An Optimal Distribution of Polluting Activities Across Space

An Application to France [*]

Martin Jégard[†]

July 2021

Preliminary and incomplete

Abstract

Should the largest cities be the most polluting? On the one hand, spatial concentration of economic activities brings welfare gains through agglomeration economies and increased local real wages. On the other hand, aggregating too much polluting activities in the same place leads to lower air quality and detrimental effects on health and productivity of local workers. Building on a spatial general equilibrium model, featuring endogenous pollution, trade and between-cities migration, I investigate welfare effects of the spatial heterogeneity of local stringencies resulting from current air pollutants regulations. I calibrate the model using French data and show that current emission policies target the most populated cities. As a result, these cities are less polluting and relatively larger than what they would be under a spatially uniform policy stringency. However I also find that taking into account productivity and amenity intrinsic local endowments could lead to higher 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

Presence of atmospheric pollutants causes adverse local health outcomes, even below regulatory standards (Zivin & Neidell, 2013; Deryugina et al., 2019). A large array of such pollutants is emitted by industrial processes for manufacturing production. In line with the common observation that economic activities are unevenly distributed across space, polluting activities are found to be concentrated in cities that also concentrate the largest shares of population (see Figure 1 for France). In the last decades, air quality regulations have been implemented to enhance local air quality and reduce the number of yearly casualties imputed to local atmospheric pollution[1]. Because air pollutant concentrations are generally higher in large cities, these policies seem to specifically target such places by imposing higher levels of regulation stringency relative to less populated areas (see section 2.1).

Considering that the spatial distributions of population and pollution are the result of a general equilibrium where people potentially react to local bad air quality by moving away but where neither firms nor workers internalize the impacts of their production and location choice on atmospheric pollutants local concentrations, it may be relevant to investigate the welfare implications of such place-based air quality regulations as well as their role in shaping the spatial distribution of polluting activities. In this paper, I answer these questions by extending recent quantitative spatial setups and provide an empirical application to France.

A recent literature has focused on the explanation for the distribution of economics activities across space, as in Allen & Arkolakis (2014) and Redding (2020), and how place-based policies can implement welfare improving spatial reallocations, as in Fajgelbaum & Gaubert (2020). Introducing endogenous pollution from manufacturing sectors in such setups, I provide a framework explaining the spatial distribution of pollution from the industry that has been identified in the empirical literature, most recently by Colmer et al. (2020).

In my model I assume a discrete system of heterogenous cities, that differ in their idiosyncratic endowments of amenities and productivity. In the tradition of Rosen (1979)-Roback (1982) a fixed national population of homogenous workers distributes itself across cities. Workers derive utility from local amenities and from the consumption of an industrial tradable composite good. In each city, continuums of heterogenous firms from distinct industrial sectors produce differentiated varieties that aggregate into the composite good. Local pollution is endogenous to the extent that it is modeled as a by-product of industrial production. Following the standard Copeland & Taylor (2004) setup, emissions can equivalently be seen as inputs for production. Firms produce tradable goods using labor and local atmospheric pollutants emissions. In the rest of the paper, I just refer to that second input as emissions and assume a unique representative industrial pollutant for simplicity. To encom-

[1] Per year, anthropogenic PM2.5 emissions are the cause of 48,000 early deaths in France (Pascal et al., 2016).

pass the adverse consequences from local air pollution, I introduce a local externality from industrial emissions by assuming that the total quantity emitted per city affects negatively local amenities, using a constant elasticity environmental damage function. Therefore, local amenities are endogenous, as in Diamond (2016). In turn, local production costs are a function of local wages, set in equilibrium, local productivity, and local emission costs. The distribution of relative emissions costs between cities is thought to represent the uneven stringency of local air pollution regulations that I document in section (2.1). Thus, the cost of emitting a unit of pollution is city-specific and set outside of the equilibrium. I retain the traditional assumption that, in the long term, workers move freely across cities so that welfare is common across space. In turn, it pins down the distribution of population across cities. Along with between-cities exogenous transport costs, this leads to an endogenous uneven distribution of industrial activities across cities.

I apply this spatial general equilibrium framework to the specific case of France in two stages. In the first one, I calibrate the model's parameters using an extensive set of city, firm, and plant-level data. I estimate key elasticities that respectively determine the strength of pollution externalities, agglomeration economies and general congestion effects. The effects are confounding in nature and the equilibrium spatial distributions of variables as well as the welfare analysis are empirical questions which results depend on the relative strengths of these parameters. In equilibrium, these elasticities govern relationships between observed endogenous variables that are pinned down by exogenous and unobserved local characteristics. Therefore, I exploit tailored instrumental variables to provide precise unbiased estimates. I also provide novel estimates of sector-specific elasticities governing the substitution of emissions and labor in production. In the second stage, I follow the tradition of revealed preferences methodologies and I invert the spatial equilibrium. From observed spatial distributions of populations, wages, and pollutants emissions I retrieve the corresponding idiosyncratic distributions of amenities, productivities, and emissions costs at the city level. Particularly, this means that I compute the distribution of emission costs, in each city relative to the others, that corresponds to the existing distribution of relative stringencies from existing local industrial pollutants emissions regulations.

Upon observing the actual distribution of local emission costs, I characterize the welfare implications of implementing air quality regulations that generate such heterogenous spatial distribution of relative local regulation stringencies. To do so, I compare the observed equilibrium with an alternative equilibrium where the level of emission regulation stringencies is constant across cities. In equilibrium, firms do not internalize the impact of their production choices on the local air quality. Specifically, they do not consider that their input choice affects local amenities and, in turn, the local labor supply. This externality creates a space for welfare improving policy. Making use of the model, I solve an optimization problem and identify the spatial distribution of relative emission costs that maximizes welfare.

My results indicate that the relatively higher levels of emission costs faced by the largest cities lead them to be relatively less polluting and larger than what they would be under a homogenous level of stringency across space. Solving the optimization problem shows that, compared to the current distribution, it is welfare-improving to apply relatively higher emission costs in cities that are more productive and have larger amenities. However, the current distribution of emission stringencies across cities does not fully coincide with the spatial distribution of city productivities and level of amenity. I show that moving to a spatial distribution of emission stringencies that takes into account these local characteristics is welfare improving. Compared to the current distribution of emission costs, the optimal distribution would impose even higher emission costs in the largest cities and relax them in the smallest cities.

The remainder of the paper is organized as follows. The next section provides some context on the French air quality regulatory framework and presents in more details the related literature to which this paper contributes. In the third section I exploit a simplified model to drive intuition on the equilibrium distribution of emissions and the pollution externality. The general model is detailed in the fourth section. The fifth section presents the calibration of the model. Finally, the last section contains results from numerical welfare optimization problems and discusses the results.

2 Elements of Context

2.1 The Geographic Component of French Air Quality Policies

As an example for France, the upper tier of Figure (1) illustrates that the most populated areas are also the most polluted. Specifically, it plots the relative 2012 mean PM 2.5 air concentration for each French commuting zones as a function of the local number of workers, which reflects the size of the area relative to the others. The middle tier of Figure (1) reveals that these most polluted and most populated commuting zones are also the most polluting. It represents the relationship between the quantity of PM 2.5 emitted by industrial activities within the commuting zone and the local number of workers. The lower tier of Figure (1) illustrates how the relative local cost of emitting pollutants (here for PM2.5 emitted by industrial processes) varies with the relative size of the local labor market. We observe that the cost of emitting PM2.5 is relatively higher in larger commuting zones.

Air pollution regulation in France is based on European standards. Limits for maximum air concentration of various pollutants and reduction objectives (for 2020 then 2030) 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 not exceed these limits and

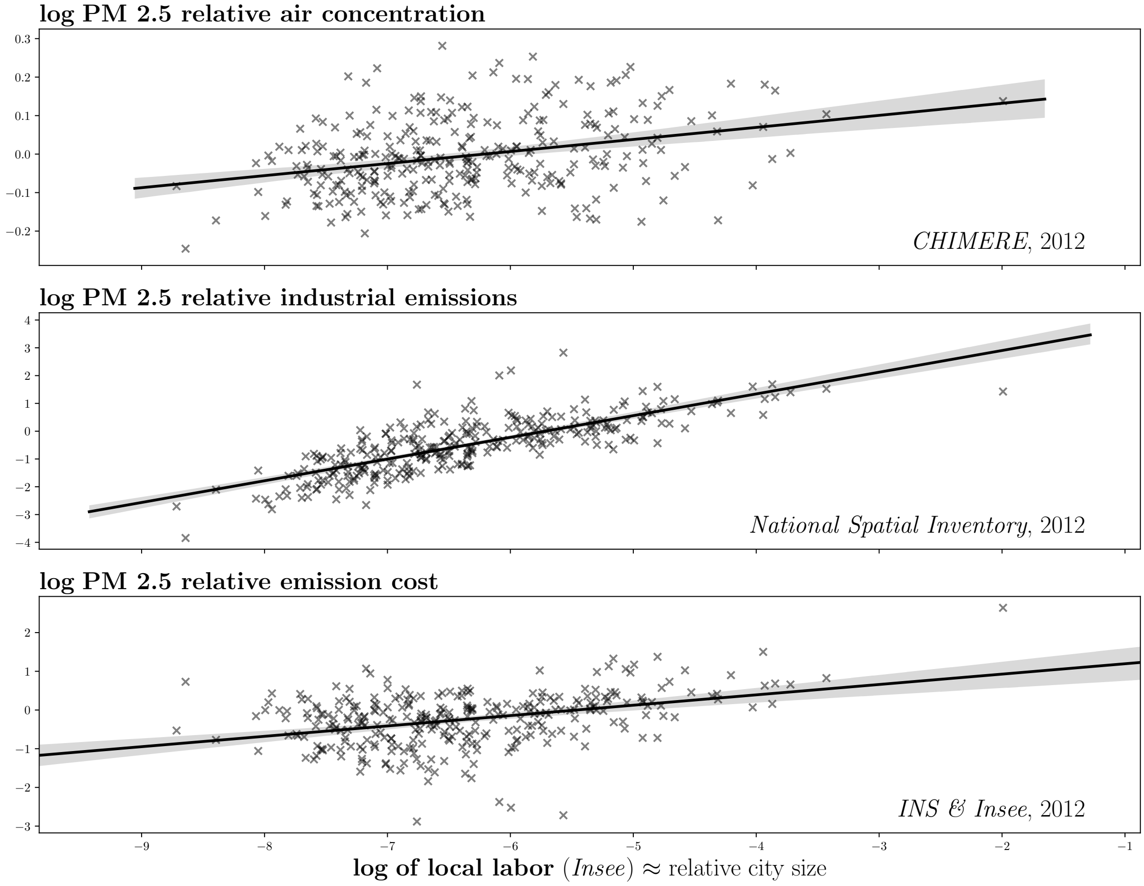


Figure 1: PM 2.5 spatial distribution

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. The relative emission cost data is the log of the ratio of total wage payments in the employment area over the quantity of PM2.5 emitted by the industry. All values are for 2012.

reach the air quality objectives. In the case of France, these actions are distributed between central and local authorities.

The French government regulates pollutant emissions through tax instruments and regulatory standards. The main tax instrument is the *TGAP*, standing for General Tax on Polluting Activities, 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[2] and does not constitute the most efficient instrument enforced to limit air pollution (IGF, 2018). 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 to their impact on the environment. Notably, it requires that a permit should be delivered by the regional authority for any polluting plant's opening in a specific area after implementing

[2]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)

specific technology standards at the request of the authority, and after conducting a local survey of the local population. Moreover, some of these classified plants may fall under the EU Directive on Industrial Emissions (IED) [3]. 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) which implies that authorization of these installations is conditional to the use of these techniques that are deemed the least pollution intensive. 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 regulate at the municipality level.

However, if the framework regulating industrial emissions is mainly set at a national level, its implementation is partly enforced by local authorities. In practice, the enforcement of the BAT for industrial plants is the responsibility of regional prefects, who are the authorities at the NUTS 3 geographic 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" (APP) 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 industries, especially in agglomerations above a certain threshold of inhabitants. For instance, the three consecutive APPs 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 (2) shows the distribution of existent APP across commuting zones[4] and highlights the positive relationship between the size of commuting zones and their probability to lie in the perimeter of an APP.

Finally, both national and local regulations mandate that when air quality reaches regulatory thresholds, emergency actions must be launched by the prefect of the department (NUTS 3 geographic level, so areas larger that the commuting zones). Data on these is publicly available for the last years. Figure (3) illustrates the positive relationship between the size of these areas and the number of emergency actions cases accounted for since 2017.

[3] 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.

[4] APP are adopted at a geographic level potentially lower than the commuting zone, which is a statistical construction. The map in Figure (2) displays commuting zone where at least one APP is implemented in its sub-areas.

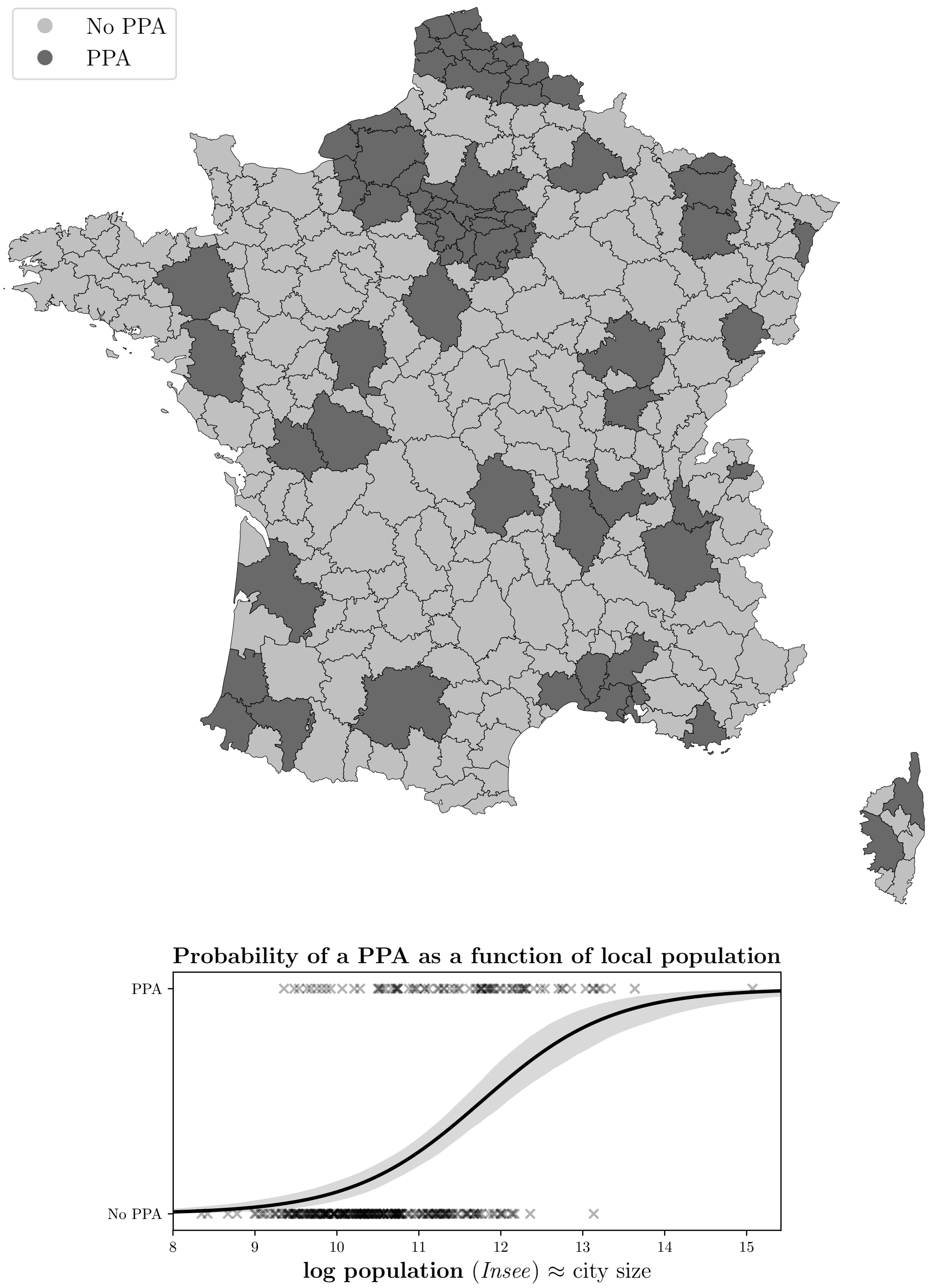


Figure 2: Distribution of Atmospheric Protection Plans across Commuting Zones
Data on PPAs is from the French Ministry to the Ecological Transition

2.2 Relation to the Literature

My paper builds on the recent strand of literature, reviewed by Redding & Rossi-Hansberg (2017) and Redding (2020), that investigates the distribution of economic activity across space. Specifically, it models space as a system of heterogenous cities with exogenous endowments in local amenities and productivities. The setup follows the tradition of Rosen (1979)-Roback (1982) and is particularly close to Diamond (2016) to the extent that I introduce endogenous local amenities (with the distinction that workers are all identical in pro-

