zivin and Neidell (2012) and Chang et al. (2019) for an analysis of the effects of pollution on productivity.

[2]Beyond the addition of risks induced by exposure to multiple pollutants, the mixture of pollutants can also generate synergistic effects between pollutants and further increase the health risk. This cocktail effect on health remains very poorly documented.

[3]i.e. days on which at least two pollutants are above the health thresholds.

[4]Measured through the standard of living: household disposable income divided by the number of consumption units.

urban areas separately, on the one hand because these areas concentrate different types of pollutants and on the other hand because the literature has already shown a heterogeneity of results between these two areas. In all our main specifications, I control for the year fixed effect and the regional fixed effect (département level).

I find contrasting results between rural and urban areas. I show that fine particulate matter (PM2.5) and nitrogen dioxide (NO2) account a greater number of pollution days compared to the other pollutants (PM10 and ozone) and they are highly concentrated in urban areas. In urban areas, our results suggest that disadvantaged municipalities accumulate more days of exposure to fine particulate matter (PM2.5) and nitrogen dioxide (NO2), all else being equal. In this area, a one standard deviation increase in income is associated with a decrease of 8.57 nitrogen dioxide (NO2) pollution days and 2.3 fine particulate matter (PM2.5) days on average over the year. In urban areas, the pollution days are composed of an equal proportion of cocktail days on average between the poorest and richest municipalities. In rural areas, the results are different. I show that the richest rural municipalities accumulate more fine particulate matter (PM2.5), particulates PM10 and nitrogen dioxide (NO2) pollution days compared to the least rich rural municipalities. However, the exposure gap between rich and poor municipalities, other things being equal, is smaller compared to urban areas. Pollution days have on average a significantly higher proportion of cocktail days for affluent rural municipalities compared to disadvantaged municipalities.

This paper relates to the literature that measures environmental inequalities in exposure to air pollution[5]. Many studies have shown the overexposure of the poorest and minorities to air pollution in the US, there is less evidence in France and Europe. While most studies consider only a limited number of pollutants for which exposure is measured in most cases through the annual average concentration of the pollutant, I estimate for the first time the unequal simultaneous exposure to air pollutants over the whole territory. I propose a new method for assessing exposure to air pollution cocktails that can be applied in studies evaluating the effects of air pollution exposure on multiple outcomes.

The rest of the paper is organized as follows. The following section presents the related literature. Section III describes the data. Section IV presents our empirical strategy. Section V presents the results, Section VI the robustness tests and the last section concludes.

[5]See Banzhaf, Ma, and Timmins (2019b) for a review of the literature on measuring these inequalities and analyzing the mechanisms of their formation.

2 Related Literature

In the US, areas with a higher proportion of minorities and disadvantaged people are exposed to higher concentrations of fine particulate matter (PM2.5) (Jbaily et al., 2022) and industrial air pollutants (Ard, 2015). Mikati et al. (2018) also find these results using fine particulate matter emissions data[6]. The results differ depending on the level of data aggregation[7] and the area considered. Hsiang, Oliva, and Walker (2019) show that the relationship between income and nitrogen dioxide (NO2) exposure is positive at the Metropolitan statistical area (MSA) level while it takes a U-shape at the census block level, the relationship becomes negative when estimated within the MSA. Rosofsky et al. (2018) use socio-demographic data at the block group unit and concentrations data for fine particulate matter (PM2.5) and nitrogen dioxide (NO2) in Massachusetts to assess environmental inequalities in rural and urban areas. They show descriptive evidence that the poorest and minorities are more exposed to pollution in urban areas, while in rural areas inequalities are lower. Finally, the extent of exposure inequalities within regions varies across the US (Zwickl, Ash, and Boyce, 2014). Several studies have looked specifically at inequalities in exposure between ethnic groups in the US. Christensen, Sarmiento-Barbieri, and Timmins (2022) show that African American or Hispanic tenants have a lower probability of response from homeowners in areas further away from emission sources where the concentration of pollutants is lower. Gillingham and Huang (2021) show that air pollution from port activity[8] disproportionately increases the number of hospitalizations of African Americans. This is due to an overexposure of African Americans but also a potential greater vulnerability. More generally, African Americans are exposed to higher concentrations of fine particulate matter (PM2.5), PM10 and nitrogen dioxide (NO2) compared to white Americans but this gap has narrowed in recent years (Currie, Voorheis, and Walker, 2020; Kravitz-Wirtz et al., 2016).

Several studies have also highlighted inequalities in exposure to air pollution in Europe. Germani, Morone, and Testa (2014) show an inverted U-shaped relationship between income and industrial pollutant emissions at the level of Italian provinces. Neier (2021) uses socio-economic data from a fine grid of the Austrian territory combined with industrial pollutant emission data. He shows that a higher proportion of foreigners in the area is associated with a higher exposure to pollution. Moreover, in urban areas the better-off are less exposed to pollution while in rural areas the relationship is reversed.

[6] These studies used geographically fine-grained data at the ZIP code tabulation area (ZCTA) and census block level.

[7] Banzhaf, Ma, and Timmins (2019b) emphasize the importance of using the least aggregated data possible to avoid the ecological fallacy problem.

[8] carbon monoxide (CO), nitrogen dioxide (NO2), fine particulate matter (PM2.5), and sulfur dioxide (SO2).

In France, several evidence suggest the overexposure of the poorest and minorities to air pollution and more particularly in urban areas. Our study is most closely related to Champalaune (2021). She uses administrative data at the neighborhood level and annual average estimates of fine particulate matter concentrations at a fine scale over the whole French territory. She shows through an individual fixed effect model that within urban and peri-urban areas, the poorest neighborhoods and those with a higher proportion of migrants are more exposed to fine particulate matter. On the other hand, in rural areas, areas with a lower share of migrants are more exposed and the relationship between income and exposure is not significant. These estimates are robust to spatial autocorrelation (using a model close to the SAR model, see empirical strategy section). Laurian and Funderburg (2014) show that towns with a higher share of migrants are more exposed to incinerators in France. Ouidir et al. (2017) assess the exposure of pregnant women in France to three air pollutants: fine particulate matter (PM2.5), particles PM10 and nitrogen dioxide (NO2). The study shows that in urban areas the most disadvantaged women are more exposed to the three pollutants, while in rural areas it is the most affluent and disadvantaged women who are most exposed. Padilla et al. (2014) use socio-economic data at the neighbourhood level in 4 large French cities and show that the strength and direction of the link between wealth, share of migrants and NO2 exposure can vary between cities[9]. Fosse, Salesse, and Viennot (2022) show descriptive evidence of a U-shaped relationship between fine particulate matter (PM2.5) emissions and income of municipalities in France and an inverted U-shaped relationship between ammonia (NH3) emissions and income.

Overall, in the US and Europe, there is an overexposure of the poorest and minorities to air pollution. However, inequalities are more pronounced in urban areas. They are also greater in some regions/cities. Most studies consider only a limited number of pollutants and exposure is measured in most cases by the annual average concentration of the pollutant. I estimate for the first time the inequality in exposure to pollutant cocktails, analysing urban and rural areas separately. I also explicitly take into account the problem of spatial autocorrelation in a series of robustness tests.

3 Data

I build a database at the communal level including about 30 000 French municipalities over the period 2012-2018[10]. I use PREV'AIR (INERIS) background pollution data as well

[9] The estimates account for spatial autocorrelation with a model similar to Champalaune (2021).

[10] Over the period studied, several municipalities merged with each other. We have therefore retained in the database the municipalities after the merger and excluded the data of the formerly autonomous municipalities. In the pre-merger period, for the new communes resulting from the mergers, we aggregate

as localized socio-economic data from the FILOSOFI (INSEE) database and the census (INSEE) that I describe below.

3.1 Pollution data

I use the PREV'AIR air pollution data produced by INERIS for the four main regulated pollutants: ozone (O3), nitrogen dioxide (NO2), fine particulate matter PM2.5 and PM10. For each day of the year, I have the average concentration in micrograms per cubic meter of air (μg/m^3) of each of these pollutants in each of the 34,000 French municipalities over the 2012-2018 period.

These data are constructed using the CHIMERE pollutant transport simulation model. From pollutant emission inventory data, sampling data, metrological data and other types of input data, the model is able to predict pollution concentrations over the entire territory at a fine scale. An additional statistical treatment based on real observations of concentrations is then applied to the predictive outputs of the CHIMERE model allowing to further increase the reliability of the concentration estimates. PREV'air data are very reliable data available on the whole French territory with a fine resolution of about 4km2.

Based on the reliability of the pollution concentration estimates and the daily availability of our data, I calculate for each municipality the number of days in the year for which the threshold defined by the WHO was exceeded for each of the 4 pollutants. I use the most recent WHO thresholds (2021).

_ Fine particulate matter (PM2,5) : 15 μg/m^3 24-hour average value

_ Coarse particles (PM10) : 45 μg/m^3 24-hour average value

_ Nitrogen dioxide (NO2) : 25 μg/m^3 24-hour average value

Concerning ozone (O3), the threshold recommended by the WHO is 100 μg/m^3, maximum daily value of 8-hour rolling averages. I do not have measurements in 8-hour rolling averages, so I define our threshold at 100 μg/m^3 24-hour average value.

I choose to use the number of days for which the concentration of the pollutant is above the WHO threshold as an exposure indicator. To our knowledge, this indicator has never been used before and I argue that it improves the quality of the analysis. Indeed, I argue that the overexposure to pollution of a part of the population is a problem insofar as this population will suffer significant negative health consequences. Indeed, environmental inequalities measured from the annual average concentration of the pollutant are not systematically associated with a significant increase in health risk for the most exposed population. If one population is exposed to an average of 3 μg/m^3 of PM2.5 particles over the year and another

the data of the municipalities forming the future merger by averaging them in order to have in the database only the municipalities newly formed by the merger over the entire period studied.

to 1 μg/m^3 of PM2.5 particles, in this case there is indeed an environmental inequality but the health consequences for the overexposed population remain limited since neither group exceeds the WHO threshold of 5 $\mu g/m^3$ annual average.

I argue that environmental inequalities must be considered in terms of the health consequences they cause, and our indicator of the number of days above the thresholds allows us to identify exposure gaps with significant health risks.

3.2 Cocktails

I exploit the daily availability of pollution data and the very high reliability of the estimates to define several cocktails of pollutants. The objective is to identify the unequal exposure of the municipalities to various mixtures of pollutants. For this purpose, for each municipality, each day in the year for which the threshold of at least one pollutant is exceeded is classified in one of the 15 possible "pure" mixtures. This method allows us to define 11 pure mixtures without interaction with other pollutants (the pollutants of the considered mixture are above the WHO thresholds and the other pollutants are below the thresholds) and 4 pure monopollutant mixtures (only the considered pollutant is above the thresholds). I sum for each year for each municipality the number of days that each mixture represents. From these data I build an exposure indicator which measures the total number of pollution days of the municipality in the year by adding all the pollution days of the 15 possible mixtures. I also construct an exposure indicator to pollutant cocktails by calculating for each municipality each year the share that each mixture represents in the total number of pollution days. I also calculate the share of the "cocktail" mixtures in the total number of pollution days.

Figure 2 shows the distribution of all pollution days in each of the possible groups. 62% of the days above the WHO thresholds, for at least one pollutant, are days where only the PM2.5 threshold is exceeded. 19% of the pollution days are a pure mixture of PM2.5 and nitrogen dioxide (NO2), 6% pure nitrogen dioxide (NO2), 5% NO2 PM2.5 and PM10, 4% O3, 3% PM2.5 PM10 (see Table 1 descriptive statistics).

3.3 Socio-economic data

I match pollution data with several types of socio-economic data. First, I use the FILOSOFI database produced by INSEE, which contains a wide variety of indicators on available income at the municipal level. The data are available for the years 2012 to 2018. Our main indicator for measuring the wealth of a municipality is the median standard of living of the municipality (the standard of living is the disposable household income divided by the number of consumption units). Data on this indicator are available for almost all the municipalities. I also mobilize other variables such as the poverty rate, the first decile and the ninth decile of standard of living. However, data for these indicators are only available for municipalities with more than 1000 households or more than 2000 inhabitants. I also use many socio-economic indicators from INSEE census data such as the share of each socio-professional category in the municipality or the share of women and men in the municipality, the composition by age and type of household in the municipality (see Table 1 descriptive statistics).

