An Optimal Distribution of Polluting Activities Across Space*

Martin Jégard†

October 2021

Preliminary and incomplete
(most recent version here)

Abstract

Should air quality policies target industries within the largest cities? On the one hand, we should seek to reduce atmospheric pollutants' emissions in places where most of the population is concentrated. On the other hand, more stringent policies can hurt local industries and targeting the cities that contribute the most economically may decrease welfare. Extending recent quantitative spatial economics models, I analyze these counteracting forces. I find that when the local damages from pollution are not internalized by the industry and workers react to low air quality through migration, the largest cities can be too small. As a result, an optimal set of policies imposes higher emission taxes in these locations relative to the rest of the country. I estimate the model using French data and find that current policies impose higher costs of emissions in larger cities but raising them even higher could achieve welfare gains.

JEL codes: R12, R13, R23, Q53, Q58, L50
Keywords: air quality, industrial pollution, spatial equilibrium, regulation

*I am grateful to Hélène Ollivier for her support and helpful comments. I also thank Geoffrey Barrows, Ariell Reshef, Katheline Schubert and Cécile Gaubert for their kind help and suggestions. I acknowledge financial support from the EUR PGSE grant ANR-17-EURE-0001 and the Université Paris 1 Panthéon-Sorbonne economics doctoral school (ED 465).

†Paris School of Economics , Address: Paris School of Economics, 48 Boulevard Jourdan 74014 Paris, France. Email: martin.jegard@psemail.eu.

1 Introduction

It is now widely acknowledged that atmospheric pollution causes substantial damage to human health.[1] Yet, most people live in large cities that also concentrate polluting industries. Recognizing this, governments have adopted air quality regulations that that are more stringent in large and populated cities. However, more stringent regulations can hurt local firms, make them less productive, and, in turn, affect local workers' income. This local tradeoff between income and pollution raises the question of whether environmental regulations should be different across heterogeneous locations. This question requires a theory of the location choice of workers and heterogeneous firms that endogenizes local air pollution as an externality from production. In this paper, I develop such a theory and explore what would be an optimal distribution of polluting activities across space in the context of France.

I first use this framework to find the set of policies that implements an optimal allocation of workers, firms, and pollution across cities. Within this framework, location choices made by individual firms and workers depend on (1) the level of amenities offered by cities, (2) agglomeration externalities that provide a productivity advantage to firms located in more populated areas, and (3) congestion forces. Local concentration of atmospheric pollutants act as a congestion force by decreasing the city-specific amenity level. Workers respond to bad air quality by moving away from it. Firms, by contrast, only internalize the negative externalities associated with their polluting emissions insofar as they have to pay a local city-specific tax on emissions. They do not internalize their impact on local labor supply that responds to the decrease in amenity level associated with bad air quality. I thus show that the largest cities, endowed with the highest level of exogenous amenities and productivities, are too small when the emission tax is uniform across space. An optimal set of city-specific pollution taxes should include higher taxes in large cities to make industrial production cleaner, air quality better, and these cities more attractive to workers.[2] Second, I explore quantitatively what would be this optimal set of city-specific pollution taxes in the context of France and compare it to the estimated current regulations on industrial emissions. I find that current regulations are differentiated across space in the right direction – they are more stringent in larger cities – but welfare could increase by raising pollution taxes in larger cities even more.

In the model, a discrete number of cities in a single country differ in terms of idiosyncratic exogenous endowments in local amenities and industrial labor productivity. Workers

[1] Health damage have been measured even when pollution levels fall below regulatory standards (Graff Zivin & Neidell, 2012, 2013; Deryugina et al., 2019).

[2] To be able to compare sets of city-specific pollution taxes, I keep the average tax level across cities constant across alternative sets. In doing so, I focus only on the effects of spatially differentiated emission taxes, not on the welfare impacts of raising tax levels.

derive utility from local amenities and from the consumption of an industrial tradable composite good. Following Rosen (1979)-Roback (1982), I assume that a given population of homogenous industrial workers choose to locate in different cities in equilibrium. In each city, continuums of heterogenous firms from distinct industrial sectors produce differentiated varieties that aggregate into the composite good. As in Copeland & Taylor (2004), firms use labor and polluting emissions as substitutable inputs for production.[3] Production costs thus depend on local wages, local labor productivity, and local emission taxes. Furthermore, firms face iceberg trade costs when exporting their varieties to other cities. As is standard in economic geography models, the model includes two externalities related to the allocation of workers across space. First, local labor productivity is endogenous and reflects agglomeration economies: when the number of worker increases in a city, so does labor productivity. Second, local welfare is influenced by general congestion effects: as local population increases, congestion appears on the housing market and in transportation, which reduces each worker's welfare.[4] Furthermore, workers are assumed to move freely across cities.[5]

The novelty of the paper rests on the assumption that the total quantity of industrial emissions per city negatively affects city-specific amenity level, using a constant elasticity environmental damage function. Doing so, I consider polluting emissions as a congestion force that affects workers' location decision, and firms do not internalize this effect. In the model, I consider for simplicity a representative pollutant that serves as an indicator of the combination of harmful pollutants emitted by industrial activities (PM10, PM2.5 and SO2). I further assume that this pollutant is essentially local in the sense that emissions only affect local air quality within the city limits.[6] I show that a central planner taking into account the trade-off between agglomeration externalities and congestion forces decides to differentiate local emission taxes by setting the relative level of tax in larger cities higher.

I apply this spatial equilibrium framework to the specific case of France. This empirical setting is interesting for three reasons. First, a large share of harmful pollutant emissions result from industrial activities in France.[7] Second, local air concentrations of pollutants in several French cities often reach levels well above World Health Organization (WHO)'s

[3]Pollution can equivalently be modelled as an input or a by-product of industrial production in the standard Copeland & Taylor (2004)'s setup.

[4]This congestion effect potentially encompasses polluting emissions related to residential energy consumption and transportation. Even though, I do not explicitly model these pollution sources, they affect welfare negatively. I assume that these sectors are not covered by the emission tax on industrial pollution.

[5]The free mobility assumption leads to an equilibrium that is informative in the long run. In Appendix A.4, I investigate the alternative assumption of no across-city migration, which is more informative in the short run. Using the simplified setup of Section 2, I find that the main result holds when the central planner maximizes the average per capita welfare of workers.

[6]In reality, air pollutants can potentially travel across space depending on climatic conditions (especially, wind), some pollutants more than others. Since I consider commuting zones instead of cities in the empirical application, the assumption is however quite realistic since most pollutants do not travel very far. Furthermore, introducing an atmospheric circulation extension is outside of this paper's scope.

[7]In 2012, industrial activities (fossil fuel combustion, chemical reactions, waste treatment or other man-

guidelines.[8] Third, air quality regulations implemented in the last decades in France tend to be more stringent in larger cities, which are also more polluted and concentrate more industrial polluting activities. Figure 1 illustrates salient stylized facts that I incorporated into the model. The upper panel shows that the larger the number of workers within a French commuting zone, the higher the mean PM 2.5 concentration in 2012. The middle panel shows that the most populated commuting zones are also the most polluting in terms of PM 2.5 emitted by industrial activities. The lower panel shows that local marginal cost of emitting PM2.5 for industrial firms tend to increase with the size of the local labor market.[9]

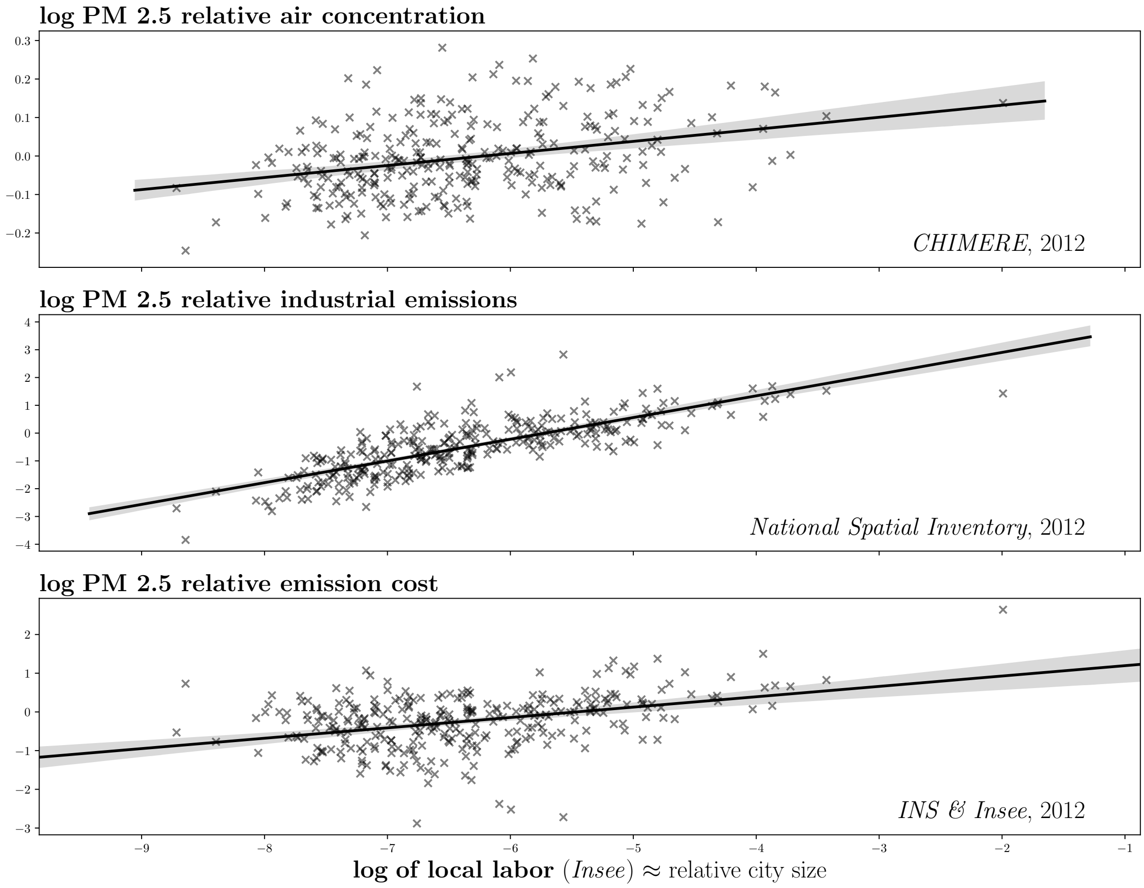


Figure 1: PM2.5 concentrations, emissions and marginal costs across commuting zones
Note: The upper panel plots the 2012 mean air concentration in PM2.5 at the commuting zone level as a function of the local number of workers. The middle panel plots the 2012 PM2.5 emissions from industrial activities at the commuting zone level as a function of the local number of workers. The lower panel plots the implicit marginal cost of emitting PM2.5 in the industry as a function of the local number of workers. This cost is the ratio of total wage payments in the commuting zone over the quantity of PM2.5 emitted by the industry. *City-level PM2.5 air concentration is from Chimere. PM 2.5 industrial emissions are from the National Spatialized Inventory. Labor data comes from the Insee and is a count of the number of workers employed in the employment area. All values are for 2012.*

ufacturing processes) emitted around 40% of PM2.5 and SO2 national emissions and around 25% of PM10 national emissions (see Figure (12) and Tables (5), (6) and (7) for details).

[8]Figure (13) shows that, in 2012, both mean and maximum air concentration measurements from ground monitors over the country for these substances have remained above WHO guidelines.

[9]Assuming that homogeneous firms in different cities have access to the same Cobb-Douglas production function between labor and PM2.5 emissions (assumptions used in the illustrative model in Section 2), I compute the marginal cost of PM2.5 emissions as the ratio of total wage payments over city-level PM2.5 emissions.

Using an extensive set of city-level, firm-level, and plant-level data, I first estimate the parameters of the model. I estimate key elasticities that determine the strength of pollution externalities, agglomeration economies and general congestion effects. I also provide novel estimates of sector-specific elasticities governing the substitution of emissions and labor in industrial production, using an instrumental variable approach. Second, I retrieve the distributions of amenities, productivities and emission taxes across cities from the observed endogenous distributions of populations, wages and emissions. Doing so, I obtain the set of current city-specific marginal costs of emitting atmospheric pollutants for industrial firms implied by current French air pollution regulations. I find that these costs are higher in large cities than in small cities. Specifically, these costs are higher in more productive cities, but not necessarily in cities endowed with more amenities.

I then characterize the welfare implications of such spatial distribution of marginal costs of emissions. I show that more stringent emission regulations in large cities enabled them to reduce pollution emissions and become larger than what they would have been under a counterfactual uniform regulation. This implies higher levels of welfare for workers living in all cities. To go further, I then identify the spatial distribution of emission taxes that would maximize workers' welfare. It would be optimal to raise emission taxes in cities with high amenities, leading to further concentration of workers in these cities, and thus exacerbating the uneven distribution of activities across space.

This paper relates to several literatures. First, it contributes to the recent quantitative spatial economics literature on the distribution of economics activities across space (Allen & Arkolakis, 2014; Redding & Rossi-Hansberg, 2017; Redding, 2020; Allen et al., 2020). Extending the framework of Rosen (1979)-Roback (1982), this literature can be decomposed into two main strands. The first one assumes that workers have heterogenous preferences over amenities and studies mainly the location decision of workers across cities (Moretti, 2013; Diamond, 2016; Almagro & Domínguez-Iino, 2021). An influential paper in this literature is Diamond (2016) that endogenizes local amenities to explain the spatial distribution of skilled and unskilled workers across U.S. cities. Diamond considers that these amenities are positively influenced by the fraction of skilled over unskilled workers. In my paper, I also introduce endogenous amenities affected by congestion forces, based on total population levels, and by the pollution externality. The second strand assumes homogenous workers to study the location decision of firms across cities and the optimal size of cities (Henderson, 1974; Eeckhout & Guner, 2015; Borck & Tabuchi, 2018; Gaubert, 2018; Albouy et al., 2019). Focusing on the supply side of the economy, namely polluting industries, I make the same assumption. In this block, Allen & Arkolakis (2014) and Allen et al. (2020) develop a spatial quantitative model where characteristics of heterogeneous locations determine the equilibrium distribution of economic activity across space. They assume endogenous amenities and agglomeration economies.[10] Even though my model is an extension of these two papers, I

[10]In this respect, I include the "productivity advantage of large cities" that have been identified for France

depart from them by focusing on a particular source of congestion externalities, namely local air pollution, and investigate its effect on the spatial distribution of economic activities.

My paper is connected to a second literature that investigates the distribution of pollution across cities (Glaeser & Kahn, 2010; Carozzi & Roth, 2019; Colmer et al., 2020; Borck & Schrauth, 2021; Eeckhout & Hedtrich, 2021) and its congestion effects (Drut & Mahieux, 2015; Leturque & Sanch-Maritan, 2019; Hanlon, 2019). Carozzi & Roth (2019) and Borck & Schrauth (2021) both show that denser cities are also more polluted (respectively for the US and Germany). Similarly, Colmer et al. (2020) uncover large disparities in the spatial distribution of PM2.5 concentration across census tracts in the US. In particular they find that increasing local population leads to higher levels of ambient PM2.5 pollution. They also identify that pollution decreased in locations where regulation became more stringent (namely census tracts in non-attainment with air quality standards). In my paper, I provide a framework explaining such spatial distribution of pollution: my setup encompasses the effect of local population, local income and local regulations on local emissions. Moreover, with a focus on France, Drut & Mahieux (2015) and Leturque & Sanch-Maritan (2019) exposed that agglomeration gains were dampened by the local levels of pollution. My framework includes this effect by assuming that local pollution actually acts as a congestion force. This mechanism is supported by findings from Hanlon (2019) who showed for the UK that local pollution industrial emissions reduced long-run city employment and population growth. Damage from pollution also have been found to be heterogenous across space depending on local characteristics (Aldeco et al., 2019; Deryugina et al., 2021; Desmet et al., 2021; Alvarez & Rossi-Hansberg, 2021). The assumption of constant elasticity of damage from pollution implicitly acknowledges this heterogeneity to the extent that large cities concentrate most of welfare loss due to pollution.

Finally, my paper contributes to the literature on optimal place-based policies chosen by a central planner. In particular, Fajgelbaum & Gaubert (2020) investigate how income transfers across cities could maximize welfare by correcting spatial externalities, assuming free mobility of workers across cities. Similarly, Suarez-Serrato & Zidar (2016) consider a tax rate on firms' profits that varies across cities. In contrast to these papers, I assume an input-specific policy instrument that induces reallocation across inputs, therefore it affects the endogenous amenity level per city through emissions. This place-based literature also contains studies including endogenous pollution (Lange & Quaas, 2007; Kyriakopoulou & Xepapadeas, 2013; Yamada, 2020; Pflüger, 2021). A paper closely related to mine is Yamada

by Combes et al. (2012) and more recently by Gaubert (2018). However, I do not make any assumption on the source of these agglomeration economies. Notably, the model does not include any sorting mechanism that may partly explain higher productivities observed in larger cities (Baldwin & Okubo, 2005; Combes et al., 2008). Another source of agglomeration comes from the costly trade assumption. Recently Bartelme (2018) showed that trade costs between US cities explained a large fraction of the spatial distribution of economic output. At the same time, my model embodies general counteracting congestion forces which limit the scale of economic concentration.

(2020) that also considers atmospheric pollutants as a congestion force and shows that imposing more stringent air quality regulations in a specific set of large Chinese cities could lead to welfare gains. In contrast to this paper, I solve for the optimal distribution of local emission regulations instead of comparing ad-hoc distributions of policies inspired by planned policy projects.

The remainder of the paper is organized as follows. In the next section, I provide a simplified framework to illustrate how introducing a local pollution externality into the standard spatial equilibrium model generates new insights. In the third section I provide some context on air quality regulations in France. A more realistic general model is detailed in the fourth section. The fifth section presents the estimation of the model's parameters. Finally, the last section contains results from numerical welfare optimization problems and discusses the results.

2 Optimal Environmental Policy in a Simple Spatial Model

Following the Rosen (1979)-Roback (1982) framework, I consider a given set of C cities in which a fixed population of workers can live. Throughout the paper, I assume homogenous workers.[11] The per capita utility of the representative worker in city j is given by:

$$u_j = a_j Z_j^{-\gamma} L_j^{-\delta} c_j, \tag{1}$$

where a_j is the local idiosyncratic endowment in amenities, L_j is the local population of workers, Z_j is the quantity of atmospheric pollutants emitted by the local industry, and c_j is the local per capita consumption of a tradable good. This utility function captures the fact that workers value the consumption of the tradable good, which depends on their income, as well as other characteristics of the locations they live in. In this paper, I focus on the local detrimental welfare effect of industrial emissions of atmospheric pollutants, as industrial activities are major emitters of several harmful pollutants.[12] In equation (1), a positive γ implies that local industrial emissions Z_j negatively affect the local air quality and therefore welfare (through health damages). For simplicity, I abstract from spatial pollution spillovers and assume that local emissions only affect local welfare. Workers also bear various agglomeration costs, which are accounted for by the general congestion term $L_j^{-\delta}$. This cap-

[11]As shown in Allen & Arkolakis (2014), my choice for workers' preferences is isomorphic to a model where heterogeneous workers have idiosyncratic utility shocks, according to a Fréchet distribution, that are independent and identically distributed across locations and individuals. Such heterogeneity across workers is a dispersion force and is captured by the parameter δ in my model.

[12]In 2012, the industry emitted 40%, 42% and 23% of total PM2.5, SO2 and PM10 emissions, see Appendix A.1.

tures local externalities from agglomeration, such as commuting costs and housing prices.[13] This general congestion term also captures the detrimental effect on welfare of atmospheric pollution emitted by non-industrial activities (transport, residential heating, or energy production). Furthermore, each city is exogenously endowed with a fixed level of amenities a_j that explains the location choices of local workers once spatial differences in consumption and local externalities are accounted for. For instance, amenities include geographical environmental attributes, available space, institutional and social installations or any local external factor that explains why some locations are more attractive than others. Workers benefit from agglomeration economies. I assume that labor productivity in location j is equal to $b_j L_j^{\nu}$. Elasticity ν governs the strength of agglomeration economies and b_j allows for idiosyncratic labor productivity differences between cities. Idiosyncratic productivity can vary across cities due to a wide array of local characteristics (better local institutions or more efficient local transport networks). Among other things, agglomeration economies can arise from local knowledge spillovers, labor markets pooling or local economies of scale.[14]

In each city, identical firms produce a homogenous tradable good using labor and emissions of a representative atmospheric pollutant with a Cobb-Douglas production function defined by an expenditure share α, with $1 > \alpha > 0$. Following Copeland & Taylor (2004), it is equivalent to assuming that, rather than being an input in production, emissions are a by-product of production. Then, firms can divert a fraction of their labor force to abate emissions. The efficiency of abatement is governed by α: the lower it is, the more efficient the abatement technology is to reduce firms' emission intensity. In location j, a unit of labor costs the local wage w_j and a unit of emissions costs the local emission tax t_j. This emission tax is a policy instrument set by a central planner. It represents the local pollution regulations (including technology standards, emission limits, or emergency responses). I assume that proceeds from the emission tax are locally redistributed to workers. These combined assumptions lead to the equilibrium distribution of industrial emissions across cities (see Appendix A.3 for computing details):

$$Z_j = \alpha \frac{w_j}{t_j} L_j. \tag{2}$$

which illustrates that industries emit more pollution in cities that are larger, and where wages are higher. Conversely, higher emission taxes reduce local emissions.

I assume that competition is perfect and trade is costless, which implies the spatial equalization of the output price and leads to the equilibrium distribution of local wages. Setting the average wage as the numeraire would add general equilibrium effects to the

[13] Allen & Arkolakis (2014) show that constant elasticity congestion costs are isomorphic to the Helpman (1998)-Redding (2016) setup where workers spend a fixed share of their income on a non-tradable good, which can include housing.

[14] Duranton & Puga (2004) provides a large range of models that deliver this constant elasticity function. Combes & Gobillon (2015) provides a survey of the empirical literature that documents agglomeration economies.

equilibrium: when regulation changes in a given city there would be spillovers on wages in other cities. I abstract from such effects to focus only the interplay between agglomeration, congestion, and industrial pollution externalities. Therefore, I normalize the output price to 1. As a result, per capita consumption of the tradable good is equal to the local wage: $c_j = w_j$. Output price normalization also pins down local wages in location j :

$$w_j = b_j L_j^{\nu} t_j^{-\frac{\alpha}{1-\alpha}}, \tag{3}$$

which implies that wages are higher in cities that are more productive and in cities that are larger, because of agglomeration economies. It also implies that raising local emission taxes has a negative effect on local wages.