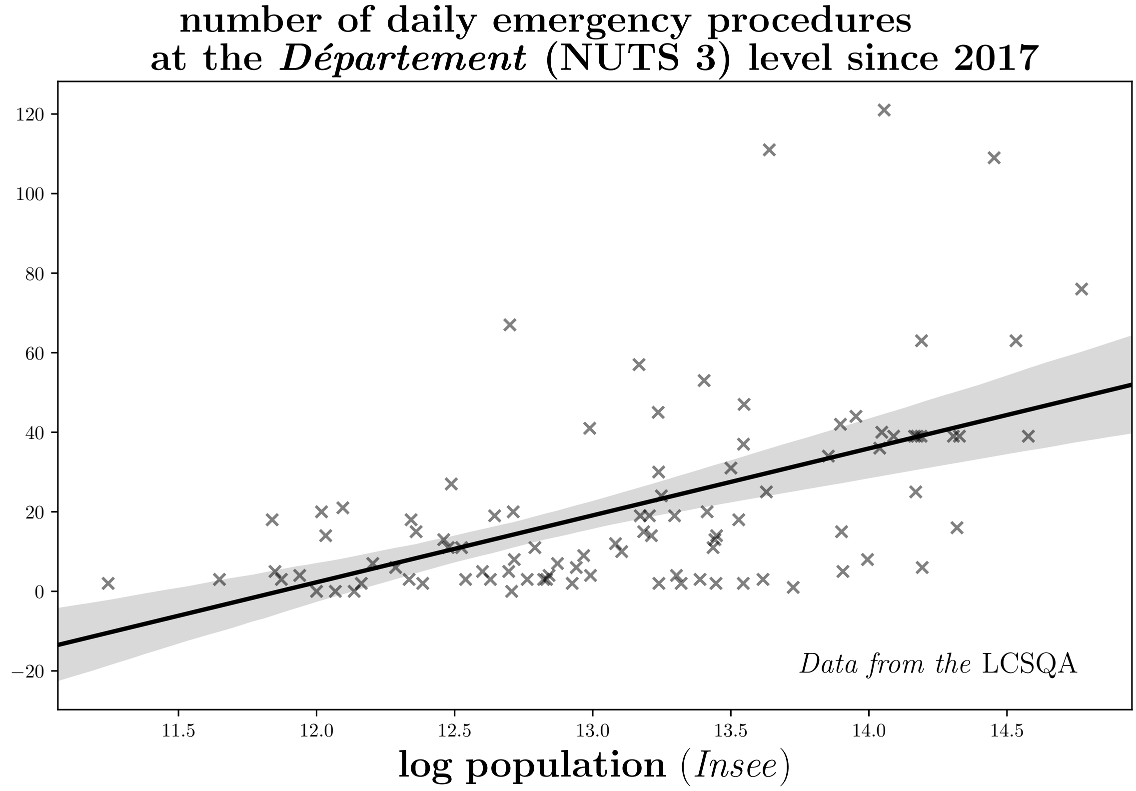


Figure 3: Number of emergency actions due to excessive measures of atmospheric pollutants

ductivity and preferences) as well as free labor mobility. The spatial equilibrium is an extension of Allen & Arkolakis (2014) where local production affects local amenities through pollution. I depart from these papers by taking a stand on a particular source of congestion externalities, namely local air pollution, and investigate its effect on the spatial distribution of economic activities. A large fraction of my assumptions is inspired by the recent universal gravity framework described in Allen et al. (2020) (AAT hereafter), including location-specific productivities. However, adding endogenous pollution as a productive input, in combination to labor, constitutes a non-trivial extension of AAT because of local reallocation of firms' expenditures across inputs.

The other major contribution of our setup is that I consider endogenous and spatially heterogenous policies. To that extent this paper also participates to the literature on place-based policies, as in Fajgelbaum & Gaubert (2020). Indeed, I assume that, when designing the national regulation on atmospheric pollutants by industrial sources, the central planner chooses the local relative levels of the policy stringency. Firms produce output using labor and atmospheric pollutants emission, and I model the relative local level of emission policy stringency as a cost imposed on the emission input. By assumption, the spatial distribution of these relative emission costs is chosen by the central planner. Alternatively, local wages are endogenous variables set in the general equilibrium. This setup is close to Suarez-Serrato

& Zidar (2016) who introduce a tax rate on firms' profit. The main difference here is that I assume an input-specific policy instrument that may induce reallocation across inputs.

When choosing the optimal local stringency level of the emission policy, the government faces several trade-offs. The model features efficiency gains from agglomeration economies, thus the central planner may be interested in fostering the spatial concentration of economic activities. In this respect, I include the "productivity advantage of large cities" that have been identified for France 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 [5]. 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.

However, I innovate by making a connection to the empirical literature that shows the negative relationship between agglomeration and local air quality. I do so by assuming that, by decreasing the city level of amenities, local emissions of atmospheric pollutants act as a congestion force. Evidence for France illustrated by Figure (1) from my paper is upheld by Carozzi & Roth (2019) and Borck & Schrauth (2019) who both find that denser cities are also more polluted (respectively for the US and Germany). A recent distinct strand of literature has shown that a local low air quality had detrimental effect on health and workers productivity (Zivin & Neidell, 2012; Deryugina et al., 2019). As a result, one can expect that, as cities grow larger (and, for a given amount of floor space, denser) the decline of local air quality slows the concentration forces down, notably by driving workers away. The 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. With a focus on France, studies by Drut & Mahieux (2015) and Leturque & Sanch-Maritan (2019) exposed that agglomeration gains were indeed dampened by the local levels of pollution.

Finally, by investigating the impact of the spatial distribution of relative stringencies of local air quality policies, I contribute to the sparse literature that explores the welfare effects of the endogenous distribution of pollution across space. Recent examples include Deryugina et al. (2020) or Desmet et al. (2021). The closest to this paper is Yamada (2020) who also considers atmospheric pollutants and shows for China that increasing the stringency of air quality regulation in a limited number of larger cities could lead to welfare gains. For tractability, I do not retain Yamada (2020)'s assumption of heterogenous workers that allows for the study of inequality in pollution exposure. I argue that the migration at no cost assumption is legitimate in the long-run, especially in a developed country. While income inequality would play a role in the heterogenous exposure to pollution in the short term to

[5] Notably, the model does not include any sorting mechanism that may partly explain higher productivities observed in larger cities.

the extent that households with lower levels of income may not be able to escape polluted cities, the costs of migrating from one commuting zone to the other are rather low in the long run. However, I go further in the welfare analysis to extent that I solve for the optimal distribution of relative stringencies of local emission policies instead of comparing ad-hoc distributions of policies inspired by empirical observations.

3 An Illustrative Framework of Spatial Externalities

To illustrate the addition of a local pollution externality to the standard spatial equilibrium model, I provide here a simplified spatial framework restricted to three spatial externalities. It features local environmental damage from pollution emissions, general congestion spillovers on local amenities and agglomeration economies[6].

3.1 Assumptions

Following the standard Roback (1982) framework, I consider a given system of cities across which a fixed population of workers is distributed in equilibrium. Let's assume that the per capita utility of the representative worker[7] in city j is given by $u_j = a_j Z_j^{-\gamma} L_j^{-\delta} c_j$, with a_j the local idiosyncratic endowment in amenities, L_j the local population of workers, Z_j the local quantity of atmospheric pollutants emitted (negatively affecting air quality, and in turn welfare), and c_j the per capita consumption of a tradable good. This setup is standard in the literature and is thought to represent the fact that workers derive welfare from consumption of goods but also from local characteristics of the locations they live in. Among these local amenities one may think of environmental attributes (rivers, forests, parks, etc.) or institutional or social installations (schools, museums, theaters, etc.) that explain why some locations are more attractive than others even after accounting for consumption. The constant elasticity congestion term, $L_j^{-\delta}$, represents the fact that when the population of a given city grows, the aggregation of workers has a detrimental effect on their welfare. This is thought to capture externalities from agglomeration that negatively affect local welfare, such as traffic congestion or noise pollution.

In this paper I also focus in a specific source of congestion, air pollution caused by local production activities, which is represented by the term $Z_j^{-\gamma}$. Given the detrimental health effect of atmospheric pollution, I assume that a worker will derive a higher welfare from living in a city less polluted, all other local characteristics being equal. The choice of

[6]These economies can potentially arise from local knowledge spillovers or labor markets pooling. They can also result from local economies of scale.

[7]Throughout the paper I assume homogenous workers.

a utility function further assumes that workers are perfectly informed on the spatial distribution of air quality indices across cities. This is a simplifying assumption that, in general, is likely to hold for France. Indeed in the recent period, local administrations in charge of air quality control have developed effective tools to inform the public on the atmospheric pollution events and on the general local air quality. However such assumption may not be valid elsewhere. A recent study by Gao et al. (2021) has found for China that in the case of particulate matter this is not the case and agents may be inaccurately informed on local pollution. Congestion effects due to other activities participating to local air pollution, such as transport for instance, are implicitly included in the general congestion term. Finally, another important implicit assumption is that air quality in city j only depends on atmospheric pollutants emissions in city j. In particular this implies that pollutants do not travel across cities. In the case of certain pollutants, this may be far from the truth. For instance particulate matter emitted in China have been found to cross the Pacific ocean. One way of reconciling this simplifying assumption with reality is to consider that, on top of actual air quality measures, workers may also react to their perceived level of local pollution that is based on the amount of local industrial activities. We come back to that assumption in the calibration section.

To encompass agglomeration economies, I further 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 productivity differences between cities. As for amenities, intrinsic productivity can vary across city due to a wide array of local characteristics that are exterior to this model and therefore considered as fixed. For instance, better local institutions or a more efficient local transport networks can explain higher level of productivity in a given city compared to another after accounting for differences in population. These exogenous endowments in amenities and productivity are the only two dimensions in which distinct cities differ.

In each city, identical firms produce the homogenous tradable good using labor and emissions of atmospheric pollutants [8] with a Cobb-Douglas production function defined by an expenditure share $1 > \alpha > 0$. 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 who observes the equilibrium. This marginal cost of polluting is thought to represent the local stringency of pollution regulations, taking into account any possible policy instrument such as technology standards, emission limits, emergency responses, etc. Note that throughout this paper, for simplicity, I only consider a representative pollutant and abstract from any tax or damage differentials that may exist between distinct polluting substances. For now, I assume that proceeds from the emission tax are locally re-

[8] Following Copeland & Taylor (2004), it is equivalent to assuming that, rather than being an input in production, emissions are a by-product. Then firms can divert a fraction of their labor force to abate emissions. The efficiency of abatement is governed by α.

distributed to workers. Assuming full employment of the local population in the industry, either for production or abatement, local expenditures $w_j L_j$ are equal to local income from labor $\frac{1-\alpha}{\alpha} t_j Z_j$[9] plus payments for emissions $t_j Z_j$[10]. In turn, local emissions are pinned down by the local labor supply and the relative factor price: $Z_j = \alpha \frac{w_j}{t_j} L_j$.

To restrict the framework to the three main externalities, I also assume that competition is perfect, trade is costless and firms are selling the tradable good on global markets so that we can normalize the price of the output to 1[11]. In turn this pins down local wages. A particularity in this setup is that only one factor price is endogenously set in equilibrium. To allow for the coexistence of aggregate factor price equalization and the exogeneity of one of the two factor prices, assuming a unique tradable sector is necessary[12].

Wages in location j are given by:

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

Substituting for wages and emissions, the indirect utility in locations j can be written as:

$$u_j = a_j b_j^{1-\gamma} L_j^{-\theta} t_j^{\frac{\gamma-\alpha}{1-\alpha}}, \quad \text{with } \theta = \delta + \gamma - \nu(1-\gamma) \tag{2}$$

The exponent θ on the population term in equation (2) illustrates the fact that, on the one hand, the pollution externality is similar to the standard congestion mechanism summarized by elasticity δ. Indeed local emissions are proportional to city size. On the other hand, because emissions are also proportional to local wages, there is an negative interaction term with agglomeration economies elasticity ν. In particular, a stronger pollution externality will dampen the concentration effect from agglomeration economies. This is explained by the fact that, when city size increases and agglomeration economies raise local productivity and local income, firms become relatively more pollution intensive, therefore increasing local emissions.

Finally, workers move freely between cities so that, in equilibrium, utility is equalized

[9]This expression come from the cost minimization implication that the distribution of expenditures across factors is given by the Cobb-Douglas share.

[10]This is equivalent to say that part of the local population works for production and the other part works for pollution abatement.

[11]Perfect competition combined with an additional assumption of costless trade leads to the spatial equalization of the aggregate factor price and in turn to the equilibrium distribution of local wages. I could choose to model a closed-economy and include general equilibrium effects: by fixing the numeraire wage, when regulation move in a given city there are spillovers in other cities. However, I restrict from including such effects to illustrate only the interplay between agglomeration, congestion and pollution spillovers.

[12]To be precise, the necessity is that, either there is a unique Cobb-Douglas expenditure share α shared by all tradable sectors, or that wages are sector-specific.

across space, and equal to a level noted $\bar{u}$. So for two cities i and j, we have:

$$\frac{L_j}{L_i} = \left(\frac{a_j}{a_i}\right)^{\frac{1}{\theta}} \left(\frac{b_j}{b_i}\right)^{\frac{1-\gamma}{\theta}} \left(\frac{t_j}{t_i}\right)^{\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}} \tag{3}$$

This ratio means that in equilibrium, for a given distribution of emission costs and if congestion effects outweigh agglomeration economies ($\theta > 0$), more workers will locate themselves in cities with higher levels of amenities. Alternatively, because local emissions are a function of income and not only of city size, whether more workers locate in more productive cities depends on the size of the pollution externality. Abstracting from agglomeration economies, wages are higher in more productive cities, so local firms are relatively more pollution intensives and pollution is higher (for a given city size). If the welfare impact of this effect is higher than the income effect, that is to say if $\gamma > 1$, then more workers will locate themselves in cites that are less productive.

In the case where atmospheric pollution from industrial activities is the only spatial externality, cities with higher levels of amenities will concentrate larger populations. This is due to the fact that workers do not internalize the effects of their location choice on local amenities. This effect is similar to regular congestion mechanisms. For a given level of local per capita consumption, cities with more amenities will attract workers up to the point where the marginal additional worker reduces local utility to the same level than the other cities. Cities with higher levels of labor productivity will provide local workers with higher incomes. This is due to the fact that firms do not internalize the effect of their employment choice [13]. Therefore these cities will also be more attractive and the effect will go in the same way as the amenity advantage. However, due to the substitutability of labor and pollutants emissions in firms' production functions, firms in cities with higher labor productivity will be more pollution intensive for a given level of emission costs. This effect will counteract the positive effect on local welfare of income from the productivity advantage of certain cities. However, as I show below, this negative effect from input reallocation can be expected to be weaker than the income effect. As a result, cities with higher levels of labor productivity will also concentrate larger populations.

3.2 Optimal Environmental Policy

Equations (1) and (3) describe the distribution of income and population across cities. Assuming that the total population to be equal to one allows me, in turn, to compute the common level of welfare $\bar{u}$ reached in equilibrium:

[13] Throughout the whole paper, I assume that all firms are input price takers.

$$\bar{u} = \left[\sum_{j \in C} a_j^{\frac{1}{\theta}} b_j^{\frac{1-\gamma}{\theta}} t_j^{\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}} \right]^{\theta} \tag{4}$$

Equation (4) represents the social welfare function that a national central planner maximizes. However, given that my setup abstracts from any general equilibrium effects, I have to choose a numeraire for the emission cost. In practice I set the constraint that the average emission cost is equal to one over the number of cities[14]. Noting $\{t_j^*\}$ the vector of emission costs that solves the problem $\max_{t_j} \bar{u}(t_j)$ so that $\sum_{j \in C} t_j = 1$, I find, that for any pair of cities :

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

The second order condition to ensure that the optimum is a maximum implies the following set of conditions on parameters γ, δ and ν:

$$\forall j \in C, \left(\frac{\gamma-\alpha}{1-\alpha} - \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}\right) L_j < 1 - \frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha} \tag{6}$$

From equation (5) I observe particular that when $\delta = \nu = 0$, that is to say when pollution is the only externality the central planner is correcting, I have $(t_j^*/t_i^*)^{\frac{\alpha}{1-\alpha}} = (a_j/a_i)^{\frac{1}{1-\gamma}}(b_j/b_i)$. This means that internalizing the impact of pollution implies higher emission costs in more productive cities. The second order condition implies that $\gamma < 1$ at least. Therefore, it also implies higher emission costs in cities with more amenities.

Without any correction (that is to say when a uniform emission cost equal to $1/C$ is applied in all cities), equation (1) shows that wages are higher in cities high higher labor productivity. However, increasing the local emission cost relative to other cities decreasing local wages. Because equation (5) indicates that the optimal emission cost is relatively higher in cities where labor is more productive, the correction dampen the local wage boost from the local intrinsic productivity. Furthermore, combining equations (1) and (5) leads to a full cancelation of the advantage of more productive cities.

Concerning the distribution of population, equation (3) shows that population concentrates in cities that are more productive and have more amenities. One can also see that, when $\gamma > \alpha$, for a given level of amenities and productivity, increasing the local level of emission cost relative to other cities attracts more worker and increases the local population. Equation (5) shows that the optimal correction applies higher emission costs in cities

[14] Another option would be to fix the total amount of emissions.

more productive and with more amenities; As a result, the optimal correction reinforce the concentration pattern that is observe in the absence of any correction.

One could argue that the environmental social planner's objective should only encompass the pollution externality. Indeed, the assumption that the central planner maximizes the full social welfare function implies that the optimal policy will also correct all externalities, not only the pollution one. From equation (5) one can further observe that in the absence of the pollution externality ($\gamma = 0$), a central planner who maximizes the social welfare function still applies heterogenous optimal emission costs across cities. This is because agglomeration economies and congestion effects are externalities that also result in inefficient spatial distribution of workers and income in equilibrium.

Starting from an emission cost common to all cities (an equal to one given our choice of numeraire), implementing the spatial distribution of relative emission costs described by (5) leads to workers migration across cities. Combining equations (3) and (5) leads to:

$$\frac{L_j^*}{L_i^*} = \left(\frac{L_j^{t=1}}{L_j^{t=1}}\right)^{\frac{\theta}{\theta - \frac{\gamma-\alpha}{1-\alpha}}} \tag{7}$$

In particular, when $\delta = \nu = 0$ the exponent on the right-hand side term is $\frac{\gamma}{1-\gamma}\frac{1-\alpha}{\alpha}$. As a result, when $\alpha < \gamma < 1$ the optimal policy relocates workers to large cities (which, in turn, become even larger).