3.4 Descriptive statistics

The richness of this study lies in the consideration of various forms of air pollution that have different origins. The correlations Table 2 show that the different pollutants are not perfectly correlated with each other, so it is relevant to study these different forms of pollution separately. The origin of the emissions of these four pollutants partly explains the imperfect correlation between them. About half of the fine particulate matter come from residential heating (wood heating is largely responsible for these emissions) and about a quarter from road transport. The concentration of nitrogen dioxide in the air comes mainly from the road sector and, to a lesser extent, from industry. PM10 emissions come from residential heating, industry but also from construction sites which produce large particles. Ozone is not emitted directly by human activities, it is formed by a chemical reaction between nitrogen oxides and other components in interaction with solar radiation[11].

The maps Figure 3 show the dispersion of the 4 pollutants over the whole French territory in 2012 and 2018 (the same scale was used from one year to another). First, on average over the period, pollution levels decreased for PM10, PM2.5 and nitrogen dioxide (NO2) but not for ozone (see Figure 3 and Figure A.3). Secondly, the geographical distribution of the four pollutants at the national level and in rural and urban areas is not similar (which confirms the relevance of studying exposure to these different pollutants without being limited to fine particulate matter).

3.5 Aggregation level & ecological fallacy

The problem of the ecological fallacy is to erroneously infer a correlation at the individual or higher level from a correlation obtained from more aggregated data. The way to avoid this problem is to use the least aggregated data possible, but also to infer conclusions only at the level of data aggregation. The PREV'AIR pollution data are background pollution data, which means that although the resolution of the data is very fine, areas of high pollution that

[11]See www.airparif.asso.fr, fine particles, origin and sources of pollution, Fine particles — Airparif.

are very localized, such as areas with heavy road traffic, have very smoothed concentration levels.

I argue that the level of aggregation of our data at the municipality level is the most relevant for our study, whose objective is to investigate unequal exposure to national background pollution at the municipality level. The municipality level of our data allows to take into account part of the exposure to pollution when individuals move close to their home (to work for example). The municipal level is suitable for background pollution data, as the concentration of background pollutants is homogeneous over most municipalities. There is a finer level of administrative division (IRIS) adapted to very localized pollution data but which is not however adapted to our study devoted to background pollution. The French municipalities are very numerous (more than 30 000) and form a fine grid of the French territory allowing to avoid the problem of ecological fallacy[12].

4 Empirical strategy

4.1 Specification

I estimate the following fixed-effects linear regression model for municipality c, in département (region) d in year t, in each of the sub-samples of urban and rural municipalities.

$$y_{cdt} = \beta x_{cdt} + \Theta z_{cdt} + \lambda_d + \alpha_t + \varepsilon_{cdt} \tag{1}$$

y_{cdt} is the number of days in year t in municipality c of department d for which the concentration levels of the pollution considered exceeded the WHO health threshold. x_{cdt} is the median standard of living in the municipality. The vector z_{cdt} includes a set of control variables such as the share of the population belonging to each socio-professional category, the share of women and men, the share of each age group, the share of each type of household. The regressions include département α_d and year λ_t fixed effects. I cluster standard errors at the département level[13]. I estimate this model in each of the 6 types of rural or urban communes defined by INSEE on the basis of the density of the municipality and its attraction to a pole. I also estimate this model at the national level in the appendix.

[12] Many studies on environmental inequality in the United States have used more aggregated data at the county level, which is equivalent to the French département.

[13] The traditional justification for clustering standard errors is to account for the correlation of residuals within département. Abadie et al. (2022) propose a new framework in which clustering is justified by two parameters: sampling and treatment design. In our case, only the treatment design is important since we have the full set of available data. Since income is partially correlated within departments, the robust standard errors (Eicker-Huber-White) are too small and the clustered standard errors at the department level too large (conservative). We prefer clustering to minimize the risk of bias.

I estimate this fixed-effect model to identify the correlation between the wealth of municipalities and their level of exposure to pollution, holding socio-demographic characteristics, department and year constant. This specification deliberately does not resolve the potential reverse causality of the income-pollution relationship. Indeed, environmental inequalities can be formed either by a variation in the wealth of the municipalities inducing a variation in pollution levels, or by the opposite phenomenon. The objective of this study is to identify environmental inequalities as they currently exist and not to identify the mechanism of formation. Identifying the causal effect of wealth on pollution is therefore not the objective of this paper.

To estimate the relationship between income and pollution, I exploit the variation within different geographical areas at different time periods in pollution levels and income of the municipalities while controlling for their socio-economic characteristics. I do not include municipality fixed effects because the objective is not to study the mechanism by which income variations influence pollution or vice versa or both at the same time within municipalities.

4.2 Spatial autocorrelation

I argue in this section that our fixed-effects model must account for the spatial dependence of our observations. Air pollution propagates continuously through space in a non-random manner (Figure 3). As a result of this spatial dependence or spatial autocorrelation, OLS estimation can be either inefficient or biased. To gauge the importance of this problem, I perform a series of tests on our OLS estimates.

There are 3 main specifications to account for different types of spatial interactions with panel data, the Spatial Error Model (SEM), the spatial autoregressive model (SAR) and the Durbin model which is a combination (Bouayad Agha, Le Gallo, and Vedrine, 2018).

To determine which of these models to use, I perform a Lagrange multiplier test on the residuals of the OLS model for error dependence (LM-error) and for the presence of the spatially lagged dependent variable in the model (LM-lag)(Anselin et al., 1996). I also perform the robust version of these tests to the presence of the other type of spatial relationship.

Depending on the results of the tests I estimate the SEM or SAR models (see appendix details). The SAR model reduces the bias of the OLS estimates while the SEM model has an impact on the efficiency of the model. These models are estimated by the maximum likelihood (ML) method with fixed effects[14].

[14] Splm package on R.

5 Results

5.1 Income and number of days of exposure to pollutants

I estimate the relationship between income and the number of days of exposure to the pollutants in the rural and urban areas in Table 3. The poorest dense urban municipalities accumulate on average significantly more days of fine particulate matter (PM2.5) and nitrogen dioxide (NO2) pollution compared to the better-off municipalities, all else being equal. The increase of one standard deviation[15] in income is associated with a decrease of 8.57 days of pollution for nitrogen dioxide (NO2) and 2.3 days for fine particulates (PM2.5) over the year. It seems that income is not significantly related to the level of PM10 and ozone (O3) pollution in this area. In urban municipalities with intermediate densities, no coefficient of income is significant for the 4 pollutants. In rural areas, income is significantly positively related to nitrogen dioxide (NO2), fine particulates (PM2.5) and PM10 pollution. Ozone level seems to be unrelated to income.

One additional standard deviation of income in rural areas under weak influence of a pole is associated with an increase of 3.68 days of nitrogen dioxide (NO2) pollution and 0.89 days of PM2.5 pollution and 0.32 days for PM10 on average. In sparsely populated rural areas, the number of additional days for nitrogen dioxide (NO2), fine particulate matter (PM2.5), and PM10 per standard deviation of income is on the order of 2.29 , 0.97 , and 0.29 days on average.

Finally, I also evaluate the link between the wealth of the municipalities and the number of days for which the threshold for at least one of the four pollutants is exceeded (ALL column). The results are similar to the previous results. One additional standard deviation of income is associated with a decrease of on average 8.25 days with at least one pollutant above the thresholds in dense urban areas and an increase of 1.57 days in sparsely populated rural areas.

Figure 1 shows that fine particulate matter (PM2.5) and nitrogen dioxide (NO2) are the pollutants for which the number of days above the thresholds is the highest and these pollutants are concentrated in urban areas (very largely for NO2). I also note that among rural municipalities, the most affluent municipalities are more exposed to PM2.5, PM10 and nitrogen dioxide (NO2). The concentration of these 3 pollutants in rural areas is on average lower compared to urban areas (Figure 1).

It should also be noted that the magnitude of exposure inequalities in terms of additional days of pollution per standard deviation of income is higher in urban municipalities. Not

[15]Calculated on the sub-sample.

only is pollution much higher in terms of level in urban municipalities, but the exposure gap is also larger in urban areas.

The literature developed mainly in the United States has identified mechanisms that can explain the inequalities in exposure to nitrogen dioxides (NO2) and fine particulate matter (PM2.5) that are observed in urban areas. Environmental inequalities can be explained both by the location choice of inhabitants, the location choice of polluting firms and a combination of both (Banzhaf, Ma, and Timmins, 2019a). Wealthier people are more willing to pay the price of not living in polluted municipalities (Banzhaf and Walsh, 2008). Highly imperfect household information on pollution also promotes exposure inequalities (Hausman and Stolper, 2021). Polluting firms may also choose to locate disproportionately in or near less affluent municipalities (Banzhaf, Ma, and Timmins, 2019b).

The relatively small difference in pollution levels that I find between the better-off and poorer municipalities in rural areas can theoretically be explained by both the level of economic activity and the specialization of the municipalities. The rural population is divided into two types of municipalities, the densely populated municipalities that concentrate the most economic activity and the less densely populated municipalities that concentrate less economic activity. Dense municipalities have higher incomes than less dense municipalities, but also higher levels of pollution, both of which are partly explained by the higher economic activity (see annex for correlations between municipality density, income, economic activity and pollution, Figure A.5)

Like rural municipalities, the densest urban municipalities have the highest economic activity and the highest pollution levels. However, the densest urban municipalities have lower incomes than the less dense urban municipalities (see figure Figure A.4). In urban municipalities, the better-off are located in the less dense and less polluted municipalities, while in rural municipalities they are located in the dense municipalities where economic activity and pollution are higher.

The origin of the pollutants may partly explain the non-significant results for ozone. Indeed, ozone, unlike other pollutants, is a so-called "secondary" pollutant, i.e. it is not directly emitted into the atmosphere but results from chemical transformations influenced in particular by solar radiation. This may explain the absence of a significant relationship with income, as ozone is determined more by meteorological than socio-economic variables (the relationship has the lowest R^2 compared to other pollutants).

In this section I have highlighted several results. Firstly, the 4 pollutants studied do not represent the same amount of days above their respective thresholds. Fine particulate matter (PM2.5) and nitrogen dioxide (NO2) have very high average levels in urban areas of around 100 and 117 days respectively over the period. PM10 and ozone (O3) have a maximum average of 13 and 5 days in the areas with the highest values. I have highlighted the significant negative relationship between income and exposure to PM2.5 and nitrogen dioxide (NO2) in urban areas (these areas being the most polluted for these pollutants). Conversely, I find a significant positive relationship between income and PM2.5, PM10 and nitrogen dioxide (NO2) pollution in rural areas. However, the average difference in exposure days according to income is greater in urban areas than in rural areas. The high level of pollution and the large exposure gap to the detriment of the less well-off in urban areas should lead to the implementation of policies to reduce pollutant emissions in these areas.

5.2 Unequal exposure to pollutant cocktails

The toxic effect of a pollutant can be multiplied in the presence of other pollutants, which may explain the differences in toxicity between different mixtures. Nevertheless, the interaction effects of pollutants on health are still poorly understood. To our knowledge, no study has assessed the toxicity associated with each mixture of air pollutants.

In this section, I define for the first time the share of multi-pollutant days compared to single-pollutant days. The impact of air pollution on health depends both on the number of days of exposure to pollutants and on the composition of these pollution days. Some mixtures of pollutants are more harmful than others, what share does each mixture of pollutants represent in the pollution days? In the first part of the analysis, I assessed the link between income and the number of exposure days. In this part, I will assess the link between income and the composition of pollution days.

As described in the data section, each pollution day is associated with one of 15 possible mixtures. The variables I used previously to estimate the pollution-income relationship are a simple measure of the number of days above the threshold for the pollutant of interest, across all mixtures. However, cocktails represent a significant number of days (Figure 2).

Figure 4 shows the average share of each mixture in the total number of days with at least one pollutant above the threshold at the national level and in urban and rural areas. According to these estimates, at the national level, on average 25% of the days with one pollutant above the threshold, the threshold of at least one other pollutant is also exceeded in the municipalities (Figure 4). The cocktail that represents the largest share is the mixture of PM2.5 and NO2. It represents on average 17% of the total pollution days. The composition of the pollution days varies between rural and urban areas. In dense urban areas 43% of pollution days are cocktail days compared to 34.6% in intermediate density urban municipalities and 23.5% in rural areas on average. Nitrogen dioxide, a very urban pollutant has a higher share in pure and mixed form in the pollution days of urban municipalities. The fine particulate matter PM2.5 are also a very present pollutant in the urban municipalities which favor the mixture with nitrogen dioxide. Fine particulate matter are also present in rural municipalities but not nitrogen dioxide which is much less present.

