NBER WORKING PAPER SERIES

FIRM SORTING AND AGGLOMERATION

Cecile Gaubert

Working Paper 24478
http://www.nber.org/papers/w24478

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2018

I am extremely grateful to Esteban Rossi-Hansberg, Steve Redding and Gene Grossman for continuous support and guidance. I thank the editor and four anonymous referees for their constructive comments, and Treb Allen, Costas Arkolakis, Thomas Chaney, Don Davis, Jan De Loecker, Gilles Duranton, Pablo Fajgelbaum, Joshua Gottlieb, Pat Kline, Oleg Itskhoki, Kiminori Matsuyama, Eduardo Morales, Dávid Nagy, Andrés Rodríguez-Clare, Felix Tintelnot and Jonathan Vogel as well as numerous seminar and conference participants for helpful discussions and comments. This research benefited from financial support from the International Economics Section at Princeton University. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Firm Sorting and Agglomeration
Cecile Gaubert
NBER Working Paper No. 24478
April 2018
JEL No. F11,R10,R30

ABSTRACT

The distribution of firms in space is far from uniform. Some locations host the most productive large firms, while others barely attract any. In this paper, I study the sorting of heterogeneous firms across locations and analyze policies designed to attract firms to particular regions (place-based policies). I first propose a theory of the distribution of heterogeneous firms in a variety of sectors across cities. Aggregate TFP and welfare depend on the extent of agglomeration externalities produced in cities and on how heterogeneous firms sort across them. The distribution of city sizes and the sorting patterns of firms are uniquely determined in equilibrium. This allows me to structurally estimate the model, using French firm-level data. I find that nearly half of the observed productivity advantage of large cities is due to firm sorting. I use the estimated model to quantify the general equilibrium effects of place-based policies. I find that policies that decrease local congestion lead to a new spatial equilibrium with higher aggregate TFP and welfare. In contrast, policies that subsidize under-developed areas have negative aggregate effects.

Cecile Gaubert
Department of Economics
University of California, Berkeley
687 Evans Hall
Berkeley, CA 94720
and NBER
cecile.gaubert@berkeley.edu

An online appendix is available at http://www.nber.org/data-appendix/w24478

1. Introduction

The distribution of economic activity across space is strikingly uneven. Location choices made by individual firms arguably play a role in shaping these spatial disparities: some areas boom, driven by the presence of large and productive firms, while other areas barely attract any. Recognizing this, governments put in place a range of policies that aim at attracting firms to specific areas of a country. How much of the productivity advantage of a region is shaped by the efficiency of the firms it attracts, and what is the aggregate impact of altering the location choice of heterogeneous firms through place-based policies? These questions require a theory of the location choice of heterogeneous firms and their impact on local and aggregate productivity. In this paper, I develop such a theory and explore its quantitative implications.

I first use this framework to decompose the productivity advantage of firms in dense areas into (1) an advantage caused by density – i.e., the extent of agglomeration externalities – and (2) the endogenous sorting of more productive firms into these areas. I estimate that nearly half the the elasticity of productivity to density comes from firm sorting. Second, I use the framework to evaluate the general equilibrium effect of spatial policies. A range of such spatial policies directly target firms, and aim to attract them to specific areas (in general, to the less developed ones). By accounting for how these policies influence the location choice of heterogeneous firms, which type of firms they attract, and how this location choice feeds back into the productivity of firms, the quantified model allows to study the general equilibrium effect of such policies on aggregate productivity and on spatial disparities between regions. I find that policies that subsidize smaller cities can have negative aggregate effects, and do not necessarily reduce spatial disparities.

In the model, cities form endogenously, on sites that are ex-ante identical. They grow in population as firms choose to locate there and increase local labor demand. Cities are the locus of agglomeration externalities such as thick labor markets or knowledge spillovers (Duranton and Puga (2003)). Firms are heterogeneous in productivity and produce in a variety of sectors with different production functions. Firms sort across cities of different sizes. Firm sorting is driven by a trade-off between gains in productivity through local externalities, and higher labor costs. This trade-off is shaped in particular by how labor-intensive is production, which varies by sector. I assume that more efficient firms benefit relatively more from local agglomeration externalities (Combes et al. (2012)). This generates positive assortative matching: more efficient firms locate in larger cities, reinforcing their initial edge. Finally, city developers compete to attract firms to their city. They act as a coordinating device in the economy, leading to a unique spatial equilibrium.

Using firm-level data, I show that the model is able to reproduce salient stylized facts about French firms. In the data, within sectors, firms have higher revenues in larger cities, but not necessarily higher employment. The model predicts this pattern. Second, across sectors, firm sorting is systematically related to sectoral intensity of input use. Similarly, in the model, in labor-intensive sectors, firms locate more in small cities, where wages are lower. Third, in the data, firm-size distribution is more thick-tailed for sectors that tend to locate in larger cities. In the model, initial differences in productivity between firms induce sorting across city sizes. This, in turn, reinforces firm heterogeneity, as firms in large cities benefit from stronger agglomeration forces, which generates a thicker-tailed firm size distribution.

I structurally estimate the model using firm-level data, and recover a model-based estimate of the shape of agglomeration externalities. This allows me to disentangle the roles played by agglomeration forces and firm sorting in explaining the productivity difference between cities of different sizes. I find that the magnitude of the productivity advantage of large cities is 4.2%, in line with existing measures of agglomeration externalities in the literature as reported by Rosenthal and Strange (2004). Using counterfactual analysis, I estimate that nearly half of this measure comes from firm sorting.

Finally, I analyze the general equilibrium impact of spatial policies that aim to influence the location choice of firms. These policies are pervasive.[1] Federal programs in the United States and in Europe provide generous tax breaks to firms that elect to locate in less developed areas. Because they induce a complex reallocation of factors across space, the aggregate impact of such programs is a priori ambiguous. First, positive local impacts may be counterbalanced by undesirable effects in other regions. Second, subsidies may attract low-productivity firms to the targeted zones while leaving high-productivity firms in the more developed regions, thereby reinforcing spatial disparities. I use the estimated model to quantify the effect of such spatial policies. Specifically, I simulate the new spatial equilibrium that results from two types of programs: a tax-relief scheme targeted at firms locating in smaller cities, and the removal of regulations that hamper city growth, such as zoning or building-height regulations.[2] The productive efficiency of the new equilibrium depends in particular on the new city-size distribution: this distribution drives the extent of agglomeration externalities leveraged by firms in the economy. I find that a policy that subsidizes less productive areas has negative aggregate effects on TFP and welfare. In contrast, a policy that favors the growth of cities leads to a new spatial equilibrium that is significantly more productive by endogenously creating agglomeration externalities and reducing the impact of market failures.

2. Relation to the Literature

The literature that studies systems of cities, pioneered by Henderson (1974), has traditionally focused on homogeneous firms. Recent contributions have introduced richer heterogeneity in the spatial setting.[3] Compared to this literature, the main contribution of the paper is to propose a general theory of mobile heterogenous firms that operate in many different sectors, without specifying firms' heterogeneity parametrically. Furthermore, the model is highly tractable and amenable to quantitative estimation.

In a seminal contribution, Behrens et al. (2014) study the spatial sorting of entrepreneurs who produce non-tradable intermediates. I study the polar case of producers of perfectly tradable goods, and show that

[1] Kline and Moretti (2014) report that an estimated \$95 billion are spent annually in the United States to attract firms to certain locations.

[2] Glaeser and Gottlieb (2008) in particular advocate in favor of reducing this type of regulation.

[3] Early studies of heterogenous firms in the spatial context include Nocke (2006) and Baldwin and Okubo (2006). They predict that more productive firms self-select into larger markets. These results are obtained in a setting where the size of regions is fixed and exogenously given. In a recent contribution, Suarez Serrato and Zidar (2016) study the incidence of the corporate tax in a spatial setting with heterogenous firms. Cities are taken as given and the paper abstracts from agglomeration externalities. In contrast, in this research, city sizes respond endogenously to the location choice of firms and the intensity of agglomeration externalities. Another strand of the literature studies the sorting of heterogenous firms within a given urban area (Brinkman et al. (2015)).

the setup is tractable enough that it can be extended to feature trade costs. Furthermore, an important difference for policy analysis is the uniqueness of the equilibrium I obtain here. Another closely related strand of the literature (Eeckhout et al. (2014), Davis and Dingel (2012) and Davis and Dingel (2013)) studies the spatial sorting of workers who differ in skill level, to shed light on patterns of wage inequality and on the spatial distribution of skills. My research uses similar conceptual tools, borrowed from the assignment literature.[4] It differs however in that, by modeling firms, I can derive novel predictions on how sectoral composition, firm size and input use vary across cities of different sizes. These dimensions are absent from models with only heterogenous labor. Modeling firms is also necessary to study the impact of place-based policies that explicitly target firms. My research is motivated by the empirical finding of Combes et al. (2012). They show that the productivity advantage of firms in large cities is not driven by tougher competition hence stronger selection in larger cities, but by agglomeration effects. Moreover, they find that the most efficient firms are disproportionately more efficient in large cities, indicating potential complementarities between firm productivity and city size. I build on this result and integrate this complementarity in a spatial equilibrium model with mobile and heterogenous firms, a feature absent from their approach. Duranton and Puga (2001) develop a lifecycle model of firm location, in order to explicitly tackle the topic of urban diversity. They propose that diversified cities serve as incubators for new ideas, which are then implemented in specialized cities. Contrary to this research, firms do not differ in productivity types, nor is there heterogeneity in city sizes.

As in Desmet and Rossi-Hansberg (2013) and Behrens et al. (2017), I use structural estimation of a model of a system of cities to assess the welfare implications of the spatial equilibrium. The focus of the analysis is different, since they do not explicitly account for sorting by heterogeneous firms. The paper also contributes to the literature that measures agglomeration externalities, as reviewed in Rosenthal and Strange (2004). There is some empirical evidence that sorting across space matters to understand the wage distribution This literature uses detailed data on workers' characteristics or a fixed effect approach to control for worker heterogeneity and sorting in a reduced form analysis (Combes et al. (2008), Mion and Naticchioni (2009), Matano and Naticchioni (2011)). I use a structural approach to explicitly account for the sorting of firms.

Finally, the counterfactual policy analysis offers a complementary approach to research that assesses the impact of specific place-based policies. The empirical literature has traditionally focused on estimating the local effects of these policies. A notable exception is Kline and Moretti (2014), who develop a methodology to estimate their aggregate effects.[5] They estimate that, following a local productivity boost, additional positive local effects due to the endogenous creation of agglomeration externalities are offset by losses in other parts of the country. My approach explicitly models the reaction of mobile firms to financial incentives, and I find a negative aggregate effect of policies attracting firms to the smallest cities. In contrast, a desirable policy is one that subsidizes firms to choose larger cities, as the

[4]The model borrows insights from the assignment model literature, such as Costinot and Vogel (2010), and in particular Eeckhout and Kircher (2012) and Sampson (2014), who focus on the matching of heterogenous firms to heterogenous workers. Here, firms match with heterogenous city sizes.

[5]Albouy (2012) focuses on a related question. He argues that federal taxes impose a de facto unequal geographic burden since they do not account for differences in local cost of living, and estimates the corresponding welfare cost. On the measurement of local effects, see, for example, Busso et al. (2013) for the US, Mayer et al. (2015) for France and Criscuolo et al. (2012) for the UK.

decentralized equilibrium features a suboptimal creation of agglomeration externalities. Finally, Glaeser and Gottlieb (2008) study theoretically the economic impact of place-based policies. My analysis brings in heterogeneous firms and the general equilibrium effect of place-based policies on the productive efficiency of the country.

The paper is organized as follows. Section 3 presents the model and its predictions. Section 4 details the empirical analysis. I show salient features from French firm-level data that are consistent with the forces at play in the model. I then structurally estimate the model using simulated method of moments. In section 5, I conduct a counterfactual policy analysis using the estimated model. Section 6 concludes.

3. A Model of the Location Choice of Heterogeneous Firms

Consider an economy in which production takes place in locations that I call cities. Cities are constrained in land supply, which acts as a congestion force. The economy is composed of a variety of sectors. Within sectors, firms are heterogeneous in productivity. They produce, in cities, using local labor and traded capital. Non-market interactions within cities give rise to positive agglomeration externalities. I assume that they have heterogeneous effects on firms, in the sense that more efficient firms are more able to leverage local externalities. Firms' choice of city results from a trade-off between the strength of local externalities and the local level of input prices. Heterogeneous firms face different incentives, which yields heterogeneity in their choice.

I first establish some key properties of an economy with these characteristics. I then close the model and endogenize city formation. To that end, I follow Henderson (1974) and postulate the existence of a class of city developers. In each potential city site, a developer represents local landowners and competes against other sites to attract firms. City developers play a coordination role in the creation of cities, which leads to a unique equilibrium of the economy.[6]

The model describes a long-run steady state of the economy and abstracts from dynamics. All derivations and proofs are detailed in the supplemental material.