Anticipating on results from section (5) where I calibrate parameters γ, δ, ν and α using French data, I illustrate predictions for the following set numerical values: $\gamma = .25$, $\alpha = .01$, $\delta = .1$ and $\nu = .01$. From these numbers one see that $\theta = .34$ and $\frac{\gamma-\alpha}{1-\alpha} = .24$. From equation (5), it leads to positive elasticities of optimal emission costs with respect to local amenity and productivity endowment respectively equal to approximately $.1$ and $.08$. In fact, our calibration exercise from section (5) show that congestion elasticities, γ and δ, are generally one order of magnitude larger that the elasticity of agglomeration economies ν and the emission elasticity α.

As illustrated per equation (5) when externalities other than pollution are included in the spatial equilibrium, defining the central planner's problem as the maximization of the social welfare function means that the resulting optimal policy will not only correct the environmental damage externality but all inefficiencies.

In fact, workers do not internalize the impact of their location choice on local air quality: as a result cities with lots of amenities are too small. Firms do not either internalize the impact of their emissions and in turn more productive cities are too small. An optimal emission cost reduces the congestion from pollution in cities with lots of amenities and relocate input use to labor in more productive cities so that firms their are less pollution intensive. All actors being atomistic, the equilibrium features three types of externalities, that are, to some extent, independent. First, workers do not internalize the congestion mechanism through which their location choice affects local amenities. Then firms do not either internalize the effects of their employment choices on local productivity through agglomeration economies. Finally, neither firms nor workers internalize the effects of location and employment choices on local air quality through atmospheric pollutants emissions. The optimal Pigouvian tax rate equalizes the marginal damage across locations. The instrument is on emissions, but not all externalities have to do with emissions.

I move now to the general specification of the spatial equilibrium. In particular I relax assumptions on perfect competition and costless trade. Therefore the equilibrium features other externalities not included in the illustrative framework. In particular one can expect costly trade and monopolistic competition[15] to have concentration effects. Also, due to general equilibrium effects there may be between cities pollution haven effect, through which increasing emission cost in a city may lead to reallocation of emissions in cities where emissions are relatively cheaper.

4 A General Spatial Model of Polluting Activities

4.1 Setup

I extend the model presented in the previous section in order to bring it closer to reality. in particular, I now assume that there are several industrial sectors that are producing goods and emitting pollution. These sectors are heterogenous in their production technologies, some are cleaner than others and consumers have distinct preferences across this sectors. Within each sector, there is now a continuum of differentiated varieties of each sector specific good that are produced by atomistic firms. I further assume that firms are heterogenous in their idiosyncratic productivity level and that within each sector they are competing monopolistically. Finally, trade between cities is costly so firms and consumer face iceberg costs on top of production cost when exporting or importing to other cities. In the following I describe these alternative specifications. The other assumptions remain the same as in the previous section.

As in the illustrative framework of spatial externalities, the representative worker living in city i has the following utility function:

$$u_j = a_j L_j^{-\delta} Z_j^{-\gamma} c_j \tag{8}$$

[15]Through the "love of variety" mechanism.

where c_j is the per capita consumption of the now composite industrial good. The industry is composed of S distinct sectors, such as c_j is the following two-tier 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} is the continuum of varieties produced in sector s in city i. I call 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 common to all workers across the country. I note P_j the aggregated price index of the industrial good in city j. I have that $P_j = \prod_{s \in S} P_{js}^{\beta_s}$ where P_{js} are city and sector specific price indices defined as:

$$P_{js} = \left(\sum_{i \in C} \int_{\omega \in \Omega_{is}} p_{ijs}(\omega)^{1-\sigma_s} d\omega \right)^{\frac{1}{1-\sigma_s}} \tag{10}$$

with $p_{ijs}(\omega)$ the unit price of variety ω produced in city i in sector s and delivered in city j. Noting L_i and w_i the population and wage, I assume that there are no friction on local labor markets, meaning that wages equalize between all sectors of production. In city i I can write the indirect utility function as:

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

In this equation one can observe the mechanisms through which local emission policies can affect welfare. A policy that would enhance local air quality (by decreasing emissions Z_i) would increase welfare. However if the policy decreases local real wages (the ratio of local wages and price indices) then it would negatively affect welfare. Finally, congestion (meaning the concentration of more population in the same area) also negatively affects welfare.

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

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

Following the literature, workers agglomeration creates productivity gains and I define the Pareto base productivity to be $\forall i \in C, b_{is} = b_i b_s L_i^{\nu_s}$. This functional form is equivalent to the assumption of agglomeration economies made in the previous section where the intrinsic labor productivity of each city $\{b_i\}_{i\in C}$ is multiplied by a factor L_i^{ν}. This factor basically shifts the productivity distribution to the right side, the more so if the city is large. On top of these agglomeration gains I assume that each sector base productivity $\{b_s\}_{s\in S}$ is common across cities. In the calibration section I empirically identify the vector $\{b_i\}_{i\in C}$ so that the model fits the observed distribution of wages across cities. The rational behind this vector comes from the empirical observation that local population heterogeneity does not explain the whole distribution of firms mean productivity across locations (**?**). Across location productivity heterogeneity may arise from match-quality, sensitivity to transportation costs, factor and input market requirements or any characteristics not included in our setup. Our specification is a way of taking this fact into account.

Production remains very close to assumptions made in the illustrative framework. Each firm then can produce a specific variety of the sectoral 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 to be sold on city j's market. Variable $a(\phi)$ corresponds to the fraction of the firm's labor force used to abate emissions cause by the production process. The production process releases pollution as a by-product according to the following function:

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

This choice of production and pollution function closely follows Copeland & Taylor (2004) and Shapiro & Walker (2018) and is standard in the literature as it accommodates various interpretation for the pollution emissions due to production. More importantly it corresponds to a Cobb-Douglas production function, with labor and emission as factors:

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

Between cities trade is costly. Firms in city i have to pay an origin-destination specific variable cost to export to city j. I model this cost τ_{ij} as an iceberg cost. I assume that the matrix $\tau = \{\tau_{ij}\}_{(i,j)\in C^2}$ represents all trade frictions. However the model can easily be extended to accommodate for origin-destination fixed trade costs as in Shapiro & Walker (2018). Productivity distributions of firms are then left-truncated by endogenous zero-profit productivity cutoffs. In this case, only the most productive firms in a given city are selling goods in all other cities. Following the Cobb-Douglas production function and CES preferences, the unit price of a variety produced in city i in sector s by a firm with productivity ϕ and delivered in city j is:

$$p_{ijs}(\phi) = \frac{\sigma_s}{\sigma_s - 1}\tau_{ij}\frac{c_{is}}{\phi^{1-\alpha_s}}, \text{ with } c_{is} = \kappa_{\alpha_s} t_i^{\alpha_s} w_i^{1-\alpha_s} \tag{16}$$

with $\kappa_{\alpha_s} = (\alpha_s^{\alpha_s}(1-\alpha_s)^{1-\alpha})^{-1}$. Revenues of a firm with productivity ϕ in city i can be written as:

$$r_{is}(\phi) = \sum_{j \in C} p_{ijs}(\phi) q_{ijs}(\phi) \tag{17}$$

and profits can be written as $\pi_{is}(\phi) = r_{is}(\phi)/\sigma_s$.

4.2 Equilibrium

Given the economy's setup, one can differentiate between a market equilibrium (ME) where cities' populations are fixed (no migration) and a full equilibrium (FE) which features between-city migration of workers. The ME is characterized by $C(2S+1)$ endogenous variables: $\{M_{os}\}_{o\in C, s\in S}$, the masses of firms in each pair of sector and city, $\{P_{ds}\}_{d\in C, s\in S}$, the sector and city specific price indices and $\{w_o\}_{o\in C}$ the city wages. These variables are pinned down by the same number of equations (the system is identified): CS free-entry conditions, CS goods market clearing conditions and C city labor market clearing conditions.

Free-entry conditions require that in a given city o and a given sector s, entrepreneurs enter production until their expected profits equal the cost $c_{os} \times f_s^e$ of drawing a productivity from the local Pareto productivity distribution. I choose to express fixed entry costs in the aggregate factor price as in Bernard et al. (2007). Thus, I have for every pair of city and sector the relation:

$$\int_\phi \pi_{os}(\phi) dG_{os}(\phi) = c_{os} f_s^e \tag{18}$$

Under my assumptions, these free-entry conditions can be written as:

$$(t_o^{\alpha_s} w_o^{1-\alpha_s})^{\sigma_s} b_o^{-(1-\alpha_s)(\sigma_s-1)} L_o^{-\nu_s(1-\alpha_s)(\sigma_s-1)} = K_s \sum_{d\in C} \tau_{od}^{1-\sigma_s} w_d L_d P_{ds}^{\sigma_s - 1} \tag{19}$$

with $K_s = \frac{\eta\beta_s\theta_s}{\theta_s-(1-\alpha_s)(\sigma_s-1)} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \kappa_{\alpha_s}^{-\sigma_s} b_s^{(1-\alpha_s)(\sigma_s-1)}$

The goods markets clearing conditions state that local expenditures must match imports and can be written as:

$$P_{ds}^{1-\sigma_s} = K_s \sum_{o \in C} \tau_{od}^{1-\sigma_s} \tilde{M}_{os} (t_o^{\alpha_s} w_o^{1-\alpha_s})^{-\sigma_s} b_o^{(1-\alpha_s)(\sigma_s-1)} L_o^{\nu_s(1-\alpha_s)(\sigma_s-1)} \tag{20}$$

with $\tilde{M}_{os} = \frac{\sigma_s}{\eta\beta_s} f_s^e M_{os} c_{os}$ for clarity[16].

The local labor markets clearing conditions state that the sum of employment over all sectors must account for the total local population. It can be written as:

$$w_o L_o = \sum_{s \in S} \beta_s \tilde{M}_{os} \tag{21}$$

Assuming that trade is costly, that is for any pair of distinct cities (o, d) I have $\tau_{oo} < \tau od$, then, as long as $\sigma_s > 1$) for each sector, matrices $\tau^{1-\sigma_s}$ are strictly diagonally dominant meaning that they are invertible. By inverting these matrices, I can combine these three sets of equations to characterize the market equilibrium with only one equation per city, linking the vector of local wages to the vector of local populations:

$$\begin{aligned} w_o L_o = \sum_{s \in S} & \beta_s t_o^{\alpha_s \sigma_s} w_o^{(1-\alpha_s)\sigma_s} b_o^{-(1-\alpha_s)(\sigma_s-1)} L_o^{-\nu_s(1-\alpha_s)(\sigma_s-1)} \\ & \times \sum_{d \in C} \gamma_{ods} \frac{w_d L_d}{\sum_{i \in C} \gamma_{ids} t_i^{\alpha_s \sigma_s} w_i^{(1-\alpha_s)\sigma_s} b_i^{-(1-\alpha_s)(\sigma_s-1)} L_i^{-\nu_s(1-\alpha_s)(\sigma_s-1)}} \end{aligned} \tag{22}$$

where $\{\gamma_{ods}\}_{o,d,s \in C^2 S}$ are the terms of the inverse of the $\tau^{1-\sigma_s}$ trade friction matrix. The other endogenous variables that characterize the STE follow once wages are pinned.

The full equilibrium is defined by $2C(S+1)+1$ endogenous variables pinned down by the same number of equations (the system is identified). Compared to the market equilibrium, I add to the endogenous variables $\{L_o\}_{o \in C}$, the city populations, and $\bar{u}$ level of local utility (common to all workers across cities under the free migration hypothesis). The additional sets of equations are the C free migration conditions and the national labor market condition. The free migration assumption states that workers can move freely across cities to maximize their utility. This leads to an equalization of the utility level across cities, to a common level noted $\bar{u}$. The national labor market states that the total number of workers in the country is fixed to an exogenous quantity N. Combining these two equations leads to another equation linking local wages to local populations:

[16]This $\tilde{M}_{is}$ aggregate represents total revenues of firms from city i and sector s. To get $\tilde{M}_{is}$, the revenues are normalized by expenditure shares beta and markups using sigma.

$$L_j^{\delta} = \frac{a_j Z_j^{-\gamma} \frac{w_j}{P_j}}{\left[\sum_{i \in C} \left(a_i Z_i^{-\gamma} \frac{w_i}{P_i} \right)^{\frac{1}{\delta}} \right]^{\delta}} \tag{23}$$

As in the illustrative setup, I assume that migration is possible at no cost for workers. Without any doubt this has a large impact on the results from the analysis I conduct in this paper. More specifically, a recent literature has focused on the unequal exposure of workers to atmospheric pollution and the fact that workers who have access to lower incomes are less able to escape pollution and move to places with cleaner places. This empirical fact is not featured in my theoretical framework to the extent that I assume zero cost of migration and only one type of worker. The reason for this choice comes from my objective to investigate across commuting zones policies. Unequal exposure to atmospheric pollution are often identified at the city level with high-skill workers being able to live in cleanest neighborhood of the city and commute and low-sill workers living in more polluted areas closer to local polluting activities. Commuting is an important within-city margin of adjustment to pollution exposure, I argue that it matters less for across cities distribution of population. Also, my setup is not dynamic and to that extent features long-term adjustments of population, income and pollution distributions across space. In the case of France, it is legitimate to assume zero-cost of domestic migration in the long-term.

In turn, the level of utility is given by the national labor marker clearing condition.

$$\bar{u} = \left[\sum_{i \in C} \bar{A}_i Z_i^{-\gamma} \left(w_i / P_i \right)^{\frac{\eta}{1-\eta}} \right]^{1-\eta} \tag{24}$$

with the local quantity of emission Z_i equal to:

$$Z_i = \sum_{s \in S} \beta_s \alpha_s \frac{\tilde{M}_{is}}{t_i} \tag{25}$$

As my model constitutes an non-trivial extension of the "universal gravity" framework (applied to an economic geography setup) presented in Allen et al. (2020) the extrapolation of their results on existence and uniqueness is not straightforward. One main distortion with this kind of model comes from the pollution intensity heterogeneity across sectors. However, this is an assumption that I should not release, as it may capture interesting reallocation effects between sectors in reaction to regulation shocks. Such reallocation would translate into sector share of income that are different across cities due to the spatial heterogeneity in the local price of pollution. Another important distortion is the introduction of the environmental damage function that is not a direct function of local population.

4.3 Definition of the Mean-Preserving Optimal Spatial Distribution of Emission Costs

I assume that the distribution of emission costs $\{t_i\}_{i\in C}$ is set by the central government of the country. More specifically, it represents the potentially heterogenous effect of national policies on local emission costs. For instance, setting a concentration limit for the whole country may correspond to a vector $\{t_i\}_{i\in C}$ with different emission cost for different cities. In this sub-section I define the optimal distribution of emission cost across cities and, in practice, I assume that the central planner has the means of choosing this distribution.

Before the definition of this optimal policy, an important observation about the equilibrium defined above is that it displays homogeneity of degree zero with respect to the level across cities of amenities, productivities, wages, population and emission costs. In practice, multiplying the vectors of city- wages, populations and emission costs does not affect the distribution of these variables across cities. As a result my framework do not allow to empirically identify the average effect of wages, population and emission costs on welfare, but only the effects of there relative distribution across cities. Another way to see this is that for each variable of the model I need to fix the numeraire which cannot be identified by the equilibrium condition. In the empirical section I set the average wage and productivity level to be equal to one and the total quantity of amenity and population (summed over all cities) to be equal to one. The fact that I cannot say anything about the numeraire also arises from my assumptions on production. I model a representative pollutant and estimate production parameters using constructed values of emissions that do not fully coincide with true emission values. Given this observation, I define the mean-preserving optimal distribution of emission policies taking the average level of emission cost $\bar{t}$ as fixed. This optimal policy is then defined as:

$$\max_{\{t_i\}_{i\in C}} \bar{u}(t_1, ..., t_C) \text{ s.t. } \sum_{i\in C} t_i = \bar{t} \times C \tag{26}$$

I argue here that welfare analyses under this mean-preserving constraint are still pertinent to the extent that, when moving from one spatial distribution of emission costs to another distribution that has the same average across cities, the percentage change of welfare do not depend on this average level.

5 Calibration of the Model

I calibrate the model presented in the previous section using a set of French data. First, I consider the set French commuting zone, as defined in 2010, as the set of cities in which the national population of workers live. This set of zone constitutes a partition of the full French metropolitan territory and is statistically defined as an area where people both work and live. Because I do not introduce commuting in the framework of my analysis this empirical definition of cities seems particularly pertinent.

Then the equilibrium is fitted using aggregated variables as observed in 2012. I combine several of data sources described in the first paragraph of this section to calibrate the model's parameters. The model's calibration mainly is composed of the following three steps:

- I calibrate the utility function (inverse labor supply elasticities γ and δ) using a regression of local mean wages on local air quality measures and local labor supply. I build specific instruments to overcome identification threats.
- I recover sector-specific elasticities of substitution $\{\sigma_s\}_{s\in S}$, Pareto shape parameters $\{\theta_s\}_{s\in S}$ and emission elasticities $\{\alpha_s\}_{s\in S}$. I use administrative firm-level tax declarations and plant-level energy use survey, from which I compute measures of emissions.
- I compute a matrix of road travel time between French commuting zones and use it as a proxy for the trade cost matrix.

Finally, I use the parameters calibrated in the previous step to retrieve vectors of local idiosyncratic productivities $\{b_i\}_{i\in C}$ and amenities $\{a_i\}_{i\in C}$ and the vector of local emission policy stringency $\{t_i\}_{i\in C}$. To do so, I exactly fit the model on the observed vectors of local wages, populations and industrial emissions using equations (22), (23) and (25).

