

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Economics

2014

Essays on Firms and Cities

Gaubert, Cecile

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp015m60qv136>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

ESSAYS ON FIRMS AND CITIES

CECILE GAUBERT

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF ECONOMICS

ADVISER: PROFESSOR ESTEBAN ROSSI-HANSBERG

SEPTEMBER 2014

© Copyright by Cecile Gaubert, 2014.

All rights reserved.

Abstract

This collection of essays investigates the interplay between the localization of a firm and its behavior. In the first chapter, I study how city size shapes firm-level volatility. I develop a dynamic extension of the classical Salop circle framework, in which firm-level volatility is driven by the intensity of local competition and by the level of specialization of firms. Both vary endogenously with market size. The model predicts that firms located in larger markets will experience higher firm-level volatility through these two channels. Using French firm-level data, I document that retailers located in larger cities indeed exhibit higher idiosyncratic volatility.

The second chapter turns to studying how city size shapes the entry and exit patterns of firms. I develop an analytically tractable model of industry dynamics at the level of a city. Larger cities endogenously foster more competition, which leads to a tougher selection of firms. On the other hand, the higher volatility of firms productivity in larger cities increases the option value of staying active. This fosters the entry of less efficient firms in larger cities and goes against the competition effect. The model proposes a coherent explanation for the two following stylized facts, documented here on a single dataset of French firms. First, minimum productivity thresholds are not related to city size. Second, turnover rates are higher in larger cities.

In the third chapter, I study a spatial equilibrium model in which firms are mobile and can choose the size of the city where to locate. Firms are heterogeneous in productivity, produce in different sectors and can sort into cities of different sizes. The distribution of city sizes and the sorting patterns of firms are uniquely determined in equilibrium. I structurally estimate the model, using French firm-level data. I use the estimated model to quantify the general equilibrium effects of place-based policies, designed to attract firms to particular regions. 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.

Acknowledgements

I am deeply indebted to Esteban Rossi-Hansberg and Steve Redding for their continuous support, guidance and encouragement. I am also grateful to Gene Grossman, Oleg Itskhoki, Jan De Loecker and Eduardo Morales for invaluable discussions and feedbacks. This work benefited greatly from the stimulating environment provided by the International Economics Section at Princeton University. I thank the IES visitors and seminar participants at the Student Trade Workshop and the International Trade Seminar at Princeton. The International Economics Section provided financial support.

Finally, I can't thank my parents enough for their constant encouragement and support throughout my Ph.D., and I thank David, Ettie and Theo: we did it.

Contents

Abstract	iii
Acknowledgements	v
List of Tables	ix
List of Figures	x
1 A note on firm-level volatility and market structure	1
1.1 Introduction	1
1.2 Market structure and volatility	5
1.2.1 Model Setup	5
1.2.2 Solving the model	9
1.2.3 Firm volatility and market size	14
1.3 Empirical analysis	20
1.3.1 Dataset	20
1.3.2 Main results	21
1.4 Conclusion	24
1.5 Tables and figures	26
1.6 Appendix	29
2 Market Size and Industry Dynamics	34
2.1 Introduction	34
2.2 Stylized facts	38

2.2.1	Data	38
2.2.2	Comparison of small and large markets	39
2.3	A simple dynamic model	45
2.3.1	Model setup	45
2.3.2	Steady-state Industry Equilibrium	47
2.3.3	Comparative statics	50
2.4	A general model of industry dynamics	54
2.4.1	General set-up	55
2.4.2	Value function	56
2.4.3	Numerical simulations	59
2.5	Conclusion	63
2.6	Tables and figures	64
2.7	Appendix	72
2.7.1	Proofs for section 2.3	72
2.7.2	Proofs for section 2.4	74
3	Firm Sorting and Agglomeration	75
3.1	Introduction	75
3.2	A Model of the Location Choice of Heterogeneous Firms	80
3.2.1	Set-up and agents' problem	81
3.2.2	Spatial equilibrium	90
3.2.3	Characterization	98
3.3	Estimation of the Model	106
3.3.1	Data	106
3.3.2	Descriptive evidence on sorting	107
3.3.3	Structural estimation	111
3.4	The Aggregate Impact of Place-Based Policies	125
3.4.1	Local tax incentives	125

3.4.2	Land-use regulation	129
3.5	Conclusion	131
3.6	Tables and figures	133
3.7	Appendix	145
3.7.1	Housing market	145
3.7.2	Extensions of the model	145
3.7.3	Proofs	146
3.7.4	General Equilibrium, TFP and welfare	153
	Bibliography	156

List of Tables

1.1	Summary statistics - retail establishments	26
1.2	Value-added volatility and market size	27
1.3	Placebo	28
2.1	Summary statistics	64
2.2	Summary statistics, by size of market	65
2.3	Comparing TFP distribution between large and small markets	65
2.4	Survival rates and market size	66
2.5	AR(1) fit for log-TFP	66
2.6	TFP volatility and market size - OLS measure	67
2.7	TFP volatility and market size - LP measure	67
2.8	TFP volatility and market size - robustness checks	68
2.9	Share of firms S that exit the market after a change in the threshold.	69
3.1	Summary statistics.	133
3.2	Share of establishment in larger cities and tradable capital intensity.	134
3.3	Tail of the firm-size distribution (ζ) vs sector location.	134
3.4	Movers.	135
3.5	Estimated parameters.	136
3.6	Elasticity of firm productivity to city size: decompositions.	137

List of Figures

2.1	Value functions, market size and volatility.	70
2.2	Exit threshold as a function of market size	70
2.3	Approximated distribution of firms' TFP.	71
3.1	Elasticity of mean value added and employment with city size.	138
3.2	Subset of targeted moments, data and model.	139
3.3	Sectoral distribution of firms revenues, model (blue) and data (red). .	140
3.4	Employment share by decile of city size, model (blue) and data (red). .	141
3.5	Dilation in firm-size distribution in revenues, large cities vs small cities.	142
3.6	City size distribution, model and data.	142
3.7	Elasticity of productivity with respect to city size: decompositions. .	143
3.8	Aggregate impact of local subsidies.	144
3.9	TFP and indirect welfare effects of increasing housing-supply elasticity.	144

Chapter 1

A note on firm-level volatility and market structure

1.1 Introduction

Firm-level fluctuations are an important determinant of firms' hiring and investment decisions. More generally, they shape the patterns of firms' entry, exit and growth, and in turn economy-wide growth. They also contribute to macro-level uncertainty, following the granular hypothesis put forward by Gabaix (2011). Despite the key importance of firm-level fluctuations, and empirical evidence that firm-level volatility varies greatly over time and across firms¹, surprisingly little is known on the determinants of firm-level volatility.

In this paper, I investigate the link between market structure, endogenous product differentiation and firm-level volatility. The empirical literature tends to show support for the intuitive hypothesis that more competition, over time or across sectors, is correlated with higher firm-level volatility.² I investigate this link formally. Tradi-

¹See Comin and Philippon (2006) and Castro, Clementi, and Lee (2011) for example.

²See Comin and Philippon (2006), Gaspar and Massa (2006) and Irvine and Pontiff (2009) for example.

tional models of large group competition have established how market structure can impact firm-level outcomes, such as their level markups, sales or profits. I extend this analysis dynamically to cover second moments of firm-level outcomes. I investigate the link between market structure and firm-level volatility of sales, prices and output. I show how observable industry and market characteristics such as market size and fixed costs shape real firm-level volatility in an imperfectly competitive market.

The model is a dynamic extension of the Salop (1979) circle framework. As in the classic framework, firms are ex-ante homogeneous and choose their location around a circle. I assume in addition that they are hit every period by idiosyncratic cost and demand shocks. Lower entry costs or larger market size leads to tougher competition. In turn, this shapes firm-level volatility, through two channels. First, in larger markets, or markets with lower entry costs, firms are pushed to endogenously serve narrower segments of demand. By a diversification argument, they are less able to smooth out demand shocks across a wide customer base. This is the ‘niche effect’: in more competitive markets, firms tend to serve niches whose demand is more volatile. Second, competition forces affect markups and pass-throughs. They change the way idiosyncratic cost shocks are passed-through to prices, and impact demand and sales. This is the ‘competition effect’: reduced market power makes firm’s sales more sensitive to idiosyncratic cost shocks.

To confront the predictions of the model to the data, I compare the volatility of sales at the firm level for firms serving the same industry,³ but in markets of different sizes and therefore arguably different degrees of competition forces. I use an extensive panel dataset of geo-localized French retailers, and use the size of the city in which they operate as a proxy for market size. By doing so, I follow a classic approach in the literature: I use measures of outcomes across cities of different sizes,

³The model does not feature the use of intermediate inputs. To confront the model with the data, the measure I use in the data is value-added.

for firm serving local demand, as a window to study how market structure impacts firm-level outcomes.⁴ The panel dimension of the data allows me to compute the time series volatility of firm’s sales, as opposed to the cross sectional variance of sales growth that is often used in empirical studies as a proxy for volatility. I find broad support for the prediction of the theory: in the data, the firm-level volatility of sales is systematically positively correlated with market size, after controlling for a range of observable composition effects.

This paper therefore documents the novel and perhaps surprising fact that firms that operate in larger cities experience more volatile sales. Large cities do not smooth out shocks experienced by firms, but instead tend to exacerbate them. The theoretical mechanism I emphasize highlights that two channels can rationalize this empirical regularity: an increased competitive pressure in larger cities, and endogenous product differentiation that leads firms to be more specialized in larger markets .

This paper complements existing research that aims at uncovering the determinants of cross sectional variation in firm-level real volatility. Empirically, Castro, Clementi, and Lee (2011) document the wide variation of firm-level volatility across sectors and argue that sectors in which firms exhibit high volatility correspond to those where creative destruction is likely to be greater. The pervasive rise of firm-level volatility in the past decade seems linked, at least in part, to an increase in competitive pressure (Comin and Philippon (2006)). Gaspar and Massa (2006) and Irvine and Pontiff (2009) find suggestive evidence that higher idiosyncratic financial volatility is linked with increased competition in the goods market. I build on these empirical observations and develop a theoretical model in which more competitive markets lead to higher idiosyncratic volatility at the firm level. Theoretically, Bachmann and Moscarini (2011) study a model of the determinants of firm-level volatility where firms price-experiment more in bad times, when there is less to loose, to explain

⁴See for example Campbell and Hopenhayn (2005), Asplund and Nocke (2006) or Syverson (2004).

the fact that firms exhibit higher volatility during recessions. D’Erasmus and Boedo (2011) study a model where more productive firms invest in intangibles to widen their market span, which leads to lower volatility by diversification. They find support for this mechanism in the data.

This research also enriches the industrial organization literature that studies how competition and market size shapes firms’ outcomes. They typically studies first moments, whereas I extend the analysis to dynamic moments, namely the volatility of firm sales, prices and output. I build on existing models of firm competition with endogenous product differentiation (Salop (1979) and Vogel (2008)) that have been studied in static settings. The model is closely related to Vogel (2008) who studies the endogenous spatial location of heterogeneous firms in a Salop circle framework. I use similar tools to study the dynamics of firm sales, in a setting where firms locate symmetrically on the circle as they are homogeneous *ex ante*. The source of firm heterogeneity only arises *ex post*, in the pricing game.

The empirical strategy I retain is close to Syverson (2004) who studies how market size impacts the distribution of firms’ productivities, and Campbell and Hopenhayn (2005) who show that establishments tend to be larger, with smaller markups, in larger cities. Close to this paper is Asplund and Nocke (2006) who show theoretically and empirically that there is more entry and exit of firms in larger markets. Here, I focus on the related question of the volatility of surviving firms.

Finally, this paper also contributes to the strand of the urban literature that studies how firm outcomes differ across cities of different sizes (see for example, on top of the papers cited above, Combes, Duranton, Gobillon, Puga, and Roux (2012) that studies the cross-sectional distribution of firm productivities across cities of different sizes). Here, I show the perhaps counterintuitive fact that firms, in non-tradable sectors, tend to be more volatile in larger cities.

The paper proceeds as follows. Section 1.2 presents the model based on the classic Salop framework. Section 1.3 confronts the theory with the data. Section 1.4 concludes.

1.2 Market structure and volatility

I extend the Salop (1979) framework to a simple dynamic setting. At each period, firms are hit by idiosyncratic cost shocks, and demand density also varies randomly. These small shocks perturb the symmetric equilibrium. The idiosyncratic volatility of each firm is shaped by how market structure and competition forces transform primitive exogenous shocks into firm-level sales, price and output fluctuations.

1.2.1 Model Setup

Time is discrete, and periods are indexed by t . Consumers are located along a circle of unit length and local density dM_l^t , where l denotes an address on the circle and t indexes the period. Each consumer buys one unit of a differentiated good, provided by a firm located on the circle. If the distance between a given consumer and firm i is h_i , consumer's indirect utility is:

$$V = u - p_i - sh_i, \tag{1.1}$$

where u is a reference level of utility, p_i is firm i 's price and s denotes a transport cost borne by the consumer. The more distant a firm is from a consumer, the higher is the cost borne by the consumer to shop at this outlet.

In the classical Salop circle framework, the local density of demand is constant both around the circle and over time, and equal to a deterministic value $\bar{M}dl$ where $\bar{M}$ is the size of the market. Here, in contrast, I assume that the local density

varies, both around the circle and over time, around this deterministic value. Each period, local density is perturbed by a set of transitory shocks that are address-specific.⁵ As in the classic Salop framework, firms in this model serve segments of demand on the circle (their ‘market span’). Here, this demand is stochastic. The demand shock that hits a firm is the sum of all the idiosyncratic demand shocks that hit the locations that this firm serves. I assume that the shocks that hit each location on the circle are independent - capturing the notion of neighborhood-specific idiosyncratic demand shocks - and identically distributed, since all neighborhoods are ex-ante identical. At each period, they have a different period-specific mean, which stands for an aggregate shock that hits all consumers in the period. Since each location is infinitesimal on the circle, I formalize these assumptions using the apparatus of continuous stochastic processes, where the process diffuses over space (instead of time, as is common): the process is indexed by an address on the circle. Consistent with the assumptions stated above, the local (infinitesimal) shocks at a given period must aggregate up along the circle to a process that has independent and stationary increments.⁶ The only continuous real-valued process with stationary independent increments is the brownian motion. Therefore, I model the stochastic process that governs the distribution of demand density along the circle, for any given period t , as a brownian motion. Formally, for a given period t , I assume that the density of demand at address l is a stochastic process governed by

$$dM_l^t = \bar{M}(dl + dX_l^t). \quad (1.2)$$

⁵Local demand is the product of consumer density and the number of units each consumer buys, usually normalized to one. The random evolution of demand can be thought of as coming from the number of local consumers being random, as I do here, or could alternatively be thought of as the number of units demanded by each consumer being random. The assumption that shocks are address-specific then captures the notion that larger markets are denser in each of the consumer types that they host. In this interpretation, all consumers of a given type are hit by the same income shock.

⁶The increments here are the sum over a given market span (m, l) of infinitesimal shocks. Since all neighborhoods along the circle are ex-ante identical, the distribution of the increments has to be stationary i.e. shouldn’t depend on the location l but only on the length of the market span $l - m$.

The deterministic market size around which demand density fluctuates is $\bar{M}$. The stochastic process X_l is a brownian motion with drift parameter γ^t and constant volatility σ , defined by:

$$dX_l^t = \gamma^t dl + \sigma dW_l. \quad (1.3)$$

Through the period-specific parameter γ^t , all locations l see their density grow at a common period-specific growth rate $(1+\gamma^t)$, which captures market-wide transitory growth shocks. Additionally, each location is hit by an (infinitesimal) normal shock σdW_l . The shock dW_l is a Wiener increment, so that for all $l > m$: $\int_m^l dW_z \sim \mathcal{N}(0, l-m)$. The demand of a firm that has a given market span (m, l) on the circle is determined by the realization of $\frac{\int_m^l dM_z^t}{M}$, which is distributed normal with mean $1 + \gamma^t$ and variance $\sigma_d^2(l-m)$. In other words, the density of demand grows in period t at rate that is distributed normal with a period-specific mean γ^t .

There is free entry to the market and an infinite supply of potential entrants that are ex ante identical. In the entry stage, firms pay a sunk cost F to enter the market, locate around the circle and set up an outlet that will produce a differentiated good with unit transport cost s . In the production stage, firms produce a differentiated good and sell it every period. At each period, the marginal cost of firm k is the realization of a random variable $\tilde{c}_{kt} = c_o(1 + \tilde{\eta}_{kt})$, where $\tilde{\eta}_{kt}$ is the realization of a random variable η which is i.i.d. across firms and time with $E(\eta) = 0$ and $var(\eta) = \sigma_c^2$. I assume that firm location is a long term strategic choice that cannot be modified at the same pace as prices. It is set once and for all at the entry stage. Prices, on the other hand, do adjust every period to the realized transitory shocks. Consequently, the market span of each firm also changes at each period even if its location does not.

Each period, the timing of the pricing game is as follows. At the end of the period, customers are hit by demand shocks. I assume that firms observe demand shocks

aggregated at the level of their current market span on the circle. The assumption that firms do not observe the whole set of realized local demand shocks is made for technical purposes⁷, but reflects the reasonable intuition that firms do not invest in costly detailed marketing research to find out about localized demand shocks that are transitory in nature. Firms know that shocks that hit a given market span are distributed normal – i.e., they know that $\int_m^l dX_z^t \sim \mathcal{N}(\gamma^t, \sigma_d^2(l - m))$ –, but they do not know the mean γ^t of the shocks, which varies over time. They infer it from the realization of the shocks that hit their current market span. Firms are then hit by idiosyncratic cost shocks. They set prices for the new period, having formed expectations about demand and observed their cost shocks. Firms’ prices at every period constitute a Nash equilibrium of the simultaneous move game where each firm chooses its price to maximize profits given its cost level, its expected demand and taking into account other firms’ prices.

As is classical in studies relying on the Salop circle framework, I focus on the case of a symmetric equilibrium in the location stage: n ex ante identical firms locate at evenly spaced locations distant by $d = \frac{1}{n}$ along the circle. The number of firms that enter the market is endogenous and pinned down by the free entry condition. Furthermore, I assume throughout that in the production game, the transitory shocks are small and only impact firms at the margin. They are sufficiently small that no firms is fully undercut, i.e. all firms retain some customers at every period. I will show by constructing it that the Nash equilibrium of the pricing game exists and is unique under these assumptions (in particular, conditional on firms’ locations being symmetric).

⁷The realization of a Brownian motion is nowhere differentiable, hence the optimal pricing strategy of firms given the realization of all local demand shocks cannot be derived from first order conditions.

1.2.2 Solving the model

I solve for the equilibrium backwards and first analyze the pricing game that takes place in the production stage at each period. I assume that there are n firms on the market, distant by $d = \frac{1}{n}$.

Let x_k be the location of firm k , and p_k its price. Let x_j and p_j denote the location and price of the firm's nearest right-hand side competitor, and $x_{j'}$ and $p_{j'}$ denote the location and price of the firm's nearest left-hand side competitor. The demand faced by firm k corresponds to the segment delimited by the consumers who are indifferent between this firm and its two neighboring competitors. I denote by h_j (resp. $h_{j'}$) the distance between firm k and the indifferent consumer located to its right-hand-side (resp. left-hand-side). For the consumer indifferent between firm k and firm i where $i = j, j'$, utility equalization implies that

$$u - p_k - sh_i = u - p_i - s(d - h_i). \quad (1.4)$$

Given these indifference curves, the distances between firm k and its furthest customers are given by

$$h_i(p_k^t, p_i^t) = \frac{p_i^t - p_k^t}{2s} + \frac{1}{2n}, \quad \text{for } i = j, j'. \quad (1.5)$$

All consumers located closer to firm k strictly maximize their utility by shopping at firm k , therefore the market span of firm k is $(x_k - h_{j'}, x_k + h_j)$. The demand faced by firm k in period t is thus the local demand density $M_l^t = \bar{M}(dl + dX_l^t)$ integrated over all addresses l that fall within the firm's market span, that is

$$D(p_k, p_j, p_{j'}) = \bar{M} \int_{x_k - h_{j'}(p_k^t, p_{j'}^t)}^{x_k + h_j(p_k^t, p_j^t)} (dl + dX_l^t).$$

Demand perceived by the firm At the end of the current period t , the economy is hit by the transitory demand shocks that will prevail in the next period. Firms do

not have full knowledge of the realization of these shocks, but only observe demand shocks aggregated at the level of their current market span on the circle. At the end of a period, firm k perceives the following perturbed demand:

$$\bar{M} \int_{x_k - h_{j'}(p_k^t, p_{j'}^t)}^{x_k + h_j(p_k^t, p_j^t)} (dl + dX_l^{t+1}) = \bar{M} H_k^t (1 + \gamma^{t+1} + \sigma^2 \bar{W}_k^{t+1}),$$

where $H_k^t = h_{j'}(p_k^t, p_{j'}^t) + h_j(p_k^t, p_j^t)$ is the firm market span in period t and

$$\bar{W}_k^{t+1} = \frac{\int_{x_k - h_{j'}(p_k^t, p_{j'}^t)}^{x_k + h_j(p_k^t, p_j^t)} dW_l^{t+1}}{H_k^t}$$

is the average realization of the mean zero Wiener increments, averaged over this market span. Firms know that these shocks are distributed normal with volatility σ_d but ignore their mean. Their best predictor of γ^{t+1} is therefore⁸

$$\hat{\gamma}^{t+1} = \gamma^{t+1} + \sigma^2 \bar{W}_k^{t+1}. \quad (1.6)$$

Firms forecast that demand shocks for period $t+1$ are distributed as $\mathcal{N}(\hat{\gamma}^{t+1}, \sigma_d^2(l - m))$. Firm k perceives the following demand for the coming period:

$$\begin{aligned} E_{k,t}[D^{t+1}(p_k, p_j, p_{j'})] &= E_{k,t}[\bar{M} \int_{x_k - h_{j'}(p_k, p_{j'})}^{x_k + h_j(p_k, p_j)} (dl + dX_l^{t+1})] \\ &= \bar{M}(1 + \hat{\gamma}^{t+1}) [h_{j'}(p_k, p_{j'}) + h_j(p_k, p_j)]. \end{aligned}$$

Every firm perceives a uniform demand, albeit with a different value of the (constant) demand density.

⁸ $\hat{\gamma}^{t+1}$ is the maximum likelihood estimator of γ^{t+1} .

Pricing Knowing the realization of its cost shock, firm k 's expected profits for the coming period is

$$\begin{aligned} E_{k,t}[\pi^{t+1}(p_k, p_j, p_{j'})] &= E_k[D^{t+1}(p_k, p_j, p_{j'})(p - c_k^{t+1})] \\ &= \bar{M}(1 + \hat{\gamma}^{t+1}) [h_{j'}(p_k, p_{j'}) + h_{j'}(p_k, p_{j'})] (p_k - c_k^{t+1}). \end{aligned} \quad (1.7)$$

There is no forward-looking pricing decision as the current choice of prices does not impact future profits. Viewed from period t , the firm's expected profits for period $t + 2$ is indeed:

$$E_{k,t}[\pi^{t+2}(p_k, p_j, p_{j'})] = \bar{M}(1 + E_{k,t}(\hat{\gamma}^{t+2}) [h_{j'}(p_k, p_{j'}) + h_{j'}(p_k, p_{j'})] (p_k - E_{k,t}(c_k^{t+2})),$$

since cost and demand shocks are independent. Given that $E_{k,t}(c_k^{t+2}) = E_{k,t}(c_k^{t+2}) = c_0$ and $E_{k,t}(\hat{\gamma}^{t+2}) = E(\gamma^{t+2})$ are not functions of p_{t+1} , it follows that $E_{k,t}[\pi^{t+2}(p_k, p_j, p_{j'})]$ does not depend on p^{t+1} either.⁹

Firm k sets its price for period $t + 1$ to maximize (1.7). Given the expression for the firm's market span (1.5), the corresponding first order condition is simply

$$p - c_k^{t+1} = \frac{p_j + p_{j'} - 2p}{2} + \frac{s}{n}. \quad (1.8)$$

Demand shocks do not enter the firm's first order condition as they are a multiplier of profits. Each firm's equilibrium price depends on its direct neighbors' prices, and on its own marginal cost. The solution to the overall system can be written in matrix notation as follows:

$$\mathbf{p}^{t+1} = \mathbf{A}(\mathbf{n})^{-1} \mathbf{c}^{t+1} + \frac{\mathbf{s}}{\mathbf{n}}, \quad (1.9)$$

⁹The only intertemporal linkage is through the variance of the distribution of the forecasted growth that does not impact profits. The distribution of $\hat{\gamma}$ in period $t + n$ is $\hat{\gamma}^{t+n} \sim \mathcal{N}(E(\gamma^{t+n}), f(p^{t+n-1}))$ where only the variance $f(\cdot)$ depends on the market span of the firm, hence on its price in the previous period $t + n - 1$.

where $\mathbf{p}^t$ is the $1 \times n$ vector of firm prices, $\mathbf{c}^t$ is the $1 \times n$ vector of firm's marginal costs, and $\mathbf{A}(n)$ is a circulant and symmetric matrix of size n :

$$\mathbf{A}(\mathbf{n}) = \begin{pmatrix} 2 & -0.5 & 0 & \cdots & 0 & -0.5 \\ -0.5 & 2 & -0.5 & \ddots & & \\ 0 & -0.5 & 2 & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & & & -0.5 & 2 & -0.5 \\ -0.5 & 0 & \cdots & 0 & -0.5 & 2 \end{pmatrix}.$$

Note that the equilibrium pricing function (1.9) shows that firm prices increase with some average cost of firms on the market, plus a constant additive markup $\frac{s}{n}$ that reflects the elasticity of demand. In markets with more firms (larger n), markups are smaller as firms face more competition. This is the classic result of the Salop circle model. Furthermore, at the symmetric equilibrium where firms are not hit by shocks, firms' prices are $p_0 = c_0 + \frac{s}{n}$ for all firms, they serve a segment of demand of width $h_0 = \frac{1}{n}$ and produce output $Q_0 = \frac{\bar{M}}{n}$.

I now turn to characterizing the elements of the matrix $\mathbf{A}(\mathbf{n})^{-1}$.

Lemma 1

(i) The matrix $\mathbf{A}(\mathbf{n})$ is invertible. Its inverse is of the form

$$\mathbf{A}(\mathbf{n})^{-1} = \text{circ}(b_0(n), b_1(n), \dots, b_{n-1}(n))$$

, i.e.

$$\mathbf{A}(\mathbf{n})^{-1} = \begin{pmatrix} b_0(n) & b_1(n) & b_2(n) & \cdots & b_{n-1}(n) \\ b_{n-1}(n) & b_0(n) & b_1(n) & \ddots & \\ & \ddots & \ddots & \ddots & \\ b_1(n) & & \cdots & & b_0(n) \end{pmatrix},$$

with $b_k(n) = b_{n-k}(n)$ for all $k = 1, \dots, n-1$. The matrix is fully determined by the $1 + \lfloor \frac{n}{2} \rfloor$ coefficients $(b_0(n), b_1(n), \dots, b_{\lfloor \frac{n}{2} \rfloor}(n))$.

(ii) All coefficients $b_i(n)$ are positive, and the coefficients of each line (or column) sum to 1.

(iii) Fixing n , the sequence $b_k(n)$ decreases with k for $k \leq \lfloor \frac{n}{2} \rfloor$.

(iv) Fixing k , the sequence $b_k(n)$ decreases with n and converges to $\frac{\beta^k}{\sqrt{3}}$ as $n \rightarrow \infty$.

Proof See the Appendix. ■

When the marginal cost of a firm is hit by a shock, it impacts this firm's price with a coefficient $b_0(n) < 1$. It also impacts other firms' prices with a coefficient that attenuates with the distance between them and the source of the shock. Firm pricing depends directly on the prices of its two immediate neighbors and by transitivity it also depends in equilibrium on the costs of all other firms. Characterization (iii) shows that the more distant a firm is from a competitor, the less it is dependent on this firm's cost in its equilibrium pricing behavior.¹⁰ Furthermore, in markets with more firms, all coefficients are dampened, i.e. the sequence $b_k(n)$ for a given k decreases as n increases. In particular, as firms face more competitors, their prices are less dependent on their own cost level, and more on their competitors' costs.

Entry stage To close the model, I solve for the number of firms entering the market in the entry stage. The expected profit of a firm upon entry is:

$$\Pi = \sum_t \beta^t E[\pi^t(p_k^t, p_j^t, p_{j'}^t)] - F.$$

The per-period profit can be expressed simply, combining equations (1.5), (1.8) and (1.7), as

$$E[\pi^t(p_k^t, p_j^t, p_{j'}^t)] = \bar{M}(1 + E(\gamma^t)) \frac{1}{s} E[(p_k^t - c_k^t)^2].$$

¹⁰Since firms are located on a circle, the distance from firm k to firm i increases until $i - k = \lfloor \frac{n}{2} \rfloor$, then decreases again. This is mirrored by the fact that $b_k(n)$ decreases with k for $k \leq \lfloor \frac{n}{2} \rfloor$ then increases.

This yields the following expression for the profits of an entrant:

$$\Pi = \frac{\bar{M}(1 + E\gamma)}{(1 - \beta)s} E[(p - c)^2] - F,$$

where $\frac{E\gamma}{1-\beta} = \sum_t \beta^t E(\gamma^t)$. Then, by (1.9), $p_i = \sum_{k=1}^n b_{(i-k) \lfloor n \rfloor} c_k + \frac{s}{n}$, where $\lfloor n \rfloor$ indicates that the indexes are computed modulo n . After some algebra, the expectation of interest can be written

$$\Pi = \frac{\bar{M}(1 + E\gamma)}{(1 - \beta)s} \left(c_0^2 \sigma_c^2 \left[\sum_1^{n-1} b_i^2 + (b_0 - 1)^2 \right] + \left(\frac{s}{n} \right)^2 \right) - F.$$

By the free entry condition, this equation pins down the number of firms n that enter the market as a function of market size $\bar{M}$.

Proposition 2 *The equilibrium number of firms increases with market size and decreases with the sunk cost of entry.*

Proof The free entry condition implies that $\frac{F}{\bar{M}} = \frac{1}{1-\beta} \left(\frac{s}{n^2} + \frac{\sigma_c^2 c_0^2 \alpha(n)}{s} \right)$, where $\alpha(n) = \sum_1^{n-1} b_i^2 + (b_0 - 1)^2$ and $\alpha'(n) < 0$.¹¹ Both terms in the right hand side decrease with n . Therefore, as $\bar{M}$ increases or F decreases, the number of firms that enter the market increases. ■

1.2.3 Firm volatility and market size

I now study how the volatility of firms' prices, output and sales vary with market size for markets with a large number of firms n . For a given outcome f (prices, quantities or sales), the volatility of f at the firm level is given by

$$vol(f)^2 = var\left(\frac{f^{t+1} - f^t}{f^t}\right).$$

¹¹see the Appendix for $\alpha'(n) < 0$

Under the maintained assumption that shocks are small, the following first order approximation holds:

$$vol(f)^2 = 2var \left(\frac{\Delta f}{f_0} \right),$$

where f_0 is the value of quantity f at the symmetric unperturbed equilibrium and $\Delta f = f^t - f^0$. Given that $\Delta c_k = c_0 \eta_k$ where the shocks η_k are i.i.d of variance σ_c^2 , the volatility of the prices of a firm is:

$$\begin{aligned} vol(p)^2 &= 2var \left(\frac{1}{c_0 + \frac{s}{n}} \sum_0^{n-1} b_{i-k[n]} \Delta c_k \right) \\ &= 2\sigma_c^2 c_0^2 \frac{\sum_0^{n-1} b_i^2}{(c_0 + \frac{s}{n})^2}. \end{aligned} \tag{1.10}$$

Proposition 3 *The volatility of firms' prices converge to a constant as market size increases.*

I summarize here the main steps of the proof; the formal proof is given in the Appendix. Consider first the impact of a firm's own idiosyncratic cost shock c_i on its prices, absent cost shocks for its competitors. The corresponding price fluctuation is $\frac{\Delta p}{p^0} = \frac{b_0(n)}{c_0 + \frac{s}{n}} \frac{\Delta c_i}{c_i}$, where b_0 measures the firm's own price reaction to a cost shock. The numerator measures the variance of firm prices. Following lemma 7, $b_0(n)$ is positive, smaller than one, and decreases with n , though it converges at a fast rate (in β^n where $\beta < 1$) to a constant. That $b_0(n)$ decreases with market size reflects the fact that in larger markets, as firms face more competitors among which consumer can choose, their prices are more dependent on other firm's competitiveness, and less on their own. Therefore, the variance of prices tends to decrease with market size, though the decrease of $b_0(\cdot)$ with n is weak as $b_0(n)$ converges very rapidly.

