



HAL
open science

Firm organization and productivity across locations

Grigorios Spanos

► **To cite this version:**

Grigorios Spanos. Firm organization and productivity across locations. *Journal of Urban Economics*, 2019, 112, pp.152-168. 10.1016/j.jue.2019.05.006 . hal-02271020

HAL Id: hal-02271020

<https://amu.hal.science/hal-02271020v1>

Submitted on 25 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Firm Organization and Productivity across Locations*

Grigorios Spanos[†]

May 24, 2019

Abstract

This paper provides a new mechanism to explain variation in firm productivity across locations: variation in the internal organization of labor into hierarchical layers, which are associated with different responsibilities within the firm. To guide my analysis, I develop a theoretical model that yields two implications. First, firms in larger markets organize into a greater number of layers. Second, because they have more layers, firms in larger markets are more productive. I then use administrative data to examine the model's implications across French employment areas and non-tradeable service industries that satisfy the model's assumptions: Clothing and Shoe Retail, Traditional Restaurants, and Hair and Beauty Salons. The findings are consistent with the model. I also observe that 8.8% to 22.4% of the productivity gains from denser areas arise from differences in the organization of firms. A separate analysis shows that results are similar across firms operating in the manufacturing sector.

KEYWORDS: firm organization, heterogeneous firms, market size, density, regional disparities, wages, firm productivity.

JEL Codes: D22, L11, L22, L23, J24, R12.

*This paper previously circulated as The Impact of Market Size on Firm Organization and Productivity. Research for this project was conducted during my stay at Aix-Marseille School of Economics (AMSE) as a visiting PhD student and a Post-Doctoral Fellow. I am indebted to Kunal Dasgupta, Gilles Duranton and Peter Morrow for their supervision, to Pierre-Philippe Combes for his comments and for kindly providing me with data and numerous programs, to Laurent Gobillon for sharing historical data, and to Tanguy van Ypersele for his many insightful theoretical comments and suggestions. This paper also benefited from comments and discussions with Nicolas Berman, Yann Bramoullé, Kevin Devereux, Stefan Dodds, Camille Hémet, Emeric Henry, Garth Frazer, Luis Garicano, Maya Judd, Nicholas Li, Joan Monras, Frank Neffke, Frédéric Robert-Nicoud, Lorenzo Rotunno, Michel Serafinelli, Marc Sangnier, Bernard Sinclair-Desgagné, Brian Tavares, Daniel Trefler, Thierry Verdier, and seminar participants at the Canadian Economics Association, the Urban Economics Association, the ADRES Doctoral Conference, the International Trade Workshop at the University of Toronto, the University of Bern, and the University of College Dublin. Finally, I would like to thank the staff of the French National Statistical Institute, INSEE, in Paris, for their kindness and support. Access to the data was funded by Aix-Marseille School of Economics (AMSE). This paper also benefited from funding from the French National Research Agency (ANR) project "Local Labour Markets: The Causes and Consequences of Spatial Differences in Labour Market Outcomes across Cities in France and Germany", from the Swiss National Foundation (FNS) project "Multinational Corporations: On the Relation Between the Complexity of the Knowledge Base and the Geographical and Functional Structure of the Business Group" (reference: FNS 100018_176445), and was supported by a public grant overseen by the ANR as part of the "Investing for the Future" program ("Le Programme d'Investissements d'Avenir", reference: ANR-10-EQPX-17 - "Centre d'Accès Sécurisé aux Données" (CASD)). All remaining errors are my own.

[†]Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE. Current address: Department of History, Economics and Society, School of Social Sciences (SDS), University of Geneva. 40 Boulevard du Pont-d'Arve, 1211 Geneva, Switzerland. E-mail: grigorios.spanos@gmail.com

1 Introduction

Ever since [Smith \(1776\)](#) and [Marshall \(1890\)](#), economists have known that there are productivity advantages to larger cities, and an extensive body of work in the economic literature has been devoted to understanding the source of these advantages (see [Rosenthal and Strange \(2004\)](#), [Duranton and Puga \(2004\)](#), [Combes and Gobillon \(2015\)](#) for reviews of the literature). While many explanations have been proposed, comparatively little attention has been devoted to the internal organization of firms. Yet, in reality, firms are production hierarchies: they organize their labor into layers of different sizes with the aim of assigning different responsibilities to different workers within the firm. Such decisions are important as they have an effect on the productivity of firms.