Substituting for (2) and (3) in (1), the indirect utility in location j can be written as:

$$u_j = a_j b_j^{1-\gamma} L_j^{-\theta} t_j^{\frac{\gamma-\alpha}{1-\alpha}}, \quad \text{with } \theta = \delta + \gamma - \nu(1-\gamma). \tag{4}$$

The exponent θ on the population term in equation (4) reflects the fact that, since $\gamma > 0$, the emission externality both reinforces congestion and weakens agglomeration economies (when cities grow, agglomeration economies increase local wages, and firms become more pollution intensive).

Finally, I assume free mobility of workers across locations.[15] As a result, utility is equalized across space in equilibrium and is equal to a level denoted $\bar{u}$. To see this, imagine two cities with distinct levels of welfare: workers living in the city with lower welfare have an incentive to move to the city with higher welfare up to the point where the marginal gain of moving is compensated by the marginal cost of congestion. Combining this assumption with (4), we obtain, for two cities i and j:

$$\frac{L_j}{L_i} = \left(\frac{a_j}{a_i}\right)^{\frac{1}{\theta}} \left(\frac{b_j}{b_i}\right)^{\frac{1-\gamma}{\theta}} \left(\frac{t_j}{t_i}\right)^{\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}}, \tag{5}$$

which implies that the distribution of workers across cities is a function of the distributions of amenities, productivities, and emission taxes. Depending on the relative strengths of the three externalities, local populations may be positively or negatively, correlated with these local characteristics. Equation (5) also illustrates that there exists a unique spatial equilibrium if and only if $\theta \neq 0$. This condition can be expressed as a condition on the elasticity of the pollution externality: $\gamma \neq -\frac{\delta-\nu}{1+\nu}$. Assuming that pollution is a congestion force (i.e. $\gamma > 0$), this condition ensures that the pollution externality does not exactly offset the combined effect of general congestions effect and agglomeration economies. When congestion effects strictly outweigh agglomeration economies, this condition always holds.

[15]This assumption is informative of long term equilibria, where within-country migration costs can be considered as low. In Appendix A.4, I consider for robustness the alternative assumption, where local populations are fixed and do not adjust in equilibrium.

Without any loss of generality, I normalize the total population to 1, so that $\sum_{j\in C} L_j = 1$, and compute the common level of welfare $\bar{u}$ reached in equilibrium:

$$\bar{u} = \left[\sum_{j\in C} a_j^{\frac{1}{\theta}} b_j^{\frac{1-\gamma}{\theta}} t_j^{\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}}\right]^{\theta} \tag{6}$$

The optimal policy for a central planner is to maximize $\bar{u}$ by adjusting the set of local emission taxes across cities. In Appendix A.3, I show that $1 \geq \gamma \geq \alpha$ and $1 \geq \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}$ are necessary and sufficient conditions for $\bar{u}$ to be concave and the optimization problem to have a unique solution. First, $1 \geq \gamma$ means that the positive direct effect of productivity on wages outweighs its indirect negative effect that causes firms to become more pollution intensive when wages increase. Second, $\alpha < \gamma$ means that the negative effect on wages of raising the local emission tax is weaker than the positive effect it has on air quality. Third, $1 \geq \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}$ ensures that the whole population is not concentrated in a unique city. The welfare function specified by equation (6) is homogeneous of degree $\frac{\gamma-\alpha}{1-\alpha}$ with respect to the set of emission taxes. This implies that, if I multiply all city level emission taxes by a common factor, workers' welfare $\bar{u}$ is multiplied by a power $\frac{\gamma-\alpha}{1-\alpha}$ of this factor. Therefore, I normalize the average emission tax $\bar{t}$ to 1 and focus on the distribution of taxes across cities.

Proposition. *Consider a set of cities C, with exogenous amenities $\{a_j\}_{j\in C}$ and productivity $\{b_j\}_{j\in C}$, with relative populations specified in equation (5). If $1 \geq \gamma > \alpha$ and $1 > \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}$, there is a unique set of emission taxes $\{t_j^*\}$ that maximizes the welfare function specified in equation (6) under the constraint that $\bar{t} = 1$ and it is defined by:*

$$\frac{t_j^*}{t_i^*} = \left[\left(\frac{a_j}{a_i}\right)^{\frac{1}{\theta}} \left(\frac{b_j}{b_j}\right)^{\frac{1-\gamma}{\theta}}\right]^{\frac{1}{1-\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}}}. \tag{7}$$

Equation (7) indicates that the relationship between optimal emission taxes and local characteristics depends on the sign of θ. In particular, $\theta > 0$ means that congestion forces outweigh agglomeration economies and more workers locate in cities with a higher level of amenities and productivity (see equation (5)). In this case, equation (7) indicates that, to maximize workers' welfare, the central planner imposes higher emission taxes in cities with a higher level of amenities and productivity.

An important implication of my model is that emission taxes should be heterogeneous across space. However, national governments usually enforce regulations that are spatially uniform. Therefore, it is interesting to compare the implications of enforcing the set of emission taxes described by (7) to the implications of enforcing a uniform set of emission taxes (each equal to 1). Denoting L_j^* and L_j^u the populations of city j under the optimal and uniform sets of emission taxes, equations (5) and (7) imply:

$$\frac{t_j^*}{t_i^*} = \frac{L_j^*}{L_i^*} = \left(\frac{L_j^u}{L_j^u}\right)^{\frac{1}{1-\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}}}. \tag{8}$$

which means that if an optimal set of emission taxes exists, taxes should be higher in cities that are larger when the tax is uniform. Moreover, the sufficient condition for an optimal set of taxes to exist implies that $\frac{1}{1-\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}} > 1$. As a result, equation (8) indicates that implementing the optimal set of emission taxes reinforces the spatial concentration pattern observed under uniform taxes. When congestion forces outweigh agglomeration forces and higher labor productivity translates in higher income, it is optimal that more workers locate in cities with high amenities and productivity endowments (to compensate the costs of agglomeration). However, firms do not fully internalize the detrimental impact of their emissions on local welfare and pollute too much in cities that could attract even more workers if they were less polluted. This is why the optimal set of emission taxes imposes a relatively higher marginal cost of polluting in these cities, compared to cities where it is not optimal to concentrate workers. As a result, firms in these cities are less emission intensive, which lowers the level of emissions and makes these cities more attractive to workers.

The set of emission taxes described by equation (7) corrects all externalities, even the ones that are not caused by industrial pollution emissions. To analyze the case where the central planner corrects for these "non-emission" externalities using external policy instruments, one can derive the optimal set of emission taxes when the emission externality is the only externality. Assuming $\delta = \nu = 0$ in equation (7), the optimal set of emission taxes still imposes higher emission costs in the same cities as in the case where all externalities are corrected by the set of emission taxes.

Finally, equation (7) reveals that if there were no industrial emissions externality, that is to say if $\gamma = 0$, the optimal set of emission taxes would still be non-uniform, to correct for congestion and agglomeration externalities. In particular, when $\gamma = 0$, equation (7) leads back to the standard result showing that large cities (with the highest productivities and amenities) should be even larger (Eeckhout & Guner, 2015; Gaubert, 2018; Albouy et al., 2019). Moreover, the average tax normalization imposes that increasing emission taxes in large cities implies decreasing them in smaller cities. Equation (3) shows that it is equivalent to decreasing wages in large cities and increasing them in small cities. This is similar to the system of optimal income transfers identified by Fajgelbaum & Gaubert (2020) when congestion costs outweigh agglomeration economies.

3 General Spatial Model

In this section I extend the model to (1) an industry composed of several polluting sectors, (2) continuums of heterogeneous firms that compete monopolistically over differentiated varieties, (3) costly trade between cities, and (4) general equilibrium effects from local changes in emission taxes.

3.1 Setup

The per capita utility of the representative worker in city j follows equation (1). The industry is now composed of S distinct sectors and c_j is the industrial composite:

$$c_j = \prod_{s\in S}\left(\sum_{i\in C}\int_{\omega\in\Omega_{is}} c_{ijs}(\omega)^{\frac{\sigma_s-1}{\sigma_s}}d\omega\right)^{\frac{\sigma_s}{\sigma_s-1}\beta_s}, \tag{9}$$

with $c_{ijs}(\omega)$ the quantity of variety ω produced in city i and consumed in city j, Ω_{is} the continuum of varieties produced in sector s in city i, σ_s the elasticity of substitution between varieties in sector $s \in S$, and β_s the share of income spent on varieties from sector s by each worker (with $\sum_{s\in S}\beta_s = 1$). These parameters are assumed to be the same across cities.

The aggregated price index of the industrial good in city j, P_j is such that:

$$P_j = \prod_{s\in S} P_{js}^{\beta_s}, \text{ and } P_{js} = \left(\sum_{i\in C}\int_{\omega\in\Omega_{is}} p_{ijs}(\omega)^{1-\sigma_s}d\omega\right)^{\frac{1}{1-\sigma_s}}, \tag{10}$$

with P_{js} are city and sector specific price indices, and $p_{ijs}(\omega)$ the unit price of variety ω produced in city i in sector s and consumed in city j. I assume that there are no friction on local labor markets, so that that wages are equal across sectors of production. As a result, per capita consumption of the tradable good is given by $c_j = \frac{w_j}{P_j}$ and workers' indirect utility is:

$$u_j = a_j Z_j^{-\gamma} L_j^{-\delta}\frac{w_j}{P_j}. \tag{11}$$

In each city j there is an infinite supply of entrepreneurs in each sector s that can choose to pay a fixed sector specific entry cost f_s^e to draw a productivity ϕ from a Pareto distribution G_{js}:

$$G_{js}(\phi) = 1 - \left(\frac{\phi}{b_{js}}\right)^{-\theta_s}, \tag{12}$$

with $b_{js} = b_j b_s L_j^{\nu_s}$. These assumptions extend section 2's assumptions of idiosyncratic labor productivity and agglomeration economies to the case of multiple sectors and heterogeneous firms. In particular, the elasticities of agglomeration economies, $\{\nu_s\}_{s\in S}$, are sector specific. In addition, idiosyncratic sector and city specific labor productivities $\{b_{js}\}_{(j,s)\in C\times S}$ are multiplicatively separable between idiosyncratic city specific productivities $\{b_i\}_{j\in C}$, that are common across sectors, and sector specific productivities $\{b_s\}_{s\in S}$, that are common across cities.

Each firm produces a specific variety of the industrial good using labor following the function:

$$q_{ijs}(\phi) = (1 - a(\phi))\,\phi l_{ijs}(\phi) \tag{13}$$

where $q_{ijs}(\phi)$ is the quantity produced by a firm of productivity ϕ in city i and sold on city j's market, and $l_{ijs}(\phi)$ is the quantity of labor used for production. Variable $a(\phi)$ corresponds to the fraction of the firm's labor force used to abate emissions caused by production. Production releases pollution as a by-product according to:

$$z_{ijs}(\phi) = (1 - a(\phi))^{\frac{1}{\alpha_s}} \phi l_{ijs}(\phi) \tag{14}$$

where $z_{ijs}(\phi)$ is the quantity of the representative pollutant emitted by a firm of productivity ϕ in city i for the production sold in city j. $\{\alpha_s\}_{s\in S}$ are the sector specific expenditure shares on emissions. Heterogeneity in these parameters across cities accounts for the fact that some industrial sectors may be more emission intensive than others. I assume that, in city j, emissions are taxed by the central planner at a rate t_j, which is the same across sectors. Local proceeds from this tax are redistributed to local workers.

These production and pollution functions closely follows Copeland & Taylor (2004) and Shapiro & Walker (2018) and are standard in the literature. Combining equations (13) and (14) implies the following:

$$q_{ijs}(\phi) = z_{ijs}(\phi)^{\alpha_s} \left(\phi l_{ijs}(\phi)\right)^{1-\alpha_s}, \tag{15}$$

which is a Cobb-Douglas function that combines two inputs: emissions and labor. As a result, total industrial emissions of the representative pollutant, in city i, are given by:

$$Z_i = \sum_{s\in S} \sum_{j\in C} \int_{\phi} z_{ijs}(\phi) dG_{is}(\phi) \tag{16}$$

Trade is costly between cities. To export to city j, firms in city i have to pay an origin-destination specific iceberg cost τ_{ij}. As a result, any firm sells to all cities.[16] Combining the assumption of monopolistic competition with equations (9) and (15) implies that the unit price of a variety produced in city i in sector s by a firm with productivity ϕ and delivered in city j follows:

$$p_{ijs}(\phi) = \frac{\sigma_s}{\sigma_s - 1} \frac{\tau_{ij} c_{is}}{\phi^{1-\alpha_s}}, \text{ with } c_{is} = \kappa_{\alpha_s} t_i^{\alpha_s} w_i^{1-\alpha_s}, \tag{17}$$

and $\kappa_{\alpha_s} = \alpha_s^{-\alpha_s}(1-\alpha_s)^{-1+\alpha}$. Revenues and profits of a firm with productivity ϕ in city i respectively follow:

$$r_{is}(\phi) = \sum_{j\in C} p_{ijs}(\phi) q_{ijs}(\phi), \text{ and } \pi_{is}(\phi) = r_{is}(\phi)/\sigma_s. \tag{18}$$

I assume that entrepreneurs enter production until their expected profits equal the fixed entry cost of drawing a productivity from the local Pareto productivity distribution. I

[16] The model could be extended to accommodate for origin-destination fixed trade costs as in Shapiro & Walker (2018). Productivity distributions of firms would then be left-truncated by endogenous zero-profit productivity cutoffs. In this case, only the most productive firms in a given city would be selling goods in all other cities.

express fixed entry costs in the aggregate factor price, as in Bernard et al. (2007). Thus, for each pair of city and sector, I have the free entry condition:

$$\int_{\phi} \pi_{js}(\phi) dG_{js}(\phi) = c_{js} f_s^e. \tag{19}$$

In equilibrium, goods markets clear for each city and sector pair so that workers' expenditures are equal to firms' sales. These goods markets clearing conditions can be written as:

$$P_{js}^{1-\sigma_s} = K_s \sum_{i \in C} \tau_{ij}^{1-\sigma_s} \tilde{M}_{is} (t_i^{\alpha_s} w_i^{1-\alpha_s})^{-\sigma_s} b_i^{(1-\alpha_s)(\sigma_s-1)} L_i^{\nu_s(1-\alpha_s)(\sigma_s-1)}, \tag{20}$$

with $\tilde{M}_{is} = \frac{\sigma_s}{\beta_s} f_s^e M_{is} c_{is}$, and M_{is} the mass of firms that produce in sector s and city i in equilibrium.

The local labor markets also clear in equilibrium, so that the sum of employment over all sectors is equal to the local population. For each city j, it can be written as:

$$w_j L_j = \sum_{s \in S} \beta_s \tilde{M}_{js}, \tag{21}$$

which can be combined with equations (19) and (20) to show that, for any distribution of population $\{L_i\}_{i \in C}$ across cities, wages in city i follow:

$$\begin{aligned} w_i L_i = &\sum_{s \in S} \beta_s t_i^{\alpha_s \sigma_s} w_i^{(1-\alpha_s)\sigma_s} b_i^{-(1-\alpha_s)(\sigma_s-1)} L_i^{-\nu_s(1-\alpha_s)(\sigma_s-1)} \\ &\times \sum_{j \in C} \mu_{ijs} \frac{w_j L_j}{\sum_{k \in C} \mu_{kjs} t_k^{\alpha_s \sigma_s} w_k^{(1-\alpha_s)\sigma_s} b_k^{-(1-\alpha_s)(\sigma_s-1)} L_k^{-\nu_s(1-\alpha_s)(\sigma_s-1)}} \end{aligned} \tag{22}$$

, where, for each sector s, $\boldsymbol{\mu_s} = \{\mu_{ijs}\}_{(i,j) \in C^2}$ are the terms of the inverse of the $\boldsymbol{\tau}^{1-\sigma_s}$, with $\boldsymbol{\tau}$ the iceberg trade costs matrix. Note that, as long as for any (i, j), $\tau_{ii} < \tau_{ij}$, and for any s, $\sigma_s > 1$, $\boldsymbol{\tau}^{1-\sigma_s}$ is strictly diagonally dominant. Therefore, it can be inverted under these conditions. Equation (22) illustrates the general equilibrium mechanisms that are included in this section: when the emission tax changes in city j, it affects wages in city j.

As in section 2, I assume free mobility and workers' utility maximization. As a result, welfare is the same across cities and equal to $\bar{u}$. From equation (1), free mobility implies that, in each city j:

$$a_j Z_j^{-\gamma} L_j^{-\delta} \frac{w_j}{P_j} = \bar{u}. \tag{23}$$

Finally, I assume that the total population of workers across cities is fixed and normalize it to 1 without loss of generality, so that:

$$\sum_{j \in C} L_j = 1, \tag{24}$$

which can be combined to equation (23) to show that:

$$L_j = \frac{\left(a_j Z_j^{-\gamma} \frac{w_j}{P_j}\right)^{1/\delta}}{\sum_{i \in C}\left(a_i Z_i^{-\gamma} \frac{w_i}{P_i}\right)^{1/\delta}}. \tag{25}$$

Given equation (16), for any set of emission taxes and masses of firms in each sector, emissions of pollution in city j follow:

$$Z_j = \sum_{s \in S} \beta_s \alpha_s \frac{\tilde{M}_{js}}{t_j}, \tag{26}$$

which, combined to equations (21), (22), and (25), implies that is an homogeneous function of degree -1 with respect to the average level of emission taxes across cities. Similarly, wages and population are homogeneous functions of degree zero with respect to the average level of emission taxes across cities. Combining equations (10) and (10) implies that local price indices are homogeneous functions of degree $\sum_{s \in S} \beta_s \alpha_s \frac{\sigma_s}{\sigma_s - 1}$ with respect to the average level of emission taxes across cities. As a result, equation (23) implies that workers' welfare is a homogeneous function of degree $\gamma - \sum_{s \in S} \beta_s \alpha_s \frac{\sigma_s}{\sigma_s - 1}$ with respect to the average level of emission taxes across cities. Consequently, when investigating the central planner's welfare maximization problem, I focus on the distribution of emission taxes across cities and keep the average emission tax to 1.

3.2 Equilibrium and Optimal Set of Emission Taxes

An equilibrium is defined as a set of populations $\{L_j\}_{j \in C}$, wages $\{w_j\}_{j \in C}$, masses of industrial firms $\{M_{js}\}_{(j,s) \in C \times S}$, price indices $\{P_{js}\}_{(j,s) \in C \times S}$, and welfare $\bar{u}$, that solve the set of free entry conditions (19), goods market clearing conditions (20), local labor markets clearing condition (21), free mobility conditions (23), and national population clearing condition (24). The equilibrium is identified, to the extent that the number of equations is equal to the number of endogenous variables. As in section 2, I define the optimal set of emission taxes as the set of taxes that maximizes workers' welfare $\bar{u}$, under the constraint that the average tax remains equal to 1. In this extended version of my model, there are no closed-form solutions, neither for the equilibrium distributions of the endogenous variables, nor for the set of emission taxes that solves the maximization problem. Therefore, I rely on numerical methods to solve the model.

4 Estimation of the Model

To perform a quantitative policy analysis, I take the model to the data. In particular, I consider the set of French commuting zones as the set of cities in which workers live.[17] This set commuting zones constitutes a partition of the full French metropolitan territory and each zone is defined statistically to be an area where local inhabitants both work and live. Using this definition of cities, I estimate the parameters of the model presented in section 3. First, I estimate equation (11) to retrieve elasticities γ and δ. Second, I use some of the model's predictions to estimate sector-specific expenditure shares $\{\beta_s\}_{s\in S}$, elasticities of substitution $\{\sigma_s\}_{s\in S}$, Pareto shape parameters $\{\theta_s\}_{s\in S}$, elasticities of agglomeration economies $\{\nu_s\}_{s\in S}$, and emission elasticities $\{\alpha_s\}_{s\in S}$. Third, I compute a matrix of road travel time between French commuting zones and use it as a proxy for the trade cost matrix. Finally, equipped with this estimated model, I retrieve the model's primitives – the level of amenities, productivity, and emission taxes – from data on the distribution of workers, wages, and industrial PM2.5 emissions across cities in 2012.

4.1 Data

In this subsection, I describe the datasets used in the quantitative exercise. Firm-level datasets come from confidential French administrative data. I use annual balance sheets and income statements for the universe of French firms from 1994 to 2016 as reported in the FICUS databases for 1994-2007 and in the FARE databases for subsequent years. A firm is identified by a stable administrative code called SIREN. The main variables of interest are total sales, average employment (number of workers and total wages paid), location, and the main sector of activity.

Plant-level information on energy consumption comes from the EACEI (*Enquête Annuelle sur la Consommation d'Energie dans l'Industrie*) surveys, which are available from 1994 to 2016. These surveys include all energy-related expenditures, with details on energy types and fuels (quantity consumed and expenditures) at the plant level. The types of energies reported are electricity (consumed and self-generated), steam, natural gas and other types of gas, coal, lignite, coke, propane and butane, domestic and heavy fuels, oil and other types of petroleum products. The surveys also provide the plant-level number of employees. The surveys cover all large plants (over 20 employees) in the industrial sectors – with the exception of the power sector – and a subset of smaller plants (between 10 and 19 employees) that is randomly selected each year. On average, between 8,000 and 11,000 plants are included in the annual survey. I use these surveys to compute atmospheric emissions at the plant-level. The Ominea database (CITEPA, 2020) provides emission factors that associate to each

[17] I use the 2010 definition of the French *"zones d'emplois"*.

fuel the corresponding amount of pollutants that are emitted. For each plant, I multiply the amount of each fuel consumed by the corresponding emission factor and sum across fuels to compute the total quantity of pollutant emitted. To ensure the validity of this approach, I compute correlation between these constructed emission values and actual values declared by large plants under the European directive and publicly available in the European Pollutant Release & Transfer Register (E-PRTR) between 2003 and 2016.[18] Table (8) in Appendix A.5.1 displays Pearson correlation coefficients along with statistical significance. It appears that my construction emissions values correlate well with the actual emissions of these large plants.

To estimate γ, the elasticity of environmental damage from pollution, I use geographic emissions from the Emissions Database for Global Atmospheric Research (EDGAR) from the European Commission's Joint Research Center (JRC). EDGAR provides annual gridded emissions of atmospheric pollutants disaggregated across polluting sectors at 0.1x0.1 degree resolution. Data is publicly available on the JRC's dedicated website [19]. Information on the methodology can be found in Crippa et al. (2018). The benefit from using this dataset rather that EACEI computed emission is that EDGAR aggregated values provide an exhaustive picture of all emissions (while EACEI surveys only account for the largest plants). I use EDGAR's data from 2000 to 2015. I spatially aggregate it at the commuting zone level by intersecting the EDGAR grid with official ESRI shapefiles for French municipalities aggregated in commuting zones.