5.1 Data

In this subsection I provide a quick description of the datasets used in the different calibration steps described in the following subsections. Firm-level datasets come from confidential French administrative data for the universe of French firms. 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), and the activity sector.

The plant-level panel rests on energy consumption data 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 several energy types and fuels (quantity consumed and expenditures), at the plant level. The types of energy 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 survey also provides 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 convert energy consumption into pollutant emissions using emissions factors from the Ominea database (CITEPA, 2020). 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[17]. Table (1) displays Pearson correlation coefficients along with statistical significance. It appears that my construction emissions values are doing a fair job in representing the actual emissions of these large plants.

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

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

For the estimation of the environmental damage, I use data on local emissions of pollutants. The National Spatial Inventory (INS, 2020) is a publicly available dataset[18] reporting emissions of around 40 pollutants from natural and anthropogenic sources distributed across municipalities. It is produced by the French environmental administration and data is available for 2004, 2007 and 2012. Basically, national emission levels are distributed across space based large emission sources emissions declaration (for instance declarations made under the Large Combustion Plant EU directive or in the E-PRTR) and on local activity data for smaller sources. We downloaded the data and aggregated municipalities in employment zones for PM10, PM2.5, NOX, SO2 and COVNM emitted by the manufacturing sectors. Activities including are combustion in the manufacturing industry and production processes. These corresponds to codes 3 and 4 in the Selected Nomenclature for Air Pollution (SNAP) which the standard European Union nomenclature for polluting activities.

[17] 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 make the estimation of sector-specific emission elasticities impossible.

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

Albeit likely more precise the National Spatial Inventory is not available for many years. For the estimation of the inverse labor supply elasticities I rely on geographic emissions data from the Emissions Database for Global Atmospheric Research (EDGAR) from the European Commission's Joint Research Center (JRC). On top of greenhouse gases emissions EDGAR also provides annual gridded emissions of atmospheric pollutants disaggregated across polluting sectors at 0.1x0.1 degree resolution. Data is publicly available of the JRC's dedicated website [19]. Information on the methodology can be found in Crippa et al. (2018).

I provide information on the spatial distribution of pollution across cities using data on local concentration of air pollutants from the CHIMERE chemistry-transport model (Chimere, 2020). This product provides a gridded dataset providing air concentration for a set of atmospheric pollutants. I aggregate mean concentration at the commuting zone level. In practice I observe daily concentration of PM10, PM2.5, NO2 and O3 on a grid covering the French metropolitan territories. On average the grid resolution is around one kilometer meaning that pollution data is available for several thousand of grid points across France. Data on local atmospheric pollution is the output of a simulation of a chemistry transport model that uses information on meteorological variables and pollutant emissions for various sources. As a robustness test of the inverse labor supply elasticities I also instrument local air quality using on NASA's MERRA-2 products providing gridded data on wind direction and speed data (see below).

Finally, for in the equilibrium inversion procedure (see below) I use a 2012 cross-section dataset of wage levels and labor for all French employment zone. I build this dataset using aggregate *Insee* data.

5.2 Calibration of the Inverse Labor Supply Elasticities

Given the free migration hypothesis, the model predicts the following log-linear relationship between wages, utility, amenities, price indices, local labor supply and emissions:

$$\log w_{it} = \log \bar{u}_t + \log a_{it} + \log P_{it} + \delta \log L_{it} + \gamma \log Z_{it} \tag{27}$$

Introducing a time variable, city-specific endowments in amenities and productivity may vary over time. The model presented in the previous section is not dynamic, however in reality local city characteristics may depend on time for reasons that are external to the mechanisms that I have chose to include in the model. For instance, enhancements to local transport networks could boost local productivity of certain cities. Local policies that support the creation of parks, or museum would also boost local amenities. As a result, en-

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

dogenous city-specific variables (wages, prices, populations and emissions) also vary across time. As explained in the theory section I assume that welfare is constant across space, so in can only vary across time. I build on this relationship to calibrate elasticities δ and γ. In practice, I only observe wages, populations and emissions for several periods but not local price nor local amenities indices. Thus I could estimate the following equation using OLS:

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

Year fixed effects eliminate annual shocks that are common to all cities. From equation (27) one can see that year fixed effects are controlling for annual shocks on the common utility level reached by all workers. I also include city fixed effects to eliminate the influence of city specific fixed natural advantages (idiosyncratic productivity or amenity). According to equation (27), the error term ϵ_{it} in the previous equation corresponds to city-year specific shocks that contains shocks on local amenities a_{it}, or on local price indices P_{it}. In the model's section I show that local price indices are functions of the vector of local productivities and of the trade cost matrix (and of the other endogenous variables). As a result, city specific shocks on local productivities or changes in city specific trade cost matrix will induce shocks on local price indices. Given that wages, populations and emissions are functions of these shocks, the simplest OLS estimation strategy will produce biased estimates. For instance, any positive shock on local amenities of a given city will positively affect the local population and potentially local emissions. Therefore the OLS identification assumption is not likely to be fulfilled. To overcome this estimation issue, I build instrumental variables for both local labor supply and local emissions. However as stated by Bartelme (2018) all observed variables are endogenous to shocks to the fundamentals (in my case productivities and amenities) so finding pertinent instrument is a challenge.

In practice, I follow Bartelme (2018) who extends the Bartik (1991) approach which is common in urban and regional economics. The idea is to approximate city specific growth rate in workers' populations and emissions using national growth rates of theses variables in disaggregated industries and interact them with city-specific shares in an initial period. The resulting shock is therefore a proxy that should be independent of unobserved shocks to city labor and emissions growth rates.

Following this idea I build the two instruments as:

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

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

Variables $\tilde{L}_{it}$ and $\tilde{Z}_{it}$ are respectively the instruments for L_{it} and Z_{it} in equation (28). As shown in their respective definition, these instruments are independent of city-year specific unobservable shocks (on local amenities, productivity or trade costs) and are only functions of national shocks, which by definition are common to all cities and therefore do not threaten identification when including year fixed effects, and of city-specific initial sector shares in employment and emissions, which by definition remain constant across time and therefore do not threaten identification when including city fixed effects. As thoroughly detailed in Goldsmith-Pinkham et al. (2020), a recent extensive analysis of Bartik 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.

From the theory section, we know that the local sectoral employment share in city i and sector s is equal to $\frac{L_{is}}{L_i} = \beta_s \frac{\tilde{M}_{is}}{w_i L_i}$. Because of trade linkages, the right hand side of this expression is a function of local characteristics of all cities. Especially, when trade costs are heterogenous across cities, meaning that some cities are better connected than other, the model implies that some sectors will benefit from it and as a result the shares will depend in local characteristics. If one expect city-specific shocks on local characteristics to be correlated to the level of said characteristics (for instance, that amenities would grow faster in cities where they are already large) then unobserved shocks could be correlated with sectoral shares in the initial period. However I argue that this threat is minimal as trade costs are rather low and solving the equilibrium I observe that sectoral shares are mainly functions of the vectors of sector parameters $\{\beta_s\}_{s\in\Omega_S}$ and $\{\alpha_s\}_{s\in\Omega_S}$.

The level of sector disaggregation plays a role in the performances of the instrument. Indeed there a trade-off in the number of distinct employment or emission activities that I introduce to construct the instruments. In the extreme case where I would only introduce one activity, the initial share would be equal to one and constant across cities. Thus it would not be correlated to city specific unobservable shocks. However, because it would be a over simplification of the sectoral structure of employment and emission one can expect that such instrument would only be a poor predictor of the actual employment or emissions. On the contrary, the more precision I introduce in the sectoral structure to build the instrument, the better the correlation between the instrument and the endogenous variable. However, such structure increases the risk of initial shares being correlated with unobservable shocks.

Table (2) reports the first stage results.

This estimation strategy build on annual firm-level data on wages, employment and activity codes, on plant-level data on localization across cities (information on the commuting zone where plants belonging to a given firm are located) and on data on annual industrial emissions aggregated at the commuting zone level.

	(1)	(2)	(3)	(4)	(5)	(6)
	$Employment$	$PM2.5$	$PM10$	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 2: First stage results for IV estimation of δ and γ

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 belong. Based on these two codes I merge the firm and plant level panels. Thus for each plant, for which I know the commuting zone where it is located, I also have a value of the mean wage paid at the firm level. I aggregate across this panel at the commuting zone level and compute a mean value of 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. On the right hand side of the estimating equation only depends on the total local employment across sectors but the instrument defined in equation (29) build on the sector local and national disaggregation across sectors. The French Activity Nomenclature distinguish 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, hostels 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 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 overlay 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 obtained a 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 choose to exclude some polluting activities that obviously do not enter my framework (aviation, road transportation and shipping).

Table (3) displays the outcome of δ and γ estimations.

	(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 3: Calibration of inverse labor supply elasticities δ and γ

As a robustness test of this calibration test, I provide an alternative estimation of γ using pollution concentration data from CHIMERE. Results are reported in the appendices.

5.3 Calibration of the Industry Parameters

In this paragraph I present the calibration of the industry specific demand and production parameters as well as the elasticity of agglomeration economies. This procedure takes several steps.

The first step closely follows the approach of Shapiro & Walker (2018) to calibrate parameters $\{\beta_s, \sigma_s, \theta_s\}_{s\in S}$. 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 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} is sales made by firm i from sector s and $\gamma_{1s} = -\frac{\theta_s}{\sigma_s - 1}$. Because Pareto distribution better fit the right part of the productivity distribution, I restrict the sample to the upper decile of the firms sample.

The next step is to estimate the sector-specific emission elasticities $\{\alpha_s\}_{s\in S}$. 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, 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 elasticities using the following empirical specification:

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

To take care of the transmission bias that bias the OLS strategy, I choose to instrument plant-level emissions z_{ics} using exogenous fuel-specific energy price variation. Basically, I use 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.

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

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. An advantage for this emission elasticity estimation strategy is that it only requires to observe plant-level inputs. Given our data constraints, a strategy building on revenue data would force me to estimate an equation at the firm-level. Because I do not observe emissions at the firm-level, I would need to restrict my estimation sample to firms that are composed of only one plant. Such restriction would threaten my ability to precisely estimate sector-specific emission elasticities.

Table (4) reports the results from this estimation for different pollutants when industrial sectors are pooled. The first-stage using the fixed-weight energy price index yields significant estimates that are not surprising, on average increasing by one percent polluting fuels prices causes a one percent decrease of atmospheric pollutant emissions. Note that both the endogenous variable and the instrument are constructed from the same data (fuel consumption at the plant level) 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. Across pollutants, the second stage yields similar values for the pooled sectors emissions intensity and one can take away an average value of 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 higher 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 discrepancy. However, the relative distribution of emission intensities across industrial sectors that I obtain and present in Table (5) is coherent with the one obtained by Shapiro & Walker (2018).

	(1)	(2)	(3)	(4)	(5)
	log (*emiPM25*)	log (*PM10*)	log (*NOx*)	log (*CO*)	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 4: Two-stages calibration of emission elasticities α (with pooled industrial sectors)

In the last step, to calibrate sector specific elasticities of agglomeration economies $\{\nu_s\}_{s\in S}$ I regress city specific mean firm productivity against city size. To estimate TFP, I simply regress firm-level log-value added on city times sector specific fixed-effects. Given my setup for production and demand 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, the coefficient estimated is precisely ν.

Table (5) displays the result from these calibration steps.

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 5: Estimated parameters

5.4 Calibration of the Trade Costs Matrix

The matrix of trade costs between cities is key in my model as it shapes the distribution of activity between locations. When bilateral trade costs between two cities decrease, then the system formed by these two cities become closer to a single-city system, in which all variables reach a common level. In terms of evolution, if bilateral trade costs between a large city and a small city decrease, then because wages are ex ante higher in the large city and lower in the small city, they are going to decrease in the large city and increase in the small city. This is a market effect. Trade costs do not directly affect the equilibrium through migration, only wages, prices and pollution do. Decreasing wages in the large city are going to a rise in local emissions and in local prices. All three market effects (decreasing wages, increasing emissions and prices) are going to decrease the overall attractiveness of the large city. Conversely, the small city is going to become more attractive. As a result, the LTE effect of decreasing trade costs between these two cities is going to be a de-agglomeration effect.

However I are limited in the ways that I can approximate trade costs as I do not observe bilateral trade flows between cities so I cannot directly estimate a gravity equations to recover bilateral friction terms. Instead, I follow Yamada (2020) and rely the methodology used in Baum-Snow et al. (2020) who assumes the following concave relationship between iceberg bilateral trade costs and bilateral travel times:

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

As noted by Baum-Snow et al. (2020) this approach is speculative so I conduct some sensibility analysis on parameters ρ and ξ[20]. Note that although some studies on the spatial distribution of activity also investigate the welfare effects of transport network development, in my setting I consider that the network is exogenously fixed. 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.

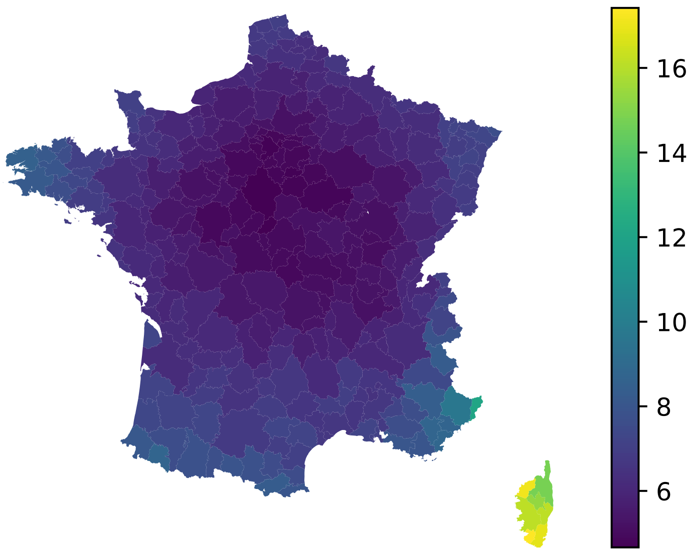


Figure 4: Average travel time across all commuting zone (in hours)
Travel times were computed using the Open Source Routing Machine (OSRM) API Python client osrm-py

Figure (4) displays the average travel time by road across all commuting zone. Without a surprise, central areas are better connected to the rest of the country.

5.5 Recovering Local Characteristics

Once all parameters are calibrated, I can use the observed equilibrium to retrieve the distribution of idiosyncratic characteristics (amenities and productivity) across cities as well as the distribution of relative air quality policy stringency. Basically, upon observing the three vectors of populations, wages and emissions ($\{L_i, w_i, Z_i\}_{i \in \Omega_C}$) I numerically solve the

[20] In Baum-Snow et al. (2020), $\rho = 0.004$ and $\xi = .8$.

non linear system of equilibrium conditions defined in the previous section by equations (22), (23) and (25). to find the three vectors of amenities, productivity and emission costs ($\{a_i, b_i, t_i\}_{i \in \Omega_C}$). Such strategy is akin to a revealed preferences methodology as is equivalent to looking for the distribution of local characteristics that would explain the observed distribution of workers, wages and emissions.

I use data from the INSEE to build the observed equilibrium vectors of population and wages and on data from the spatialized national inventory for the distribution of emissions. Both data sources report values for the year 2012 at the municipality level. I aggregate values at the commuting zone level using the 2010 definition of commuting zones which defines a correspondence table between municipality and commuting zone cades. I then 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.2.1) provides details on the sources of data for observed equilibrium distributions of wages, populations and emissions across French cities. Appendix (A.2.2) provides descriptive statistics on the observed equilibrium distributions of wages, populations and emissions across French cities. Finally appendix (A.2.3) provides descriptive statistics on the computed exogenous distributions of amenities, productivities and emission costs across French cities.

6 An Application to France

6.1 Local Endowments in Idiosyncratic Characteristics

Results from the last step described in the calibration section are displayed by Figure (5). By inverting the equilibrium equations using observed vectors of population L, wages w, and emissions Z, I recover the corresponding city characteristics: amenities a, productivities b and emission costs t. Therefore I am able to characterize how the spatial distribution of relative emission costs correlates with the distribution of city sizes (sub graph (a) which is similar to the lowest subgraph of Figure (1)) but also with the spatial distribution of amenities (subgraph (b)) and productivities (subgraph (c)). In Figure (5) I plot the city-level emission costs obtained through the equilibrium inversion as a log-linear function of the observed city sizes and the amenities and productivities also computed through the equilibrium inversion.