The result is similar when taking into account the shocks faced by all the firm's competitors in the market. The variance of firm prices is shaped by a weighted average of the shocks faced by all firms, with weights summing to one and decreasing in the

distance between two firms. This reflects the result that firms are more sensitive to their immediate neighbors' competitiveness than to more distant firms'. Their sensitivity to other firms' competitiveness fades as the distance between two firms increases. Furthermore, holding distance between two firms constant, the weights decrease as the number of competitors increases. The weights converge quickly to their limit level. In large as in small markets, firm's own cost shock and its immediate neighbors' shock have a dominant effect on a firm's prices. Therefore, the variance of a firm's prices, caused by the system of cost shocks that hits all firms, doesn't decrease in $\frac{1}{\sqrt{n}}$ as the law of large numbers would suggest but instead decreases at a quickly decreasing rate and converges to a limit $V = \lim_{n \rightarrow \infty} \sum_0^n b_i^2$.

Price volatility depends on this variance of prices, relative to the price level. The price level also decreases with market size, because markups are smaller there as firms face more competition, and converges to c_0 , the perfect-competition benchmark. This is a classic result in a Salop framework. Overall, the volatility of prices (1.10) converges (from below) to the constant $2\sigma_c^2 V$. For large n , as market size increase, the volatility of prices increases slowly towards this limit.

Second, firm's output in period t is given by

$$Q^t = \bar{M} \int_{x_k - h_{j'}(p_k^t, p_{j'}^t)}^{x_k + h_j(p_k^t, p_j^t)} (dl + dX_l^t).$$

Therefore,

$$\frac{\Delta Q}{Q_0} = \frac{\Delta h_j + \Delta h_{j'}}{2h_0} + \gamma^t + \sigma^2 \bar{W}_0^t,$$

where $\bar{W}_0^t = \frac{\int_{x_k-h_0}^{x_k+h_0} dW_l}{2h_0}$ and $h_o = \frac{1}{2n}$. Equation (1.8) yields $\Delta h_j + \Delta h_{j'} = \frac{\Delta p - \Delta c_k}{s}$ for firm k in equilibrium, so that

$$\begin{aligned} \text{vol}(Q)^2 &= 2\text{var} \left\{ \frac{n}{s} \left((b_0 - 1)\Delta c_i + \sum_{k \neq i} b_{i-k \lfloor n \rfloor} \Delta c_k \right) + \gamma^t + \sigma^2 \bar{W}_0^t \right\} \\ &= 2\sigma_c^2 c_0^2 \left(\frac{n}{s} \right)^2 \left((b_0 - 1)^2 + \sum_{k=1}^{n-1} b_k^2 \right) + \text{var}(\gamma) + \sigma^2 \text{var}(\bar{W}_0). \end{aligned}$$

The volatility of output is driven by demand shocks captured by $\text{var}(\gamma) + \sigma^2 \text{var}(\bar{W}_0)$ and by the response of demand to price shocks, driven by cost shocks. I examine them in turn.

Lemma 4 *The volatility of firm-level demand, $\text{var}(\gamma) + \sigma^2 \text{var}(\bar{W}_0)$, increases with market size.*

Proof The random variable $\bar{W}_0 = \frac{\int_{x_k-h_0}^{x_k+h_0} dW_l}{2h_0}$ is the average of the realized idiosyncratic shocks over the typical market span of a firm, i.e. a segment of length $2h_0 = \frac{1}{n}$. Since dW_l is a Wiener increment, $\int_m^l dW_z \sim \mathcal{N}(0, l - m)$, and $\text{var}(\int_{x_k-\frac{1}{2n}}^{x_k+\frac{1}{2n}} dW_l) = \frac{1}{n}$. It follows that

$$\text{var}(\bar{W}_0) = \text{var}\left(n \int_{x_k-\frac{1}{2n}}^{x_k+\frac{1}{2n}} dW_l\right) = n.$$

Therefore, $\text{var}(\gamma) + \sigma^2 \text{var}(\bar{W}_0) = \text{var}(\gamma) + n\sigma^2$. The primitive processes being identical between large and small markets, $\text{var}(\gamma)$ and σ^2 are constant across market sizes, but as market size increases, n increases by proposition 2, and the volatility of firm-level demand increases. ■

This is the diversification effect. In denser markets, firms target narrower segments of demand whereas in smaller markets, they serve a wider span of demand. At the level of a given firm, idiosyncratic demand shocks are more smoothed out in

smaller markets, where they are averaged over a wider span of neighborhoods (in the geographic interpretation of the Salop circle), or customer types (in its product space interpretation).

The volatility of firm's output is driven by both demand shocks and cost shocks. We have established that demand shocks are more volatile, at the firm-level, in larger markets, and now examine how the volatility of firm's output varies with market size, taking into account both demand and cost shocks.

Proposition 5 *The volatility of firms' output increases with market size, for large n .*

I summarize here the main steps of the proof; the formal proof is given in the Appendix. Consider first the impact of a firm's own idiosyncratic cost shock c_i on its output, absent cost shocks for its competitors ($\Delta c_k = 0$ for $k \neq i$). The volatility of firm's output is then $\frac{\Delta h_i}{h^0} = \frac{n}{s}(b_0(n) - 1)\Delta c_i$. Demand, and hence output, respond negatively to a positive cost shock, all the more as demand is elastic. The price elasticity of demand $\frac{n}{s}$ increases endogenously in larger markets where the number of firm n is higher. Output volatility increases in turn.¹²

I now take into account the cost shocks that hit all firms. Firm's demand responds positively to a cost increase that hits the firm's competitors. The effect is magnified in larger markets where the elasticity of demand is higher. Again, this leads to an increase in output volatility in larger markets, for large n .

Finally, as shown in Lemma 4, the demand shocks also lead to higher volatility of demand, and output, in larger markets. Therefore, both the cost and the demand channel tend to increase the volatility of firms' output in larger markets.

¹²Here $(b_0(n) - 1)^2$ increases as well in larger markets, though at a very low rate. This effect goes in the same direction as the markup effect.

I finally turn to the volatility of firm sales,

$$\begin{aligned} \text{var}\left(\frac{\Delta S}{S_0}\right) &= 2 \text{var}\left(\frac{\Delta p}{p_0} + \frac{dQ}{Q_0}\right) \\ &= 2 \text{var}(B_0(n) \Delta c_i + B_1(n) \sum_{k \neq i} b_{i-k[n]} \Delta c_k + \gamma + \sigma^2 \bar{W}_0), \end{aligned}$$

where $B_0(n) = \frac{n}{s} (b_0(n) - 1) + \frac{b_0(n)}{c_o + \frac{s}{n}}$ and $B_1(n) = \frac{2+c_o \frac{n}{s}}{c_o + \frac{s}{n}}$.

This leads to the following expression for the volatility of sales

$$\sigma_S^2 = 2 c_o^2 \sigma_c^2 \left[B_0^2(n) + B_1^2(n) \sum_{k \neq i} b_{i-k[n]} \right] + \text{var}(\gamma) + \sigma^2 \text{var}(\bar{W}_0).$$

Proposition 6 *The volatility of firm sales increases with market size, for large n .*

Proof See the Appendix. ■

This result takes into account the correlation structure between prices and output. Firm's own cost shocks generate a negative correlation between prices and output, and firm's competitors shocks generate a positive correlation between prices and output. Examining the volatility of sales leads to the conclusion that the impact of market size on sales volatility is positive, for large n

The volatility of sales comes, first, from the reaction of sales to a given set of cost shocks. It itself comes from (a) the reaction of a firm's price to cost shocks and (b) the reaction of demand to a change in prices. The higher elasticity of demand in larger markets leads overall to a higher reaction of sales to a given cost shock. Second, demand shocks directly affect sales, and they also lead to higher volatility in larger markets, through a diversification argument.

I now turn to studying how patterns of firm volatility vary with market size in the data.

1.3 Empirical analysis

I choose to study firm-level volatility in the retail industry, for which demand and competition forces are arguably local. I compare markets of different sizes within a single country, France, an approach often followed in the literature (see for example Syverson (2004), Asplund and Nocke (2006), Campbell and Hopenhayn (2005) or Combes, Duranton, Gobillon, Puga, and Roux (2012)). This allows to minimize the differences in environment faced by firms located in different markets, other than competition forces, that could otherwise confound the results. I analyze how the firm-level volatility of value-added varies for firms located in markets of different sizes, as proxied for by the density of the local area where they are located.

1.3.1 Dataset

I use a panel data of French firms from 1993 to 2006 in the retail and personal services industries. It is an administrative data set which contains annual information on the balance sheet of French firms, declared for tax purposes. Summary statistics are reported in Table 1.1. The dataset includes yearly firm-level data on employment, capital, value-added, production, and 3-digit industry classification. It is matched with establishment-level data, which indicates the geographical location (at the postal code level) of each establishment of a given firm-year. I only keep firms with two employees or more, to minimize measurement error. Data for firms with fewer employees tend to be very noisy.

The geographical markets considered here are the 341 French commuting zones, or ‘*Zones d’emploi*’ (employment zones), within metropolitan France. The indicator of market size that I use is employment density. Using density as an indicator of local scale is common in the empirical economic geography, and is discussed in particular in Combes, Duranton, Gobillon, and Roux (2010). As a robustness check, I show

that the results are robust to alternative specifications of market size, e.g. the total employment or the total population of the commuting zone.

The model does not feature the use of intermediate inputs. Input cost differences in large and small cities could be another source of variation in firm sales volatility across markets, not accounted for by the model. Therefore, I use measures of value-added and not sales to confront the theory with the data.

1.3.2 Main results

Measuring firm-level volatility The data features value-added and employment information at the level of the firm (which can have several establishments), as well as employment and localization information at the level of each establishment. I have to take a stance on how to treat firms that have establishments in several locations, a dimension not accounted for in the model. I describe how I compute two different measures of firm-level volatility, then I discuss the benefits and limitations of each approach.

For the first measure, I treat each establishment that I observe in the data as a firm in the model: establishments are local and respond to their local competitive environment. This measure projects the value-added of a firm on its various establishments, proportionally to the establishment's employment. This gives a proxy for establishment-level value-added. To compute establishment volatility, I then only keep a balanced panel of establishments that survive over the whole period. This allows me to abstract from turnover. Asplund and Nocke (2006) have shown that turnover rates are higher in larger markets; instead, I focus here on the extent to which firm volatility, within surviving firms, differ in larger markets versus small markets, consistent with the focus of the theoretical analysis. This measure of establishment-

level volatility, σ_1 , is then computed as follows:

$$\sigma_{1,i} = \sqrt{\text{var} \left(\frac{\tilde{S}_{i,t+1} - \tilde{S}_{i,t}}{\tilde{S}_{i,t}} \right)}, \text{ for } t = 1 \dots T - 1.$$

The index $i \in \{1, N_e\}$ indexes establishments, where N_e is the number of establishment in the sample. The number of firms in the sample is denoted N_f . Finally, $\tilde{S}_{i,t} = S_{j(i),t} \frac{E_{i,t}}{\sum_{k|j(k)=j(i)} E_{k,t}}$ is the measure of establishment-level value-added. The index $j(i) \in \{1, N_f\}$ corresponds to the index of the firm to which establishment $i \in \{1, N_e\}$ belongs.

Second, I also compute firm-level volatility of value-added for single-establishment firms, for which a direct measure of value-added is available. For this measure, I drop all multi-establishment firms which leaves me with 69% of firms and 55% of establishments in my sample. Finally, I retain a balanced panel of firms that survive over the whole period in order to abstract from firm turnover as noted above. Formally,

$$\sigma_{2,j} = \sqrt{\text{var} \left(\frac{S_{j,t+1} - S_{j,t}}{S_{j,t}} \right)}, \text{ for } t = 1 \dots T - 1$$

where j indexes mono-establishment firms.

The measure of establishment value-added used in σ_1 is indirect, based on employment data. This could induce measurement error. For example, it is likely that the sales per employee are higher in larger cities, as larger cities are more productive on average and have higher prices. The projection I use therefore arguably underestimates the sales in large cities, and overestimates the ones in small. Nevertheless, as volatility is a scale free measure (the variance of the growth rate of value-added), this measurement error should not, per se, systematically bias the measures of volatility that I use. In particular, this error shouldn't lead to volatility being systematically over- or under-estimated in large markets vs small markets. The second measure

restricts the sample to single-establishment firms for which direct measures of value-added are available, which alleviates this first concern. On the other hand, limiting the analysis to mono-establishment firms could lead to a selection bias. If, for example, chains tend to locate in larger markets and have a wide positioning, leaving niche positioning to single-establishment stores, then the strategy and positioning of single-establishment firms are different by nature in large versus small market, leading to composition effects as a possible confounding factor. Because of this, the first measure σ_1 is my preferred measure of analysis.

Results: firm-level volatility and market size I run the following regression to investigate the link between market size and the firm-level volatility of value-added:

$$\sigma_{a,j} = \beta_{a,0} + \beta_{a,1} \ln(MarketSize)_j + X_j + \epsilon_j, \quad \text{for } a = 1, 2$$

$$X_j = \{\ln(FirmSize)_j; Sector_j\}$$

where *MarketSize* is a measure of the size of the market where the establishment operates (employment density, total employment, or total population depending on the specification), X_j is a set of controls that varies depending on the specification, and a indexes the specification used to measure volatility. The standard errors are clustered at the city level. The covariate $\ln(FirmSize)$ controls for the level of firm's value-added. I also include a sectoral fixed effect to control for composition effects: sectors that are more volatile could be over-represented in large or small cities.

Results are presented in Table 1.2. There is a significant positive correlation between firm-level volatility and market size. It is robust to controlling for firm size (column II), which is expectedly negatively correlated with volatility. It is also robust to adding sectoral fixed effects (column III), or changing the measure of market size from employment density (columns I-III) to total employment or total population (columns IV and V). These results correspond to the preferred specification for firm

volatility. The last columns (VI-VIII) show that using σ_2 leads to similar results both quantitatively and qualitatively.

These results are consistent with the qualitative predictions of the model. Larger cities host firms that are more volatile. A concern is that other factors, rather than the role played by the size of local demand and hence competition forces, might be at play and confound the results. To alleviate this concern, I run the following placebo regression. I study the volatility of firms in the tourism industry (hotels, tourist attractions...). This industry is similar to retail in many respects, but differs in that the demand it faces is not only driven by the size of the local demand. Therefore, the forces at play in the model should not be well captured by a measure of city size in this industry. Table 1.3 reports the results. For these sectors, the correlation between city size and volatility is very small and not significantly different from zero. Given the relatively small sample used here, the estimates are noisy, but the point estimate is an order of magnitude smaller than for retailers.

I also consider firms in the manufacturing sector that produce goods in export-intensive industries. The idea here is also to study sectors for which demand is not local. Again, Table 1.3 shows that in contrast to the retail sector, there is no significant positive correlation between the size of the city where the firm is located and its volatility. This provides suggestive evidence that local demand and local competition, as opposed to other factors not captured in the model, play a role in shaping the observed patterns of firm volatility in the non tradable sector.

1.4 Conclusion

This paper documents a perhaps surprising fact, that the firm-level volatility of value-added tends to be higher in larger cities, for firms producing in the non tradable sector. I provide a simple explanation for this fact. Larger cities foster more competition,

which magnifies the idiosyncratic shocks that hit firms. The model also predicts that firms tend to focus on narrower segments of demand in larger cities. By positioning themselves on niche markets, they are subject to higher demand volatility, which reinforces the effect. Disentangling empirically the two effects - competition and niche - is an interesting question left for future research.

1.5 Tables and figures

Table 1.1: Summary statistics - retail establishments

	mean	sd	p50	p10	p90	count
Panel A - Small cities						
Empl	10.17	21	6	2	20	12,426
Ln(va)	7.60	1	8	6	9	12,404
Panel B - Large cities						
Empl	11.27	38	6	2	20	12,278
Ln(va)	7.68	1	8	6	9	12,268

Summary statistics for a balanced sample of establishments that survives from 1993 to 2006. Ln(va) is the value-added of the firm pro-rated to each establishment. "Large cities": cities with the highest employment density, hosting half of the firms in the sample."Small cities": cities with the lowest employment density, hosting half of the firms in the sample.

Table 1.2: Value-added volatility and market size

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Measure of volatility	σ_1	σ_1	σ_1	σ_1	σ_1	σ_2	σ_2	σ_2
MarketDensity	0.027*** (0.004)	0.027*** (0.004)	0.028*** (0.003)			0.035*** (0.006)		
MarketPop				0.050*** (0.010)			0.051** (0.019)	
MarketEmp					0.045*** (0.007)			0.048** (0.015)
Ln(va)		-0.119*** (0.009)	-0.114*** (0.009)	-0.114*** (0.009)	-0.114*** (0.009)	-0.179*** (0.012)	-0.179*** (0.012)	-0.179*** (0.012)
Sector FE	N	N	Y	Y	Y	Y	Y	Y
R-squared	0.004	0.023	0.068	0.068	0.069	0.113	0.109	0.110
Observations	22148	21492	21492	21492	21492	16041	16041	16041
Cluster	city	city	city	city	city	city	city	city

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard error are clustered at the city ("Zone d'emploi") level.

Table 1.3: Placebo

	(1)	(2)	(3)	(4)	(5)
	Retail	Export intensive		Tourism	
MarketDensity	0.035*** (0.006)	0.005 (0.012)		0.001 (0.009)	
MarketPop			-0.006 (0.018)		0.006 (0.022)
Ln(va)	-0.179*** (0.012)	-0.142*** (0.012)	-0.144*** (0.013)	-0.236*** (0.044)	-0.238*** (0.044)
R-squared	0.113	0.110	0.110	0.071	0.071
Observations	16271	2071	2071	492	492
Cluster	city	city	city	city	city

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard error are clustered at the city ("Zone d'emploi") level.

1.6 Appendix

Derivation of profits $p_i = \sum_{k=1}^n b_{(i-k)[n]} c_k + \frac{s}{n}$ and $E[c_k^t] = c_0$ for all k . Using that $\sum_{k=1}^n b_{(i-k)[n]} = 1$ (see proof of Lemma 1), we get: $E[(p^t - c_i^t - \frac{s}{n})] = 0$. It follows that:

$$\begin{aligned} E[(p^t - c_i^t)^2] &= E[(p^t - c_i^t - \frac{s}{n} + \frac{s}{n})^2] \\ &= E\left[\left(\sum_{k=1}^n b_{(i-k)[n]} c_k - c_i\right)^2\right] + \left(\frac{s}{n}\right)^2 \\ &= c_0^2 \sigma_c^2 \left[\sum_1^{n-1} b_i^2 + (b_0 - 1)^2\right] + \left(\frac{s}{n}\right)^2 \end{aligned}$$

where the last line uses that the η_k are independent variables of variance σ_c^2 .

Proof of Lemma 1

(i) Since $\mathbf{A} = \text{circ}(2, -.5, 0, \dots, 0, -.5)$ is circulant and symmetric,¹³ its inverse is also circulant and symmetric (the invertibility of $\mathbf{A}$ is guaranteed as it is diagonally dominant). Writing $\mathbf{A}^{-1}(n) = \text{circ}(b_0(n), b_1(n), \dots, b_{n-1}(n))$, this yields $b_k(n) = b_{n-k}(n) \forall k = 1, \dots, n-1$.

(ii) The system of equations that define $b_i(n), i = 0 \dots n-1$ is:

$$2b_0(n) - .5b_{n-1}(n) - .5b_1(n) = 1 \quad \text{and}$$

$$2b_i(n) - .5b_{i-1}(n) - .5b_{i+1}(n) = 0, \quad \text{for } i = 1 \dots n-1,$$

where the subscripts are modulo n . Summing these expressions gives $\sum_0^{n-1} b_i(n) = 1$. Formally solving for the $b_k(n)$ gives¹⁴:

$$\begin{aligned} b_k(n) &= \frac{1}{2} \left[(2 + \sqrt{3})^k (b_0(n) - \frac{1}{\sqrt{3}}) + (2 - \sqrt{3})^k (b_0(n) + \frac{1}{\sqrt{3}}) \right] \quad \text{for } k = 1 \dots n-1 \quad \text{and} \\ b_0(n) &= \frac{1}{\sqrt{3}} \frac{(2 + \sqrt{3})^{\frac{n-3}{2}} (5 + 3\sqrt{3}) - (2 - \sqrt{3})^{\frac{n-3}{2}} (5 - 3\sqrt{3})}{(2 + \sqrt{3})^{\frac{n-3}{2}} (5 + 3\sqrt{3}) + (2 - \sqrt{3})^{\frac{n-3}{2}} (5 - 3\sqrt{3})}. \end{aligned}$$

¹³A circulant matrix $\text{circ}(a_0, a_1, a_2, \dots, a_{n-1})$ is also symmetric iff $a_k = a_{n-k}, \forall k = 1, \dots, n-1$

¹⁴Details of the computation are as in Vogel (2008).

Noting that $(2 - \sqrt{3})(2 + \sqrt{3}) = 1$ and writing $\beta = 2 - \sqrt{3}$ (in particular, $\beta > 0$ and $\beta < 1$) and $K = \frac{3\sqrt{3}-5}{3\sqrt{3}+5}$ (in particular, $K > 0$ and $K < 1$), the expression for $b_0(n)$ simplifies to:

$$b_0(n) = \frac{1}{\sqrt{3}} \frac{1 + \beta^{n-3}K}{1 - \beta^{n-3}K} \quad (1.11)$$

It follows that, for $k = 1 \dots n - 1$,

$$b_k(n) = \frac{1}{2}(\beta^{-k} + \beta^k)b_0(n) + \frac{1}{2\sqrt{3}}(-\beta^{-k} + \beta^k) \quad (1.12)$$

Since $0 < \beta < 1$, it is readily seen from (1.11) that $b_0(n) > \frac{1}{\sqrt{3}}$ and from (1.12) that $b_k(n) > 0$ in turn, for $k > 0$.

(iii) I now show that fixing n , $b_k(n)$ decreases with k iff $k \leq \frac{n}{2}$ (then, since $b_k(n) = b_{n-k}(n)$, $b_k(n)$ increases with k for $k \geq \frac{n}{2}$). Fix n . Then the derivative of $b_k(n)$ with respect to k is of the sign of:

$$\frac{db_k(n)}{dk} \propto \left(\frac{b_0(n)}{2} - \frac{1}{2\sqrt{3}}\right)\beta^{-k} - \left(\frac{b_0(n)}{2} + \frac{1}{2\sqrt{3}}\right)\beta^k$$

Both terms in parenthesis being positive, this derivative increases with k . It is equal to zero when $\beta^{2k} = \frac{\frac{b_0(n)}{2} - \frac{1}{2\sqrt{3}}}{\frac{b_0(n)}{2} + \frac{1}{2\sqrt{3}}} = \beta^{n-3}K$, using equation (1.11). Since $\frac{\log K}{\log \beta} = 3$, this corresponds to $k = \frac{n}{2}$. Therefore, for $k < \frac{n}{2}$, $\frac{db_k(n)}{dk} < 0$.

(iv) I finally show that fixing k , $b_k(n)$ decreases with n and converges towards $\frac{\beta^k}{\sqrt{3}}$. Differentiating equation (1.11) with respect to n shows that $b_0(n)$ decreases with n . In turn, as $b_k(n)$ simply depends on n through $b_0(n)$, it is readily seen that $b_k(n)$ also decreases with n . Finally, as $n \rightarrow \infty$, $\beta^n \rightarrow 0$ and a Taylor expansion of equation (1.11) yields $b_0(n) \sim \frac{1}{\sqrt{3}}(1 + 2\beta^{n-3}K)$. In turn, equation (1.12) yields $b_k(n) \sim \frac{\beta^k}{\sqrt{3}} + \beta^{n-3} \frac{K}{\sqrt{3}}(\beta^{-k} + \beta^k) + o(\beta^{n-3})$. Therefore, $b_k(n) \rightarrow \frac{\beta^k}{\sqrt{3}}$ as $n \rightarrow \infty$. ■

Proof of Proposition 2

I first derive a lemma that will be useful in all proofs.

Lemma 7 *The sum $\sum_1^{n-1} b_i(n)^2$ equals $\frac{(\beta^2 - \beta^{2n})(\frac{K^2}{\beta^6 - \beta^8} + \frac{1}{1 - \beta^2}) + 2K\beta^{n-3}}{(1 - \beta^{n-3}K)^2}$. It decreases with n and converges to $\beta^2(\frac{K^2}{\beta^6 - \beta^8} + \frac{1}{1 - \beta^2}) \sim .15$.*

The expression for $\sum_1^{n-1} b_i(n)^2$ comes from straightforward algebra, given the expression for $b_i(n)$. The denominator increases with n , as $\beta < 1$. Differentiating the numerator with respect to n leads to a derivative of the sign of $\log(\beta)[-2\beta^n(\frac{K^2}{\beta^6 - \beta^8} + \frac{1}{1 - \beta^2}) + \frac{2K}{\beta^3}]$, with $\log(\beta) < 0$. The term in brackets increases with n and is positive at $n = 1$ (~ 1.69). Therefore the derivative of the numerator is negative for all n . This proves that $\sum_1^{n-1} b_i(n)^2$ decreases with n . Given that $\beta^n \rightarrow 0$, $\sum_1^{n-1} b_i(n)^2 \rightarrow \beta^2(\frac{K^2}{\beta^6 - \beta^8} + \frac{1}{1 - \beta^2})$ as $n \rightarrow \infty$. ■

Going back to proposition 2, I compute $\alpha(n) = \sum_0^{n-1} b_i^2 + 1 - 2b_0$ and find, after some algebra, that

$$\alpha(n) = \frac{cst + \beta^{2n} \left[\frac{K^2(1+\sqrt{3})^2}{\beta^3} - 3\left(\frac{K^2}{\beta^6 - \beta^8} + \frac{1}{1 - \beta^2}\right) \right] + 10K\beta^{n-3}}{(1 - \beta^{n-3}K)^2},$$

where the constant in brackets is positive. This yields the first result: $\alpha'(n) < 0$. ■

Volatility with small shocks Let f stand for firm's price, sales or output. The following approximation holds to the first order:

$$\begin{aligned} \frac{f_{t+1} - f_t}{f_t} &\sim \left(\frac{f_{t+1} - f_0}{f_t} - \frac{f_t - f_0}{f_t} \right) \\ \frac{f_{t+1} - f_t}{f_t} &\sim \left(\frac{f_{t+1} - f_0}{f_0} - \frac{f_t - f_0}{f_0} \right) \left(1 - \frac{f_t - f_0}{f_0} \right) \\ &\sim \frac{f_{t+1} - f_0}{f_0} - \frac{f_t - f_0}{f_0} \end{aligned}$$

Moreover, f_{t+1} and f_t are independent. This comes from the fact that pricing decisions depend only on current cost shocks, which are uncorrelated, and do not depend on past shocks despite the fact that demand shocks are observed through past demand. Also, demand shocks are independent. The quantities f_{t+1} and f_t are also identically distributed

as all primitives are. This leads to:

$$vol(f) \sim 2 var\left(\frac{f^t - f_0}{f_0}\right) = var\left(\frac{\Delta f}{f_0}\right)$$

Proof of Proposition 3

Price volatility is given by $vol(p)^2 = 2 \sigma_c^2 c_0^2 \frac{\sum_0^{n-1} b_i^2}{(c_0 + \frac{s}{n})^2}$.

(i) By lemmas 7 and 1, $\sum_0^{n-1} b_i^2$ converges to a limit L_1 . Since $(c_0 + \frac{s}{n})^2 \rightarrow c_0^2$, $vol(p)^2 \rightarrow cst$ as $n \rightarrow \infty$.

(ii) A Taylor expansion at $n \sim \infty$ gives $\sum_0^{n-1} b_i^2 \sim L_1 + 2\beta^n + o(\beta^n)$, therefore $\frac{\sum_0^{n-1} b_i^2}{(c_0 + \frac{s}{n})^2} \sim \frac{1}{c_0^2} (L_1 - \frac{2L_1 s}{c_0 n} + o(\frac{1}{n}))$ and the convergence is from below.

Proof of Lemma 4 The proof is in the main text.

Proof of Proposition 5

Output volatility is given by

$$vol(Q)^2 = 2\sigma_c^2 c_0^2 \left(\frac{n}{s}\right)^2 \left((b_0 - 1)^2 + \sum_1^{n-1} b_i^2 \right) + var(\gamma) + \sigma^2 var(\bar{W}_0).$$

A Taylor expansion at $n \sim \infty$ gives $(b_0 - 1)^2 + \sum_0^{n-1} b_i^2 \sim (\frac{1}{\sqrt{3}} - 1)^2 + L_1 + K_2 \beta^n + o(\beta^n)$ with $K_2 > 0$. Therefore, $\left(\frac{n}{s}\right)^2 \left((b_0 - 1)^2 + \sum_1^{n-1} b_i^2 \right) \sim \left(\frac{n}{s}\right)^2 \left((\frac{1}{\sqrt{3}} - 1)^2 + L_1 \right) + o(1)$ and, together with lemma 4, this yields that $vol(Q)^2$ increases with n for large n . ■

Proof of Proposition 6

Sales volatility is given by

$$\sigma_S^2 = 2 c_o^2 \sigma_c^2 \left[B_0^2(n) + B_1^2(n) \sum_{k \neq i} b_{i-k[n]} \right] + var(\gamma) + \sigma^2 var(\bar{W}_0).$$

I first examine the derivative of $B_0(n)^2$ with respect to n . Writing $B_0(n) = \frac{2b_0(n)-1+c_0\frac{n}{s}(b_0(n)-1)}{c_0+\frac{s}{n}}$, it follows that $B_0(n)^2 = \frac{A_0(n)^2}{(c_0+\frac{s}{n})^2}$. The denominator decreases with n . Then $\frac{dA_0(n)}{dn} = \frac{c_0}{s}(b_0(n)-1) + b'_0(n)(c_0\frac{n}{s}+2) < 0$ as $b_0(n) < 1$ and $b'_0(n) < 0$ (from equation (1.11), $b_0(n)$ decreases monotonically from .6 to $\frac{1}{\sqrt{3}}$). Also, $A_0(n) < 0$ when $\frac{2b_0(n)-1}{n(1-b_0(n))} < \frac{c_0}{s}$. Since $\frac{2b_0(n)-1}{1-b_0(n)}$ decreases towards 0 for large n , $B_0(n)^2$ increases with n for large n .

I then examine $B_1(n)^2$. We have established in lemma 1 that $\sum_{k \neq i} b_{i-k[n]}^2$ decreases and converges to $\sum_1^{n-1} b_i(n)^2 \rightarrow \beta^2(\frac{K^2}{\beta^6-\beta^8} + \frac{1}{1-\beta^2}) \sim .15$. Write l this limit. For large n , $\sum_1^{n-1} b_i(n)^2 = l + 4K\beta^{n-3} + o(\beta^n)$. Also, $B_1(n)^2 = n^2\frac{2}{c_0^2s^2} + 4n\frac{1}{c_0s} + o(1)$. Therefore $B_1(n)^2 \sum_{k \neq i} b_{i-k[n]}^2 = n^2\frac{2l}{c_0^2s^2} + o(n^2)$ increases with n for large n . ■

Chapter 2

Market Size and Industry Dynamics

2.1 Introduction

Models of firm dynamics are traditionally based on constant elasticity of substitution (CES) preferences, which imply constant mark-ups. In contrast, a robust prediction of oligopoly theory is that larger or more open markets are more competitive and have lower price-cost markups. Yet, the implications of this force for firm dynamics are not well known. In this paper, I propose a tractable framework of firm dynamics with variable markups. The model delivers new insights on the forces that shape firms' entry and exit decisions in markets of different sizes. First, firm selection is 'tougher' in larger markets that are endogenously more competitive, as in the static workhorse model of Melitz and Ottaviano (2008). In addition, firm selection is also governed by the option value that firms derive from staying in operation in the market. This value depends on the amount of uncertainty firms face, which systematically increases with market size in the data.

The model offers a coherent explanation of two stylized facts that I document here on a single, extensive dataset of French firms. Consistent with the analysis of Combes, Duranton, Gobillon, Puga, and Roux (2012), I document that markets of different sizes do not exhibit significant differences in productivity thresholds, even in industries where demand is inherently local. In other words, a simple cross-sectional analysis fails to detect that competition is ‘tougher’ in larger markets. Second, I analyze the link between market size and firms’ survival rates. Attrition is significantly and robustly larger in larger markets. If the first fact is consistent with a model where the intensity of competition is not impacted by market size, this second fact is not.

A dynamic analysis offers a possible explanation for this apparent puzzle. Entry and exit decisions are not only shaped by competition forces, but also by the extent of firm-level uncertainty. The uncertainty faced by firms is a key determinant of productivity thresholds in any *dynamic* model of firm behavior. There is an option value attached to uncertainty: relative to a static model, less efficient firms can stay active in hope of large positive productivity shocks in the future. In other words, in more volatile markets, selection on current productivity is not as severe as it is in more certain environments. Crucially, the data reveals that firm-level volatility is significantly higher in larger markets. This fact is robust to different measures of firm-level volatility and different specifications. A consequence of this new stylized fact is that the option value of remaining active or entering a larger market is higher than in a smaller market. Thus, while larger markets can exhibit stronger competitive forces, the larger uncertainty observed in larger markets also fosters entry by less efficient firms, so that the overall link between market size and productivity thresholds is a priori ambiguous.