This paper aims to bridge this gap by investigating how the internal organization of firms varies with the size of local markets. To guide my analysis, I first develop a theoretical model that allows the size of local markets to affect the hierarchical structure of firms. Using administrative data, I then examine the model's implications across French employment areas and the following service industries: Clothing and Shoe Retail, Traditional Restaurants, and Hair and Beauty Salons. Although these industries account for a small portion of the economic activity of areas, they are consistent with the model's underlying assumptions. In my empirical analysis, I find that firms in denser areas organize into a greater number of layers. A firm operating in an area with a 100% higher density is 2.8% less likely to operate with only one layer in its organization, 2.0% more likely to have three layers, and 0.5% more likely to be a four-layer firm. I also find that firms with a greater number of layers are on average more productive. Finally, I find that differences in the number of layers across firms account for a meaningful share of the productivity gains from operating in denser areas. For instance, with firm productivity measured as value-added per worker, organization accounts for roughly 22.4% of the productivity gains from operating in denser areas. A separate empirical analysis shows that results are similar across firms operating in the manufacturing sector.

The theoretical model combines the knowledge-based management hierarchy framework of [Garicano \(2000\)](#) and [Caliendo and Rossi-Hansberg \(2012\)](#) with the monopolistically competitive framework of [Melitz and Ottaviano \(2008\)](#). Production is modeled as in [Caliendo and Rossi-Hansberg \(2012\)](#) and requires labor to be combined with knowledge. Firms are production hierarchies and organize workers into layers of different sizes. Production workers always form the firm's lowest layer, whereas managers occupy all the other layers. The number of layers in a firm is also determined in the model. Different organizational structures are associated with different menus of costs and, as demonstrated by [Caliendo and Rossi-Hansberg \(2012\)](#), the decision to add a layer is similar to a tradeoff between fixed and variable costs. An extra layer increases costs

because workers are compensated for the knowledge they acquire and the time they devote to the firm. Above a certain scale, however, an additional layer allows firms to use workers' knowledge more efficiently, lowering average costs. Thus, when a firm adds a layer, it is in a sense increasing its fixed costs and lowering its variable costs. Ultimately, a firm's number of layers is determined by its profit maximizing output, which depends on the demand for its product and the degree of competition between firms.

To highlight how firm organization varies with the size of local markets, I embed this structure into an economy in which agents have quasi-linear preferences, and firms have heterogeneous product demand. Differences in the size of markets generate differences in the level of competition between firms. Larger markets increase the competition between firms and lower markups, which in turn forces firms with low product demand to exit the market, and the remaining firms to restructure their organization. As a consequence of these two effects, the cumulative distribution of firm organization in larger markets first-order stochastically dominates the distribution of smaller markets.

My model also implies that differences in organization bring about differences in the productivity of firms. Firms with a greater number of layers manage the knowledge of workers with greater efficiency and are consequently able to reduce average costs. Thus, the theoretical model implies that in larger markets, firms are more productive because they organize into a greater number of hierarchical layers. To the best of my knowledge, this is the first paper to document this new channel through which the size of local markets affects the productivity of firms.

Employing high quality French administrative data, I examine the implications of the model. The primary data sources are the Déclarations Annuelles des Données Sociales (DADS) and the Fichier Complet Unifié de Suse (FICUS). From the DADS, I classify firms into one-layer, two-layer, three-layer, and four-layer firms, employing the method of [Caliendo et al. \(2015b\)](#). Using both sources, I calculate different estimates of firm productivity, including value-added per worker and several estimates of total factor productivity. The data also contain a wealth of information on the characteristics of firms and the demographic composition of local areas.