The first observation from Figure (5) is the heterogeneity of emission costs across

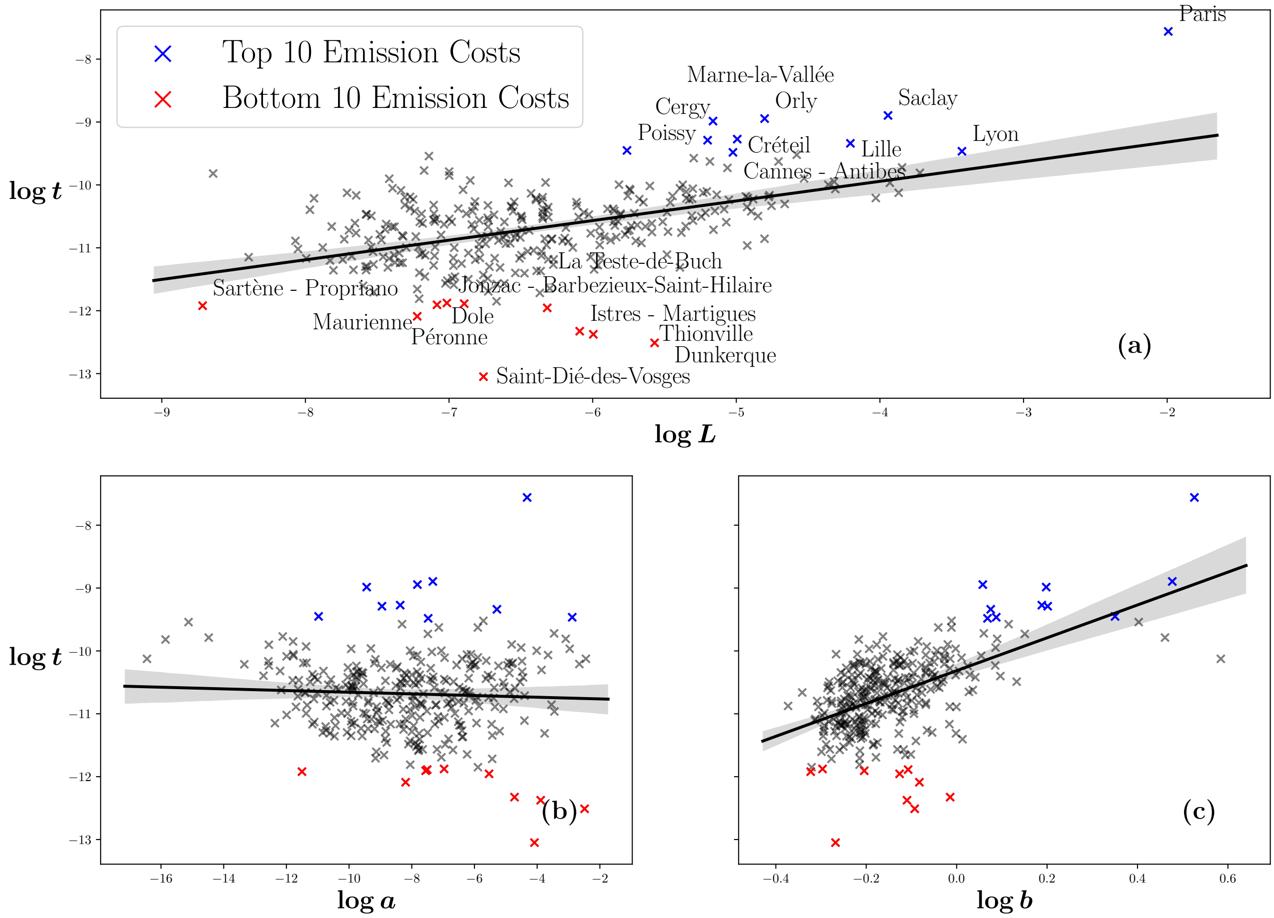


Figure 5: Log-linear relationship between local relative emission cost and local relative population, productivity and amenity.

Local relative emission costs, productivities and amenities estimated by inverting the observed equilibrium using one tradable sector and data for 2012.

cities. It appears that the set of policies currently implemented to tackle local atmospheric pollution imposes distinct levels of emission regulation stringency across French cities, relative to each other. As explained in the introduction, this distribution of emission costs is not the central planner policy instrument per se, but rather the result of the aggregate policies that are implemented and that are summarized here as a local measure of relative stringencies. Second, I observe a general positive relationship between these relative stringencies and relative city sizes. This means that the current policies impose more stringent emissions regulations in larger cities. Third, local levels of emission stringencies do not seem to vary with the local levels of amenities but do increase with local levels of productivities levels.

One can compare these observations to the closed form results of equation (5) in section (3). Within the simplified framework, the optimal levels of relative emission costs are given by a constant elasticity function of local levels of amenities and productivities. If the optimal distribution of relative emission costs in the general model follows the same rationales as in the illustrative model, then, according to equation (5), the elasticity of the optimal emission costs with respect to local amenities and productivities should respectively be close to $\frac{1}{\theta-\frac{\gamma-\alpha}{1-\alpha}}$ and $\frac{1-\gamma}{\theta-\frac{\gamma-\alpha}{1-\alpha}}$. For a back-of-the-envelope numerical application, considering calibrated values $\gamma = .49$, $\delta = .91$ and mean values for sector specific parameters $\nu = .01$ and $\alpha = .05$, one can compute values of 1.02 and $.51$ for these two elasticities.

This simple numerical application has an informative value as it allows us to observe that the current distribution of emission costs across French cities seems to potentially be off from what the model indicates to be the optimal distribution. Indeed, given that the simplified model's elasticities of optimal local emission costs with respect to local amenities and local productivities are empirically found to be positive, intuition would suggest that in the more general model optimal emission costs should also be higher in cities that have higher amenities and productivities relative to cities that are relatively less endowed in these characteristics. However, subgraph (b) from Figure (5) shows that the current distribution of emission costs does not display such positive correlation between local emission costs and local amenities. However subgraph (c) from Figure (5) displays a strong positive correlation between local levels of productivities and emission costs that must not be too far from the optimum.

Figure (5) also highlights in blue the top ten cities with the highest levels of emission costs and in red the bottom ten cities with the lowest levels of emission costs. The fact that cities with the highest levels of emission costs are also the most productive (and that the cities with the lowest levels of emissions costs are the less productive) is obvious. However among the cities with the highest levels of emission costs, some have very high levels of amenities but others have very low levels of amenities.

Based on the arguments developed in section 5, I argue that even if values for calibrated elasticities are subjects to discussion, the estimation strategies are nonetheless informative about their relative sizes. As a result the fact that the elasticities of optimal local relative emission costs with respect to local amenities and productivity levels are positive is likely to be robust to other pollutants, additional economics sector inclusion or other alternative assumptions.

6.2 Consequences from the Heterogenous Distribution of Emissions Costs Across Cities

Results from the previous subsection highlighted the spatial heterogeneity of local emission costs across French cities. In this subsection, I provide an analysis of the role that this heterogenous stringency across space plays in the distribution of emissions, population and wages.

To that end, I compare the observed equilibrium with another equilibrium where the level of emission costs is uniform across all cities. The gap (in the distribution of endogenous variables across cities) between the two equilibrium can be thought of as representing the consequences of the spatial heterogeneity of emission costs. Using the values of idiosyncratic amenities and productivities obtained through the observed equilibrium inversion I compute a new equilibrium by solving the equilibrium conditions (22), (23) using the same emission cost in every city. This uniform mission cost value is the mean of the vector of emission costs obtained in the previous computation and has been normalized to one without loss of generality throughout the analysis. Using a similar approach as for the observed equilibrium inversion, the uniform equilibrium is solved numerically using the Levenberg–Marquardt algorithm. Figure (6) 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 costs that is heterogenous across cities. Darker and lighter areas correspond to respectively a decrease and an increase of the variable.

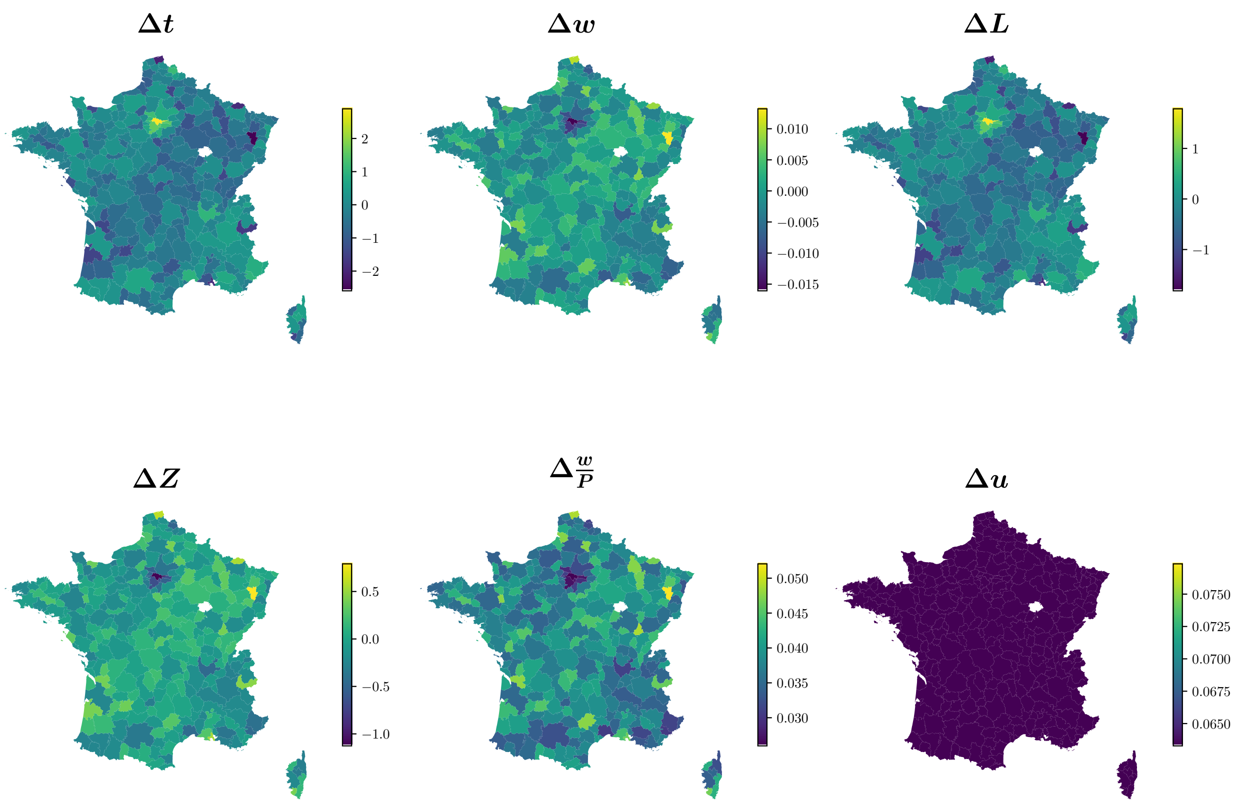


Figure 6: Spatial reallocations and welfare effect of imposing spatially heterogenous emission costs.

Because emission costs are higher in larger cities, such as Paris, Lyon, Bordeaux, Toulouse, mapping Δt leads to a positive variation in these areas and a decrease of the relative emission costs in less populated cities. In turn, mapping the ΔZ reveals a decrease in the local quantity of emissions in the largest cities and an increase in the smallest cities. That is to say, when the central planner applies 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. In this framework a decrease of emissions is a “good thing” and can be seen as being equivalent to an improvement of local air quality.

Figure (6) also plots ΔL to illustrate the reallocation of population across cities due to moving from the uniform emission costs to the current heterogenous stringency. Indeed, according to the assumption of free migration, workers react to the change in local air quality (and prices which are affected by the change in local emission costs) and move away if the local utility has decreased so that the new equilibrium welfare level is equalized across space. From the map illustrating this ΔL one can observe that the current distribution of stringencies that puts higher emission costs in larger cities relative to smaller cities leads to a reallocation of workers from small cities where emissions have increased to larger cities where emissions have decreased. That is to say that larger cities become cleaner so that even more people have incentives to live there relatively to smaller cities.

It may be important here to clarify why I describe cities as being small or large. Indeed, in my spatial framework the size of a given city is an endogenous variable: fractions of its population can move to another city as a response to changes in air quality of real wages. The only two idiosyncratic characteristics of a city are its amenities and its productivity which are fixed and exogenous. In order to be completely coherent with my model when designating a type of cities, one should distinguish between “more or less productive cities” and ”cities with high or low amenities” rather than “large or small cities”. This is possible because, albeit not observed directly, these local characteristics can be computed from observables. However, for simplicity, I often choose to designate a city as being “large” or “small”. This is possible because in general, in the alternative equilibria that I present (the observed and uniform equilibrium and in the next section the optimal equilibrium) the ranking of city sizes does not change and larger cities remain larger and smaller cities remain smaller.

The spatial reallocations of populations and emissions are also illustrated by Figures (7) and (**??**). Figure (7) plots the log-difference between the relative local populations under the current distribution of emission costs and under a uniform emission costs. For a given city, a positive number indicates that the local population is higher under the current distribution of emission costs compared to what it would be if emission costs were the same across all cities. On subgraph (a) the reallocation of population is presented as a function the current population distribution (current city sizes) and on sub graphs (b) and (c) as a function of the distribution of amenities and productivities. In particular it shows that the top ten cities which population benefits the most from the current distribution of emission costs are also highly productive cities but not cities with particularly high amenities.

Figure (**??**) is the parallel of Figure (7) for the effect of the current distribution of emission costs on the distribution of emissions across cities. It plots the log-difference between the relative local emissions under the current distribution of emission costs and under a uniform emission costs. Subgraphs (b) and (c) illustrate the fact that higher relative emission costs in more productive cities leads to lower levels of emissions there compared to a uniform emission costs distribution would lead to. In turn, there are no clear correlation with the distribution of local amenities.

More generally, these two figures suggests (both subgraphs (a)'s) that , currently, air quality policies implement more stringent regulations in larger cities allow these cities to be cleaner (less emissions) and in turn larger (higher population) than what they would be they faced the same level of emission regulations stringency.

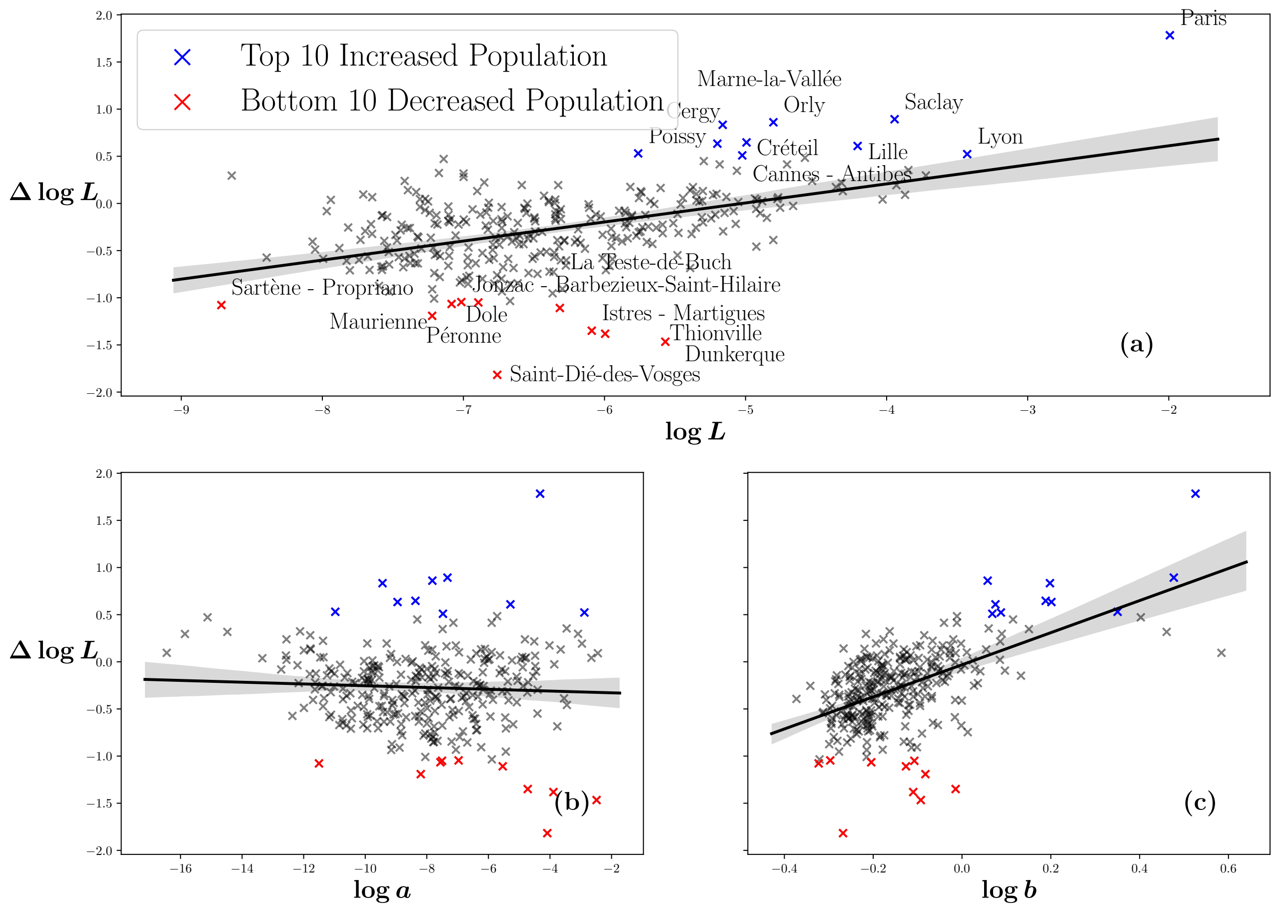


Figure 7: Effects of spatial heterogeneity emission stringency on agglomeration: under the current policy, large cities emit less pollution and are therefore, larger than what they would be if they all faced the same emission cost.