I have just shown that our measure of exposure to a pollutant in terms of the number of days above the threshold is composed of a significant proportion of "cocktail days". The shares of each mixture in the total number of pollution days that I have estimated in Figure 4 are averages. Some municipalities may therefore have a higher or lower share of cocktail days than the average.

Table 4 presents the results in each of the rural[16] and urban areas. I use Z-score on the income variable to facilitate interpretation of the coefficients[17]. The composition of pollution days can be influenced by the number of pollution days in the municipality. Since income is related to the number of pollution days, I add it as a control variable to avoid an omitted variable bias. I also include all control variables and fixed effects from previous regressions. In dense and intermediate density urban areas, the share of cocktail days is on average similar between poor and more affluent municipalities. In contrast, among rural municipalities, wealthier municipalities have a significantly higher share of cocktail days on average. One additional standard deviation of income is associated with a 1.75 percentage point increase in the cocktail share.

To investigate these results further, I now evaluate the relationship between income and the share of each of the major mixtures (those with a significant share, see Figure 4) represents. I calculate the share that each mixture represents in the total number of pollution days. Table 5 shows a non-significant relationship between the richness and the share of each mixture in urban areas. This result is consistent with our previous results. The pollution composition between rich and poor municipalities is on average similar in urban areas. Table 5 shows which mixes are related to income in rural areas. On average in this area, the pollution days in the richer municipalities are composed of a significantly lower share of pure fine particulate matter (PM2.5) and a higher share of a mixture of nitrogen dioxide (NO2) and fine particulate matter (PM2.5) , nitrogen dioxide (NO2), fine particulate matter (PM2.5) and PM10 but also pure nitrogen dioxide (NO2). One additional standard deviation of income is associated with an average decrease of 2.8 percentage points in the share of pure fine particulate matter and an increase of 1.5 points in the share of NO2 PM2.5, 0.2 points in NO2 PM2.5 PM10 and 1 point in pure NO2.

In the wealthy rural areas the nitrogen dioxide (NO2) takes a more important place in

[16] The results are similar in the different types of rural areas (as in the first part), so we consider the rural municipalities as a whole.

[17] The variable is standardized at the national level, which implies that the standard deviation of the variable in the subsamples is not equal to one.

pure and mixed form compared to the poor areas. As I have shown, in this area the better-off accumulate more nitrogen dioxide (NO2) pollution days and the coefficient is the most important compared to the other pollutants (Table 3). This surplus of NO2 is associated with a change in the composition of pollution days.

As NO2 is a pollutant from urban emissions, this result is consistent with our theoretical explanation of the overexposure of the better-off in rural areas. Affluent rural municipalities are denser and more urban, concentrating more economic activity favoring the formation of NO2, in a proportion nevertheless much lower than urban areas (see appendix the correlations between density of the municipalities, income, economic activity and composition of the pollution Figure A.5).

Overall, among urban municipalities, pollution days have on average the same composition in the poorest and most affluent municipalities. On the other hand, in rural areas, pollution days are composed of a larger share of cocktail days in the better-off municipalities.

6 Robustness

6.1 Measures of wealth

Income is more heterogeneous in urban municipalities compared to rural municipalities. The median standard of living is a good socio-economic indicator, but information about the top and bottom of the standard of living distribution is lost, particularly in urban municipalities. I use a set of data on the poverty rate, the first and ninth decile of the standard of living, the share of social benefits in income and the share of income from assets in the total income of the municipality. These data from the FILOSOFI database are only available for municipalities with more than 1,000 households or more than 2,000 people, a condition met by almost 100% of dense urban municipalities and 66% of intermediate density urban municipalities.

Table 6 reports the results for the subsample of dense urban municipalities. Each coefficient is associated with a regression. The coefficients for the poverty rate, the first decile and the share of assets in income are highly significant for nitrogen dioxide (NO2), particulate matter (PM2.5) and for the measurement of days with at least one pollutant above the threshold. The direction and significance of these coefficients point to an inequality of exposure to the disadvantage of the poorest for nitrogen dioxide (NO2) and PM2.5 particles, which confirms our previous results. PM10 and ozone (O3) have non-significant coefficients which further affirms our previous results.

Table 7 presents the results for the subsample of intermediate density urban municipalities. The vast majority of the coefficients are insignificant in accordance with our previous estimates, with the exception of ozone (O3). Ozone (O3) appears to be more concentrated in the more affluent intermediate density urban municipalities. Estimation by new wealth indicators generally supports our results with the exception of ozone (O3).

Table 8 shows the results for the cocktail share in urban municipalities. No coefficients are significant for all wealth measures which confirms our results.

6.2 Spatial autocorrelation

In this section I explicitly take into account spatial autocorrelation in our estimates. First, as described in the empirical strategy section, I perform a series of tests on the residuals of our OLS estimates. For each regression performed on each subsample the results are Table 9. For computational reasons I perform the tests on the error term of the national regression taking data only for the year 2018, (the number of municipalities at the national level is very large which requires many computational resources). The test results for each year are similar to the year 2018 (results not shown). For the same reason of computational limitation, I do not perform the tests on the errors of the subsample of sparsely populated rural municipalities (which has the most municipalities among the subsamples).

I begin by performing a Moran test on the error terms of the national estimates (these regressions are weighted by the population of each municipality, see appendix). For all specifications I reject the hypothesis of no spatial autocorrelation of the error terms. The Moran test allows us to formally confirm the presence of spatial autocorrelation, the following tests allow us to identify the spatial model best suited to our specifications. The results of the LM-error and LM-lag tests on the different fixed effects models do not allow us to decide in favor of a particular model, I reject the null hypothesis of both tests. On the other hand, the robust version of these tests leads us to favor the implementation of a spatially autocorrelated error model (SEM). This model consists in estimating a spatial autocorrelation parameter in the error term (see appendix E). I can already conclude from the results of these tests that the majority of our specifications estimated by OLS are not biased because of autocorrelation, but they are less efficient.

By specifying the SEM model I model the spatial correlation of errors as being caused by the spatial correlation of variables affecting pollution levels but which are not correlated with the independent variables in our model. Meteorological variables are a plausible explanation for these spatial relationships in our error term. The results of the SEM model estimates are presented in Table 10. The spatial autocorrelation coefficients ρ are positive and highly significant. The SEM model estimates support our previous results.

7 Conclusion

The health risk associated with exposure to air pollution depends on the time of exposure but also on the toxicity of the pollutants. In this study, I evaluate the relationship between income, the number of days of exposure to air pollutants and the relationship between income and the composition of these pollution days.

First, I showed that urban areas concentrate the most days of fine particulate matter (PM2.5) and nitrogen dioxide (NO2) pollution. Among the dense urban municipalities, the poorest are significantly overexposed to these two pollutants in terms of number of days, all other things being equal. In urban area, income is not related to the number of PM10 and ozone (O3) pollution days. The share of cocktail days is on average similar between the poorest and the most affluent municipalities in the urban area.

In rural areas, the most affluent municipalities accumulate more PM2.5, PM10 and nitrogen dioxide (NO2) pollution days, all else being equal. However, rural areas are less polluted and the exposure gap between affluent and disadvantaged municipalities is smaller compared to urban areas. Among rural municipalities, the most affluent municipalities have on average a significantly higher share of cocktail days.

I showed in this paper the high concentration of pollutants in urban areas and the important overexposure of the poorest within these areas. Several studies have shown the vulnerability of the poorest to air pollution (Deguen et al., 2015; Morelli et al., 2016; Hsiang, Oliva, and Walker, 2019; Deryugina et al., 2021) which reinforces the need to prioritize the implementation of public policies to reduce pollution in urban areas and particularly in the poorest urban areas. It is also necessary to take environmental justice into account in public policies aimed at reducing pollution so as not to further reinforce inequalities in exposure. For example, recently in Paris, the closure of roads has created a phenomenon of displacement of road traffic on roads close to the homes of the poorest and therefore increased the pollution exposure of this population (Bou Sleiman, 2022). Thus, although pollution levels have tended to decrease in France in recent years, public policies aimed at reducing pollution must also take into account the distributional aspect.

References

Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* .

Aguilar-Gomez, Sandra, Holt Dwyer, Joshua S. Graff Zivin, and Matthew Neidell. 2022. "This is Air: The "Non-Health" Effects of Air Pollution." *National Bureau of Economic Research* .

Anderson, Michael 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.

Anselin, Luc, Anil K. Bera, Raymond Florax, and Mann J. Yoon. 1996. "Simple diagnostic tests for spatial dependence." *Regional Science and Urban Economics* 26 (1):77–104.

Ard, Kerry. 2015. "Trends in exposure to industrial air toxins for different racial and socioeconomic groups: A spatial and temporal examination of environmental inequality in the U.S. from 1995 to 2004." *Social Science Research* 53:375–390.

Banzhaf, H. Spencer, Lala Ma, and Christopher Timmins. 2019a. "Environmental Justice: Establishing Causal Relationships." *Annual Review of Resource Economics* 11 (1):377–398.

Banzhaf, H. Spencer and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout." *American Economic Review* 98 (3):843–63.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins. 2019b. "Environmental Justice: The Economics of Race, Place, and Pollution." *Journal of Economic Perspectives* 33 (1):185–208.

Bou Sleiman, Lea. 2022. "Displacing Congestion: Evidence from Paris." .

Bouayad Agha, Salima, Julie Le Gallo, and Lionel Vedrine. 2018. "Chapter 7. Spatial econometrics on panel data." *Handbook of Spatial Analysis, Theory and Application with R, Insee Méthodes N°131* .

Carneiro, Juliana, Matthew A. Cole, and Eric Strobl. 2021. "The Effects of Air Pollution on Students' Cognitive Performance: Evidence from Brazilian University Entrance Tests." *Journal of the Association of Environmental and Resource Economists* 8 (6):1051–1077.

Champalaune, Pascale. 2021. "Inequality in Exposure to Air Pollution in France: Evidence from a City-Level Public Policy." .

Chang, Tom Y., Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2019. "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China." *American Economic Journal: Applied Economics* 11 (1):151–72.

Christensen, Peter, Ignacio Sarmiento-Barbieri, and Christopher Timmins. 2022. "Housing Discrimination and the Toxics Exposure Gap in the United States: Evidence from the Rental Market." *The Review of Economics and Statistics* 104 (4):807–818.

Currie, Janet, John Voorheis, and Reed 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." *National Bureau of Economic Research* .

Deguen, Séverine, Claire Petit, Angélique Delbarre, Wahida Kihal, Cindy Padilla, Tarik Benmarhnia, Annabelle Lapostolle, Pierre Chauvin, and Denis Zmirou-Navier. 2015. "Neighbourhood Characteristics and Long-Term Air Pollution Levels Modify the Association between the Short-Term Nitrogen Dioxide Concentrations and All-Cause Mortality in Paris." *PLOS ONE* 10 (7).

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review* 109 (12):4178–4219.

Deryugina, Tatyana, Nolan Miller, David Molitor, and Julian Reif. 2021. "Geographic and Socioeconomic Heterogeneity in the Benefits of Reducing Air Pollution in the United States." *Environmental and Energy Policy and the Economy* 2:157–189.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution." *American Economic Journal: Applied Economics* 8 (4):36–65.

Fosse, Julien, Camille Salesse, and Mathilde Viennot. 2022. "Inégalités environnementales et sociales se superposent-elles ?" *La note d'analyse de France Stratégie* 112 (7):1–16.

Germani, Anna Rita, Piergiuseppe Morone, and Giuseppina Testa. 2014. "Environmental justice and air pollution: A case study on Italian provinces." *Ecological Economics* 106:69–82.

Gillingham, Kenneth and Pei Huang. 2021. "Racial Disparities in the Health Effects from Air Pollution: Evidence from Ports." *National Bureau of Economic Research* .

Graff Zivin, Joshua and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." *American Economic Review* 102 (7):3652–73.

Hausman, Catherine and Samuel Stolper. 2021. "Inequality, information failures, and air pollution." *Journal of Environmental Economics and Management* 110.

Hsiang, Solomon, Paulina Oliva, and Reed Walker. 2019. "The Distribution of Environmental Damages." *Review of Environmental Economics and Policy* 13 (1):83–103.

Jbaily, Abdulrahman, Xiaodan Zhou, Jie Liu, Ting-Hwan Lee, Leila Kamareddine, Stéphane Verguet, and Francesca Dominici. 2022. "Air pollution exposure disparities across US population and income groups." *Nature* 601 (7892):228–233.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health." *The Review of Economics and Statistics* 98 (2):350–366.