3.1. Set-up and agents' problem

3.1.1. Cities

Each city is built on a given stock of land, normalized to 1. All city sites are identical ex-ante. The number of workers that live in a city is noted L. The economy features an distribution of city sizes, which is left arbitrary for now.

In what follows, I index cities and all the relevant city-level parameters by L. City size is sufficient to characterize all the economic forces at play, in the tradition of models of systems of cities pioneered by Henderson (1974). In particular the distance between two cities plays no role as goods produced in the economy are either freely traded between cities within the country, or are, in the case of housing, non

[6]In the model, the one conclusion that relies on the presence of these city developers is the uniqueness of equilibrium. All the other characterizations, and in particular the various comparative statics with respect to city size, remain valid if the mechanism for city formation is left unspecified i.e. if the city size distribution is taken as exogenous.

tradable.[7]

Land is used to build housing, which is divisible and consumed by workers. Atomistic landowners construct housing h^S by combining their land γ with local labor ℓ, according to the housing production function

$$h^S = \gamma^b \left(\frac{\ell}{1-b}\right)^{1-b}. \tag{1}$$

Landowners compete in the housing market, taking both the housing price $p_H(L)$ and the local wage $w(L)$ as given.

3.1.2. Workers

Set-up There is a mass N of identical workers. Each worker is endowed with one unit of labor. A worker lives in the city of his choosing, consumes a bundle of traded goods and housing, and is paid the local wage $w(L)$. Workers' utility is

$$U = \left(\frac{c}{\eta}\right)^{\eta} \left(\frac{h}{1-\eta}\right)^{1-\eta}, \tag{2}$$

where h denotes housing and c is a Cobb-Douglas bundle of goods across S sectors and a CES bundle of varieties within sector, defined as

$$c = \prod_{j=1}^{j=S} c_j^{\xi_j} \text{ , with } \sum_{j=1}^{j=S} \xi_j = 1 \text{ and } c_j = \left[\int c_j(i)^{\frac{\sigma_j-1}{\sigma_j}} di\right]^{\frac{\sigma_j}{\sigma_j-1}}.$$

I denote by $P = \left[\prod_{j=1}^{S} (\frac{P_j}{\xi_j})^{-\xi_j}\right]^{-1}$ the aggregate price index for the composite good c. Since goods are freely tradable, the price index is the same across cities. Workers are perfectly mobile and ex ante identical.

Workers' problem Workers in city L consume $c(L)$ units of the good and $h(L)$ units of housing to maximize their utility (2), under the budget constraint $P\,c(L) + p_H(L)h(L) = w(L)$. Given (1) and the housing market clearing condition, the quantity of housing consumed by each worker in equilibrium in city L is

$$h(L) = (1-\eta)^{1-b} L^{-b}. \tag{3}$$

[7] The assumption of free trade, if convenient, is not necessary here. I show in Appendix C that when trade between cities is costly, city size L is still sufficient to characterize all the economic forces at play. In turn, the characterizations of the economy with heterogeneous firms sorting that I propose below in section 3.2 still hold in the case of costly trade. Note that in this extension, cities of the same population have in equilibrium the same market access. To break this systematic correlation, one would have to introduce more sources of heterogeneity to the model - for instance, heterogeneity in local amenities. This is left for future research.

Housing consumption is lower in more populous cities because cities are constrained in space. This congestion force counterbalances the agglomeration-inducing effects of positive production externalities in cities and prevents the economy from complete agglomeration into one city.

Since workers are freely mobile, their utility must be equalized in equilibrium across all inhabited locations to a level $\bar{U}$. In equilibrium, wages must increase with city size to compensate workers for congestion costs, according to

$$w(L) = \bar{w}((1-\eta)\,L)^{b\frac{1-\eta}{\eta}}, \tag{4}$$

where $\bar{w} = \bar{U}^{\frac{1}{\eta}}\,P$ is an economy-wide constant to be determined in the general equilibrium.

3.1.3. Firms

Production The economy consists of S sectors that manufacture differentiated tradable products. Sectors are indexed by $j = 1, ..., S$. Firms produce varieties using two factors of production that have the following key characteristics. One has a price that increases with city size; the other has a constant price across cities. For simplicity, I consider only one factor whose price depends on city size: labor. In particular, I do not consider land directly in the firm production function. I call the other factor capital, as a shorthand for freely tradable inputs. Capital is provided competitively by absentee capitalists. The price of capital is fixed exogenously in international markets, and the stock of capital in the country adjusts to the demand of firms.[8]

Within their sectors, firms differ exogenously in efficiency z. A firm of efficiency z in sector j and city of size L produces output according to the following Cobb-Douglas production function

$$y_j(z,L) \;=\; \psi(z,L,s_j)\; k^{\alpha_j}\; \ell^{1-\alpha_j},$$

where ℓ and k denote labor and capital inputs, α_j is the capital intensity of all firms in sector j and $\psi(z, L, s_j)$ is a firm-specific Hicks-neutral productivity shifter detailed below. It is determined by firm's 'raw' efficiency, the extent of the local agglomeration externalities and a sector-specific parameter s_j.

Firms engage in monopolistic competition. Varieties produced by firms are freely tradable across space: there is a sectoral price index that is constant through space. Firms take it as given. What matters for location choice is the trade-off between production externalities and costs of production. The relative input price varies with city size. Wages increase with L (equation (4)), whereas capital has a uniform price. Therefore, the factor intensity of a firm shapes, in part, its location decision. A more labor-intensive firm faces, all else equal, a greater incentive to locate in a smaller city where wages are lower.

Productivity and agglomeration The productivity of a firm $\psi(z, L, s_j)$ increases with its own 'raw' efficiency z and with local agglomeration externalities that depend on city size L. The productivity

[8]Featuring this input to production allows the model to capture that input-use intensity is one of the determinants of location choice of firms. Beyond allowing to capture these type of effects, capital is not necessary to build the equilibrium of the model.

function is also indexed by a sector-specific parameter s_j. I explain the roles of these parameters in turn.

A key assumption of the model is that the productivity of a firm $\psi(z, L, s_j)$ exhibits a strong complementarity between local externalities and the 'raw' efficiency of the firm. This assumption is driven by the findings of Combes et al. (2012), who study a wide set of French industries and provide evidence that more efficient firms are disproportionately more productive in larger cities, pointing to such a complementarity as a potential explanation for this fact. Moreover, in Section 4, I present a set of stylized facts on French firms' location and production patterns. They are consistent with sorting, a consequence of the assumed complementarity.

Knowledge spillovers can arguably exhibit this type of complementarity. More efficient firms can better leverage the local information they obtain. A similar idea, though for individual agents, is provided by Davis and Dingel (2012). In their model, more able individuals optimally spend less time producing and more time leveraging local knowledge, which increases their productivity, leading to such a complementarity.

In what follows, I remain agnostic on the source of agglomeration externalities and their specific functional form. This allows me to highlight the generic features of an economy with such complementarities. I let the productivity $\psi(z, L, s)$ have the following properties:

Assumption A *$\psi(z, L, s)$ is log-supermodular in city size L, firm raw efficiency z and sectoral characteristic s. Furthermore, the log-supermodularity in (z, L) is strict, and ψ is is twice differentiable. That is,*

$$\frac{\partial^2 log\, \psi(z, L, s)}{\partial L \partial z} > 0, \quad \frac{\partial^2 log\, \psi(z, L, s)}{\partial L \partial s} \geq 0, \quad and \quad \frac{\partial^2 log\, \psi(z, L, s)}{\partial z \partial s} \geq 0.$$

I introduce a sector-specific parameter s_j that allows sectors to vary in the way they benefit from local urbanization externalities. Rosenthal and Strange (2004) note that empirical studies suggest that the force and scope of agglomeration externalities vary across industries. More specifically, Audretsch and Feldman (1996) suggest that the benefits from agglomeration externalities are shaped by an industry's life-cycle and that highly innovative sectors benefit more strongly from local externalities than mature industries. I index industries such that, in high s sectors, firms benefit from stronger agglomeration forces, for a given city size. In the estimation of the model, I allow for parameter values that shut down the heterogeneous effect between agglomeration externalities and firm efficiency. The specification I retain for ψ nests the typical specification considered in the literature, where only agglomeration forces of the form $\psi = zL^s$ are at play.[9]

Finally, I restrict the analysis to productivity functions $\psi(z, L, s)$ for which the firms' problem is well defined and concave for all firms. In other words, I assume that the positive effects of agglomeration externalities are not too strong compared to the congestion forces. A sufficient condition for this, given that the congestion forces increase with city size with a constant elasticity, is that agglomeration externalities have decreasing elasticity with respect to city size for any firm type z.

Entry and location choice There is an infinite supply of potential entrants who can enter the sector of their choosing. Firms pay a sunk cost f_{Ej} in terms of the final good to enter sector j, then draw a raw

[9] In that case $\frac{\partial^2 log\, \psi(z,L,s)}{\partial L \partial z} = 0$, $\frac{\partial^2 log\, \psi(z,L,s)}{\partial s \partial z} = 0$ and $\frac{\partial^2 log\, \psi(z,L,s)}{\partial L \partial s} > 0$.

efficiency level z from a distribution $F_j(.)$.[10] Once firms discover their raw efficiency, they choose the size of the city where they want to produce. Contrary to the setting in Melitz (2003), the model abstracts from any selection of firms at entry, since there is no fixed cost to produce. I focus instead on *where* firms decide to produce once they discover their efficiency, and how this shapes the spatial equilibrium of the economy. That is, rather than selection on entry, I focus on selection on city size.

Firms' problem A firm's choice of city size is influenced by two factors. First, relative input prices vary by city size. Second, firm productivity increases with city size, through greater agglomeration externalities. The firm's problem can be solved recursively. For a given city size, the problem of the firm is to hire labor and capital and set prices to maximize profits, taking as given the size of the city (and hence the size of the externality term) and input prices. Then, firms choose location to maximize this optimized profit.[11]

Consider a firm of efficiency z producing in sector j and in a city of size L. Firms hire optimally labor and capital, given the relative factor prices $\frac{w(L)}{\rho}$ – where ρ denotes the cost of capital – and their local productivity $\psi(z, L, s_j)$. Firms treat local productivity as exogenous, so that the agglomeration economies take the form of external economies of scale. Given the CES preferences and the monopolistic competition, firms set constant markups over their marginal cost. This yields optimized profits for firm z in sector j as a function of city size L

$$\pi_j(z, L) = \kappa_{1j} \left(\frac{\psi(z, L, s_j)}{w(L)^{1-\alpha_j}} \right)^{\sigma_j - 1} R_j \, P_j^{\sigma_j - 1}, \tag{5}$$

where P_j is the sectoral price index, R_j is the aggregate spending on goods from sector j and $\kappa_{1j} = \frac{((\sigma_j-1)\alpha_j^{\alpha_j}(1-\alpha_j)^{1-\alpha_j}(\rho P)^{-\alpha_j})^{\sigma_j-1}}{\sigma_j^{\sigma_j}}$ is a sector-specific constant.

Note that firm employment, conditional on being in a city of size L, is given by

$$\ell_j(z, L) = (1 - \alpha_j)(\sigma_j - 1)\frac{\pi_j(z, L)}{w(L)}. \tag{6}$$

The proportionality between profits and the wage bill is a direct consequence of constant factor shares, implied by the Cobb-Douglas production function, and of constant markup pricing. The problem of the firm thus is to choose the city size L to maximize (5).

3.2. Properties

I examine here the properties of an equilibrium of this economy, taking as given an arbitrary support of city sizes $\mathcal{L}$. These properties are robust to extending the model to imperfect sorting or costly trade, as I detail at the end of the analysis.

[10] I assume that this distribution is an interval (possibly unbounded) on the real line. This assumption is made for tractability; the results carry through without it, although the notation is more cumbersome.

[11] In reality, there are two types of sorting: ex-post sorting - that is, firms that are already established and decide to change location - and ex-ante sorting - that is, new firms, or new establishments of an existing firm, being created somewhere. Since the model is static, it conflates both types of sorting.