My analysis follows [Ciccone and Hall \(1996\)](#) and measures the size of markets using the population density of employment areas. Since the model makes specific assumptions on the structure of competition between firms, it directly applies only to non-tradeable monopolistically competitive industries. In order to precisely investigate its implications, the main part of the analysis focuses on industries that satisfy these conditions: Clothing and Shoe Retail, Traditional Restaurants, and Hair and Beauty Salons. I also take into account several empirical concerns. First, consumer demand and the cost of hiring workers may vary with the density of the area, affecting the level of compe-

tition and firm organization. Second, it is likely that firms with a greater number of layers attract more workers to an area, or that local shocks simultaneously determine the size of areas and the outcomes of firms. Such issues could confound estimates of the relationship between density and the different outcomes of firms. To deal with these issues, I also control for the local characteristics of areas and rely on an instrumental variable estimation, using historical data from as far back as 1831.

A main prediction of the model is that the distribution of firm organization in high density areas first-order stochastically dominates the distribution of low density areas. The findings from the Mann-Whitney stochastic dominance test and the regression results confirm this prediction. Consistent with the model, I also find that firms with a greater number of layers are on average more productive. Finally, the main conclusion of my analysis is that differences in organization are economically significant. From the various measures of firm productivity, I find that between 8.8% and 22.4% of the productivity gains from operating in denser areas arise from differences in firm organization. While the model directly applies to specific market structures, a separate analysis also investigates whether results are similar in the manufacturing sector, and reveals that my findings are robust to other industries and that differences in firm organization contribute a similar amount to the productivity gains from denser areas.

The paper is organized as follows. Section 2 discusses the related literature. Section 3 presents the theoretical model and its main implications. Section 4 introduces the data, and Section 5 empirically examines the main implications of the model. The proofs of propositions are included in the Online Appendix.

2 Literature Review

This paper contributes to two strands of the economic literature: research on firm organization and that exploring the productivity advantages of larger areas.

A growing body of work has integrated the internal organization of firms in the study of the economy (see [Antras and Rossi-Hansberg \(2009\)](#) and [Garicano and Rossi-Hansberg \(2014\)](#) for a review). While most research in this area focuses on the interaction between technology, international trade, and the hierarchical structure of firms, my analysis considers the size of local markets.¹ This paper also contributes to the current literature in other ways. First, it provides a new channel through which the economic environment effects the internal organization of firms. Second,

¹Papers examining how changes in information technology affect firm organization include: [Garicano \(2000\)](#), [Garicano and Rossi-Hansberg \(2004\)](#), [Garicano and Rossi-Hansberg \(2006\)](#), [Raghuram and Wulf \(2006\)](#), [Garicano and Heaton \(2010\)](#), [Bloom et al. \(2014\)](#), [Charnoz et al. \(2018\)](#) and [Caicedo et al. \(2019\)](#). Studies analyzing how trade, or offshoring, affect firm organization include: [Antras et al. \(2006\)](#), [Antras et al. \(2008\)](#), [Guadalupe and Wulf \(2010\)](#), [Dasgupta \(2012\)](#), [Caliendo and Rossi-Hansberg \(2012\)](#), [Caliendo et al. \(2015b\)](#), [Friedrich \(2015\)](#), [Spanos \(2016\)](#) and [Gumpert \(2017\)](#).

it documents and explains the differences in the distribution of firm organization across locations. Finally, it shows that firm organization is important for understanding spatial differences in the productivity of firms.

Within this literature my analysis is most closely related to [Caliendo and Rossi-Hansberg \(2012\)](#) and [Caliendo et al. \(2015b\)](#). [Caliendo and Rossi-Hansberg \(2012\)](#) embed the knowledge-based management hierarchy model developed by [Garicano \(2000\)](#) into a monopolistic competition framework with heterogeneous firms. In their model, the elasticity of substitution between product varieties does not vary with the size of local markets. Thus, in the absence of trade the distribution of firm organization is the same across locations. As shown below, however, firm organization changes with the size of local markets. Empirically, [Caliendo et al. \(2015b\)](#) develop a method to construct the organization of firms using French administrative data. They focus on the manufacturing sector and show that their method is consistent with several predictions of the management hierarchy models. As documented below, their method is also applicable to firms operating in the service sector.