Kravitz-Wirtz, Nicole, Kyle Crowder, Anjum Hajat, and Victoria Sass. 2016. "The Long-Term Dynamics of Racial/Ethnic Inequality in Neighborhood Air Pollution Exposure, 1990-2009." *Du Bois Review: Social Science Research on Race* 13 (2):237–259.

Laurian, Lucie and Richard Funderburg. 2014. "Environmental justice in France? A spatio-temporal analysis of incinerator location." *Journal of Environmental Planning and Management* 57 (3):424–446.

Mikati, Ihab, Adam F. Benson, Thomas J. Luben, Jason D. Sacks, and Jennifer Richmond-Bryant. 2018. "Disparities in Distribution of Particulate Matter Emission Sources by Race and Poverty Status." *American Journal of Public Health* 108 (4):480–485.

Morelli, Xavier, Camille Rieux, Josef Cyrys, Bertil Forsberg, and Rémy Slama. 2016. "Air pollution, health and social deprivation: A fine-scale risk assessment." *Environmental Research* 147:59–70.

Moretti, Enrico and Matthew Neidell. 2011. "Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles." *The Journal of Human Resources* 46 (1):154–175.

Neier, Thomas. 2021. "Austrian Air – Just Clean for Locals: A Nationwide Analysis of Environmental Inequality." *Ecological Economics* 187:107100.

Ouidir, Marion, Johanna Lepeule, Valérie Siroux, Laure Malherbe, Frederik Meleux, Emmanuel Rivière, Ludivine Launay, Cécile Zaros, Marie Cheminat, Marie-Aline Charles, and Rémy Slama. 2017. "Is atmospheric pollution exposure during pregnancy associated with individual and contextual characteristics? A nationwide study in France." *Journal of Epidemiology and Community Health* 71 (10):1026.

Padilla, Cindy M., Wahida Kihal-Talantikite, Verónica M. Vieira, Philippe Rossello, Geraldine Le Nir, Denis Zmirou-Navier, and Severine Deguen. 2014. "Air quality and social deprivation in four French metropolitan areas—A localized spatio-temporal environmental inequality analysis." *Environmental Research* 134:315–324.

Rosofsky, Anna, Jonathan I. Levy, Antonella Zanobetti, Patricia Janulewicz, and M. Patricia Fabian. 2018. "Temporal trends in air pollution exposure inequality in Massachusetts." *Environmental Research* 161:76–86.

Zwickl, Klara, Michael Ash, and James K. Boyce. 2014. "Regional variation in environmental inequality: Industrial air toxics exposure in U.S. cities." *Ecological Economics* 107:494–509.

Tables

Table 1: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
Pollutions					
O3	217,889	4.375	5.324	0	53
NO2	217,889	26.554	29.645	0	325
PM25	217,889	77.409	31.645	0	279
PM10	217,889	6.854	7.019	0	61
ALL	217,889	85.941	35.314	5	331
Cocktails					
Pure PM2.5	217,889	52.623	20.848	0	194
Pure NO2	217,889	5.357	12.308	0	210
Pure O3	217,889	3.162	4.746	0	50
NO2 PM2.5	217,889	16.736	16.528	0	167
NO2 PM2.5 PM10	217,889	4.436	5.472	0	55
PM2.5 PM10	217,889	2.407	2.959	0	20
PM2.5 O3	217,889	1.184	1.923	0	23
Social variables					
Median standard of living	217,681	20,770	3,016.255	9,958	53,500
poverty rate (60 % threshold)	30,667	13.219	6.417	1.000	54.000
share of social benefits	36,145	5.137	2.567	0.400	27.800
first decile	36,145	12,235	1,975	6,153	21,800
ninth decile	36,145	36,497	8,370	21,863	145,730
share of income from assets	36,145	10.235	4.685	2.100	74
Control variables					
other household without family percentage	217,840	0.018	0.024	0.000	0.273
household 1 person percentage	217,840	0.275	0.078	0.000	0.667
household with family with child percentage	217,840	0.311	0.089	0.000	0.700
household with single-parent family percentage	217,840	0.071	0.045	0.000	0.455
household with family without children percentage	217,840	0.324	0.067	0.000	0.727
pop 0 14 percentage	217,840	0.184	0.041	0.000	0.364
pop 15 29 percentage	217,840	0.137	0.032	0.000	0.525
pop 30 44 percentage	217,840	0.188	0.037	0.023	0.382
pop 45 59 percentage	217,840	0.178	0.049	0.010	0.545
pop 60 74 percentage	217,840	0.178	0.049	0.010	0.545
pop 75 89 percentage	217,840	0.086	0.037	0.000	0.441
pop 90 plus percentage	217,840	0.010	0.010	0.000	0.179
pop femme percentage	217,840	0.501	0.022	0.242	0.711
inactive percentage	217,840	0.244	0.050	0.030	0.743
unemployed percentage	217,840	0.079	0.031	0.000	0.465
intermediate profession percentage	217,837	0.163	0.067	0.000	0.625
employee percentage	217,837	0.181	0.058	0.000	0.683
management and intellectual profession percentage	217,837	0.071	0.055	0.000	0.538
artisans business owners percentage	217,837	0.053	0.039	0.000	0.480
farmers percentage	217,837	0.038	0.055	0.000	0.609
worker percentage	217,837	0.171	0.074	0.000	0.697
sparsely populated autonomous rural area	217,868	0.251	0.434	0	1
very sparsely populated autonomous rural area	217,868	0.180	0.384	0	1
rural under weak influence of a pole	217,868	0.217	0.412	0	1
rural under strong influence of a pole	217,868	0.219	0.414	0	1
dense urban	217,868	0.025	0.155	0	1
urban intermediate density	217,868	0.108	0.310	0	1
weighting					
population scale	217,840	0.128	0.992	0.004	139.530

Table 2: correlations between pollutants

	NO2	PM2.5	PM10	O3
NO2	1			
PM2.5	0.569	1		
PM10	0.570	0.696	1	
O3	-0.094	-0.124	-0.220	1

Note: Data from the period 2012-2018 are used. Variables are number of days of pollution in the year in each municipality.

Table 3: Income and number of pollution days in rural and urban areas

	ALL	NO2	PM25	PM10	O3
	(1)	(2)	(3)	(4)	(5)
DENSE URBAN					
Income	-0.002^{***}	-0.002^{**}	-0.0005^{***}	-0.00003	-0.00000
	(0.001)	(0.001)	(0.0002)	(0.0001)	(0.00004)
Observations	5,362	5,362	5,362	5,362	5,362
R^2	0.792	0.794	0.815	0.822	0.651
URBAN INTERMEDIATE DENSITY					
Income	0.0002	0.001	-0.0002	-0.00003	0.00003
	(0.0005)	(0.001)	(0.0002)	(0.0001)	(0.0001)
Observations	23,471	23,471	23,471	23,471	23,471
R^2	0.790	0.683	0.809	0.707	0.614
RURAL UNDER STRONG INFLUENCE OF A POLE					
Income	0.001^{***}	0.002^{***}	0.001^{***}	0.0001^{***}	-0.00001
	(0.0001)	(0.0002)	(0.0001)	(0.00002)	(0.00002)
Observations	47,721	47,721	47,721	47,721	47,721
R^2	0.828	0.728	0.847	0.734	0.597
RURAL UNDER WEAK INFLUENCE OF A POLE					
Income	0.001^{***}	0.001^{***}	0.0003^{**}	0.0001^{***}	-0.0001
	(0.0002)	(0.0003)	(0.0001)	(0.00003)	(0.0001)
Observations	47,094	47,094	47,094	47,094	47,094
R^2	0.842	0.727	0.842	0.717	0.612
SPARSELY POPULATED AUTONOMOUS RURAL AREA					
Income	0.001^{***}	0.001^{***}	0.0005^{**}	0.0001^{***}	0.00002
	(0.0002)	(0.0003)	(0.0002)	(0.00003)	(0.0001)
Observations	54,681	54,681	54,681	54,681	54,681
R^2	0.832	0.710	0.828	0.712	0.588
VERY SPARSELY POPULATED AUTONOMOUS RURAL AREA					
Income	0.0004^{***}	0.0004^{***}	0.0003^{**}	0.0001^{***}	0.00002
	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.00005)
Observations	38,988	38,988	38,988	38,988	38,988
R^2	0.870	0.746	0.875	0.685	0.601
Département dummies	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes
controls	Yes	Yes	Yes	Yes	Yes

Note: Table reports OLS estimates of equation (1) in each of the sub-samples of urban and rural municipalities. The dependent variable is the number of days in the year for which the health threshold concentration (WHO) of the considered pollutant is exceeded. The dependent variable ALL is the number of days in the year when at least one of the pollutants exceeds its respective threshold. Robust standard errors are clustered at the local département level. *p<0.1; **p<0.05; ***p<0.01

Table 4: Income and share of cocktail days in rural and urban areas

	Share of cocktail days		
	DENSE URBAN	URBAN INTERMEDIATE DENSITY	RURAL
	(1)	(2)	(3)
Income	0.002	0.005	0.019^{***}
	(0.002)	(0.005)	(0.003)
Number of pollution days	0.0001	0.001^{***}	0.001^{***}
	(0.0002)	(0.0002)	(0.0002)
Département dummies	Yes	Yes	Yes
year dummies	Yes	Yes	Yes
controls	Yes	Yes	Yes
Observations	5,362	23,471	188,484
R^2	0.742	0.744	0.752

Note: Table reports OLS estimates in each of the sub-samples of urban and rural municipalities. The dependent variable is the the share of cocktail days (days with at least two pollutants above the health thresholds) among pollution days (at least one pollutant above its health threshold). I use Z-score on the income variable to facilitate the interpretation of the coefficients. Robust standard errors are clustered at the local departement level. *p<0.1; **p<0.05; ***p<0.01

Table 5: Income and share of mixture in total pollution days

	Share of mixture in total pollution days						
	PM2.5	NO2 PM2.5	PM2.5 PM10	PM2.5 O3	NO2 PM2.5 PM10	NO2	O3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DENSE URBAN							
Income	-0.005	0.004	-0.0005^{*}	-0.001	-0.001	0.003	-0.00005
	(0.004)	(0.003)	(0.0003)	(0.0005)	(0.001)	(0.004)	(0.001)
Observations	5,362	5,362	5,362	5,362	5,362	5,362	5,362
R^2	0.809	0.637	0.524	0.346	0.760	0.850	0.651
URBAN INTERMEDIATE							
Income	-0.011	0.005	0.0003	0.0001	-0.001	0.006	0.001
	(0.007)	(0.005)	(0.0002)	(0.001)	(0.001)	(0.005)	(0.002)
Observations	23,471	23,471	23,471	23,471	23,471	23,471	23,471
R^2	0.687	0.701	0.539	0.380	0.650	0.601	0.648
RURAL							
Income	-0.031^{***}	0.017^{***}	0.0002	-0.0001	0.002^{***}	0.011^{***}	0.0002
	(0.003)	(0.002)	(0.0003)	(0.0002)	(0.001)	(0.002)	(0.002)
Observations	188,484	188,484	188,484	188,484	188,484	188,484	188,484
R^2	0.664	0.706	0.573	0.371	0.658	0.505	0.635
Département dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of pollution days	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Table reports OLS estimates in each of the sub-samples of urban and rural municipalities. The dependent variable is the the share of the mixture (days when the pollutant(s) in the mixture are above the thresholds but no other pollutant exceeds its threshold) among pollution days (at least one pollutant above its health threshold). I use Z-score on the income variable to facilitate the interpretation of the coefficients. Robust standard errors are clustered at the local departement level. *p<0.1; **p<0.05; ***p<0.01

Table 6: Robustness checks : economic indicators in dense urban area (number of pollution days)