The structure of the model is as follows. Demand follows the linear demand system of Ottaviano, Tabuchi, and Thisse (2002), which generates variable markups. Firms are heterogeneous in their marginal cost of production and are forward-looking in

their entry and exit decision. Their cost evolves over time according to a stochastic process. To build intuitions and derive analytically tractable comparative statics, I first consider a simple, discrete time process for marginal costs based on Asplund and Nocke (2006). Each period, firms either keep their last period cost or draw a new, uncorrelated cost from an exogenous distribution. At the steady state, the cost threshold above which firms decide to exit declines with market size: the tougher competition induced in larger markets puts pressure on markups and drives high-cost firms out. In contrast, more uncertainty on future cost shocks benefits firms with current high costs. This dynamic effect, novel to my analysis, raises the cost threshold. When larger markets are also more volatile, which is a robust stylized fact in my sample, the effect of market size on firms' entry and exit decisions becomes ambiguous. Survival rates, however, are always lower in larger markets, which is consistent with the second stylized fact that I document.

I then generalize the analysis to a more realistic process. Firms' costs evolve according to a geometric brownian motion, a traditional assumption in the firm dynamics literature.¹ I solve for the steady-state general equilibrium of the industry. A simple calibration of the model shows that, as in the discrete-time model, larger markets have lower cost thresholds and more volatile markets have higher cost thresholds. Furthermore, when volatility increases with market size at the rate I measure in the data, the net effect of an increase of market size on the cost threshold is small and, for a range of parameter values, within the margin of error of the procedure of Combes, Duranton, Gobillon, Puga, and Roux (2012).

The first contribution of the paper is methodological and, in this sense, close to the analysis made in Asker, Collard-Wexler, and De Loecker (2013). In the dynamic framework I propose, larger cities can feature tougher competition and lower markups, and yet exhibit similar productivity cutoffs compared to smaller markets. This ap-

¹see Luttmer (2007) and Arkolakis (2011) for example.

parent contradiction emanates from the different firm dynamics that prevail in larger markets. At the steady-state, these dynamics shape ‘static’ market outcomes, such as the firm-size distribution. It can therefore be misleading to analyze such outcomes through the lens of a static model. In short, the minimum productivity threshold is a sufficient statistic to measure competition effects in a static model, but it is not sufficient in a dynamic model. Additional dynamic moments are needed (such as exit rates) to measure competition effects.

The paper also contributes to the literature on firm dynamics. I propose a tractable framework with variable markups *a la* Melitz and Ottaviano (2008) extended with rich firm dynamics. In contrast, existing models of industry dynamics, notably Luttmer (2007) or Arkolakis (2011) rely on CES demand and constant markups. A notable exception is Asplund and Nocke (2006). Their analysis attributes the higher turnover rate in larger markets to the lower cost thresholds that prevail in these markets. However, this explanation directly contradicts the evidence in Combes, Duranton, Gobillon, Puga, and Roux (2012), which fail to detect significant differences in productivity thresholds between small and large markets.

More broadly, this research contributes to the urban and industrial organization literatures that study how city- or market- size shapes firm-level outcomes. Existing empirical evidence on the pro-competitive effect of market size is ambiguous. On the one hand, Syverson (2004) and Campbell and Hopenhayn (2005) provide evidence that market size has pro-competitive effects. On the other hand, Combes, Duranton, Gobillon, Puga, and Roux (2012) conclude that the difference in firm outcomes in large versus small cities is driven by agglomeration externalities and not competition forces. While I confirm the empirical findings of Combes, Duranton, Gobillon, Puga, and Roux (2012), I show that they can be consistent with a model where competition forces are tougher in larger markets.

The remainder of the paper is organized as follows. Section 2.2 presents the stylized facts, Section 2.3 presents the stylized model, Section 2.4 presents the main model and Section 2.5 concludes.

2.2 Stylized facts

This section presents a list of stylized facts based on French firm-level data. The theoretical framework developed in the next sections aims at proposing a coherent interpretation of these facts. These stylized facts summarize the relationship between market size and a set of firm-level outcomes, with a particular focus on dynamic outcomes such as firm turnover and firm-level volatility. I follow the strategy commonly employed in the literature and use the size of the city where establishments are located as a proxy for the size of the market.

2.2.1 Data

I use a panel dataset of French firms from 1993 to 2006. The dataset includes yearly firm-level data on employment, capital, value added, production, and 3-digit industry classification. It is matched with establishment-level data, which indicates the geographical location at the postal code level of each establishment of a given firm-year. I only keep firms with two employees or more, to minimize measurement error. Data for firms with fewer employees tend to be very noisy.

The geographical markets considered here are the 341 French commuting zones, or “*Zones d’emploi*” (employment zones), within metropolitan France. Employment density is used as a proxy for market size. Using density as an indicator of local scale is common in the empirical economic geography literature and is discussed in particular in Combes, Duranton, Gobillon, and Roux (2010).

Firm entry and exit are measured as follows. Firms are considered entrants when they enter the panel strictly after 1993. Similarly, exiting firms are the ones that leave the panel strictly before 2006. This is an imprecise measure of exit, as firms may come out of the panel if they experience a change in tax reporting obligations for example. To get an alternative measure of firm exit, I also use a dataset recording the official bankruptcy date of a firm, as ruled by a bankruptcy court. These courts have a saying when a firm is in debt. This definition of exit suffers from a different caveat: firms that do not finance themselves through debt – which tend to be larger firms – are not represented in this dataset. I use both indicators of exit in the analysis.

I assume that the production function is Cobb-Douglas, for each 2-digit industry. Firm-level TFP ϕ_{it} is the residual of the OLS regression, run separately for each industry:

$$q_{it} = \alpha_l l_{it} + \alpha_k k_{it} + \phi_{it}, \quad (2.1)$$

where i indexes firms, t indexes years, and q , l and k are the logs of firm value added, number of employees and capital stock. As is well known, this simple procedure may lead to biased estimates of α_l and α_k . I also report as robustness checks the results obtained with the method proposed by Levinsohn and Petrin (2003) to account for the potential endogeneity of capital and labor. The results are very similar under these two specifications.

Summary statistics Table 2.1 reports summary statistics. There are 2,910,000 firm-year observations, for 13 years of data. The vast majority of firms (2,330,000 firm-year observations) have either a single establishment or all of their establishments in the same zone. I retain only these firms in the analysis, as their localization information can be matched with balance sheet data. Table 2.2 reports additional summary statistics for this subset of firms, by size of the market where firms are located. Firms located in larger market have consistently higher TFP levels.

2.2.2 Comparison of small and large markets

Firms are more productive in larger cities. This is an established and consensual empirical regularity. A traditional explanation put forward in the literature is that there is a tougher selection of firms at play in larger cities, where competition is fiercer. As firms face more competitors, markups are pressured down and only the most productive firms survive. This class of models predicts that the minimum productivity thresholds required for entry is higher in larger markets. I study empirically how firms' minimum productivity threshold differ in cities of different sizes.

Productivity threshold Combes, Duranton, Gobillon, Puga, and Roux (2012) develop an econometric procedure to compare minimum productivity thresholds between large and small cities. It is an infinite moment procedure that compares the whole distribution of log-TFP between two given markets.

They examine three possible reasons why firms are measured to be more productive, on average, in larger cities. First, it could come from a selection effect driven by tougher competition in larger markets. In that case, the log-TFP distribution of firms in the two markets should differ in their left tail, where the selection occurs. Second, it could be that all firms are made more productive in larger cities by positive agglomeration externalities. In that case, the log-TFP distribution of firms in large versus small cities should be shifted versions of one another. Third, it could be that more productive firms are disproportionately made more productive in larger markets by agglomeration externalities. In that case, the log-TFP distribution of firms in larger markets should be a 'dilated' version of the log-TFP distribution of firms in smaller markets, with more dispersion in the right tail.

The procedure compares two distributions of firm-level log-TFP, and backs out three parameters that measure these three types of differences between distributions, i.e. parameters measuring the extent to which the two distributions differ in the

left tail (‘truncation’ parameter), differ in level (‘shift’ parameter) and differ in the right tail (‘dilation’ parameter). Combes, Duranton, Gobillon, Puga, and Roux (2012) study manufacturing firms in France, and compare the distribution of firms located in large cities to the one in small cities. They fail to detect a difference in their minimum productivity thresholds. Consequently, they attribute the higher productivity of firms in large cities solely to TFP-enhancing agglomeration externalities.

A limitation of the approach is that the empirical analysis is made on manufacturing firms, for which city size is an imperfect proxy of market size. I therefore apply the same infinite moment procedure to retail and personal services industries. These industries are the ones for which one should expect the effect of local demand and local competition to be the strongest. They have high, if not infinite, trade costs: people shop at their local grocery store and go to their local hairdresser. Therefore, firms in these sectors are heavily dependent on local market conditions.

Table 2.3 presents the estimates for retail and personal services industries. Surprisingly, the estimates are strikingly similar to what was found in Combes, Duranton, Gobillon, Puga, and Roux (2012) for manufacturing. The procedure fails to detect any difference in minimum productivity threshold between large and small markets. This finding constitutes the first stylized fact that guides my theoretical analysis:

Fact 1: Large and small markets have similar minimum productivity thresholds.

There could be two reasons for this finding. First, it could be that the forces that drive firm selection do not differ between large and small markets. For example, in two prominent models of industry dynamics based on a CES demand structure, Melitz (2003a) and Luttmer (2007), larger markets are not more competitive. Market size plays no role on the productivity distribution of firms.

Second, it could be that larger market are more competitive, but that other forces leads to counterbalancing effects on the exit threshold.

To disentangle further between these two hypothesis, I now examine a dynamic dimension of the data. I investigate the exit dynamics of firms, noting that CES-based models predict no differences in firm dynamics between small and large markets, as market size plays no role on firm selection or firm attrition.

Survival rates Asplund and Nocke (2006) have shown that firm turnover is higher in larger markets, based on a dataset of swedish hair salons. I examine the French data in that dimension, in order to establish the fact on a more representative sample of firms, and on the same dataset as the one used to establish Fact 1. Furthermore, their explanation for the higher turnover of firms in larger markets relies on the dynamic implications of a differential productivity cutoff in large versus small markets. I establish here that firm survival rates are lower in larger markets, even though the minimum productivity cutoff is identical in large versus small markets, on the same dataset.

Table 2.4 reports the estimates of the following probit regression :

$$\theta_i^5 = MarketSize_i + Sector_i + Year_Entry_i + X_i + \epsilon_i, \quad (2.2)$$

where θ_i^5 is an indicator variable equal to 1 if and only if a given firm survives 5 years or more. $MarketSize_i$ indicates the size of the market where the firms operate, i.e. its log-employment density. The regression includes sector and year of entry fixed effects. In specification (2), I add controls to alleviate the concern that the results could be driven by composition effects. For example, firms may be systematically larger and more able to survive negative shocks in larger or smaller markets. The controls added are year of entry log-TFP, log-Employment, and log-Capital at the firm level. Finally, specifications (3) and (4) estimate the same regression but for alternative measures of survival, namely (3) the probability of surviving 10 years or more, and (4) the probability of being bankrupt before 5 years.

Table 2.4 presents estimates for these four specifications. The probability of firm exiting is higher in larger markets, and significantly so. This holds when controlling for observable composition effects between markets of different sizes, and is robust to alternative specifications of firm exit. I summarize this robust relationship between market size and probability of survival in Fact 2.

Fact 2: Survival rates are lower in larger markets.

Taken together, Facts 1 and 2 are puzzling. Models where market size plays no role can account for Fact 1 but not Fact 2. An alternative explanation is that larger market *are* more competitive, but that other forces lead to counterbalancing effects on the exit threshold. One such possible force is the level of uncertainty on the market, a key determinant of exit thresholds in any dynamic model of firm behavior (see for example Dixit and Pindyck (1994)). There can be an option value to this uncertainty: relatively inefficient firms can stay active to leverage their option value and hope for high-return positive shocks in the future. Following this intuition, I investigate the dynamics of firm-level productivity in the data, and in particular how its level of uncertainty compares in large and small markets.

Volatility of productivity I exploit the panel dimension of the data and measure the volatility of the firm-level productivity process. Table 2.5 shows the results of an AR(1) regression for the productivity process, as is common in the literature (see for example Asker, Collard-Wexler, and De Loecker (2013) or Castro, Clementi, and Lee (2011)). The persistence coefficient ρ and the value of the volatility of the process σ are assumed to be constant across time and industries:

$$\phi_{i,t+1} = a_s + a_t + \rho\phi_{i,t} + \sigma\epsilon_{i,t}, \quad (2.3)$$

where $\phi_{i,t}$ is the firm-level log-TFP for firm i in year t , a_s is a 3-digit sector fixed effect that controls for a sector-specific trend in productivity growth, a_t is a year fixed effect and $\epsilon_{i,t}$ is an idiosyncratic productivity shock of mean 0 and unit variance. The volatility $\hat{\sigma}$ is the standard deviation of the residuals. To check that the sample I use in analysis is not systematically different from the whole sample, I compare the results for all firms (specification (2)) and single establishment firms only (specification (3)). The volatility is the same in both cases, at 0.37. As a robustness check, I report in columns (4)-(6) the same regressions for an alternative measure of firm level TFP based on the Levinsohn and Petrin (2003) procedure. The results are similar.

Table 2.7 reports the estimates of regression (2.3), estimated separately by quartile of market size, to allow for coefficients and variance of shocks to vary across markets of different sizes. The volatility of the productivity process is higher in larger markets. Levene's and Brown and Forsythe's robust tests for variance equality strongly reject the null hypothesis of an identical variance between the four groups of firms, and between each pair of markets.

The summary statistics reported in Table 2.2 points at the fact that beyond productivity, firms are different in different markets. To check whether this composition effect is not what is driving the result, I conduct additional regressions in Table 2.8. In specification (1), I control for firm size in labor and capital, according to the following regression :

$$\phi_{i,t+1} = a_s + a_{size} + a_t + \rho_{size}\phi_{i,t} + \beta_l l_{i,t} + \beta_k k_{i,t} + \epsilon_{i,t}, \quad (2.4)$$

where a_{size} is a market size fixed effect, the persistence coefficient ρ_{size} is specific to each quartile of market size, and l and k are log-employment and log-capital. I then compute the variance of the residuals, by quartile of market size (σ_1 to σ_4), and conduct a Levene test for equal variance between all four groups, as well as two by two. In specification (2), the same regression is conducted on the subsample of

firms that have 5 to 10 employees, and in specification (3), for firms with 30 to 50 employees². Across all specifications, the volatility of the TFP process is globally increasing in the size of the market. This is robust to the inclusion of several controls for the size of firms, which seems to indicate that these results are not driven by composition effects. I summarize these findings on the volatility of the productivity process into Fact 3:

Fact 3: The volatility of the productivity process is higher in larger markets.

I now present a theoretical framework of firm dynamics, where market size shapes firm dynamics. After I present the model, I shall return to the three facts presented above and interpret them coherently in the light of the theory.

2.3 A simple dynamic model

I first study a simple dynamic version of the Melitz and Ottaviano (2008) framework, following Asplund and Nocke (2006). The simple model guides intuitions and allows for comparative statics that can be analytically derived. I assume that time is discrete, and use the time index t to index periods when necessary.

2.3.1 Model setup

Demand structure There is a mass L of identical consumers in a given market. Each consumer has preferences over a consumption stream $\{U_t\}$ of goods from which she derives utility according to:

$$\sum U_t \left(\frac{1}{1+r} \right)^t,$$

²These are meant to be illustrative. The results stay globally robust for other firm sizes as well, not reported here.

where consumers discount the future at rate r . The per-period demand structure is similar as the one in Melitz and Ottaviano (2008). Consumers derive utility from the consumption of an homogeneous good ($q_{0,t}$) taken as numeraire and a continuum of different varieties ω of a differentiated product ($q_t(\omega)$, for $\omega \in \Omega_t$ the set of varieties offered). The per-period utility function for each individual consumer is:

$$U_t = q_{0,t} + a \int q_t(\omega) d\omega - \frac{1}{2} \gamma \int q_t(\omega)^2 d\omega - \frac{1}{2} \eta \left(\int q_t(\omega) d\omega \right)^2, \quad (2.5)$$

where $\gamma > 0$ governs the degree of product differentiation between varieties, and $\eta > 0$ and $a > 0$ index the substitution between the differentiated varieties and the homogenous good. This per-period utility leads to the following per-period aggregate demand for variety ω :

$$q_t(\omega) = \frac{aL}{\eta N_t + \gamma} - \frac{L}{\gamma} p_t(\omega) + \frac{\eta N_t}{\eta N_t + \gamma} \frac{L}{\gamma} \bar{p}_t, \quad (2.6)$$

where N_t is the measure of varieties consumed in period t , and $\bar{p}_t$ is the average price paid for varieties that are consumed. This average price is $\bar{p}_t = \frac{1}{N_t} \int_{\omega \in \tilde{\Omega}_t} p_t(\omega) d\omega$, where $\tilde{\Omega}_t$ is the set of varieties consumed. As a result, there is no demand for the good above the following choke price A_t :

$$A_t \equiv \frac{\gamma a + \eta N_t \bar{p}_t}{\eta N_t + \gamma}. \quad (2.7)$$

Production Labor is the only factor of production. The homogeneous product is offered by identical firms with constant unit cost in a perfectly competitive environment. This pins down the wage at a unit level for every period. In the differentiated sector, there is free entry to the market, and an infinite supply of potential entrants. Entry is costly and takes place every period. Firms incur a sunk cost of s upon entering the market, then draw a marginal cost c_t (equal to a unit labor requirement) from an exogenous cost distribution $G(\cdot)$. Thereafter, every period, c_t^i evolves according to a random process described below. Once firms know their cost draw, they must

incur a common fixed continuation cost of λ to be able to produce. If they do not, they exit the market at no cost and join the pool of potential entrants from which they are indistinguishable. Surviving firms maximize their profits using the residual demand function (2.6). Firms compete according to monopolistic competition. They take market level data $(\tilde{\Omega}_t, \bar{p}_t, N_t)$ as given.

Productivity evolution Each period after they have entered the market, incumbents learn their marginal cost for the current period. It is drawn from the following Markov process $F(c_t|c_{t-1})$:

$$c_t^i = \begin{cases} c_{t-1}^i & \text{with probability } \alpha \\ \sim G(.) & \text{with probability } 1 - \alpha. \end{cases} \quad (2.8)$$

This formulation is a stylized way of getting at a productivity process that exhibits both persistence and random fluctuations. This simple formulation allows us to derive closed-form solutions and simple comparative statics.

2.3.2 Steady-state Industry Equilibrium

Every period, firm i chooses its profit-maximizing price and quantity given the demand structure and market-level variables. The level of the choke price governs firms' markups m_t^i , as profit-maximizing markups are

$$m_t^i(c_t^i; A_t) = \frac{1}{2}(A_t - c_t^i). \quad (2.9)$$

If firm i decides to produce in the current period t , its operating profit is

$$L\pi(c_t^i; A_t) = \begin{cases} \frac{L}{4\gamma}(A_t - c_t^i)^2 & \text{if } c_t^i \leq A_t \\ 0 & \text{if } c_t^i > A_t, \end{cases} \quad (2.10)$$

where A_t is the choke price defined in (2.7), and $\pi(c; A)$ is the per-consumer, per-period profit of a cost c firm. Note that the choke price A_t is endogenously determined

in the model in general equilibrium and can vary with market characteristics, typically market size. Therefore, a firm's operating profit depends on market size not only via a direct multiplicative effect through L , but also via the choke price $A_t(L)$.

I focus on the steady-state equilibrium of the industry. In the steady state, the productivity distribution for the industry is stationary, and the industry-level variables A , N and $\bar{p}$ are constant. Since a firm can choose to exit at no cost, or stay in the market and produce, the value function of a firm with current cost c is:

$$V(c; A) = \max \left\{ 0, L\pi(c; A) - \lambda + \rho \left(\alpha V(c; A) + (1 - \alpha) \int V(c; A) dG(c) \right) \right\}, \quad (2.11)$$

where ρ is the firm's discount factor. $V(c)$ is non-increasing and continuous in c , as $\pi(c; A)$ is.³ Firms decide to stay in the market as long as their value is positive, i.e. as long as their cost is under a threshold B , defined as

$$B = \sup_{c \in (0, \infty)} \{c \mid V(c; A) > 0\}. \quad (2.12)$$

The free entry condition is such that new entrants make zero profit in expectation:

$$\int_0^B V(c; A) dG(c) = s. \quad (2.13)$$

Note that this expression also pins down the expected value of an incumbent firm for next period conditional on redrawing a fresh cost draw from $G(\cdot)$. Therefore, combining equations (2.11) and (2.13) leads to the following closed form expression

³Since $\pi(c; A)$ is non-increasing and continuous in c , and next period distribution of costs is such that $F(c'|c_1) \geq F(c'|c_2)$ if $c_1 < c_2$, $V(c)$ is also non-increasing and continuous in c (see for example Dixit and Pindyck (1994))

for firms' value function

$$V(c) = \begin{cases} \max \left\{ 0, \bar{V}(c; A) \right\} & \text{if } c < A \\ \max \left\{ 0, \frac{-\lambda + \rho(1 - \alpha)s}{1 - \rho\alpha} \right\} & \text{if } c \geq A, \end{cases}$$

where

$$\bar{V}(c; A) = \frac{\frac{L}{4\gamma}(A - c)^2 - \lambda + \rho(1 - \alpha)s}{1 - \rho\alpha}$$

Note that if the continuation costs are small enough that $\lambda < \rho(1 - \alpha)s$, firms of all costs have strictly positive values. No firm ever exits the market. This is because the option value of staying in the market is large enough for firms of any cost level to stay in the market without producing, and wait for a better cost draw.

I therefore restrict the analysis to the empirically relevant case where there is some exit in the steady state, i.e. I assume that $\lambda \geq \rho(1 - \alpha)s$. The value function further simplifies to:

$$V(c; A) = \begin{cases} \bar{V}(c; A) & \text{if } c < B \\ 0 & \text{if } c \geq B, \end{cases}$$

The maximum cost level under which firm stay in the market is the smallest positive root of $\bar{V}(c; A)$. Finally, the choke price is the only A such that the free entry condition (2.13) holds.⁴ The exit cutoff B is below the choke price A and can be expressed as a simple function of A ,

$$B = A - \sqrt{\frac{4\gamma(\lambda - \rho(1 - \alpha))s}{L}}.$$

In this economy, the general equilibrium quantity A summarizes competition forces and drives firm's markups and operating profits. Contrary to the case in Melitz and

⁴ The left-hand side of equation (2.13) is continuous and non-decreasing in A – as both B and $V(c; A)$ are. It is negative at $A = 0$ and goes to infinity with A . The intermediate value theorem ensures that there exists a unique A (and $B(A)$) such that the free entry condition is met.

Ottaviano (2008), it is distinct from the cost cutoff B , which is shaped in part by competition forces, but also by the parameter α governing the randomness in the firms' cost evolution.

Stationary distribution of firms The mass of firms N is determined by equations (2.7) and (2.9):

$$N = \gamma \frac{a - A}{A - \bar{c}},$$

where $\bar{c} = \int_0^B c dG(c)$. A stationary equilibrium with positive mass of firms exists when $a - A > 0$. It is straightforward to see that the stationary cost distribution is $G(c)/G(B)$ for $c \leq B$ and $G(c) = 1$ for $c > B$, i.e. a truncated version of the new entrant distribution $G(\cdot)$ where the truncation occurs at the cost cutoff B .

2.3.3 Comparative statics

I analyze how the steady-state firm distribution and firm dynamics are impacted by two main characteristics of the market, namely its size L and its level of persistence α . As shown in section 2.2, firms in larger markets are characterized by a more volatile productivity process. Therefore, industry dynamics in larger markets are shaped by (1) competition forces driven by market size L and (2) the level of volatility in firm dynamics driven by α .

The effect of market size and persistence

Exit threshold I first examine the impact of market size and cost persistence on the exit threshold, a key determinant of firm dynamics and of the cross-sectional distribution of firms' productivities.

Proposition 8 *a) As market size increases, all else being equal, the exit cost threshold decreases. Namely, if $L_1 > L_0$, then $B_0 > B_1$.*

b) As the persistence of the cost process decreases, all else being equal, the exit cost threshold increases. Namely, if $\alpha_1 < \alpha_0$, then $B_0 < B_1$.

The formal proofs are detailed in the Appendix. When market size increases – holding the persistence level constant –, it has two effects on profits. First, all firms benefit from increased demand. Operating profits $L\pi(c; A_0)$ increase. Therefore, firms' value function $V(c; A_0)$ increases for all c , including for the marginal firm with cost B_0 that now makes positive profits. Then, for the free entry condition (2.13) to keep on holding, the choke price A has to adjust downwards: a larger market size leads to an increase in the competitive pressure faced by firms. Overall, given the form of the profit function, firms that have higher cost benefit less (or suffer more) from a larger, more competitive market. The exit threshold B is lower in larger markets. When competition is fiercer, selection tougher, as is the case in the static model.

A decrease in the persistence parameter α leads to a higher option value of waiting for high cost firms, which tends to loosen the exit threshold. When the persistence of the productivity process decreases, low-cost firms loose: they face a higher risk of loosing their current better-than-average cost draw. On the other hand, inefficient firms are willing to stay around more (i.e. at worse cost levels), because they have a better option value through a greater chance of a fresh draw. In particular, the firm that was marginal between producing and exiting stays active in a world with lower persistence. This increases the minimum cost threshold above which firms exit. In addition, decreasing persistence may have an indirect effect, through A , on profits. The toughness of competition may have to adjust to preserve the free entry condition. Depending on parameter values, and in particular on the distribution $G(\cdot)$, it could be that A has to adjust down – if the persistence effect increases firms' profits on average –, or up in the opposite case. The impact of this indirect effect on the exit threshold is dominated by the direct effect, as the formal proof shows. Therefore, the exit threshold B is higher in markets with less persistent productivity process.

Taken together, these two effects give rise to an ambiguous aggregate impact on the exit cutoff of firms. If larger markets are also more volatile, as the data suggests, the cost cutoff is pushed down in larger markets by competition forces, but also pushed up because of the higher option value of staying around. I now examine how survival rates and firm volatility are impacted by a change in market size L or in persistence parameter α .

Survival rate It is straightforward to see that the probability of an active firm surviving one more year is

$$\theta_1 = \alpha + (1 - \alpha)G(B). \quad (2.14)$$

An active firm survives if it keeps its current cost level, or if it gets a fresh cost draw for which the cost is below the threshold B . The probability to survive more than n years is $\theta_n = (\theta_1)^n$.

Corollary 9 *If $L_1 > L_0$, all else equal, $\theta_1 < \theta_0$. If $\alpha_1 < \alpha_0$, all else equal, the impact on θ is ambiguous.*

In general, the impact of a change in α on θ is ambiguous : as α goes down, the probability of a firm keeping its current (good) cost draw decreases. On the other hand, if it gets a fresh cost draw, the probability of it being good enough to survive is higher, since the threshold B is higher by proposition 8. In contrast, an increase in market size unambiguously reduces survival rates. This comes from the fact that the cost threshold B is lower, holding α constant. This explanation of a lower survival rate in larger cities is therefore inconsistent with Fact 1.

Volatility I relate here the parameter α , that governs the randomness in the cost process, with firm-level volatility. The volatility of firm-level idiosyncratic shocks,

conditional on survival, is directly related to the persistence parameter α . Write $\sigma^2 = \text{var}(\frac{c_{t+1}-c_t}{c_t} | c_t, c_{t+1} > B)$ the volatility of firm's costs, conditional on observing a firm two periods in a row.

Corollary 10 *Assume that $G(\cdot)$ is log-concave⁵. If $L_1 > L_0$, all else equal, $\sigma_1 < \sigma_0$. If $\alpha_1 < \alpha_0$, all else equal, $\sigma_1 > \sigma_0$.*

The volatility of costs is given by

$$\sigma^2 = \text{var}(c_{t+1} | c_t, c_{t+1} > B) = (1 - \alpha)^2 \text{var}(G(c) | c < B). \quad (2.15)$$

It is closely related to the variance of the cross-sectional distributions of marginal costs, given the random process at play. In larger markets, for a given persistence parameter, the distribution of costs has a narrower support, hence a lower variance. In markets with less persistence in the cost process, the distribution of costs has a wider support and there is more chance of redrawing a new cost shock. Overall, firm volatility is driven *down* by the competition effect, whereas it is driven *up* by a decrease in α . Only the latter is consistent with the empirical regularity that firm-level volatility is higher in larger markets, as documented in section 2.2.

Comparative static under constant threshold

When larger markets also exhibit higher firm-level volatility as the data suggest (see section 2.2), the overall impact of market size on exit threshold is ambiguous : fiercer competition in bigger markets leads to a lower cost threshold, but better option value there pushes the threshold up.

This model, therefore, can be consistent with the finding of Combes, Duranton, Gobillon, Puga, and Roux (2012) that there is no differential exit threshold between

⁵This is true eg if $G(\cdot)$ is lognormal, which is what Combes, Duranton, Gobillon, Puga, and Roux (2012) find in the data

large and small cities. In what follows, I re-examine the comparative statics presented above under the assumption that this is indeed true: the two effects compensate each other, so that large and small markets have the same exit threshold.

Assumption A *Large and small markets exhibit the same exit threshold.*

Under the assumption that B is constant across markets of different sizes, I now examine how survival rates of firms and the volatility of firms' costs is impacted by market size.

The impact of market size on survival rates is examined in Asplund and Nocke (2006), who show empirically that firm turnover is higher in larger markets. They propose a theoretical explanation based on the fact that larger markets have lower cost threshold for exit, akin to Corollary 9 above. This explanation is at odds, though, with the findings of Combes, Duranton, Gobillon, Puga, and Roux (2012). The findings below offer a possible explanation to reconcile the two observations. Under assumption A that the threshold is overall stable between large and small cities, the impact of α on survival rates is unambiguous.

Corollary 11 *If α decreases and L increases such that assumption A holds, the probability θ_n of surviving n additional years or more, conditional on being active on the market, decreases. Survival rates are lower.*

As the productivity process is subject to more volatility, the probability of surviving decreases for all incumbent firms at all time horizons, even absent a difference in exit thresholds in large versus small markets. This can be easily seen from expression (2.14). Finally, conditional on observing a firm two years in a row, the variance of a firm's year $t+1$ productivity knowing year t 's will be higher in a larger market with lower persistence α .

Corollary 12 *If α decrease and L increase such that assumption A holds, the volatility of the cost process σ increases.*

This is readily seen from equation (2.15). The model therefore offers a possible explanation for the stylized facts highlighted in section 2.2: minimum cost cutoffs are not related to market size, but turnover rates are higher in larger markets and productivities are more volatile there.

2.4 A general model of industry dynamics

In this section, I study a more general and realistic random process for productivity. Instead of either keeping their own current cost or redrawing a fresh draw uncorrelated from the precedent, as I assumed above, firms' marginal cost follows a geometric brownian motion, as is assumed for example in Luttmer (2007) or Arkolakis (2011).

2.4.1 General set-up

Demand Each consumer has preferences over a consumption stream $\{U_t\}$ of goods from which she derives utility according to:

$$\int U_t e^{-rt} dt,$$

where future consumption is discounted at rate r . The composite good is defined as in equation (3.3):

$$U_t = q_{o,t} + a \int q_t(\omega) d\omega - \frac{1}{2} \gamma \int q_t(\omega)^2 d\omega - \frac{1}{2} \eta \left(\int q_t(\omega) d\omega \right)^2.$$

Aggregate demand is characterized by an instantaneous choke price A_t (see equation (2.7)), above which the demand for the differentiated good is zero.

Production Production follows the same logic as in section 2.3, except that entry, exit and production are instantaneous. Furthermore, the stochastic process that governs the evolution of firms' marginal costs is different.

There is free entry in the market. At each point in time, an infinite supply of potential entrants can pay a sunk entry cost s and enter the market. Upon entry, they draw a initial marginal cost level from a distribution $G(\cdot)$. Once they have entered, firms must incur an instantaneous fixed cost λ to produce with their current marginal cost c_t^i . They price according to monopolistic competition. Finally, they can exit at any time at no cost. If they do so, they join the pool of potential entrants from which they are indistinguishable.

Cost process Firm's marginal cost follows a geometric Brownian motion

$$\frac{dc_t}{c_t} = \mu dt + \sigma dz, \quad (2.16)$$

where dz is a brownian motion with independent increments, the drift parameter μ governs the rate at which a firm's costs improves on average and σ governs the volatility of the cost improvements.