To retrieve the model's primitives, I use a 2012 cross-section dataset of wages, labor, and industrial PM2.5 emissions at the commuting zone level. For wages and labor, I use data from the *Insee*'s website. For PM2.5 industrial emissions, I use the National Spatial Inventory (INS) INS (2020). The INS is a publicly available dataset reporting emissions of around 40 pollutants from natural and anthropogenic sources.[20] The data is available at the municipality level, which is a more disaggregated spatial unit than EDGAR's grid cells. However, data is only available for 2004, 2007, and 2012. Therefore, I use the INS only to retrieve the model's primitives, which are identified through cross-sectional variation, and rely on EDGAR for the panel estimation. Precisely, I use PM2.5 emissions from the manufacturing industry and from production processes. These corresponds to codes 3 and 4 in the Selected Nomenclature for Air Pollution.

[18]Although actual plant-level measures of atmospheric pollutants emissions are publicly available at the plant level in the European Pollutant Release and Transfer Register (E-PRTR), it only includes the largest plants resulting in a much restricted sample of industrial plants than the EACEI sample. Using the E-PRTR data would not allow me to estimate sector-specific emission elasticities.

[19]`https://edgar.jrc.ec.europa.eu/index.php/dataset_ap50`

[20]`http://emissions-air.developpement-durable.gouv.fr/index.html`

4.2 Estimation of Congestion and Pollution Damage Elasticities

Equation (11) can be re-written as a log-linear relationship between wages, utility, amenities, price indices, local labor supply and emissions:

$$\log w_{it} = \log u_{it} - \log a_{it} + \log P_{it} + \delta \log L_{it} + \gamma \log Z_{it}. \tag{27}$$

Introducing a time variable t, city i endowments in amenities and productivity may vary over time (for instance, due to enhancements to local transport networks or to the creation of museums, parks, etc.). Furthermore, given existing migration frictions, the assumption of free mobility may not hold in the short term (from one year to another). Therefore, welfare is potentially different across cities. In practice, I only observe wages, populations, and emissions. Thus, I run the following equation:

$$\log w_{it} = \delta \log L_{it} + \gamma \log Z_{it} + \chi_i + \mu_t + \epsilon_{it}. \tag{28}$$

Year fixed effects μ_t eliminate annual shocks that are common to all cities. City fixed effects χ_i absorb city-specific fixed characteristics. The error term ϵ_{it} corresponds to city-year specific shocks, including shocks to local amenities a_{it}, local welfare u_{it}, and local price indices P_{it}. Equation (20) implies that local price indices are functions of the set of local productivities, of the trade cost matrix and of the other endogenous variables. Hence, city-specific shocks to local productivities or to the trade cost matrix induce shocks on local price indices. Wages, populations and emissions are also functions of these shocks. As a result, a standard OLS estimation strategy would produces biased estimates. Indeed, any positive shock to local amenities in a given city positively affects local population and emissions. Therefore, the OLS identification assumption is unlikely to hold. To overcome this identification issue, I build shift-share instrumental variables for both local labor supply and local emissions.

To build my instruments, I follow the recent literature on shift-share instruments that has extended the Bartik (1991) approach (Card, 2001; Autor et al., 2013; Nunn & Qian, 2014; Bartelme, 2018; Bombardini & Li, 2020; Barrows & Ollivier, 2021). The idea is to approximate city-specific growth rates in populations and emissions using national growth rates of theses variables in disaggregated industries and interact them with city-specific shares in an initial period. Thus, I build two instruments as:

$$\tilde{L}_{it} = \sum_s L_{it_0 s} \frac{L_{st}}{L_{st_0}}, \text{ with } s \in \Omega_L, \tag{29}$$

and

$$\tilde{Z}_{it} = \sum_s Z_{it_0 s} \frac{Z_{st}}{Z_{st_0}}, \text{ with } s \in \Omega_Z. \tag{30}$$

Variables $\tilde{L}_{it}$ and $\tilde{Z}_{it}$ are respectively the instruments for L_{it} and Z_{it} in equation (28). These instruments are independent of city-year specific unobservable shocks to local amenities, productivity or trade costs. They are only functions of year-specific national shocks and city-specific initial sector shares in employment and emissions. National shocks are common to all cities and do not threaten identification when year fixed effects are included. Initial sector shares remain constant across time and do not threaten identification when city fixed effects are included. As thoroughly detailed in Goldsmith-Pinkham et al. (2020), a recent extensive analysis of shift-share instruments, in order for these two instruments to be exogenous and the identification assumption to be respected, my strategy implicitly assumes that unobserved shocks on productivity, amenities or trade costs are uncorrelated with initial industry shares.[21]

My estimation strategy builds on annual data on industrial emissions aggregated at the commuting zone level, on firm and plant-level data on wages, employment and activity codes also aggregated at the commuting zone level. Each firm is identified by a unique 9 digits identification code (SIREN) and each plant is identified by a 15 digits identification code (SIRET) with the 9 first digits corresponding to the SIREN of the firm to which the plant belongs. Based on these two codes I merge the firm and plant level panels. Thus, for each plant I have a value of the mean wage paid at the firm level. Based on plant-level zip codes, I compute average wages across local plants at the commuting zone level. This results in a panel of local wages across commuting zones from 2000 to 2015. The plant-level panel also includes the number of workers employed on average over the year as well as an activity code (in the French Activity Nomenclature). I sum across local plants for each sector to build a panel of total sectoral employment at the commuting zone level from 2000 to 2015. The right hand side of the estimating equation only depends on the total local employment across sectors, but the instrument defined in equation (29) builds on the sector local and national disaggregation across sectors. The French Activity Nomenclature distinguishes between hundreds of very precisely defined industries. I aggregate codes at a larger level which I call industries and end up with 16 different categories so that Ω_L is composed of agriculture, extraction activities, manufacturing activities, the energy sector, waste management, construction, trade, transport, hotels and restauration, telecommunications, finance, real-estate, public administrations, teaching sector, health, arts and other activities.

I build a commuting zone panel of pollutants emissions from 2000 to 2015 disaggregated across polluting sectors using the geographic emissions dataset from EDGAR. Data is available as annual 0.1 times 0.1 degree grid sets for each sectors. Each data point corresponds to the quantity of pollutant emitted within the grid cell annually per unit of area. Based on geographic coordinates I attribute each cell from the EDGAR grid to French commuting zones (based on publicly available geographic information on municipalities and compositions of commuting zones). When a cell overlays several commuting zones, I attribute emissions based on the surface share of the grid overlaying each commuting zone. Finally I sum emissions over all grid cells within each commuting zone so as to obtain a

[21] An alternative identification assumption is that the common shocks are exogenous (Adão et al., 2019; Borusyak et al., 2021).

panel of sectoral emissions across years and commuting zones. I build such panel for particulate matters PM10 and PM2.5, nitrous gases NOx and ozone precursors CO and COVNM. Emissions are disaggregated across 16 polluting sectors: power industry, oil refineries and transformation industry, combustion for manufacturing, energy for buildings, fuel exploitation, non-metallic minerals production, chemical processes, iron and steel production, non-ferrous metals production, non energy use of fuels, solvents and products use, food and paper, manure management, agriculture (3 distinct activities), waste management and disposal (3 distinct activities) and fossil fuel fires. These are polluting activities defined by the Selected Nomenclature for Air Pollution recommended in the emission register guidebooks implemented by the IPCC. I exclude some polluting activities that my framework does not include (aviation, road transportation and shipping).

Table (1) reports the first stage results. Table (2) displays the outcome of δ and γ estimations. I observe that the OLS coefficients, without and with city and year fixed-effects, underestimate the values of elasticities δ and γ. Results from the preferred strategy, reported in columns (3) to (7) of Figure (2) are all positive and statistically significant (except for the elasticity of environmental damages from emissions of NMVOC, that stands for Non Methanic Volatile Organic Components). Estimated values of δ have the same interpretation in columns (3) to (7), and corresponds to the strength of congestion effects. Estimated values of γ correspond to the elasticity of environmental damage from alternative atmospheric pollutants. The effects of PM2.5 and PM10 on welfare are the stronger compared to other pollutants, and very close to each other. This is in line with evidence on the negative effect of particulate matter on human health. NOx, nitrogen dioxide, has also been found to have a negative effect on health. However, this pollutant is not primarily emitted by industrial activities (rather by the transport sector). This explains the lower value of γ in column (5). Finally, the main health effect from NMVOC and CO is due to their chemical transformation in ozone. The fact that NMVOC and CO do not directly affect health may explain the lower values obtained for γ in columns (6) and (7). To conduct my quantitative exercise, I focus on PM2.5 as the representative pollutant and use coefficients from column (3) in my numerical analysis.

	(1)	(2)	(3)	(4)	(5)	(6)
	$Labor$	$PM2.5$	$PM10$	NOx	$NMVOC$	CO
Instrument	0.181^{***}	0.889^{***}	0.868^{***}	1.211^{***}	0.960^{***}	1.126^{***}
	(0.015)	(0.028)	(0.026)	(0.023)	(0.035)	(0.035)
Observations	5136	4864	4864	4864	4864	4864
City & Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1: First stage results for IV estimation of δ and γ

Note: Column (1) reports the first-stage regression coefficient for local levels of employment (number of workers). Columns (2) to (6) report first-stage regression coefficients for local emissions of a set of pollutants.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
δ	-0.01	0.09	0.91^{**}	0.89^{**}	1.00^{***}	0.88^{**}	0.89^{**}
	(0.02)	(0.05)	(0.35)	(0.35)	(0.38)	(0.34)	(0.35)
γ	0.05^{**}	-0.06	0.49^{***}	0.49^{***}	0.08^{*}	0.08	0.31^{*}
	(0.02)	(0.06)	(0.17)	(0.17)	(0.05)	(0.24)	(0.16)
Observations	4864	4864	4864	4864	4864	4864	4864
City & Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
IV	No	No	Yes	Yes	Yes	Yes	Yes
Pollutant	$PM2.5$	$PM2.5$	$PM2.5$	$PM10$	NOx	$NMVOC$	CO

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Estimation of δ and γ

Note: This table presents estimated coefficients δ and γ from equation (28). Columns (1) reports coefficients estimated through a standard OLS approach, using emissions of PM2.5. Column (2) reports coefficients estimated using city and year fixed-effects, using emissions of PM2.5. Column (3) reports coefficients estimated using city and year fixed-effects and instrumenting local labor and emissions of PM2.5. Columns (4) to (7) report coefficients using emissions of other pollutants instead of PM2.5.

4.3 Estimation of Sector-Specific Parameters

To calibrate parameters $\{\beta_s, \sigma_s, \theta_s\}_{s\in S}$, I closely follows the approach of Shapiro & Walker (2018). First, I compute Cobb-Douglas parameters $\{\beta_s\}_{s\in S}$ using the model's prediction that these parameters are the national shares of revenues of each sectors. I sum firm-level revenues and compute these parameters. Then, elasticities of substitution are recovered using the prediction that, for each sector, the ratio of total payment to labor on value-added is equal to $(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}$. Finally, to estimate the Pareto shape parameters, I use the prediction that the distribution of firm-level sales is Pareto with shape parameter $\theta_s/(\sigma_s-1)$ and I estimate the following equation:

$$\log Pr(x > X_{is}) = \gamma_{0s} + \gamma_{1s}\log(X_{is}) + \epsilon_{is} \tag{31}$$

where X_{ics} represents sales made by firm i from sector s in city c and $\gamma_{1s} = -\frac{\theta_s}{\sigma_s-1}$. Because Pareto distribution better fits the right part of the productivity distribution, I restrict the sample to the upper decile of the firms sample.[22]

I estimate sector-specific emission elasticities $\{\alpha_s\}_{s\in S}$ at the plant level. Assuming that for a plant i, installed in city c, and producing goods from sector s, with a productivity ϕ_{ics}, production follows $q_{ics} = z_{ics}^{\alpha_s} l_{ics}^{1-\alpha_s}\phi_{ics}^{1-\alpha_s}$ and that demand is $q_{ics} = k_{cs}p_{ics}^{-\sigma_s}$, where p_{ics} is the price charged, the relationship between plant-level employment and emissions follows:

$$l_{ics} = \tilde{k}_{cs} z_{ics}^{\frac{\alpha_s(\sigma_s-1)}{1+\alpha_s(\sigma_s-1)}} \phi_{ics}^{\frac{(1-\alpha_s)(\sigma_s-1)}{1+\alpha_s(\sigma_s)}}, \tag{32}$$

where $\tilde{k}_{cs} = \left(\left[(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}\right] w_c^{-\sigma_s} k_{cs}\right)^{\frac{1}{1+\alpha_s(\sigma_s-1)}}$, which, in log, allows us to recover emission

[22]This is an approach developed in Hsieh & Ossa (2016) and Antràs et al. (2017).

elasticities using the following empirical specification:

$$\log l_{icst} = \beta_{0s} + \beta_{1s} \log(z_{icst}) + \epsilon_{icst}. \tag{33}$$

Appendix (A.5.2) provides the algebra corresponding to the above equation. Using a standard OLS strategy to estimate equation (33) would yield biased estimates. Indeed, unobserved plant specific productivity shocks are correlated with the levels of employment and emissions. This is the "transmission bias" identified in the literature on production functions estimations. To circumvent this problem, I instrument plant-level emissions z_{ics} with exogenous fuel-specific energy price variation. Building on Sato et al. (2019), I use a fixed-weight energy price index that measures the plant-specific exposure to variation in fuel prices based on each plant energy consumption across fuel types in the first period when it is observed. I define this instrument as:

$$FEPI_{ist} = \sum_{f \in \Omega_{fuels}} \omega_{f,ist_0} p_{f,st}, \tag{34}$$

where $FEPI_{ist}$ is the plant-specific energy price index built from plant i share of energy expenditures in fuels $f \in \Omega_{fuels}$ (coal, natural gas, electricity, etc.) in period t_0 and $p_{f,st}$ the specific fuel price common to all plants in sector s in period t. An advantage of this emission elasticity estimation strategy is that it only requires to observe plant-level inputs. Conveniently, the EACEI surveys combine plant-level labor and energy consumption across fuels. Note that both the endogenous variable and the instrument are constructed from the EACEI surveys but are not aggregated across fuels using the same weights: the instrument is the sum of consumption weighted by fuel prices and the emissions are the sum of consumption weighted by emission factors that are also fuel specific. I estimate equation (33) for each sector using PM2.5 emissions. Table (4) displays all sector-specific coefficients that I use in my numerical analysis.

To my knowledge, the only other paper that has estimated these parameters is Shapiro & Walker (2018). However, they pool observations from different sectors and only estimate on coefficient. To allow the comparison with their results, Table (3) reports the results that I obtain for different pollutants when industrial sectors are pooled. The first-stage using the fixed-weight energy price index yields significant estimates that are intuitive: on average increasing polluting fuels prices by one percent causes a one percent decrease in atmospheric pollutant emissions. Across pollutants, the second stage yields similar estimates for the pooled sectors emissions intensity, approximately 5%. To my knowledge the only other recent estimates of industrial emission intensities are computed by Shapiro & Walker (2018). They use plant level emission data combined with information on emissions abatement costs from PACE surveys in the United States. Their pooled estimates are lower than mine (1.1% for PM2.5, 1.1% for PM10, .1% for NOx, .08% for CO and .8% for COVNM). The fact that their sample includes a wider range of economic activities that are less pollution intensive may explain this difference.

	(1) log (*emiPM25*)	(2) log (*PM10*)	(3) log (*NOx*)	(4) log (*CO*)	(5) log (*COVNM*)
First Stage:					
$\log FEPI$	-1.00^{***}	-1.00^{***}	-0.98^{***}	-1.07^{***}	-1.06^{***}
	(0.012)	(0.012)	(0.010)	(0.009)	(0.009)
Second Stage:					
$\frac{\alpha(\sigma-1)}{1+\alpha(\sigma-1)}$	0.049^{***}	0.049^{***}	0.051^{***}	0.046^{***}	0.046^{***}
	(0.004)	(0.004)	(0.004	(0.003)	(0.003)
Observations	$223,401$	$223,401$	$223,402$	$223,406$	$223,406$
Year, Region & Industry FE	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Two-stages calibration of emission elasticities α (with pooled industrial sectors)

Finally, I estimate sector-specific elasticities of agglomeration economies $\{\nu_s\}_{s\in S}$ using a regression of city-specific mean firm productivity against city size. To estimate productivity, I regress firm-level log value added on city times sector specific fixed-effects. From equations (12) and (18), the residual from such regression corresponds to firm specific TFP (up to a proportionality factor $(1-\alpha_s)(\sigma_s-1)$). Finally, I compute the average of this residual within each city and regress it on the logarithm of local population to obtain ν.

Sectors	Sales share (β) (1)	Elasticity of substitution (σ) (2)	Pareto shape parameter (θ) (3)	Elasticity of Agglo. Economies (ν) (4)	Pollution elasticity (α) (5)
Automobile & transport	.02	2.27	1.85 (.19)	.05 (.006)	.036 (.001)
Chemicals	.05	3.48	1.33 (.13)	.02 (.005)	.080 (.012)
Communications & Electronics	.01	3.45	2.85 (.16)	.00 (.004)	.025 (.010)
Electrical Equipment	.01	3.91	3.59 (.31)	.00 (.004)	.079 (.010)
Extraction	.01	2.22	1.58 (.08)	.02 (.006)	.126 (.015)
Food, beverages & Tobacco	.05	3.82	2.89 (.02)	-.01 (.001)	.030 (.003)
Machinery & Equipment	.01	3.10	3.88 (.08)	.01 (.002)	.015 (.005)
Metal	.02	3.06	2.29 (.02)	.00 (.002)	.084 (.006)
Rubber & Plastic	.02	2.92	2.66 (.15)	.00 (.002)	.120 (.011)
Textile & Apparel	.01	2.99	2.21 (.04)	.02 (.002)	.089 (.006)
Wood & Paper	.01	2.90	2.42 (.03)	.01 (.002)	.035 (.005)
Other Manufacturing	.02	2.47	1.49 (.01)	.00 (.002)	.035 (.008)
Non manufacturing	.75	2.69	1.48 (.00)	.00 (.000)	.021 (.016)
Pooled (except NM)				.01 (.001)	

Table 4: Estimated parameters

4.4 Trade Costs Matrix

Because I do not observe bilateral trade flows between cities, I cannot estimate a gravity equations to recover bilateral friction terms. Therefore, I follow Yamada (2020) and rely on the methodology used in Baum-Snow et al. (2020) that assumes the following concave relationship between iceberg bilateral trade costs and bilateral travel times:

$$\tau_{ij} = 1 + \rho \times (\text{hours of travel time}_{ij})^{\xi}. \tag{35}$$

I compute travel time using the Open Source Routing Machine (OSRM) API Python client *osrm-py*. The OSRM is a C++ routing engine for shortest paths in road networks building in the road network data of the project OpenStreetMap. Figure (2) displays the average travel time by road across all commuting zone. As expected, central areas are better connected to the rest of the country.

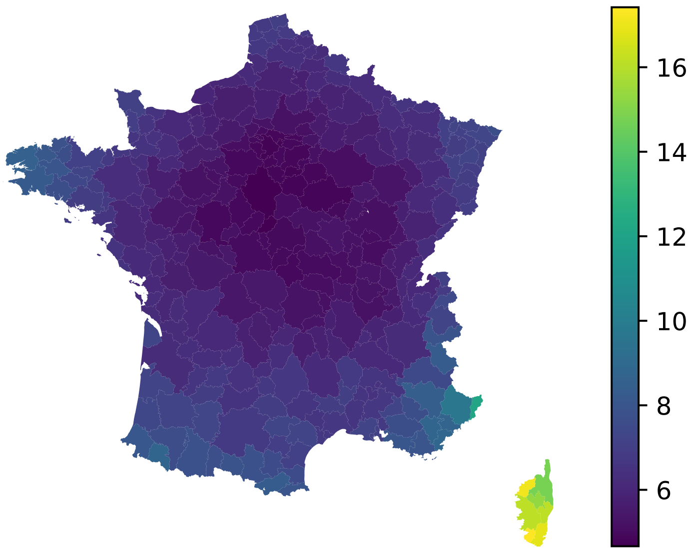


Figure 2: Average travel time across all commuting zone

Note: On this map I plot the average travel time from each commuting zone toward all the other commuting zones. Travel time is in hours. Travel times were computed using the Open Source Routing Machine (OSRM) API Python client *osrm-py*

4.5 Recovering Local Characteristics

Once all parameters are estimated, I numerically solve the non-linear system of equilibrium conditions defined in section 3 by equations (22), (25), and (26). To do so, I use data on the distribution of workers, wages, and industrial PM2.5 emissions across cities in 2012 (see subsection 4.1 for a description of the data). This procedure allows me to retrieve the set of idiosyncratic characteristics – amenities and productivities – as well as the set of emission taxes across cities. Without loss of generality, I set the mean emission tax, productivity and the sum of amenities across cities to be equal to one. In practice, I invert the equilibrium conditions using the Levenberg–Marquardt algorithm as implemented in the *scipy* library in Python 3. This algorithm is an interpolation between the standard Gauss–Newton algorithm and the method of gradient descent that is more robust to the choice of initial values. Results from this step are described in the next section. Appendix A.6.1 provides further details on the sources of data for observed equilibrium distributions of wages, populations and emissions across French cities. Appendix A.6.2 provides descriptive statistics on wages, populations and emissions across French cities. Finally, appendix (A.6.3) provides descriptive statistics on the computed exogenous distributions of amenities, productivities and emission taxes across French cities.

5 Spatially Heterogeneous Emission Regulations

5.1 Evidence of Spatially Heterogeneous Emission Regulations

In Figure (3), I plot the city-level emission taxes obtained through the numerical solution of equilibrium conditions described in subsection 4.5. More precisely, panel (a) plots local emission taxes as a function of population in 2012, panel (b) plots local emission taxes as a function of amenities, and panel (c) plots local emission taxes as a function of productivities. Throughout this section, I plot the logarithm of all variables. This is done for visual clarity, as most distributions are very skewed. I emphasize that these local emission taxes are not actual policy instruments implemented in France. As in Shapiro & Walker (2018), a way to see these values is as follows: if all French emission regulations were to be replaced by local emission taxes, Figure (3) would be the set of taxes that leads to the distribution of PM2.5 emissions that we observe across French cities. Equivalently, these local emission taxes can be seen as measures of how stringent are emission regulations in each city relative the others.

Figure 3 uncovers a large heterogeneity of emission regulations across cities. The positive correlation between local taxes and populations also reveals that existing regulations impose more stringent emissions regulations in larger cities. This observation is supported by anecdotal evidence presented in Appendix A.2. Indeed, several French regulations, that seek to enhance local air quality, impose higher emission standard in larger cities. Is this set of regulations optimally distributed across cities? In 2, the simple model predicted that an optimal set of emission taxes was positively correlated with both amenities and productivities.[23] Panel (c) from Figure (3) displays a strong positive correlation between local levels of productivities and emission taxes that must not be too far from the optimum. However, panel (b) from Figure (3) shows that the current distribution of emission taxes does not dis-

[23]In section 2, the optimal set of emission taxes is positively correlated with both amenities and productivities when $\theta > 0$, which is the case given values computed in section 4.