	DENSE URBAN				
	NO2	PM25	PM10	O3	ALL
	(1)	(2)	(3)	(4)	(5)
poverty rate	1.375***	0.231*	0.049	−0.020	1.227***
	(0.453)	(0.135)	(0.039)	(0.016)	(0.357)
Observations	5,129	5,129	5,129	5,129	5,129
R^2	0.799	0.813	0.821	0.648	0.797
first decile	−0.006***	−0.001***	−0.0001	0.00004	−0.005***
	(0.002)	(0.0003)	(0.0001)	(0.00005)	(0.002)
Observations	5,267	5,267	5,267	5,267	5,267
R^2	0.798	0.814	0.821	0.649	0.797
ninth decile	−0.0003	−0.0001	−0.00002	−0.00000	−0.0003
	(0.0002)	(0.0001)	(0.00002)	(0.00002)	(0.0002)
Observations	5,267	5,267	5,267	5,267	5,267
R^2	0.794	0.814	0.821	0.649	0.793
share of social benefits	−1.117	0.172	0.077	0.007	−0.566
	(0.897)	(0.331)	(0.105)	(0.049)	(0.620)
Observations	5,267	5,267	5,267	5,267	5,267
R^2	0.794	0.813	0.821	0.649	0.792
share of income from assets	−0.751**	−0.368**	0.036	0.022	−0.679***
	(0.348)	(0.163)	(0.107)	(0.021)	(0.191)
Observations	5,267	5,267	5,267	5,267	5,267
R^2	0.795	0.814	0.821	0.649	0.793
Département dummies	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes
controls	Yes	Yes	Yes	Yes	Yes

Note: Table reports OLS estimates of equation (1) in the dense urban sub-sample. The dependent variable is the number of days in the year for which the health threshold concentration (WHO) of the considered pollutant is exceeded. The dependent variable ALL is the number of days in the year when at least one of the pollutants exceeds its respective threshold. Robust standard errors are clustered at the local departement level. *p<0.1; **p<0.05; ***p<0.01

Table 7: Robustness checks : economic indicators in urban intermediate density area (number of pollution days)

	URBAN INTERMEDIATE DENSITY				
	NO2	PM25	PM10	O3	ALL
	(1)	(2)	(3)	(4)	(5)
poverty rate	0.055	-0.103	0.016	-0.061^{**}	-0.072
	(0.212)	(0.145)	(0.036)	(0.029)	(0.174)
Observations	13,175	13,175	13,175	13,175	13,175
R^2	0.711	0.813	0.715	0.599	0.808
first decile	-0.0001	0.0001	-0.0001	0.0001	0.00003
	(0.001)	(0.0004)	(0.0001)	(0.0001)	(0.001)
Observations	15,490	15,490	15,490	15,490	15,490
R^2	0.712	0.812	0.715	0.600	0.803
ninth decile	0.0001	-0.0002^{**}	-0.0001^{***}	0.00003	0.00001
	(0.0003)	(0.0001)	(0.00002)	(0.00003)	(0.0002)
Observations	15,490	15,490	15,490	15,490	15,490
R^2	0.712	0.813	0.716	0.600	0.803
share of social benefits	-0.124	0.777^{*}	0.160	-0.189^{**}	0.115
	(0.596)	(0.400)	(0.119)	(0.090)	(0.512)
Observations	15,490	15,490	15,490	15,490	15,490
R^2	0.712	0.813	0.716	0.600	0.803
share of income from assets	0.151	-0.109	0.002	-0.023^{**}	0.035
	(0.096)	(0.072)	(0.031)	(0.012)	(0.077)
Observations	15,490	15,490	15,490	15,490	15,490
R^2	0.712	0.812	0.715	0.600	0.804
Département dummies	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes
controls	Yes	Yes	Yes	Yes	Yes

Note: Table reports OLS estimates of equation (1) in the urban intermediate density sub-sample. The dependent variable is the number of days in the year for which the health threshold concentration (WHO) of the considered pollutant is exceeded. The dependent variable ALL is the number of days in the year when at least one of the pollutants exceeds its respective threshold. Robust standard errors are clustered at the local departement level. *p<0.1; **p<0.05; ***p<0.01

Table 8: Robustness checks : economic indicators in urban intermediate density area (share of cocktail days)

	Share of cocktail days	
	DENSE URBAN	URBAN INTERMEDIATE DENSITY
	(1)	(2)
poverty rate	−0.0004	−0.0001
	(0.001)	(0.001)
Observations	5,129	13,175
R^2	0.741	0.754
first decile	0.00000	0.00000
	(0.00000)	(0.00000)
Observations	5,267	15,490
R^2	0.741	0.761
ninth decile	0.00000	0.00000
	(0.00000)	(0.00000)
Observations	5,267	15,490
R^2	0.741	0.761
share of social benefits	0.001	0.0003
	(0.001)	(0.002)
Observations	5,267	15,490
R^2	0.741	0.761
share of income from assets	−0.0002	0.0001
	(0.001)	(0.001)
Observations	5,267	15,490
R^2	0.741	0.761
Département dummies	Yes	Yes
year dummies	Yes	Yes
controls	Yes	Yes
Number of pollution days	Yes	Yes

Note: Table reports OLS estimates in each of the sub-samples of dense urban and urban intermediate density municipalities. The dependent variable is the the share of cocktail days (days with at least two pollutants above the health thresholds) among pollution days (at least one pollutant above its health threshold). Robust standard errors are clustered at the local departement level. *p<0.1; **p<0.05; ***p<0.01

Table 9: Tests of spatial auto-correlation (p-value)

	Test Moran	Test LM-Error	Test LM-Lag	Test Robust LM-Error	Test Robust LM-Lag
NO2					
National (2018)	0.000	0.000	1	0.000	1
Very sparsely populated autonomous rural area		0.000	0.000	0.000	0.003
Rural under weak influence of a pole		0.000	0.000	0.000	0.1279
Rural under strong influence of a pole		0.000	0.000	0.000	0.663
Dense urban		0.000	0.002	0.000	0.7633
Urban intermediate density		0.000	0.000	0.000	0.2936
PM25					
National (2018)	0.000	0.000	1	0.000	1
Very sparsely populated autonomous rural area		0.000	0.000	0.000	0.160
Rural under weak influence of a pole		0.000	0.000	0.000	0.9745
Rural under strong influence of a pole		0.000	0.000	0.000	0.572
Dense urban		0.000	0.000	0.000	0.7394
Urban intermediate density		0.000	0.000	0.000	0.0349
PM10					
National (2018)	0.000	0.000	1	0.000	1
Very sparsely populated autonomous rural area		0.000	0.000	0.000	0.546
Rural under weak influence of a pole		0.000	0.000	0.000	0.4273
Rural under strong influence of a pole		0.000	0.000	0.000	0.8724
Dense urban		0.000	0.000	0.000	0.3276
Urban intermediate density		0.000	0.000	0.000	0.3208
O3					
National (2018)	0.000	0.000	1	0.000	1
Very sparsely populated autonomous rural area		0.000	0.000	0.000	0.011
Rural under weak influence of a pole		0.000	0.000	0.000	0.002
Rural under strong influence of a pole		0.000	0.000	0.000	0.004
Dense urban		0.000	0.000	0.000	0.07969
Urban intermediate density		0.000	0.000	0.000	0.4042
Share of cocktail					
National (2018)	0.000	0.000	1	0.000	1
Very sparsely populated autonomous rural area		0.000	0.000	0.000	0.3973
Rural under weak influence of a pole		0.000	0.000	0.000	0.000
Rural under strong influence of a pole		0.000	0.000	0.000	0.0005
Dense urban		0.000	0.000	0.000	0.5108
Urban intermediate density		0.000	0.000	0.000	0.0003

Note: Each coefficient corresponds to the p-value of the associated test performed on OLS error terms. To perform these tests we use the spatial weighting matrix of the k nearest neighbors, using k=5. Regressions at the national level are weighted by the population of the municipalities.

Table 10: Robustness checks : spatial regression models (SEM)

	NO2	PM2.5	PM10	O3	Share of cocktail
	(1)	(2)	(3)	(4)	(5)
DENSE URBAN					
Income	-0.002^{***}	-0.0005^{***}	-0.0001	-0.00000	0.002
	(0.0003)	(0.0001)	(0.00004)	(0.00003)	(0.0019)
ρ	0.126^{***}	0.422^{***}	0.504^{***}	0.437^{***}	0.364^{***}
	(0.021)	(0.016)	(0.015)	(0.016)	(0.017)
Observations	5,362	5,362	5,362	5,362	5,362
URBAN INTERMEDIATE DENSITY					
Income	0.001^{***}	-0.0001^{**}	-0.00003^{**}	0.00002	0.005^{***}
	(0.0001)	(0.0001)	(0.00002)	(0.00001)	(0.0009)
ρ	0.127^{***}	0.405^{***}	0.490^{***}	0.406^{***}	0.301^{***}
	(0.0102)	(0.0079)	(0.007)	(0.008)	(0.009)
Observations	23,471	23,471	23,471	23,471	23,471
RURAL UNDER STRONG INFLUENCE OF A POLE					
Income	0.002^{***}	0.001^{***}	0.0001^{***}	-0.00001	0.020^{***}
	(0.00004)	(0.00003)	(0.00001)	(0.00001)	(0.0006)
ρ	0.283^{***}	0.536^{***}	0.581^{***}	0.506^{***}	0.436^{***}
	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)
Observations	47,509	47,509	47,509	47,509	47,509
RURAL UNDER WEAK INFLUENCE OF A POLE					
Income	0.001^{***}	0.0003^{***}	0.0001^{***}	-0.0001^{***}	0.018^{***}
	(0.00003)	(0.00003)	(0.00001)	(0.00001)	(0.00057)
ρ	0.270^{***}	0.487^{***}	0.535^{***}	0.476^{***}	0.394^{***}
	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)
Observations	46,879	46,879	46,879	46,879	46,879
VERY SPARSELY POPULATED RURAL AREA					
Income	0.0004^{***}	0.0003^{***}	0.0001^{***}	0.00002^{*}	0.010^{***}
	(0.00002)	(0.00003)	(0.00001)	(0.00001)	(0.00057)
ρ	0.399^{***}	0.461^{***}	0.471^{***}	0.401^{***}	0.411^{***}
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	38,318	38,318	38,318	38,318	38,318
Département dummies	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes
controls	Yes	Yes	Yes	Yes	Yes
Number of pollution days	No	No	No	No	Yes

Note: Table reports maximum likelihood estimates of the Spatial Error Model (SEM) described in the empirical strategy section and in the appendix section, in urban and rural sub-samples. The dependent variables are the number of days in the year for which the health threshold concentration (WHO) of the considered pollutant is exceeded (columns 1 to 4) and the share of cocktail days (days with at least two pollutants above the health thresholds) among pollution days (at least one pollutant above its health threshold). ρ is the spatial autocorrelation parameter in the error term as described in the spatial aurocorrelation section in the appendix. I use the k nearest neighbors spatial weight matrix, using k=5. We use the Z-score on the income variable for the cocktail share regressions for ease of interpretation of the coefficients. *p<0.1; **p<0.05; ***p<0.01

Figures

Figure 1: Average number of days in the year above health thresholds in each area (2012-2018 period)

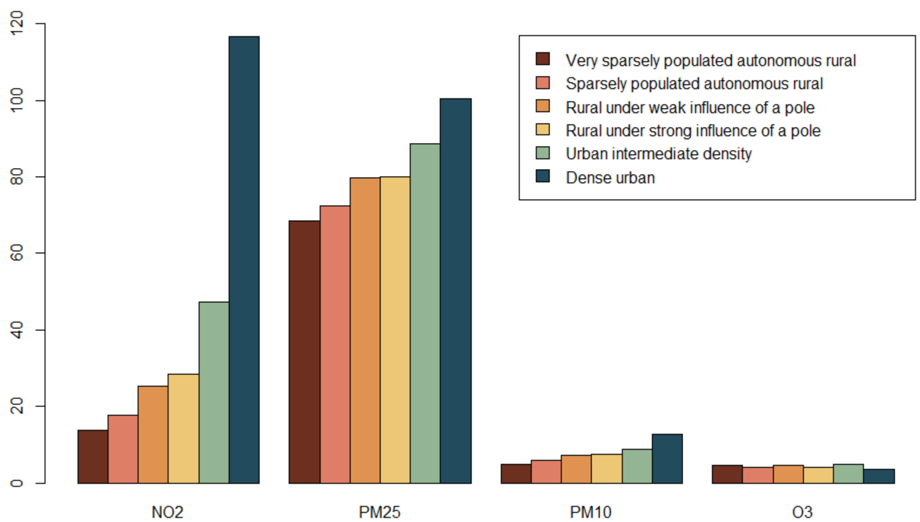


Figure 2: Share of each mixture in the total number of pollution days (all pollutants combined)