3.2.1. Firm Sorting

Firms choose city size to maximize profits given by (5). This is summarized by the following first-order condition[12], given the wage in equation (4) and writing $\psi_2(z, L, s_j) = \frac{\partial \psi(z,L,s_j)}{\partial L}$:

$$\frac{\psi_2(z, L, s_j)\, L}{\psi(z, L, s_j)} = (1 - \alpha_j)\, b\, \frac{1 - \eta}{\eta} \tag{7}$$

This condition states that the elasticity of productivity to city size is equal to the elasticity of labor costs with respect to city size. At the optimal city size for a given firm, its marginal gain to choosing a larger city equals the marginal cost of doing so. This first-order condition defines implicitly the matching function:

$$L_j^*(z) = \arg\max_{L \in \mathcal{L}} \pi_j(z, L). \tag{8}$$

There is a unique profit-maximizing city size for a firm of type z in sector j, under the regularity conditions I have assumed. Furthemore, it is readily seen from (5) that the profit function of the firm inherits the log-supermodularity of the productivity function in z and L. Therefore, the following lemma holds.

Lemma 1 *The matching function $L_j^*(z)$ is non-decreasing in z.*

This result comes from a classic theorem in monotone comparative statics (Topkis (1998)). The benefit to being in larger cities is greater for more productive firms and only they are willing in equilibrium to pay the higher factor prices there. Furthermore, the matching function is fully determined by the firm maximization problem, conditional on the set of city sizes $\mathcal{L}$. This optimal choice does not depend on general equilibrium quantities that enter the profit function proportionally for all city sizes.

3.2.2. Within-sector patterns

Within a given sector j, the revenue, production and employment distributions are all determined by the matching function $L_j^*(z)$. In the sorting equilibrium, for a firm of efficiency z, productivity, revenues and employment are given by

$$\psi_j^*(z) = \psi(z, L^*(z), s_j),$$

$$r_j^*(z) = \sigma_j \kappa_{1j} \left(\frac{\psi(z, L^*(z), s_j)}{w(L_j^*(z))^{1-\alpha_j}} \right)^{\sigma_j - 1} P_j^{\sigma - 1} R_j, \tag{9}$$

$$\ell_j^*(z) = \kappa_{2j} \frac{\psi(z, L^*(z), s_j)^{\sigma - 1}}{w(L_j^*(z))^{(\sigma-1)(1-\alpha_j)+1}} P_j^{\sigma_j - 1} R_j, \tag{10}$$

where the starred variables denote the outcomes in the sorting equilibrium. Since there is positive assortative matching between a firm's raw efficiency and city size (lemma 1), firm-level observables also exhibit complementarities with city size.

[12] I assume here for simplicity that $\mathcal{L}$ is a convex set, so that one can take derivative. I examine the case where $\mathcal{L}$ is not convex (e.g., a discrete collection of city sizes) in the Appendix

Proposition 2 *In equilibrium, within each sector, firm revenues, profits and productivity increase with city size, in the following sense. For any $L_H, L_L \in \mathcal{L}$ such that $L_H > L_L$, take z_H such that $L^*_j(z_H) = L_H$ and $L^*_j(z_L) = L_L$. Then, $r^*_j(z_H) > r^*_j(z_L)$, $\pi^*_j(z_H) > \pi^*_j(z_L)$, and $\psi^*_j(z_H) > \psi^*_j(z_L)$.*

These strong predictions on the ranking of the size of firms (in revenues or productivity) vis a vis the city size are a direct consequence of the perfect sorting of firms. In contrast, employment can be either positively or negatively associated with city size through the effect of wages. Within a sector, $\ell^*(z) \propto \frac{r^*(z)}{w(L^*(z))}$, where both revenues and wages increase with city size. Firms may have lower employment in larger cities, even though they are more productive and profitable. More precisely, if $\epsilon_l = \frac{d \log \bar{\ell^*}(L)}{d \log L}$ and $\epsilon_r = \frac{d \log \bar{r^*}(L)}{d \log L}$ are the elasticities of mean employment and mean revenues with respect to city size in equilibrium, then ϵ_l is not necessarily positive, since:

$$\epsilon_l = \epsilon_r - b\frac{1-\eta}{\eta},, \tag{11}$$

3.2.3. Comparative statics across sectors

I now compare the predicted distribution of firm outcomes across sectors. Sectors differ in their capital intensity α_j and in the strength of their benefit from agglomeration externalities s_j. Both impact the sorting process, leading in turn to differences in observed outcomes. The following comparative statics exercises examine how the geographic and size distribution of firms in a sector vary with each parameter holding all other sectoral characteristics constant, in particular the distribution of raw efficiencies $F(.)$.[13]

Geographic distribution Define the *geographic distribution* of firms in a sector as the probability that a firm from the sector is in a city of size smaller than L. That is, let

$$\tilde{F}(L; \alpha_j, s_j) = P(\text{firm from sector } (\alpha_j, s_j) \text{ is in a city of size smaller that } L).$$

Proposition 3 *The geographic distribution $\tilde{F}_j$ of a high α_j sector first-order stochastically dominates that of a lower α_k sector, all else equal. The geographic distribution $\tilde{F}_j$ of a high s_j sector first-order stochastically dominates that of a lower s_k sector, all else equal.*

These results stem from the following observation. As shown before, the matching function $L^*_j(z)$ is always increasing, but its slope and absolute level depend on the capital intensity α_j and the strength of agglomeration externalities s_j in the sector. In labor-intensive sectors, the weight of the wage effect is heavier in the trade-off between the benefits of agglomeration externalities and labor costs. This pushes the matching function down, towards smaller cities. For any city size threshold, there are more firms from a labor-intensive sector that choose to locate in a city smaller than the threshold. In contrast, in sectors with strong agglomeration externalities, firms benefit more from a given city size, which pushes the matching function up for all firms. All else equal, they locate more in larger cities.

[13] Note that if firms where not heterogeneous, both these distributions would be degenerate. In every sector, there would be one firm size, and firms would be located in one city size.

Firm-size distribution The intensity of sorting, which reinforces initial differences between firms, impacts the dispersion of the observed sectoral firm-size distribution. Let $Q_j(p)$ denote the p-th quantile of the firm revenue distribution in sector j.

Proposition 4 *All else equal, if $(\alpha_j, s_j) \geq (\alpha_i, s_i)$, the observed firm-size distribution in revenues is more spread in Sector j than in Sector i. For any $p_1 < p_2 \in (0, 1)$, $\frac{Q_i(p_2)}{Q_i(p_1)} \leq \frac{Q_j(p_2)}{Q_j(p_1)}$.*

In other words, if one normalizes the median of the revenue distribution to a common level, all higher quantiles in the revenue distribution of Sector 2 are strictly higher than in Sector 1, and all lower quantiles are below. This comes from the fact that the distribution of firm revenues is shaped not only by the distribution of raw efficiencies (held constant across sectors in this comparative statics exercise), but also by the complementarity between z and city size, whose choice is endogeneous. In higher-s sectors for example, there is more to be gained by more productive firms to locate in larger cities. As a consequence, the difference in city size choice, and in turn in firm revenues, is larger between high- and low-z firms in a higher-s sector compared to a lower-s sector. The distribution of firms outcomes is more unequal.

In particular, higher α_j or higher s_j sectors have thicker upper-tails in their firm-size distributions. This leads to a characterization that will prove useful empirically. Firm-size distributions are empirically well approximated by power law distributions, in their right tail. The exponent of this distribution characterizes the thickness of the tail of the distribution. Assume that the revenue distribution of firms in two sectors 1 and 2 can be approximated by a power law distribution in the right tail, with respective exponents ζ_1 and ζ_2. Then the following corollary holds:

Corollary 5 *Let $(\alpha_2, s_2) \geq (\alpha_1, s_1)$. The tail of the firm-size distribution in Sector 2 is thicker than the tail of the firm-size distribution in sector 1: $\zeta_2 \leq \zeta_1$.*

3.2.4. Model Extensions

I detail several extensions to the baseline setup in the supplemental material. The general properties described in this section are robust to these extensions. In particular, the model can be extended to feature *costly trade* between cities. Even with costly trade and unequal price indexes across cities, I show that city size L remains a sufficient statistic for the local economic conditions. The result that more efficient firms sort into larger cities still obtains.

Second, I examine the properties of the model in the presence of imperfect sorting as hypothesized in the empirical specification of section 4. The properties described above are either unchanged, or hold true on average, rather than systematically, in that case.

3.3. Closing the Model

In this section, I specify a mechanism for endogenous city formation, and close the model. I postulate the existence of a class of large local players who develop cities. This leads to a unique equilibrium of this economy, characterized by a specific distribution of city sizes.

3.3.1. City developers

Set-up There is one city-developer for each potential city site. City developers fully tax local landowners. They are therefore the residual claimants on local land value. Their objective is to maximize these revenues, net of the cost of the policies they put in place. They compete to attract firms to their city by subsidizing firms' profits. Absent these developers, there would be a coordination failure as atomistic agents alone - firms, workers or landowners - cannot create a new city. City developers are, in contrast, large players at the city level. As in Henderson (1974), city developers act as a coordinating device that allows a unique equilibrium to emerge in terms of city-size distribution. There is perfect competition and free entry among city developers, which drives their profits to zero in equilibrium.

City developers' problem Each city developer announces a subsidy to local firms' profits in sector j, $T_j(L)$, which may depend on city size L.[14] Developers are funded by fully taxing profits made on the housing market. As the housing market clears in each city, aggregate landowner profits at the city level are

$$\pi_H(L) = b(1-\eta)Lw(L). \tag{12}$$

It will prove useful when solving for the equilibrium to note that a constant share of the local labor force is hired to build housing, namely

$$\ell_H(L) = (1-b)(1-\eta)L. \tag{13}$$

A city developer developing a city of size L faces the following problem:

$$\begin{aligned} &\max_{\{T_j(L)\}_{j\in 1,\dots,S}} \Pi_L = b(1-\eta)w(L)L - \sum_{j=1}^{S} \int_z T_j(L)\pi_j(z,L)\,\mathbb{1}_j(z,L)dF_j(z), \\ &\text{such that} \\ &\qquad \mathbb{1}_j(z,L) = 1 \qquad \text{if firm } z \text{ chooses their city}, \\ &\qquad \mathbb{1}_j(z,L) = 0 \qquad \text{otherwise.} \end{aligned} \tag{14}$$

In this expression, $F_j(.)$ is the distribution of firm's raw efficiencies in sector j and $\pi_j(z,L)$ is the local profit of a firm of efficiency z in sector j, as defined in (5).

Note that subsidies enter the firm's optimal location choice problem:

$$\max_L (1+T_j(L))\pi_j(z,L). \tag{15}$$

Hence, one may worry that they impact firms' choices and the properties described in section 3.2. As I show below though, equilibrium subsidies are independent of L, hence they do not alter the location choice of firms compared to a world without subsidy, nor the distribution of firm profits (up to a multiplicative constant). That is, the city developers subsidies allow the economy to coordinate on a unique equilibrium without affecting the sorting of firms.

[14] Local policies often offer a reduced corporate tax rate for eligible firms, which is equivalent to this profit subsidy. Alternatively, one could consider firm-specific subsidies rather than ad-valorem subsidies. The equilibrium I find here is still an equilibrium in this case, as shown in online Appendix D.2.

3.3.2. Spatial equilibrium

Having set up the problems of workers, firms, landowners and city developers, I am now ready to solve for the equilibrium of the economy. I show that this equilibrium exists and is unique.

Definition 1 *An equilibrium is a set of cities $\mathcal{L}$ characterized by a city-size distribution $f_L(.)$, a wage schedule $w(L)$, a housing-price $p_H(L)$ and for each sector $j = 1, ..., S$ a location function $L_j(z)$, an employment function $\ell_j(z)$, a capital-use function $k_j(z)$, a production function $y_j(z)$, a price index P_j and a mass of firms M_j such that*

(i) workers maximize utility (equation (2)) given $w(L), p_H(L)$ and P_j,

(ii) utility is equalized across all inhabited cities,

(iii) firms maximize profits (equation (5)) given $w(L), \rho$ and P_j,

(iv) landowners maximize profits given $w(L)$ and $p_H(L)$,

(v) city developers choose $T_j(L)$ to maximize profits (equation (14)) given $w(L)$ and the firm problem,

(vi) factors, goods and housing markets clear; in particular, the labor market clears in each city,

(vii) capital is competitively allocated, and

(viii) firms and city developers earn zero profits.

In what follows, I present a constructive proof of the existence of a such an equilibrium. Furthermore, I show that the equilibrium is unique, and stable. As is standard in the literature, I allow for the possibility of a non-integer number of cities of any given size (see Abdel-Rahman and Anas (2004) for a review and more recently Rossi-Hansberg and Wright (2007b) or Behrens et al. (2014)).

Proposition 6 *There exists a unique equilibrium of this economy.*

The equilibrium is unique in terms of distribution of outcomes, such as firm-size distribution, city-size distribution and matching functions between firms and city sizes. It is not, of course, unique in terms of *which site* is occupied by a city of a given size, as all sites are identical ex ante.