2.4.2 Value function

Given the form of demand, firm's instantaneous operating profits are

$$L\pi(c_t^i; A_t) = \begin{cases} \frac{L}{4\gamma}(A_t - c_t^i)^2 & \text{if } c_t^i \leq A_t \\ 0 & \text{if } c_t^i > A_t. \end{cases} \quad (2.17)$$

I study the necessary conditions that prevail in the stationary equilibrium. The aggregate firm-productivity distribution is stationary at its steady-state level and

market-wide aggregate quantities such as $A_t = A$ are also constant. Because the stochastic process is Markovian, we can relabel time using $t = 0$ for the current time.

Consider a firm that has already entered the market. At each point in time, the firm can either stay or make the irreversible decision to exit the market at no exit cost and get a zero value. Writing τ the optimal exit time for the firm, the value of a firm with current cost c is therefore:

$$V(c; A) = \sup_{\tau > 0} E \left\{ \int_0^{\tau} (L\pi(c_t; A) - \lambda) e^{-\rho t} dt \mid c_0 = c \right\}, \quad (2.18)$$

where ρ is the firms' discount rate. The problem can be solved using classical tools described for example in Dixit and Pindyck (1994). Firm's operating profits are decreasing in marginal costs. Because of the classical form of the problem, we conjecture that the value function will be strictly positive up to a threshold B , above which all firms exit and firms' value is therefore constant, equal to zero.

In $[0, B]$, the Bellman equation that $V(\cdot; A)$ must satisfy is

$$-\rho V(c; A) + \mu c V'(c; A) + \frac{1}{2} \sigma^2 c^2 V''(c; A) + L\pi(c; A) - \lambda = 0, \quad (2.19)$$

where the derivatives are taken with respect to c . Furthermore, for all $c \geq B$,

$$-\rho V(c) + \mu c V'(c) + \frac{1}{2} \sigma^2 c^2 V''(c) + \pi(c) - \lambda \leq 0.$$

At the threshold $c = B$, two boundary conditions must hold. First, by definition of B ,

$$V(B) = 0. \quad (2.20)$$

Second, the smooth pasting condition dictates that

$$V'(B) = 0. \quad (2.21)$$

I make the additional assumption that the instantaneous continuation cost λ is high enough so that a firm that makes zero operating profits exits the market. This ensures that $B \leq A$.

The homogenous solutions to (2.19) are of the form $K_1 c^{-\alpha} + K_2 c^\beta$, where $\alpha > 0$, $\beta > 1$ and

$$\beta = \frac{\sigma^2 - 2\mu + \sqrt{(\sigma^2 - 2\mu)^2 + 8\rho\sigma^2}}{2\sigma^2}. \quad (2.22)$$

A particular solution to (2.19) can be found by guessing a second order function in c that I write $ac^2 + bc + d$. Finally, in the vicinity of $c = 0$, the current operating profit of firms is finite, and the value option of exiting must be close to zero, since there is a very low probability that the firm will have high costs in the future. Therefore, the value function cannot go to infinity at $c = 0$, and $K_1 = 0$.

The value function is therefore of the form

$$V(c) = ac^2 + bc + d + Kc^\beta,$$

with $a = \frac{L/4\gamma}{\rho - 2\mu - \sigma^2}$, $b = \frac{LA/2\gamma}{\mu - \rho}$ and $d = \frac{LA^2/4\gamma - \lambda}{\rho}$.

Finally, the boundary conditions (2.20) and (2.21) jointly pin down the values of K and of the exit threshold B . This yields

$$K = -\frac{aB^2 + bB + d}{B^\beta}, \quad (2.23)$$

with B the smallest root of the quadratic function $\phi(x)$

$$\phi(x) = x^2 + \frac{2(1 - \beta)A(\rho - 2\mu - \sigma^2)}{(\mu - \rho)(2 - \beta)}x - \frac{\beta(LA^2 - \lambda 4\gamma)(\rho - 2\mu - \sigma^2)}{\rho L(2 - \beta)}.$$

Finally, the equilibrium choke price A is pinned down by the free entry condition

$$V_e = \int_0^B V(c'; A) dG(c') - s = 0. \quad (2.24)$$

This equation implicitly defines A . $V(c; A)$ is increasing in A for all c , therefore the threshold B will also be increasing in A (all else equal). The whole expression for V_e is therefore increasing in A as well. It varies from $-s$ for $A = 0$ to infinity as A goes to infinity. Because of the continuity of V , there exists a unique choke price A that characterizes the steady state equilibrium.

Proposition 13 *Larger markets have lower choke price in equilibrium: $\frac{\partial A}{\partial L} < 0$. Markets with higher volatility also have lower choke price in equilibrium: $\frac{\partial A}{\partial \sigma} < 0$.*

The market-wide choke price A adjusts in equilibrium to ensure that the free entry condition (2.24) holds in every market. The direct effect of increasing market size is to increase firm's profits for all active firms, as operating profits are proportional to market size. In turn, it also increases the value functions of all active firms, as well as the expected stream of profits for any new entrant. To ensure that the free condition holds, it must be that A is reduced in larger markets, so that markups and profits decline through this indirect effect.

Because of the convexity of firm's current profits, higher volatility increases the value of all firms. Again, A has to adjust downwards for the free condition to keep on holding.

Following proposition 13, markups unambiguously go down in larger markets compared to smaller markets, when larger markets have both higher aggregate demand and higher volatility.

To go further and study how the exit threshold B changes with market size and market volatility, I now illustrate these comparative statics numerically at calibrated parameter values.

2.4.3 Numerical simulations

I calibrate the model using the following parameters. The values for μ and σ that govern the productivity process are taken from Arkolakis (2011). Noting that the cost process (2.16) is equivalent to

$$d\theta_t = (-\mu + \frac{\sigma^2}{2})dt + \sigma dz, \quad (2.25)$$

where θ_t is the log-productivity of the process, the parameters used in Arkolakis (2011) correspond to $\mu = -1.46\%$ and $\sigma = 6.68\%$. Firms' discount rate ρ is 4% as is standard in the literature. I normalize market size to $L = 1$. The ratio of entry cost to fixed costs is fixed at 5, to reflect the fact that sunk costs of entry are large compared to recurring fixed costs of production. The remaining parameters to be calibrated are the fixed costs λ , as well as the preference parameter γ on which the literature provides no guidance. Since the model restricts parameters to be such that firms do not enter a 'mothball' state, i.e. do not stay on the market without producing, waiting for a better cost draw, this limits the value λ can take for a given value of γ . If fixed costs are low, firms have an incentive to stay on the market without producing as long as they are not too far from a cost level that would give them positive profits. This is excluded for now in my analysis. Therefore, I normalize the preference parameter $\gamma = 1$ and take $\lambda = 10$ as the baseline value for fixed cost. This ensures that firms will not enter a 'mothball' state at the equilibrium of the industry.

I first consider how the exit threshold B changes, in general equilibrium, when L increases or σ increases. I find numerically that the exit threshold decreases systematically with market size, and that it increases systematically with volatility. The exit threshold behaves as in the stylized model. Figure 2.1 illustrates the behavior of value functions as market size (resp. volatility) increases. As market size increases (and markups decline through the feedback effect of a lower A), low-cost firms benefit and

high-cost firms suffer on net. The firm that was previously the marginal entrant exits in a larger market. Similarly, as volatility increases (and markups decline through the feedback effect of a lower A), the low cost firms benefit on net, but the high-cost firms loose. This is intuitive. Increasing volatility increases the option value of exiting at no cost. This is valuable to high cost firms, but barely for low cost firms who on average will draw high productivity levels in the future. Therefore, volatility barely impacts the value of low-cost firm, but the markup effect does. On the other hand, volatility is valuable for high-cost firms, and they are less harmed by the markup effect.

I finally study whether the ‘competition effect’ and the ‘volatility effect’ can quantitatively have canceling impacts on the exit threshold. To do so, I compare how the exit threshold moves as market size increases, holding σ constant ($\frac{\partial B(L, \sigma)}{\partial L}$), to what happens to the exit threshold when both market size *and* volatility increase. Volatility increases with market size at the rate recovered from the data ($\frac{dB(L, \sigma(L))}{dL}$).⁶ Figure 2.2 reports the results. The magnitude of the elasticity of B with respect to market size and volatility depends on parameter values. It depends in particular on the value of fixed costs (for a given preference parameter γ).

For low fixed costs (panel A), the volatility effect is strong, and makes up for about 2/3 of the competition effect. In the example reported in panel A, doubling market size decreases B by 3%. Once the volatility effect is accounted for, the decline in B is only 1%.

Panel B illustrates that as fixed costs get higher, the ‘volatility effect’ becomes weaker. Fixed costs are three times higher than they are in panel A: $\lambda_2 = 3\lambda_1$. Doubling market size decreases B by 5%. Once the volatility effect is accounted for, the decline in B is 3%. These simulations never lead to a full cancellation of the

⁶Estimates reported in Table 2.7 correspond to an elasticity of volatility with respect to city size of $\frac{\partial \log \sigma}{\partial \log L} = 0.12$.

two effects, though the volatility effect greatly dampens the change in cost threshold when fixed costs are low.⁷

Using these simulations, I assess whether these differences in threshold could be compatible with the quantitative findings of Combes, Duranton, Gobillon, Puga, and Roux (2012). They find small and insignificant values for the parameter that detects a difference in threshold between two markets, with a preferred estimate of $\hat{S} = 0.001$ with a standard error of 0.001. This parameter S doesn't measure the difference in thresholds directly, but the share of firms that were present in the small market but are absent of the more competitive market as the productivity threshold increases (or equivalently, since I use marginal costs here, when the cost threshold decreases). In the calibrations I present above, the difference in cost threshold range from 1% to 5% as market size doubles. The corresponding impact in terms of the share of firms that are pushed out of the market depends on the distribution of firms. Intuitively, if there is a large mass of firms near the threshold, this impact can be sizable. On the other hand, if the mass of firms near the threshold is relatively low, this impact is likely to be small.

To give an order of magnitude for this impact, I use the typical distribution of firms' productivities estimated in Combes, Duranton, Gobillon, Puga, and Roux (2012). They show that the distribution of TFP of French firms is well approximated by a lognormal. More precisely, the distribution is a mixture of lognormal and Pareto, with respective weights of .95 and .05. Figure 2.3 plots the corresponding probability distribution function at their estimated parameters for the distribution. I use their parameter estimate for the distribution of productivities, and compute what fraction of firm would be pushed out of the market as the productivity threshold increases

⁷What happens as fixed costs get even lower is a topic of ongoing research. In this region of the parameter space, firms with cost close to – but higher than – the choke price A enter a mothball state. They stay on the market, maintaining their presence by paying the fixed cost, without producing. In this situation, the volatility effect is likely to be strong and could perhaps revert the competition effect.

in the proportions found in the calibration. The table below reports the fraction of firms that are pushed out of the market by an increase in threshold of 1%, 3% and 5%, depending on the value of the productivity threshold in the reference market.

The share of firms that exit is higher if the distribution in the reference market is already fairly truncated, since more firms are located around the threshold there. Most of the values are consistent with the quantitative findings of Combes, Duranton, Gobillon, Puga, and Roux (2012) of $\hat{S} = 0.001$ with standard error 0.001. Note that in contrast, if the distribution of firms was instead assumed to be Pareto, increasing the threshold by 5% (resp. 1%) would lead to a a sizable share of firms exiting the market, of 9.3% (resp 2%), which would be one to two orders of magnitude larger than the estimated parameter $\hat{S} = 0.1\%$.⁸ If the distribution of firms' productivities is often approximated by a Pareto in the literature, this approximation is motivated by the typical behavior of most productive firms, i.e. the right tail of the distribution. It does poorly in approximating low productivity firms, i.e. the left tail of the distribution which is precisely of interest here.

These simple computations show that the findings of Combes, Duranton, Gobillon, Puga, and Roux (2012) can be consistent with a model where competition is indeed 'tougher' in larger markets. Taking into accounts the dynamics of firm productivity, the net threshold effects of increasing market size (and firm-level volatility) is within the margin of error of the procedure of Combes, Duranton, Gobillon, Puga, and Roux (2012) for a range of parameter values. Going further than this possibility result requires a structural estimation of the model, which is the topic of ongoing research.

⁸I use the parameter estimates of the Pareto part of the mixture reported in Combes, Duranton, Gobillon, Puga, and Roux (2012) to compute these numbers.

2.5 Conclusion

This paper develops a stochastic firm dynamics framework, where markets differ endogenously in the ‘toughness’ of their competitive forces. In this dynamic context, the minimum productivity threshold required for firm entry depends both on competition forces, captured by market size, and on the volatility of the stochastic process that governs firms’ productivities. The minimum productivity threshold is not a sufficient statistic for the toughness of competition, in contrast to what a static analysis would suggest. Because larger markets are characterized by a higher volatility of firms’ productivities, a fact I document from the data, the way the exit threshold moves with market size is ambiguous. The model proposes a possible explanation for two empirical regularities - productivity thresholds for entry do not seem to be systematically related to market size, but turnover rates are higher in larger markets - that are at odds with existing static or dynamic industry models. Understanding the mechanism that drives the higher volatility of firms’ productivities in larger markets is an interesting question that is left for future research.

2.6 Tables and figures

Table 2.1: Summary statistics

	mean	sd	p50	p10	p90	count
Employment	23.00	246.47	7.00	2.00	35.00	2912450
Capital	25994.10	1425674.15	1287.00	171.00	10200.13	2912450
log(TFP) (OLS)	4.52	0.69	4.52	3.74	5.34	2831256
log(TFP) (LP)	4.19	0.75	4.13	3.39	5.20	2817541
Number of establishments	1.50	6.67	1.00	1.00	2.00	2912450

Source: Panel of French Firms over 1993-2006. *Note:* p10 and p90 are respectively the 10th and 90th percentile of the distribution.

Table 2.2: Summary statistics, by size of market

	mean	sd	p50	p10	p90	count
Panel A - Market Size Q1						
Employment	15.47	32.48	7.00	2.00	34.00	523174
Capital	5664.11	28667.25	1410.31	213.00	9632.00	523174
log(TFP) (OLS)	4.42	0.60	4.44	3.74	5.10	512307
log(TFP) (LP)	4.12	0.72	4.05	3.34	5.14	510296
Panel B - Market Size Q2						
Employment	14.20	35.09	7.00	2.00	30.00	537150
Capital	5618.70	47961.13	1266.00	183.67	8372.00	537150
log(TFP) (OLS)	4.47	0.64	4.50	3.74	5.21	525004
log(TFP) (LP)	4.15	0.73	4.09	3.38	5.17	522587
Panel C - Market Size Q3						
Employment	12.30	32.43	6.00	2.00	25.00	518894
Capital	6091.55	119328.16	1062.65	146.00	6789.15	518894
log(TFP) (OLS)	4.51	0.69	4.54	3.72	5.33	505410
log(TFP) (LP)	4.19	0.73	4.14	3.40	5.17	502833
Panel D - Market Size Q4						
Employment	11.54	47.64	6.00	2.00	22.00	752312
Capital	27979.37	1622662.80	905.22	113.00	6738.00	752312
log(TFP) (OLS)	4.65	0.79	4.67	3.74	5.60	723975
log(TFP) (LP)	4.24	0.78	4.21	3.40	5.24	719531

Source: Panel of French Firms over 1993-2006. Market Size Q1- Q4 corresponds to the quartile of employment density of a market (ie commuting zone). MSize Q1 corresponds to the smallest markets, and MSize Q4 to the largest. *Note:* p10 and p90 are respectively the 10th and 90th percentile of the distribution.

Table 2.3: Comparing TFP distribution between large and small markets

	Δ Threshold versus Δ Shift		
	(S)	(A)	(D)
Manufacturing	0.000	0.099	1.182
Retail and Services	0.000	0.054	1.145

Note: (S) estimates the difference in threshold between large (above median of employment density) and small (below median) market. (A) measures the shift of the overall TFP distribution between the 2, and (D) measure the dilation of the large market distribution versus the small market distribution. The procedure used is the one developed in Combes, Duranton, Gobillon, Puga, and Roux (2012).

Table 2.4: Survival rates and market size

	Probability of survival			Probability of
	5 years	10 years	Bankruptcy 5 yrs	
	(1)	(2)	(3)	(4)
MSize	-.0077*** (.00014)	-.0072*** (.00014)	-.013*** (.00031)	.0018*** (.00014)
Log TFP		.086*** (.00055)	.12*** (.0013)	-.065*** (.00048)
Log(Capital)		.032*** (.00027)	.022*** (.00054)	-.018*** (.00023)
Log(Empl)		.02*** (.00046)	.0037*** (.00089)	.013*** (.00041)
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
Observations	1497718	1458128	608137	812377

Source: Panel of French Firms over 1993-2006. *Note:* This table presents the marginals from the probit regression 2.2. The dependent variable is a dummy equal to 1 if the firm still operates 5 years after its creation, for (1)-(2), still operate 10 years after its creation for (3), and is bankrupt before 5 years for (4). Robust standard errors are in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.5: AR(1) fit for log-TFP

Measure of TFP	Log TFP _t					
	OLS (1)	OLS (2)	OLS (3)	LP (4)	LP (5)	LP (6)
Log TFP _{t-1} (OLS)	0.850*** (0.001)	0.732*** (0.001)	0.732*** (0.001)			
Log TFP _{t-1} (LP)				0.888*** (0.001)	0.749*** (0.001)	0.747*** (0.001)
Std error of residuals	0.383	0.372	0.372	0.380	0.368	0.368
Year FE	No	Yes	Yes	No	Yes	Yes
Sector FE	No	Yes	Yes	No	Yes	Yes
R-squared	0.675	0.692	0.692	0.733	0.750	0.745
Observations	1795150	1795150	1795150	2285363	2285363	1810236

Note: Estimates of regression (2.3). (1), (2), (4) and (5) are over the full panel of firms, (3) and (6) are restricted to single-establishment firms. Columns (1)-(3) correspond to an OLS measure of TFP, columns (4)-(6) to the LP measure. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6: TFP volatility and market size - OLS measure

	Log TFP _t (OLS measure)			
	Market Size 1 (smallest)	Market Size 2	Market Size 3	Market Size 4 (largest)
Log TFP _{t-1}	0.741*** (0.003)	0.735*** (0.003)	0.730*** (0.003)	0.718*** (0.002)
σ	0.331	0.347	0.367	0.420
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
R-squared	0.671	0.680	0.693	0.694
Observations	415499	421135	399868	558648
Test of equality for variance		.001***		
Test of $\sigma_1 = \sigma_2$ $Pr > F$		.001***		
Test of $\sigma_2 = \sigma_3$ $Pr > F$		.001***		
Test of $\sigma_3 = \sigma_4$ $Pr > F$		.001***		

Note: Results of regression (2.3) run for each market size. Markets size are defined by their quartile of employment density. σ is the standard error of the residuals. Tests of equality of variance are p-values of Levene test for equality of variance. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: TFP volatility and market size - LP measure

	Log TFP _t (LP measure)			
	Market Size 1 (smallest)	Market Size 2	Market Size 3	Market Size 4 (largest)
Log TFP _{t-1}	0.759*** (0.003)	0.751*** (0.003)	0.748*** (0.003)	0.738*** (0.002)
σ	0.327	0.344	0.363	0.411
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
R-squared	0.786	0.769	0.742	0.705
Observations	413241	418414	397030	553959

Note: Results of regression (2.3) run for each market size. Markets size are defined by their quartile of employment density. σ is the standard error of the residuals. Tests of equality of variance are p-values of Levene test for equality of variance. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8: TFP volatility and market size - robustness checks

	(1)	(2)	(3)
Panel A - Regression results			
	Log TFP _t		
Log TFP _{t-1}	0.732*** (0.002)	0.729*** (0.003)	0.700*** (0.010)
Log TFP _{t-1} * Q2_MSize	-0.002 (0.003)	-0.001 (0.004)	-0.010 (0.013)
Log TFP _{t-1} * Q3_MSize	0.001 (0.003)	-0.001 (0.004)	-0.002 (0.012)
Log TFP _{t-1} * Q4_MSize	0.003 (0.002)	0.003 (0.004)	0.001 (0.012)
Log employment	-0.023*** (0.000)	0.011*** (0.002)	0.025*** (0.007)
Log capital	0.010*** (0.000)	0.007*** (0.001)	0.010*** (0.001)
Year FE	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes
Market Size FE	Yes	Yes	Yes
R-squared	0.693	0.711	0.700
Observations	1795150	629876	104473
Panel B - Variance of the residuals			
σ_1	0.330	0.318	0.273
σ_2	0.347	0.326	0.285
σ_3	0.367	0.341	0.308
σ_4	0.420	0.386	0.375
Test of equality for variance	0.000***	0.000***	0.000***
Test of $\sigma_1 = \sigma_2$.	0.000***	0.001***	0.001***

Note: Estimates of regression (2.4), where TFP is computed with OLS. The results for TFP-LP, not reported, are very similar. (1) is over the full sample of single-establishments firms ; (2) is restricted to firm with 5 to 10 employees, (3) to firms with 30 to 50 employees. Indexes 1 to 4 on σ and *Msize* refer to the market size, as defined by the quartile of employment density. σ is the standard deviation of the residuals Tests of equality of variance are p-values of Levene test for equality of variance. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.9: Share of firms S that exit the market after a change in the threshold.

		(I)	(II)	(III)
Productivity threshold in the reference market		Percentage change in threshold		
		1%	3%	5%
0.1	(0.0%)	0.000	0.000	0.000
0.2	(0.3%)	0.000	0.000	0.001
0.3	(2.0%)	0.001	0.003	0.005
0.4	(6.0%)	0.002	0.007	0.012
0.5	(12.2%)	0.004	0.012	0.021

Notes. In the reference (small) market, the productivity follows the distribution reported in figure 2.3, truncated below the mode of the distribution at values ranging from 0.1 to 0.5. The percentage in parenthesis reports the share of firms that are below this truncation point. Columns (I)-(III) report the share of firms from this reference market that exit as the truncation point increases.

Figure 2.1: Value functions, market size and volatility.

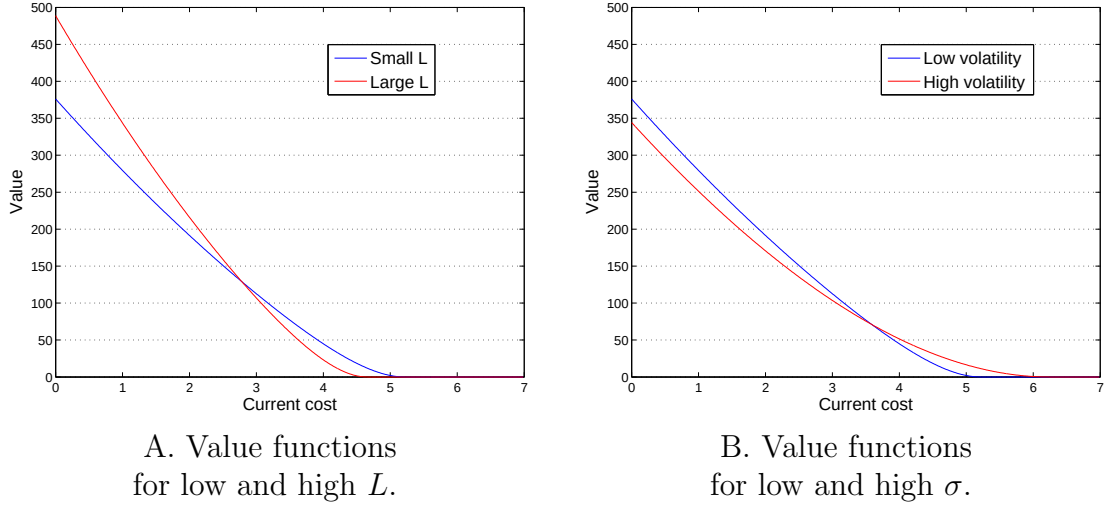
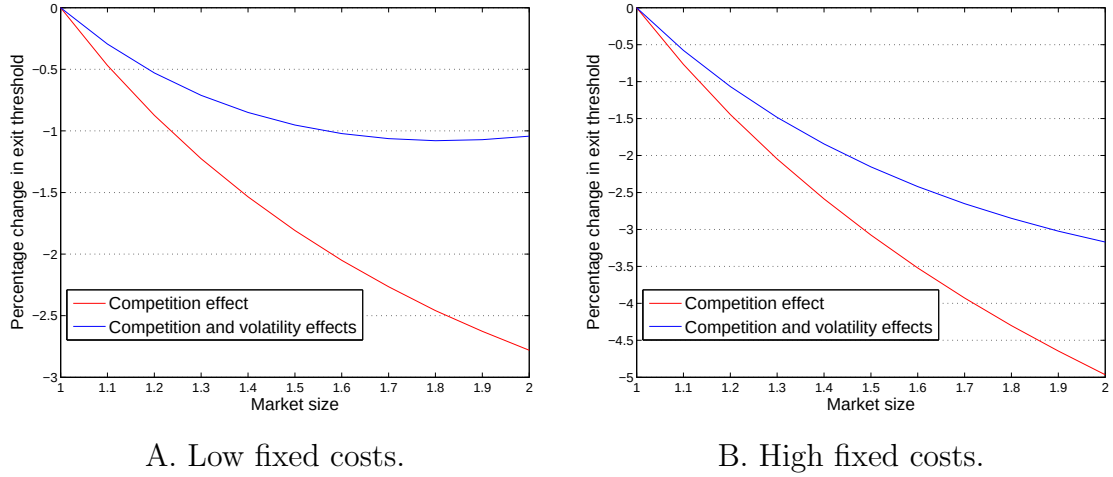
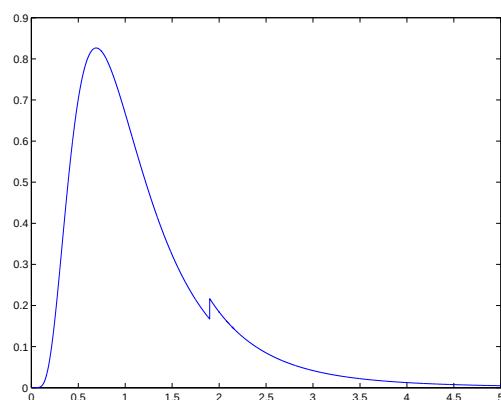


Figure 2.2: Exit threshold as a function of market size



In panel A, $\lambda = 10$, in panel B, $\lambda = 30$. In both cases, the blue line represents the case where volatility increases with L following $\frac{\partial \log \sigma}{\partial \log L} = 0.12$, the red line has market size increasing at constant volatility.

Figure 2.3: Approximated distribution of firms' TFP.



Notes. The distribution is a mixture of a lognormal distribution with mean $\hat{m} = -0.05$ and variance $\hat{v} = 0.32$, and a Pareto distribution with shape $\hat{z} = 1.89$ and minimum $\hat{b} = 1.90$, with respective weights 95% and 5%.

2.7 Appendix

2.7.1 Proofs for section 2.3

Proposition 1 Market size:

It is readily seen from the expression of $V(c, L, A)$ that $\frac{\partial V(c, L, A)}{\partial L} \geq 0$, where the inequality is strict for all $c \leq B$. For the free entry condition 2.13 to hold, it must be that either (1) A stays constant and the exit cutoff B decreases or (2) A decreases. The former is excluded, as for all $c < B$ $V(c, L_1, A_0) > V(c, L_0, A_0) > 0$ which contradicts $V(B_1, L_1, A_0) = 0$ that would have to hold if $A_1 = A_0$. Therefore, $A_1 < A_0$ and $A'(L) < 0$. As a corollary, it must be that not all firms see their operating profit increase in a larger market, otherwise the free entry condition would not be met. Therefore, by continuity of the profit function, there exists (at least) a $\underline{c} \in \min(B_0, B_1)$ such that $\frac{L_1}{4\gamma}(A_1 - \underline{c})^2 = \frac{L_0}{4\gamma}(A_0 - \underline{c})^2$. Furthermore, $\frac{d \log[\frac{L}{4\gamma}(A(L) - c)^2]}{dLdc} = \frac{2A'(L)}{(A(L) - c)^2} < 0$. The percentage increase in operating profits decreases with c . For $c > \underline{c}$, this percentage is smaller than one, so that $V(c, L_1, A_1) < V(c, L_0, A_0)$. Necessarily, $V(c, L_1, A_1)$ hits 0 first and $B_1 < B_0$.

Persistence:

The direct effect of α on value functions for $c < B$ is of the sign of :

$$\frac{\partial V}{\partial \alpha}(c; L; A; \alpha) \propto s(\rho - 1) + \frac{L}{4\gamma}(A - c)^2 - \lambda \quad (2.26)$$

$V(c; \alpha)$ exhibits decreasing differences in (c, α) : $\frac{\partial V}{\partial \alpha \partial c} < 0$. The direct effect of decreasing α is a rotation of $V(c)$ around c_s , from $V_0(c)$ to $\tilde{V}(c)$: low cost firms loose, as they have a decreased probability of keeping their 'better than average' cost draw. High cost firms gain, as they have a better chance of a fresh start. The firms that are not affected by a change in α are the one who are indifferent between their current draw and redrawing a cost, i.e. those for which $V(c) = s$, as s is the expected value a firm gets from redrawing. If A does not adjust in equilibrium, the direct effect of a decrease in α is an increase in B : since high cost firms gain, $V(B_0, \alpha_1, A_0) > V(B_0, \alpha_0, A_0) = 0$. Since $V(c)$ is a decreasing function, the minimum cost such that $V(c, \alpha_1, A_0) = 0$ is therefore $\tilde{B}_1$ for which $\tilde{B}_1 > B_0$.

Since some firms gain and other loose, the sign of $\int_0^{\tilde{B}_1} V(c; A_0) dG(c) - s$ is ambiguous - and in particular depends on the distribution $G(c)$. It could be positive if high cost firms are over-represented in the distribution. In that case, A_1 has to adjust downwards ($A_1 < A_0$) to preserve the free entry condition (2.13)). Conversely, if firms loose in expectation following the change in persistence, it must be that $A_1 > A_0$ to make up for the average loss in profits.

If $A_1 \geq A_0$, $V(c; \alpha_1, A_1) \geq V(c; \alpha_1, A_0)$ for $c \in (0, \tilde{B}_1)$. In particular, $V(\tilde{B}_1, \alpha_1, A_1) > V(\tilde{B}_1, \alpha_1, A_0) = 0$. Since $V(c)$ is a decreasing function, the minimum cost such that $V(c, \alpha_1, A_1) = 0$ is therefore B_1 for which $B_1 \geq \tilde{B}_1 > B_0$.

If $A_1 < A_0$, write:

$$\frac{dV}{d\alpha}(c, \alpha, A(\alpha)) = \frac{\partial V}{\partial \alpha}(c, \alpha, A(\alpha)) + \frac{dA}{d\alpha} \frac{\partial V}{\partial A}(c, \alpha, A(\alpha))$$

It is straightforward to check that $\frac{\partial}{\partial c} \left(\frac{dV}{d\alpha}(c, \alpha, A(\alpha)) \right) < 0$. This comes from $\frac{dA}{d\alpha} > 0$ in the case we study here, $\frac{\partial V}{\partial \alpha \partial c} < 0$ from above and $\frac{\partial V}{\partial A \partial c} = -\frac{2L}{\gamma(1-\rho\alpha)} < 0$. For the free entry condition to hold, it must be that $V(c, \alpha_1, A_1)$ does not lie strictly above or strictly below $V(c, \alpha_0, A_0)$: they cross at least once at $\underline{c}$. For $c > \underline{c}$ a decrease in α lead to positive gains: $V(c, \alpha_1, A_1) > V(c, \alpha_0, A_0)$. In particular, $V(B_0, \alpha_1, A_1) > V(B_0, \alpha_0, A_0) = 0$. Since $V(c)$ is a decreasing function, the minimum cost such that $V(c, \alpha_1, A_1) = 0$ is therefore B_1 for which $B_1 > B_0$.

Corollary 3

$$\sigma^2 = \text{var}(c_{t+1} | c_t, c_{t+1} > B) = (1 - \alpha)^2 \text{var}(G(c) | c < B)$$

Since B decreases when L increases, $\text{var}(G(c) | c < B)$ decreases as well as long as the distribution $G(\cdot)$ is log-concave. Decreasing α has the opposite effect and increases unambiguously σ , as it leads to more chance of redrawing a fresh draw, and to a higher entry/exit threshold B .