My analysis also complements the findings of [Garicano and Hubbard \(2007\)](#), who examine how market size determines the degree of field specialization and the organization of law firms. Using data from the U.S., they find that larger markets have more specialized firms and a greater share of workers employed in management hierarchies. The mechanism they propose, however, is different from that of this paper. While [Garicano and Hubbard \(2007\)](#) focus on uncertainty over demand, I examine the level of competition between firms. That said, my analysis relates to their results in two additional ways. First, the model allows firms to organize into any number of layers and thus provides a general comparison of the distribution of firm organization across locations. Second, I quantify the extent to which differences in firm organization contribute to spatial differences in the productivity of firms.

This paper also builds on an extensive urban literature exploring the productivity advantages of larger areas. It is a well-established fact that larger areas increase the productivity of workers and firms (see, for example, [Ciccone and Hall \(1996\)](#), [Syverson \(2004\)](#), [Greenstone et al. \(2010\)](#) and [Combes et al. \(2012\)](#)). I provide a new micro-founded theory that contributes to understanding this phenomenon. The theory is based on the knowledge-based management hierarchy framework of [Garicano \(2000\)](#), which generates internal economies of scale endogenously from the efficient organization of knowledge in firms.

While my analysis focuses on the productivity of firms, the model can also shed light on other economic features of larger areas. The organizational problem of [Garicano \(2000\)](#) involves determining the optimal division of labor into managerial and production tasks and the sharing of

knowledge in firms. The solution to this problem implies that with more layers, tasks are more specialized and workers share more information with one another. This paper thus contributes to understanding various economic characteristics of areas: the division of labor, the specialization of tasks and the sharing of knowledge, all of which are greater in larger areas (Duranton and Puga (2004)).

Finally, there are several policy implications to my findings. In particular, the results suggest that there is less scope for local intervention to increase the productivity of areas. To the extent that city planners are interested in policies that promote agglomeration economies through an increase in the size of areas, the analysis indicates that the benefits of agglomeration are actually lower than previous estimates in the literature. A share of the productivity gains from larger areas, commonly attributed to external economies of scale, are already internalized and priced into firms.

3 Model

The economy contains N homogeneous individuals who supply their unit of labor inelastically. Let q_o^c and q_i^c denote individual consumption of the homogeneous good and variety i , and Ω a connected real interval representing the set of varieties available in the economy. As in Foster et al. (2008), individual utility has a linear quadratic form and is equal to:

$$U^c = q_o^c + \int_{\Omega} \alpha_i q_i^c di - \frac{\gamma}{2} \int_{\Omega} (q_i^c)^2 di - \frac{\eta}{2} \left(\int_{\Omega} q_i^c di \right)^2, \quad (1)$$

where the parameters α_i , η and γ are positive, and α_i varies across varieties.² I assume individual demand for the homogeneous good is positive.³ Let $\Omega^* \subset \Omega$ denote the connected set of varieties produced in equilibrium, with positive measure M . Maximizing utility, substituting, and isolating terms yields the linear demand of variety $i \in \Omega^*$:

$$q_i = Nq_i^c = \frac{N}{\gamma} \alpha_i - \frac{N}{\gamma} p_i - \frac{N}{\gamma} \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}), \quad (2)$$

where $\bar{\alpha} = \frac{1}{M} \int_{\Omega^*} \alpha_i di$ and $\bar{p} = \frac{1}{M} \int_{\Omega^*} p_i di$ are the average variety draws and prices of varieties in Ω^* . Equation (2) clarifies the role of α_i : for a given price there is greater demand for varieties with a greater α_i . Moreover, from this structure the price elasticity of demand of variety i , ϵ_i , is equal to $[p_i^{max} / p_i - 1]^{-1}$, where $p_i^{max} = \alpha_i - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})$ denotes the maximum chargeable price. The elasticity is monotonically decreasing with respect to α_i and increasing with $\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})$. The latter

²These parameters determine the level of competition between varieties and the homogeneous good. Varieties are more substitutable as γ approaches zero, and a greater α_i or a lower η increase the demand of variety i relative to the homogeneous good.