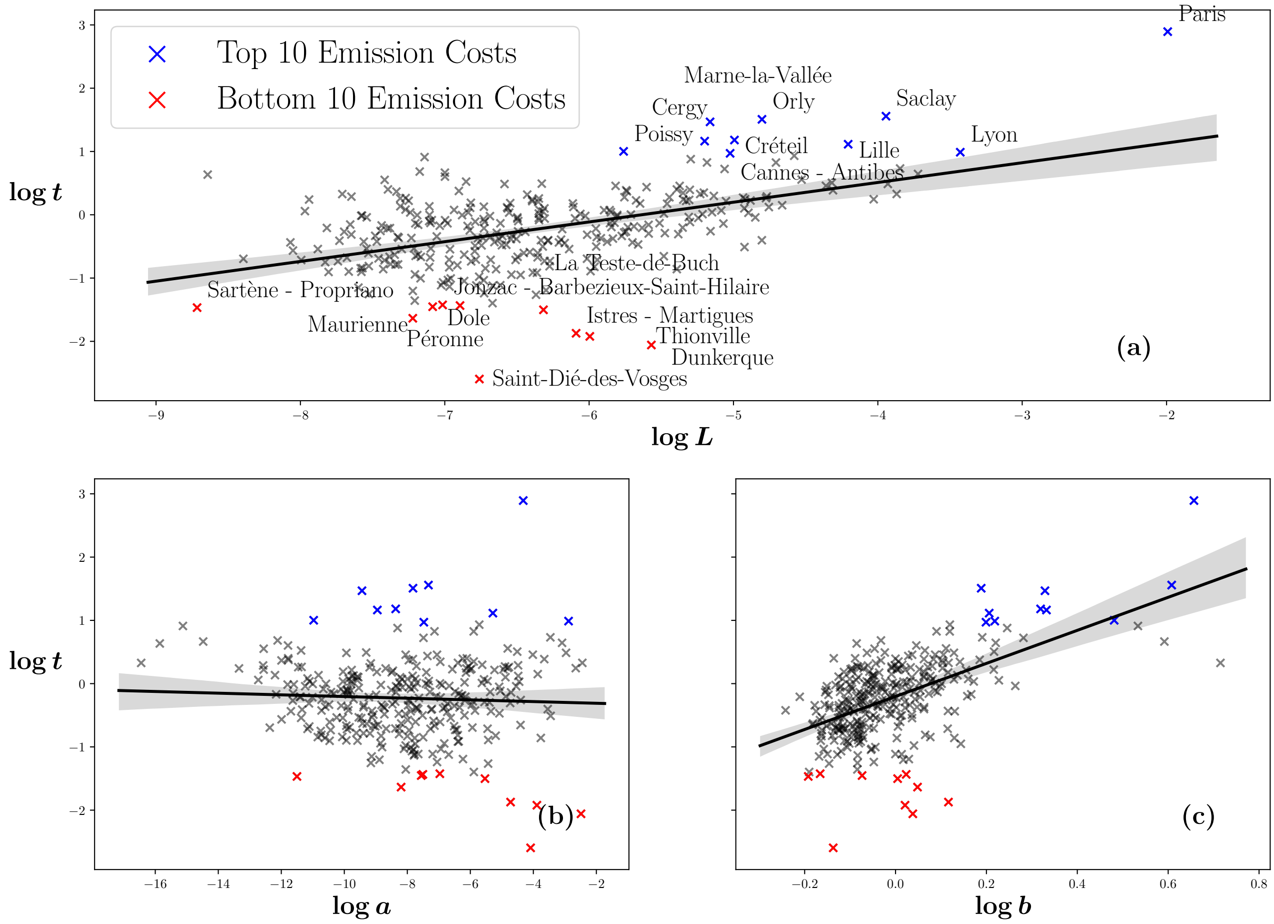


Figure 3: Log-linear relationship between local relative emission taxe and local relative population, productivity and amenity

Note: On this figure I plot the current distribution of emission taxes as a function of current populations on panel (a), of local amenities on panel (b), and of local productivity on panel (c). The distributions of local relative emission taxes, productivities and amenities are estimated by inverting the observed equilibrium using one tradable sector and data for 2012.

play such positive correlation between emission taxes and local amenities. Therefore, the current set of emission taxes across French cities seems to potentially be off from what the model indicates to be the optimal distribution.

Figure (3) also highlights in blue the top ten cities where emission regulations are the most stringent and in red the bottom ten cities where they are the less stringent. As expected, cities with the highest levels of emission taxes are also the most productive and that cities with the lowest levels of emissions costs are the less productive. Yet, among cities with the highest levels of emission taxes, some have very high levels of amenities, whereas others have very low levels of amenities.

5.2 Welfare Impacts of Spatially Heterogeneous Emission Regulations

Results from the previous subsection highlight the spatial heterogeneity of local emission taxes across French cities. In this subsection, I analyze how this heterogeneity affects the distribution of emissions, population and wages. To that end, I compare the observed equilibrium with a counterfactual equilibrium where the level of emission taxes is uniform across all cities. The differences between the two equilibrium can help understanding the consequences of the spatial heterogeneity of emission taxes. Using the values of idiosyncratic amenities and productivities obtained previously, I compute a new equilibrium by solving equilibrium conditions (22), (25) using a uniform emission tax across cities. This uniform emission tax value is the mean of all emission taxes obtained in the previous computation, which has been normalized to one without loss of generality throughout the analysis. The counterfactual equilibrium is solved numerically using the Levenberg–Marquardt algorithm. Figure (4) displays the results from this exercise. For each endogenous variable of interest in the model, I plot a map showing the variation in level of moving from a uniform emission cost to the current distribution of emission taxes that is heterogeneous across cities.

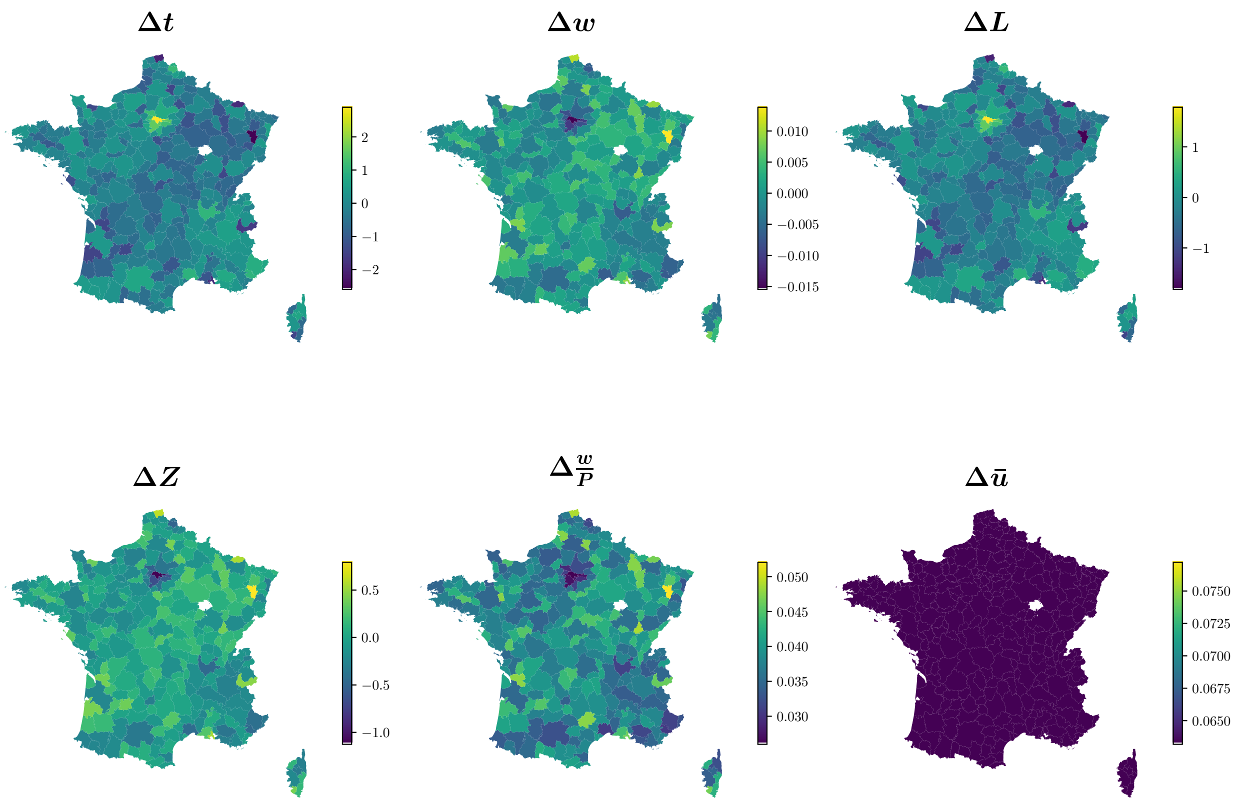


Figure 4: Spatial reallocations and welfare effect of imposing spatially heterogeneous emission taxes

Note: On this figure I map the changes in emission taxes, wages, industrial workers' populations, emissions, real wages, and welfare du to a move from a spatially uniform emission tax to the current set of emission tax.

Because emission taxes are actually higher in larger cities, such as Paris, Lyon, Bordeaux, Toulouse, the map of changes in t reveals an increase in these cities, and a decrease in less populated cities. In turn, the map of ΔZ reveals a decrease in the largest cities, and an increase in the smallest cities. Hence, when the central planner implements more stringent air quality policies in largest cities relative to smaller cities, emissions are relocated from the most populated areas to the least populated areas. Figure (4) also plots the changes in the level of population (ΔL) to illustrate the reallocation of population across cities due to the move from uniform to heterogeneous stringency. Given the free migration assumption, workers react to changes in local air quality (and prices which are affected by the change in local emission taxes) and move away from more polluted areas. The panel on ΔL show that there is a reallocation of workers from small cities (where emissions have increased) to larger cities where emissions have decreased. As larger cities become cleaner, more people have an incentive to live there.

Figure (5) plots the log-difference in local populations between the current equilibrium and the counterfactual one with uniform emission taxes, as a function of initial population (in panel (a)), of amenities (in panel (b)), and of productivities (in panel (c)). For a given city, a positive number indicates that the local population is higher under the current distribution of emission taxes compared to the counterfactual. Panel (a) reveals that the largest cities have a larger population because of the spatial heterogeneity in regulations. Panels (b) and (c) show that the top ten cities for which population increases the most are also highly productive cities but not cities with particularly high amenities. Figure (6) is similar to Figure (5), except that it shows changed in emissions instead of populations. Panel (a) shows that the largest cities emit less pollution because of the spatial heterogeneity in regulations. Panels (b) and (c) illustrate the fact that more stringent regulations in more productive cities leads to lower levels of emissions there, compared to the counterfactual. There are no clear correlation with the distribution of local amenities. Panels (a) from Figures (5) and (6) suggest that more stringent regulations in larger cities allowing them to be cleaner (less emissions) and in turn larger (higher population) than what they would be under a uniform regulation.

5.3 Optimal Mean-Preserving Set of Heterogeneous Regulations

The model can be used to assess which set of emission taxes across cities maximizes welfare. To do so, I numerically solve the $\bar{u}$ maximization problem under the constraint that the mean emission tax across cities remains equal to one. This optimization exercise yields a new spatial equilibrium that can be compared to the current equilibrium. With this numerical solution, I present normative and empirical results on the gap between the current distribution of relative stringencies across French cities and the distribution that would maximize workers' welfare. In practice, I identify for each city an optimal level of regulation that I compare to the current level of regulation.

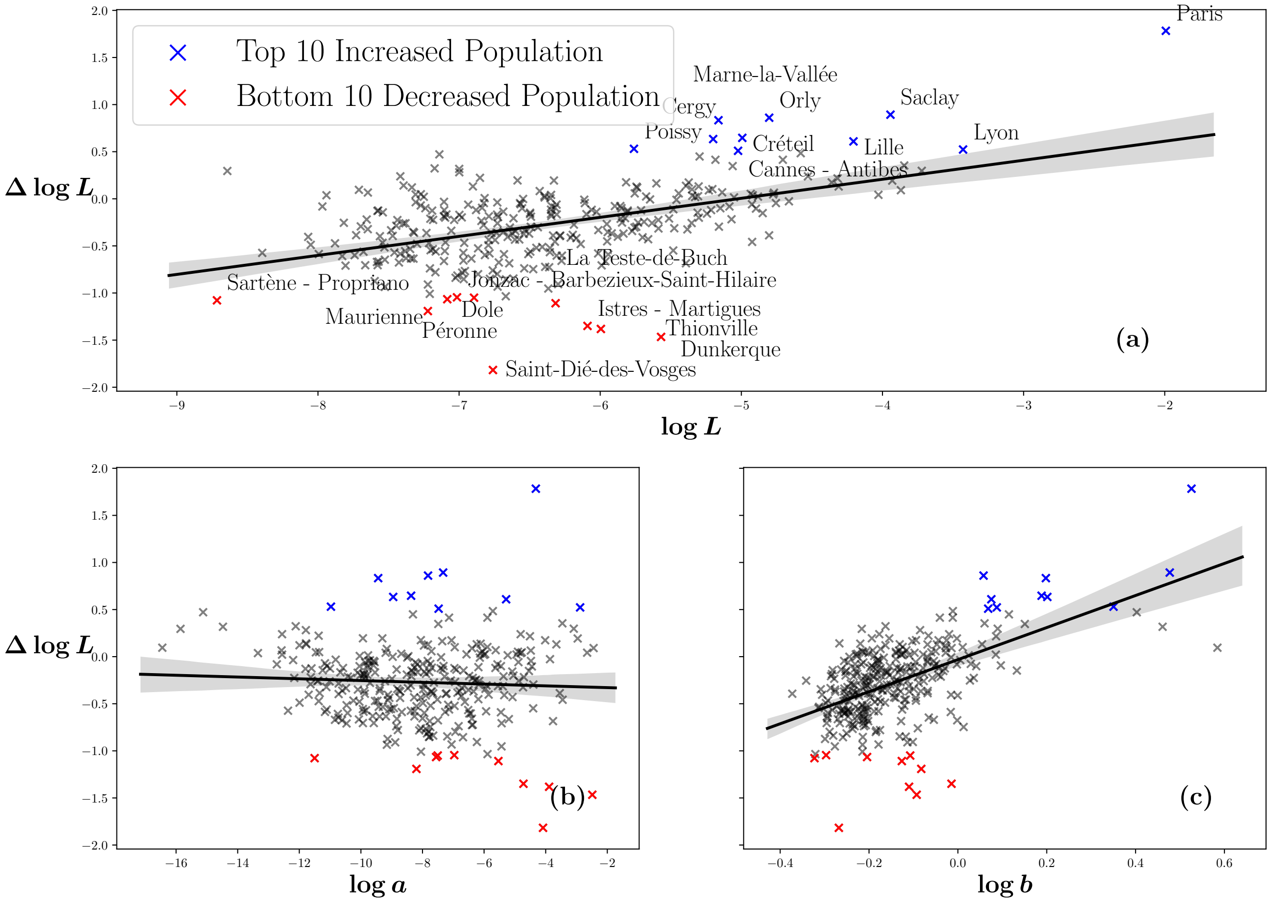


Figure 5: Effects of spatial heterogeneity emission stringency on local populations: under the current policy, large cities are larger than what they would be if they faced a uniform emission tax

Note: On this figure I plot the relative change in industrial workers' population at the commuting zone level as a result of a move from a uniform to the current distribution of emission taxes. This change is plotted as a function of current populations on panel (a), of local amenities on panel (b), and of local productivity on panel (c).

Figure (7) displays the results from the maximization exercise. It plots the optimal distribution of the logarithm of emission taxes as a function of the log of: current city population (panel (a)), amenities (panel (b)), and productivities (panel (c)). The current distributions from Figure (3) are depicted in black and the optimal distributions are depicted in red. Figure (7) reveals that increasing the relative emission regulations in larger cities would achieve welfare gains. Figure (7) shows that the distributions in red display less variation on the left part than on the right part. This is because set of emission taxes results from a constrained numerical optimization algorithm: small cities are so much smaller than large cities, the constraint to keep the mean emission tax fixed implies that optimal emission taxes are put to the minimum level possible in small cities. These points are corner solutions. However, because distributions of cities' characteristics (amenities and productivities) are so skewed, only the largest cities matter for aggregate welfare.

As discussed in section 2 and illustrated in panel (b) from Figure (7), increasing the relative stringency of regulations in cities with higher amenities would welfare-improving. This is because, currently, cities with higher amenities are generally too small because work-

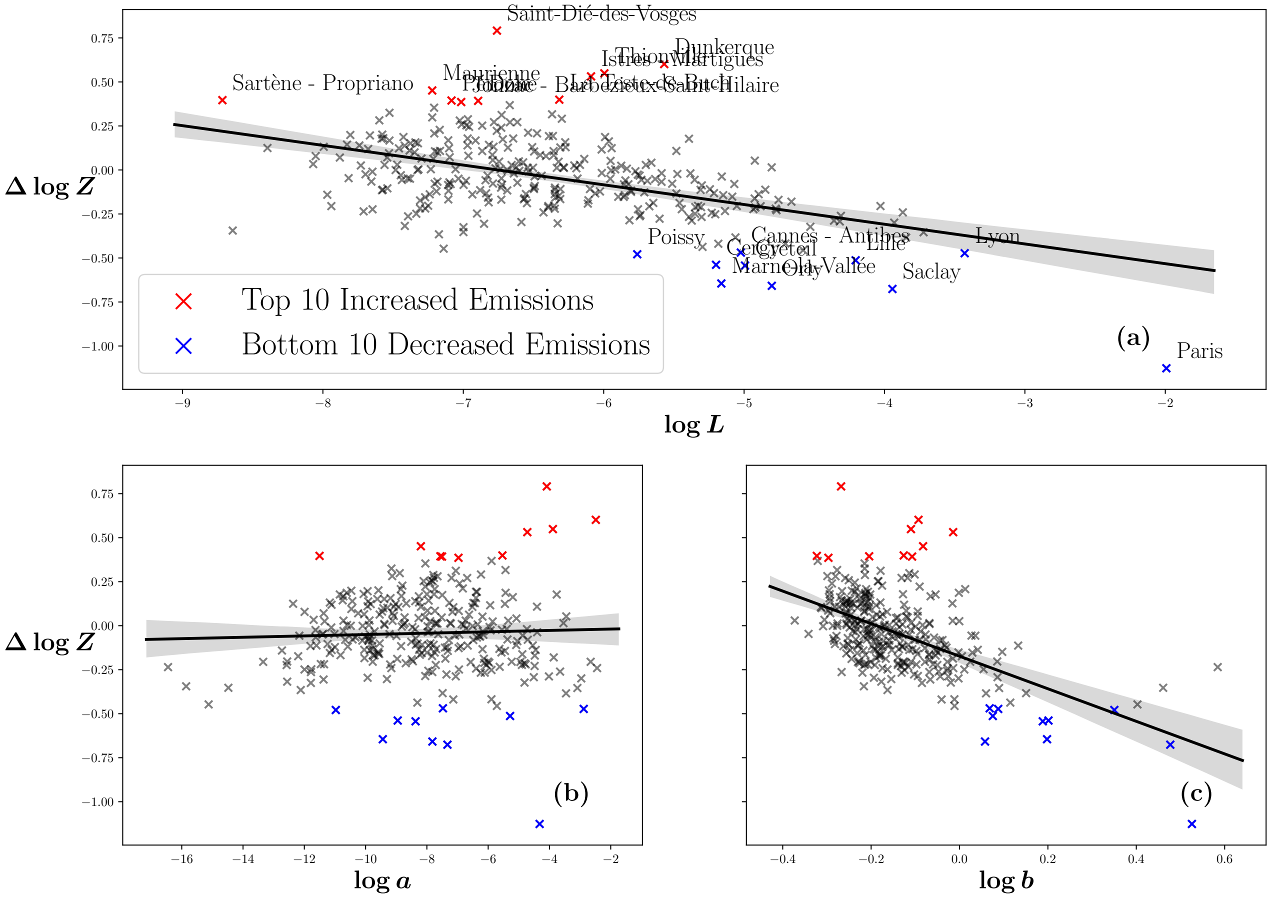


Figure 6: Effects of spatial heterogeneity emission stringency on local emissions: under the current policy, large cities emit less pollution than what they would if they faced a uniform emission tax

Note: On this figure I plot the relative change in industrial emissions at the commuting zone level as a result of a move from a uniform to the current distribution of emission taxes. This change is plotted as a function of current emissions on panel (a), of local amenities on panel (b), and of local productivity on panel (c).

ers and firms do not internalize the impact of their location choices on local air quality. Panel (a) shows the reallocation necessary to move from the current distribution of relative emission taxes to the optimal one. These taxes need to be raised in the currently largest cities, and decreased in the smallest cities. Panel (c) from Figure (7) indicates that there is already a strong positive correlation between the spatial distribution of emission taxes and the distribution of idiosyncratic productivities. Solving the optimization problem confirms that such correlation is optimal. Indeed, panel (c) shows that the coefficient of the log-linear relationship between the local idiosyncratic levels of productivities and the relative emission taxes is currently not too far from the optimal one.