3.3.3. Constructing the spatial equilibrium

The equilibrium is constructed in four steps. First, I solve for the equilibrium subsidy offered by city developers. Second, I show that it pins down how firms match with city sizes, as well as the set of city sizes generated in equilibrium by city developers. Third, general equilibrium quantities are determined by market clearing conditions and free entry conditions in the traded goods sectors, once we know the equilibrium matching function from step 2. Finally, the city-size distribution is determined by these quantities, using labor-market clearing conditions. In each step, the relevant functions and quantities are uniquely determined; hence, the equilibrium is unique.

Step 1: Equilibrium subsidy

Lemma 7 *In equilibrium, city developers offer and firms take-up a constant subsidy to firms' profit* $T_j^* = \frac{b(1-\eta)(1-\alpha_j)(\sigma_j-1)}{1-(1-\eta)(1-b)}$ *for firms in sector* j*, irrespective of city size* L *or firm type* z*.*

Formally, each city developer announces a subsidy $T_j^*\delta(L-L_i)$ for a city size L_i where $\delta(0)=1$, and $\delta(x)=0$ for $x \neq 0$. A city developer is indifferent between attracting firms from one or many sectors. Therefore, there is an indeterminacy in equilibrium as to which sector(s) each city developer targets.

I sketch the proof in the case of an economy with only one traded goods sector. The formal proof with S sectors follows the same logic and is given in online Appendix B. City developers face perfect competition, which drives their profits down to zero in equilibrium. Their revenues correspond to the profits made in the housing sector (equation (12)), which are proportional to the aggregate wage bill in the city $w(L)L$. They compete to attract firms by subsidizing their profits. In equilibrium, irrespective of which firms choose to locate in city L, these profits will also be proportional to the sectoral wage bill $w(L)\tilde{L}$, where $\tilde{L}$ is the labor force hired in the traded goods sector locally, as can be seen from equation (6). Finally, the local labor force works either in the housing sector (equation (13)) or the traded goods sector, so that $\tilde{L} = L(1-(1-b)(1-\eta))$. Profits given by (14) simplify to $b(1-\eta)w(L)L - T\frac{(1-(1-b)(1-\eta))}{(\sigma-1)(1-\alpha)}w(L)L$. The choice of city size is irrelevant, and T^* is the only subsidy consistent with zero profits. City developers that offer lower subsidies will not attract any firm, hence will not create cities. City developers that offer higher subsidies attract firms but make negative profits. Note that the equilibrium subsidy does not depend on city size, hence equilibrium subsidies do not alter the location choice of firms compared to a world without subsidy, conditional on the same cities existing.

Step 2: Equilibrium city sizes and the matching function The city developers' problem determines the equilibrium city sizes generated in the economy. Cities are opened up when there is an incentive for city developers to do so, i.e. when there exists a set of firms and workers that would be better off choosing this city size. Workers are indifferent between all locations, but firms are not, since their profits vary with city size. Given the equilibrium subsidy T_j^* offered by city developers, firms choose city size to maximize (15). There is a unique profit-maximizing city size for a firm of type z in sector j, under the regularity conditions I have assumed. Define the optimal city size as follows

$$L_j^{**}(z) = \arg\max_{L\geq 0} \pi_j^*(z,L). \tag{16}$$

Assume that, for some firm type z and sector j, no city of size $L_j^{**}(z)$ exists. There is then a profitable deviation for a city developer on an unoccupied site to open up this city. It will attract the corresponding firms and workers, and city developers will make a positive profit by subsidizing firms at a rate marginally smaller than T_j^*. The number of such cities adjusts so that each city has the right size in equilibrium. This leads to the following lemma, letting $\mathcal{L}$ denote the set of city sizes in equilibrium:

Lemma 8 *The set of city sizes* $\mathcal{L}$ *in equilibrium is the optimal set of city sizes for firms.*

Given this set of city sizes, the optimal choice of each firm is fully determined. Under the regularity assumptions made on ψ as well as on the distribution of z, $F_j(.)$, the optimal set of city sizes for firms in a given sector is an interval (possibly unbounded). The sectoral matching function is invertible over this support. For a given sector, I use the notation $z_j^*(L)$ to denote the inverse of $L_j^*(z)$. It is increasing in L. The set of city sizes $\mathcal{L}$ available in equilibrium is the union of the sector-by-sector intervals.

Step 3: General equilibrium quantities The equilibrium has been constructed up to the determination of general equilibrium values. The reference level of wages $\bar{w}$ defined in equation (4) is taken as the numeraire. The remaining $2S+1$ unknowns are the aggregate revenues in the traded goods sector R, the mass of firms M_j and the sectoral price indexes P_j for all $j \in \{1, ..., S\}$. As detailed in online Appendix E, these are uniquely pinned down by the system of $2S+1$ equations corresponding to the free entry conditions for firms, the goods market clearing conditions in each sector, and the national labor market clearing condition.

Step 4: Equilibrium city-size distribution The city developers' problem and the firms' problem jointly characterize (1) the set of city sizes that necessarily exist in equilibrium and (2) the matching function between firm type and city size. Given these, the city-size distribution is pinned down by the labor market clearing conditions. The population living in a city of size smaller than any L must equal the number of workers employed by firms that have chosen to locate in these same cities, plus the workers hired to build housing. That is,$\forall L > L_{min}$,:

$$\int_{L_{min}}^{L} u f_L(u)\, du = \sum_{j=1}^{S} M_j \int_{z_j^*(L_{min})}^{z_j^*(L)} \ell_j(z, L_j^*(z))\, dF_j(z) + (1-\eta)(1-b) \int_{L_{min}}^{L} u f_L(u)\, du,$$

where $L_{min} = \inf(\mathcal{L})$ the smallest city size in the equilibrium.

Differentiating this with respect to L and dividing by L on both sides gives the city size density ($f_L(L)$ is not normalized to sum to 1)

$$f_L(L) = \kappa_4 \frac{\sum_{j=1}^{S} M_j \mathbb{1}_j(L)\, \ell_j(z_j^*(L))\, f_j(z_j^*(L))\, \frac{dz_j^*(L)}{dL}}{L}, \tag{17}$$

where $\kappa_4 = \frac{1}{1-(1-\eta)(1-b)}$ and $\mathbb{1}_j(L) = 1$ if sector j has firms in cities L, and 0 otherwise. The equilibrium distribution of city sizes $f_L(.)$ is uniquely determined by equation (17), hence the following lemma:

Lemma 9 $f_L(.)$ *is the unique equilibrium of this economy in terms of the distribution of city sizes.*

Several remarks are in order here. First, the city-size distribution is shaped by the distribution of firm efficiency and by the sorting mechanism. This offers a static view of the determination of the city-size distribution, driven by heterogeneity in firm types. In the empirical exercise, I compute the city-size distribution obtained with equation (17), where firm heterogeneity is estimated from French firm-level data but the city-size distribution is not used in the estimation. It exhibits Zipf's law, consistent with the data on cities.

Second, for each city size, the share of employment in each sector can be computed using the same method, now sector by sector. For a given city size, the average sectoral composition over all cities of a given size size L is determined by the model. On the other hand, the model is silent on the sectoral composition of any *individual* city of size L, which is irrelevant for aggregate outcomes. City developers in particular are indifferent to the sectoral composition of their city.[15]

Third, the equilibrium features cities that host a variety of sectors, and sectors that spread out over a range of cities of different sizes. This is contrast to classic urban models that rely on homogenous firms and hence generally feature fully specialized cities and a single city size for each sector.[16]

Finally, I verify in the online Appendix F that this equilibrium is stable and provide there a detailed discussion of stability.

3.3.4. Welfare analysis

To close this theoretical analysis, I explore the welfare properties of this equilibrium. This is complemented in section 5 by a set of quantitative counterfactual policy analyses, in which I quantify the welfare implications of typical spatial policies.

I consider the problem of a benevolent social planner who freely chooses allocations in this economy so as to maximize total welfare in spatial equilibrium, i.e. under the constraint of free mobility of workers.[17] In order to focus on the inefficiencies that arise in the traded good sector, I take the housing sector as given in what follows, i.e. a constant fraction of the local labor force is used to build housing as in the competitive equilibrium. Welfare could potentially be improved beyond what is laid out here through an intervention on the housing market.[18]

The problem of the social planner, formally stated and solved in online Appendix G, is to choose allocations of firms and workers - in particular, she chooses firm's location, firm's employment and firm's production, as well as the consumption and location of workers, so as to maximize welfare. In her choice of city sizes she faces a trade-off between increasing productive efficiency by creating larger cities to leverage agglomeration externalities on the one hand, and limiting the disutility from congestion borne by workers on the other hand. The first-order condition for the location choice of firm z in the social planner's problem writes:

$$\frac{\psi_2(z,L,s)L}{\psi(z,L,s)} = b\frac{1-\eta}{\eta}(1-\alpha)\chi(z), \tag{18}$$

where $\chi(z) < 1$ is a wedge between the first-order condition in the market equilibrium (7) and the one of the social planner. Firms choose cities that are too small in the market equilibrium relative to the

[15]This city-level indeterminacy comes from the fact that agglomeration externalities depend on the overall size of the city, and not on its sectoral composition. To lift this, the model could easily be extended to accommodate *localization* externalities. The agglomeration externality depends in that case on the size of a given (set of) sector(s), and not the overall city size. Cities would then be perfectly specialized in that sector(s), since the congestion costs depend on the overall city size, but the benefits are sector(s)-specific. This would not change any other characterization of the equilibrium. In particular, the city-size distribution defined in equation (17) and lemma 9 would still hold.

[16]A notable exception is Helsley and Strange (2014) who propose a model of coagglomeration of a range of industries in cities.

[17]Total welfare could be further improved if this constraint of equal utility of workers across inhabited cities was lifted. I do not consider this case since this equilibrium would not be stable. Some workers would always have an incentive to move to increase their utility. This alternative equilibrium would also raises equity issues as identical workers would have different levels of utility in equilibrium.

[18]A benefit of this approach is that these results hold irrespective of the source of congestion that I consider, as long as the congestion force increases log-linearly with city size.

optimal spatial equilibrium.[19]

Proposition 10 *The equilibrium is suboptimal. Firms locate in cities that are too small. To reach the optimum, the first-best policy taxes wages offered by firms in smaller cities and subsidizes firms' wages in larger cities, according to a tax/subsidy schedule that varies monotonically with city size.*

The intuition for this result is as follows.The social marginal benefit of choosing a larger city is higher than the private benefit perceived by firms through their profit function. The benefit that is not internalized by firms is that by choosing a larger city they increase the productivity of the economy as a whole which decreases the cost of entry for all firms. Fostering entry increases welfare, by the love of variety effect. Firms ignore the effect of their choice of city size on the cost of entry, and therefore choose cities that are too small compared to the optimum. This general equilibrium cross-city effect is not internalized by firms nor by city developers who, despite being large local players, are still atomistic at the national level.

To align firms' incentives to the solution to the social planner's problem, wages have to be subsidized/taxed so the wage schedule paid by firms is

$$w(L) \propto (L^{b\frac{1-\eta}{\eta}} + A) \tag{19}$$

where A is a constant, as opposed to $w(L) \propto L^{b\frac{1-\eta}{\eta}}$ in the competitive equilibrium.

This analysis helps see that "threshold-type" spatial policies, in which firms are subsidized when they locate in cities smaller than a given size, are not intuitively welfare enhancing. They tend to distort the choice of city size in the wrong direction. They attract firms to cities that are smaller (rather than larger) than the one they choose in the competitive equilibrium. Desirable policies on the other hand are ones that tend to flatten out the wage schedule, making the wage increase less fast with city size than it would otherwise. I explore these points further, quantitatively, in section 5.

4. Estimation of the Model

I now take the model to the data, in order to be able to perform a quantitative policy analysis. Using French firm-level data, I first show that sectors display location patterns and firm-size distribution characteristics that are consistent with the theoretical predictions. I then structurally estimate the model.

4.1. Data

The firm-level data set of French firms that I use contains information on the balance sheets of French firms, declared for tax purposes. All firms with revenues over 730,000 euros are included. It reports information on employment, capital, value added, production, and 3-digit industry classification. It is matched with establishment-level data, which indicate the geographical location at the postal code level of each establishment of a given firm-year. As is standard in the literature, the geographical areas I use to measure city size are the 314 French commuting zones, or *"Zones d'emploi"* (employment zones), within

[19] See also Albouy et al. (2016) for a discussion of another theoretical setting where large cities can be too small.

metropolitan France. They are defined with respect to the observed commuting patterns of workers and cover all of France. They are designed to capture local labor markets and are better suited than administrative areas, which they abstract from, to capturing the economic forces at play in the model.[20] To measure the size of the city, I use the total local employment of the area, since I need a proxy for externalities such as knowledge spillovers or labor market pooling that depend on the size of the workforce. I use the data for the year 2000 in the estimation procedure.