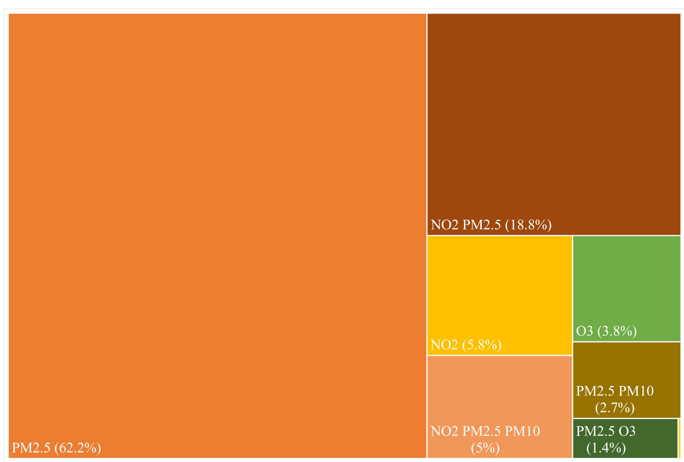
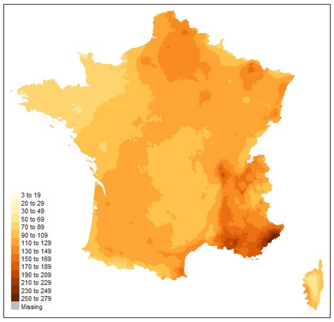
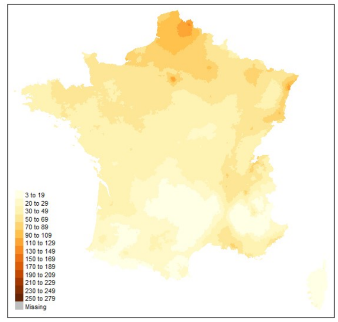


Figure 3: Number of pollution days in 2012 and 2018

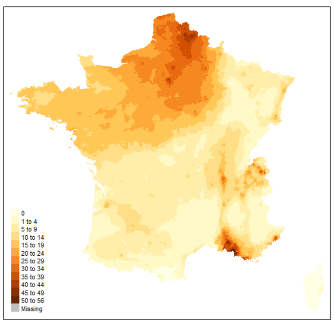
2012 **2018**



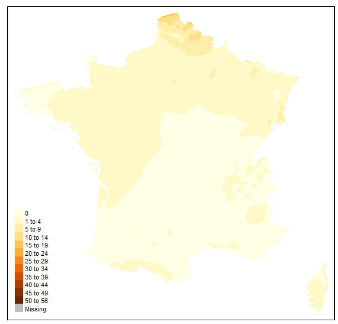
(a) PM2.5



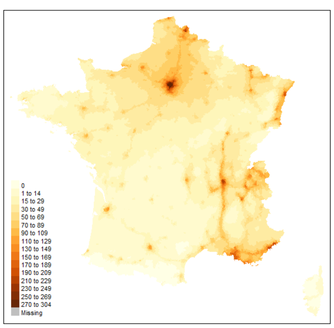
(b) PM2.5



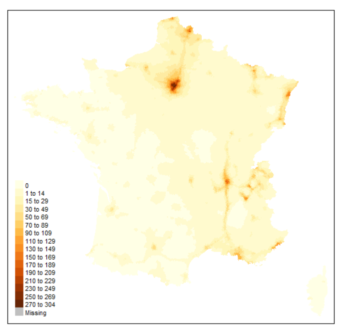
(c) PM10



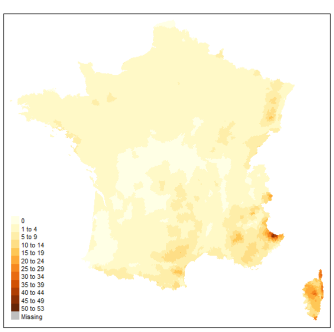
(d) PM10



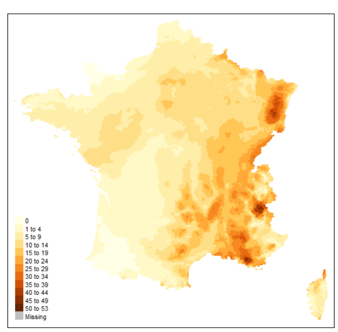
(e) NO2



(f) NO2

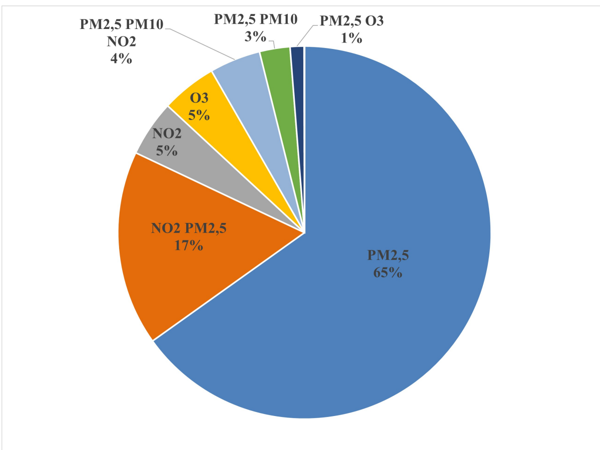


(g) O3

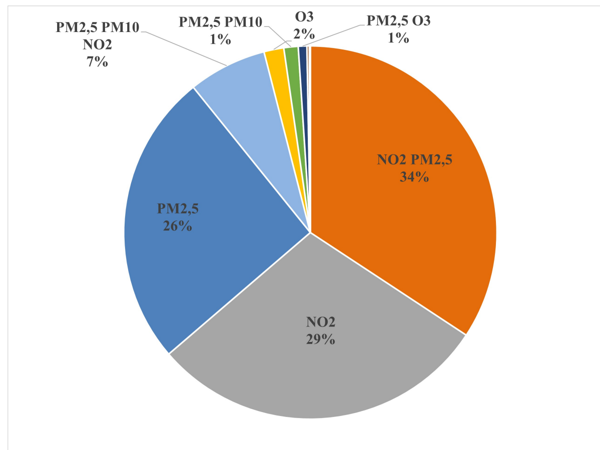


(h) O3

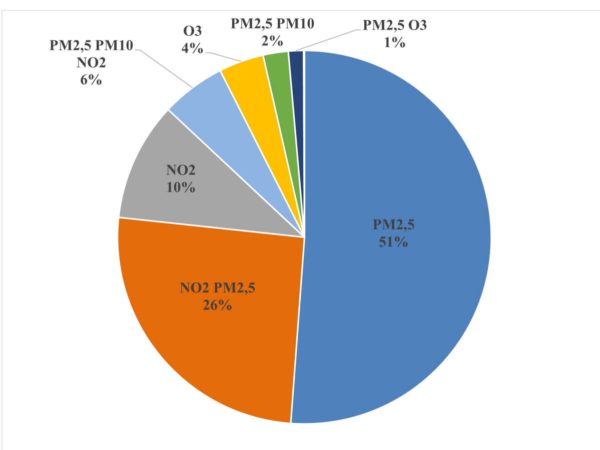
Figure 4: Average share of each pollutant mixture in the total number of pollution days (2012-2018)



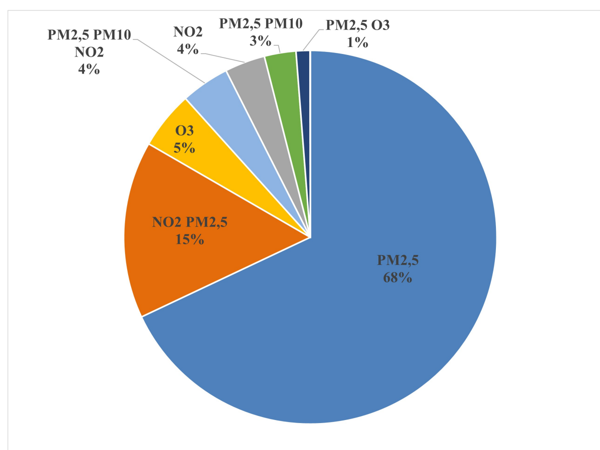
(a) National



(b) Dense urban



(c) Urban intermediate density



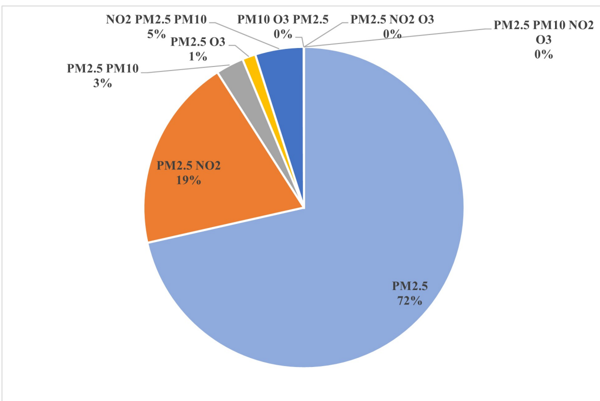
(d) Rural

Appendix

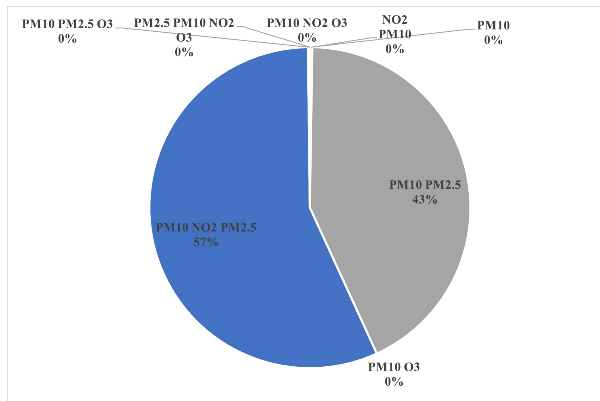
A Composition of each pollutant

Figure A.1 shows for each pollutant the average share of each mixture in the total number of days above the threshold for that pollutant. According to these estimates, on 29% of the days on which the PM2.5 threshold is exceeded, the threshold for at least one other pollutant is also exceeded (Figure A.1). Thus, studies assessing the effect of PM2.5 concentration on health actually capture the effect of pollutant cocktails and not the pure effect of PM2.5. Similarly, studies assessing exposure inequalities do not differentiate between exposure to pure PM2.5 particles and cocktails. In the case of PM10 particles, in 100% of the cases, the days above the threshold for PM10 particles are mixed with other pollutants (the PM10 PM2.5 and NO2 mixture represents 57% of PM10 pollution days on average, the PM10 PM2.5 mixture represents 43% of days on average). The PM2.5 NO2 mixture represents 67% of the nitrogen dioxide (NO2) pollution days on average, while the pollution days for this mixture represent only 19% of the PM2.5 pollution days (Figure A.1). Ozone (O3) is pure in 64% of cases on average. Figure A.2 shows the composition of pollution days in three types of areas. Overall, the composition of pollution days changes between urban and rural areas. In the case of fine particles and ozone, the share of cocktail days is higher in urban than in rural areas.