5.3.1 Reallocating Population Across Space

Figure (8) displays the population reallocation effect that would follow from the adoption of the optimal distribution of relative emission taxes. In particular, it highlights in blue the cities that would see their local population increase. Only a small subset of cities (around

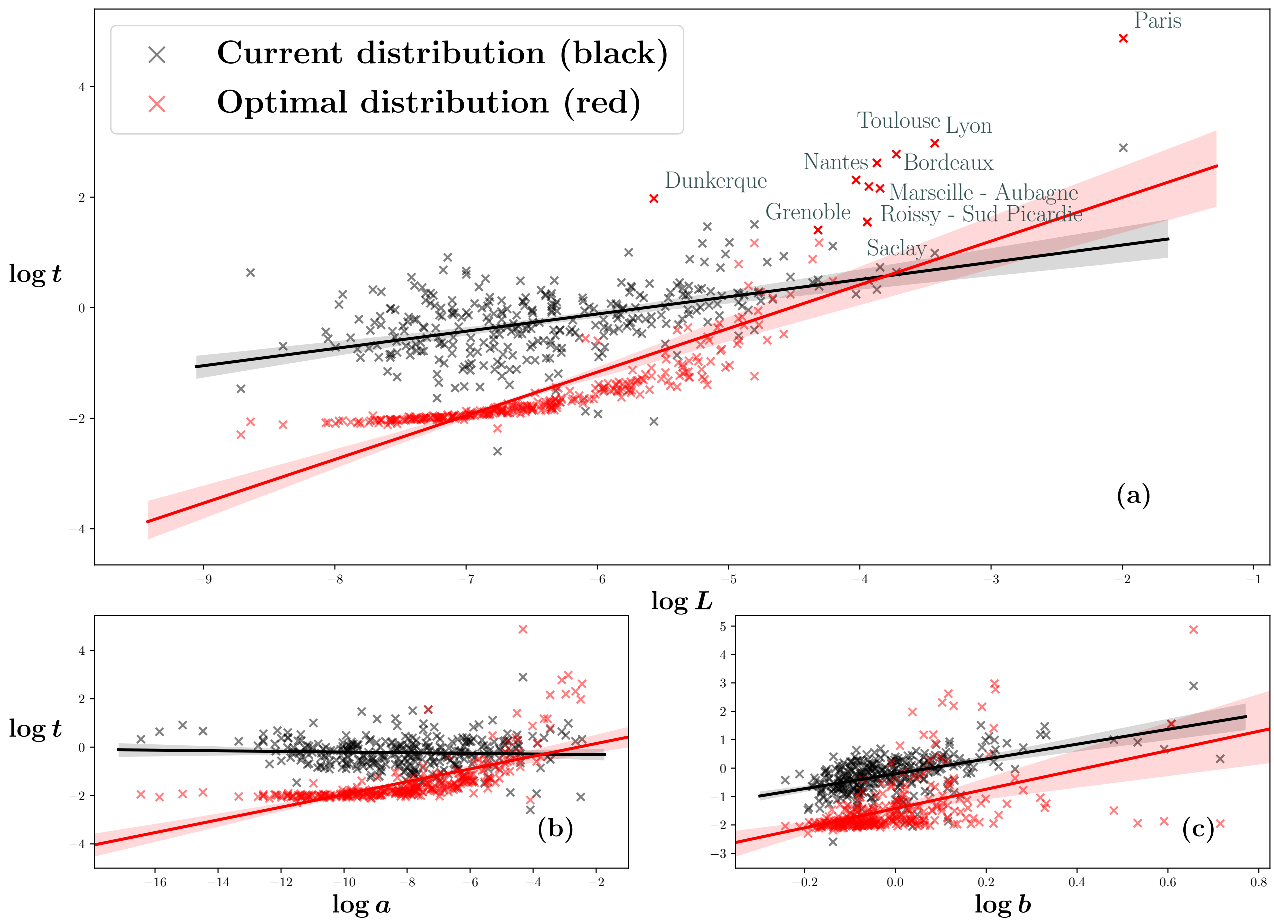


Figure 7: Comparison between current distribution of relative emission taxes (black) and the mean-preserving optimal distribution (red): emission taxes are lower in smaller cities and increased in larger cities

Note: On this figure I plot the current distribution of emission taxes in black and the optimal distribution of emission taxes as functions of current populations on panel (a), of local amenities on panel (b), and of local productivity on panel (c). The current distribution is depicted in black and the optimal distribution is depicted in red.

15 of them, out of 300 cities) would grow larger and the other cities would become smaller. Specifically, panel (a) shows that population concentration is reinforced: the largest cities grow even larger. Panel (b) indicates that this reallocation mainly comes from a reallocation of workers toward cities with higher amenities.

Figure (9) reveals the concentration effect of adopting the optimal set of emission taxes. Ranking cities according to their current populations, the cumulative population starting from the smallest city is depicted by the black line. The red line depicts the same cumulative sum using what local populations would be under the optimal set of emission taxes. Since the red line would be more skewed to the right than the black line, workers become more concentrated in the largest cities under the optimal distribution of emissions costs than under the current set of policies.

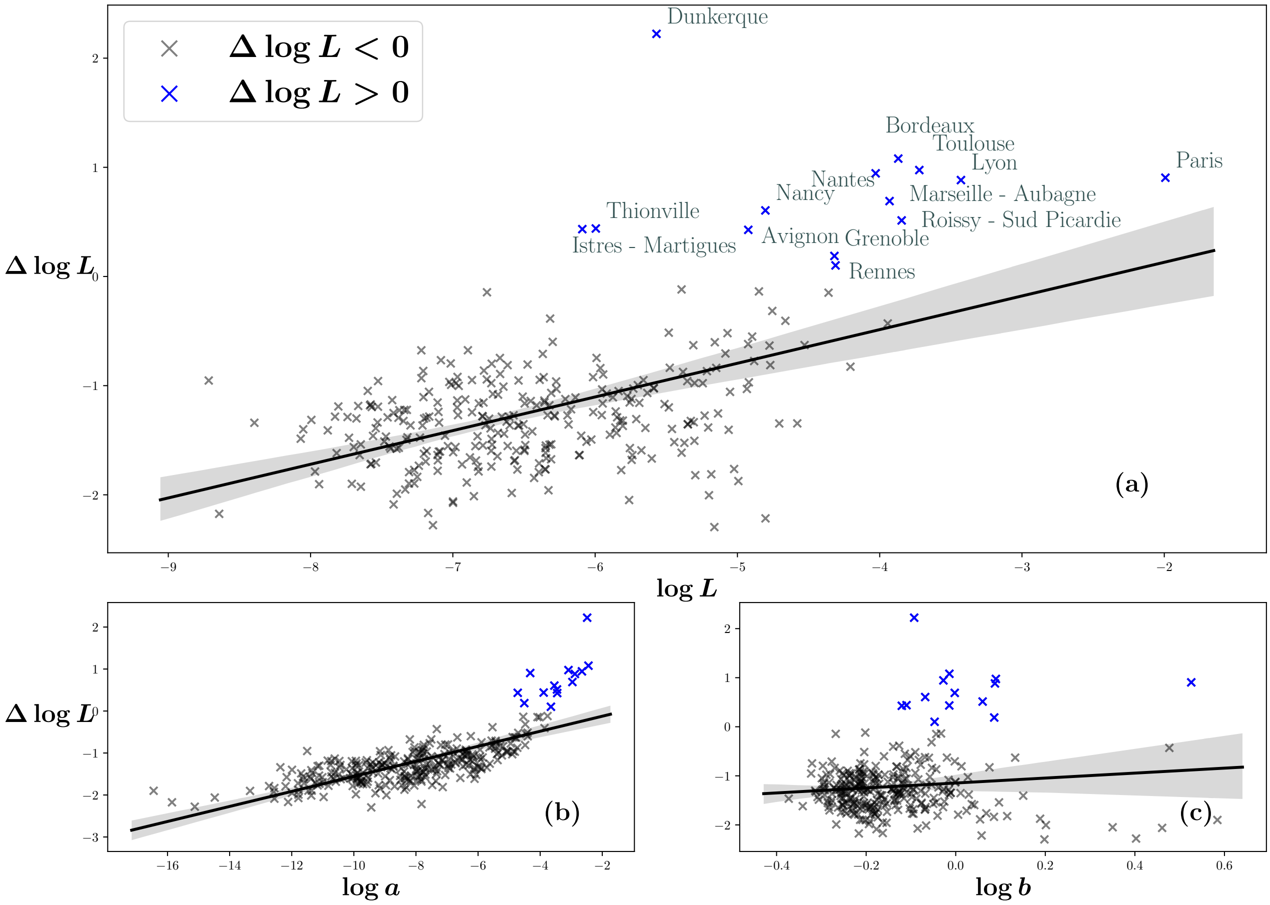


Figure 8: Population reallocation when moving from the current to the optimal distribution of relative emission taxes

Note: On this figure I plot the relative change in industrial workers' population at the commuting zone level as a result of a move from the current to the optimal distribution of emission taxes. This change is plotted as a function of current populations on panel (a), of local amenities on panel (b), and of local productivity on panel (c). In blue are the cities that would see their population increase as a result of this emission taxes distribution shift.

5.3.2 Reallocating Emissions Across Space

Figure 10 displays the pollution reallocation effect of moving to the optimal set of emission taxes. In particular, it highlights in green the cities that would see their local emissions decrease. In blue are the cities for which emission would decrease and population would decrease. As expected, it is optimal to decrease emissions in the currently more polluted cities (panel (a)), corresponding to a reallocation of emissions from cities with high amenities to cities with low amenities (panel (b)).

Figure (11) compares the cumulative distributions of emissions under the current (black line) and the optimal (red line) distributions of emission taxes across cities. Since the red line is less skewed to the right than the black line, the optimal reallocation of emissions across cities reduces spatial concentration of emissions.

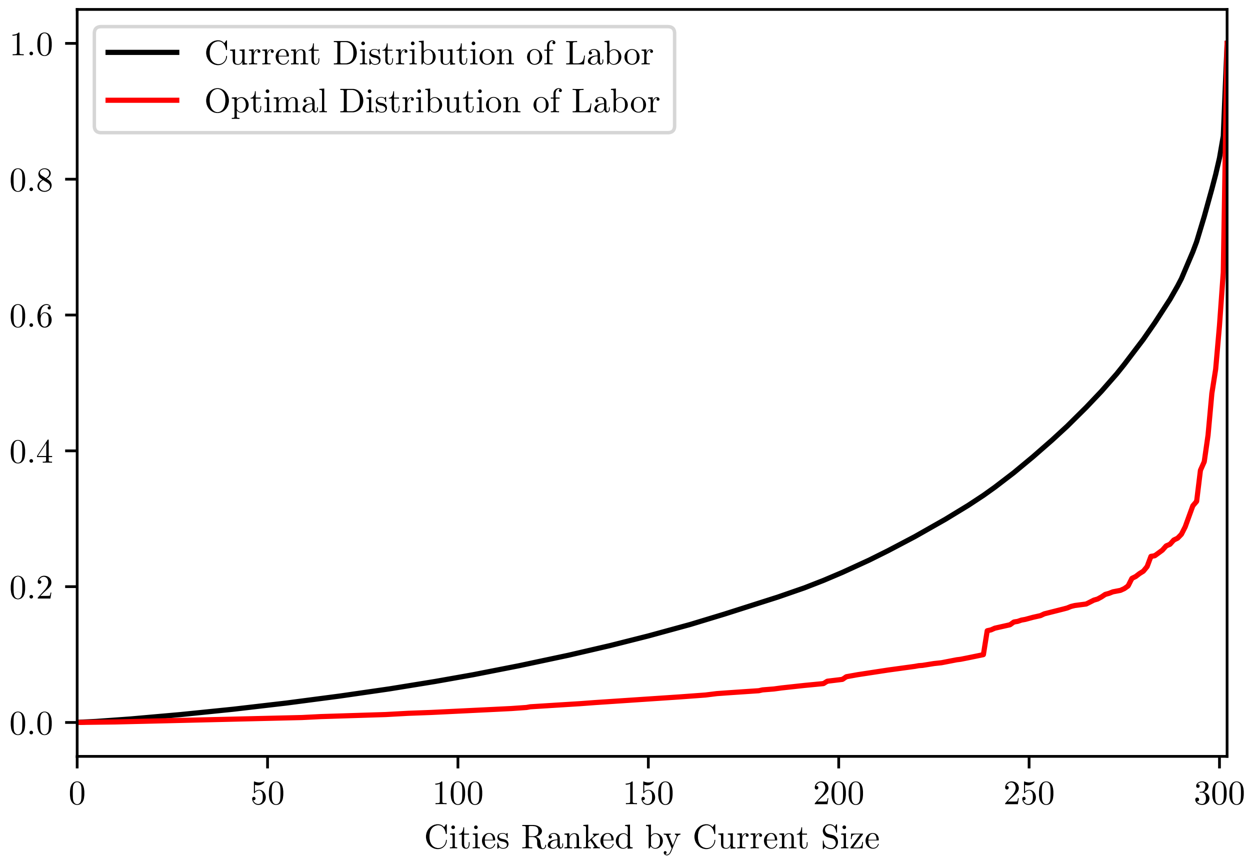


Figure 9: Cumulative distributions of population under the current and the optimal distributions of emission taxes

Note: On this graph, I plot the cumulative sum of industrial workers' population across French commuting zones as a function of their respective ranks by industrial workers' population. I depicted this distribution in black for the current distribution and in red for the optimal distribution. A 45 degree line who mean that emissions are evenly spread across cities. A curve skewed to the right indicates spatial concentration of workers in a few places.

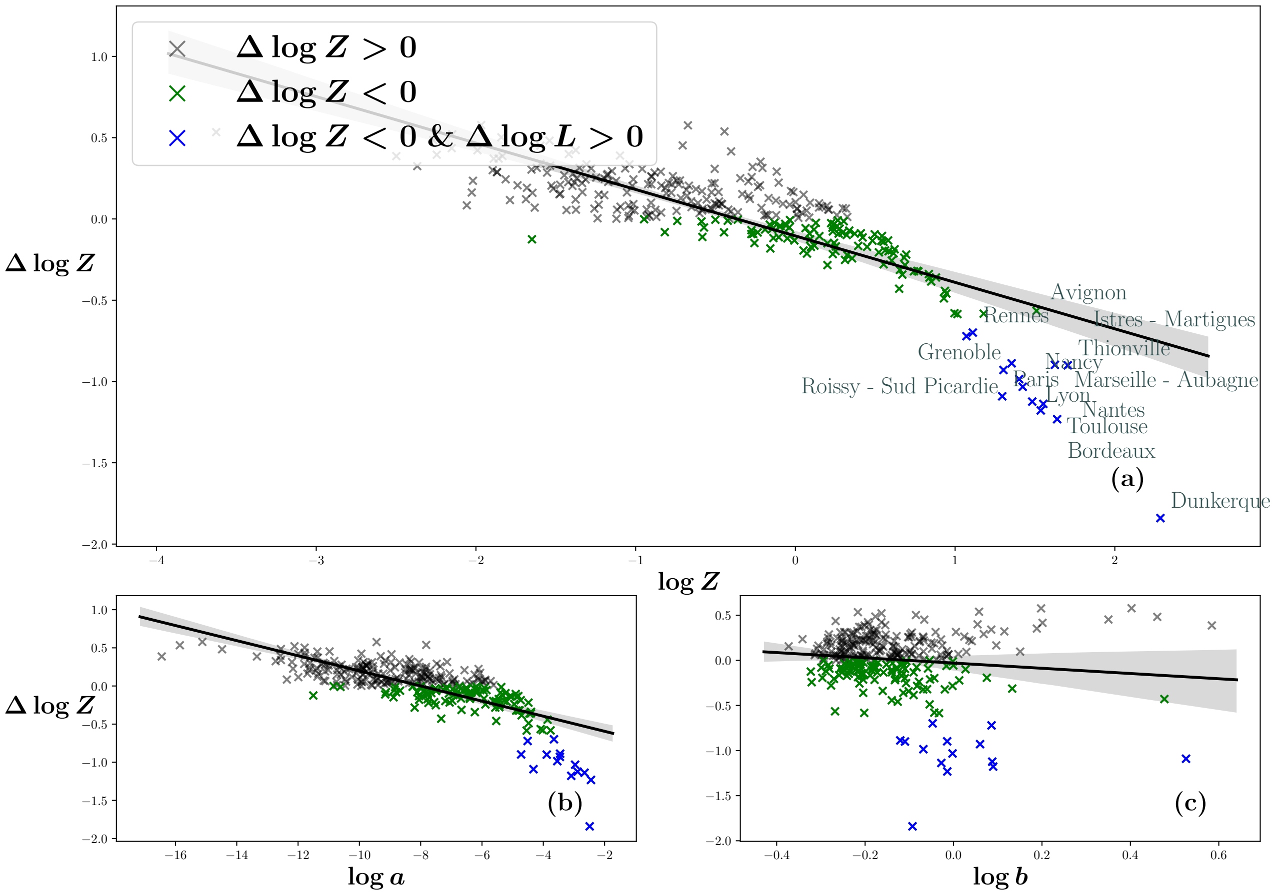


Figure 10: Emissions reallocation when moving from the current to the optimal distribution of relative emission taxes

Note: On this figure I plot the relative change in emissions at the commuting zone level as a result of a move from the current to the optimal distribution of emission taxes. This change is plotted as a function of current emissions on panel (a), of local amenities on panel (b), and of local productivity on panel (c). In blue are the cities that would see their population increase as a result of this emission taxes distribution shift. In green are the cities that would see their emissions decrease.

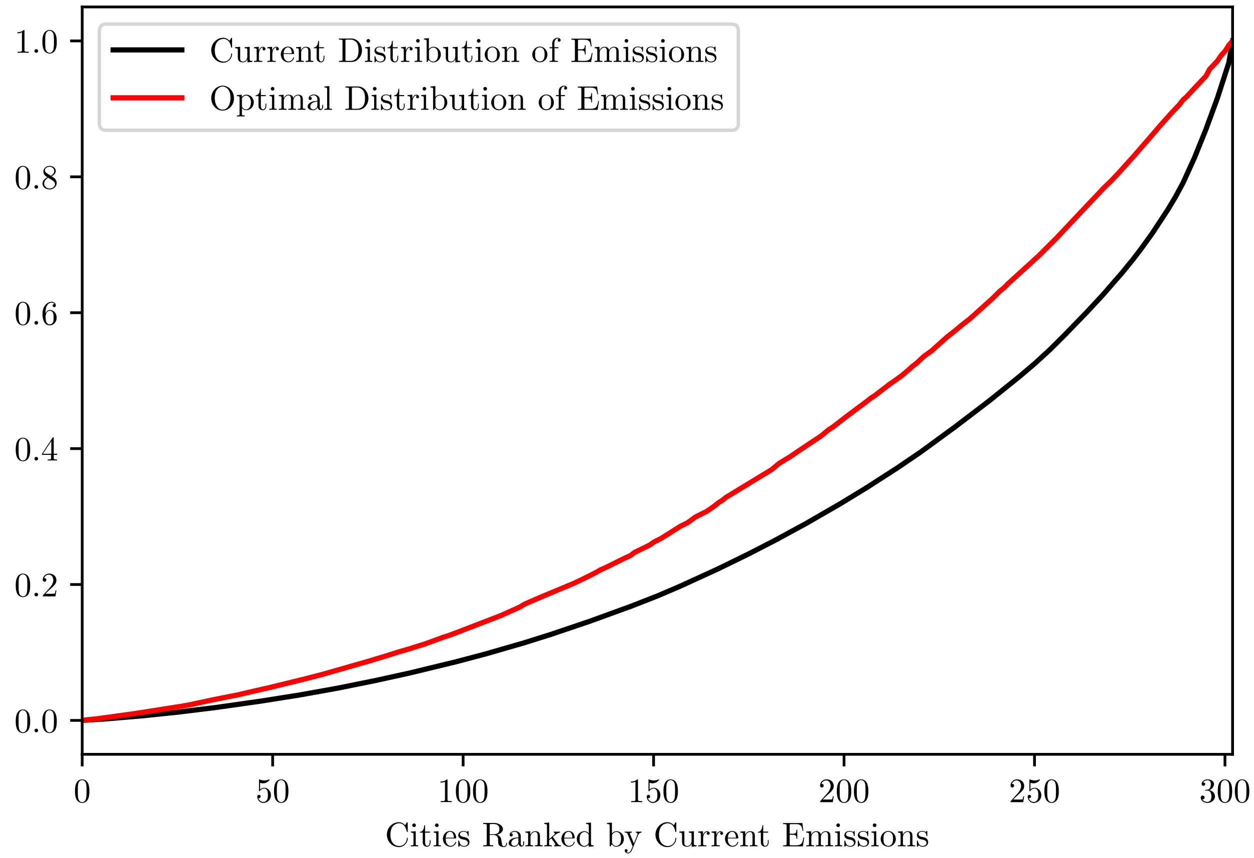


Figure 11: Cumulative distributions of pollution emissions under the current and the optimal distributions of emission taxes

Note: On this graph, I plot the cumulative sum of industrial emissions across French commuting zones as a function of their respective ranks by quantity of emission. I depicted this distribution in black for the current distribution and in red for the optimal distribution. A 45 degree line who mean that emissions are evenly spread across cities. A curve skewed to the right indicates spatial concentration of emissions in a few places.

6 Conclusion

In this paper, I analyze how the spatial distribution of local air quality policies can improve welfare. I incorporate in a simple spatial model endogenous industrial emissions of atmospheric pollutants. As pollution has a negative effect on welfare and workers move away from polluted areas, I show that the central planner should adopt higher emission taxes in cities that have higher levels of amenities and productivity. I build a general version of the model that includes heterogeneous industrial sectors and I estimate the model's parameters using an extensive set of data on French firms and cities. Then, I retrieve the model's primitives – local amenities, productivities, and emission taxes – from data on the distribution of workers, wages, and industrial PM2.5 emissions across cities. In the French context, I show that current regulations correspond to a set of emission taxes that are higher in the largest cities. Finally, I numerically solve the central planner's welfare maximization problem. I find that further increasing emission taxes in the largest French cities could achieve welfare gains.

This paper sheds light on a spatial tradeoff between clean air and productivity. My results indicate that policy makers should not implement environmental policies uniformly across locations. Rather, they should be more stringent in large cities. This has important implications in terms of environmental regulations design. The optimal set of relative emission taxes that I identify in this paper is independent of the average level of the tax. This means that it is applicable to any new air quality regulation: for any given objective of pollution reduction, local regulatory burdens should follow the distribution described in this paper. Future work should incorporate additional factors such as inequalities in exposure to atmospheric pollution (Agyeman et al., 2016; Banzhaf et al., 2019; Colmer et al., 2020). Extending this work to include distributional effects of air quality regulations and their impacts on inequalities in exposure to pollution (Currie et al., 2020; Shapiro & Walker, 2021) could further help investigate optimal place-based environmental policies.

References

Adão, R., Kolesár, M., & Morales, E. (2019). Shift-Share Designs: Theory and Inference. *The Quarterly Journal of Economics*, 134(4), 1949–2010.

Agyeman, J., Schlosberg, D., Craven, L., & Matthews, C. (2016). Trends and Directions in Environmental Justice: From Inequity to Everyday Life, Community, and Just Sustainabilities. *Annual Review of Environment and Resources*, 41(1), 321–340.

Albouy, D., Behrens, K., Robert-Nicoud, F., & Seegert, N. (2019). The Optimal Distribution of Population across Cities. *Journal of Urban Economics*, 110, 102–113.

Aldeco, L., Barrage, L., & Turner, M. A. (2019). Equilibrium Particulate Exposure. *IGC Working Paper C-89451-NOC-1*.

Allen, T. & Arkolakis, C. (2014). Trade and the Topography of the Spatial Economy. *Quaterly Journal of Economics*, 129(3), 1085–1140.

Allen, T., Arkolakis, C., & Takahashi, Y. (2020). Universal Gravity. *Journal of Political Economy*, 128(2), 393–433.

Almagro, M. & Domínguez-Iino, T. (2021). Location Sorting and Endogenous Amenities: Evidence from Amsterdam. *mimeo*.

Alvarez, J. L. C. & Rossi-Hansberg, E. (2021). The Economic Geography of Global Warming. Working Paper 28466, National Bureau of Economic Research.

Antràs, P., Fort, T. C., & Tintelnot, F. (2017). The Margins of Global Sourcing: Theory and Evidence from US Firms. *American Economic Review*, 107(9), 2514–2564.