2.7.2 Proofs for section 2.4

Proposition 6 $V(c)$ is decreasing and convex in c , because π is decreasing and convex and the stochastic process that governs the evolution of c preserve these properties. Because of the convexity of $V(\cdot)$ a higher variance in the process leads to a higher value of $V(c)$ for all c . Call $C(X)$ the space of bounded and continuous functions on X . For $i \in \{0, 1\}$, define the operator T_i as follows:

$$T_i v(x) = \max\{0; \pi(c) - \lambda + \frac{1}{1+\rho} E_i(V(x'|x))\} \quad \text{for } v \in C(X),$$

where E_i denotes the expectation at time $t+dt$ corresponding to a volatility of the process of σ_i . Then, T_i maps $C(X)$ to itself, and for any two convex functions v and w ,

$$v \leq w \Rightarrow E_0(v) \leq E_0(w) \leq E_1(w) \Rightarrow T_0 v \leq T_1 w$$

It follows that the fixed point of T_0 lies below the fixed point of T_1 . Therefore, to keep the free entry condition (2.24) holding, A needs to decrease as σ increases: $A_1 < A_0$. ■

Chapter 3

Firm Sorting and Agglomeration

3.1 Introduction

The distribution of firms in space is far from uniform. Some locations host the most productive large firms, while others barely attract any. Recognizing the key role played by firms in the development of local activity, governments offer financial incentives to attract firms to specific areas. Such interventions aim, in general, to reduce spatial disparities and to increase welfare by fostering local agglomeration externalities. In practice, though, positive local impacts may be counterbalanced by undesirable effects in other regions. In addition, spatial disparities could be reinforced if subsidies attract low-productivity firms while leaving high-productivity ones in booming regions. To understand these effects, I propose a theory of the distribution of heterogeneous firms in a variety of sectors across cities. The model has a unique equilibrium, which is affected by the implementation of these place-based policies. This allows me to conduct a counterfactual policy analysis and to quantify their impact in terms of aggregate productivity and welfare.

Place-based policies are pervasive. Kline and Moretti (2013) report that an estimated 95 billion dollars are spent annually in the United States to attract firms to

certain locations. Massive subsidies are targeted to individual firms in a nationwide competition to attract plants.¹ Federal programs provide generous tax breaks in an effort to attract firms to less developed areas, as analyzed most recently by Busso, Gregory, and Kline (2013). The European Union regulates these aids and reports a yearly figure of 15 billion euros.² In a broader sense, place-based policies also encompass regulations that impact the growth of cities, such as zoning regulations or restrictions on the height of buildings. I analyze the impact of both types of policies in the light of my framework.

To analyze the long-run impact of these policies, one needs to understand how heterogeneous firms react to them and how targeted cities grow or shrink as a result. I develop a model that is uniquely suited to answering these questions. Heterogeneous firms sort across cities. Cities are endogenously impacted by this sorting. They grow when firms choose to locate there and increase the local labor demand. Since firms' location choices are distorted by local policies, cities are, in turn, endogenously affected. Importantly, the model admits a unique equilibrium, which is impacted by local policies. The model captures several forces. Cities are *ex ante* identical. They are the locus of intangible agglomeration externalities such as thick labor markets or knowledge spillovers (Duranton and Puga (2003)). The sorting of firms into cities of different sizes is driven by a trade-off between gains in productivity, through local externalities, and higher labor costs. I assume that more efficient firms benefit relatively more from local externalities. This generates positive assortative matching: more efficient firms locate in larger cities, reinforcing their initial edge. A class of city

¹Examples abound and are compiled in Story (2012). Recent examples include Apple who obtained a reported 88 million dollars in tax cuts to locate a data center in Reno, Nevada, Caterpillar with a reported 77 million dollars incentives to build a plant in Georgia, or Boeing getting a 120 million dollars subsidy from South Carolina to expand its operations in Charleston.

²In contrast to the United States, the European Union regulates the amount and geographical scope of these aids. They are restricted to *less advantaged* regions. Both in Europe and in the US, there exists a wider range of policies that are designed with spatial impact in mind but are not accounted for in these figures, such as the design of local tax reporting for multi-location firms. Bartik (2004) reports a taxonomy of such policies.

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 features of the French economy. First, 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. The resulting firm-size distribution is more thick-tailed for sectors that tend to locate in larger cities. Second, across sectors, firm sorting is shaped by the intensity of input use. In labor-intensive sectors, firms locate more in small cities where wages are lower. Third, within sectors, firms have higher revenues in larger cities. They may have lower employment, though, since they face higher labor costs. I show that these outcomes of the model are broadly consistent with stylized facts about French firms.

I structurally estimate the model. This structure is needed to quantify the impact of place-based policies on the spatial equilibrium, absent a real-life counterfactual. The structural estimation recovers a model-based estimate of the shape of agglomeration externalities. It allows me to disentangle the roles played by agglomeration forces on the one hand, and firm sorting on the other hand, in shaping productivity gains associated with city size. I estimate that nearly two thirds of the measured productivity advantage of large cities comes from the sorting of firms based on their efficiency. The magnitude of the agglomeration economies I estimate are in line with existing estimates in the literature as reported by Rosenthal and Strange (2004).

Finally, I simulate the new spatial equilibrium that results from two policies: 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, as advocated by Glaeser and Gottlieb (2008). Firms make different location choices as a result, which modifies local labor demand and in equilibrium impacts the city-size

distribution. The productive efficiency of the new equilibrium depends in particular on the new city-size distribution: it 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.

The paper is related to several strands of the literature. The main contribution of the paper is to propose a model of spatial equilibrium that is well suited to the analysis of place-based policies, since it features freely mobile, heterogeneous firms and endogenous city sizes, and has a unique equilibrium. The literature that studies systems of cities, pioneered by Henderson (1974), has traditionally focused on homogenous firms. Recent contributions have introduced richer heterogeneity in the spatial setting.³ In a seminal contribution, Behrens, Duranton, and Robert-Nicoud (2010) study the spatial sorting of entrepreneurs who produce non-tradable intermediates. I study the polar case of producers of perfectly tradable goods. An important difference for policy analysis is the uniqueness of the equilibrium I obtain here. Another closely related strand of the literature (Eeckhout, Pinheiro, and Schmidheiny (2010), 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 paper uses similar conceptual tools, borrowed from the assignment literature, to focus on how *firms* sort and impact local

³Early studies of heterogeneous firms in the spatial context include Nocke (2006) and Baldwin and Okubo (2006). They predict that more productive firms self-select in larger markets. These results are obtained in a setting where the size of regions is fixed and exogenously given. In contrast, I focus on how local policies endogenously change the spatial equilibrium, including city sizes themselves, as they attain their objective to see targeted cities grow.

labor demand, motivated by the fact that firms are directly targeted by place-based policies.⁴

Second, the model shows novel theoretical and empirical evidence of a link between the geographical pattern of economic activity and the sectoral firm-size distribution. It contributes to the literature that aims to explain the determinants of firm-size distribution by exploiting cross-sectoral heterogeneity. Rossi-Hansberg and Wright (2007a) focus on the role of decreasing returns to scale in determining the shape of the establishment-size distribution, while di Giovanni, Levchenko, and Rancière (2011) consider the role played by endogenous selection into exporting. I focus instead on the role played by location choice. The model also contributes to the literature on city-size distribution. City-size distribution is traditionally explained by random growth models, as in Gabaix (1999) and Eeckhout (2004), for example. Here, as in Behrens, Duranton, and Robert-Nicoud (2010), the distribution of city sizes is endogenous to the sorting of heterogeneous firms in a static spatial equilibrium, since cities are the collection of the labor force employed by local firms. The city-size distribution is shaped by properties of the firm-size distribution.

As in Desmet and Rossi-Hansberg (2013) and Behrens, Mion, Murata, and Südekum (2013), 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). Most analyses do not take into account the sorting of heterogeneous agents, workers or firms, across locations. A notable exception is Combes, Duranton, and Gobillon (2008), who use detailed data on

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

workers characteristics to control for worker heterogeneity and sorting in a reduced form analysis. I use a structural approach to explicitly account for the sorting of firms. Combes, Duranton, Gobillon, Puga, and Roux (2012) also emphasize the heterogeneous effect of agglomeration externalities on heterogeneous firms. They show that the most efficient firms are disproportionately more efficient in large cities, indicating potential complementarities between firm productivity and city size. Their analysis is based on a model without firm location choice, in contrast to the approach I take.

Finally, the counterfactual policy analysis offers a complementary approach to research that assesses the impact of specific place-based policies (see, for example, Busso, Gregory, and Kline (2013) for the US, Mayer, Mayneris, and Py (2012) for France and Criscuolo, Martin, Overman, and Reenen (2012) for the UK). The literature has traditionally focused on estimating the local effects of these policies. A notable exception is Kline and Moretti (2013), who develop a methodology to estimate their aggregate effects.⁵ For the policy they consider, they estimate that positive local effects are offset by losses in other parts of the country. My approach is similar in spirit, with the difference that I account explicitly for firm sorting. I find a negative aggregate effect of policies targeting the smallest cities. 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 2 presents the model and its predictions. Section 3 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

⁵Albouy (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.

estimate the model using indirect inference. In section 4, I conduct a counterfactual analysis using the estimated model. Section 5 concludes.

3.2 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, the local level of input prices and, possibly, the existence of local subsidies. Heterogeneous firms face different incentives, which yields heterogeneity in their choice. 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. The model describes a long-run steady state of the economy and abstracts from dynamics.

3.2.1 Set-up and agents' problem

Cities

Each potential city site has a given stock of land, normalized to 1. Sites are identical ex-ante. Cities emerge endogenously on these sites, potentially with different popula-

tion levels L . In what follows, I index cities and all the relevant city-level parameters by city size 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 tradable.

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

$$h = \gamma^b \left(\frac{\ell}{1-b} \right)^{1-b}. \quad (3.1)$$

The land-use intensity b governs the amount of housing that can be produced with a given stock of land. Lowering b increases the housing supply in a city where the stock of land is fixed. This parameter summarizes the technological and regulatory constraints on housing supply in a city. In the policy experiments, I analyze how it impacts productivity and welfare in the spatial equilibrium.

Landowners compete in the housing market, taking both the housing price $p_H(L)$ and the local wage $w(L)$ as given. Since there is a fixed total supply of land equal to 1 in the city, the housing supply equation is

$$H(L) = \left(\frac{p_H(L)}{w(L)} \right)^{\frac{1-b}{b}}, \quad (3.2)$$

where $H(L)$ is the total quantity of housing supplied in a city of size L .

Workers

Set-up There is a mass N_0 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

$$U = \left(\frac{c}{\eta}\right)^\eta \left(\frac{h}{1-\eta}\right)^{1-\eta}, \quad (3.3)$$

where h denotes housing and c is a Cobb-Douglas bundle of goods across S sectors of the economy defined as

$$c = \prod_{j=1}^{j=S} c_j^{\xi_j}, \quad \text{with} \quad \sum_{j=1}^{j=S} \xi_j = 1.$$

Within a sector j , consumers choose varieties according to the CES aggregator

$$c_j = \left[\int c_j(i)^{\frac{\sigma_j-1}{\sigma_j}} di \right]^{\frac{\sigma_j}{\sigma_j-1}}.$$

I denote by P 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 (3.3), under the budget constraint

$$P c(L) + p_H(L)h(L) = w(L).$$

As a result, the aggregate local demand for housing is

$$H(L) = \frac{(1-\eta)w(L)L}{p_H(L)}. \quad (3.4)$$

The local housing market clears. Equations (3.2) and (3.4) pin down prices and quantities of housing produced.⁶ The quantity of housing consumed by each worker in city L is

$$h(L) = (1 - \eta)^{1-b} L^{-b}. \quad (3.5)$$

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

$$w(L) = \bar{w}((1 - \eta) L)^{b \frac{1-\eta}{\eta}}, \quad (3.6)$$

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

Firms

Production The economy consists of S sectors that manufacture differentiated tradable products. Sectors are indexed by $j = 1, \dots, S$. Firms have the same factor intensities within their sectors but differ in productivity. 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

⁶Appendix 3.7.1 gives the equilibrium housing prices and labor hired.

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.

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 firms in sector j and $\psi(z, L, s_j)$ is a firm-specific Hicks-neutral productivity shifter, as detailed below. It is determined by the 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 (3.6)), 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 $\psi(z, L, s_j)$ depends on its own 'raw' efficiency z , on local agglomeration externalities that increase with

city size L , and on a sector-specific parameter s_j . I explain the roles of these parameters in turn.

A key assumption of the model is that ψ exhibits a strong complementarity between local externalities and the efficiency of the firm. Firms that are more efficient at producing are also more efficient at leveraging local agglomeration externalities, such as knowledge spillovers or labor-market pooling. The empirical literature offers some suggestive evidence pointing to such complementarity between firm efficiency and agglomeration externalities. Arzaghi and Henderson (2008) present an empirical study of the advertising industry. They measure the benefit agencies derive from interactions in local networks and provide evidence that agencies that were larger in the first place are more willing than smaller ones to pay higher rents in order to have access to a better local network. Combes, Duranton, Gobillon, Puga, and Roux (2012) study a wide set of French industries and provide suggestive evidence that more efficient firms are disproportionately more productive in larger cities, pointing to such a complementarity as a potential explanation for this fact. Finally, in Section 3.3, 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 Appendix 3.7.2, I extend this idea to the set-up of my model where workers are hired to produce the blueprint of the firm. The firm can have workers work at the plant or spend time improving the blueprint by leveraging local information, e.g., on production processes or product appeal. It is optimal for more efficient firms to

encourage workers to spend more time discovering local ideas. As a result, the productivity of the firm exhibits complementarity between city size and the firm's own efficiency.⁷

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 B $\psi(z, L, s)$ is log-supermodular in city size L , firm raw efficiency z and sectoral characteristic s , and is twice differentiable. In addition, $\psi(z, L, s)$ is strictly log-supermodular in (z, L) . That is,

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

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.⁸ 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

⁷The productivity function is log-supermodular in firm efficiency and city size.

⁸In addition, because $\frac{\partial^2 \log \psi(z, L, s)}{\partial z \partial s}$ is non-negative, for any city size L , the elasticity of firm productivity to its initial efficiency is either equal across sectors, or larger in high s sectors.

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, absent any local subsidies, for all firms. In other words, I assume that the positive effects of agglomeration externalities are not too strong compared to the congestion forces. In particular, given that the congestion forces increase with city size with a constant elasticity, agglomeration externalities must have decreasing elasticity to city size to prevent a degenerate outcome with complete agglomeration of firms in the largest city.

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 efficiency level z from a distribution $F_j(\cdot)$. I assume that this distribution is an open 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. Once firms discover their raw efficiency, they choose the size of the city where they want to produce.¹⁰

Firms' problem A firms' choice of city size is influenced by three factors. First, relative input prices vary by city size. Second, firm productivity increases with city size, through greater agglomeration externalities. Third, local city developers compete to attract firms to their cities by subsidizing profits at rate $T_j(L)$, which varies by city-size and sector.¹¹

⁹In 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 $\frac{\partial^2 \log \psi(z, L, s)}{\partial L \partial s} > 0$.

¹⁰Contrary to the setting in Melitz (2003b), 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.

¹¹The subsidy offered by city developers $T_j(L)$ may in principle differ across cities for cities of the same size. As I show in the next section, this is not the case in equilibrium. Anticipating this, I abstract from introducing cumbersome notation denoting different cities of the same size.

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), input prices, and subsidies. Then, firms choose location to maximize this optimized profit.

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 $\psi(z, L, s_j)$. This choice is not distorted by local subsidies to firms' profits. 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

$$\pi_j(z, L) = \kappa_{1j} (1 + T_j(L)) \left(\frac{\psi(z, L, s_j)}{w(L)^{1-\alpha_j}} \right)^{\sigma_j-1} R_j P_j^{\sigma_j-1}, \quad (3.7)$$

where P_j is the sectoral price index, R_j is the aggregate spending on goods from sector j and κ_{1j} is a sector-specific constant.¹²

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

$$\ell_j(z, L) = (1 - \alpha_j)(\sigma_j - 1) \frac{\pi_j(z, L)}{w(L)(1 + T_j(L))}. \quad (3.8)$$

The proportionality between profits - net of subsidies - 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 (3.7).

¹²The sectoral constant κ_{1j} is $\kappa_{1j} = \frac{((\sigma_j-1)\alpha_j^{\alpha_j}(1-\alpha_j)^{1-\alpha_j}\rho^{-\alpha_j})^{\sigma_j-1}}{\sigma_j}$

City developers

Set-up There is one city-developer for each potential city size. City developers fully tax local landowners and compete to attract firms to their city by subsidizing their 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. Developers have limited information over firms. They observe the industry of the firm as well as the profits made by firms locally, but they do not observe the level of efficiency of every individual firm. As a result, they do not have the full set of instruments that would allow them to internalize the externalities in this economy, contrary to the case in Henderson (1974). The equilibrium outcome will not be efficient. Nevertheless, as in Henderson (1974), city developers act as a coordinating device that allows a unique equilibrium to emerge in terms of city-size distribution.¹³ There is perfect competition and free entry among city developers, which drives their profits to zero in equilibrium.

City developers' problem Each city developer i announces a city size L and a level of subsidy to local firms' profits in sector j , $T_j^i(L)$. Developers are funded by the profits made on the housing market. As the housing market clears in each city, aggregate landowner profits at the city level are:

$$\pi_H(L) = b(1 - \eta)Lw(L), \quad (3.9)$$

¹³The model is silent on which location is used by city developers. The equilibrium is unique in terms of city-size distribution, not in terms of the actual choice of sites by developers, as all sites are ex ante identical.

as detailed in Appendix 3.7.1. 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

$$\ell_H(L) = (1 - b)(1 - \eta)L. \quad (3.10)$$

This impacts the local labor-market-clearing conditions used in equilibrium.

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

$$\max_{L, \{T_j^i(L)\}_{j \in 1, \dots, S}} \Pi_L = b(1 - \eta)w(L)L - \sum_{j=1}^S \int_z \frac{T_j^i(L)}{1 + T_j^i(L)} \pi_j(z, L) \mathbb{1}_j(z, L, i) M_j dF_j(z),$$

such that

$$\begin{aligned} \mathbb{1}_j(z, L, i) &= 1 && \text{if } L = \arg \max_L \pi_j(z, L) \text{ and firm } z \text{ chooses city } i, \\ \mathbb{1}_j(z, L) &= 0 && \text{otherwise.} \end{aligned} \quad (3.11)$$

In this expression, M_j denotes the mass of firms in sector j , $F_j(\cdot)$ is the distribution of 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 (3.7).

3.2.2 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.

Equilibrium definition

Definition 1 *An equilibrium is a set of cities $\mathcal{L}$ characterized by a city-size distribution $f_L(\cdot)$, a wage schedule $w(L)$, a housing-price schedule $p_H(L)$ and for each sector $j = 1, \dots, S$ 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 (3.3)) given $w(L)$, $p_H(L)$ and P_j ,
- (ii) utility is equalized across all inhabited cities,
- (iii) firms maximize profits (equation (3.7)) given $w(L)$, ρ 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 (3.11)) 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 optimally 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. 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, Duranton, and Robert-Nicoud (2010)).

Proposition 14 *There exists a unique equilibrium of this economy.*

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 15 *In equilibrium, city developers offer 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.*

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 Appendix 3.7.3. 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 (3.9)), 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 $w(L)N$, where N is the labor force hired in the traded goods sector locally, as can be seen from equation (3.8). Finally, the local labor force works either in the housing sector (equation (3.10)) or the traded goods sector, so that $N = L(1 - (1-b)(1-\eta))$. Profits given by (3.11) 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.

The equilibrium subsidy does not depend on city size, perhaps surprisingly. This comes from the fact that city developers only have a limited set of tools to influence firms' choices, since they only subsidize profits and do not observe firm efficiencies. These tools do not allow them to influence firm choices and internalize the local

production externalities. In this economy, the existence of a class of city developers helps pin down the set of cities that is opened up in equilibrium - as seen below - but does not influence firm choices. Therefore, the characterization of the equilibrium in terms of firms' outcomes, which I detail later, does not depend on the existence of city developers. They are also valid in an economy where cities are exogenously given.

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, the profit function of firm z in sector j is:

$$\pi_j^*(z, L) = \kappa_{1j} (1 + T_j^*) \left(\frac{\psi(z, L, s_j)}{w(L)^{1-\alpha_j}} \right)^{\sigma_j-1} R_j P_j^{\sigma_j-1} \quad (3.12)$$

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

$$L_j^{**}(z) = \arg \max_{L \geq 0} \pi_j^*(z, L). \quad (3.13)$$

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 16 *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. Define the matching function

$$L_j^*(z) = \arg \max_{L \in \mathcal{L}} \pi_j^*(z, L). \quad (3.14)$$

It is readily seen that the profit function of the firm (equation (3.12)) inherits the strict log-supermodularity of the productivity function in z and L . Therefore, the following lemma holds.

Lemma 17 *The matching function $L_j^*(z)$ is increasing 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 $\mathcal{L}$. As seen from equation (3.12), this optimal choice does not depend on general equilibrium quantities that enter the profit function proportionally for all city sizes. Finally, under the regularity assumptions made on ψ as well as on the distribution of z , $F_j(\cdot)$, 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 the following general equilibrium values. The reference level of wages $\bar{w}$ defined in equation (3.6) is taken as the numeraire. The remaining

unknowns are the aggregate revenues in the traded goods sector R , the mass of firms M_j and the sectoral price indexes P_j .

I use the following notation:

$$E_j = \int \frac{\psi(z, L_j^*(z), s_j)^{\sigma_j-1}}{L_j^*(z)^{\frac{b(1-\eta)(1+(1-\alpha_j)(\sigma_j-1))}{\eta}}} dF_j(z), \text{ and} \quad (3.15)$$

$$S_j = \int \left(\frac{\psi(z, L_j^*(z), s_j)}{L_j^*(z)^{\frac{b(1-\eta)(1-\alpha_j)}{\eta}}} \right)^{\sigma_j-1} dF_j(z), \quad (3.16)$$

where E_j and S_j are sector-level quantities that are fully determined by the matching functions $L_j^*(z)$ for each sector j . They are normalized measures of employment and sales in each sector.¹⁴

There is free entry of firms in each sector. Therefore, for all $j \in \{1, \dots, S\}$,

$$f_{Ej} P = (1 + T_j^*) \frac{\kappa_{1j}}{\kappa_{3j}} S_j \xi_j R P_j^{\sigma_j-1}, \quad (3.17)$$

where f_{Ej} is the units of final goods used up in the sunk cost of entry.

Aggregate production in each sector is the sum of individual firms' production. Therefore, for all $j \in \{1, \dots, S\}$,

$$\xi_j R = \frac{\sigma_j \kappa_{1j}}{\kappa_{3j}} M_j S_j \xi_j R P_j^{\sigma_j-1}. \quad (3.18)$$

The mass of workers N_o works either in one of the traded goods sectors or in the construction sector. The national labor market clearing condition is therefore

$$N_o = \sum_{j=1}^S M_j \frac{\kappa_{2j}}{\kappa_{3j}} E_j \xi_j R P_j^{\sigma_j-1} + N_o(1-b)(1-\eta), \quad (3.19)$$

¹⁴Given the wage equation (3.6) and the expression for operating profits (3.7), aggregate operating profits in sector j are $\frac{\kappa_{1j}}{\kappa_{3j}} M_j S_j R_j P_j^{\sigma_j-1} (1 + T_j^*)$ where $\kappa_{3j} = (1-\eta)^{b(1-\eta)(1-\alpha_j)(\sigma_j-1)}$. Similarly, aggregate revenues in sector j are $\frac{\sigma_j \kappa_{1j}}{\kappa_{3j}} M_j S_j R_j P_j^{\sigma_j-1}$ and aggregate employment in sector j is $\frac{\kappa_{2j}}{\kappa_{3j}} M_j E_j R_j P_j^{\sigma_j-1}$, where $\kappa_{2j} = \kappa_{1j}(1-\alpha_j)(\sigma_j-1)$.

where the last term derives from equation (3.10).

Inverting this system of $2S + 1$ equations gives the $2S + 1$ unknowns P_j , M_j for all $j \in \{1, \dots, S\}$ and R , the aggregate revenues in the traded goods sector, as detailed in Appendix 3.7.4.

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,

$$\forall L > L_{min},$$

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

where $L_{min} = \inf(\mathcal{L})$ the smallest city size in the equilibrium.¹⁵

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)

$$f_L(L) = \kappa_4 \frac{\sum_{j=1}^S M_j \mathbb{1}_j(L) \ell_j(z_j^*(L)) f_j(z_j^*(L)) \frac{dz_j^*(L)}{dL}}{L}, \quad (3.20)$$

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 $f_L(\cdot)$ is uniquely determined by equation (3.20).

¹⁵I abuse notations here since $z_j^*(L)$ is not defined over the entire set $\mathcal{L}$ but over a convex set strictly included in $\mathcal{L}$.

Lemma 18 *$f_L(.)$ 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 (3.20), 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. I show in Appendix 3.7.3 that under some parametric restrictions for ψ , consistent with the ones used in the empirical section, the city-size distribution follows Zipf's law if the firm-size distributions follow Zipf's law.

Second, as all sites used by city developers are identical ex ante, the model has no predictions on which sites are used to build these cities. Finally, 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 such cities is determined by the model. On the other hand, the model is mute on the sectoral composition of an *individual* city within a given class, which is irrelevant for aggregate outcomes. The model could easily be extended to accommodate *localization* externalities, by assuming that, for a given sector, the agglomeration externality depends on the size of this sector and not the overall city size. This would lift this city-level indeterminacy. Cities would be perfectly specialized in that sector, since the congestion costs depend on the overall city size, but the benefits are sector-specific. This would not change any other characterization of the equilibrium. In particular, the city-size distribution defined in equation (3.20) and lemma 18 would still hold.

This step completes the full characterization of the unique equilibrium of the economy.

3.2.3 Characterization

I proceed to derive several properties of the equilibrium, following two objectives. The first is to verify that the assumptions of the model lead to predictions that are broadly consistent with salient features of the data. I derive the theoretical predictions here and, in the empirical Section 3.3, I present a set of stylized facts from French firm-level data that are broadly consistent with these predictions. My second objective is to propose a model-based decomposition of the impact of place-based policies on welfare and the productive efficiency of the equilibrium, which will guide the counterfactual policy analysis I conduct in Section 3.4.

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

$$\begin{aligned}\psi_j^*(z) &= \psi(z, L^*(z), s_j), \\ r_j^*(z) &= \sigma_j \kappa_{1j} \left(\frac{\psi(z, L^*(z), s_j)}{w(L_j^*(z))^{1-\alpha_j}} \right)^{\sigma_j-1} P_j^{\sigma-1} R_j, \end{aligned} \tag{3.21}$$

$$\ell_j^*(z) = \kappa_{2j} \frac{\psi(z, L^*(z), s_j)^{\sigma-1}}{w(L_j^*(z))^{(\sigma-1)(1-\alpha_j)+1}} P_j^{\sigma-1} R_j, \tag{3.22}$$

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 17), firm-level observables also exhibit complementarities with city size. Let $\mathcal{L}$ denote the set of city sizes in the economy. In all that follows, the characterizations

hold regardless of whether $\mathcal{L}$ is the unique set determined in equilibrium by the city developers problem, or $\mathcal{L}$ is exogenously given.

Proposition 19 *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.¹⁶ 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

$$\epsilon_l = \epsilon_r - b \frac{1 - \eta}{\eta}, \quad (3.23)$$

so that ϵ_l is not necessarily positive.

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 , and both impact the sorting process, leading in turn to differences in observed outcomes. The model delivers predictions on the geographic distribution

¹⁶In the data, the sorting is imperfect. I allow for imperfect sorting in the estimation by specifying an error structure around the baseline model.

of sectors as well as on the heterogeneity of the firm-size distribution across sectors, which I use to guide the estimation in Section 3.3.

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

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

Proposition 20 *In a competitive equilibrium, all else equal, the geographic distribution $\tilde{F}_j$ of a high α_j sector first-order stochastically dominates that of a lower α_k sector. The geographic distribution $\tilde{F}_j$ of a high s_j sector first-order stochastically dominates that of a lower s_k sector.*

The formal proofs are in Appendix 3.7.3. 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.

Note that the model predicts that firms of different sectors coexist in equilibrium within each type of city size.¹⁷ The support of city sizes chosen by firms in each sector

¹⁷As noted above, the share of employment in each sector is pinned down on average over all cities of the same size, but is indeterminate for individual cities. If, in contrast to what I assume, exter-

is shaped by s_j and α_j ; in particular, if the productivity distribution has convex and bounded support, the support of city sizes chosen by firms is summarized by an interval, whose bounds increase with α_j or s_j . The intersection of such supports is not, a priori, empty. Holding city size fixed, a variety of sectors may be present in some types of cities at equilibrium, but each with a different efficiency level, realized productivity, and realized revenues, as the matching function differs by sector.

Firm-size distribution Because firm sorting reinforces initial differences between firms, the intensity of sorting 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 21 *If $(\alpha_2, s_2) \geq (\alpha_1, s_1)$, the observed firm-size distribution in revenues is more spread in Sector 2 than in Sector 1. For any $p_1 < p_2 \in (0, 1)$, $\frac{Q_1(p_2)}{Q_1(p_1)} \leq \frac{Q_2(p_2)}{Q_2(p_1)}$.*

In other words, if one normalizes the median of each revenue distribution to the same level, all higher quantiles in the revenue distribution of Sector 2 are strictly higher than in Sector 1, and all lower quantiles are below. The distribution in Sector 1 is more unequal. As a consequence, 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

nalities were sector-specific (localization externalities), the model would predict the same sectoral shares on average over all cities of a given size. The only difference would be that individual cities would then be fully specialized.

Corollary 22 *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: $\zeta_2 \leq \zeta_1$.*

Policy implications

Having characterized the equilibrium, I now turn to proposing a model-based decomposition of welfare, which will provide tools to guide the quantitative counterfactual policy analyses I conduct in Section 3.4.

In the quantitative exercise, I analyze two types of policies. First, federal governments subsidize firms to modify their location choice. In general, these policies are designed to attract firms to underdeveloped areas.¹⁸ They are advocated for reasons of equity – policy makers want to smooth out spatial inequalities – and efficiency, since helping cities grow can have multiplying effects on productivity in the presence of agglomeration externalities. In what follows, I discuss how to analyze their aggregate welfare effects in the context of the model. These policies do not affect spatial inequalities in terms of welfare, since utility is equalized across all cities in equilibrium in the model. Nevertheless, they do impact other measures of spatial inequality, such as spatial inequality in real wage or in productivity. I quantify the impact of place-based policies on such measures, together with their aggregate welfare impact, in Section 3.4.

I model place-based policies as follows. In contrast to the constant subsidies given out by local developers, these policies subsidize firms more when they choose to locate in less productive cities, i.e., smaller cities in the context of the model.¹⁹ This takes the form of subsidies or tax breaks that are linked with the size of the city chosen by firms. This scheme shapes the incentives of firms and distorts the

¹⁸I discuss these policies in more detail in the policy analysis section.

¹⁹Smaller cities host the least productive firms in the model

equilibrium matching functions $L_j^*(z)$, leaving the wage schedule unchanged. This policy is funded by raising a lump-sum tax on firms.

Second, I study policies that aim to remove some of the existing constraints on city growth. This class of policies encompasses a broad range of instruments which make the housing supply more elastic, such as lifting existing zoning regulations or building-height regulations. In the model, the housing supply elasticity is driven by the land-use intensity b in the housing production function (3.1). I model an easing of zoning regulations as a decrease in land-use intensity b . This leads to a direct, mechanical, positive welfare effect that is not the focus here, and for which I control in the quantitative analysis: all else equal, more housing can be built with the same amount of land, which directly increases workers' utility. I focus instead on the indirect effects this policy has on the differentiated goods sectors. Because the utility of workers is equalized across city sizes, the wage elasticity to city size $b^{\frac{1-\eta}{\eta}}$ is also shaped by the land-use intensity. In turn, this elasticity impacts the firm matching function, and, ultimately, the productive efficiency of the differentiated goods sectors. Firms match with larger cities, and cities are larger in equilibrium.

Welfare depends on the productivity of the economy, as well as on the congestion costs borne out by workers. The policies I study impact both. I first examine the channels through which policies impact aggregate productivity in the differentiated goods sectors, before turning to the expression for welfare.

Productivity Let Y_j , N_j and K_j denote, respectively, sectoral levels of output, employment and capital. As I show in Appendix 3.7.4, sectoral productivity $TFP_j \equiv \frac{Y_j}{K_j^{\alpha_j} N_j^{1-\alpha_j}}$ can be expressed

$$TFP_j = M_j^{\frac{1}{\sigma_j-1}} \frac{S_j^{\frac{\sigma_j}{\sigma_j-1}-\alpha_j}}{E_j^{1-\alpha_j}}, \quad (3.24)$$

where E_j and S_j are defined in equations (3.15) and (3.16).