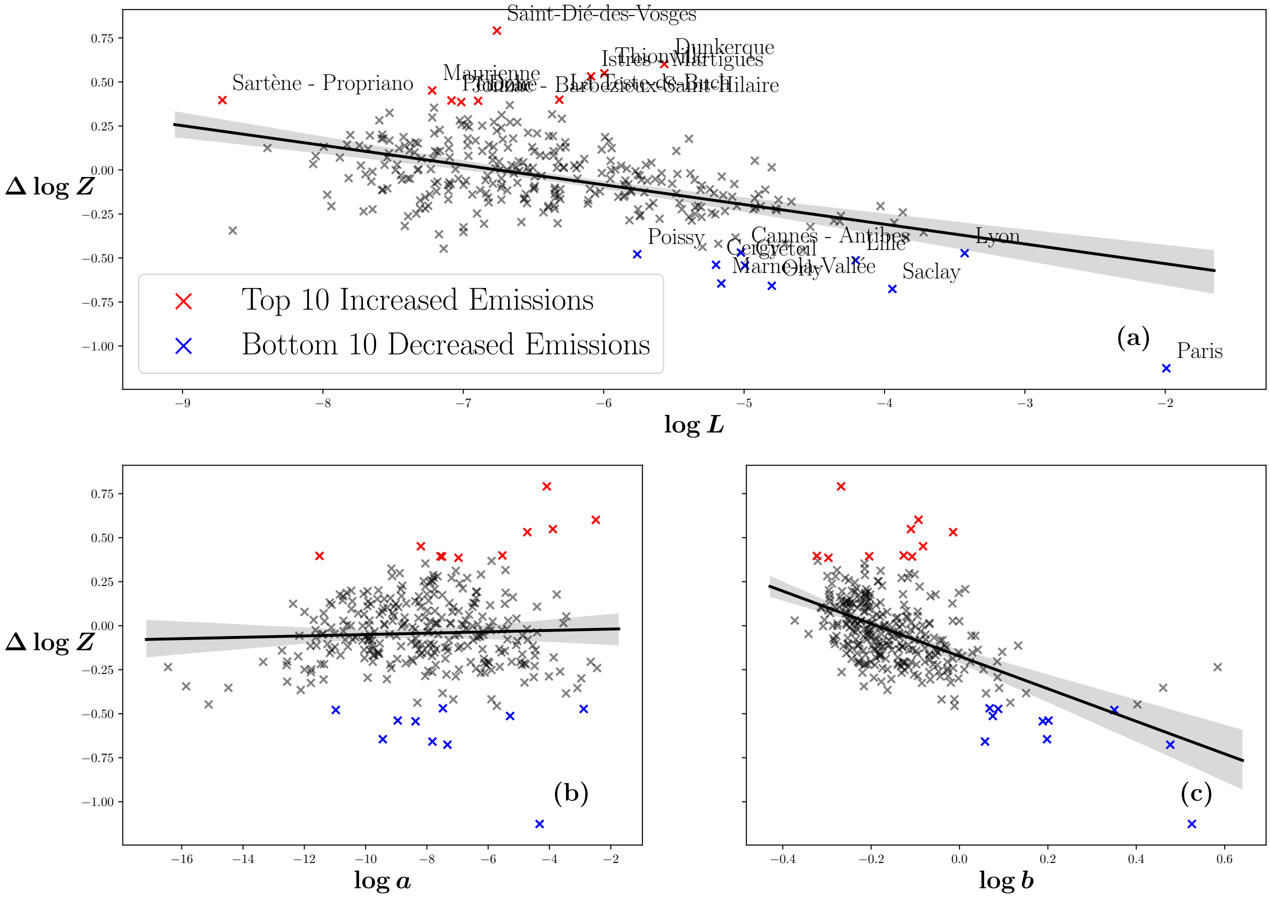


Figure 8: Effects of spatial heterogeneity emission stringency on agglomeration: under the current policy, large cities emit less pollution and are therefore, larger than what they would be if they all faced the same emission cost.

6.3 Identification of the Mean-Preserving Optimal Regulation

The model can not only be informative of the current distribution of relative emission stringencies across cities, but it can also be used to assess whether this distribution maximizes welfare. To do so I numerically solve the optimization problem defined in equation (26) in order to identify the optimal distribution of relative emission costs. This optimization problem yields a new spatial equilibrium that can be compared to the current equilibrium. In the illustrative framework I provide some closed-form results on the relationship between cities' idiosyncratic characteristics and their respective optimal level of relative emission costs. In the present subsection 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' utility function. In practice, I identify for each city an optimal level of emission stringency that I can use to assess whether the current stringency of air quality policy is "too high" or "too low" relative to other cities.

Naturally these results are only a consequence of the choice for an objective function maximized by the central planner. In particular I assume that they aim at maximizing the aggregate welfare of workers, considering the local trade-off between income, congestion and air quality. However one could also argue that because the central planner's prerogatives are only to set the relative emissions costs, their objective should only be to correct the pollution externality. In this case an alternative objective function could be to minimize the maximum pollutant concentration (emission per area or per capita for instance) across cities. I leave such optimization problem outside of the scope of the paper and focus on the case where the central planner uses the distribution of relative emission costs to correct all externalities.

Figure (9) is the main illustration for this numerically solved optimization problem with a simultaneous comparison with the current equilibrium. It plots the distribution of the logarithm of relative emission costs first as a function of the logarithm of the current city sizes (subplot (a)), then of the levels of amenities (subplot (b)) and finally of the productivity (subplot (c)). The current distributions are figured in black, and are the same as in Figure (5), and the optimal distributions are figured in red. The main observation from figure (9) is that increasing the relative emission stringency faced by larger cities compared to smaller cities should bring welfare gains.

One can also observe that in the optimal distribution, a lot of smaller cities should roughly face the same level of regulation stringency as there seems to be a lower threshold. Equivalently, the red distributions display less variation on the left part than on the right part. The reason for this comes from the fact that the set of red dots results from the constrained numerical optimization algorithm. Because smaller cities are so much smaller than larger cities (most of the population is concentrated in a few large cities) and I constrain the mean emission cost to be fixed, their optimal emission costs are put to the minimum amount possible. Another way to explain this is to see these points as corner solutions. In the end, because the distributions of cities' characteristics (amenities and productivities, so not only size) are so skewed that only the largest cities really matter for the aggregate welfare.

As explained in the first part of this section, according to intuitions drawn from the illustrative setup's section, it seems that increasing the relative stringency of regulations in cities with higher amenities would improve welfare. This intuition is supported by Figure (5)'s subgraph (b): in cities with lower amenities, the relative emission cost is set to its minimum and it is increased for cities with the highest amenities. This comes from the fact that, currently, cities with higher amenities are generally too small because workers and firms do not internalize the impact of their location choices on local air quality. There was already a strong positive correlation between the spatial distribution of emission costs and the distribution of idiosyncratic productivities. Solving the optimization problem confirms that such correlation is optimal. Indeed, subgraph (c) shows that the coefficient of the log-linear relationship between the local idiosyncratic levels of productivities and the relative

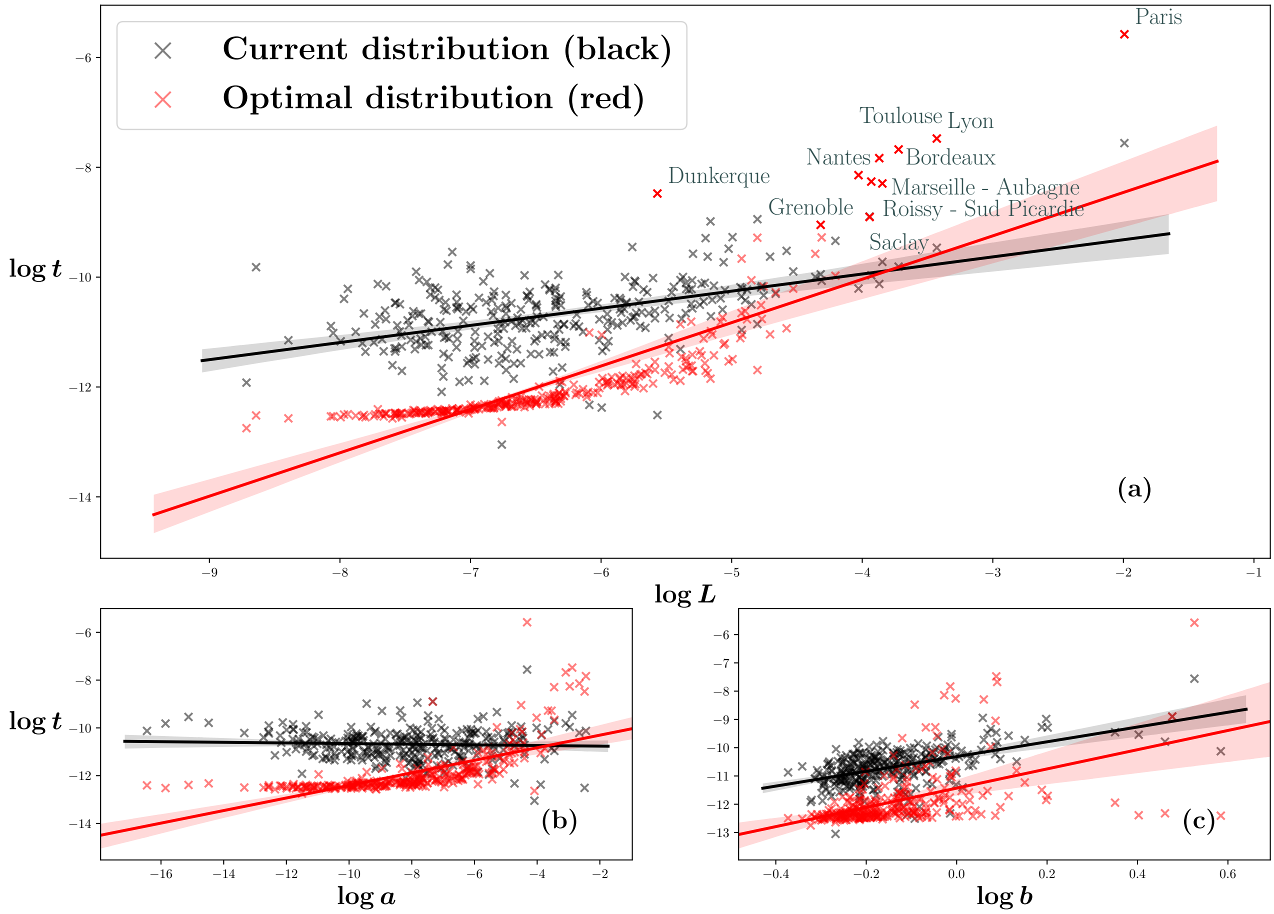


Figure 9: Comparison between current distribution of relative emission costs (black) and the mean-preserving optimal distribution (red): emission costs are lower in smaller cities and increased in larger cities.

emission costs is currently not too far from the optimal. Subgraph (a) summarizes the reallocation necessary to move from the current distribution of relative emission costs to the optimal one. Basically, the relative stringency of air quality policies needs to be raised in the currently largest cities and decreased in the smallest cities.

6.3.1 Population Reallocations

Moving from the current distribution of relative emission costs to the optimal one would trigger a reallocation of populations across cities. Figure (10) displays the population reallocation effect of moving to the optimal distribution of relative emission costs. In particular it highlights in blue the cities that would have their local population increased as a result of this new stringency distribution. Only a small subset of cities (around 15 of them over the 300 cities) grows larger and the rest grows smaller. Specifically, subgraph (a) illustrate the fact that the population concentration mechanism is reinforced and the largest cities grow even larger. From subgraph (b) and (c) one can realize that this reallocation mainly comes from the increased emission costs in cities with higher amenities. As cities with high amenities become less polluted, more workers move there.

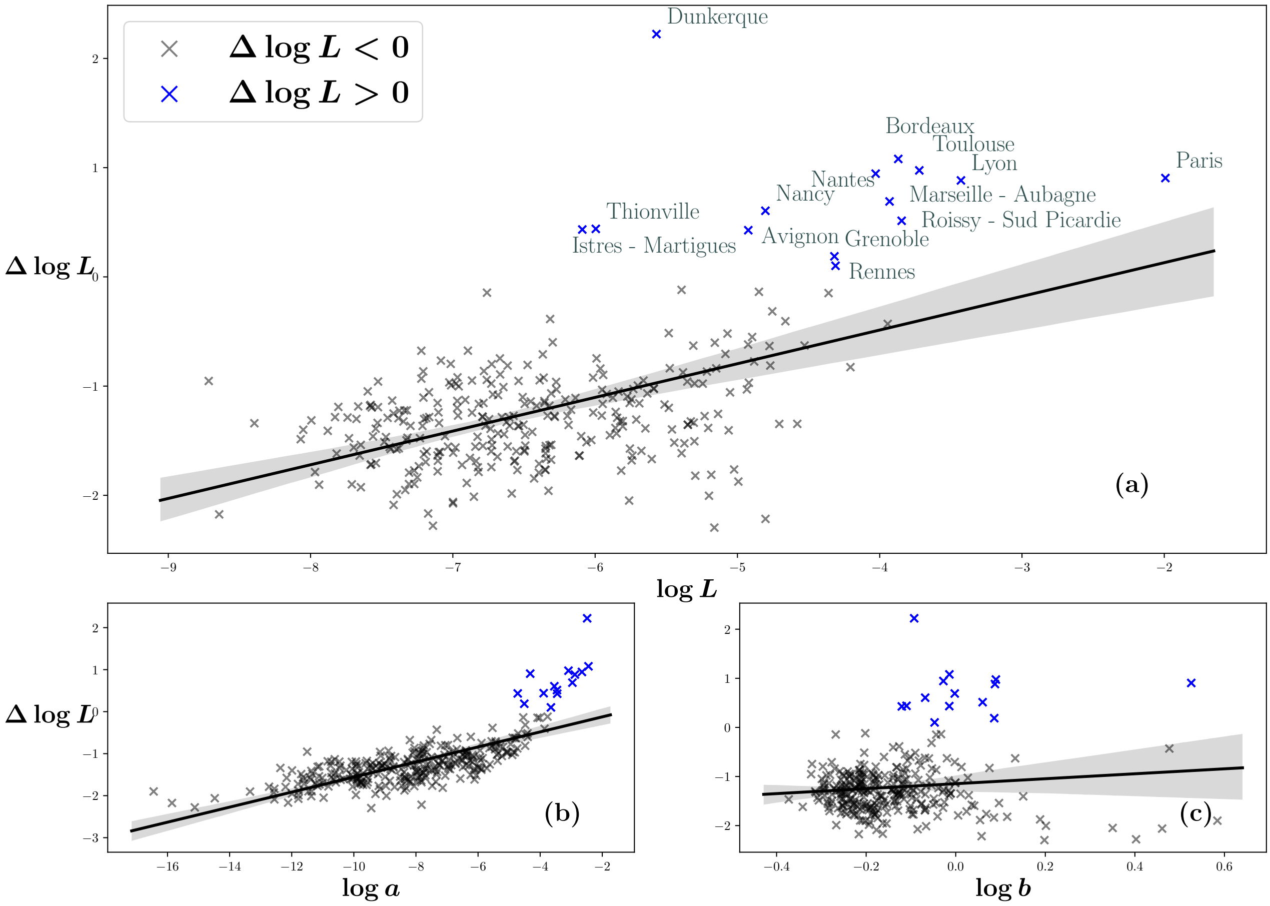


Figure 10: Population reallocation when moving from the current to the optimal distribution of relative emission regulation local stringencies

In blue are the cities that would see their population increase as a result of this emission costs distribution shift.

Moving from the current to the optimal distribution of local emission costs not only induces a reallocation of workers across cities but it also changes the degree of spatial concentration of population. Figure (11) compares the cumulative distribution of population across cities before and after moving to the optimal distribution of emission costs. In practice I rank cities according to their current populations and compute the cumulative population starting from the smallest city. This yields the black line. Then I do the same cumulative sum using what local populations would be under the optimal distribution of emission costs and obtain the red line. The fact that the red line is more skewed to the right than the black line is evidence that the national population is more concentrated in the largest cities under the optimal distribution of emissions costs than under the current set of policies.

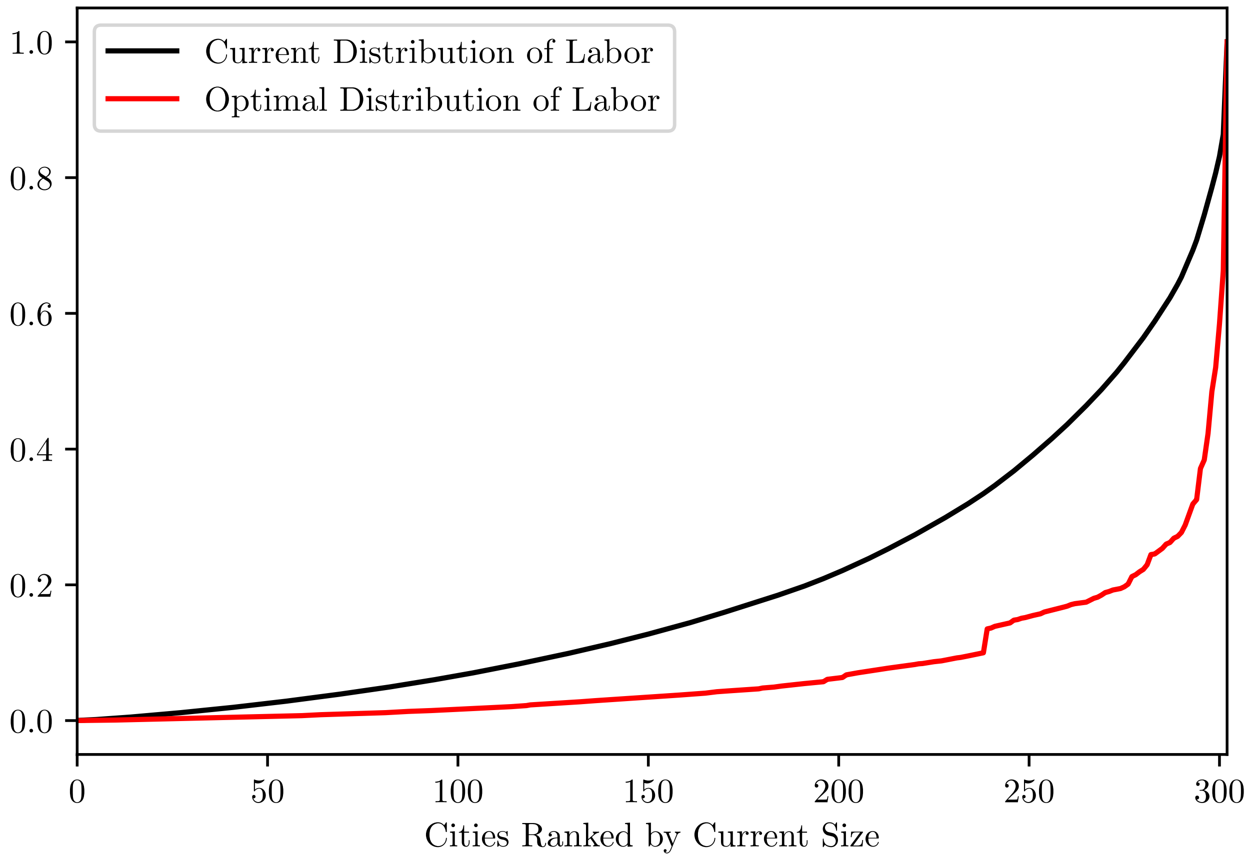


Figure 11: Cumulative Distributions of Population Under the Current and the Optimal Distributions of Emission Costs.