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6), 2121–2168.

Baldwin, R. E. & Okubo, T. (2005). Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography*, 6(3), 323–346.

Banzhaf, S., Ma, L., & Timmins, C. (2019). Environmental Justice: The Economics of Race, Place, and Pollution. *Journal of Economic Perspectives*, 33(1), 185–208.

Barrows, G. & Ollivier, H. (2021). Foreign demand, developing country exports, and CO2 emissions: Firm-level evidence from India. *Journal of Development Economics*, 149, 102587.

Bartelme, D. (2018). Trade Costs and Economic Geography: Evidence from the U.S. *mimeo*.

Bartik, T. J. (1991). Who Benefits from State and Local Economic Development Policies? *Upjohn Institute for Employment Research*.

Baum-Snow, N., Henderson, J. V., Turner, M. A., Zhang, Q., & Brandt, L. (2020). Does investment in national highways help or hurt hinterland city growth? *Journal of Urban Economics*, 115, 103124.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2007). Comparative Advantage and Heterogeneous Firms. *Review of Economic Studies*, 74(1), 31–66.

Bombardini, M. & Li, B. (2020). Trade, pollution and mortality in China. *Journal of International Economics*, 125, 103321.

Borck, R. & Schrauth, P. (2021). Population density and urban air quality. *Regional Science and Urban Economics*, 86, 103596.

Borck, R. & Tabuchi, T. (2018). Pollution and city size: can cities be too small? *Journal of Economic Geography*, 19(5), 995–1020.

Borusyak, K., Hull, P., & Jaravel, X. (2021). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies*.

Bougon, P. & Lavergne, R. (2019). Des outils réglementaires et fiscaux pour réduire les émissions de polluants atmosphériques de l'industrie. *Annales des Mines - Responsabilité et environnement*, N 96(4), 81.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1), 22–64.

Carozzi, F. & Roth, S. (2019). Dirty Density: Air Quality and the Density of American Cities . *CEP Discussion Papers. Centre for Economic Performance, LSE,London, UK.*

CITEPA (2020). Rapport OMINEA – 17ème édition. *https://www.citepa.org/fr/ominea/*.

Colmer, J., Hardman, I., Shimshack, J., & Voorheis, J. (2020). Disparities in PM 2.5 air pollution in the United States. *Science*, 369(6503), 575–578.

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723–742.

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

Combes, P.-P. & Gobillon, L. (2015). Chapter 5 - The Empirics of Agglomeration Economies. volume 5 of *Handbook of Regional and Urban Economics* (pp. 247–348). Elsevier.

Copeland, B. R. & Taylor, M. S. (2004). Trade, Growth, and the Environment. *Journal of Economic Literature*.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V., & Janssens-Maenhout, G. (2018). Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. *Earth System Science Data*, 10(4), 1987–2013.

Currie, J., Voorheis, J., & Walker, R. (2020). What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality. *Working Paper CES-20-02*.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction. *American Economic Review*, 109(12), 4178–4219.

Deryugina, T., Miller, N., Molitor, D., & Reif, J. (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.

Desmet, K., Kopp, R. E., Kulp, S. A., Nagy, D. K., Oppenheimer, M., Rossi-Hansberg, E., & Strauss, B. H. (2021). Evaluating the Economic Cost of Coastal Flooding. *American Economic Journal: Macroeconomics*, 13(2), 444–486.

DGEC (2016). Panorama Energie-Climat. *DICOM-DGEC/CHE/16188*.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review*, 106(3), 479–524.

Drut, M. & Mahieux, A. (2015). Correcting agglomeration economies: How air pollution matters. *HAL Working Papers hal-01007019*.

Duranton, G. & Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. volume 4 of *Cities and Geography* (pp. 2063–2117). Elsevier.

Eeckhout, J. & Guner, N. (2015). Optimal Spatial Taxation: Are Big Cities too Small? *mimeo*.

Eeckhout, J. & Hedtrich, C. (2021). Green Urbanization. *mimeo*.

Fajgelbaum, P. D. & Gaubert, C. (2020). Optimal Spatial Policies, Geography, and Sorting. *The Quarterly Journal of Economics*, 135(2), 959–1036.

Gaubert, C. (2018). Firm Sorting and Agglomeration. *American Economic Review*, 108(11), 3117–3153.

Glaeser, E. L. & Kahn, M. E. (2010). The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics*, 67(3), 404–418.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8), 2586–2624.

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

Graff Zivin, J. & Neidell, M. (2013). Environment, Health, and Human Capital. *Journal of Economic Literature*, 51(3), 689–730.

Hanlon, W. W. (2019). Coal Smoke, City Growth, and the Costs of the Industrial Revolution. *The Economic Journal*, 130(626), 462–488.

Hauvuy, E. & Riedinger, N. (2005). Une estimation du coût d'abattement de la pollution atmosphérique pour les entreprises françaises. *Économie & prévision*, 168(2), 63.

Helpman, E. (1998). The Size of Regions. *Topics in Public Economics. Theoretical andApplied Analysis*, (pp. 33–54).

Henderson, J. V. (1974). The Sizes and Types of Cities. *The American Economic Review*.

Hsieh, C.-T. & Ossa, R. (2016). A global view of productivity growth in China. *Journal of International Economics*, 102, 209–224.

IGF (2018). Évaluation de l'impact environnemental et économique de la Taxe générale sur les activités polluantes (TGAP) sur les émissions de polluants atmosphériques. *Rapport IGF n 2018-M-027-03*.

INS (2020). Inventaire national spatialisé des émissions de polluants dans l'air. *http://emissions-air.developpement-durable.gouv.fr*.

Kyriakopoulou, E. & Xepapadeas, A. (2013). Environmental policy, first nature advantage and the emergence of economic clusters. *Regional Science and Urban Economics*, 43(1), 101–116.

Lange, A. & Quaas, M. F. (2007). Economic Geography and the Effect of Environmental Pollution on Agglomeration. *The B.E. Journal of Economic Analysis & Policy*, 7(1).

Leturque, E. & Sanch-Maritan, M. (2019). Do agglomeration economies are lower for polluting sectors? *HAL Working Papers hal-02347595*.

Millock, K. & Nauges, C. (2006). Ex Post Evaluation of an Earmarked Tax on Air Pollution. *Land Economics*, 82(1), 68–84.

Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics*, 5(1), 65–103.

Nunn, N. & Qian, N. (2014). US Food Aid and Civil Conflict. *American Economic Review*, 104(6), 1630–1666.

Pflüger, M. (2021). City Size, Pollution and Emission Policies. *Journal of Urban Economics*, 126, 103391.

Redding, S. (2020). Trade and Geography. *NBER Working Paper 27821*.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics*, 101, 148–167.

Redding, S. J. & Rossi-Hansberg, E. (2017). Quantitative Spatial Economics. *Annual Review of Economics*, 9(1), 21–58.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6), 1257–1278.

Rosen, S. (1979). *Wage-Based Indexes of Urban Quality of Life*, (pp. 74–104). Johns Hopkins University Press.

Sato, M., Singer, G., Dussaux, D., & Lovo, S. (2019). International and sectoral variation in industrial energy prices 1995–2015. *Energy Economics*, 78, 235–258.

Shapiro, J. S. & Walker, R. (2018). Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade. *American Economic Review*, 108(12), 3814–3854.

Shapiro, J. S. & Walker, R. (2021). Where Is Pollution Moving? Environmental Markets and Environmental Justice. *AEA Papers and Proceedings*, 111, 410–414.

Suarez-Serrato, J. C. & Zidar, O. (2016). Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms. *American Economic Review*, 106(9), 2582–2624.

Yamada, E. (2020). A Spatial Equilibrium Analysis of Air Pollution in China. *JICA-RI Working Paper*.

A Appendices

A.1 Descriptive Statistics on Pollution from the Industry

Over the last decades, several regulations have been implemented in France to reduce atmospheric pollution. As a result, national levels of emissions have decreased in the past. Figure (12), constructed from the European Emissions Database for Global Atmospheric Research (EDGAR), shows this decrease for particulate matters 10 and 2.5 as well as for sulfur dioxide. However, observed levels of pollution in France are still well above World Health Organization (WHO) guidelines in several areas over the country. Using data on air concentration in PM10, PM2.5 and SO2 from ground monitors installed in diverse locations, I illustrate this situation on Figure (13) and show that a substantial fraction of monitors measures concentrations above the 24-hour and annual guidelines from the WHO guidelines. There are many sources for this pollution. Figure (12) also illustrates the distribution of the pollution sources. One can observe that for the three pollutants highlighted, PM10, PM2.5 and SO2, industrial activities represent a large fraction of the emissions. Tables (5), (6) and (7) present the respective contributions of the main polluting activities: in 2012, industrial activities contributed to respectively 23%, 40% and 42% of PM10, PM2.5 and SO2 national emissions.

% Share from:	1990	2000	2012
Industry	19.4	21.4	23.0
Energy	6.7	4.8	3.8
Transport	12.2	15.4	9.4
Residential	40.1	34.6	39.1
Agriculture	17.2	22.5	24.6
Waste	4.4	1.3	0.1

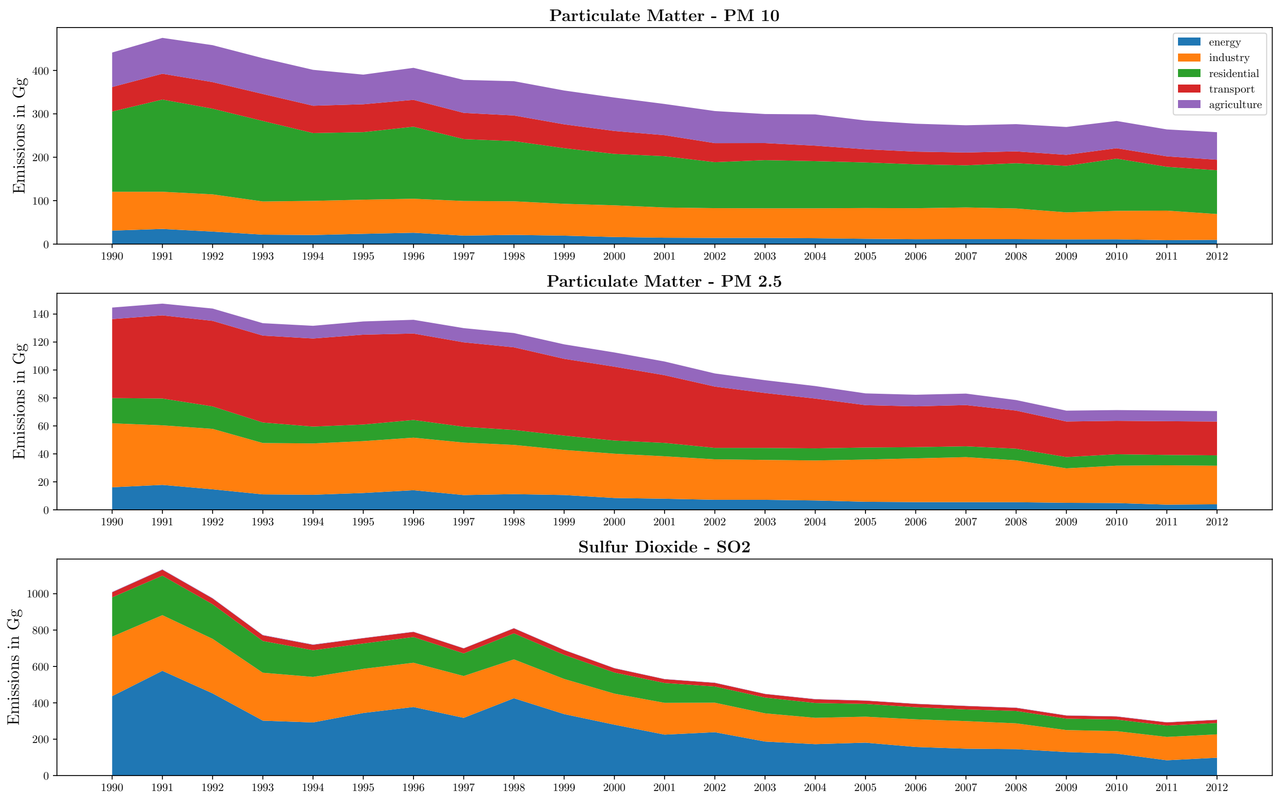
Table 5: Distribution of PM10 Emissions Across Activities (data from EDGAR timeseries): in 2012, 23% of PM10 emissions result from industrial activities.

% Share from:	1990	2000	2012
Industry	30.5	28.0	39.6
Energy	10.8	7.5	5.7
Transport	37.6	46.3	33.2
Residential	12.0	8.3	10.6
Agriculture	5.5	9.0	10.8
Waste	3.6	1.0	0.1

Table 6: Distribution of PM2.5 Emissions Across Activities (data from EDGAR timeseries): in 2012, almost 40% of PM2.5 emissions result from industrial activities.

% Share from:	1990	2000	2012
Industry	32.3	28.8	41.7
Energy	43.0	47.2	31.9
Transport	2.7	3.9	5.5
Residential	21.3	19.6	20.3
Agriculture	0.2	0.2	0.5
Waste	0.5	0.4	0.2

Table 7: Distribution of SO2 Emissions Across Activities (data from EDGAR timeseries): in 2012, almost 42% of SO2 emissions result from industrial activities.



Data are from the EDGAR timeseries (see Crippa et al. (2019)). Industry sector includes IPPC codes 1A2 & 2. Units are Gg for all pollutants.

Figure 12: 1990-2012 Evolution of PM10, PM2.5 and SO2 emissions by Activities (data from EDGAR timeseries): for these pollutants, a substantial fraction of national emissions still result from industrial activities.

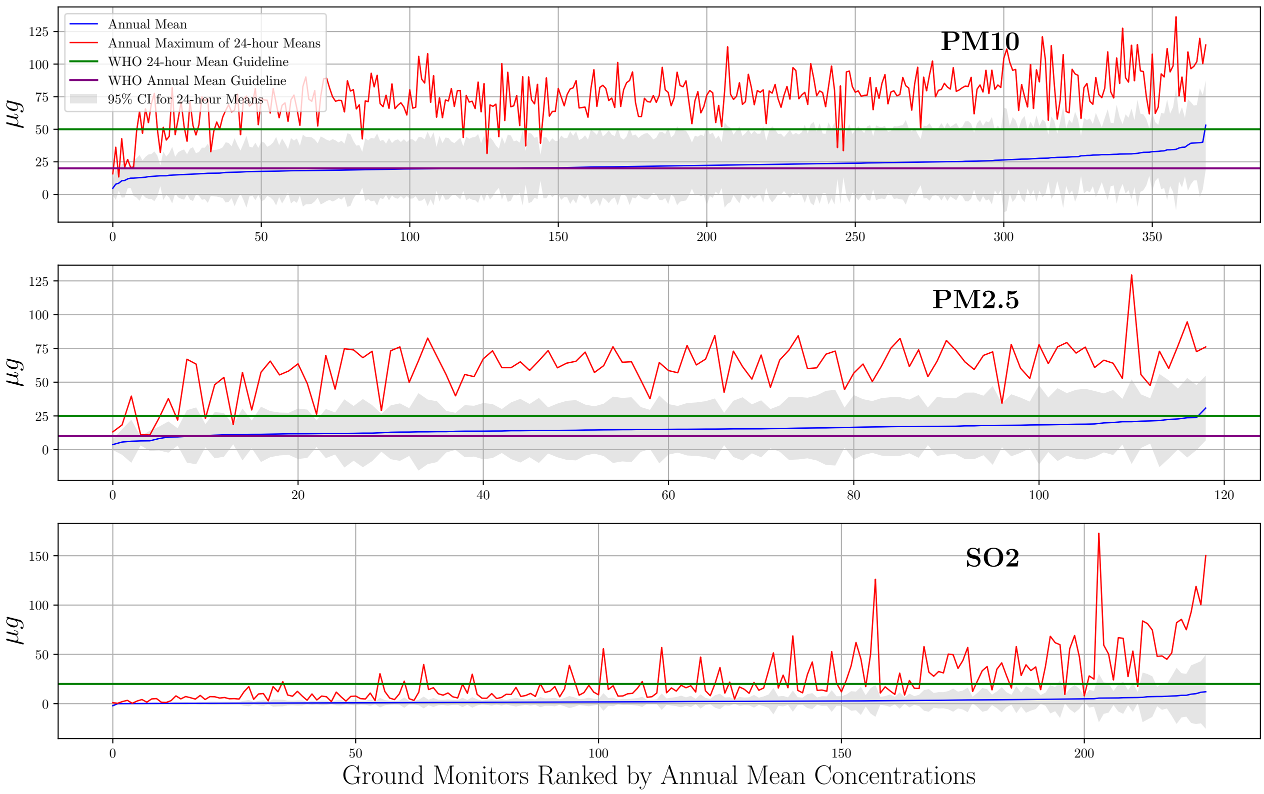


Figure 13: Annual and 24-hour Mean Concentration Measurement from French Ground Monitors in 2012 (data from the EEA AirBase V8): for PM10, PM2.5 and SO2, a substantial fraction of ground monitors measures concentrations above WHO guidelines.

A.2 Air Quality Regulations in France

Air pollution regulation in France is based on European standards. Limits for maximum air concentration of various pollutants and reduction objectives are set at the EU level (directives 2004/107 and 2008/50/CE) to be enforced in each Member States. In turn, they must implement specific action plans to remain within these limits and reach the air quality objectives. In the case of France, these actions are taken both by central and local authorities.

The French government regulates pollutant emissions through tax instruments and regulatory standards (Hauvuy & Riedinger, 2005; Bougon & Lavergne, 2019). The main tax instrument is the General Tax on Polluting Activities (*TGAP*), which is a national tax on the quantity of pollutant emitted per year. In 2016, more than a thousand installations were subjects to this tax (IGF, 2018). However, the air pollution component of the TGAP is limited,[24] and does not constitute the most efficient instrument enforced to limit air pollution (Millock & Nauges, 2006; IGF, 2018). Frameworks regulating industrial emissions are mainly set at a national level. However, their implementation is usually enforced by local authorities. For instance, the main regulatory framework for polluting activities is the *ICPE* regulation, standing for Plants Classified for the Protection of the Environment. The *ICPE* is a set of norms governing polluting plants activities in relation with their impact on the environment. Notably, it requires that an opening permit is delivered by the authority of the departmental prefects for any polluting plant under the condition of implementing specific technology standards and after conducting a local survey of the local population.[25] Moreover, some of these classified plants may fall under the EU Directive on Industrial Emissions (IED).[26] This is the case if they are running polluting installations with capacities above thresholds set in the IED. The main obligation enforced under this regulation is the Best Available Technique (BAT): authorizations for industrial installations are conditional to the use of the least pollution intensive techniques. In that respect, the current national regulation focuses mainly on the largest plants, but any smaller polluting plant not covered by the *ICPE* regulation is regulated at the municipality level.

Furthermore, since the LAURE law, adopted in 1996, other local authorities can use specific measures to improve local air quality and reach air quality national targets. For instance, starting in the early 2000's, several "Atmospheric Protection Plans" (*PPA*) have been implemented in different areas. Within their respective application zone, many of these plans adopted differentiated measures for more versus less densely populated areas. Some of these plans introduced more stringent environmental standards for the manufacturing

[24] In 2016, €50 millions were collected compared to the €3.8 billions collected for the carbon component of the energy consumption tax in 2016 (DGEC, 2016)

[25] These prefects are the regional authorities for the *"départements"* which are NUTS 3 geographic units (larger than commuting zones).

[26] Since 2010, the Industrial Emissions Directive has replaced the Integrated Pollution Prevention and Control Directive. The guidelines are similar and aim at preventing air, water, and soil pollution.

industries, especially in agglomerations above a certain threshold of inhabitants. For instance, the three consecutive *PPA*s for the Paris region mandated lower industrial NOx and PM emission caps relative to the national caps. These plans also implement emergency responses when air concentration of certain pollutant exceed national and European standards. Figure (14) shows the distribution of existent *PPA* across commuting zones[27] and highlights the positive relationship between the size of commuting zones and their probability to lie in the perimeter of an *PPA*.

Finally, both national and local regulations mandate that when air quality reaches regulatory thresholds, emergency actions must be launched by the departmental prefect. Figure (15) illustrates the positive relationship between the size of these areas and the number of emergency actions cases accounted for since 2017.

[27] *PPA* are adopted at a geographic level potentially lower than the commuting zone, which is a statistical construction. The map in Figure (14) displays commuting zone where at least one *PPA* is implemented in its sub-areas.

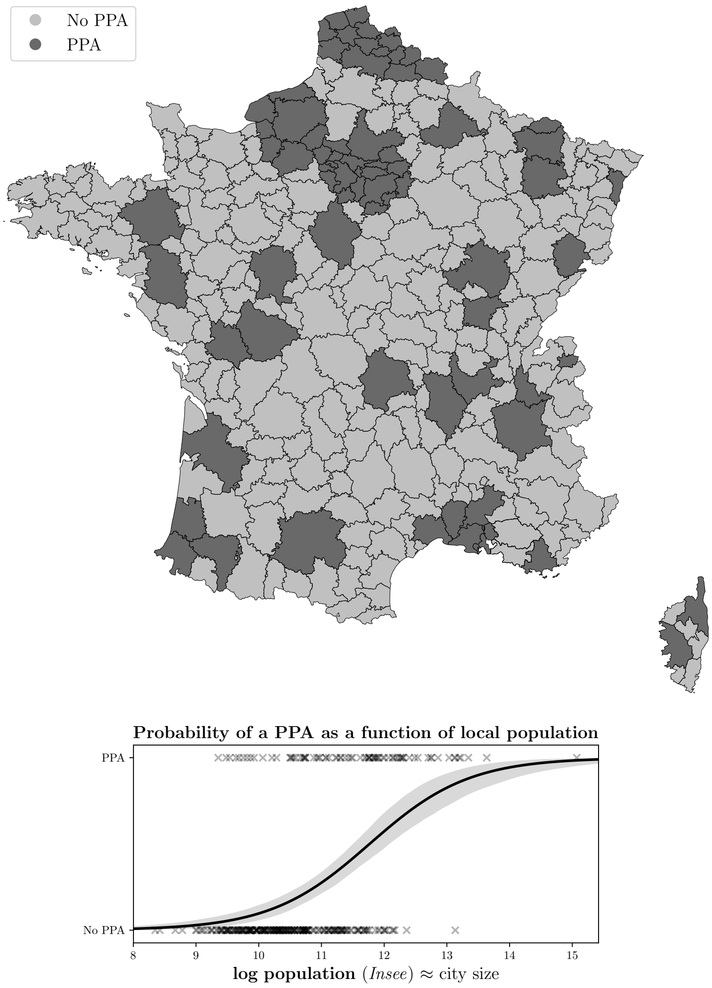


Figure 14: Distribution of Atmospheric Protection Plans across commuting zones
Note: The upper map shows the set of commuting zones that contains municipalities that adopted PPAs since the LAURE law. The lower panel plots the probability that a commuting zone contains municipalities that adopted a PPA as a function of its population. The black line is the result of a logit function fitted on the data on PPA adoption across French commuting zones.
Data on PPAs is from the French Ministry to the Ecological Transition

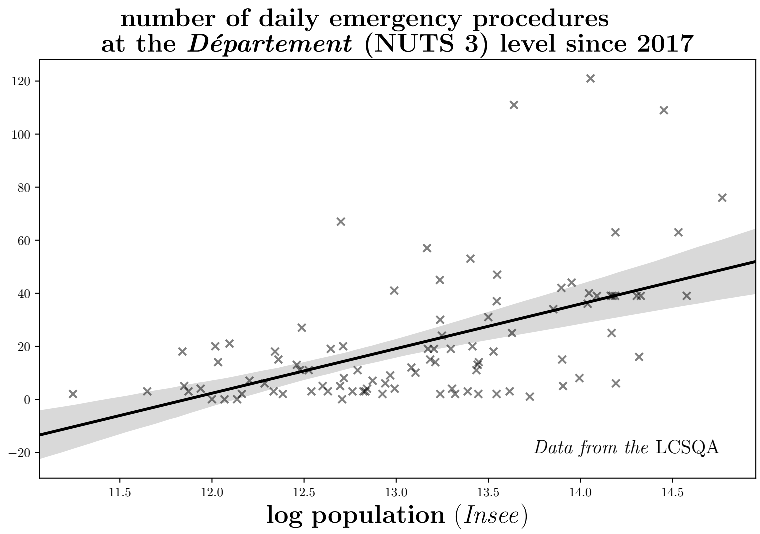


Figure 15: Number of emergency actions due to excessive measures of atmospheric pollutants
Note: This figure plots the number of daily emergency procedures at the *"département"* level since 2017 as a function of the *"département"* population.