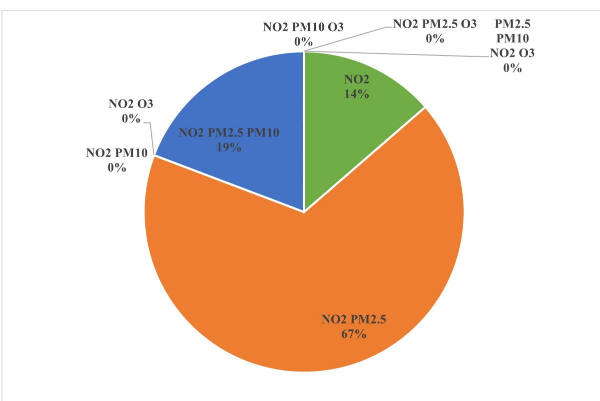
Figure A.1: Average share of each pollutant mixture in the number of pollution days for the considered pollutant (national level, period 2012-2018)



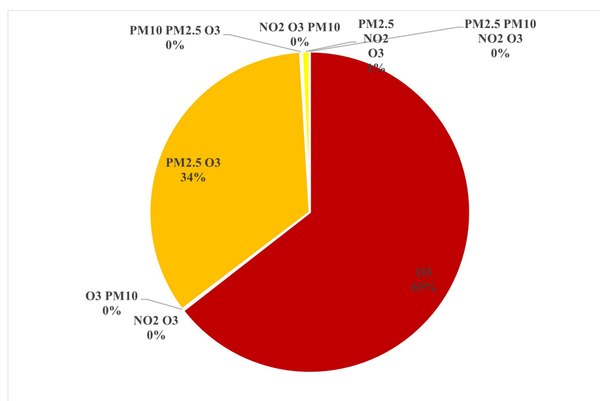
(a) PM2.5



(b) PM10

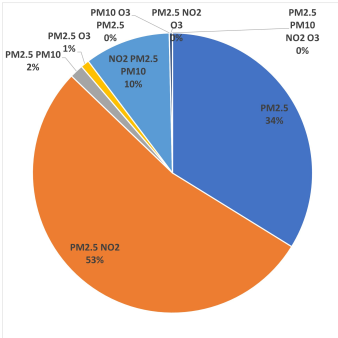


(c) NO2

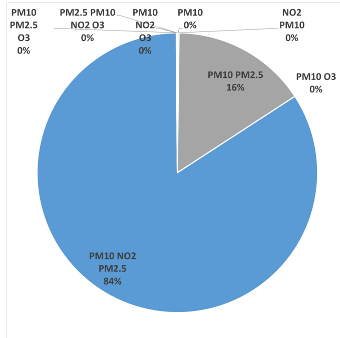


(d) O3

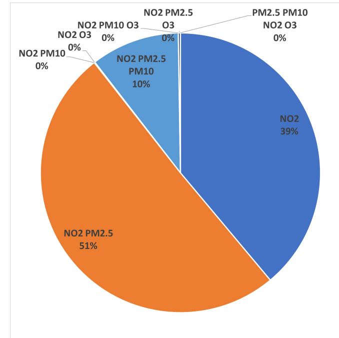
Figure A.2: Average share of each mixture in the number of pollution days for each pollutant in dense urban areas (first row), intermediate dense urban areas (second row) and rural areas (third row)



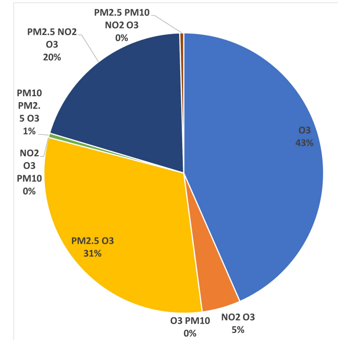
(a) PM2.5



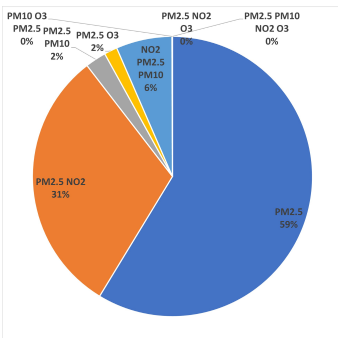
(d) PM10



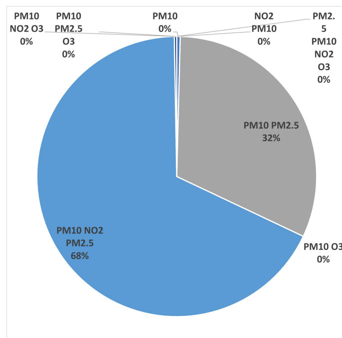
(g) NO2



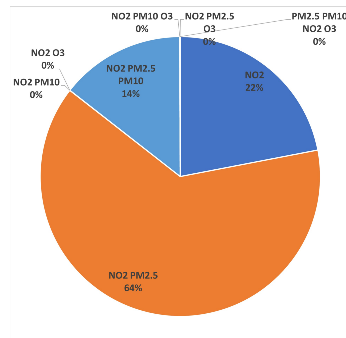
(j) O3



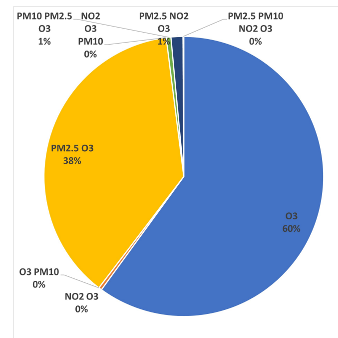
(b) PM2.5



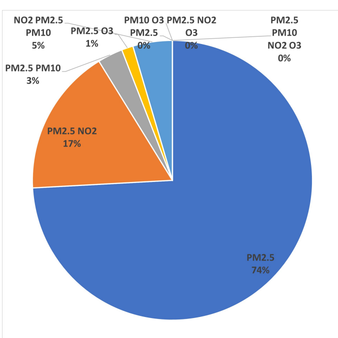
(e) PM10



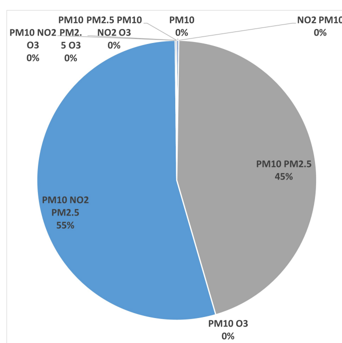
(h) NO2



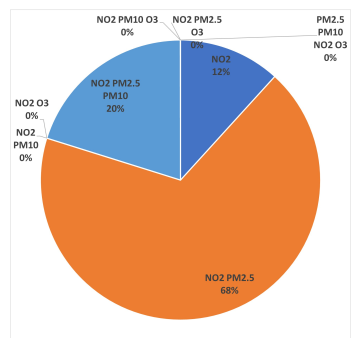
(k) O3



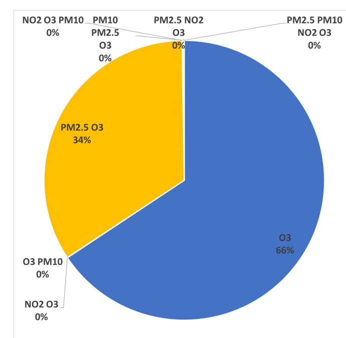
(c) PM2.5



(f) PM10



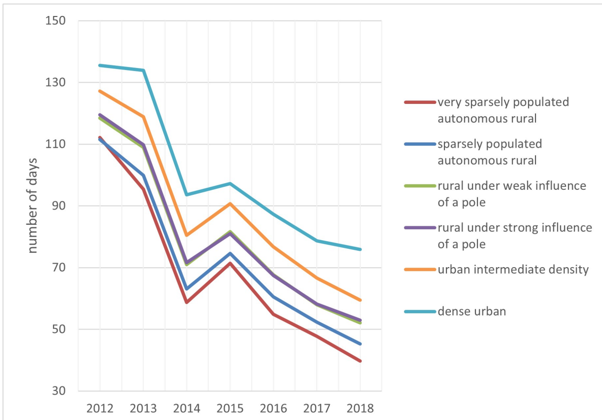
(i) NO2



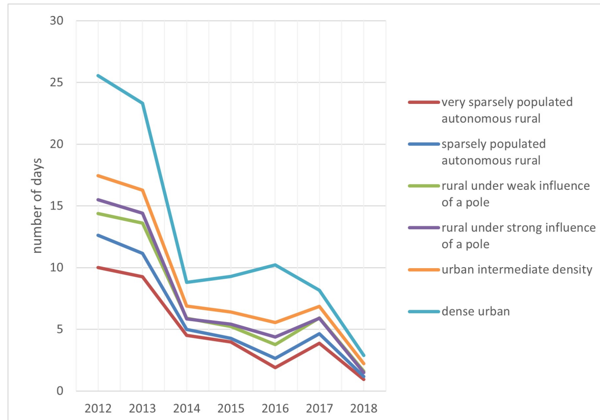
(l) O3

B Figures

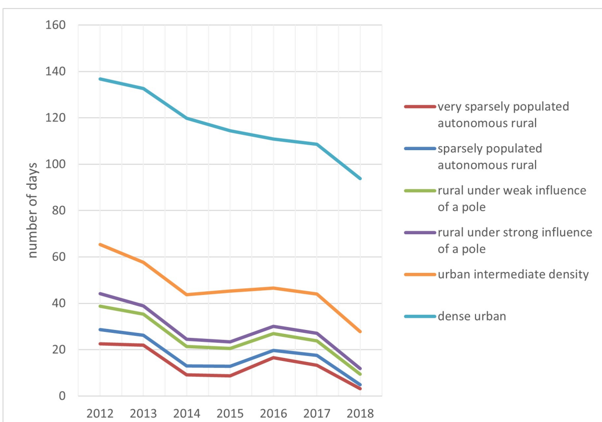
Figure A.3: Evolution of the average number of pollution days in urban and rural areas



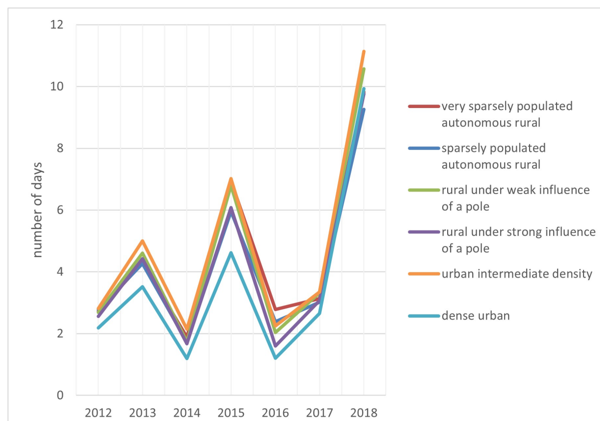
(a) PM2.5



(b) PM10



(c) NO2



(d) O3

C Estimation at the national level

We estimate a linear regression model with fixed effects at the national level. France has a large number of small municipalities with very small populations, which gives them more weight in the analysis. We weight the regressions at the national level according to the population in each municipality at each period. This weighting allows us to identify a link that is more representative of the situation of the largest number of inhabitants. We cluster the standard errors at the department level.

The results are presented in Table A1. The estimates suggest a significant negative relationship between municipality income and the exposure to PM2.5 and NO2 pollution. Ozone (O3), on the other hand, is concentrated in affluent and rural communes although the effect is not significant. Income is negatively associated with the level of PM10 pollution although the coefficient is not significant.

A one standard deviation increase in income in the municipality is associated with a reduction of 3 days above the thresholds in the year for nitrogen dioxide (NO2) and of 1.2 days for PM2.5 particles on average. The coefficients for PM10 and ozone (O3) are lower. We slightly underestimate the number of days above the threshold for ozone because the daily average ozone concentration is not measured as a rolling 8-hour average as measured by the WHO but as a 24-hour average, which may slightly influence the estimated coefficient for ozone (O3). For ozone (O3), a one standard deviation increase in income is associated with a 0.3 increase in the number of days above the threshold, and a 0.04 decrease in the number of days for PM10.

The coefficients of the dummy years indicate for nitrogen dioxide (NO2), PM2.5 and PM10 a continuous and significant decrease in pollution levels between 2012 and 2018. For ozone (O3) on the other hand, the overall levels of ozone (O3) pollution above the threshold seem to evolve randomly over the period.

Our national results reflect relationships in urban areas. Indeed, since a large proportion of the population lives in urban municipalities, the population weighting of municipalities in the national regressions gives more weight to this type of municipality.