³In other words, individuals spend income on the homogeneous good: $I^c > \int_{i \in \Omega^*} p_i q_i^c di$, where I^c denotes individual income.

term is common to all varieties, and as will be shown below, increases with the size of markets, N . As in [Melitz and Ottaviano \(2008\)](#), I define a tougher competitive environment as an equilibrium with a more elastic demand for every variety.

Labor is the only input used in production and is inelastically supplied in a competitive market. Each individual has one unit of time and can work in one of three sectors: the homogeneous good sector, the differentiated goods sector, or the schooling sector. Because they are identical, in equilibrium all individuals earn the same wage w , irrespective of their sector and occupation.

In the homogeneous good sector, each good is produced with a constant returns to scale technology and a unit of output requires a unit of labor. In the analysis that follows, the price of the homogenous good is normalized to 1, and as a result, in equilibrium wages are equal to 1 in the economy.

In the differentiated goods sector, individuals create firms. For simplicity, I refer to an agent that creates a new firm as an entrepreneur.⁴ To enter the sector an entrepreneur first pays a fixed cost f_E , in units of labor, to develop her product. The product is associated with a demand draw α from a known cumulative distribution $G(\alpha)$, which determines the demand schedule of her firm. Given the local environment, if the demand draw is low she may decide to immediately exit the sector; otherwise she creates a firm to produce her product.

I model production in the differentiated goods sector as in [Garicano \(2000\)](#) and [Caliendo and Rossi-Hansberg \(2012\)](#). Firms are knowledge-based management hierarchies and organize workers into layers, so as to use their knowledge efficiently. In general, workers employed in a firm are divided into two broad categories: production workers and managers. Production workers are located in the lowest layers of firms, while managers occupy all other layers. Entrepreneurs are in the highest layer, and perform the same tasks as agents in that layer.⁵

In the differentiated goods sector, production requires labor and knowledge. Agents solve problems to produce output and are compensated by the firm for any knowledge they acquire. Every production worker spends her one unit of time generating a problem, z , from a known exponential distribution, $F(z)$, with parameter $\lambda > 0$. A given problem drawn, z , can be solved by a worker if her knowledge set contains z , at which point A units of output are produced. The type of problem a worker draws, however, is unobservable. For a given realization of z , the only information available to a worker is whether or not she can solve the problem.⁶

⁴As I use the same production technology as [Caliendo and Rossi-Hansberg \(2012\)](#), I also adopt their nomenclature with the exception of the labeling of firm organization. [Caliendo and Rossi-Hansberg \(2012\)](#) refer to firm organization by the number of management layers, whereas I use the total number of layers.

⁵In other words, if an entrepreneur creates a firm with zero management layers, then she performs the same tasks as a production worker. Otherwise she performs the tasks of a manager in the top layer.

⁶In the literature, this is referred to as the "labeling assumption", which is crucial for the existence of organizations (see [Garicano \(2000\)](#) for details). The parameter λ determines how common problems are in production. When λ is

In a one-layer firm, or a firm with no layers of managers, if a worker cannot solve the problem drawn, z , nothing is produced. This is not the case in a firm with managers. If the worker cannot solve the problem, she asks her manager one layer above. The manager spends h units of her time listening and solves the problem if her knowledge set includes z . If the manager cannot solve the problem then the worker asks a manager one layer higher. This process continues until the problem is solved and A units of output are produced, or the problem reaches the highest layer of the organization, which is occupied by the entrepreneur. The entrepreneur also spends h units of her time listening to the worker's problem and solves it if her knowledge set includes z . If, however, the entrepreneur cannot solve the problem, nothing is produced.