I retain only tradable sectors in the analysis, consistent with the assumptions of the model. The set of industries is the one considered in Combes et al. (2012), i.e., manufacturing sectors and business services, excluding finance and insurance[21], which correspond to 157,070 firms. Summary statistics are reported in Table 1.

4.2. Descriptive evidence on sorting

Before proceeding to the structural estimation of the model, I present a first look at the raw data. My objective is to check that the comparative statics of the model are broadly consistent with the patterns exhibited in the data. Recall that in the model, the complementarity between firm efficiency and agglomeration forces leads to the sorting of firms across cities of different sizes. This impacts the elasticity of firm-level observables with respect to city size within industries (prop. 2). Furthermore, firm sorting is shaped by two key sectoral parameters, namely, the sectoral strength of agglomeration externalities s_j as well as the sectoral intensity of use of traded inputs α_j. The model shows how these parameters shape (i) the location patterns of firms in a given industry (prop. 3) and (ii) the dispersion of the sectoral firm-size distribution (prop. 4). I turn to examining the raw data in these dimensions.

To do so, I use the most disaggregated level of industry available in the data. I keep sectors with more than 200 observations, for a total of 146 industries, and present correlations between different sectoral characteristics, guided by the theory. These correlations could be driven by explanations alternative to the ones I propose in the model. To mitigate these concerns, I check that the patterns I find are robust to a set of sectoral controls that I detail below. The broad consistency of the data with the salient features of the model are only suggestive evidence that sorting forces may be at play.

I first investigate how, in each sector, average firm value added and average firm employment change as city size increases.[22] In the model, the elasticity of firm revenues to city size is positive within industries whereas the elasticity of employment to city size is strictly lower and possibly negative. Empirically, I compute the average firm-level value added and employment by industry and city and compute their elasticity with respect to city size.[23] Figure 1 plots the distributions of these elasticities. The elasticity of employment to city size almost always lies below the elasticity of value added to city size. For value

[20] They are presented as areas where *"most workers live and work, and where establishments can find most of their workforce"*. The previous definition of these zones was constrained by some administrative borders. I use a new definition of these zones published by INSEE in 2011 that abstracts from these constraints. These zones are a collection of towns ("code commune"). I use a concordance table between these and postal codes to classify the firm data in terms of zone d'emploi.

[21] for which establishment-level data is not available

[22] Because the model does not feature the use of intermediates, I use value added as the measure of firm output.

[23] For this measure, I restrict the sample to single-establishment firms as the data on value added is only available at the firm level for multi-establishment firms. Single establishment account for 83% of firms and 44% of employment in the sample. I verify using the full sample that, reassuringly, the employment/city size relationship is not systematically biaised for multi-establishment firms compared to mono-establishment firms.

added, it is positive for 85% of industries, corresponding to 93% of firms, and is significantly negative for only one industry, the manufacture of kitchen furniture. This is broadly consistent with the intuition of the model.

Second, the model suggests that firm location choices are linked to sectoral characteristics and, in particular, the intensity of input use, which I can measure in the data. To proxy for inputs whose price does not systematically increase with city size, I use a measure of "tradable capital" defined as total capital net of real estate assets. I measure tradable capital intensity, α^K, as the Cobb-Douglas share of this tradable capital in value added. I then run the following regression:

$$share_j = \beta_0 + \beta_1 \alpha_j^K + \beta_2 X_j + \epsilon_j,$$

where j indexes sectors, $share_j$ is the share of establishments in sector j located in large cities (i.e., the largest cities that hosts half of the population) and X_j is a set of control variables varying at the industry level. Table 2 reports the coefficient estimates. It shows that industries that use more tradable capital are significantly more likely to be located in larger cities. However, these industries could also be the ones with higher skill intensity, driven to larger cities in search of skilled workers. To control for that, I use an auxiliary data set to measure industry-level skill intensity.[24] Specification (III) in Table 2 shows that controlling for industry-level skill intensity does not affect the results. Specification (IV) runs the same regression, limiting the sample to export-intensive industries. This control aims at mitigating the concern that location choice may be driven by demand-side explanation, whereas the model focuses on the supply side. Again, the results are robust to using this reduced sample. Overall, Table 2 is consistent with the idea that firms location choices are shaped by the intensity of input use in their industry.[25]

Third, the model predicts that firms that locate in large cities benefit disproportionally from agglomeration externalities. As a consequence, the sectoral firm-size distribution is more fat-tailed for industries located in larger cities. Table 3 correlates the thickness of the industry-level firm-size distribution, summarized by its shape parameter ζ_j, with the share of establishments located in large cities in industry j.[26] Table 3 shows a negative correlation between ζ_j and the fraction of establishments in industry j located in large cities (defined as in Table 2). In other words, industries that locate more in large cities also have thicker-tailed firm-size distributions. This negative correlation is robust to controlling for the number of firms and the average value added in industry i (Specification II), as well as reducing the sample to export-intensive industries (Specification III).

Finally, the model predicts that more efficient firms self-select into larger cities. I investigate this question by focusing on the relocation pattern of movers, i.e. mono-establishment firms that change location from one year to the next. The nature of this question leads me to extend the sample period

[24] I use the random sample of 1/12 of the French workforce published by INSEE. It contains information on workers' skill level, salary and industry. For each industry, I measure the share of the labor force that is high-skilled. I define a dummy variable that equals one for sectors with above-median skill-intensity.

[25] I further check that these results are robust to alternative specifications of what constitute a large city (top cities hosting 25% and 33% of the workforce), as well as to dropping business services, which are arguably less tradable, altogether from the analysis.

[26] The shape parameter ζ_j is estimated by running the following regression, following Gabaix and Ibragimov (2011): $\log\left(rank_{ij} - \frac{1}{2}\right) = \alpha_j - \zeta_j \log(\text{value added}_{ij}) + \epsilon_{ij}$, where j indexes industries, i indexes firms and $rank_{ji}$ is the rank of firm i in industry j in terms of value added.

to 1999-2006. There is no direct way to measure a firm's raw efficiency from the data. However, in the model, within a city-industry pair, firm revenues increase with firm efficiency. I thus compute the following firm-level residual ω_{ijt} and use it to proxy for firm efficiency:

$$\log(\text{value added}^c_{ijt}) = \delta^c + \delta^t + \delta^j + \omega_{ijt},$$

where δ^c, δ^t and δ^j are sets of, respectively, city, year and industry fixed-effects, and i is a firm in industry j located in city c in year t. For all firms relocating from year t to $t+1$, I define $\Delta_t\text{City Size}_i$ as $\log(L_{i,t+1}/L_{i,t})$, where $L_{i,t}$ is the size of the city where firm i is located in year t. I then estimate:

$$\Delta_t\text{City Size}_i = \alpha + \beta\omega_{ijt} + X_{it} + \epsilon_{it}$$

where X_{it} includes an industry fixed effect and the logarithm of $L_{i,t}$ or a set of initial city fixed effects.[27] Table 4 shows that, conditional on moving, firms that are initially larger tend to move into larger cities. Similar results obtain when I drop firms that switch industry when they move. I emphasize that this result is a simple correlation and cannot be interpreted causally in the absence of a valid instrument for the selection into the sample of movers. Table 4 simply shows that, among the set of movers, there exists a positive correlation between initial firm size and the size of the city the firm moves into, a correlation pattern that is consistent with sorting.

4.3. Structural estimation

I now turn to the estimation of the model. The model is estimated industry by industry, on 23 aggregated industries. I minimize the distance between moments of the data and their simulated counterparts to estimate the sectoral parameters that govern the model.

4.3.1. Model specification

I first lay out the econometric specification of the model. The literature has traditionally assumed that agglomeration externalities were of the form $\psi(z, L, s_j) = zL^{a_j}$, where a_j measures the strength of externalities. In such a framework, firm productivity is not log-supermodular in z and L. In contrast, I postulate the following functional form of the productivity function, for a firm i of raw efficiency z_i operating in sector $j \in 1 \ldots S$:

Assumption B

$$\log(\psi_j(z_i, L; s_j, a_j)) = a_j \log L + \log(z_i)(1 + \log \frac{L}{L_o})^{s_j} + \epsilon_{i,L} \quad \textit{for } \log(z_i) \geq 0 \textit{ and } L \geq L_o \tag{20}$$

$$\log(\psi_j(z_i, L; s_j, a_j)) = 0 \quad \textit{for } L < L_o$$

The parameter a_j measures the classic log-linear agglomeration externality. The strength of the complementarity between agglomeration externalities and firm efficiency is captured by s_j. When $s_j = 0$, the

[27] These controls absorb the mechanical relationship by which firms in large (resp. small) cities are more likely, conditional on moving, to move to smaller (resp. larger) cities.

model nests the traditional model of agglomeration externalities without complementarity. L_o measures the minimum city size below which a city is too small for a firm to produce in. In what follows, I write $\tilde{L} = \frac{L}{L_o}$, and $\mathcal{L}$ the set of normalized city sizes in the simulated economy.[28] I assume that $\log(z)$ is distributed according to a normal distribution with variance $\nu_{Z,j}$, truncated at its mean to prevent $\log(z)$ from being negative. This restriction is needed for the productivity of firms to be increasing in city size in the specification of assumption (B). I introduce an error structure by assuming that firms draw idiosyncratic productivity shocks $\epsilon_{i,L}$ for each city size, where $\epsilon_{i,L}$ is i.i.d. across city sizes and firms. It is distributed as a type-I extreme value, with mean zero and variance $\nu_{R,j}$. This shock captures the fact that an entrepreneur has idiosyncratic motives for choosing a specific location: for example, he could decide to locate in a city where he has a lot of personal connections that make him more efficient at developing his business. These idiosyncratic motives for location generate imperfect sorting. The predictions of the theory are still relevant to the case of imperfect sorting, once adapted to reflect the fact that they hold for firms *on average* within a city size rather than systematically for all firms in a given city, as shown formally in online Appendix C.

I assume that idiosyncratic shocks are city-size specific, with mean zero and a constant variance across city size bins, and not city-specific as would perhaps be more natural. Still, these shocks can themselves represent the maximum of shocks at a more disaggregated level (e.g., at the city level). The maximum of a finite number of independent draws from a type-I extreme value distribution is also distributed as a type-I EV, with the same variance. Aggregating at the city-size level does not impact the estimation of the variance of the draws. I normalize the mean to be zero. If the model is misspecified and in reality, there is a systematic difference in mean idiosyncratic shocks across different city-size bins, this mean value is not separately identified from the log-linear agglomeration externality term a_j, which will capture both in the estimation.

I estimate the model under the assumption, made in section 3, that city sites are all ex ante identical. In reality, sites differ in their natural amenities. This can contaminate the estimation of the model if there is a correlation in the data between these amenities, local firms' productivities and city sizes. Under the assumption that these natural advantages benefit all firms in the same way, this correlation will be captured in estimation by the log-linear agglomeration externality term a_j. This will tend to bias upward the coefficient a_j but importantly does not affect the estimation of the log-supermodular term s_j.[29] Furthermore, the bias is likely to be small. Combes et al. (2008) show that the role played by natural endowments on the productivity of French cities is quite limited. Michaels and Rauch (2016) argue that French cities locations are strongly path dependent. They were efficiently chosen at the time of the Roman Empire, but remained largely unchanged over time. Cities' locations are unlikely to reflect

[28] L_o is a normalization parameter in levels that changes proportionally the size of all cities but does not affect the estimation, which relies on *relative* measures. $\tilde{L}$ is the relevant measure for firm choices. L_o is calibrated to match the actual level of city sizes in the data.

[29] 'The model does not feature consumption amenities either. In principle, the presence of consumption amenities can systematically dampen the relationship between city size and wages. The estimation directly controls for the observed relationship between city size and wages though. This should capture the effect of consumption amenities on city size. Therefore, even though consumption amenities are not directly modelled, they should not biais estimation.

strong exogenous comparative advantage from the perspective of modern technologies.[30] Finally, note that I estimate the model under the maintained assumption that labor is homogeneous. Insofar as there is systematic cross-city variation in labor quality, the estimation will attribute this to variation in firm productivity.

4.3.2. Estimation procedure

The estimation is conducted in two stages. In the first stage, I start by calibrating for each industry its capital intensity α_j and elasticity of substitution σ_j. The capital intensities are calibrated to the share of capital in sectoral Cobb-Douglas production functions, and the elasticity of substitution is calibrated to match the average revenue to cost margin in each sector.[31] I then calibrate the composite parameter $b\frac{1-\eta}{\eta}$, equal to the elasticity of wages to city size in the model. To do so, I follow equation (11) and use the difference between the elasticities of average firm revenue to city size and the elasticities of average firm employment to city size across all sectors. Finally, I calibrate the Cobb-Douglas share of each industry ξ_j by measuring its share of value-added produced.