Holding fixed the matching function $L_j^*(z)$ and the mass of firms M_j , this expression would be maximized if $w(L)$ did not depend on city size. In that case, TFP_j is simply

$$\left(\int \psi(z, L_j^*(z), s_j)^{\sigma_j-1} dF_j(z) \right)^{\frac{1}{\sigma_j-1}}.$$

This is reminiscent of the results in the misallocation literature (Hsieh and Klenow (2009)). Productivity is maximized when all firms face the same input price, whereas in the spatial equilibrium, wages increase with city size. Second, given the wage schedule and the mass of firms M_j , productivity increases if the matching function $L_j(z)$ is pushed up, since firms benefit from more externalities. Finally, TFP_j increases as more firms enter the market, through a standard love of variety effect.²⁰

These three channels offer a way to decompose the effects of place-based policies on the productive efficiency of the economy. I quantify each in the quantitative section. Qualitatively, the first type of policies I study affects productivity by pushing down the equilibrium matching function. Firms tend to be pushed down toward smaller cities, but on the other hand, in the new spatial equilibrium, city sizes adjust to local labor demand and small cities grow. The aggregate effect on productive efficiency is ambiguous, as is the entry effect. The second type of policies has positive effects both through the misallocation channel, since the wage schedule is flatter, and through a new matching function in which firms sort into larger cities in equilibrium. The entry effect is a priori ambiguous.

²⁰The expression for the mass of firms is $M_j \propto \left(\left(\sum_{j=1}^S \xi_j \frac{E_j}{S_j} \right)^{1-\bar{\alpha}} \prod_{j=1}^S S_j^{\frac{\xi_j}{\sigma_j-1}} \right)^{\frac{\bar{\sigma}-1}{(1-\bar{\alpha})(\bar{\sigma}-1)-1}}$ where $\bar{\alpha} = \sum_{j=1}^S \alpha_j \xi_j$ and $\frac{1}{\bar{\sigma}-1} = \sum_{j=1}^S \frac{\xi_j}{\sigma_j-1}$.

Welfare The model lends itself naturally to welfare analysis. Given the choice of numeraire, welfare is given by

$$\bar{U} = \left(\frac{1}{P} \right)^\eta$$

As shown in Appendix 3.7.4, welfare depends positively on a measure of aggregate productivity, and negatively on a term that summarizes the aggregate congestion costs in the economy. Comparing two equilibria,

$$\bar{U} \propto \left(\prod_{j=1}^S TFP_j^{\xi_j} \right)^{\frac{\eta}{1-\bar{\alpha}}} \left(\prod_{j=1}^S \left(\frac{S_j}{E_j} \right)^{\xi_j(1-\alpha_j)} \right)^{-\frac{\eta}{1-\bar{\alpha}}},$$

where $\bar{\alpha} = \sum_{j=1}^S \alpha_j \xi_j$ is an aggregate measure of the capital intensity of the economy.

The term $\prod_{j=1}^S TFP_j^{\xi_j}$ is a model-based measure of aggregate productivity. Take the example of a policy that increases TFP by pushing firms to larger cities. It has a direct positive impact on welfare, magnified by the term $\frac{1}{1-\bar{\alpha}}$ that captures the fact that capital flows in response to the increased TFP in the economy, making workers more productive.

This effect is dampened by the second term, which captures the congestion effects that are at play in the economy. Recall that in this economy, wages increase to compensate workers for increased congestion costs in larger cities. Here, $\frac{S_j}{E_j}$ measures the ratio of the average sales of firms to their average employment in a given sector. It is a model-based measure of the representative wage in the economy, since $\frac{r_j^*(z)}{\ell_j^*(z)} \propto w(L^*(z))$ for each firm. Coming back to the example of a policy that tends to push firms into larger cities, such a policy will also tend to increase aggregate congestion in the economy by pushing workers more into larger cities. Individual workers are compensated for this congestion by increased wages, in relative terms across cities,

so that all workers are indifferent across city sizes. But the level of congestion borne by the *representative* worker depends on how workers are distributed across city sizes. It increases as the economy is pushed toward larger cities. This negative effect is captured by the second term in the welfare expression that decreases with the representative wage.

3.3 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 presented above. I then use these theoretical characterizations of the equilibrium to guide the structural estimation of the model.

3.3.1 Data

I use a firm-level data set of French firms (BRN). It 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 (SIRENE), 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 metropolitan France. They are defined with respect to the observed commuting patterns of workers and cover all of France.²¹ They are designed to capture local labor markets and are better suited than administrative areas, which they abstract from, to capturing the eco-

²¹They are presented as areas where "*most workers live and work, and where establishments can find most of their workforce*".

nomic forces at play in the model.²² 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, Duranton, Gobillon, Puga, and Roux (2012), i.e., manufacturing sectors and business services, excluding finance and insurance. Finally, I retain firms with two employees or more. Data for smaller firms tend to be very noisy. I trim the bottom and top 1% of the data. This leaves me with 130,018 firms. Summary statistics are reported in Table 3.1.

3.3.2 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 (lemma 19). Furthermore, firm sorting is shaped by two key sector-level 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 (lemma 20) and (ii) the dispersion of the sectoral firm-size distribution (lemma 21). I turn to examining the raw data in these dimensions.

²²The previous definition of these zones was constrained by some administrative borders. A new definition of these zones was published by INSEE in 2011. Updated measures of zone sizes were published for all years after 1997.

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 148 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.²³ 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.²⁴ Figure 3.1 plots the distributions of these elasticities. For value added, it is positive for 95% of the observations, and is significantly negative for only one industry, sawmills. This is intuitive since natural resources drive location in this industry—a dimension the sorting model abstracts from. The elasticity of employment to city size is shifted to the left compared to the elasticity of value added to city size. 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. The model distinguishes between inputs whose price varies locally and inputs with constant price across space. To proxy for inputs whose price does not systematically increase with city size, I use a measure of “tradable capital” defined as total

²³ Because the model does not feature the use of intermediates, I use value added as the measure of firm output.

²⁴For 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 81% of firms and 45% of employment in the sample.

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 3.2 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, since the main data set does not have information on the composition of the workforce.²⁵ Specification (III) in Table 3.2 shows that controlling for industry-level skill intensity does not affect the results. Specification (IV) runs the same specification, 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 supply-side explanations. Again, the results are robust to using this reduced sample. Overall, Table 3.2 is consistent with the idea that firms location choices are shaped by the intensity of input use in their industry.

Third, the model predicts that firms that locate in large cities benefit disproportionately from agglomeration externalities. As a consequence, the sectoral firm-size distribution is more fat-tailed for industries located in larger cities. Table 3.3 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 in-

²⁵I 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.

dustry j . The shape parameter ζ_j is estimated by running the following regression, following Gabaix and Ibragimov (2011):

$$\log\left(rank_{ij} - \frac{1}{2}\right) = \alpha_j - \zeta_j \log(\text{value added}_{ij}) + \epsilon_{ij}, \quad (3.25)$$

where j indexes industries, i indexes firms and $rank_{ji}$ is the rank of firm i in industry j in terms of value added. Table 3.3 shows a negative correlation between ζ_j and the fraction of establishments in industry j located in large cities (defined as in Table 3.2). In other words, industries that locate more in large cities 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 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:

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

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} is the logarithm of $L_{i,t}$ or a set of initial city fixed effects.²⁶ Table 3.4 shows that, conditional on moving, firms that are initially larger tend to move into larger cities. 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 3.4 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.

3.3.3 Structural estimation

I now turn to the estimation of the model. The model is estimated industry by industry, on 23 aggregated industries. I use an indirect inference method and minimize the distance between moments of the data and their simulated counterparts to estimate the sectoral parameters that govern the model. Because, in the theory, the strength of agglomeration externalities shapes (i) the location patterns of firms ; (ii) the elasticity of firm-level revenues to city size and (iii) the dispersion of the sectoral firm-size distribution, I use moments that describe these dimensions of the data as targets of the estimation procedure.

Model specification

I first lay out the econometric specification of the model. The literature has traditionally assumed that agglomeration externalities were of the form $\psi(z, L, s_j) = zL^{s_j}$, where s_j measures the strength of externalities. However, in such a framework, externalities enter multiplicatively in the profit function and there is no complementarity between firm productivity and city size. In contrast, the model presented in Section 3.2 assumes such a complementarity. I thus postulate the following functional form of the productivity function, for each sector $j \in 1 \dots S$:

²⁶These 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.

Assumption C

$$\log(\psi_j(z, L, s_j)) = a_j \log L + \log(z)(1 + \log \frac{L}{L_o})^{s_j} + \epsilon_{z,L} \quad \text{for } \log(z) \geq 0 \text{ and } L \geq L_o \quad (3.26)$$

$$\log(\psi_j(z, L)) = 0 \quad \text{for } L < L_o$$

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 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. It 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. In what follows, I write $\tilde{L} = \frac{L}{L_o}$, and $\mathcal{L}$ the set of normalized city sizes in the simulated economy.²⁷ I assume that $\log(z)$ is distributed according to a normal distribution with variance ν_Z , truncated at its mean to prevent $\log(z)$ from being negative. This restriction is needed for the productivity of a firm to be increasing in city size in equation (3.26).

I introduce an error structure by assuming that firms draw idiosyncratic productivity shocks $\epsilon_{z,L}$ for each city size, where $\epsilon_{z,L}$ is i.i.d. across city sizes and firms. It is distributed as a type-I extreme value, with mean zero and variance ν_R . This generates imperfect sorting. 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.²⁸

²⁷ L_o changes proportionally the size of all cities but does not alter the relative matching function and firm-size distributions used in the estimation nor the relative prices. $\tilde{L}$ is the relevant measure for firm choices. L_o is calibrated to match the actual level of city sizes in the data.

²⁸ I show in Appendix 3.7.2 that under Assumption C, firms in sectors with a lower labor intensity $1 - \alpha_j$ or a higher agglomeration parameter s_j locate with a higher probability in larger cities, whereas in the theoretical section this statement was not probabilistic.

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.

Estimation procedure

The estimation is conducted in two stages. In the first stage, a set of parameters is calibrated from the data. In the second stage, I use indirect inference to back out the parameters that require a simulation method to be estimated. The indirect inference 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. I still am able to capture relevant heterogeneity across sectors.

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.²⁹ I then calibrate the land-use intensity b and housing share $1 - \eta$. b and

²⁹In each sector, $\hat{\sigma}_j$ and $\hat{\alpha}_j$ are calibrated using $\frac{\hat{\sigma}_j}{\hat{\sigma}_j - 1} = \text{mean}(\frac{\text{v.a.}}{\text{costs}})$ 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.

$1 - \eta$ jointly determine (1) the elasticity of wages to city size and (2) the elasticity of the housing supply. I thus calibrate them to match the median elasticity of housing supply measured in the US by Saiz (2010) and the wage elasticity to city size measured in the data.³⁰ To measure the elasticity of wages to city size, I follow equation (3.23) 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 simply measuring its share of value-added produced.

In the second stage, I estimate the following parameters for each sector: a_j , which measures the log-linear agglomeration term; s_j , which measures the intensity of complementarity between agglomeration externalities and firm efficiency; ν_{Zj} , which measures the variance of the firms' raw efficiency; and ν_{Rj} , which is the variance of the firm-city size idiosyncratic shock. Firms make a discrete choice of (normalized) city size, according to the following equation

$$\log \tilde{L}_j^*(z) = \arg \max_{\log \tilde{L} \in \mathcal{L}} \log(z) \left(1 + \log \tilde{L}\right)^{s_j} + (a_j - b(1 - \alpha_j) \frac{1 - \eta}{\eta}) \log \tilde{L} + \epsilon_{z,L}, \quad (3.27)$$

which is the empirical counterpart of equation (3.14). Because the choice equation involves unobservable heterogeneity across firms and is non-linear, I have to use a simulation method to recover the model primitives. I use an indirect inference method (Gouriéroux and Monfort (1997)) to estimate the true parameter $\theta_0^j = (a^j, s^j, \nu_R^j, \nu_Z^j)$ for each sector j . The general approach is close to the one in (Eaton, Kortum, and Kramarz (2011)), except that I also use indirect moments in the inference (namely, regression coefficient estimates). The estimate θ_{II}^j minimizes an indirect inference loss

³⁰I use the median measure of housing supply elasticity for the US as there has been no similar study done on French data, to my knowledge.

function

$$\|m_j - \hat{m}_j(\theta)\|_{W_j^2} = (m_j - \hat{m}_j(\theta))' W_j (m_j - \hat{m}_j(\theta)), \quad (3.28)$$

where m_j is a vector stacking a set of moments constructed using French 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.³¹ I compute standard errors using a bootstrap technique.³²

I simulate an economy with 100,000 firms and 200 city sizes. I follow the literature in using a number of draws that is much larger than the actual number of firms in each sector, to minimize the simulation error. I use a grid of 200 normalized city sizes $\tilde{L}$, ranging from 1 to M where M is the ratio of the size of the largest city to the size of the smallest city among the 314 cities observed in the French data. This set of city-sizes $\mathcal{L}$ is taken as exogenously given.³³ In contrast, the corresponding city-size distribution is not given a priori: the number of cities of each size will adjust to firm choices in general equilibrium to satisfy the labor-market clearing conditions.

³¹ 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 . I calculate Ω_j using the following bootstrap procedure. For a sector with N_j firms, (1) I resample with replacement N_j firms from the initial set of firms in this sector (2) for each resampling, I compute m_j^b , the value of the moments for this set of firms (3) I compute

$$\Omega_j = \frac{1}{2000} \sum_{b=1}^{2000} (m_j^b - m_j)(m_j^b - m_j)'$$

I take its generalized inverse to compute W_j .

³²I run the estimation procedure 30 times; for each iteration, I take a new set of normalized draws for the firm productivity and the firm-city size idiosyncratic shocks. This accounts for simulation error. For each iteration, I also recompute the targeted moments by resampling firms in the data. This accounts for sampling error. I then estimate the parameter value θ_b^j , which minimizes (3.28). The standard errors I report for parameters in sector j are the square root of the diagonal elements of

$$V_j = \frac{1}{30} \sum_{b=1}^{30} (\theta_b^j - \theta_{II}^j)(\theta_b^j - \theta_{II}^j)'$$

³³As pointed out in the theory section, the characterizations of the economy provided in Section 3.2 hold if the set of possible city sizes is exogenously given.

The algorithm I use to simulate the economy and estimate the parameters for each sector is as follows:

Step 1: I draw, once and for all, a set of 100,000 random seeds and a set of $100,000 \times 200$ random seeds from a uniform distribution on $(0, 1)$.

Step 2: For given parameter values of ν_R and ν_z , I transform these seeds into the relevant distribution for firm efficiency and firm-city size shocks.

Step 3: For given parameter values of a and s , I compute the optimal city size choice of firms according to equation (3.27).

Step 4: I compute the 17 targeted moments described below.

Step 5: I find the parameters that minimize the distance between the simulated moments and the targeted moments from the data (equation (3.28)) using the simulated annealing algorithm.

The estimation is made in partial equilibrium, given the choice set of normalized city-sizes $\mathcal{L}$. It relies on measures that are independent of general equilibrium quantities, namely the sectoral matching function between firm efficiency and city size, and relative measures of firm size within a sector.³⁴

Moments

I use four sets of moments to characterize the economy, guided by the predictions of the model, and provide intuition for how they help identify the parameters of the model.

³⁴Specifically, as detailed in the theoretical model, the optimal choice of city size by a firm depends only on its productivity function and on the elasticity of wages with respect to city size. It does not depend on general equilibrium quantities. The sizes of all firms in a given sector depend proportionally on a sector-level constant determined in general equilibrium (see equations (3.21) and (3.22)). Normalized by its median value, the distribution of firm sizes within a sector is fully determined by the matching function.

(i) Sectoral location. The first set of moments summarizes the geographic distribution of economic activity within a sector. I use the share of employment in a given sector that falls into one of 4 bins of city sizes. 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. I normalize these sizes by the size of the smallest city, and call these thresholds t_i^L . I compute the fraction of employment for each sector in each of the city bins, both in the data and in the simulated sample. The corresponding moment for sector j and bin i is

$$s_i^{L,j} = \frac{\sum_{t_i^L \leq L < t_{i+1}^L} \int \ell_j^*(z) \mathbb{1}_{L_j^*(z)=L} dF_j(z)}{\int \ell_j^*(z) dF_j(z)},$$

where $\ell_j^*(z)$ is the employment of firm z and $\mathbb{1}_{L_j^*(z)=L}$ is a characteristic function which equals 1 if and only if firm z in sector j chooses to locate in city size L . These moments depend directly on the matching function between firms and city size. This procedure defines 4 moments for each industry.

(ii) Elasticity of revenues to city size. I use the elasticity of average firm revenues to city size as an additional moment. In the model, absent an idiosyncratic productivity shocks, this elasticity would have a closed-form expression that does not depend on the distribution of firm efficiency and would only be a function of the strength of agglomeration externalities. In the presence of an error structure, the elasticity of average firm revenues to city size is also shaped by the presence of idiosyncratic productivity shocks, which dampens the elasticities predicted by perfect sorting. This moment is computed as follows in sector j . Define

$$\bar{r}_j(L) = \frac{\int r_j^*(z) \mathbb{1}_{L_j^*(z)=L} dF_j(z)}{\int \mathbb{1}_{L_j^*(z)=L} dF_j(z)}$$

the average revenues of sector j firms that locate in city L . The elasticity of average firm revenue to city size in sector j is the regression coefficient estimate $\varepsilon_{r,j}$ of the following equation:

$$\log(\bar{r}_j(L_i)) = c_j + \varepsilon_{r,j} \log L_i + \nu_i,$$

where i indexes the city sizes in $\mathcal{L}$. In the data, I compute this moment on the sub-sample of mono-establishment firms, as the information on value-added is only available at the firm level and not at the establishment level.

(iii) Firm-size distribution in revenues. Third, I use moments that characterize the firm-size distribution in revenues. As discussed in Section 3.2, these are directly impacted by the sorting mechanism. To compute these moments, I first normalize firms value-added within a given sector by their median value. I retrieve from the data the 25, 50, 75 and 90th percentiles of the distribution and denote them $t_i^{r,j}$.³⁵ These percentiles define 5 bins of normalized revenues. I then count the fraction of firms that fall into each bin

$$s_i^{r,j} = \frac{\int \mathbb{1}_{t_i^{r,j} \leq \tilde{r}_j(z) < t_{i+1}^{r,j}} dF_j(z)}{\int dF_j(z)},$$

where $\tilde{r}_j(z)$ are the normalized revenues of firm z in sector j . I also measure the tail of the distribution, captured by a log rank-log size regression as in equation 3.25. This procedure define 6 moments for each sector.

If the matching function alone does not allow me to identify separately ν_R from ν_Z , these moments do. 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, namely, firm employment and value-added. In particular, these relationships allow me to identify the variance of idiosyncratic shocks separately

³⁵As in Eaton, Kortum, and Kramarz (2011), higher quantiles are emphasized in the procedure, since they capture most of the value added, and the bottom quantiles are noisier.

form the variance of firm's raw efficiency. Intuitively, ν_Z impacts the relative quantiles of the firm-size distribution both indirectly, through the matching function, and directly, through the distribution of raw efficiency z . In contrast, ν_R impacts the relative quantiles of the firm-size distribution only indirectly, through the matching function.

(iv) Firm-size distribution in employment. Fourth, I construct the same set of moments as in (iii) for employment. In a model with perfect sorting, there would be a perfect correlation between ex-ante firm efficiency and city size and hence between wages and revenues. Thus, with perfect sorting, the information contained in the distribution of revenues and employment would be redundant. However, the presence of idiosyncratic shock leads to imperfect sorting and shapes the difference between the distribution of employment and the distribution of revenues. Having both these distributions as moments contributes to quantifying the amount of imperfect sorting, captured by ν_R .

Model fit

Targeted moments. I first examine the model fit for the set of moments targeted by the estimation procedure. Figure 3.2 shows the cross-sectoral fit between the estimated moments and their targeted value in the data, for the following moments: (1) the shape parameter of the distribution in revenues (2) the shape parameter of the distribution of employment (3) the elasticity of revenues to city size and (4) the share of employment located in cities below the median city (as defined above). The 45° line represents the actual moments in the data. Overall, as seen in Figure 3.2, the model captures well the cross-sectional heterogeneity in these targeted moments. The estimation relies on five other 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

3.3 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. The fit for the employment distribution is similar and is not reported. Finally, the estimation relies on the share of sectoral employment 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 3.4. 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. Formally, I measure the variance explained by the model as an R^2 of a regression of the simulated share on the actual shares, forcing an origin of 0 and a slope of 1. The R^2 is 0.38 (0.46 when weighting by sector size).

Moments not targeted. Next, I investigate the model fit for moments not used in the estimation. First, I examine how the firm-size distribution differs for the set of firms located in small cities and the set of firms located in large cities. This is motivated by the analysis in Combes, Duranton, Gobillon, Puga, and Roux (2012), who show that the productivity distribution of firms is more dilated for the set of firms that locate in larger cities. I measure the ratio of the 75th percentile to the 25th percentile of the firm-size distribution in revenues, both for firms located in cities above the median threshold and in cities below. I define “dilation” as the ratio of this two measures. Figure 3.5 plots for each sector the log of “dilation” computed on the simulated data against the same variable computed on the actual data. I measure the fit of the model by computing the R^2 of a regression of the simulated dilation on the actual dilation, forcing an origin of 0 and a slope of 1. The R-squared is 0.39 (0.45 when weighting by sector size). The model captures the existence of a dilation effect

in firm-size distribution, between small and large cities, as well as the cross-sectoral heterogeneity in dilation.

Finally, another moment not targeted in the estimation 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 3.2). 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 3.6, where the city-size distribution is plotted for the simulated data and the actual data. 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.³⁶

Analysis of the parameter estimates

I turn to the analysis of the parameter estimates of the model. Table 3.5 reports the estimated parameters industry by industry, with standard errors in parenthesis. The sectoral estimate of s_j , the parameter that governs the strength of the complementarity between firm efficiency and agglomeration externalities, is positive for all but three industries. These three industries correspond to the shoes and leather industry, the manufacture of glass and ceramics (where the coefficient is non significantly different from 0) and metallurgy. All three are relatively mature industries. Three other sectors have positive but insignificant point-estimates for s_j , namely paper products, rubber and plastic products, and office machinery. That more mature industries tend to exhibit different agglomeration forces is reminiscent of the argument in Audretsch

³⁶The levels are arbitrary and chosen so that the figure is readable.

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. I present average measures across sectors in the main text. Table 3.6 reports these decompositions industry by industry.

A first raw measure of the observed elasticity of firm productivity to city size (or city density, as all cities have the same amount of land in the model) can be computed by running the following simple OLS regression:

$$\log Prod_{i,j} = \beta_0 + \beta_1 \log L_i + \delta_j + \mu_i, \quad (3.29)$$

where $Prod_{i,j}$ stands for $\psi_{i,j}$, the equilibrium productivity of 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 5.4%. 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 (3.29).³⁷ In the estimated model, these observed 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 simulate the model using the same distribution of firm efficiency and idiosyncratic shocks as in the baseline equilibrium, but constrain firms to choose their city size as if they all had the average efficiency in their sector. I recover from this simulation a new distribution of firm-level observables. In this counterfactual, the difference in firms' location choice is only driven by firm-city size specific iid productivity shocks.³⁸ This counterfactual thus allows me to compute what would be realized productivities if firms did not sort across cities according to their raw efficiencies. Panel A of Figure 3.7 compares the relationship between firm-level productivity and city size, in the baseline model and in this counterfactual simulation. This relationship is flatter in the counterfactual simulation. Estimating equation (3.29) on the counterfactual data leads to an elasticity of firm productivity to city size of 2.1%. By this account, firm sorting accounts for almost two thirds of the productivity gains measured in equilibrium between cities of different sizes.

I now turn to a different exercise and compute a model-based estimate of what would be the impact on firm productivities of an exogenous increase in city size, all else equal. To do so, I first decompose the observed equilibrium firm-level productivities

³⁷An exception is Combes, Duranton, and Gobillon (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.

³⁸I also keep the set of possible city sizes constant.

into a “systematic” component and an idiosyncratic component as follows:

$$\log \psi_{i,j} = \log \tilde{\psi}_{i,j} + \epsilon_{i,L_i}.$$

The “systematic” component in productivity is defined as

$$\log \tilde{\psi}_{i,j} = a_j \log L_i + \log(z_i) \left(1 + \log \frac{L_i}{L_o}\right)^{s_j}$$

from equation (3.26), where L_i is the optimal city size chosen by firm z_i in equilibrium. The idiosyncratic shocks are orthogonal to z and L *ex ante*. Because firms select their optimal city size based on their draws, there is a correlation *ex post* between the realized ϵ_{z,L_i} and the firm-level observables. Panel B of Figure 3.7 plots the observed firm productivity against city size at equilibrium, as well as its systematic and its idiosyncratic components. As Panel B of Figure 3.7 shows, the magnitude of the realized idiosyncratic shocks decreases with city size. This result can be intuitively explained as follows. Absent idiosyncratic shocks, firms tend to be more profitable in larger cities, on average. To locate in smaller cities, they need to draw a relatively large idiosyncratic shock there. This selection effect tends to dampen the OLS estimate of β_1 in equation (3.29). Using the structure of the model, I can estimate the effect of an exogenous increase in city size on local firms productivity, by computing the sensitivity of the systematic component of firm productivity $\tilde{\psi}$ to city size L .³⁹ To do so, I measure $\tilde{\psi}$ in the baseline equilibrium for all firms that share a common efficiency level z but have chosen in equilibrium to locate in different city sizes, driven by the iid shocks ϵ_{zL} . Panel C of Figure 3.7 plots, for a given level of firm efficiency z , the systematic component of firm productivity as a function of city size. Results are shown for the 25th percentile, the median and the 75% percentile of the raw ef-

³⁹This exercise is by definition partial equilibrium in that I hold constant firms’ location and realization of idiosyncratic productivity shocks.

efficiency distribution z . The corresponding elasticity of firm systematic productivity to city size are estimated, respectively, at 5.8%, 7.2% and 8.3%. These estimates offer a model-based measure of what would be the increase in firm productivity if a city size doubled (and local agglomeration externalities increased accordingly), all else equal. It also illustrates the heterogeneous effects of agglomeration externalities, at the estimated parameter values.

3.4 The Aggregate Impact of Place-Based Policies

Equipped with the estimates of the model's parameters, I now turn to the evaluation of the general equilibrium impact of a set of place-based policies.

3.4.1 Local tax incentives

I first study policies that subsidize firms locating in less developed cities. This type of federal program is widespread and has been studied, for example, in Busso, Gregory, and Kline (2013) for the US or Mayer, Mayneris, and Py (2012) for France. They aim at reducing spatial disparities and are 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. This argument, however, needs to be refined. As has been pointed out in the literature (Glaeser and Gottlieb (2008), Kline and Moretti (2013)) 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. Nevertheless, the economy

is characterized by spatial disparities in real wages, and in the productivity of firms located in different cities. Place-based policies impact these spatial inequalities. They tend to benefit the targeted areas, but the extent to which they reduce aggregate measures of inequality 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. I first calibrate the policy to match the key characteristics of the French “ZFU” program (*Zones Franches Urbaines*), a policy similar 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. It is targeted to geographically delimited zones that cover an overall population of 1.5M, or 2.3% of the French population. I retain this number as the geographical scope of these policies and implement 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 15% of firms’ profits in these areas, paid for by a lump-sum tax levied on all firms in the country. The gross cost of this subsidy in the model is 0.03% of GDP, which matches the one reported for France for the ZFU program. A 15% subsidy on profits correspond to a 45% tax break on the French corporate tax, whereas the French ZFU program offers full corporate tax breaks for 5 years as well as other generous labor tax and property tax relief.

To compute the counterfactual equilibrium, I proceed as follows.

Step 1: I start from the equilibrium estimated in the data. I hold fixed the number of workers in the economy, the real price of capital, the set of idiosyncratic productivity shocks for each firm and city-size bin, and the distribution of firms’ initial raw efficiencies.

Step 2: I recompute the optimal choice of city-size by firms, taking into account the altered incentives they face in the presence of the subsidy.

Step 3: Because the composition of firms within a given city-size bin changes, total labor demand in a city-size bin is modified. I hold constant the number of cities in each bin and allow the city size to grow (or shrink) so that the labor market clears within each city-size bin. This methodology captures the idea that these policies are intended to “push” or jump-start local areas, which in addition grow through agglomeration effects.⁴⁰

Step 4: As city sizes change, the agglomeration economies and wage schedules are modified, which feeds back into firms’ location choice.

Step 5: I iterate this procedure from step 2, using the interim city-size distribution.

The fixed point of this procedure constitutes the new counterfactual equilibrium.

Local effects The model predicts large effects of this policy on the targeted cities. In targeted cities, the number of establishments grows by 21%. The corresponding local increase in population is, however, only 6%. 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, Mayneris, and Py (2012) on the effect of the French ZFU; Mayer, Mayneris, and Py (2012) 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” 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.

⁴⁰I maintain the subsidy to the cities initially targeted as they grow.

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. Figure 3.8 reports the TFP and welfare in the counterfactual equilibrium, relative to the reference equilibrium, as a function of the gross cost of the subsidy expressed in percentage of GDP. In the model, the subsidy is paid for by lump-sum transfers and there is no shadow cost of public funds; hence, the welfare estimates are upper bounds on the actual effects of the policy.

The simulation show 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 1.6% in TFP in the aggregate, and a loss of 2.2% in welfare. While such a policy allows to decrease congestion overall (see Section 3.2), the welfare gain from decreasing congestion is only 0.25%. It is largely dominated by the TFP effect.⁴¹

I use the counterfactual economy to study the impact of these place-based policies on inequality. I look at two measures of inequality: (1) the Gini coefficient for the distribution of real income across workers (2) a Gini coefficient for the distribution of production outcomes across cities. This second measure summarizes to what extent the fraction of total output produced in each city is distributed (un)equally across cities and therefore captures the inequality in the relative contribution of cities to the aggregate economy. A reason to focus on such a measure is that policy makers may want to smooth out this type of inequality across cities.

Surprisingly, the type of place-based policies I study leads to an increase in spatial inequalities as measured by these two Gini indices (see Panel B of Figure 3.8). The

⁴¹As shown in section 3.2, TFP has a magnified impact on welfare as capital flows in and out of the economy in response to the TFP shock.

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. That larger city grows 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 two Gini coefficients.

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

3.4.2 Land-use regulation

Glaeser and Gottlieb (2008) forcefully argue against policies that limit the growth of cities. In this section, I study one such policy, land-use regulation. At a general level, land-use regulation consists in policies that 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.⁴² A rationale for these restrictions on land-use development is that they may increase the quality of life for existing residents. On the 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 land-use regulation by decreasing the land-use intensity parameter b in the housing production function (equation (3.1)). Decreasing this

⁴²Empirically, Gyourko, Saiz, and Summers (2008) develop an index of differences in the local land-use regulatory climate across US cities based on various measures of the local regulatory environment.

parameter increases the elasticity of housing supply. To quantify the impact of land-use regulation policies, I 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, as estimated by 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 (3.6)), which leads firms in the heterogeneous goods sectors to locate in larger cities. This indirect effect enhances the productive efficiency of these sectors. 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.

Figure 3.9 reports TFP and welfare, relative to the reference equilibrium, for various levels of the housing supply elasticity. An increase in the housing supply elasticity from the 25th to the 75th percentile leads to a 3.1% gain in TFP and a 3.3% indirect gain in welfare.⁴³

The 3.1% gain in TFP can be decomposed into an externality effect, an entry effect and a misallocation effect, as described in Section 3.2. The externality effect

⁴³The aggregate welfare gain, including both the direct and the indirect effects of an increase in housing supply, is 20.5%

stems from the fact that an increase in the housing supply elasticity shifts the spatial equilibrium toward larger cities, that offer larger agglomeration externalities. This increase in productivity drives up firm profits, which in turn attracts more entry, resulting in increased variety for consumers. Combined together, these effects are estimated to lead to a 2.5% increase in aggregate TFP. The remaining effect can be attributed to reduced misallocation in the economy: since the increase in housing supply elasticity flattens out the wage schedule, firms face less distortion in input prices. Factors are allocated to heterogeneous firms in a more efficient way, which leads to an increase in TFP. The effect of reduced misallocation is quantitatively smaller than the externality and the entry effect combined (0.6% vs. 2.5%).

This policy experiment illustrates how increasing housing supply in cities can have positive effects beyond directly reducing congestion costs. They allow a more efficient spatial organization of production in the differentiated goods sectors by endogenously creating agglomeration externalities and reducing the extent of misallocation.

3.5 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. This sorting shapes the productivity of each firm and the amount of agglomeration externalities in the economy. It also shapes the allocation of factors across firms: cities are a medium through which the production of heterogeneous firms is organized within a country. The sorting of firms, mediated by the existence of city developers who act as a coordinating device for the creation of cities, lead 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. This complements the existing literature, which has traditionally focused on estimating local effects of such policies. Using the structure of the model, I estimate the general equilibrium effects of two types of place-based policies. A policy 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 2.2% and does not reduce the observed spatial disparities. 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 3.3% welfare gain through a spatial reorganization of production.