The schooling sector is composed of agents who spend their time providing knowledge to agents in the differentiated goods sector. I assume that a unit interval of knowledge requires c units of teachers' time in the schooling sector. Teachers earn the same wage for their time as other workers in the economy; thus the cost of a unit of knowledge is simply c . In addition, knowledge in the differentiated goods sector is not cumulative: managers do not need to learn to solve the most common problems before they can solve the exceptional ones.⁷

3.1 Entrepreneur's Problem

Consider an entrepreneur who does not exit the differentiated goods sector. Given her demand schedule she creates a firm to produce a unique variety of the differentiated good with the objective of maximizing profits. In making this decision she also selects the number of layers, the knowledge, and the number of agents in each layer of the firm. The subsections that follow address these choices. I first describe the properties of the cost functions, which depend on how firms organize production. The cost functions are determined from the firm's cost minimization problem and because the latter is analyzed in [Caliendo and Rossi-Hansberg \(2012\)](#), I simply summarize the results. For details and proofs I refer the reader to their paper. I then describe the solution that maximizes the profits of a firm.

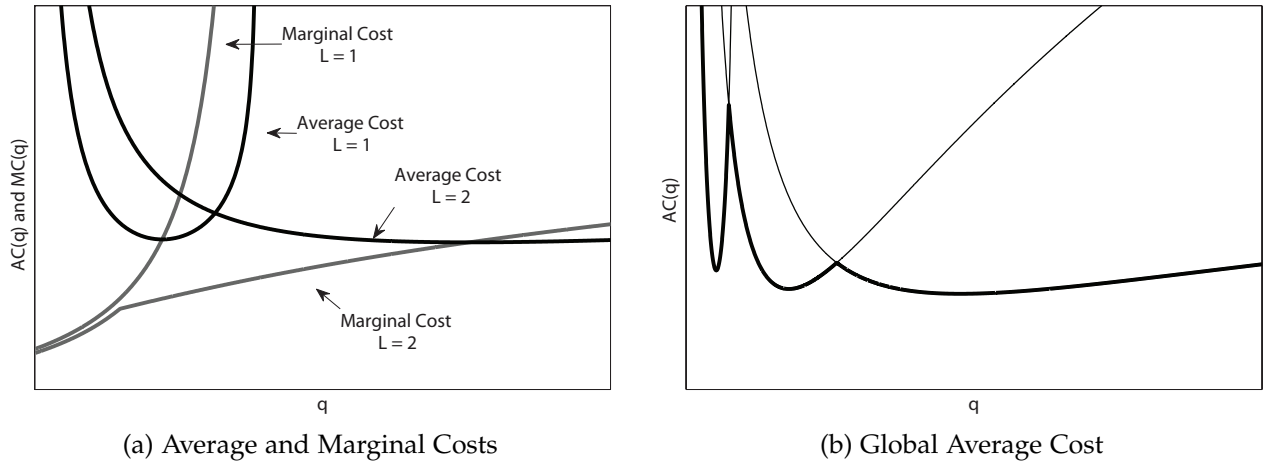
Cost Minimization: Properties of Cost Functions

Figure 1a illustrates the marginal and average cost curves as a function of the quantity produced, in one and two-layer firms (L denotes the total number of layers in firms). Figure 1b shows the global average cost curve when the organization of production is endogenous and firms can have

high, problems are less costly to solve because their distribution is more concentrated around zero, and as λ approaches infinity, knowledge becomes unimportant for production.

⁷To solve a problem z , an agent's knowledge set must include z , but the agent is not required to be able solve all problems between 0 and z . If the length of an agent's knowledge set is t , then the cost of acquiring that knowledge is ct .

Figure 1: Cost Functions of a Firm with L Total Number of Layers



Notes: Panel (a) illustrates the marginal (gray lines) and average (black lines) cost curves of a firm producing with one or two layers. Panel (b) illustrates the average costs curves (thin black lines) of a firm producing with one, two or three layers, and the associated global average cost curve (thick black line). The parameters used to simulate the cost functions are as follows: $A = 5$, $w = 1$, $h = 0.95$, and $c/\lambda = 2$.

either one, two, or three layers. Both figures are reproduced as in [Caliendo and Rossi-Hansberg \(2012\)](#) and illustrate the main properties of the cost functions described in this subsection.