6.3.2 Pollutants Emissions Reallocations

There is also an optimal way of reallocating emissions across cities that is implied by the optimal distribution of local emission costs. Figure 12 displays the pollution reallocation effect of moving to the optimal distribution of relative emission costs. In particular it highlights in green the cities that would have their local emissions decreased as a result of this new stringency distribution. In blue are the cities that also grow in population. In accordance with results on the optimal reallocation of emission costs, it is optimal that emissions are decreased in the currently more populated cities (subgraph (a)), corresponding to a reallocation of emissions from cities with high amenities to cities with low amenities (subgraph (b)).

As for population, reallocation of emissions across cities has implications for the degree of spatial concentration of emissions across cities. Figure (13) compares the cumulative distributions of emissions under the current (black line) and the optimal (red line) distributions of emission costs across cities. The red line being less skewed to the right it is clear that the optimal reallocation of emissions across cities implies a lower degree of spatial concen-

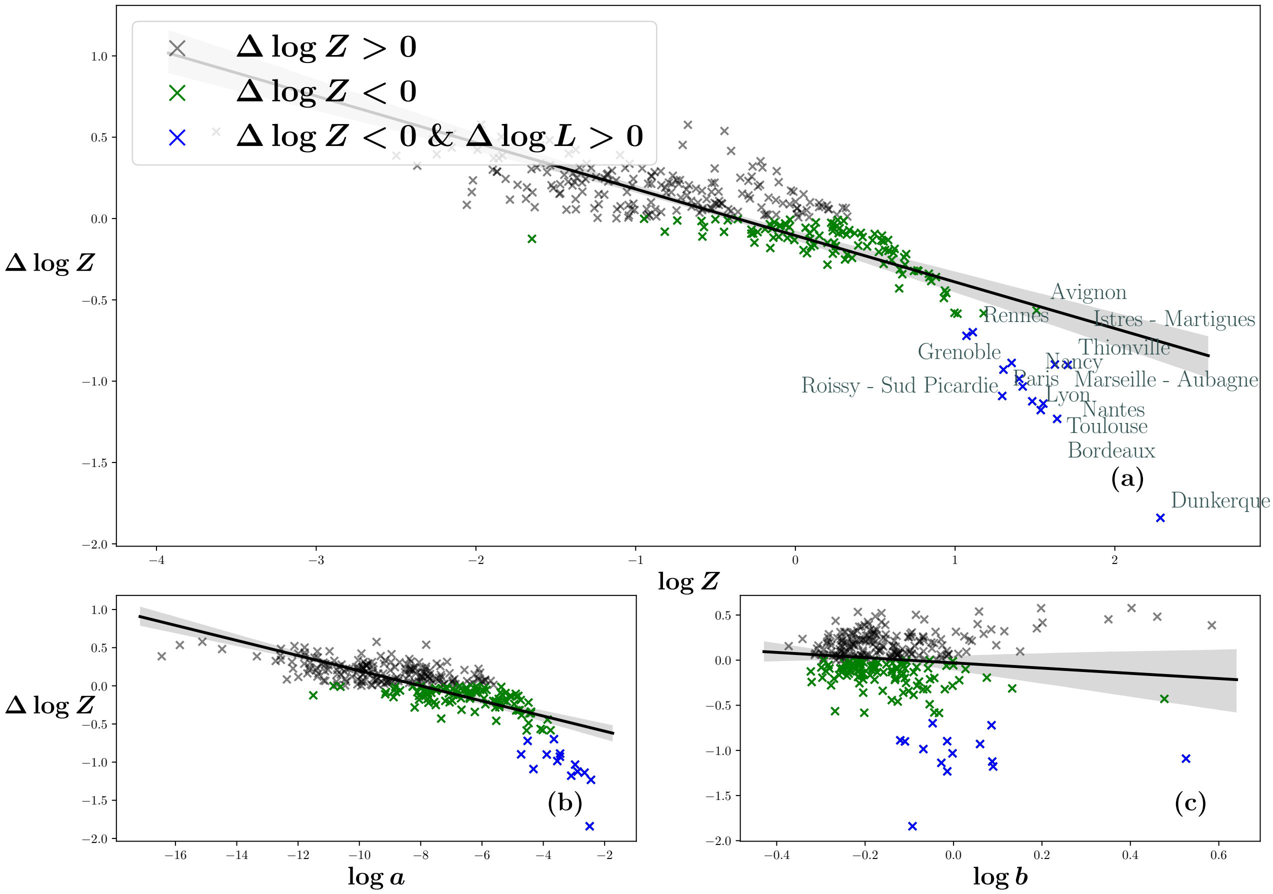


Figure 12: Emissions reallocation when moving from the current to the optimal distribution of relative emission regulation local stringencies

In blue are the cities that would see their population increase as a result of this emission costs distribution shift.

tration of emissions.

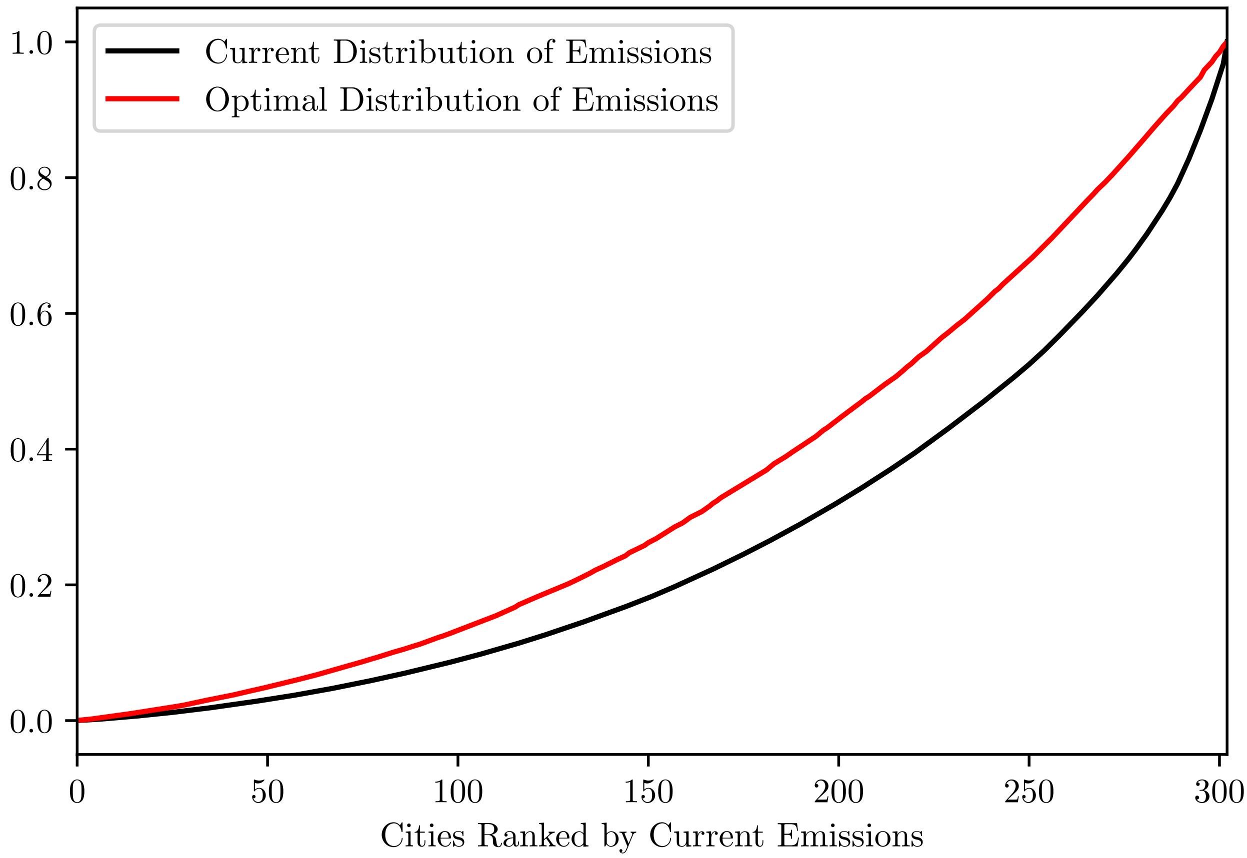


Figure 13: Cumulative Distributions of Pollution Emissions Under the Current and the Optimal Distributions of Emission Costs.

7 Conclusion

In this paper, I analyze how the distribution of polluting activities across space affects workers' welfare and to what extent the spatial distribution of relative stringency of local air quality policies can correct for this externality. First I build an illustrative model of spatial externality featuring endogenous air quality affected by local polluting activities. Because pollution has adverse effects on consumers utility it acts as a centrifugal force and workers have the opportunity to move away from heavily polluted cities. Simplifying assumptions allow me to show closed form solution and in general a central planner that maximizes workers' welfare will apply relatively more stringent regulations in cities that have higher amenities and productivity.

Then I extend this illustrative setup to a general spatial equilibrium model featuring industrial agglomeration economies, free mobility of workers and industrial emissions of pollution by heterogenous sectors. Competition is imperfect and trade is costly. I derive the conditions for the spatial equilibrium and define the optimization that the central planner has to solve in order to maximize welfare.

I calibrate this general framework using an extensive set of data sources on French firms and French cities. After calibrating inverse labor supply elasticities, demand and production sector specific parameters and a proxy for between cities iceberg trade costs, I invert the equilibrium conditions to retrieve cities idiosyncratic characteristics, amenities and productivity, and relative air quality policy local stringency from the observed equilibrium quantities, populations, wages and emissions.

Using this calibrated model, I investigate to what extent spatial emission cost heterogeneity shapes the distribution of population and pollution and compute the optimal distribution of emission cost across space. I find that current air quality policies match with a spatial distribution of relative emission costs that are heterogenous across cities. Moreover I observe that the relative stringency of air quality policies is higher in the largest cities. Based on the computation of cities' local characteristics I also find that emission costs are higher in more productive cities but not, on average, in cities with more amenities.

Finally, I solve the central planner's optimization problem and identify the distribution of relative local emission costs that maximizes workers' welfare. I find that the optimal policy sets high emission costs in large cities in order to enhance air quality there which, in turn, reinforce spatial agglomeration and brings welfare gains.

References

Allen, T. & Arkolakis, C. (2014). Trade and the topography of the spatial economy. *Quaterly Journal of Economics*.

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

Bartelme, D. (2018). Trade costs and economic geography: Evidence from the u.s. *mimeo*.

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.

Borck, R. & Schrauth, P. (2019). Population density and urban air quality. *CEPA Discussion Papers*, (pp.8̃).

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.*

Chimere (2020). Chemistry-transport model v2020r1 - documentation. *http://www.lmd.polytechnique.fr/chimere/*.

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., Puga, D., & Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6), 2543–2594.

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.

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. (2020). Geographic and socioeconomic heterogeneity in the benefits of reducing air pollution in the united states.

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.

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*.

Fajgelbaum, P. D. & Gaubert, C. (2020). Optimal spatial policies, geography, and sorting. *The Quarterly Journal of Economics*, 135(2), 959–1036.

Gao, X., Song, R., & Timmins, C. (2021). The role of information in the rosen-roback framework. *National Bureau of Economic Research Working Paper No. 28943*.

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

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8), 2586–2624.

Hanlon, W. W. (2019). Coal smoke, city growth, and the costs of the industrial revolution. *The Economic Journal*, 130(626), 462–488.

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

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

Pascal, M., de Crouy Chanel, P., Wagner, V., Corso, M., Tillier, C., Bentayeb, M., Blanchard, M., Cochet, A., Pascal, L., Host, S., Goria, S., Tertre, A. L., Chatignoux, E., Ung, A., Beaudeau, P., & Medina, S. (2016). The mortality impacts of fine particles in france. *Science of The Total Environment*, 571, 416–425.

Redding, S. (2020). Trade and geography.

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*.

Rosen, S. (1979). *Wage-Based Indexes of Urban Quality of Life*, (pp. 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.

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*.

Zivin, J. G. & Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7), 3652–3673.

Zivin, J. G. & Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51(3), 689–730.

A Appendices

A.1 Appendices for the Illustrative Framework

A.1.1 Proofs for the Illustrative Setup Results

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:

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

Using the Lagrangian, the First Order Conditions 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:

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

So that 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:

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

, 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 we 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:

$$\mathscr{L} = \bar{u}(t_1, ..., t_C) + \lambda \left(\sum_{j \in C} t_j - 1 \right) \tag{49}$$

with λ the Lagrangian multiplier. And we find:

$$\frac{\partial \mathscr{L}}{\partial t_j} = 0 = \lambda + \bar{u}^{\frac{\theta-1}{\theta}} \frac{\gamma - \alpha}{1 - \alpha} a_j^{\frac{1}{\theta}} b_j^{\frac{1-\gamma}{\theta}} t_j^{*\frac{1}{\theta}\frac{\gamma-\alpha}{1-\alpha}-1} \tag{50}$$

from which one can get equation (5).

A.2 Appendices for the Application to France

A.2.1 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(14) 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.

A.2.2 Descriptive Statistics on Equilibrium Data

A.2.3 Descriptive Statistics on Computed Local Characteristics

A.2.4 Descriptive Statistics on Between-Cities Travel Times

	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 6: 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 7: 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 8: 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 9: 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 10: 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 11: Top Ten Cities with Lowest Amenities, Productivities and Emission Costs

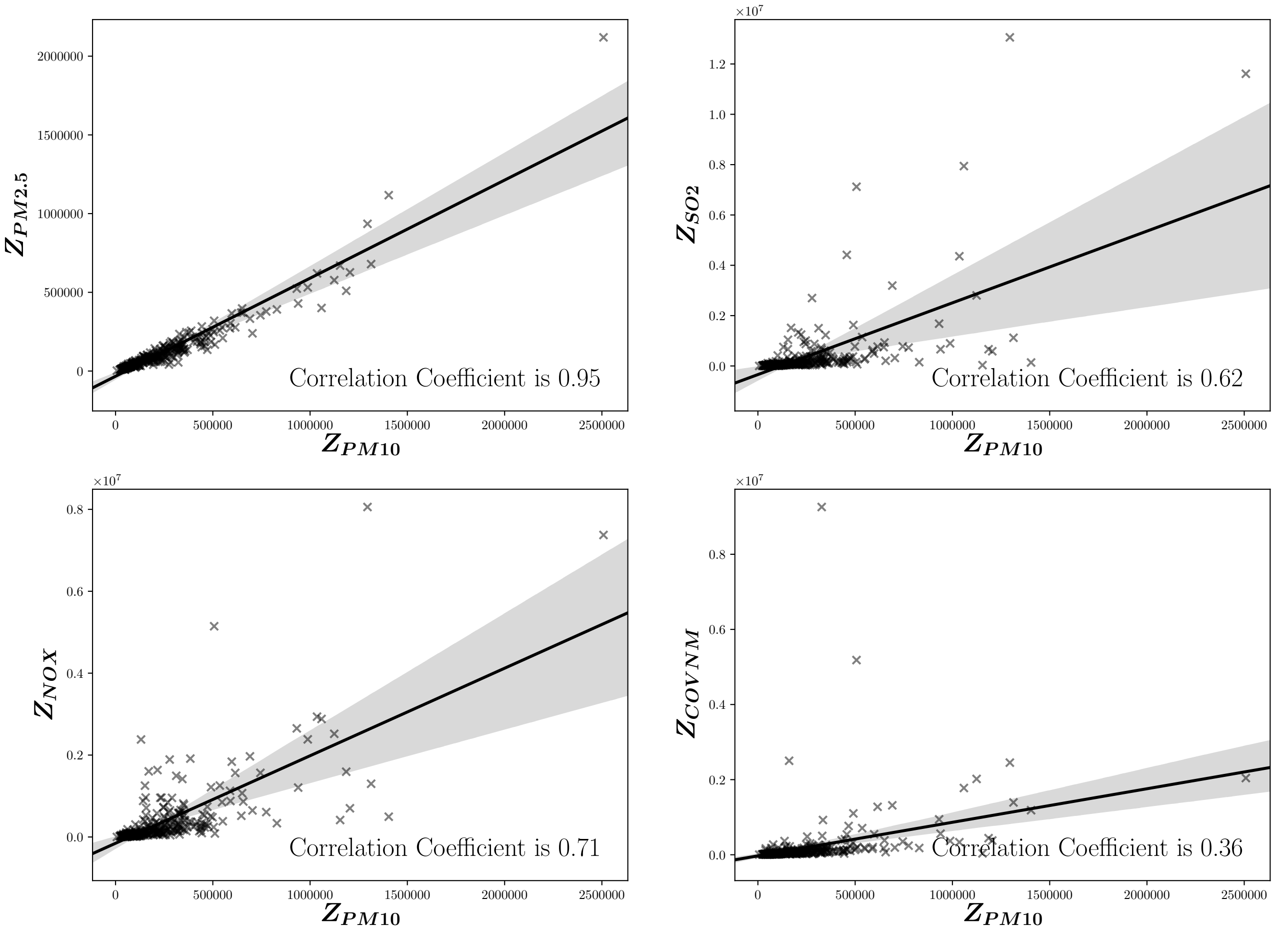


Figure 14: Correlation Between Emissions of Harmful Industrial Pollutants at the Commuting Zone Level