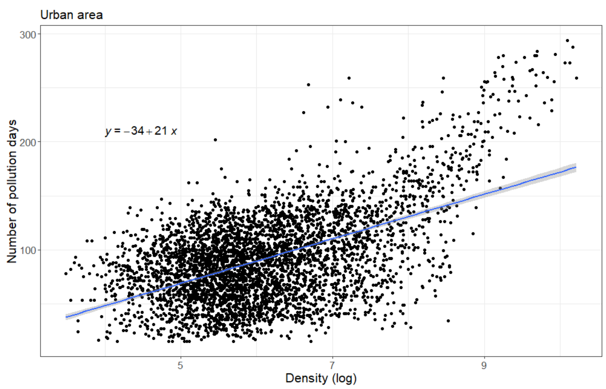
Table A1: Income and number of pollution days at national level

	NO2	PM25	PM10	O3	ALL
	(1)	(2)	(3)	(4)	(5)
Income	-0.001^{**}	-0.0004^{**}	-0.00001	0.0001	-0.001^{***}
	(0.0005)	(0.0001)	(0.00004)	(0.00005)	(0.0004)
2012	REF	REF	REF	REF	REF
2013	-4.345^{**}	-8.380^{**}	-2.440^{**}	1.814^{***}	-4.961^{*}
	(1.903)	(3.318)	(0.979)	(0.338)	(2.874)
2014	-17.101^{***}	-46.864^{***}	-13.339^{***}	-0.657^{***}	-35.897^{***}
	(2.273)	(3.542)	(1.313)	(0.242)	(3.419)
2015	-17.788^{***}	-38.229^{***}	-13.259^{***}	2.794^{***}	-31.055^{***}
	(1.912)	(2.030)	(1.421)	(0.570)	(1.545)
2016	-17.754^{***}	-51.161^{***}	-13.790^{***}	-0.495^{*}	-43.960^{***}
	(1.883)	(3.332)	(1.207)	(0.294)	(2.783)
2017	-19.347^{***}	-58.698^{***}	-13.793^{***}	0.189	-51.386^{***}
	(1.882)	(2.617)	(1.469)	(0.291)	(2.306)
2018	-34.000^{***}	-65.145^{***}	-18.219^{***}	7.341^{***}	-50.663^{***}
	(2.217)	(4.810)	(1.410)	(0.656)	(4.144)
sparsely populated autonomous rural	REF	REF	REF	REF	REF
very sparsely populated autonomous rural	0.221	-0.398	-0.018	-0.022	-0.480
	(1.443)	(0.509)	(0.145)	(0.121)	(1.170)
rural under weak influence of a pole	-1.012	0.590	0.174^{*}	-0.025	-0.319
	(1.035)	(0.468)	(0.097)	(0.103)	(0.871)
rural under strong influence of a pole	1.687	1.900^{***}	0.529^{***}	-0.239^{**}	1.471
	(1.436)	(0.608)	(0.138)	(0.121)	(1.136)
dense urban	39.723^{***}	10.535^{***}	2.564^{***}	-1.100^{***}	26.686^{***}
	(3.929)	(1.238)	(0.360)	(0.350)	(3.056)
urban intermediate density	7.459^{***}	3.807^{***}	0.871^{***}	-0.249^{*}	5.162^{***}
	(1.469)	(0.576)	(0.156)	(0.142)	(1.030)
Departement dummies	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	217,317	217,317	217,317	217,317	217,317
R^2	0.907	0.825	0.741	0.588	0.900

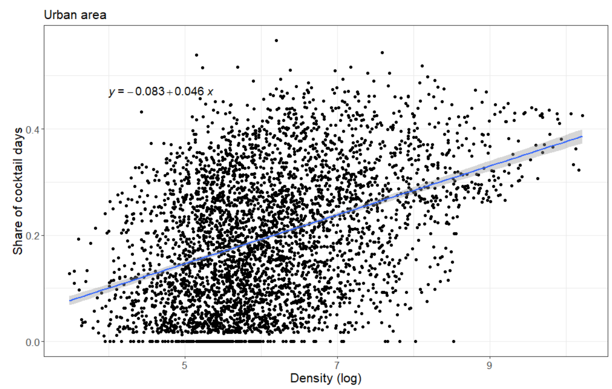
Note: Robust standard errors are clustered at the local departement level. All regressions are weighted by the population of the municipalities. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

D Density, income, economic activity and pollution

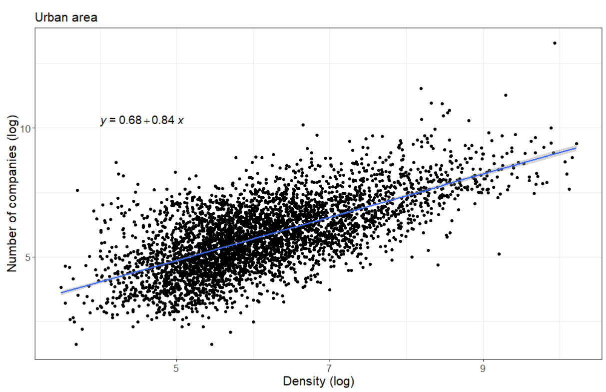
Figure A.4: correlations between municipal density, income and pollution (ALL) in urban area (2018)



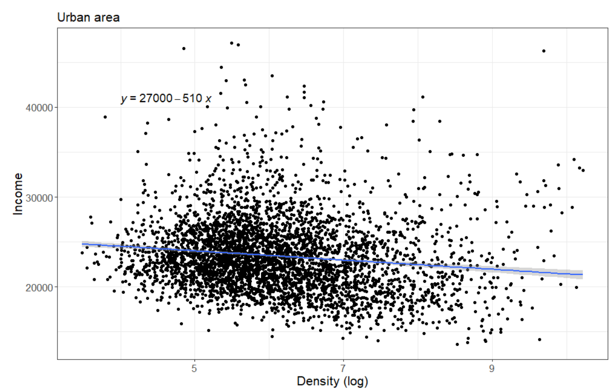
(a)Pollution days



(b) Share of cocktail days



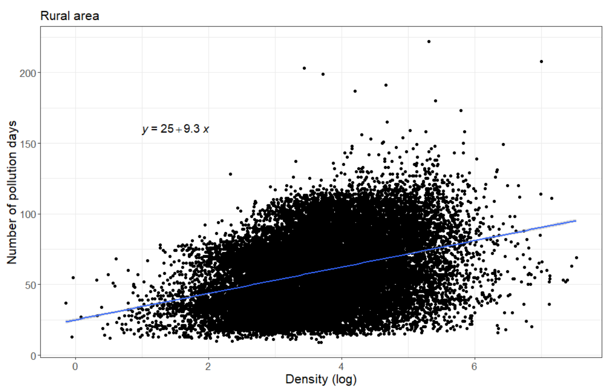
(c) Number of companies



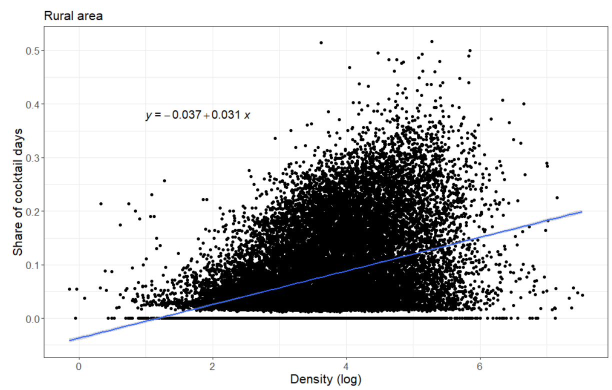
(d) Income

Note : Density is the number of inhabitants per km^2 calculated by the "Observatoire des territoires" from the INSEE RP 2018 data. The data on the number of companies comes from the INSEE's "Démographie des entreprises et des établissements" database. We use the variable that counts the number of establishments in the municipality in 2020.

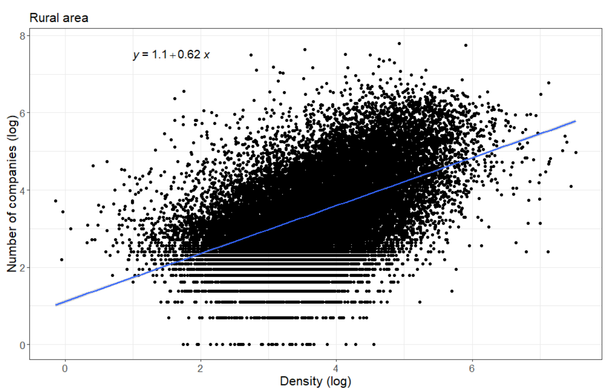
Figure A.5: correlations between municipal density, income and pollution (ALL) in rural area (2018)



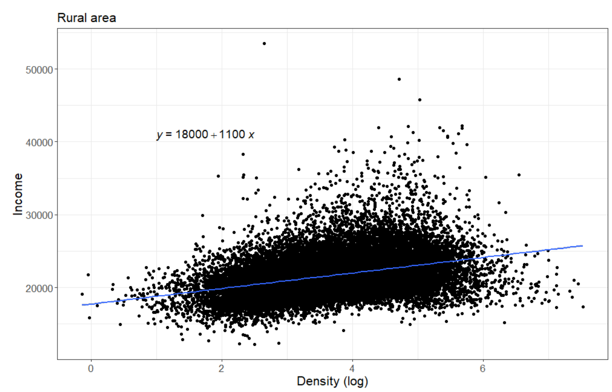
(a)Pollution days



(b) Share of cocktail days



(c) Number of companies



(d) Income

Note : Density is the number of inhabitants per km^2 calculated by the "Observatoire des territoires" from the INSEE RP 2018 data. The data on the number of companies comes from the INSEE's "Démographie des entreprises et des établissements" database. We use the variable that counts the number of establishments in the municipality in 2020.

E Spatial autocorrelation

In order to test the robustness of our results, we estimate a spatial error model to explicitly take into account the spatial autocorrelation problem. Following Bouayad Agha, Le Gallo, and Vedrine (2018) the fixed effects SEM with two dimensional data can be written as :

$$y_{ct} = \beta x_{ct} + \lambda_t + \varepsilon_{ct} \tag{2}$$

$$\varepsilon_{ct} = \rho \sum_{c \neq j}^{N} w_{cj} \varepsilon_{jt} + \varphi_{ct} \tag{3}$$

With y_{ct}, the number of pollution days in the year of municipality i in year t, with c = 1,2,...,N and t = 1,2,...,T, λ_c an time fixed effect and ε_{ct} the error term. ρ is the spatial error autocorrelation parameter and $\sum_{j=1}^{N} w_{cj}\varepsilon_{jt}$ represents the spatially displaced error term. w_{cj} is a spatial weighting matrix of dimension (N,N) that describes the spatial relationship between the municipalities, with c and j two different municipalities. The spatial relationships described by the spatial weighting matrix W can take several forms. We retain the k nearest neighbors matrix by retaining k=5. For each municipality, the 5 closest municipalities are identified geographically. The 5 closest municipalities have a value (weight) of 0.2 with the commune considered in the weighting matrix W.

Following Bouayad Agha, Le Gallo, and Vedrine (2018), we extend this model to data in three dimensions:

$$y_{cdt} = \beta x_{cdt} + \alpha_d + \lambda_t + \varepsilon_{cdt} \tag{4}$$

$$\varepsilon_{cdt} = \rho \sum_{g=1}^{M} \sum_{h=1}^{N_g} w_{cd,hg} \varepsilon_{hgt} + \varphi_{cdt} \tag{5}$$

With y_{cdt}, the number of pollution days in the year of municipality c of département d in year t, with c = 1,2,...,N, d = 1,2,...,M and t = 1,2,...,T. ρ is the spatial parameter to be estimated and $\sum_{g=1}^{M}\sum_{h=1}^{N_g} w_{cd,hg}\varepsilon_{hgt}$ the spatially displaced error term. $w_{cd,hg}$ are the elements of the spatial weighting matrix of dimension (N,N), with cd representing municipality c in département d and hg a municipality h in département g.

CEE-M Working Papers[1] - 2022

WP 2022-01	Denis Claude & **Mabel tidball** « Taking firms' margin targets seriously in a model of competition in supply functions »
WP 2022-02	Cauê D. Carrilhoa, **Gabriela Demarchi**, Amy E. Duchelle, Sven Wunder, & Marla Morsello « Permanence of avoided deforestation in a Transamazon REDD+ initiative (Pará, Brazil)functions »
WP 2022-03	Francisco Cabo, Alain Jean-marie & **Mabel tidball** « Positional effects in public good provision. Strategic interaction and inertia »
WP 2022-04	**Fabien Prieur**, Weihua Ruan & Benteng Zou « Optimal lockdown and vaccination policies to contain the spread of a mutating infectious disease »
WP 2022-05	**Claudia Kelsall**, Martin F. Quaas & **Nicolas Quérou** « Risk aversion in renewable resource harvesting »
WP 2022-06	**Raphaël Soubeyran, Nicolas Quérou** & Mamadou Gueye « Social Preferences and the Distribution of Rewards »
WP 2022-07	**Antoine Pietri** « Plus de voiture pour davantage de vélo ? Une étude sur la ville de Tours »
WP 2022-08	**Sebastien Desbureaux**, Frederic Mortier, Esha Zaveri, Michelle van Vliet, Jason Russ, Aude Sophie Rodella & Richard Damania « Mapping global hotspots and trends of water quality (1992-2010): a data driven approach »
WP 2022-09	**Gabriela Demarchi**, Caue D. Carrilho, Thibault Catry, Stibniati Atmadja & **Julie Subervie** « Beyond reducing deforestation: impacts of conservation programs on household livelihoods »
WP 2022-10	**Robbert-Jan Schaap**, Exequiel Gonzalez-Poblete, Karin Loreto Silva Aedo, & Florian Diekert « Beyond reducing deforestation: impacts of conservation programs on household livelihoods »

[1] **CEE-M Working Papers / Contact** : laurent.garnier@inrae.fr

- RePEc https://ideas.repec.org/s/hal/wpceem.html
- HAL https://halshs.archives-ouvertes.fr/CEE-M-WP/

WP 2022-11	**Camille Salesse** « Inequality in exposure to air pollution in France: bringing pollutant cocktails into the picture»