Fixing the number of organizational layers in a firm, the cost functions have several properties. First, marginal costs are positive and increasing in output, which implies that the knowledge of agents in all layers, and the size of each layer is increasing in output. Second, because marginal costs are positive, total costs are strictly increasing in output. Third, the average cost curves are convex in output, attain their minimum when they intersect the marginal cost curves, and converge to infinity as output approaches zero or infinity.⁸

The intuition for these results is the following. Given the number of layers, an increase in production requires more knowledge in the firm. Some of this additional knowledge is acquired by the entrepreneur, while the remainder is acquired by agents in the other layers. As agents in a given layer acquire more knowledge their managers can supervise more of them, which implies that the chosen organizational structure becomes larger and the cost of producing an extra unit of output increases. In addition, average costs are not a monotonic function of quantity. When production is small, increases in output lead to a less than linear increase in total costs, as the firm does not need to provide too much knowledge to agents in the lower layers of the firm. Average costs are reduced until the minimum efficiency scale (MES) is reached. Beyond the MES, however,

⁸The parameters used to simulate the cost functions as follows: $A = 5$, $w = 1$, $h = 0.95$, and $c/\lambda = 2$. Throughout the paper, I assume, as in [Caliendo and Rossi-Hansberg \(2012\)](#), that the inequality $\frac{c}{\lambda} \leq \frac{h}{1-h}$ holds. This ensures that a firm would rather decrease its number of layers before choosing to hire employees with zero knowledge (see [Caliendo and Rossi-Hansberg \(2012\)](#) for additional details). For a proof of these statements, and those made in the paragraphs below, see Propositions 1 and 2 in [Caliendo and Rossi-Hansberg \(2012\)](#).

the opposite occurs and an increase in output raises the average costs of the firm.

Further, at the level of output where a firm is indifferent between two organizational structures and adds a layer of management, marginal costs decrease discontinuously. This implies that the knowledge of agents in all layers decreases discontinuously, while the number of employees in each layer increases discontinuously. Intuitively, because agents are compensated for their time and knowledge, adding a layer of management is costly. Yet by doing so, firms allow a manager with the ability to solve less frequent problems to better use her knowledge in the firm. To attain a certain level of output, a firm that adds a layer of management requires more agents, but with less knowledge, in its existing layers. The firm, therefore, reduces its cost per worker. In other words, for a given level of output, adding a new layer of management can be understood as the firm paying higher “fixed” costs; it hires more agents and, beyond a level of output, can produce at lower average costs because each agent acquires less knowledge (Caliendo and Rossi-Hansberg (2012)).

Let $C(q_i)$ denote the global minimum total cost of producing q_i units of the differentiated variety. The global total cost function of a firm is the lower envelope of the total cost functions of producing q_i units with a given number of layers, and is equal to: $C(q_i) = \min_{L \geq 1} \{C_L(q_i)\}$, where L denotes the total number of layers in firms. The global total cost function contains the MES of all organizations as well as the regions around them. This is because the MES of each organization achieves a greater quantity and a lower average cost when the number of layers are increased. Therefore, for a given organizational form, global cost functions adopt the properties of local cost functions, whereas when the number of layers increase, global cost functions are either discontinuous or non-differentiable. All of these results are summarized below:

Summary 1 *The global cost functions have the following properties:*

- i. Fixing the number of layers, marginal cost, $MC(q_i)$, is positive and increasing in output. When the number of layers increases, $MC(q_i)$ decreases discontinuously.*
- ii. Fixing the number of layers, average cost, $AC(q_i)$, is convex and attains the minimum when $MC(q_i) = AC(q_i)$. $AC(q_i)$ is globally continuous but not globally convex and converges to infinity when q_i approaches zero or infinity.*
- iii. Fixing the number of layers, total cost, $C(q_i)$, is continuous and convex. $C(q_i)$ is globally continuous but not globally convex.*

Profit Maximization

Caliendo and Rossi-Hansberg (2012) embed this production framework in a monopolistically com-

petitive model with heterogeneous firms and a market structure characterized by a constant elasticity of demand. In their setting, the elasticity of demand does not vary with the size of the market. Thus, in the absence of trade the level of competition, the optimal quantity, and the organization of firms remains the same across markets of different sizes (see Proposition 7 from [Caliendo and Rossi-Hansberg \(2012\)](#)). Here I depart from [Caliendo and Rossi-Hansberg \(2012\)](#) and combine their production framework with the competitive framework of [Melitz and Ottaviano \(2008\)](#), which allows the elasticity of demand to change with the size of markets.