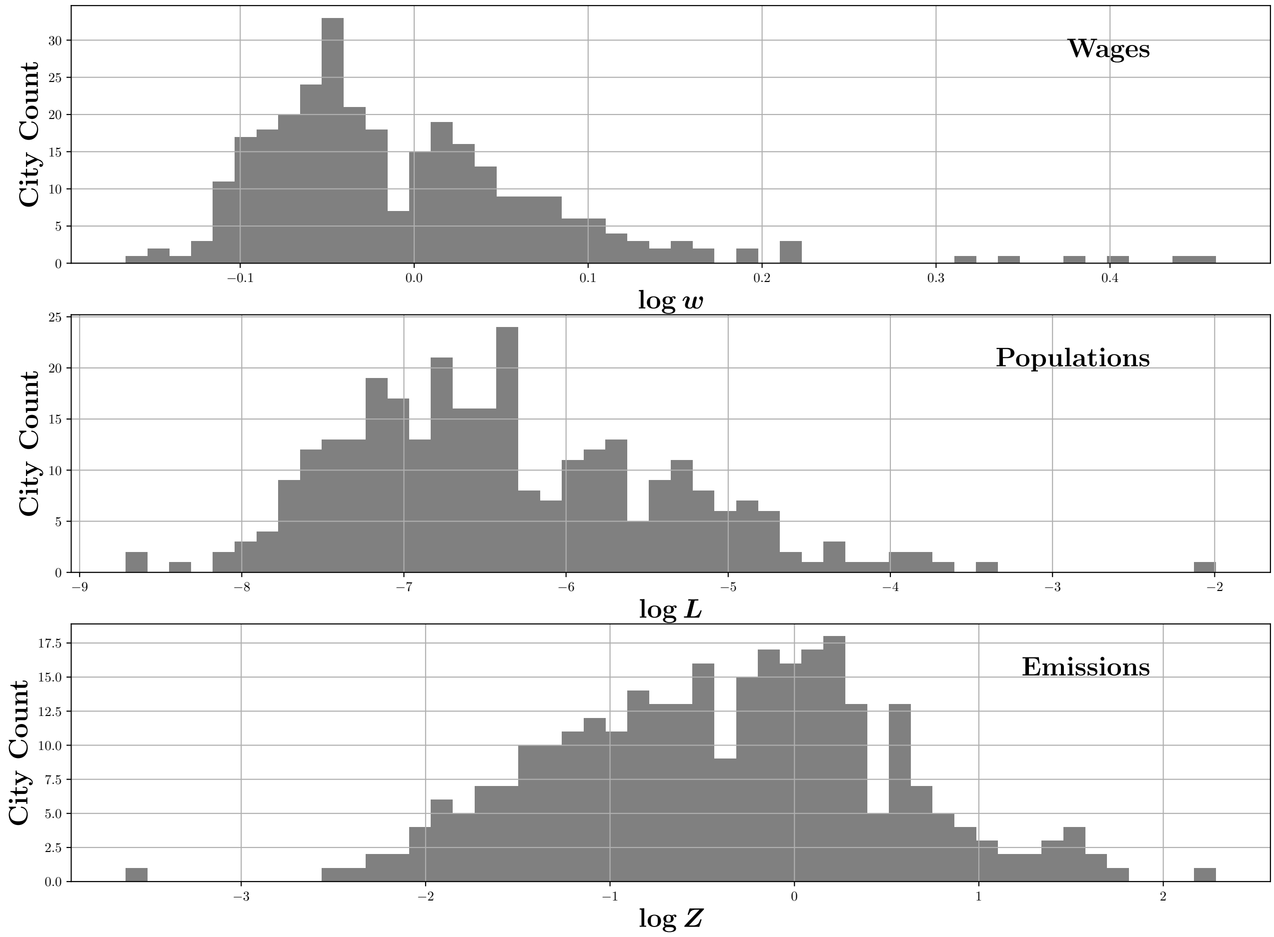


Figure 15: Histograms of Spatial Distributions of Wages, Populations and Emissions across French Cities

All variables are in log.

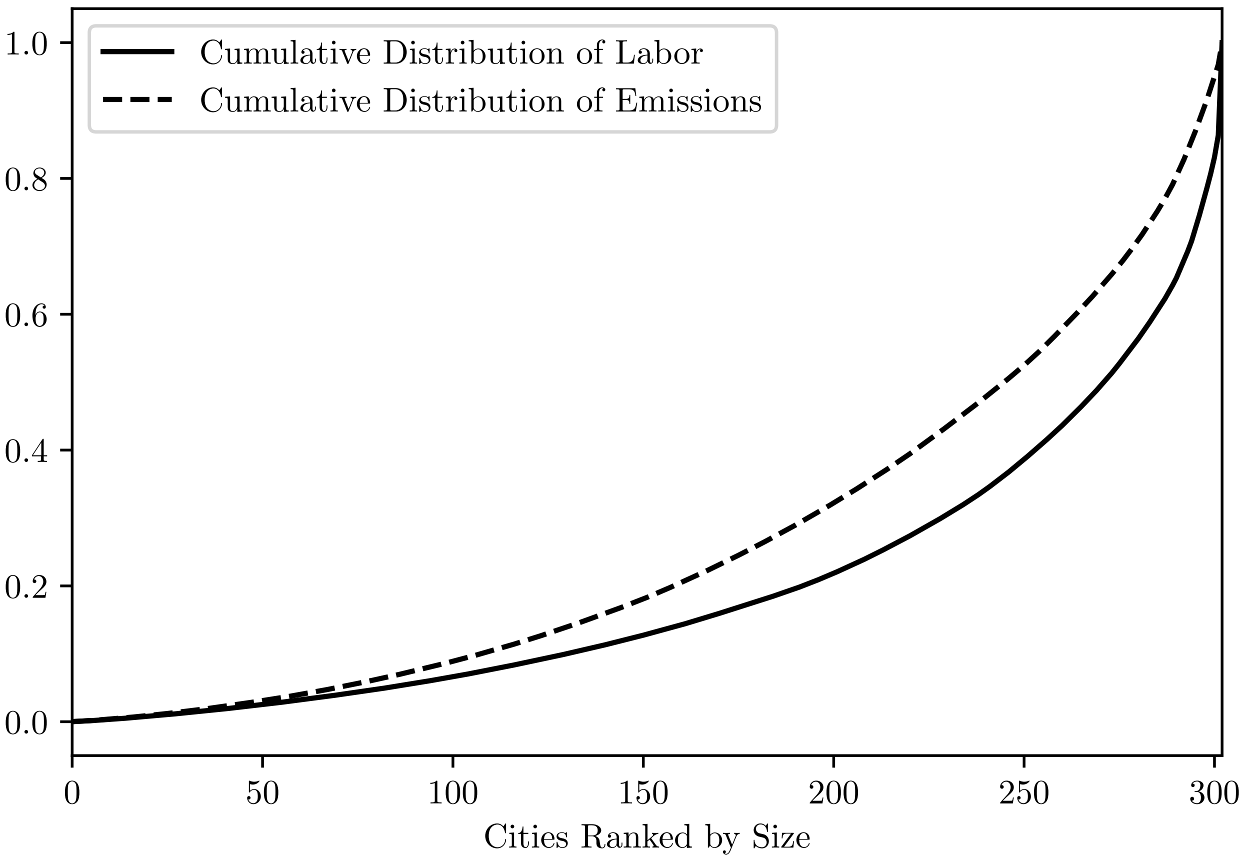


Figure 16: Cumulative Distributions of Populations and Emissions across French Cities
Cities are ranked by size, from smallest to largest, and the normalized cumulative sum of populations and emissions is plotted.

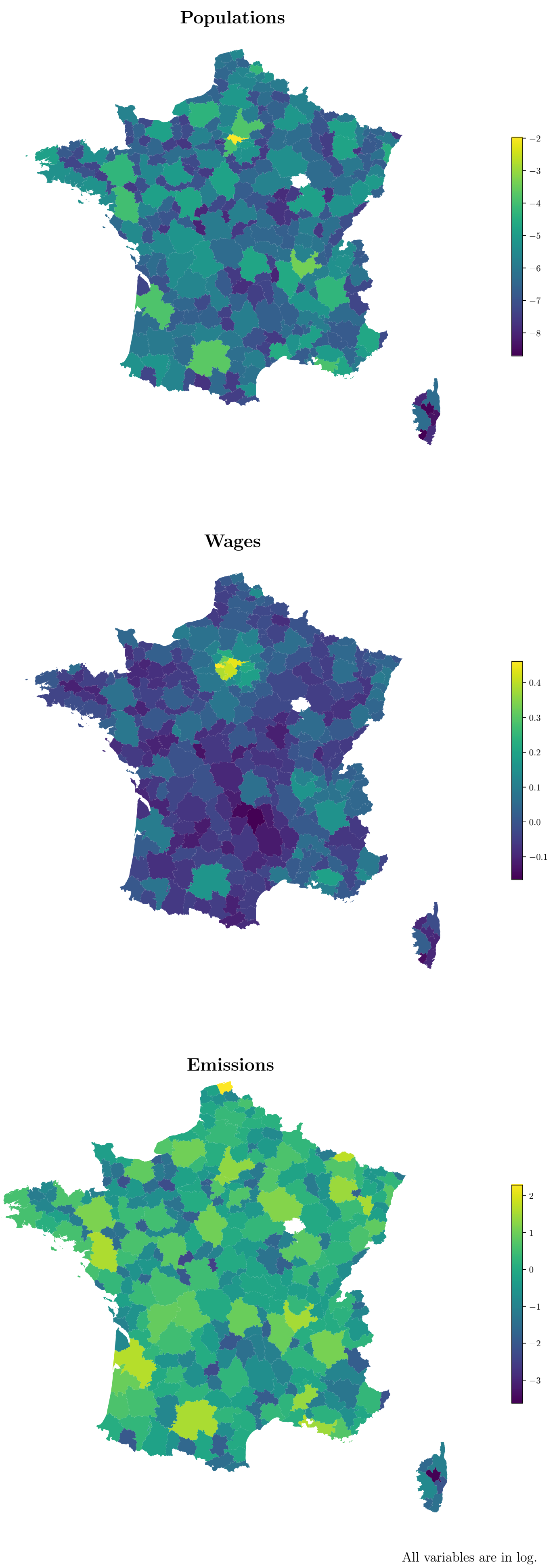


Figure 17: Spatial Distributions of Wages, Populations and Emissions across French Cities

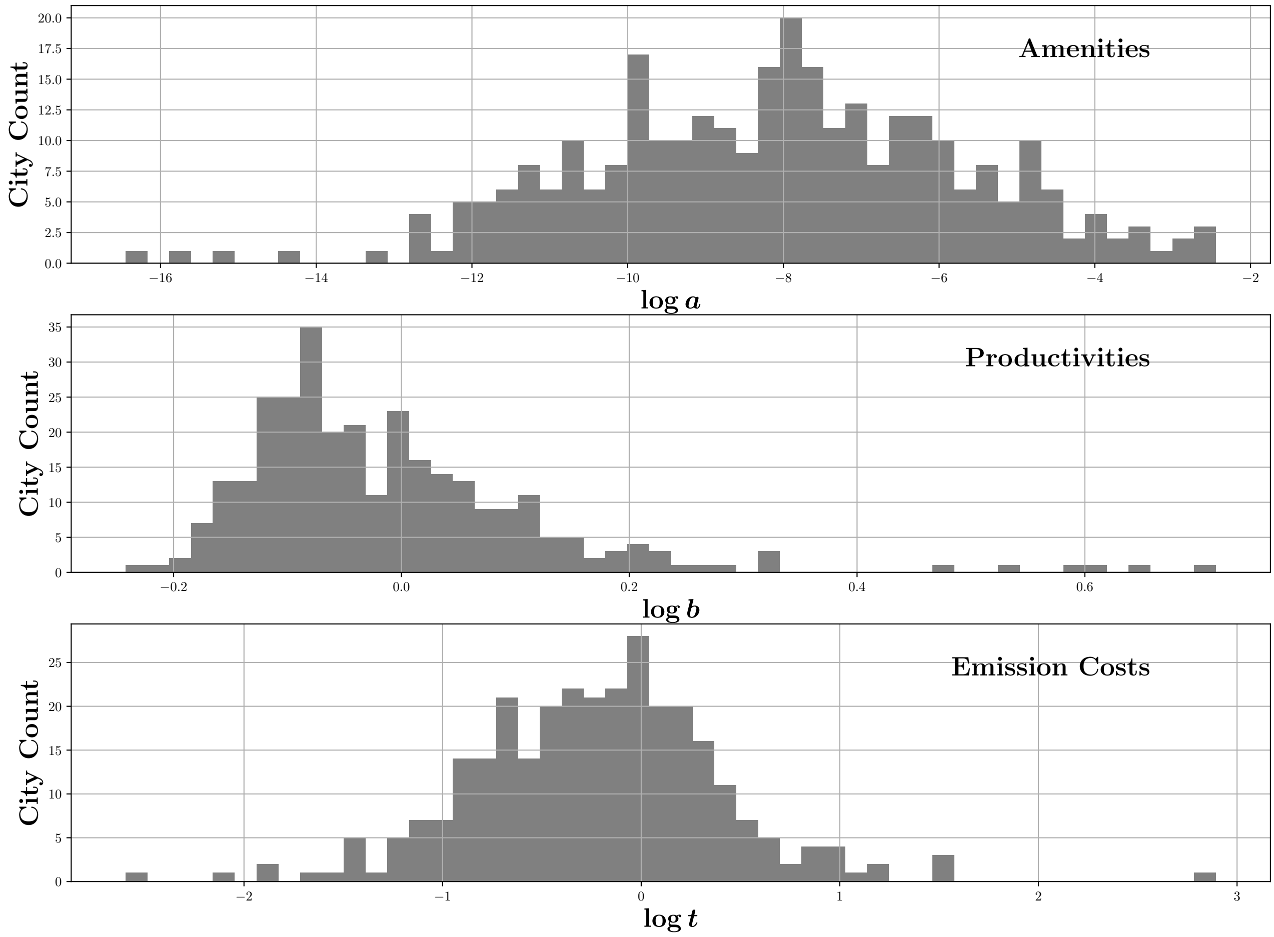
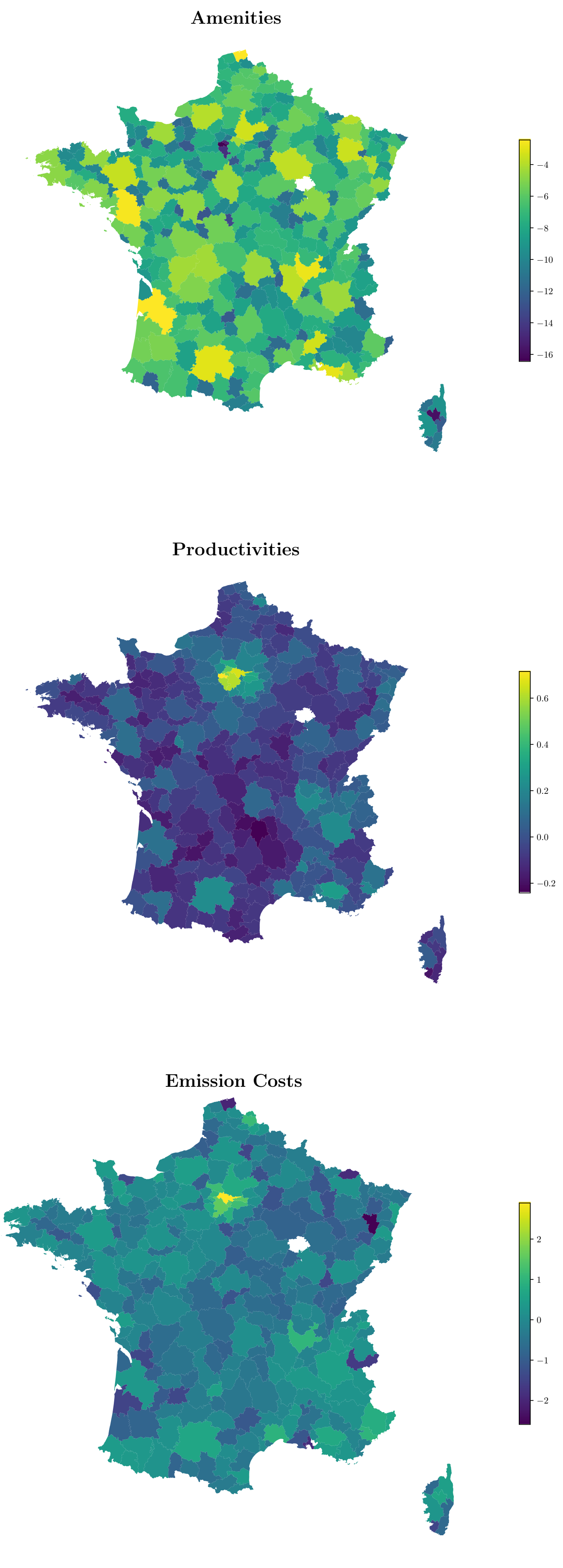


Figure 18: Histograms of Spatial Distributions of Wages, Populations and Emissions across French Cities

All variables are in log.



All variables are in log.

Figure 19: Spatial Distributions of Amenities, Productivities and Emission Costs across French Cities