3.6 Tables and figures

Table 3.1: Summary statistics.

	log value added			log employment			N
	mean	p10	p90	mean	p10	p90	
Manufacture of food products and beverages	7.82	6.27	9.81	2.44	1.10	4.09	12,613
Manufacture of textiles	8.30	6.62	10.10	2.91	1.39	4.56	2,741
Manufacture of wearing apparel	7.83	6.13	9.54	2.63	1.10	4.22	2,858
Manufacture of leather goods and footwear, leather tanning	8.10	6.24	10.06	2.90	1.10	4.87	824
Manufacture and products of wood, except furniture	7.94	6.60	9.40	2.52	1.39	3.83	3,454
Manufacture of pulp, paper and paper products	8.85	6.97	11.01	3.22	1.39	5.25	1,214
Publishing, printing and reproduction of recorded media	7.79	6.28	9.52	2.18	0.69	3.69	8,829
Manufacture of chemicals and chemical products	9.17	6.74	11.98	3.24	1.10	5.57	2,342
Manufacture of rubber and plastic products	8.60	6.92	10.44	3.02	1.39	4.77	3,359
Manufacture of glass, ceramic, brick and cement products	8.23	6.65	10.07	2.61	1.10	4.28	2,848
Manufacture of basic metals	9.24	7.26	11.57	3.58	1.79	5.81	793
Manufacture of fabricated metal products, except machinery	8.17	6.81	9.57	2.59	1.39	3.89	15,269
Manufacture of machinery	8.24	6.64	10.09	2.61	1.10	4.34	7,040
Manufacture of office machinery and computers	8.41	6.49	10.76	2.79	1.10	4.75	263
Manufacture of electrical machinery	8.36	6.65	10.55	2.84	1.10	4.85	2,090
Manufacture of radio, television and communication equipment	8.53	6.75	10.49	2.93	1.10	4.82	1,376
Manufacture of medical, precision and optical instruments	8.02	6.59	9.78	2.37	1.10	4.04	3,704
Manufacture of motor vehicles	8.66	6.76	11.25	3.16	1.39	5.46	1,269
Manufacture of other transport equipment	8.25	6.24	10.72	2.77	1.10	4.98	872
Manufacture of furniture	7.72	6.15	9.47	2.36	0.69	3.94	4,413
Recycling	7.88	6.53	9.39	2.23	1.10	3.58	1,237
Information technology services	8.01	6.30	9.97	2.30	0.69	3.97	7,749
Business services, non I.T.	7.65	6.33	9.17	1.89	0.69	3.26	39,915

Table 3.2: Share of establishment in larger cities and tradable capital intensity.

Dep. variable	Share of establishments in large cities			
	I	II	III	IV
<i>Sample</i>		<i>all tradables</i>		<i>export intensive</i>
Tradable capital intensity	0.467** (0.146)	0.591** (0.142)	0.566** (0.162)	0.488** (0.178)
High skill intensity			0.057** (0.029)	0.010 (0.029)
Nb firms	no	yes	yes	yes
Mean va	no	yes	yes	yes
R-squared	0.066	0.220	0.173	0.128
Observations	148	148	118	89

(*) $p < 0.10$, (**) $p < 0.05$. 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. 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 3.3: Tail of the firm-size distribution (ζ) vs sector location.

	ζ , tail of firm-size distribution		
	I	II	III
<i>Sample</i>		<i>all tradables</i>	<i>export intensive</i>
Share in large cities	-0.533** (0.142)	-0.681** (0.105)	-0.263** (0.113)
Nb firms	no	yes	yes
Mean value added	no	yes	yes
R-squared	0.088	0.574	0.487
Observations	148	148	89

(*) $p < 0.10$, (**) $p < 0.05$. Pareto Shape: ζ estimated by $\text{Log}(\text{Rank}_i - 1/2) = a - \zeta \text{Log}(va_i) + \epsilon_i$, on firms above median size, for industries with more than 200 firms. Export intensive: industry above median for all sectors in the economy in export intensity, proxied by the ratio of export to domestic sales.

Table 3.4: Movers.

<i>Sample</i>	Δ_t City Size			
	I <i>all tradables</i>	II	III <i>export intensive</i>	IV
ln(firm size)	0.100** (0.027)	0.081** (0.024)	0.092** (0.030)	0.080** (0.029)
Initial City Size	-0.950** (0.066)		-0.950** (0.059)	
Constant	11.824** (0.776)	-0.243** (0.005)	11.761** (0.688)	-0.290** (0.006)
Initial city F.E.		yes		yes
Cluster	city	city	city	city
R-squared	0.504	0.608	0.505	0.611
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 = $\ln(\frac{C_{t+1}}{C_t})$, where C_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 $\ln(VA)_i = DS_i + DT_i + DC_i + \epsilon_i$ 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 3.5: Estimated parameters.

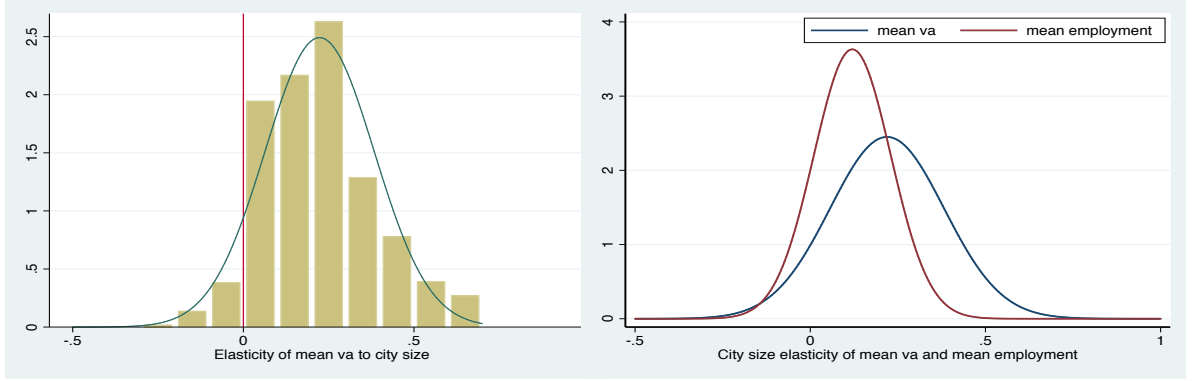
	$\hat{\nu}_z$	$\hat{\nu}_R$	$\hat{a}$	$\hat{s}$
Manufacture of food products and beverages	0.366 (0.009)	0.310 (0.022)	0.025 (0.034)	0.143 (0.033)
Manufacture of textiles	0.340 (0.023)	0.102 (0.016)	0.014 (0.009)	0.130 (0.014)
Manufacture of wearing apparel	0.052 (0.005)	0.166 (0.009)	0.015 (0.001)	0.817 (0.056)
Manufacture of leather goods and footwear, leather tanning	0.241 (0.027)	0.166 (0.014)	0.053 (0.012)	-0.117 (0.015)
Manufacture and products of wood, except furniture	0.287 (0.016)	0.167 (0.010)	0.006 (0.002)	0.047 (0.008)
Manufacture of pulp, paper and paper products	0.339 (0.023)	0.230 (0.025)	0.033 (0.001)	0.091 (0.054)
Publishing, printing and reproduction of recorded media	0.109 (0.004)	0.312 (0.016)	0.012 (0.009)	0.892 (0.073)
Manufacture of chemicals and chemical products	0.939 (0.058)	0.115 (0.055)	0.000 (0.001)	0.071 (0.017)
Manufacture of rubber and plastic products	0.371 (0.052)	0.279 (0.024)	0.052 (0.028)	0.053 (0.042)
Manufacture of glass, ceramic, brick and cement products	0.409 (0.006)	0.348 (0.035)	0.069 (0.034)	-0.179 (0.113)
Manufacture of basic metals	0.187 (0.062)	0.301 (0.028)	0.049 (0.016)	-0.729 (0.031)
Manufacture of fabricated metal products, except machinery	0.067 (0.004)	0.180 (0.005)	0.007 (0.004)	0.798 (0.025)
Manufacture of machinery	0.140 (0.013)	0.208 (0.011)	-0.005 (0.024)	0.512 (0.135)
Manufacture of office machinery and computers	0.095 (0.053)	0.288 (0.024)	0.236 (0.052)	0.095 (0.170)
Manufacture of electrical machinery	0.148 (0.009)	0.214 (0.016)	0.063 (0.010)	0.433 (0.040)
Manufacture of radio, television and communication equipment	0.073 (0.014)	0.254 (0.015)	0.025 (0.014)	0.899 (0.257)
Manufacture of medical, precision and optical instruments	0.128 (0.014)	0.276 (0.014)	0.100 (0.007)	0.441 (0.068)
Manufacture of motor vehicles	0.159 (0.042)	0.289 (0.021)	0.099 (0.010)	0.258 (0.053)
Manufacture of other transport equipment	0.172 (0.078)	0.236 (0.024)	-0.033 (0.021)	0.634 (0.036)
Manufacture of furniture	0.225 (0.011)	0.240 (0.013)	0.022 (0.021)	0.271 (0.070)
Recycling	0.271 (0.019)	0.377 (0.025)	0.018 (0.012)	0.505 (0.078)
Information technology services	0.544 (0.017)	0.096 (0.008)	0.093 (0.013)	0.090 (0.031)
Business services, non I.T.	0.130 (0.001)	0.255 (0.006)	0.098 (0.013)	0.673 (0.041)

Table 3.6: Elasticity of firm productivity to city size: decompositions.

	(I) With sorting	(II) Without sorting	(III) at 25th perc.	(IV) Systematic Component at median	(V) at 75th perc.
Manufacture of food products and beverages	0.04	0.03	0.03	0.04	0.05
Manufacture of textiles	0.05	0.02	0.02	0.03	0.03
Manufacture of wearing apparel	0.04	0.02	0.03	0.05	0.07
Manufacture of leather goods and footwear, leather tanning	0.03	0.03	0.05	0.05	0.04
Manufacture and products of wood, except furniture	0.03	0.03	0.01	0.01	0.01
Manufacture of pulp, paper and paper products	0.03	0.02	0.04	0.04	0.05
Publishing, printing and reproduction of recorded media	0.08	0.01	0.06	0.09	0.14
Manufacture of chemicals and chemical products	0.10	0.01	0.01	0.02	0.03
Manufacture of rubber and plastic products	0.03	0.02	0.05	0.06	0.06
Manufacture of glass, ceramic, brick and cement products	0.01	0.03	0.06	0.05	0.04
Manufacture of basic metals	0.02	0.02	0.09	0.09	0.08
Manufacture of fabricated metal products, except machinery	0.05	0.02	0.03	0.05	0.07
Manufacture of machinery	0.05	0.02	0.01	0.03	0.05
Manufacture of office machinery and computers	0.03	0.03	0.24	0.24	0.24
Manufacture of electrical machinery	0.04	0.02	0.08	0.09	0.10
Manufacture of radio, television and communication equipment	0.06	0.02	0.05	0.08	0.11
Manufacture of medical, precision and optical instruments	0.04	0.03	0.11	0.12	0.14
Manufacture of motor vehicles	0.03	0.02	0.11	0.11	0.12
Manufacture of other transport equipment	0.08	0.02	0.00	0.03	0.07
Manufacture of furniture	0.04	0.02	0.03	0.04	0.06
Recycling	0.07	0.01	0.05	0.08	0.12
Information technology services	0.06	0.02	0.10	0.10	0.11
Business services, non I.T.	0.07	0.02	0.13	0.15	0.18

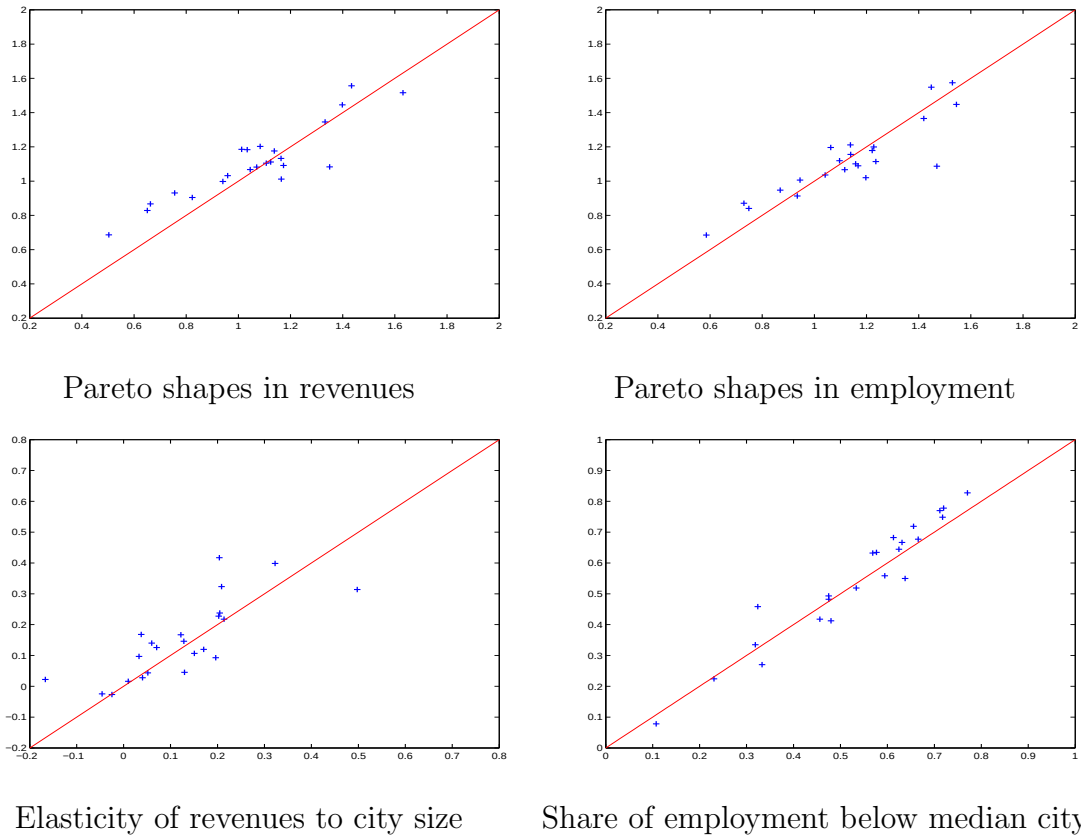
Notes. The elasticity of productivity to city size is computed as the regression coefficient β_j for sector j in the firm-level regression $\log Prod_{i,j} = \beta_0 + \beta_j \log L_i + \nu_i$ run industry by industry, where L_i is the size of the city chosen by firm i and $Prod_{i,j}$ is a measure of the productivity of firm i in sector j evaluated as follows. In Specification (I), $Prod_{i,j}$ is the productivity of firm i in sector j from the baseline model where heterogeneous firms sort across city sizes according to their raw efficiency. In Specification (II), $Prod_{i,j}$ is the productivity of firm i in sector j from a counterfactual exercise where heterogeneous firms do not sort across city sizes according to their raw efficiency. In Specifications (III-V), $Prod_{i,j}$ measures the “systematic” component of firm i productivity, and the regression is run only on firms that share the same raw efficiency z . In Specification (III),(IV), (V) z is, respectively, at the 25th, 50th and 75th percentile of the sectoral distribution of raw efficiencies.

Figure 3.1: Elasticity of mean value added and employment with city size.



Note: Histogram by firm of β in the regression: $\log \text{mean va}(L_j) = \alpha + \beta \log L_j + \epsilon_j$ (resp. $\log \text{mean empl}(L_j)$), ran sector by sector at the NAF600 level for industries with more than 200 mono-establishment firms.

Figure 3.2: Subset of targeted moments, data and model.



Note: The blue points correspond to sectoral estimates for each moment, plotted against the sectoral values in the data. The red line is the 45° line.

Figure 3.3: Sectoral distribution of firms revenues, model (blue) and data (red).

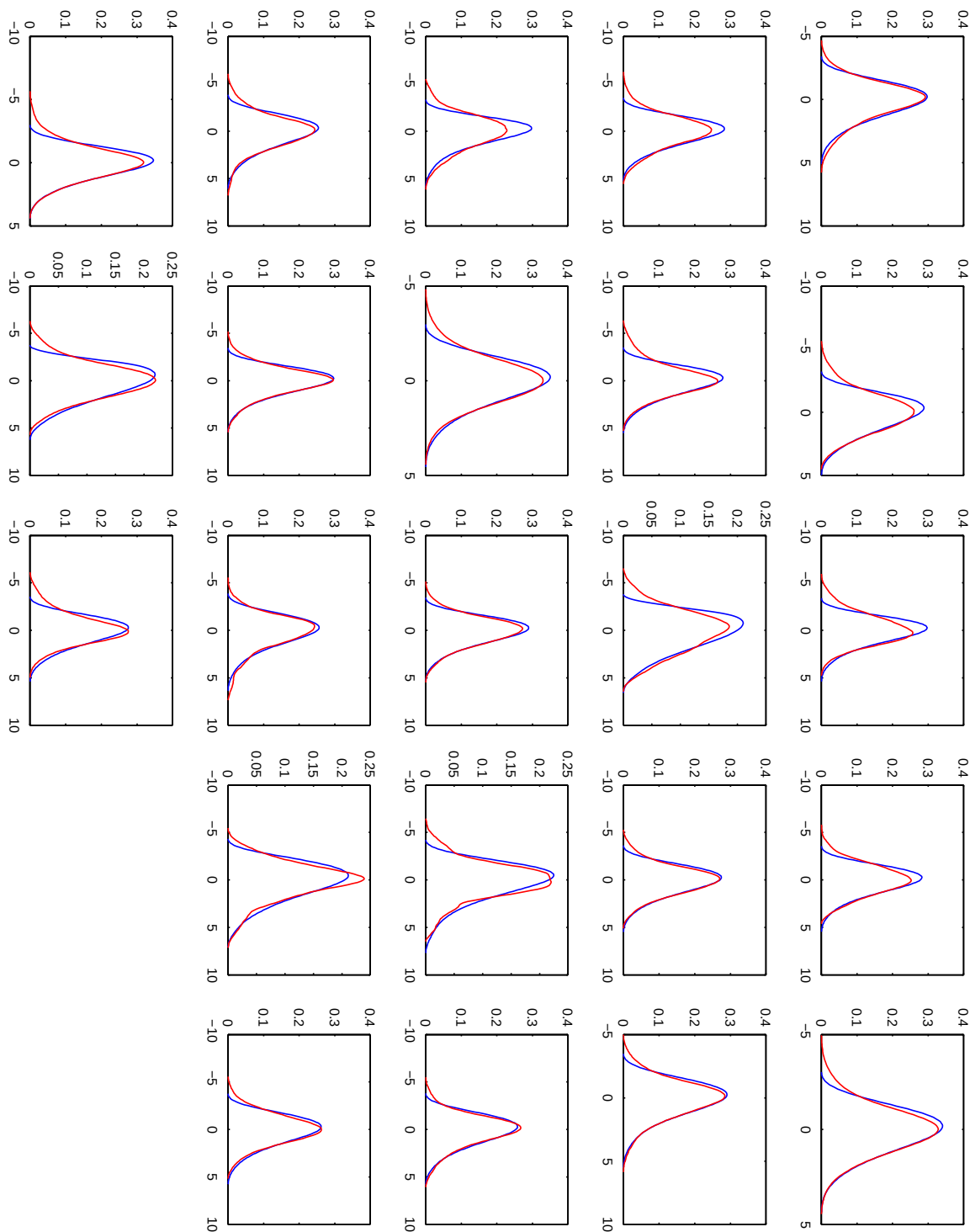


Figure 3.4: Employment share by decile of city size, model (blue) and data (red).

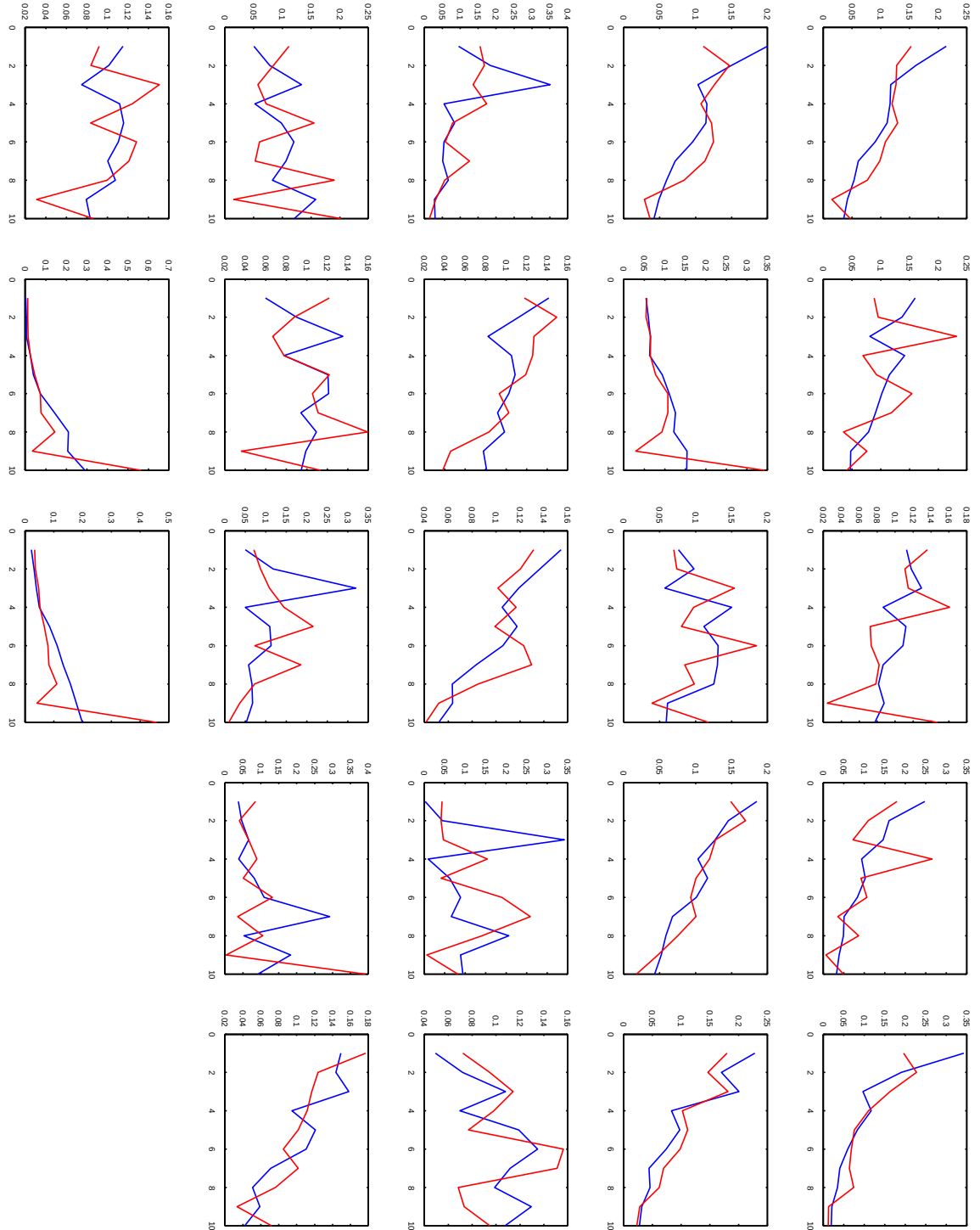
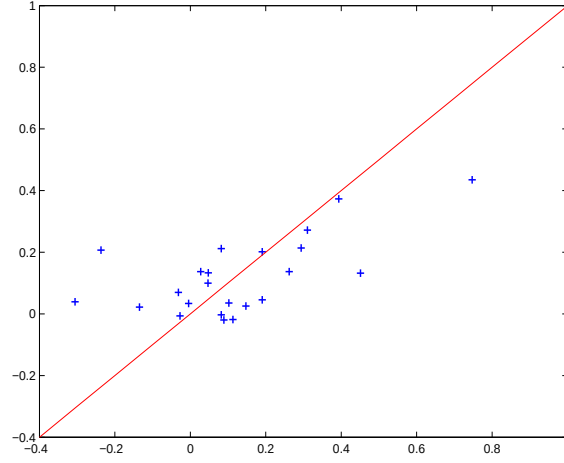


Figure 3.5: Dilation in firm-size distribution in revenues, large cities vs small cities.



Note: For each sector j , $d_j = \log(\frac{q_{75}}{q_{25}})_{large} - \log(\frac{q_{75}}{q_{25}})_{small}$, where 'large' (resp. 'small') indexes the set of firms that locate above (resp. below) the median threshold of city size. $q(p)$ indicates the p th quantile of revenue distribution. The graph plots this dilation measure $\hat{d}_j$ from the estimated model against d_j in the data.

Figure 3.6: City size distribution, model and data.

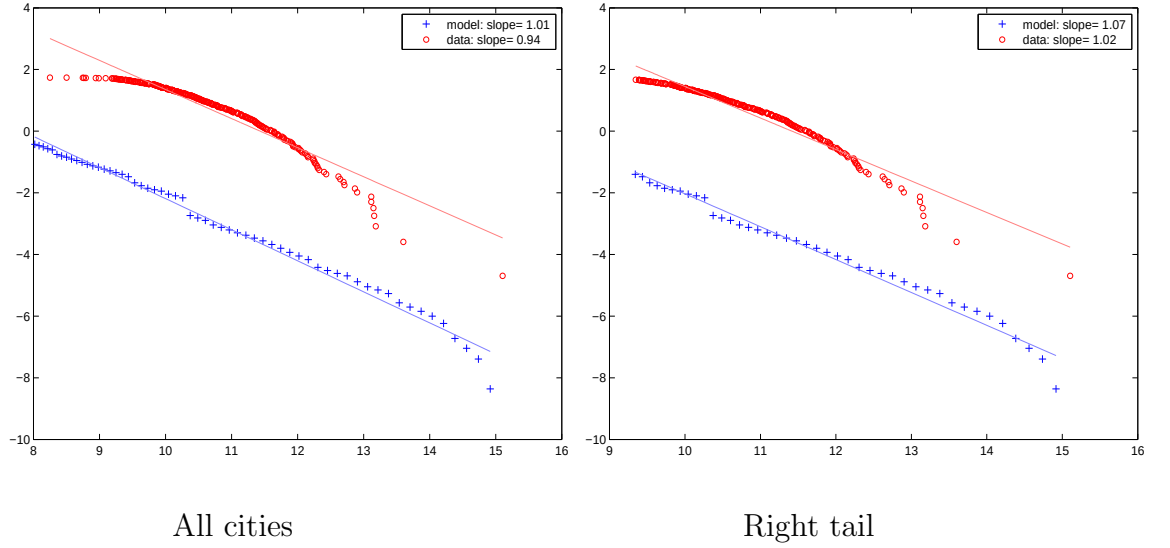
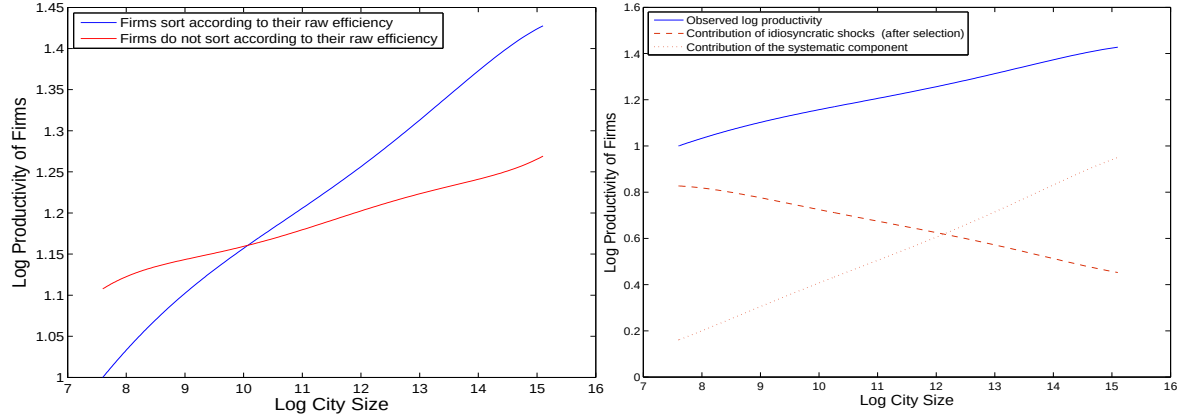
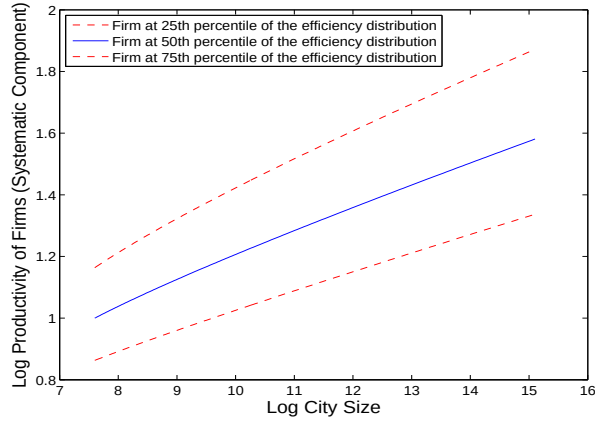


Figure 3.7: Elasticity of productivity with respect to city size: decompositions.



A. Productivity and city size, with and without firm sorting

B. Productivity and city size, systematic and idiosyncratic components



C. Systematic productivity and city size, for three levels of firm efficiency

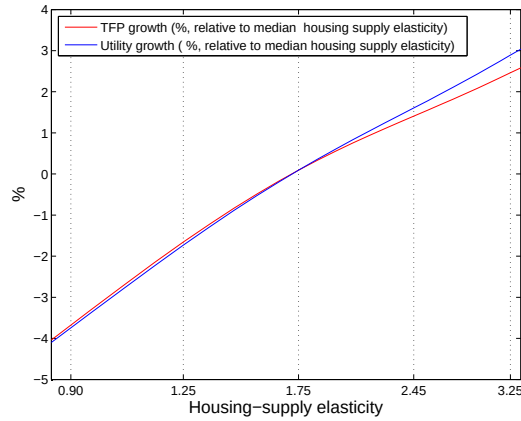
Note: these plots represent averages across sectors. Lines are smoothed for readability. The simulated data have a variance around this line as there is not a unique productivity level for each city-size-sector pair.

Figure 3.8: Aggregate impact of local subsidies.



Note: Firms profits are subsidized when they locate in the smallest cities of the reference equilibrium. The targeted area represents 2.3% of the population. The policy is financed by a lump-sum tax on firms.

Figure 3.9: TFP and indirect welfare effects of increasing housing-supply elasticity.



The horizontal axis measures housing supply elasticity in the economy $\frac{d \log H}{d \log p_H}$. Saiz (2010) reports that median elasticity of housing supply is 1.75, the 25th percentile is at 2.45 and the 75th percentile at 1.25.

3.7 Appendix

3.7.1 Housing market

Housing supply (equation (3.2)) and demand (equation (3.4)) equate so that $p_H(L) = (1 - \eta)^b w(L) L^b$. This yields the following labor use and profits in the housing sector:

$$\ell_H(L) = (1 - b)(1 - \eta)L, \quad \text{and} \quad (3.30)$$

$$\pi_H(L) = b(1 - \eta)Lw(L). \quad (3.31)$$

The housing supply elasticity is given by $\frac{d \log H(L)}{d \log p_H(L)} = \eta \frac{1-b}{b}$.

Anticipating on the policy discussion, note that a decrease in b increases the housing supply elasticity and also leads to a flatter wage schedule across city sizes, as $\frac{d \log w(L)}{d \log L} = b \frac{1-\eta}{\eta}$.

3.7.2 Extensions of the model

Log supermodular productivity function

I adapt the model of Davis and Dingel (2012) who microfound complementarity between entrepreneur skill and city size to a setting where entrepreneurs hire workers to produce output. Workers supply one efficiency unit of labor. They can also spend a share ψ of their time outside of the firm to gather information that will make them better at producing the firm blueprint z , such as information on efficient processes of production, or on the appeal of a product design. The return to this activity is $z \phi f(L)$ where the return increases in city size through $f(L)$: information gathering is more productive when there is more local information available. As a result of this process, the worker's amount of labor in efficiency units is: $g(z, \phi, L) = (1 - \phi)(1 + z \phi f(L))$. The firm chooses the time spent on productive activity $1 - \phi$ by maximizing their profit, given their production function $Y = z k^\alpha (\ell g(z, \phi, L))^{1-\alpha}$. As the firm profit function is multiplicatively separable in $g(z, \phi, L)$, they choose optimally ϕ to maximize this term. This yields an optimized $g^*(z, L) = \frac{1}{f(L)z} (f(L)z + 1 - \alpha)(1 + f(L)z)^{1-\alpha}$ at the optimal ϕ^* , which can be shown

after some algebra to be log-supermodular in z and L as long as $f(L)$ increases with L ($\frac{\partial^2 \log g^*(z, L)}{\partial z \partial L} > 0$).