A.3 Illustrative Framework Appendix

In each city j, the quantity of the tradable good produced is:

$$Q_j = Z_j^{\alpha}(\phi_j^L L_j^{prod})^{1-\alpha}, \tag{36}$$

with L_j^{prod} the local population employed for production, ϕ_j^L the local labor productivity and Z_j the local quantity of emissions. Each input has the respective unit price w_j and t_j. Moreover, remember that due to the assumption of agglomeration economies local labor productivity ϕ_j^L is equal to $b_j L_j^{\nu}$. Local firms solve the cost minimization program:

$$\min_{Z_j, L_j} L_j^{prod} w_j + Z_j t_j \text{ s.t. } Q_j = Z_j^{\alpha}(\phi_j^L L_j^{prod})^{1-\alpha}. \tag{37}$$

FOCs with respect to labor and emissions yield the relative labor to emissions intensity, which is city specific and depends on the local ratio of input prices:

$$\frac{Z_j}{L_j^{prod}} = \frac{\alpha}{1-\alpha}\frac{w_j}{t_j}. \tag{38}$$

Agglomeration economies are an externality and are not taken into account by firms in their cost minimization problem. That is to say that each firm is atomistic and does not consider that fact that the higher is local employment the larger labor productivity will be. From this proportion of each input in production, one can compute the marginal production price c_j:

$$c_j = \kappa(\alpha) t_j^{\alpha}\left(\frac{w_j}{\phi_j^L}\right)^{1-\alpha}, \tag{39}$$

with $\kappa(\alpha) = \frac{(1-\alpha)^{1-\alpha}}{\alpha_\alpha}$ which I drop in the following computations for simplicity and without loss of generality (α remains a constant in the equilibrium and the central planner's optimization problem).

Given the assumption of perfect competition and not mark-up with a fixed output price fixed to 1 on foreign markets, we then have in equilibrium that $c_j = 1$ which gives the expression for equilibrium wages:

$$w_j = b_j L_j^{\nu} t_j^{-\frac{\alpha}{1-\alpha}}. \tag{40}$$

The total income accounts for payments for production labor and labor employed for abatement (or). As a result one can write:

$$w_j L_j = w_j L_j^{prod} + t_j Z_j, \tag{41}$$

which yields the formula for local emissions as function of local population, local wages and local emission cost:

$$Z_j = \alpha \frac{w_j L_j}{t_j}. \tag{42}$$

Given that the output price is the same across cities and normalized to 1, each worker's budget constraint can be written as:

$$w_j \times 1 = 1 \times c_j. \tag{43}$$

Using equations (40), (42) and (43) to substitute in the representative worker's utility function we get:

$$u_j = a_j b_j^{1-\gamma} L_j^{-\theta} t_j^{\frac{\gamma-\alpha}{1-\alpha}}, \quad \text{with } \theta = \delta + \gamma - \nu(1-\gamma). \tag{44}$$

Given the assumption of free migration of workers, utility is equalized across cities and I note its equilibrium level $\bar{u}$. The national population is fixed and normalized to 1, so that:

$$\sum_{j \in C} L_j = 1. \tag{45}$$

Using the fact that equation (44) can be rewritten as:

$$L_j = \left[\bar{u}^{-1} a_j b_j^{1-\gamma} t_j^{\frac{\gamma-\alpha}{1-\alpha}}\right]^{\frac{1}{\theta}}, \tag{46}$$

which illustrates that there exist a unique spatial equilibrium if and only if $\theta \neq 0$. This condition can be expressed as a condition on the elasticity of the pollution externality: $\gamma \neq -\frac{\delta-\nu}{1+\nu}$. Assuming that pollution is a congestion force (i.e. $\gamma > 0$), this condition ensures that the pollution externality does not exactly offset the combined effect of general congestions effect and agglomeration economies. When congestion effects strictly outweigh agglomeration economies, this condition always holds. I substitute in equation (45) and extract the equilibrium welfare:

$$\bar{u} = \left[\sum_{j \in C} a_j^{\frac{1}{\theta}} b_j^{\frac{1-\gamma}{\theta}} t_j^{\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}}\right]^{\theta}. \tag{47}$$

Finally I solve that optimization program:

$$\max_{\{t_i\}_{i \in C}} \bar{u}(t_1, ..., t_C) \text{ s.t. } \sum_{j \in C} t_j = 1, \tag{48}$$

by computing the first and second order conditions of the Lagrangian:

$$\mathscr{L} = \bar{u}(t_1, ..., t_C) + \lambda \left(C - \sum_{j \in C} t_j\right), \tag{49}$$

with λ the Lagrangian multiplier. Computing the derivatives of $\bar{u}$, I find:

$$\frac{t_j}{\bar{u}} \frac{\partial \bar{u}}{\partial t_j} = \frac{\gamma-\alpha}{1-\alpha} L_j \tag{50}$$

and

$$\frac{\partial^2 \bar{u}}{\partial t_j^2} = \frac{1}{t_j} \frac{\partial \bar{u}}{\partial t_j} \left[\left(\frac{\gamma-\alpha}{1-\alpha} - \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}\right) L_j - \left(1 - \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}\right)\right]. \tag{51}$$

Then, the first order conditions are:

$$\forall j \in C, \bar{u}^{\frac{\theta-1}{\theta}} \frac{\gamma - \alpha}{1 - \alpha} a_j^{\frac{1}{\theta}} b_j^{\frac{1-\gamma}{\theta}} t_j^{*\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}-1} = \lambda, \tag{52}$$

from which wet get equation (7).

The concavity condition that ensures that the set of emission taxes identified in equation (7) is unique and corresponds to a maximization of workers' welfare is:

$$\forall j \in C, \frac{\partial^2 \bar{u}}{\partial t_j^2} < 0. \tag{53}$$

which can be expressed as:

$$\left(\frac{\gamma - \alpha}{1 - \alpha}\right)^2 \left(1 - \frac{1}{\theta}\right) L_j - \left(\frac{\gamma - \alpha}{1 - \alpha}\right) \left(1 - \frac{1}{\theta} \frac{\gamma - \alpha}{1 - \alpha}\right) < 0. \tag{54}$$

In particular, the left-hand side term of inequation (54) is a linear function of each L_j with variations that only depend on the sign of $\theta - 1$. We are looking for the set of parameters for which the welfare function is concave over any set of cities (in particular, the concavity condition must hold for city sizes infinitely close to 0 or equal to 1, if there is a unique city). Consequently, if $\theta > 1$, we only need $\left(\frac{\gamma-\alpha}{1-\alpha}\right)^2 \left(1 - \frac{1}{\theta}\right) < \left(\frac{\gamma-\alpha}{1-\alpha}\right) \left(1 - \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}\right)$. Respectively, if $\theta \leq 1$, we only need $0 \leq \left(\frac{\gamma-\alpha}{1-\alpha}\right) \left(1 - \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}\right)$.

If $\theta > 1$, the condition holds if and only if $1 > \gamma > \alpha$. If $\theta \leq 1$, the condition holds if and only if

$\theta > 1$ means pe is strong $\theta \leq 1$ means pe is weak

A.4 Relaxing the Free Migration Assumption

Assume all the hypotheses of the simplified framework. The only difference is that, now, the distribution of workers across locations is fixed, meaning that for every location j, L_j is exogenously given. This assumption means that there are not adjustment of populations to local changes of wage or air quality. This corresponds to a "very" short-term version of the model where movement frictions are infinitely high.

In this small extension I consider that the central planner still tries to maximize welfare by choosing the distribution of relative emission taxes that maximizes the weighted average welfare per capita. Indeed, without free migration, welfare is not equalized across locations anymore. Still, the central planner has to maximize an objective function. Summing the local welfare levels using local populations as weight corresponds to computing average welfare per capita. It puts the same weight on each worker welfare. In the equilibrium with free migration, it corresponds exactly to the common level of welfare reached in all locations and $\tilde{u} = \bar{u}$.

We note this alternative objective function $\tilde{u}$ and define it as:

$$\tilde{u} = \sum_{j \in C} L_j u_j = \sum_{j \in C} L_j a_j L_j^{-\delta} Z_j^{-\gamma} c_j \tag{55}$$

This extension constitute the alternative to the free migration model. Both are extreme and in reality one may find a situation that is between these two cases. Introducing costs of moving that may allow endogenous distribution of labor across space would require adding a dynamic side to the model. Such assumption would also call for the companion extension to heterogenous workers to encompass the empirical fact that workers with higher income are more likely to adjust to local pollution than low income workers;

Keeping the other assumptions of the simple model, we have:

$$w_j = b_j L_j^{\nu} t_j^{-\frac{\alpha}{1-\alpha}} \tag{56}$$

$$Z_j = \alpha \frac{w_j L_j}{t_j} \tag{57}$$

Replacing the $\tilde{u}$, I get that the weighted average per capita welfare is (without constant):

$$\tilde{u} = \sum_{j \in C} L_j^{1-\theta} a_j b_j^{1-\gamma} t_j^{\frac{\gamma-\alpha}{1-\alpha}} \tag{58}$$

with $\theta = \gamma + \delta - \nu(1-\gamma)$.

Assuming that the central planner solves:

$$\max_{\{t_j\}_{j \in C}} \tilde{u} \quad \text{s.t.} \sum_{j \in C} t_j = 1 \tag{59}$$

She finds the following distribution of relative emission taxes:

$$\forall (j,i) \in C^2, \quad \frac{t_j^*}{t_i^*} = \left(\frac{a_j}{a_i}\right)^{\frac{1-\alpha}{1-\gamma}} \left(\frac{b_j}{b_i}\right)^{1-\alpha} \left(\frac{L_j}{L_i}\right)^{\frac{1-\alpha}{1-\gamma}(1-\theta)} \tag{60}$$

Using estimated elasticities this imply that optimal taxes should be higher in larger cities and cities more productive and with higher amenities which is in line with the result obtained when free migration is assumed.

A.5 Estimation Appendix

A.5.1 Building Plant-Level Pollutants Emissions from Energy Surveys

Pollutant	PCC	Nb. Obs.
SO2	.1929*	*1060*
PM10	.7758*	*45*
NOX	.8192*	*1734*
COVNM	.2143*	*4689*
CO	.7898*	*313*

Table 8: Pearson correlation coefficients between actual emissions values from the E-PRTR and estimated values from the EACEI survey *(* indicates significance at the 1% level)*.

A.5.2 Estimation of the Sector Specific Emission Intensities

Assuming that for a plant i, installed in city c and producing goods from sector s, that has an intrinsic productivity equal to ϕ_{ics} production follows $q_{ics} = z_{ics}^{\alpha_s} l_{ics}^{1-\alpha_s} \phi_{ics}^{1-\alpha_s}$ and that it faces demand $q_{ics} = k_{cs} p_{ics}^{-\sigma_s}$, where p_{ics} is the price charged.

As I do not observe the distribution of inputs across destination, I abstract from the fact the plants may face distinct demand functions in each destination (in the case where trade between cities is costly). This assumption is mainly due to data constraint but considering that, in practice, within-country trade costs are low, it is also acceptable. The variable k_{cs} encompasses both the size of the output market and the level of competition on this market. What is important is that all plants within the same region face the same k_{cs}.

Writing the expression for plant revenues:

$$r_{ics} = p_{ics} q_{ics} \tag{61}$$

$$= k_{cs}^{\frac{1}{\sigma_s}} q_{ics}^{\frac{\sigma_s - 1}{\sigma_s}} \tag{62}$$

$$= k_{cs}^{\frac{1}{\sigma_s}} \left(z_{ics}^{\alpha_s} l_{ics}^{1-\alpha_s} \phi_{ics}^{1-\alpha_s} \right)^{\frac{\sigma_s - 1}{\sigma_s}} \tag{63}$$

Based on the standard results from monopolistic competition with CES demand, markups are fixed and depend on σ_s. Based on the standard results of cost minimization under Cobb-Douglas production functions, plant-level expenditure shares across inputs are fixed and given by α_s. These two results combined allow the following formulation of the relationship between revenues and employment:

$$r_{ics} = \frac{\sigma_s}{\sigma_s - 1}\frac{1}{1-\alpha_s} w_{ics} l_{ics} \tag{64}$$

with w_{ics} the plant-specific unit cost of labor. In the main model of this paper I make the assumption that this cost is the wage and is the same across sectors and firms within a city. However, in this appendix on the estimation of emission elasticities, I temporarily relax this assumption to show that if, in reality, there are plant specific shocks on wages (which is likely to be the case) it does not threaten my estimation strategy.

Substituting revenues in (61) using (64):

$$\frac{\sigma_s}{\sigma_s - 1}\frac{1}{1-\alpha_s} w_{ics} l_{ics} = k_{cs}^{\frac{1}{\sigma_s}} \left(z_{ics}^{\alpha_s} l_{ics}^{1-\alpha_s} \phi_{ics}^{1-\alpha_s} \right)^{\frac{\sigma_s - 1}{\sigma_s}} \tag{65}$$

$$\Rightarrow l_{ics}^{1-(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}} = \frac{\sigma_s - 1}{\sigma_s}(1-\alpha_s) k_{cs}^{\frac{1}{\sigma_s}} w_{ics}^{-1} \phi_{ics}^{(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}} z_{ics}^{\alpha_s \frac{\sigma_s-1}{\sigma_s}} \tag{66}$$

$$\Rightarrow l_{ics} = \left[\frac{\sigma_s - 1}{\sigma_s}(1-\alpha_s) k_{cs}^{\frac{1}{\sigma_s}} \right]^{\frac{1}{1-(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}}} z_{ics}^{\frac{\alpha_s \frac{\sigma_s-1}{\sigma_s}}{1-(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}}} \left(w_{ics}^{-1} \phi_{ics}^{(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}} \right)^{\frac{1}{1-(1-\alpha_s)\frac{\sigma_s-1}{\sigma_s}}} \tag{67}$$

which can be simplified as:

$$l_{ics} = \tilde{k}_{cs}(k_{cs}) z_{ics}^{\frac{\alpha_s(\sigma_s-1)}{1+\alpha_s(\sigma_s-1)}} \tilde{\phi}_{ics}(\phi_{ics}, w_{ics}) \tag{68}$$

with $\tilde{k}_{cs}$ a function only of k_{cs}, so common to all plants within a city-sector pair, and $\tilde{\phi}_{ics}$ a function of plant specific productivity and wage shocks.

Using the previous relationship, I estimate:

$$\log l_{icst} = \beta_{0s} + \beta_{1s} \log z_{icst} + \mu_t + \rho_{ics} + \epsilon_{icst} \tag{69}$$

to retrieve $\beta_{1s} = \frac{\alpha_s(\sigma_s-1)}{1+\alpha_s(\sigma_s-1)}$. In introduce time fixed-effects, μ_t, and plant fixed effects, ρ_{ics}.

However, unobserved shocks on productivity and wages at the plant level are implicitly included in the residual ϵ_{icst} and are correlated with input use l_{icst} and z_{icst}, OLS estimates are biased. This is a sort of "transmission bias" which is frequent in the literature estimating production functions. Th only difference here is that I do not estimate the production function using revenues on the left-hand side. This is due to not observing revenues at the plant level but only at the firm level. Indeed, I also only observe emissions of atmospheric pollutants at the plant level for subsets of plants within firms. Using plant and firm level observations in the same equation would force me to drop observations from multi-plants firms. As a result, the estimation sample would be too small.

I rely on an IV approach to estimate β_{1s}. I build an instrument that uses exogenous shocks on fuel prices as source of variation of plant-level emissions. This instrument is a fixed-weight energy price index that measures the plant-specific exposure to variation in fuels prices based on each plant distribution of energy consumption across fuel types in the first period where it is observed. I build on Sato et al. (2019) to compute this instrumental variable.

$$FEPI_{ist} = \sum_{f \in \Omega_{fuels}} \omega_{f,ist_0} p_{f,st} \tag{70}$$

where $FEPI_{ist}$ is the plant-specific energy price index build from plant i share of energy expenditures in fuels $f \in \Omega_{fuels}$ (coal, natural gas, electricity, etc.) in period t_0 and $p_{f,st}$ the specific fuel price common to all plants in sector s in period t. The exclusion restriction requires in this case that plant specific shares of expenditures across fuels are not correlated with unobserved shocks that are included in ϵ_{icst}. Given (68), shocks in the residual only comes from demand shocks that are common to all plants, shocks on plant specific wage and shocks on plant specific productivity. Therefore, the exclusion restriction is, in this case, likely to be verified.

A.6 Application to France Appendix

A.6.1 Sources of Equilibrium Data

Empirical Distribution of Wages Across Commuting Zones: I use the distribution of average hourly wages at the commuting zone level for 2012 from the INSEE dataset publicly available at `https://www.insee.fr/fr/statistiques/2021266`. This data is aggregated geographically by the INSEE using worker level data from the Social Data Annual Declaration (DADS - *Déclaration annuelle des données sociales*). I use the 2010 commuting zones (*Zones d'Emploi*) definition and normalize mean hourly local wages so that the average over all commuting zones is equal to one.

Empirical Distribution of Populations Across Commuting Zones: I use the distribution of working population at the *"Communes"* level for 2012 from the INSEE dataset publicly available at `https://www.insee.fr/fr/statistiques/2128672`. The *"Communes"* is one of the most disaggregated administrative geographic unit in France. There are around 36,000 of them over the territory. The 2010 commuting zone's (*Zones d'Emploi*) definition is a partition of the set of French *"Communes"* in around 300 areas where inhabitants both work and live. I use the correspondence table between *"Communes"* and *Zones d'Emploi* from the INSEE and publicly available at `https://www.insee.fr/fr/information/2114596`. I sum working populations at the *"Communes"* level to get aggregated working populations at the *Zones d'Emploi* level. Finally I normalize local populations so that the total national working population over the whole set of commuting zones is equal to one.

Empirical Distribution of Emissions Across Commuting Zones: I use data from the National Spatialized Inventory (*"Inventaire National Spatialisé"*) which is built by the French ministry in charge of environmental issues. The inventory is publicly available at `http://emissions-air.developpement-durable.gouv.fr/`. I exported data for the 2012 platform (data is also available for 2004 and 2007) for PM10, PM2.5, SO2, NOx and COVNM. These datasets provide amounts of pollutants emitted disaggregated at the *"Communes"* level and across broad sectors of activities following the Selected Nomenclature for Air Pollution (SNAP). I aggregate emissions at the commuting zone level by using the correspondence table corresponding to the 2010 definition of commuting zones and summing emissions within commuting zones across *"Communes"*. I keep emissions from codes 3 and 4 of the SNAP which correspond to emissions from industrial combustion plants and industrial processes without combustion. These two codes broadly correspond to emissions due to the manufacturing industries. I aggregate over these two emission sectors by summing emissions. Figure(16) displays the correlation between levels of emissions of available pollutants. One can observe that the correlation is pretty strong. I only retain the distribution of PM10 emissions as the empirical data for the distribution of my model's representative pollutant across commuting zones.