In the second stage, I use simulated method of moments (SMM) to back out the quadruple $\theta_0^j = (a^j, s^j, \nu_R^j, \nu_Z^j)$ for each sector $j \in (1..S)$. Firms make a discrete choice of (normalized) city size, according to the following equation

$$\log \tilde{L}_j^*(z_i) = \underset{\log \tilde{L} \in \mathcal{L}}{\arg\max} \quad \log(z_i)\big(1 + \log\, \tilde{L}\big)^{s_j} + (a_j - b(1-\alpha_j)\frac{1-\eta}{\eta}) \log \tilde{L} + \epsilon_{i,L}, \tag{21}$$

which is the empirical counterpart of equation (8). Because the choice equation involves unobservable heterogeneity across firms and is non-linear, I have to use a simulation method (Gouriéroux and Monfort (1997)) to recover the model primitives. The SMM method is carried through sector by sector. I retain a rather aggregated definition of sectors, corresponding to 23 industries of the French NAF classification, in order to limit the computing requirements of the procedure. The general approach is close to the one in Eaton et al. (2011). The estimate θ_{II}^j minimizes the loss function

$$\|m_j - \hat{m}_j(\theta)\|_{W_j^2} = \big(m_j - \hat{m}_j(\theta)\big)' W_j \big(m_j - \hat{m}_j(\theta)\big), \tag{22}$$

where m_j is a vector stacking a set of moments constructed using firm data, as detailed below; $\hat{m}_j(\theta)$ is the vector for the corresponding moments constructed from the simulated economy for parameter value θ; W_j is a matrix of weights.[32] Details of the estimation procedure are reported in the online appendix.

[30] Note that for practical reasons I also estimate the model without taking into account existing place-based policies. That is, the estimation takes the city sizes as given, and assumes that firms' choice of location is not systematically distorted by policies. This corresponds to the model of section 3 with city developers, where equilibrium subsidies are independent of city sizes. It is also the case if city sizes are taken as exogenous, without policies. To the extent that - in reality – there are policies that tend in general to favor smaller cities, this assumption could lead to an under-estimation of agglomeration effects in quantification.

[31] In each sector, $\hat{\sigma_j}$ and $\hat{\alpha_j}$ are calibrated using $\frac{\hat{\sigma_j}}{\hat{\sigma_j}-1} = mean(\frac{\text{v.a.}}{\text{costs}})$ where costs exclude the cost of intermediate inputs, and $\hat{\alpha_j} = \alpha_j^{CD} \frac{\sigma_j}{\sigma_j - 1}$ where α_j^{CD} is the sectoral revenue-based Cobb-Douglas share of capital.

[32] The weighting matrix W_j for sector j is a generalized inverse of the estimated variance-covariance matrix Ω_j of the moments calculated from the data m_j.

4.3.3. Identification and choice of moments

I use three sets of non-parametric moments, for each sector, to characterize the economy. The first set of moments describes non-parametrically how average firm value-added increases with city size, sector by sector. I use 4 moments for each sector, capturing average firm size for each quartile of city size. Intuitively, how fast firm size increases with city size helps pin down the agglomeration parameters a and s. The parameters a and s both impact firm productivity and value-added, but a impacts them log-linearly with city size, and s impacts them more than log-linearly because it entails the sorting of more productive (high z) firms into larger cities.

To help identify the parameters that govern firm-level heterogeneity ν_z and ν_R, I also use 5 moments that characterize non-parametrically the firm-size distribution in value added.[33] If the distribution of value added conditional on city size does not allow me to identify separately ν_R from ν_Z, these moments do. The parameter ν_R governs the variance of the noise in a classic discrete choice setting problem (equation (21)). This parameter is usually normalized, as it cannot be inferred from simply observing the choice of the agent. Here, in contrast to a classic discrete choice setting, I observe not only the choice of city size made by firms, but also additional outcomes that are impacted by this choice, for instance, firm value-added. This last part is unusual, and these additional moments allow me to identify the variance of idiosyncratic shocks separately form the variance of firm's raw efficiency.

In addition, I use moments that describe non parametrically the distribution of sectoral value-added across city sizes. I measure the share of value-added in a given sector that falls into one of 4 bins of city sizes.[34] These moments summarize the geographic distribution of economic activity within a sector. Together with the first set of moments, they give information on the density of firms located in different city sizes. They help inform the strength of sorting forces, since they summarize the firms' location choices. They also contribute to identifying the distribution of raw efficiencies, conditional on the agglomeration parameters.

Online Appendix H shows identification of the parameters a, s and ν_z in a simple case where the variance of shocks ν_R is fixed. Formally, the distribution of value added conditional on city size alone is sufficient in that case for identification. Intuitively though, the other set of moments are economically important in estimation.[35]

4.3.4. Model fit

Figures in Appendix J report the model fit for the set of moments targeted by the estimation procedure. Specifically, the way firm average value-added increases with city size is generally well captured by the estimated model (Figure J.1). Further, the estimation relies on five moments of the sectoral firm-size distribution in revenues. To get a sense of the fit of the model fit in this dimension, I show in figure

[33] These bins are defined by the 25, 50, 75 and 90th percentiles of the distribution in the data, normalized by the median. As in Eaton et al. (2011), higher quantiles are emphasized in the procedure, since they capture most of the value added, and the bottom quantiles are noisier.

[34] I order cities in the data by size and create bins using as thresholds cities with less than 25%, 50% and 75% of the overall workforce.

[35] Identification is shown conditional on the specification I retain in assumption B. A proof of identification a for more general specification is beyond the scope of this paper and left for future research.

J.2 how the *whole* firm-size distribution compares in the data and in the model. In general, the fit is better for the upper tail than the lower tail, which is intuitive since the estimation focuses on upper-tail quantiles and the initial distribution of z is truncated to the left. Finally, the estimation relies on the share of sectoral value-added in four given city-size bins. I compute more generally for each sector the share of employment by decile of city size and represent the simulated vs. actual shares on Figure J.3. The model accurately captures the cross-sectoral heterogeneity in location patterns. The within-sector patterns are noisier, but still follow well the overall trends in the data.

Finally, I focus on a moment not directly targeted in the estimation, which is the city-size distribution. The estimation is made on a grid of possible city sizes that have the same maximum to minimum range as in the data. I make, however, no assumption on the number of cities in each size bin, i.e., on the city-size distribution. Armed with sectoral estimations, I can solve for the general equilibrium of the model and in particular compute the city-size distribution that clears labor markets at the estimated parameter values (see Section 3). The estimated city-size distribution exhibits Zipf's law and follows quite well the actual city-size distribution measured here in total local employment of the city, consistent with the data used in estimation. The fit is shown in Figure 2, where the city-size distribution is plotted for the simulated data and the actual data.[36]

4.3.5. Analysis of the parameter estimates

The estimated parameters industry by industry are reported in table 5. The sectoral estimate of s_j, the parameter that governs the strength of the complementarity between firm efficiency and agglomeration externalities, is positive for a vast majority of industries. It is negative for two industries, the shoes and leather industry and metallurgy, which are relatively mature industries.[37] That more mature industries tend to exhibit different agglomeration forces is reminiscent of the argument in Audretsch and Feldman (1996), who argues that the nature of agglomeration forces depend on the life cycle of industries and show that agglomeration forces tend to decline as industries get more mature and less innovative.

Together, the agglomeration parameters and the variance parameters *jointly* determine the distribution of the realized productivity of firms and, crucially, the productivity gains associated with city size in equilibrium. These gains have been used in the literature as a proxy to measure agglomeration externalities. Here, the productivity gains associated with city size depend not only on the strength of agglomeration externalities, but on the sorting of firms, and on their selection on local idiosyncratic productivity shocks. In what follows, I present direct and counterfactual measures of the elasticity of firm productivity to city size to highlight how these forces interplay and understand how the parameter estimates translate into economic forces. For simplicity, I present average measures across sectors.

A first raw measure of the observed elasticity of firm productivity to city size can be computed by running the following simple OLS regression:

$$\log \psi_{i,j} = \beta_0 + \beta_1 \log L_i + \delta_j + \mu_i, \tag{23}$$

[36] The fitted lines correspond to a log rank-log size regression run on each of these distributions. Parallel slopes indicate that both distributions have the same tail. The levels are arbitrary and chosen so that the figure is readable.

[37] In the food manufacturing sector, the best fitting parameters correspond to no variance in firm type, hence no sorting per se. The log-supermodular coefficient that governs sorting is not well defined in this sector.

where $\psi_{i,j}$ is the equilibrium productivity of simulated firm i with efficiency z_i in industry j, $L_i = L_j^*(z_i)$ is the size of the city where firm i has chosen to produce and δ_j is an industry fixed effect. The OLS estimate of β_1, the elasticity of observed firm productivity to city size, is 4.2%. Interestingly, this measure falls within the range of existing measures of agglomeration externalities, as reported in Rosenthal and Strange (2004). They typically range from 3% to 8%. Rosenthal and Strange (2004) note that most studies do not account for sorting or selection effects when estimating the economic gains to density - they are therefore broadly comparable, in scope, to the OLS estimation of β_1 in equation (23).[38] In the estimated model, these productivity gains are driven only in part by the existence of agglomeration externalities. Part of these gains come from the sorting of more efficient firms into larger cities, which I examine now.

To measure the contribution of firm sorting in the observed economic gains to density, I conduct the following counterfactual analysis. I recalibrate the model with firms constrained to choose their city size as if they all had the average efficiency in their sector, thereby shutting down systematic sorting. In this exercise, the difference in firms' location choice is only driven by firm-city size specific iid productivity shocks. I find that the relationship between firm-level productivity and city size is flatter in the counterfactual simulation than in the baseline model. Estimating equation (23) on this counterfactual data leads to an elasticity of firm productivity to city size of 2.3%. By this account, firm sorting accounts for almost half of the productivity gains measured in equilibrium between cities of different sizes.

Finally, to gauge the importance of the sorting forces emphasized in the model I decompose the variance of productivity between the contribution of the systematic component of productivity on which firms sort, and the contribution of the idiosyncratic part. To that end, I regress the log of each of these two components on firm's log productivity, in each sector, as follows:

$$\begin{aligned} \log(z_i)(1 + \log \frac{L}{L_o})^{s_j} &= \beta_{j,systematic} \log(\widetilde{\psi}_i) + \nu_{1,i} \\ \epsilon_{iL} &= \beta_{j,random} \log(\widetilde{\psi}_i) + \nu_{2,i}, \end{aligned}$$

where $\log(\widetilde{\psi}_i) = \log(z_i)(1 + \log \frac{L}{L_o})^{s_j} + \epsilon_{iL}$. Mechanically, this procedure yields coefficients that sum to one and give us a metric for the relative importance of sorting vs. random shocks to shape the distribution of firms' productivities. I find that, on average across sectors, the systematic component explains 51% of the variance of firm productivities. Both dimensions of productivity contribute roughly equally to explaining firm productivity in the estimated model. This points at sorting as an important mechanism to rationalize the data.

5. The Aggregate Impact of Place-Based Policies

Equipped with the estimates of the model's parameters, I finally turn to the evaluation of the general equilibrium impact of a set of place-based policies. I use the real wage, constant across space, as a measure of welfare. Details of the implementation are reported in the online appendix.

[38] An exception is Combes et al. (2008), who estimate agglomeration externalities using detailed French worker-level data and control for the sorting of workers across locations. They find an estimate of 3.7% of the elasticity of productivity to employment density.

5.1. Local tax incentives

I first study policies that subsidize firms locating in less developed cities or regions. This type of federal program aims at reducing spatial disparities and is also advocated for reason of efficiency. The case for increased efficiency relies on the idea that in the presence of agglomeration externalities, jump-starting a local area by attracting more economic activity can locally create more agglomeration externalities, enhancing local TFP and attracting even more firms.[39] This argument, however, needs to be refined. As has been pointed out in the literature (Glaeser and Gottlieb (2008), Kline and Moretti (2014)) this effect depends in particular on the overall shape of agglomeration externalities. While smaller cities may in fact benefit from these policies, larger cities marginally lose some resources – and therefore benefit from less agglomeration economies. The net effect on the overall economy is a priori ambiguous. Turning to spatial disparities, since utility is equalized across all workers, there is no welfare inequality in equilibrium in the model.[40] Nevertheless, the economy is characterized by other spatial disparities, in city size or GDP per capita for example, that may matter for political economy reasons. Place-based policies impact these spatial differences. They tend to benefit the targeted areas, but the extent to which they reduce aggregate measures of spatial differences depends on the overall reallocation of economic activity in space, which I examine in the quantitative exercise below. To evaluate these policies in the context of my model, I consider a set of counterfactuals in which firms are subsidized to locate in the smallest cities of the country, which are also the least productive ones. To calibrate the intensity of the simulated policy, I choose as a reference point an example of a specific policy put in place in France, which targets disadvantaged neighborhoods (rather than smaller cities). It covers an overall population of 1.5M or 2.3% of the French population, the French "ZFU" program (*Zones Franches Urbaines*). I simulate a scheme that subsidizes firms locating in the smallest cities corresponding to 2.3% of the population in the simulated data. I implement a subsidy of 12% of firms' profits in these areas, paid for by a lump-sum tax levied on all firms in the country.[41]

Local effects The model predicts large effects of subsidizing small cities on the targeted areas. In targeted cities, the number of establishments grows by 19%. The corresponding local increase in population is, however, only 4%. This is because the firms attracted by the policy in these areas are small and have low productivity. These results are roughly consistent with the order of magnitude estimated in Mayer et al. (2015) on the effect of the French ZFU; Mayer et al. (2015) find a 31% increase in the entry rate of establishments in the three years following the policy's implementation and note that these new establishments are small relative to existing establishments. Of course, this is just a "plausibility check"

[39] A long line of research following Henderson (1974) has argued on theoretical grounds that cities are too big. Though this argument could justify this type of redistribution policy, recent research has qualified this result (see in Albouy (2012), Albouy et al. (2016) and the theoretical results of section 3).