Model with imperfect sorting

Let $p(z, L_i; j)$ the probability that a firm of type z in sector j chooses city size L_i against other cities L_k for $k = 1, \dots, n$. As the idiosyncratic shocks to productivity follows a type I extreme value distribution,

$$p(z, L; j) = \frac{\pi(z, L_i, j)}{\sum_{k=1}^n \pi(z, L_k, j)}.$$

It is readily seen that p inherits the log-supermodularity of π in (z, L, j) (abusing notations, j standing for s_j or α_j). Therefore the probability distribution of establishments in a high j vs a low j sector follows the monotone likelihood ratio property. First order stochastic dominance follows.

3.7.3 Proofs

Lemma 15

Proof Consider a given city of size L developed by city developer i . Let

$$N_j(L, i) = \int_z \ell_j(z, L) \mathbb{1}(z, L, i) M_j dF_j(z)$$

denote the number of workers working in sector j in this specific city i . Equation (3.8) shows that, for a given city size and a given sector, labor hired by local firms is proportionate to the ratio of firms profit to the local wage. The total local profits of firms in sector j is therefore proportional to the sectoral wage bill $w(L)N_j$. The city developers problem (3.11) then simplifies to

$$\max_{L, \{T_j(L)\}_{j=1, \dots, S}} \Pi_L = b(1 - \eta)w(L)L - \sum_{j=1}^S T_j(L) \frac{w(L)N_j}{(1 - \alpha_j)(\sigma_j - 1)} \quad (3.32)$$

Free entry pushes the profit of city developers to zero in equilibrium, which drives $T_j(L)$ to T_j^* . The problem is akin to a Bertrand game. Consider a given city size L . First, any deviation from T_j^* downwards leads a city developer to loose all firms from the corresponding sector, as new city developers could offer T_j^* for sector j and 0 for all other sectors, attract all firms for sector j for whom this subsidy is more attractive, and make exactly zero profit. A city L that attracts only firms from sector j attracts $N_j = L - \ell_H(L) = L(1 - (1 - \eta)(1 - b))$ workers from sector j , where $\ell_H(L)$ is the labor force hired in the construction sector and the second equality stems from equation (3.10). It is readily seen from this expression that the subsidy T_j^* for sector j and 0 for all other sectors leads to zero profits. Second, any deviation from T_j^* upwards in any sector j leads to negative profits. To see this, consider all cities of size L , and take the one that offers the highest subsidy city to firms in sector j . Call this city i , and assume that $T_j^i > T_j^*$. From the first step of the proof, we know that in any given city, for all sectors k , either $T_k \geq T_k^*$ and $N_k \geq 0$ or $T_k < T_k^*$ and $N_k = 0$. City developer profits in city i are therefore :

$$\begin{aligned}\Pi^i &= b(1 - \eta)w(L)L - \sum_{k=1}^S T_k^i \frac{w(L)N_k^i}{(1 - \alpha_k)(\sigma_k - 1)} \\ &< b(1 - \eta)w(L)L - \sum_{k=1}^S T_k^* \frac{w(L)N_k^i}{(1 - \alpha_k)(\sigma_k - 1)} \\ &< b(1 - \eta)w(L)L - \frac{b(1 - \eta)}{1 - (1 - \eta)(1 - b)} w(L) \left(\sum_{j=1}^S N_k \right) < 0\end{aligned}$$

where the last inequality comes from $\sum_{j=1}^S N_k = L - \ell_H(L)$ and (3.10). ■

Lemma 16

Proof Let L_o denote the suboptimal city size where firms of type (z, j) are located. They get profit $\pi_j^*(z, L_o)$. Denote $\Delta = \frac{\pi_j^*(z_j, L^*(z))}{\pi_j^*(z, L_o)} - 1 > 0$. A city developer can open a city of size $L^*(z)$ by offering a subsidy $\tilde{T}_j = \frac{1 + \frac{\Delta}{2}}{1 + \Delta} (1 + T_j^*) - 1$, which will attract firms as they make a higher profit than at L_o , and allows the city developer to make positive profits. City size distribution adjusts in equilibrium to determine the number of such cities. ■

Lemma 17

Proof Fix s . Since $\pi(z, L, s)$ is strictly LSM in (z, L) , it follows that for all $z_1 > z_2$ and $L_1 > L_2$, $\frac{\pi(z_1, L_1, s)}{\pi(z_1, L_2, s)} > \frac{\pi(z_2, L_1, s)}{\pi(z_2, L_2, s)}$. So if z_2 has higher profits in L_1 than in L_2 , so does z_1 . Necessarily, $L^*(z_1) \geq L^*(z_2)$.

Moreover, under the technical assumptions made here, $L_j^*(z)$ is a strictly increasing function. Since the set of z is convex, and $\psi(z, L, s)$ is such that the profit maximization problem is concave for all firms, the optimal set of city sizes is itself convex. It follows that $L_j^*(z)$ is invertible. It is locally differentiable (using in addition that $\psi(z, L, s)$ is differentiable), as the implicit function theorem applies and $\frac{dL_j^*(z)}{dz} = -\frac{\frac{\partial(\frac{\psi_2 L}{\psi})}{\partial z}(z, L_j^*(z), s)}{\frac{\partial(\frac{\psi_2 L}{\psi})}{\partial L}(z, L_j^*(z), s)}$. ■

Lemma 18

The proof is in the main text.

City size distribution

I establish here the following characterization of the city-size distribution.

Characterization 23 *If the firm size distribution in revenues follow Zipf's law, a sufficient condition for the city size distribution to follow Zipf's law is that revenues have constant elasticity with respect to city size in equilibrium, under realistic parameter values.*

When the productivity function is $\log(\psi_j(z, L, s_j)) = a_j \log L + \log(z)(1 + \log \frac{L}{L_o})^{s_j}$, equilibrium firm-level revenues have constant elasticity with respect to city size in equilibrium.

Proof There is a bijection between z and firm level observable at equilibrium r^*, ℓ^*, L^* . By an abuse of notation, this functional relationship will be denoted $z(r), z(L), \ell(z)$. All of these relationships pertain to the sorting equilibrium, but I omit the star to keep the notations light.

Step 1: Write $g_j(r)$ the distribution of revenues in sector j . The firm-size distribution in revenues is readily computed through a change of variable, starting from the raw efficiency

distribution:

$$g_j(r) = f_j(z(r)) \frac{dz}{dr}(r) \quad (3.33)$$

Step 2: Assume that in all sectors - or at least for the ones that will locate in the largest cities in equilibrium -, firm-size distribution in revenues is well approximated by a Pareto distribution with Pareto shape ζ_j , close to 1. Write $g_j(r)$ the distribution of revenues in sector j :

$$\exists r_{j,o} : \forall r > r_{j,o}, g_j(r) \propto r^{-\zeta_j-1} \quad (3.34)$$

Step 3: The city size distribution $f_L(\cdot)$ as defined by equation 3.20 can be written as:

$$f_L(L) = \sum_{j=1}^S M_j \tilde{f}^j(L)$$

where $\tilde{f}^j(L) = \frac{\ell_j(z_j^*(L)) f_j(z_j^*(L)) \frac{dz_j^*(L)}{dL}}{L}$. Then,

$$\begin{aligned} \tilde{f}^j(L) &\propto g_j(r^*(L)) \frac{dr_j^*}{dL} \frac{r_j(L)}{L^{b_{\frac{\eta}{1-\eta}}+1}} \\ &\propto \frac{r_j(L)^{-\zeta_j+1}}{L^{b_{\frac{\eta}{1-\eta}}+2}} \end{aligned}$$

The first line uses (3.33), $\ell_j \propto \frac{r_j(L)}{w(L)}$ and $w(L) \propto L^{b_{\frac{\eta}{1-\eta}}}$. The second line uses the assumption that r^* has constant elasticity with respect to L at the sorting equilibrium, and (3.34). As $\zeta \sim 1$,

$$\tilde{f}^j(L) \propto \frac{1}{L^{b_{\frac{\eta}{1-\eta}}+2}} \quad \text{and} \quad f_L(L) \propto \frac{1}{L^{b_{\frac{\eta}{1-\eta}}+2}}.$$

Finally, $b_{\frac{\eta}{1-\eta}} < 1$, as it measures the elasticity of wages with city size, on the order of magnitude of 5%. The tail of the city size distribution is well approximated by a Pareto of shape close to 1. City size distribution follows Zipf's law. ■

Proposition 19

Proof Fix j . For productivity, the results comes from the facts that (1) $L_j^*(z)$ is non decreasing in z and (2) that $\psi(z, L, s_j)$ is increasing in L . Revenues are proportional to profits ($r_j * (z) = \frac{\sigma_j}{1+T_j^*} \pi_j^*(z)$). The proof for profits is as follows. $\psi(z_H, L_L, s_j) > \psi(z_L, L_L, s_j)$ as ψ is increasing in z , which leads to $\pi(z_H, L_L) > \pi(z_L, L_L)$, as firms face the same wage in the same city. Finally, $\pi(z_H, L_H, s_j) \geq \pi(z_H, L_L, s_j)$ as L_H is the profit maximizing choice for z_H . Therefore, $\pi(z_H, L_H, s_j) > \pi(z_L, L_L, s_j)$. ■

In addition, $\epsilon_l = \epsilon_r - (1 - \alpha) \frac{1-\eta}{\eta}$.

Proof For a given city size L and a given sector j , $\bar{r}_j^*(L) = \sum_{z \text{ in } L} r_j^*(z) \propto \sum_{z \text{ in } L} w(L) \ell_j^*(z) \propto w(L) \bar{\ell}_j^*(L)$, where their proportion is constant across city sizes. Therefore $\frac{d \log \bar{\ell}_j^*(L)}{d \log L} = \frac{d \log \bar{r}_j^*(L)}{d \log L} - \epsilon_w$, where the elasticity of wages with respect to city sizes is $\epsilon_w = b \frac{\eta}{1-\eta}$. ■

Proposition 20

Proof The proof here covers both the case of the main assumptions of the model (continuity and convexity of the support of z and L), and the case where the set of city sizes is exogenously given, and in particular discrete. Let $\mathcal{Z} : \mathcal{L} \times A \times E \rightarrow Z$ be the correspondence that assigns to any $L \in \mathcal{L}$ and $\alpha \in A$ a set of z that chooses L at equilibrium. (It is a function under the assumptions made in the main text (see proof of Lemma 17).) Define $\bar{z}(L, \alpha, s) = \max_z \{z \in \mathcal{Z}(L, \alpha, s)\}$ as the maximum efficiency level of a firm that chooses city size L in a sector characterized by the parameters (α, s) . I will use the following lemmas:

Lemma 24 *$\log \pi$ is supermodular with respect to the triple (z, L, α)*

It is readily seen that: $\frac{\partial^2 \log \pi(z, L, \alpha, s)}{\partial z \partial L} > 0$, $\frac{\partial^2 \log \pi(z, L, \alpha, s)}{\partial z \partial \alpha} = 0$ and $\frac{\partial^2 \log \pi(z, L, \alpha, s)}{\partial L \partial \alpha} = \frac{(\sigma-1)b(1-\eta)}{\eta L} > 0$. This result does not rely on an assumption on the convexity of $\mathcal{L}$. Checking the cross partials are sufficient to prove the supermodularity even if L is taken

from a discrete set, as π can be extended straightforwardly to a convex domain, the convex hull of $\mathcal{L}$.

Lemma 25 $\bar{z}(L, \alpha, s)$ is non decreasing in α, s .

The lemma is a direct consequence of the supermodularity of $\log \pi$ with respect to the quadruple (z, L, α, s) . Using a classical theorem in monotone comparative statics, if $\log \pi(z, L, \alpha, s)$ is supermodular in (z, L, α, s) , and $L^*(z, \alpha, s) = \max_L \log \pi(z, L, \alpha, s)$ then $(z_H, \alpha_H, s_H) \geq (z_L, \alpha_L, s_L) \Rightarrow L^*(z_H, \alpha_H, s_H) \geq L^*(z_L, \alpha_L, s_L)$. Note that everywhere, the $\geq$ sign denotes the lattice order on $\mathcal{R}^3$ (all elements are greater or equal than).

Coming back to the proof of the main proposition, we can now write:

$$\begin{aligned} \tilde{F}(L; \alpha, s) &= P(\text{firm from sector}(\alpha, s) \text{ is in a city of size smaller than } L) \\ &= F(\bar{z}(L, \alpha, s)) \end{aligned}$$

where $F(\cdot)$ is the raw efficiency distribution of the firms in the industry. Let $\alpha_H > \alpha_L$.

For any $z \in Z$, the previous lemma ensures that $L^*(z, \alpha_H, s) \geq L^*(z, \alpha_L, s)$. In particular, fix a given L and s and write using shorter notation: $\bar{z}_{\alpha_L} = \bar{z}(L, \alpha_L, s)$. Then $L^*(\bar{z}_{\alpha_L}, \alpha_H, s) \geq L^*(\bar{z}_{\alpha_L}, \alpha_L, s) = L$. Because $L^*(z, \alpha_H, s)$ is increasing in z , it follows that:

$$z \in \mathcal{Z}(L, \alpha_H, s) \Rightarrow z \leq \bar{z}_{\alpha_L}$$

and therefore $\bar{z}_{\alpha_H} \leq \bar{z}_{\alpha_L}$ or using the long notation: $\bar{z}(L, \alpha_H, s) \leq \bar{z}(L, \alpha_L, s)$

It follows that $F(\bar{z}(L, \alpha_H)) \leq F(\bar{z}(L, \alpha_L))$ and that $F(L; \alpha, s)$ is decreasing in α . This completes the proof of the first order stochastic dominance of the geographic distribution of a high α sector vs that of a lower α .

The proof is exactly the same for the comparative static in s , we just have to verify that $\pi(z, L, s)$ is log supermodular in (z, L, s) . Since $\pi(z, L, s_j) = \kappa \left(\frac{\psi(z, L, s_j)}{w(L)^{1-\alpha}} \right)^{\sigma-1} \frac{R_j}{P_j^{1-\sigma}}$ and $w(L)$ doesn't depend on s , $\pi(z, L, s)$ directly inherits the log supermodularity of $\psi(z, L, s)$ in its parameters. ■

Proposition 21

Proof The proof here covers both the case of the main assumptions of the model (continuity and convexity of the support of z and L), and the case where the set of city sizes is exogenously given, and in particular discrete. Within sectors, the revenue function $r_j^*(z)$ at the sorting equilibrium is an increasing function for any j . Let $p_1 < p_2 \in (0, 1)$. Under the assumption, maintained throughout the comparative static exercise, that sectors draw z from the same distribution, there $\exists z_1 < z_2$ such that $Q_{j_1}(p_1) = r_{j_1}^*(z_1)$ and $Q_{j_2}(p_1) = r_{j_2}^*(z_1)$ (same thing for z_2 and p_2), ie. the quantiles of the $r_{j_1}^*$ and $r_{j_2}^*$ distributions correspond to the same quantile of the z distribution. This yields $\frac{Q_{j_1}(p_2)}{Q_{j_1}(p_1)} = \frac{r_{j_1}^*(z_2)}{r_{j_1}^*(z_1)}$, and $\frac{Q_{j_2}(p_2)}{Q_{j_2}(p_1)} = \frac{r_{j_2}^*(z_2)}{r_{j_2}^*(z_1)}$.

Finally, it is a classic result in monotone comparative statics (Topkis (1998)) that if $\pi(z, L, \alpha)$ is log-supermodular in (z, L, α) , then $\pi^*(z, \alpha) = \max_L \pi(z, L, \alpha)$ is log supermodular in (z, α) , or $\frac{\pi_{j_2}^*(z_2)}{\pi_{j_2}^*(z_1)} \geq \frac{\pi_{j_1}^*(z_2)}{\pi_{j_1}^*(z_1)}$. Revenues are proportional to profits within sectors, which completes the proof. The same proof applies for s . ■

Corollary 22

Proof Let $p_j \in (0, 1)$ be a threshold above which the distribution is well approximated by a Pareto distribution in sector j , and r_j the corresponding quantile of the distribution. The distribution of r conditional on being larger than r_j is:

$$\forall r > r_j, H_j(r | r \geq r_j) \approx 1 - \left(\frac{r}{r_j}\right)^{-\zeta_j},$$

where ζ_j is the shape parameter of the Pareto distribution for sector j . Thus, if $F_j(r) = p$, one can write:

$$\begin{aligned} \forall p > p_j, \quad p &\approx F_j(r_j) + H_j(r) \approx p_j + 1 - \left(\frac{r}{r_j}\right)^{-\zeta_j} \\ \frac{r}{r_j} &\approx (1 + p_j - p)^{-\frac{1}{\zeta_j}} \end{aligned}$$

Letting $p_0 = \max(p_1, p_2)$ and writing $r_j = Q_j(p_0)$ for $j = 1, 2$, and using proposition (21) gives:

$$\begin{aligned} \frac{Q_{j1}(p)}{Q_{j1}(p_0)} &\leq \frac{Q_{j2}(p)}{Q_{j2}(p_0)} \\ (1 + p_0 - p)^{-\frac{1}{\zeta_1}} &\leq (1 + p_0 - p)^{-\frac{1}{\zeta_2}} \quad \text{for all } p > p_0 \text{ and } p < 1 \\ \zeta_1 &\geq \zeta_2, \end{aligned}$$

where the last inequality comes from $1 + p_0 - p \in (0, 1)$. ■

3.7.4 General Equilibrium, TFP and welfare

General equilibrium quantities

Total manufacturing revenues The national labor market clearing condition (3.19) together with equation (3.18) leads to the aggregate revenues in manufacturing,

$$R = N_o \frac{(1 - (1 - b)(1 - \eta))}{\sum_{j=1}^S \xi_j \frac{E_j}{S_j}}. \quad (3.35)$$

Price index and mass of firms Combining equations (3.18) and (3.17) leads to the sectoral mass of firms⁴⁴

$$M_j = \frac{\sigma_j \xi_j}{f_{Ej}(1 + T_j^*)} \frac{R}{P}. \quad (3.36)$$

This equation is still valid when a federal government subsidizes firm's profits depending on their location choice, as the policy is financed for by a lump sum tax on all firms profits.

Using equations (3.36) and (3.17)

$$P_j^{\sigma_j - 1} = \frac{1}{\sigma_j \frac{\kappa_{1j}}{\kappa_{3j}} M_j S_j} \propto \frac{P^{1 + \alpha_j(\sigma_j - 1)}}{R S_j}$$

so that

⁴⁴The price of capital ρ is exogenously given on international market. As I do not choose the price index P as the numeraire, ρ is not fixed in absolute value but relative to P .

$$P^{-1} = \prod_{j=1}^S \left(\frac{P_j}{\xi_j} \right)^{-\xi_j} \propto \left(\left(\sum_{j=1}^S \xi_j \frac{E_j}{S_j} \right)^{-\frac{1}{\bar{\sigma}-1}} \prod_{j=1}^S S_j^{\frac{\xi_j}{\bar{\sigma}-1}} \right)^{\frac{\bar{\sigma}-1}{(1-\bar{\alpha})(\bar{\sigma}-1)-1}}$$

where $\bar{\alpha} = \sum_{j=1}^S \alpha_j \xi_j$ and $\frac{1}{\bar{\sigma}-1} = \sum_{j=1}^S \frac{\xi_j}{\sigma_j-1}$. Hence, $M_j \propto \left(\left(\sum_{j=1}^S \xi_j \frac{E_j}{S_j} \right)^{1-\bar{\alpha}} \prod_{j=1}^S S_j^{\frac{\xi_j}{\bar{\sigma}-1}} \right)^{\frac{\bar{\sigma}-1}{(1-\bar{\alpha})(\bar{\sigma}-1)-1}}$.

TFP and welfare

TFP Let Y_j , N_j and K_j denote the sectoral output, employment and capital. Sectoral productivity $TFP_j \equiv \frac{Y_j}{K_j^{\alpha_j} N_j^{1-\alpha_j}}$ can be expressed as follows.

$$\begin{aligned} Y_j &= M_j^{\frac{1}{\sigma_j-1}} \left(\int (\psi(u, L_j^*(u), s_j) k_j^*(u)^{\alpha_j} \ell_j^*(u)^{1-\alpha_j})^{\frac{\sigma_j-1}{\sigma_j}} dF_j(u) \right)^{\frac{\sigma_j}{\sigma_j-1}} \\ &= K_j^{\alpha_j} N_j^{1-\alpha_j} M_j^{\frac{1}{\sigma_j-1}} \left(\int (\psi(u, L_j^*(u), s_j) o_j^k(u)^{\alpha_j} o_j^\ell(u)^{1-\alpha_j})^{\frac{\sigma_j-1}{\sigma_j}} dF_j(u) \right)^{\frac{\sigma_j}{\sigma_j-1}}, \end{aligned} \quad (3.37)$$

where

$$\begin{aligned} o_j^\ell(z) &= \frac{\ell_j^*(z)}{N_j} = \frac{\frac{\psi(z, L_j^*(z), s_j)^{\sigma_j-1}}{w(L_j^*(z))^{(1-\alpha_j)(\sigma_j-1)+1}}}{\int \left(\frac{\psi(u, L_j^*(u), s_j)^{\sigma_j-1}}{w(L_j^*(u))^{(1-\alpha_j)(\sigma_j-1)+1}} \right) dF_j(u)}, \text{ and} \\ o_j^k(z) &= \frac{k_j^*(z)}{K_j} = \frac{\frac{\psi(z, L_j^*(z), s_j)^{\sigma_j-1}}{w(L_j^*(z))^{(1-\alpha_j)(\sigma_j-1)}}}{\int \left(\frac{\psi(u, L_j^*(u), s_j)^{\sigma_j-1}}{w(L_j^*(u))^{(1-\alpha_j)(\sigma_j-1)}} \right) dF_j(u)}. \end{aligned} \quad (3.38)$$

After some algebra, this leads to

$$\begin{aligned} TFP_j &= M_j^{\frac{1}{\sigma_j-1}} \frac{\left(\int \frac{\psi(u, L_j^*(u), s_j)^{\sigma_j-1}}{w(L_j^*(u))^{(1-\alpha_j)(\sigma_j-1)+1}} dF_j(u) \right)^{\frac{\sigma_j}{\sigma_j-1}-\alpha_j}}{\left(\int \frac{\psi(u, L_j^*(u), s_j)^{\sigma_j-1}}{w(L_j^*(u))^{(1-\alpha_j)(\sigma_j-1)+1}} dF_j(u) \right)^{1-\alpha_j}} \\ TFP_j &= M_j^{\frac{1}{\sigma_j-1}} \frac{S_j^{\frac{\sigma_j}{\sigma_j-1}-\alpha_j}}{E_j^{1-\alpha_j}}, \end{aligned} \quad (3.39)$$

where S_j and E_j are defined in the main text. Sectoral TFP only depends on the matching function $L_j^*(z)$ and the wage schedule elasticity with respect to city size. To see this,

consider the wage schedule defined by equation 3.6. The wage level $\bar{w}(1 - \eta)^{b\frac{1-\eta}{\eta}}$ does not enter expressions (3.38). The absolute wage level cancels out in the expression of the ratios $o_j^\ell(z)$ and $o_j^k(z)$. It follows that the absolute wage level does not impact equations (3.37) or (3.39).

Decomposing Welfare

$$TFP_j = M_j^{\frac{1}{\sigma_j-1}} \left(\frac{S_j}{E_j} \right)^{1-\alpha_j} S_j^{\frac{1}{\sigma_j-1}}$$

Given equation (3.36) and the expression for the price index,

$$P_j^{-1} = TFP_j \left(\frac{E_j}{S_j} \right)^{1-\alpha_j} \frac{1}{P^{\alpha_j}}$$

$$P^{-1} \propto \left(\prod_{j=1}^S TFP_j^{\xi_j} \left(\frac{E_j}{S_j} \right)^{\xi_j(1-\alpha_j)} \right)^{\frac{1}{1-\bar{\alpha}}}$$

Bibliography

- ABDEL-RAHMAN, H., AND A. ANAS (2004): “Theories of systems of cities,” *Handbook of Regional and Urban Economics*, 4, 2293–2339.
- ALBOUY, D. (2012): “Evaluating the efficiency and equity of federal fiscal equalization,” *Journal of Public Economics*, 96(9-10), 824–839.
- ARKOLAKIS, C. (2011): “A unified theory of firm selection and growth,” Discussion paper, National Bureau of Economic Research.
- ARZAGHI, M., AND V. HENDERSON (2008): “Networking off Madison Avenue,” *Review of Economic Studies*, 75(4), 1011–1038.
- ASKER, J., A. COLLARD-WEXLER, AND J. DE LOECKER (2013): “Dynamic Inputs and Resource (Mis) Allocation,” *Journal of Political Economy*, *forthcoming*.
- ASPLUND, M., AND V. NOCKE (2006): “Firm Turnover in Imperfectly Competitive Markets,” *The Review of Economic Studies*, 73(2), pp. 295–327.
- AUDRETSCH, D., AND M. FELDMAN (1996): “Innovative clusters and the industry life cycle,” *Review of Industrial Organization*, 11(2), 253–273.
- BACHMANN, R., AND G. MOSCARINI (2011): “Business cycles and endogenous uncertainty,” *manuscript*.
- BALDWIN, R., AND T. OKUBO (2006): “Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting,” *Journal of Economic Geography*, 6(3), 323–346.
- BARTIK, T. (2004): “Evaluating the Impacts of Local Economic Development Policies on Local Economic Outcomes: What Has Been Done and What Is Doable?,” in *Evaluating Local Economic and Employment Development*, Book chapters authored by Upjohn Institute researchers, pp. 113–142. W.E. Upjohn Institute for Employment Research.
- BEHRENS, K., G. DURANTON, AND F. ROBERT-NICOUD (2010): “Productive cities: Sorting, selection and agglomeration,” CEPR Discussion Papers 7922, C.E.P.R. Discussion Papers.

- BEHRENS, K., G. MION, Y. MURATA, AND J. SÜDEKUM (2013): “Spatial frictions,” Discussion paper, Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit.
- BUSO, M., J. GREGORY, AND P. KLINE (2013): “Assessing the Incidence and Efficiency of a Prominent Place Based Policy,” *American Economic Review*, 103(2), 897–947.
- CAMPBELL, J. R., AND H. A. HOPENHAYN (2005): “Market Size Matters,” *The Journal of Industrial Economics*, 53(1), pp. 1–25.
- CASTRO, R., G. L. CLEMENTI, AND Y. LEE (2011): “Cross-Sectoral Variation in The Volatility of Plant-Level Idiosyncratic Shocks,” Discussion paper, National Bureau of Economic Research.
- COMBES, P., G. DURANTON, L. GOBILLON, AND S. ROUX (2010): “Estimating Agglomeration Economies with History, Geology, and Worker Effects,” *Agglomeration Economics*, p. 15.
- COMBES, P.-P., G. DURANTON, AND L. GOBILLON (2008): “Spatial wage disparities: Sorting matters!,” *Journal of Urban Economics*, 63(2), 723–742.
- COMBES, P.-P., G. DURANTON, L. GOBILLON, D. PUGA, AND S. ROUX (2012): “The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection,” *Econometrica*, 80(6), 2543–2594.
- COMIN, D. A., AND T. PHILIPPON (2006): “The rise in firm-level volatility: Causes and consequences,” in *NBER Macroeconomics Annual 2005, Volume 20*, pp. 167–228. MIT Press.
- COSTINOT, A., AND J. VOGEL (2010): “Matching and Inequality in the World Economy,” *Journal of Political Economy*, 118(4), 747–786.
- CRISCUOLO, C., R. MARTIN, H. OVERMAN, AND J. V. REENEN (2012): “The Causal Effects of an Industrial Policy,” CEP Discussion Papers dp01113, Centre for Economic Performance, LSE.
- DAVIS, D., AND J. DINGEL (2012): “A Spatial Knowledge Economy,” NBER Working Papers 18188, National Bureau of Economic Research.
- (2013): “The Comparative Advantage of Cities,” .
- D’ERASMO, P. N., AND H. J. M. BOEDO (2011): “Intangibles and endogenous firm volatility over the business cycle,” *manuscript, University of Virginia*.
- DESMET, K., AND E. ROSSI-HANSBERG (2013): “Urban Accounting and Welfare,” *American Economic Review*, 103(6), 2296–2327.

- DI GIOVANNI, J., A. LEVCHENKO, AND R. RANCIÈRE (2011): “Power laws in firm size and openness to trade: Measurement and implications,” *Journal of International Economics*, 85(1), 42–52.
- DIXIT, A., AND R. PINDYCK (1994): *Investment under uncertainty*, vol. 15. Princeton University Press Princeton, NJ.
- DURANTON, G., AND D. PUGA (2003): “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,” .
- EATON, J., S. KORTUM, AND F. KRAMARZ (2011): “An anatomy of international trade: Evidence from French firms,” *Econometrica*, 79(5), 1453–1498.
- ECKHOUT, J. (2004): “Gibrat’s Law for (All) Cities,” *American Economic Review*, 94(5), 1429–1451.
- ECKHOUT, J., AND P. KIRCHER (2012): “Assortative matching with large firms: Span of control over more versus better workers,” .
- ECKHOUT, J., R. PINHEIRO, AND K. SCHMIDHEINY (2010): “Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled,” .
- GABAIX, X. (1999): “Zipf’s Law For Cities: An Explanation,” *Quarterly Journal of Economics*, 114(3), 739–767.
- (2011): “The granular origins of aggregate fluctuations,” *Econometrica*, 79(3), 733–772.
- GABAIX, X., AND R. IBRAGIMOV (2011): “Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents,” *Journal of Business & Economic Statistics*, 29(1), 24–39.
- GASPAR, J.-M., AND M. MASSA (2006): “Idiosyncratic Volatility and Product Market Competition*,” *The Journal of Business*, 79(6), 3125–3152.
- GLAESER, E., AND J. GOTTLIEB (2008): “The economics of place-making policies,” *Brookings Papers on Economic Activity*, 2008(1), 155–239.
- GOURIÉROUX, C., AND A. MONFORT (1997): *Simulation-based Econometric Methods*, OUP/CORE Lecture Series. OUP Oxford.
- GYOURKO, J., A. SAIZ, AND A. SUMMERS (2008): “A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index,” *Urban Studies*, 45(3), 693–729.
- HENDERSON, V. (1974): “The Sizes and Types of Cities,” *American Economic Review*, 64(4), 640–56.

- HSIEH, C.-T., AND P. KLENOW (2009): “Misallocation and manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(4), 1403–1448.
- IRVINE, P. J., AND J. PONTIFF (2009): “Idiosyncratic return volatility, cash flows, and product market competition,” *Review of Financial Studies*, 22(3), 1149–1177.
- KLINE, P., AND E. MORETTI (2013): “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority,” .
- LEVINSOHN, J., AND A. PETRIN (2003): “Estimating production functions using inputs to control for unobservables,” *The Review of Economic Studies*, 70(2), 317–341.
- LUTTMER, E. G. J. (2007): “Selection, Growth, and the Size Distribution of Firms,” *The Quarterly Journal of Economics*, 122(3), 1103–1144.
- MAYER, T., F. MAYNERIS, AND L. PY (2012): “The Impact of Urban Enterprise Zones on Establishments’ Location Decisions: Evidence from French ZFUs,” .
- MELITZ, M. J. (2003a): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), pp. 1695–1725.
- (2003b): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695–1725.
- MELITZ, M. J., AND G. I. P. OTTAVIANO (2008): “Market Size, Trade, and Productivity,” *The Review of Economic Studies*, 75(1), pp. 295–316.
- NOCKE, V. (2006): “A Gap for Me: Entrepreneurs and Entry,” *Journal of the European Economic Association*, 4(5), 929–956.
- OTTAVIANO, G., T. TABUCHI, AND J.-F. THISSE (2002): “Agglomeration and trade revisited,” *International Economic Review*, pp. 409–435.
- ROSENTHAL, S., AND W. STRANGE (2004): “Evidence on the nature and sources of agglomeration economies,” *Handbook of Regional and Urban Economics*, 4, 2119–2171.
- ROSSI-HANSBERG, E., AND M. WRIGHT (2007a): “Establishment size dynamics in the aggregate economy,” *American Economic Review*, 97(5), 1639.
- (2007b): “Urban structure and growth,” *Review of Economic Studies*, 74(2), 597–624.
- SAIZ, A. (2010): “The geographic determinants of housing supply,” *Quarterly Journal of Economics*, 125(3), 1253–1296.
- SALOP, S. C. (1979): “Monopolistic competition with outside goods,” .

- SAMPSON, T. (forthcoming): “Selection into trade and wage inequality,” *American Economic Journal: Microeconomics*.
- STORY, L. (2012): “As Companies Seek Tax Deals, Governments Pay High Price,” New York Times.
- SYVERSON, C. (2004): “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy*, 112(6), pp. 1181–1222.
- TOPKIS, D. (1998): *Supermodularity and complementarity*. Princeton University Press.
- VOGEL, J. (2008): “Spatial competition with heterogeneous firms,” *Journal of Political Economy*, 116(3), 423–466.