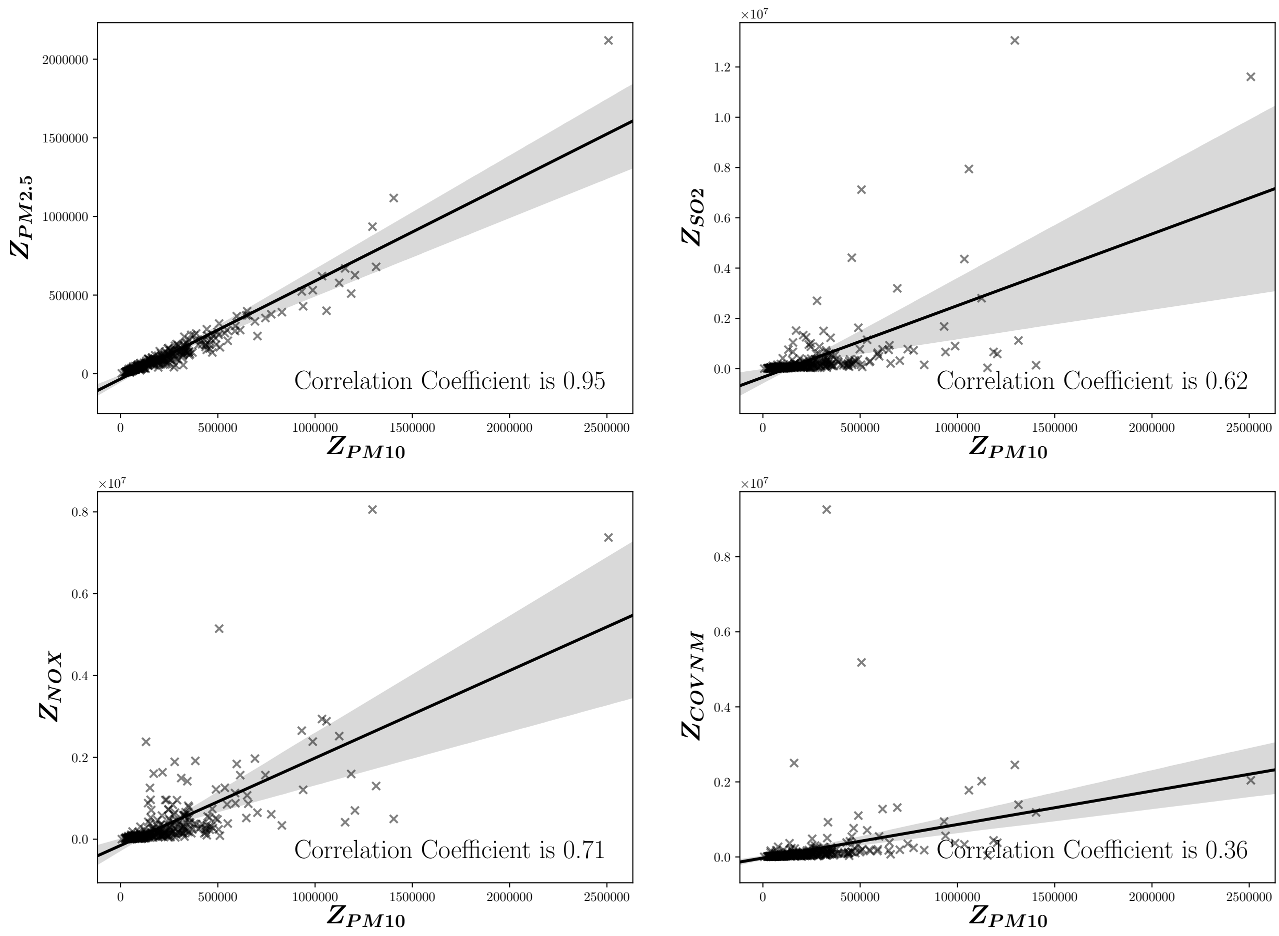


Figure 16: Correlation Between Emissions of Harmful Industrial Pollutants at the Commuting Zone Level

A.6.2 Descriptive Statistics on Equilibrium Data

A.6.3 Descriptive Statistics on Computed Local Characteristics

	Wages	Populations	PM2.5 Emissions
count	303.000	303.000	303.000
mean	1.000	0.003	1.000
std	0.101	0.009	1.072
min	0.847	0.000	0.027
25%	0.938	0.001	0.352
50%	0.973	0.001	0.697
75%	1.037	0.003	1.256
max	1.585	0.136	9.830

Table 9: Descriptive Statistics on the Spatial Distributions of Wages, Populations and Emissions across French Cities

Highest Populations	Highest Wages	Highest Emissions
Paris	Saint-Quentin-en-Yvelines	Dunkerque
Lyon	Paris	Thionville
Toulouse	Saclay	Bordeaux
Roissy - Sud Picardie	Rambouillet	Istres - Martigues
Bordeaux	Versailles	Nantes
Marseille - Aubagne	Poissy	Toulouse
Saclay	Cergy	Saint-Dié-des-Vosges
Nantes	Marne-la-Vallée	Lyon
Lille	Créteil	Marseille - Aubagne
Rennes	Aix-en-Provence	Nancy

Table 10: Top Ten Cities with Highest Populations, Wages and Emissions

Lowest Populations	Lowest Wages	Lowest Emissions
Sartène - Propriano	Saint-Flour	Corte
Corte	Mauriac	Issoudun
Ghisonaccia - Aléria	Sartène - Propriano	Loches
Le Blanc	Le Blanc	Figeac
Ambert	Sarlat-la-Canéda	L'Aigle
Calvi - L'Île-Rousse	Brioude	Vire
Loches	Lozère	Menton - Vallée de la Roya
Issoudun	Villeneuve-sur-Lot	Ghisonaccia - Aléria
Avallon	Saint-Amand-Montrond	Le Blanc
Porto-Vecchio	Ussel	Mauriac

Table 11: Bottom Ten Cities with Lowest Populations, Wages and Emissions

	Amenities	Productivities	Emission Costs
count	303.000	303.000	303.000
mean	0.003	1.000	1.000
std	0.010	0.160	1.173
min	0.000	0.785	0.075
25%	0.000	0.906	0.521
50%	0.000	0.956	0.812
75%	0.002	1.045	1.153
max	0.087	2.044	18.059

Table 12: Descriptive Statistics on the Spatial Distributions of Amenities, Productivities and Emission Costs across French Cities

Highest Amenities	Highest Productivities	Highest Emission Costs
Bordeaux	Saint-Quentin-en-Yvelines	Paris
Dunkerque	Paris	Saclay
Nantes	Saclay	Orly
Lyon	Rambouillet	Marne-la-Vallée
Marseille - Aubagne	Versailles	Créteil
Toulouse	Poissy	Cergy
Avignon	Cergy	Lille
Roissy - Sud Picardie	Marne-la-Vallée	Poissy
Nancy	Créteil	Lyon
Rennes	Aix-en-Provence	Cannes - Antibes

Table 13: Top Ten Cities with Highest Amenities, Productivities and Emission Costs

Lowest Amenities	Lowest Productivities	Lowest Emission Costs
Saint-Quentin-en-Yvelines	Saint-Flour	Saint-Dié-des-Vosges
Corte	Mauriac	Dunkerque
Versailles	Sartène - Propriano	Thionville
Rambouillet	Villeneuve-sur-Lot	Istres - Martigues
Issoudun	Sarlat-la-Canéda	Maurienne
Loches	Brioude	La Teste-de-Buch
L'Aigle	Lozère	Sartène - Propriano
Figeac	Le Blanc	Péronne
Étampes	Saint-Amand-Montrond	Dole
Ghisonaccia - Aléria	Bressuire	Jonzac - Barbezieux-Saint-Hilaire

Table 14: Top Ten Cities with Lowest Amenities, Productivities and Emission Costs

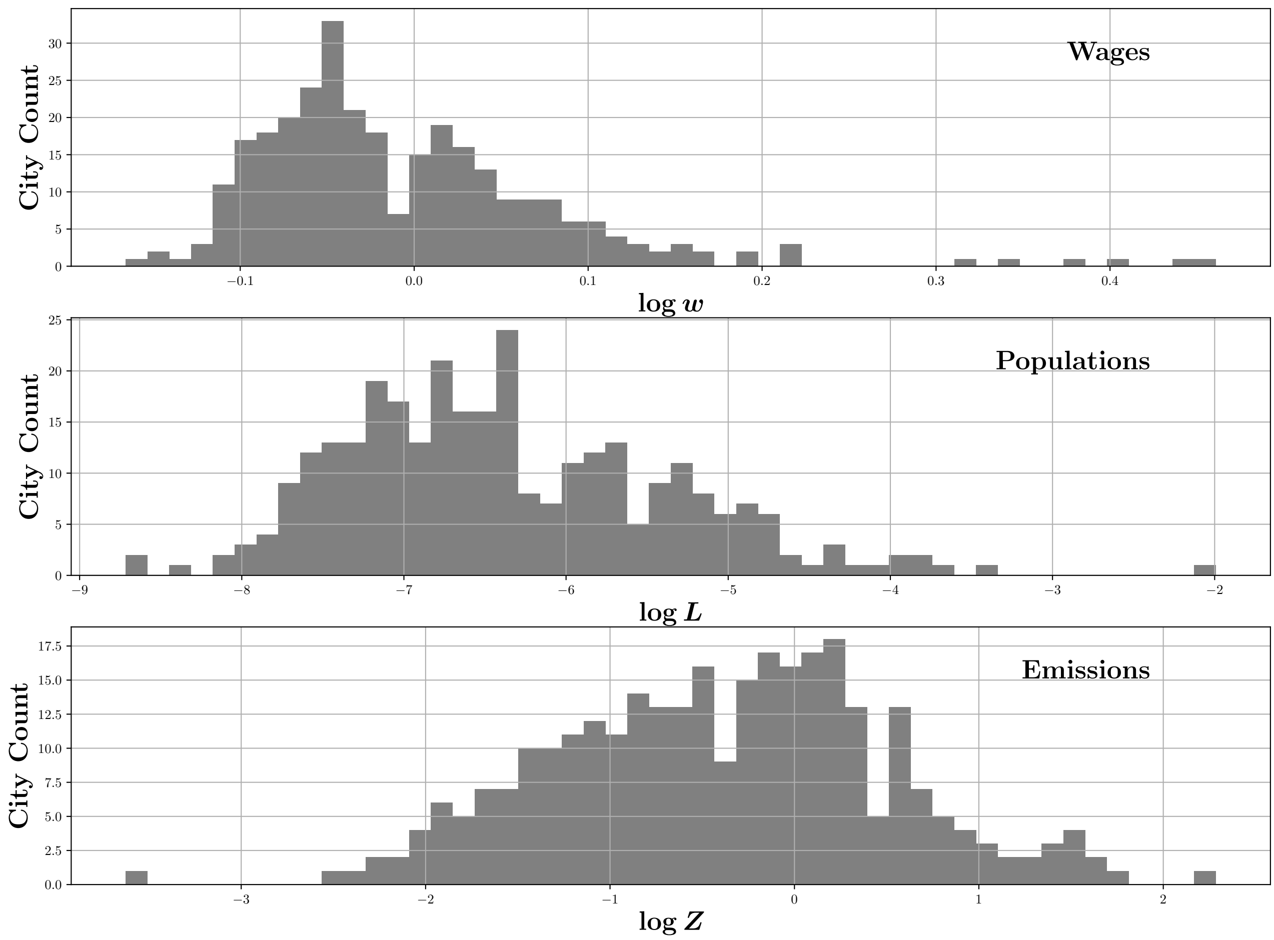


Figure 17: Histograms of Spatial Distributions of Wages, Populations and Emissions across French Cities

All variables are in log.

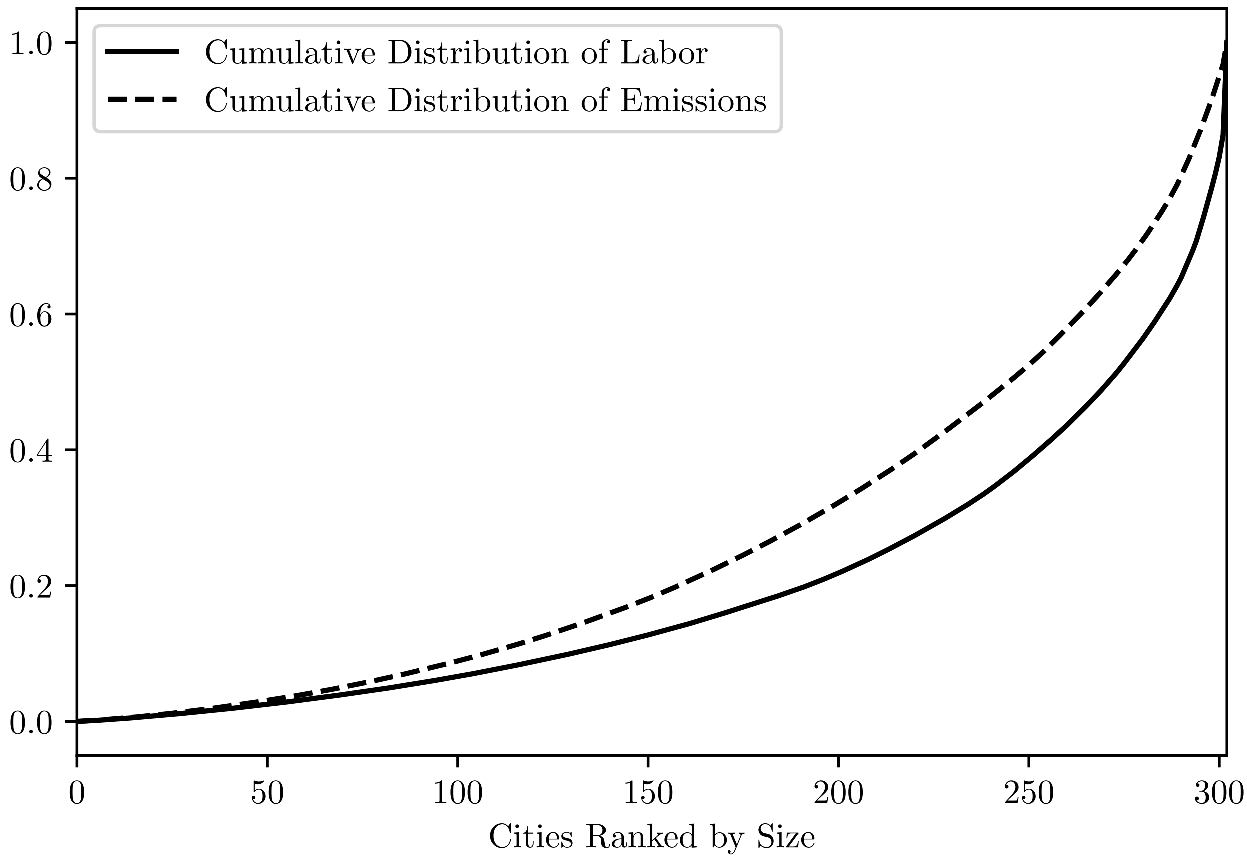
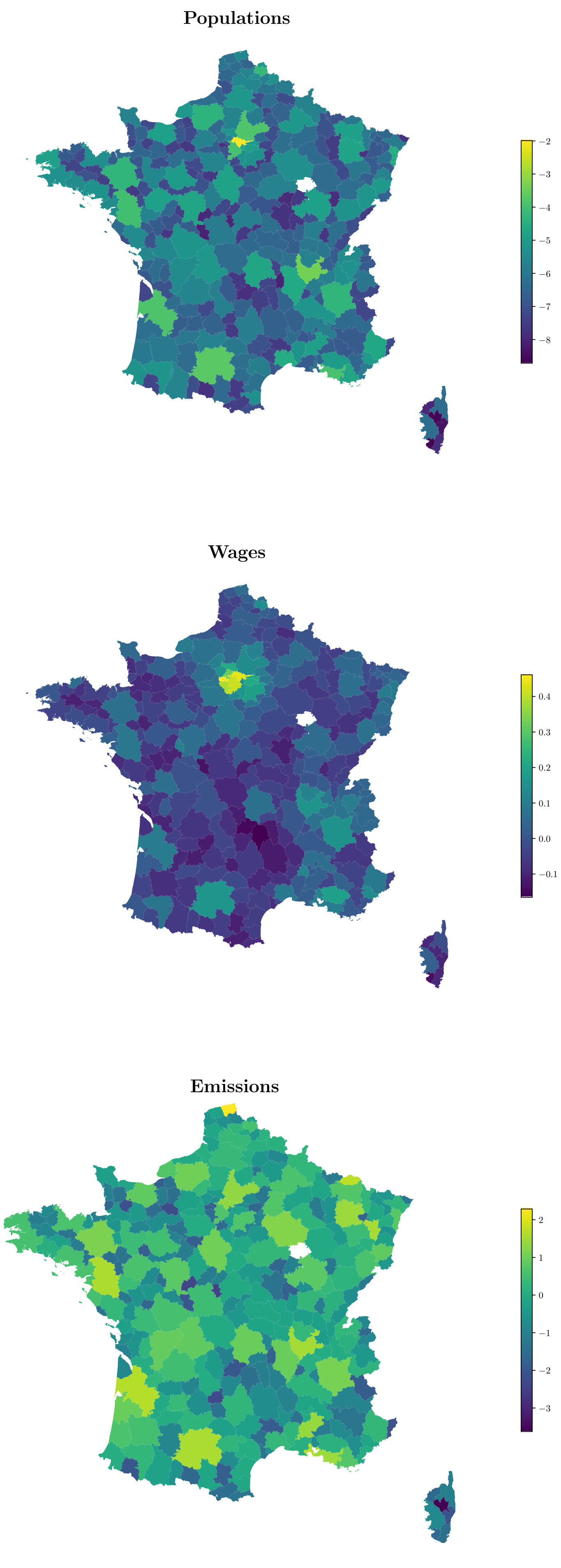


Figure 18: Cumulative Distributions of Populations and Emissions across French Cities
Cities are ranked by size, from smallest to largest, and the normalized cumulative sum of populations and emissions is plotted.



All variables are in log.

Figure 19: Spatial Distributions of Wages, Populations and Emissions across French Cities

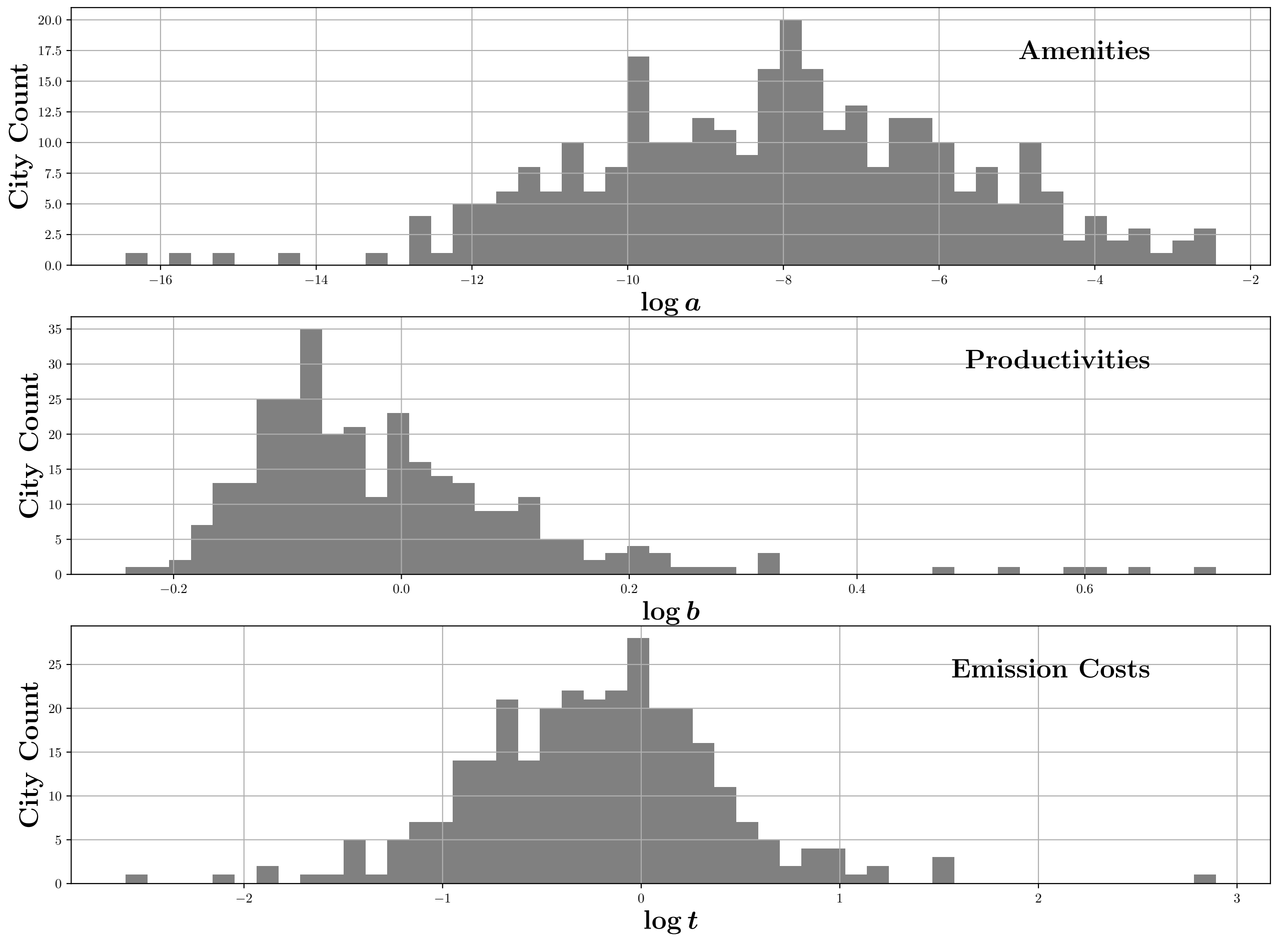


Figure 20: Histograms of Spatial Distributions of Wages, Populations and Emissions across French Cities

All variables are in log.

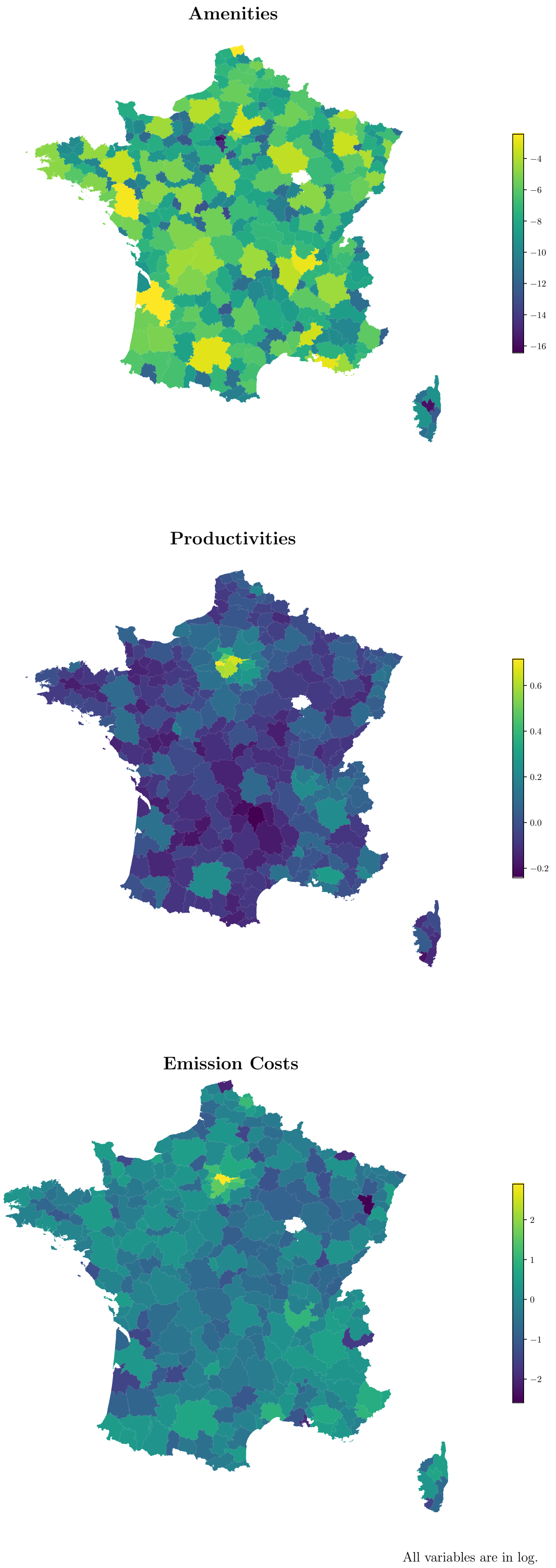


Figure 21: Spatial Distributions of Amenities, Productivities and Emission Costs across French Cities