[40] Since workers are identical and freely mobile, there is no welfare inequality in the model, with or without policies. If, in reality, spatial policies lead to a reduction in well-being inequality, then the efficiency costs I estimate should be traded off against these benefits. The quantitative analysis provides a way to gauge the aggregate efficiency costs of place-based policy.

[41] The French "ZFU" program (*Zones Franches Urbaines*) is a policy similar to to the Empowerment Zone program in the US. This policy costs 500 to 600 million euros in a typical year, corresponding to extensive tax breaks given to local establishments. A 12% subsidy on profits in the model corrresponds to a a cost of 0.04% of GDP, which matches the one reported for France for the ZFU program.

since the two exercises are not directly comparable - the ZFU targets sub-areas smaller than the cities of my model, and the model does not have dynamic effects.

Aggregate effects Beyond evaluating the effects of the policy on the targeted cities, the counterfactual exercise allows me to compute the general equilibrium effect of this type of policy. I compute the aggregate TFP and welfare effects of the policy for different levels of the subsidy, holding constant the targeted areas. The welfare measure I use is the real income of the representative worker. It does not account for other elements that can arguably be in the objective function of the decision maker, such as measures of equity, and which often motivate these policies. In that sense, the (negative) welfare effects reported here can be seen as costs of this policy, to be weighted against potential benefits that are outside of the model.

The simulation shows that these place-based policies have negative long-run effects, both on the productive efficiency of the economy and in terms of welfare. A subsidy to smaller cities that amounts to 1% of GDP leads to a loss of 1.05% in TFP in the aggregate, and a loss of 1.4% in welfare. While such a policy allows to decrease congestion overall, the welfare gain from decreasing congestion is largely dominated by the negative TFP effect.[42]

I then use the counterfactual economy to study the impact of these place-based policies on the dispersion of spatial outcomes, by measuring how the Gini coefficient for the distribution of GDP per capita in the economy reacts to the policy. A reason to focus on such a measure is that policy makers may want to smooth out this type of disparity across cities. Perhaps surprisingly, the type of place-based policies I study leads to an increase in spatial disparities as measured by this Gini index. The intuition for this result is as follows. The counterfactual equilibrium is characterized by (1) growth in the size of smaller cities, (2) a decrease in the population of mid-size cities, and (3) an increase in the population of the largest cities. This change in city size distribution is plotted in figure 3. That larger cities grow in the counterfactual economy comes from the fact that, as mid-size cities lose population in favor of smaller cities, they offer less agglomeration externalities. As a consequence, these mid-size cities become less attractive than larger cities for a set of firms that were previously indifferent between these mid-size cities and *larger* cities. Small and large cities thus expand at the expense of mid-size cities. Quantitatively, this leads to a rise in the Gini coefficient.

According to these results, place-based policies may have general equilibrium effects that run counter to their rationale.

5.2. Land-use regulation

Glaeser and Gottlieb (2008) forcefully argue against policies that limit the growth of cities and constrain the available housing supply. Zoning regulations or regulations on the type or height of buildings that can be built within a city constitute examples of such land-use regulations. A rationale for these restrictions on land-use development is that they may increase the quality of life for existing residents. On the

[42] TFP has a magnified impact on welfare as capital flows in and out of the economy in response to the TFP shock. The corresponding formulas are explicited in the supplemental material (I.2).

other hand, by constraining the housing supply and limiting the size of cities, they may dampen the agglomeration effects at play in the economy.

I model the loosening of local land-use regulation that could be mandated by a federal government by decreasing the land-use intensity parameter b in the housing production function (equation (1)).[43] Decreasing this parameter increases the elasticity of housing supply. To quantify the impact of land-use regulation policies, I first need to calibrate b. To do so, I assume that b is such that the housing supply elasticity is at the median measure across US cities, as estimated by Saiz (2010).[44] I then compare the aggregate TFP and welfare of two counterfactual economies: one where the housing supply elasticity is set at the 25th percentile of the housing supply elasticity distribution in Saiz (2010) and one where it is set at the 75th percentile.

Increasing the housing supply elasticity has two separate effects on welfare. First, a direct – mechanical – effect on utility. All else equal, as the housing sector becomes more productive and the housing supply elasticity increases, the housing units available to households increase, which directly raise their utility. This mechanical effect is not the focus here. Beyond this direct effect, an increase in the housing supply elasticity flattens out the wage schedule (see equation (4)), which leads firms in the heterogeneous goods sectors to locate in larger cities. This indirect effect enhances the productive efficiency of these sectors.[45]

Figure J.5 reports TFP and welfare, relative to the reference equilibrium, for various levels of the housing supply elasticity. An overall increase in the housing supply elasticity from the 25th to the 75th percentile leads to a 1.6% gain in TFP and a 1.8% indirect gain in welfare.

This policy experiment illustrates how increasing housing supply in cities can have positive effects beyond directly reducing congestion costs. They allow for a more efficient spatial organization of production in the differentiated goods sectors by endogenously creating agglomeration externalities and enhancing the way labor is allocated to heterogeneous firms in the economy.

6. Conclusion

I offer a new general equilibrium model of heterogeneous firms that are freely mobile within a country and can choose the size of the city where they produce. I show that the way firms sort across cities of different sizes is relevant to understanding aggregate outcomes. The sorting of firms, mediated by the existence of city developers who act as a coordinating device for the creation of cities, leads to a unique spatial equilibrium of this economy. Therefore, the model can be used to conduct policy analysis. It allows the quantification of the complex spatial equilibrium effects of spatial policies. Using the structure of the model, I estimate the general equilibrium effects of two types of place-based policies. A policy

[43] Land-use regulation have been largely delegated to local municipalities in France over the past 30 years, but the national government can still decide on general rules that apply everywhere (RNU, *Reglement National d'Urbanisme*).

[44] For France, Combes et al. (2016) propose a range of estimates for the (mean) price elasticity of housing with respect to city size. The measure I use is well within this range.

[45] To focus on this indirect effect, I control for the direct effect on utility of an increase in housing supply as follows. For each value of the housing supply elasticity, I simulate the equilibrium of the economy as described above. To measure welfare per capita, I take into account the spatial reallocation of economic activity, but hold constant b, hence the price of housing, in the utility of workers. Fixing the price of housing mutes the mechanical welfare effect coming from an increase in housing supply. The aggregate welfare gain, including both the direct and the indirect effects of an increase in housing supply, is 20.5%

that explicitly targets firms locating in the least productive cities tends to hamper the productivity of the economy as a whole. For the specific policy I study, spending 1% of GDP on local tax relief leads to an aggregate welfare loss of 1.4% and does not reduce observed spatial dispersions that may matter for political economy reasons. On the other hand, policies that encourage the growth of all cities - not just the smallest ones - can enhance equilibrium productivity and welfare: moving the housing-supply elasticity from the 25th to the 75th percentile of housing-supply elasticity leads to a 1.6% gain in TFP and 1.8% in welfare through a spatial reorganization of production.

References

Abdel-Rahman, Hesham and Alex Anas, "Theories of systems of cities," *Handbook of Regional and Urban Economics*, 2004, *4*, 2293–2339.

Albouy, David, "Evaluating the efficiency and equity of federal fiscal equalization," *Journal of Public Economics*, 2012, *96* (9-10), 824–839.

_ **, Kristian Behrens, Frédéric Robert-Nicoud, and Nathan Seegert**, "The optimal distribution of population across cities," Technical Report, National Bureau of Economic Research 2016.

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

Athey, Susan, "Monotone comparative statics under uncertainty," *quarterly Journal of Economics*, 2002, pp. 187–223.

Audretsch, David and Maryann Feldman, "Innovative clusters and the industry life cycle," *Review of Industrial Organization*, 1996, *11* (2), 253–273.

Baldwin, Richard and Toshihiro Okubo, "Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting," *Journal of Economic Geography*, June 2006, *6* (3), 323–346.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

_ **, Giordano Mion, Yasusada Murata, and Jens Suedekum**, "Spatial frictions," *Journal of Urban Economics*, 2017, *97*, 40 – 70.

Brinkman, Jeffrey, Daniele Coen-Pirani, and Holger Sieg, "Firm dynamics in an urban economy," *International Economic Review*, 2015, *56* (4), 1135–1164.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, April 2013, *103* (2), 897–947.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial wage disparities: Sorting matters!," *Journal of Urban Economics*, 2008, *63* (2), 723–742.

_ **,** _ **, and** _ , "The costs of agglomeration: Land prices in French cities," Technical Report 2016.

_ **,** _ **,** _ **, Diego Puga, and Sébastien Roux**, "The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection," *Econometrica*, November 2012, *80* (6), 2543–2594.

Costinot, Arnaud, "An Elementary Theory of Comparative Advantage," *Econometrica*, 07 2009, *77* (4), 1165–1192.

_ **and Jonathan Vogel**, "Matching and Inequality in the World Economy," *Journal of Political Economy*, 08 2010, *118* (4), 747–786.

Criscuolo, Chiara, Ralf Martin, Henry Overman, and John Van Reenen, "The Causal Effects of an Industrial Policy," Technical Report, Centre for Economic Performance, LSE 2012.

Davis, Donald and Jonathan Dingel, "A Spatial Knowledge Economy," NBER Working Papers 18188, National Bureau of Economic Research June 2012.

_ and _ , "The Comparative Advantage of Cities," Working Paper 2013.

Desmet, Klaus and Esteban Rossi-Hansberg, "Urban Accounting and Welfare," *American Economic Review*, 2013, *103* (6), 2296–2327.

Duranton, Gilles and Diego Puga, "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 2001, pp. 1454–1477.

_ and _ , "Micro-foundation of urban agglomeration economies, in Henderson V. J. and Thisse JF.(eds) "Handbook of Regional and Urban Economics Vol. 4 Cities and Geography," 2003.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, "An anatomy of international trade: Evidence from French firms," *Econometrica*, 2011, *79* (5), 1453–1498.

Eeckhout, Jan and Philipp Kircher, "Assortative matching with large firms: Span of control over more versus better workers," Working Paper 2012.

_ , Roberto Pinheiro, and Kurt Schmidheiny, "Spatial sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Gabaix, Xavier and Rustam Ibragimov, "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," *Journal of Business & Economic Statistics*, 2011, *29* (1), 24–39.

Glaeser, Edward and Joshua Gottlieb, "The economics of place-making policies," *Brookings Papers on Economic Activity*, 2008, *2008* (1), 155–239.

Gourieroux, Christian, Alain Monfort, and Eric Renault, "Indirect inference," *Journal of applied econometrics*, 1993, *8* (S1), S85–S118.

Gouriéroux, Christian and Alain Monfort, *Simulation-based Econometric Methods* OUP/CORE Lecture Series, OUP Oxford, 1997.

Helsley, Robert W. and William C. Strange, "Coagglomeration, Clusters, and the Scale and Composition of Cities," *Journal of Political Economy*, 2014, *122* (5), 1064–1093.

Henderson, Vernon, "The Sizes and Types of Cities," *American Economic Review*, September 1974, *64* (4), 640–56.

Kline, Patrick and Enrico Moretti, "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

Matano, Alessia and Paolo Naticchioni, "Wage distribution and the spatial sorting of workers," *Journal of Economic Geography*, 2011, *12* (2), 379.

Mayer, Thierry, Florian Mayneris, and Loriane Py, "The impact of urban enterprise zones on establishment location decisions and labor market outcomes: evidence from france," *Journal of Economic Geography*, 2015, p. lbv035.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725.

Michaels, Guy and Ferdinand Rauch, "Resetting the Urban Network: 117-2012," *The Economic Journal*, 2016.

Milgrom, Paul R, "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 1981, pp. 380–391.

Mion, Giordano and Paolo Naticchioni, "Triage et arrimage des compétences et des entreprises dans l'espace.," *Canadian Journal of Economics/Revue canadienne d'économique*, 2009, *42* (1), 28–55.

Nocke, Volker, "A Gap for Me: Entrepreneurs and Entry," *Journal of the European Economic Association*, 09 2006, *4* (5), 929–956.

Rosenthal, Stuart and William Strange, "Evidence on the nature and sources of agglomeration economies," *Handbook of Regional and Urban Economics*, 2004, *4*, 2119–2171.

Rossi-Hansberg, Esteban and Mark Wright, "Urban structure and growth," *Review of Economic Studies*, 2007b, *74* (2), 597–624.

Saiz, Albert, "The geographic determinants of housing supply," *Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Sampson, Thomas, "Selection into trade and wage inequality," *American Economic Journal: Microeconomics*, 2014, *6* (3), 157–202.

Serrato, JC Suarez and Owen Zidar, "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms," *American Economic Review*, 2016, *106* (9), 2582–2624.

Topkis, D, *Supermodularity and complementarity*, Princeton University Press, 1998.

Table 1: Summary statistics.

	log value added			log employment			N
	mean	p25	p75	mean	p25	p75	
Manufacture of food products and beverages	7.57	5.81	8.43	2.30	1.39	3.00	14,102
Manufacture of textiles	8.06	6.10	9.12	2.80	1.95	3.66	2,955
Manufacture of wearing apparel	7.50	5.36	8.57	2.47	1.61	3.43	3,219
Manufacture of leather goods and footwear, leather tanning	7.90	5.88	8.95	2.79	1.79	3.78	878
Manufacture and products of wood, except furniture	7.77	6.20	8.55	2.43	1.79	3.09	3,688
Manufacture of pulp, paper and paper products	8.66	6.55	9.69	3.15	2.20	4.04	1,284
Publishing, printing and reproduction of recorded media	7.20	4.96	8.26	1.94	1.10	2.71	11,238
Manufacture of chemicals and chemical products	8.77	6.03	10.33	3.11	1.79	4.30	2,647
Manufacture of rubber and plastic products	8.41	6.50	9.40	2.94	2.08	3.81	3,563
Manufacture of glass, ceramic, brick and cement products	7.99	6.12	8.89	2.55	1.61	3.30	3,143
Manufacture of basic metals	9.06	6.90	10.18	3.52	2.30	4.51	834
Manufacture of fabricated metal products, except machinery	8.02	6.50	8.81	2.50	1.79	3.22	16,160
Manufacture of machinery	8.00	6.16	8.96	2.48	1.61	3.33	7,689
Manufacture of office machinery and computers	7.91	5.52	9.04	2.54	1.61	3.40	312
Manufacture of electrical machinery	8.14	6.22	9.09	2.70	1.61	3.56	2,273
Manufacture of radio, television and communication equipment	8.17	5.88	9.25	2.74	1.61	3.69	1,544
Manufacture of medical, precision and optical instruments	7.70	5.89	8.55	2.18	1.39	2.94	4,235
Manufacture of motor vehicles	8.48	6.39	9.42	3.06	1.95	3.83	1,346
Manufacture of other transport equipment	7.86	5.45	8.98	2.60	1.39	3.56	1,007
Manufacture of furniture	7.28	5.24	8.30	2.14	1.10	3.00	5,269
Recycling	7.62	5.89	8.45	2.11	1.39	2.77	1,394
Information technology services	7.24	4.84	8.48	1.95	0.69	2.83	10,617
Business services, non I.T.	6.93	4.79	7.95	1.55	0.69	2.20	57,673

Table 2: Share of establishment in larger cities and tradable capital intensity.

Dep. variable	Share of establishments in large cities			
	I	II	III	IV
Sample		*all tradables*		*export intensive*
Tradable capital intensity	0.479**	0.613**	0.596**	0.551**
	(0.152)	(0.147)	(0.168)	(0.200)
High skill intensity=1			0.058**	0.042
			(0.029)	(0.033)
Nb firms	no	yes	yes	yes
Mean va	no	yes	yes	yes
R-squared	0.065	0.219	0.174	0.139
Observations	146	146	117	84

(*)$p < 0.10$, (**)$p < 0.05$. Tradable capital intensity: share of capital net of real estate assets in a Cobb-Douglas production function with labor, tradable capital and non tradable capital. Large cities: larger cities representing 50% of workers. Nb firms: number of firms. Mean va: average value added per firm. High skill intensity are sectors above median of skill intensity. Export intensive: industry above median for all sectors in the economy in export intensity, proxied by the ratio of export to domestic sales.

Table 3: Tail of the firm-size distribution (ζ) vs sector location.

	ζ, tail of firm-size distribution		
	I	II	III
Sample		*all tradables*	*export intensive*
Share in large cities	-0.544**	-0.686**	-0.263**
	(0.142)	(0.122)	(0.115)
Nb firms	no	yes	yes
Mean va	no	yes	yes
R-squared	0.092	0.578	0.487
Observations	146	146	89

(*)$p < 0.10$, (**)$p < 0.05$. Pareto Shape: ζ estimated by $Log(Rank_i - 1/2) = a - \zeta Log(va_i) + \epsilon_i$, on firms above median size, for industries with more than 200 firms. Nb firms: number of firms. Mean va: average value added per firm. Export intensive: industry above median for all sectors in the economy in export intensity, proxied by the ratio of export to domestic sales. Heteroskedasticity-robust standard errors.

Table 4: Movers.

	Δ_t City Size			
	I	II	III	IV
Sample	*all tradables*		*export intensive*	
log(firm size)	0.089**	0.073**	0.090**	0.080**
	(0.020)	(0.019)	(0.026)	(0.025)
Initial City Size	-0.987**		-0.986**	
	(0.055)		(0.048)	
Constant	12.296**			
	(0.651)			
Industry F.E.	yes	yes	yes	yes
Initial city F.E.		yes		yes
R-squared	0.537	0.629	0.540	0.635
Observations	6103	6103	3675	3675

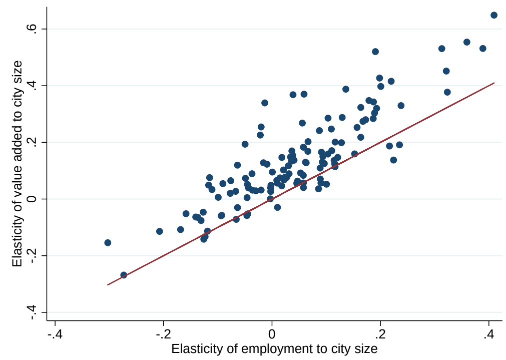
(*)$p < 0.10$, (**)$p < 0.05$. Set of mono-establishment firms which move between 2 years, between 1999 and 2005. Δ_t City Size= $log(\frac{L_{t+1}}{L_t})$, where L_t is the size of the city where the firm locates at time t. Size is measured by the firm value added relative to other firms in the same sector-year-city, as the residual of $log(VA)_i = DS_i + DT_i + DC_i + \epsilon_i$ where DS is a sector fixed effect, DT a year fixed effect, DC a city fixed effect. Export intensive: industry above median for all sectors in the economy in export intensity, proxied by the ratio of export to domestic sales.

Table 5: Estimated parameters.

	$\hat{s}$	$\hat{\nu_R}$	$\hat{\nu_z}$	$\hat{a}$
Manufacture of food products and beverages	n.d	0.476	0.000	0.142
	n.d.	(0.016)	(0.089)	(0.058)
Manufacture of textiles	0.038	0.294	0.274	0.009
	(0.788)	(0.036)	(0.297)	(0.080)
Manufacture of wearing apparel	0.147	0.219	0.252	0.040
	(0.233)	(0.020)	(0.138)	(0.034)
Manufacture of leather goods and footwear, leather tanning	-0.102	0.162	0.465	0.033
	(0.068)	(0.037)	(0.027)	(0.006)
Manufacture and products of wood, except furniture	0.043	0.176	0.397	-0.020
	(0.021)	(0.009)	(0.021)	(0.004)
Manufacture of pulp, paper and paper products	0.049	0.243	0.589	0.019
	(0.014)	(0.005)	(0.013)	(0.008)
Publishing, printing and reproduction of recorded media	0.210	0.408	0.407	0.171
	(0.639)	(0.403)	(0.958)	(0.220)
Manufacture of chemicals and chemical products	0.217	0.430	0.977	-0.001
	(0.305)	(0.205)	(0.694)	(0.029)
Manufacture of rubber and plastic products	0.001	0.137	0.738	0.021
	(0.003)	(0.006)	(0.005)	(0.001)
Manufacture of glass, ceramic, brick and cement products	0.056	0.172	0.741	-0.019
	(0.010)	(0.016)	(0.021)	(0.005)
Manufacture of basic metals	-0.037	0.172	0.790	0.027
	(0.007)	(0.005)	(0.009)	(0.004)
Manufacture of fabricated metal products, except machinery	0.065	0.178	0.317	0.027
	(0.017)	(0.006)	(0.011)	(0.003)
Manufacture of machinery	0.070	0.137	0.496	0.024
	(0.006)	(0.006)	(0.012)	(0.002)
Manufacture of computers and office machinery	-0.009	0.176	0.529	0.123
	(0.076)	(0.022)	(0.088)	(0.022)
Manufacture of electrical machinery	0.033	0.071	0.552	0.034
	(0.005)	(0.004)	(0.006)	(0.001)
Manufacture of radio, television and communication equipment	0.060	0.191	0.536	0.051
	(0.058)	(0.143)	(0.235)	(0.024)
Manufacture of medical, precision and optical instruments	0.138	0.196	0.432	0.044
	(0.009)	(0.010)	(0.013)	(0.004)
Manufacture of motor vehicles	0.743	0.281	0.147	-0.062
	(0.019)	(0.008)	(0.010)	(0.031)
Manufacture of other transport equipment	0.045	0.201	0.536	0.103
	(0.021)	(0.011)	(0.023)	(0.015)
Manufacture of furniture	0.553	0.340	0.091	0.013
	(0.244)	(0.015)	(0.028)	(0.065)
Recycling	0.178	0.482	0.567	0.052
	(0.039)	(0.015)	(0.042)	(0.016)
Information technology services	0.426	0.280	0.301	0.152
	(0.086)	(0.020)	(0.068)	(0.053)
Business services, non I.T.	0.058	0.300	0.548	0.199
	(0.021)	(0.037)	(0.063)	(0.017)

Note: s is the log-supermodular agglomeration coefficient, a the log-linear agglomeration coefficient, ν_R the variance of iid shocks, ν_z the variance of firms raw efficiency. The log-supermodular coefficient s_1 is not defined for the first sector, as the estimation backs out a degenerate distribution for firms productivity in that sector.

Figure 1: Elasticity of mean value added and employment with city size.



Note: This figure plots for β in the regression: $\log \text{mean va}(L_i) = \alpha + \beta \log L_i + \epsilon_i$ against β in the regression β: $\log \text{mean empl}(L_i) = \alpha + \beta \log L_i + \epsilon_i$, ran sector by sector at the NAF600 level for industries with more than 200 mono-establishment firms.

Figure 2: City size distribution, model and data.

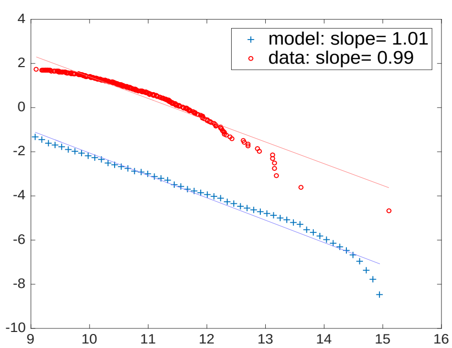
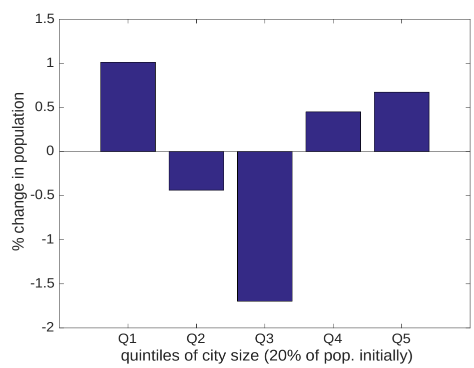


Figure 3: Change in city size distribution, after policy.



Note: This figure plots the change in city size distribution after implementing the policy described in section 4.1. Smallest cities corresponding to 2.3% of pop. are subsidized ; the subsidy amounts to 1% of GDP.