Consider an entrepreneur who creates a firm. Given her demand, she competes in a monopolistically competitive market and chooses the quantity that maximizes the profits of her firm. In other words, she solves the following problem: $\pi(\alpha) = \max_q p(q(\alpha))q(\alpha) - C(q(\alpha))$. Since $C(q)$ is not globally convex with respect to q , the first-order condition to this problem only describes local solutions, which depend on the number of layers in a firm. Comparing the profits from these solutions, one then arrives at the global optimal profits of a firm.⁹ Consequently, in this subsection I first describe the local solutions to the entrepreneur's maximization problem and then return to the global solution.

From the first-order condition, the quantity that maximizes local profits solves the equation:

$$q_L(\alpha) = \frac{N}{2\gamma} \left[\alpha - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC_L(q_L(\alpha)) \right], \quad (3)$$

where the subscript L indicates that the number of layers is held constant at L .¹⁰ Let $p_L(\alpha)$ and $\pi_L(\alpha)$ denote the optimal price and profit of a firm with demand draw α and operating with L layers. The following equations describe these variables in terms of the model's parameters:

$$\begin{aligned} p_L(\alpha) &= \mu_L^{AC}(\alpha) + AC_L(q_L(\alpha)), \\ &= \mu_L^{MC}(\alpha) + MC_L(q_L(\alpha)) = \frac{1}{2} \left[\alpha - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) + MC_L(q_L(\alpha)) \right], \end{aligned} \quad (4)$$

$$\pi_L(\alpha) = r_L(\alpha) - C_L(q_L(\alpha)) = \frac{N}{4\gamma} \left[\left(\alpha - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) \right)^2 - MC_L(q_L(\alpha))^2 \right] - C_L(q_L(\alpha)), \quad (5)$$

where $\mu_L^{MC}(\alpha)$, $\mu_L^{AC}(\alpha)$, and $r_L(\alpha)$ are the expressions describing markup over marginal costs, markup over average costs, and revenue of a firm with demand draw α and producing with L layers.

⁹For given L , the local optimal profit function of a firm is strictly concave in q , $\pi(0) = -1$ and $\lim_{q \rightarrow \infty} \pi(q) = -\infty$ and therefore the local profit maximization problem is well defined.

¹⁰The left-hand side is increasing with respect to q , while the right-hand side is decreasing. Thus, for every α there exists a unique quantity, $q_L(\alpha)$, such that equation (3) holds with equality. A closed form solution for $q_L(\alpha)$ is unavailable because marginal costs also depend on quantity.

These local solutions vary with the demand schedule of firms. From equations (3) and (4) it follows that firms with greater demand produce more and set higher prices. Naturally, since quantities and prices increase with demand, revenues are also an increasing function of α . Additionally, it follows from equation (4) that firms with greater demand also set higher markups.¹¹ These results are summarized below:

Proposition 2 *Holding the number of layers constant, $q_L(\alpha)$, $p_L(\alpha)$, $\mu_L^{MC}(\alpha)$, $\mu_L^{AC}(\alpha)$, and $r_L(\alpha)$, are increasing with respect to α . Moreover, $\pi_L(\alpha)$ is continuous, strictly increasing with respect to α and strictly concave in q .*

In addition, profits are continuous and increasing with respect α . Now, consider an entrepreneur who is indifferent between producing with L or with $L + 1$ layers.¹² When she increases the number of layers in her firm, marginal costs decrease discontinuously and this has an effect on the remaining outcomes of her firm.¹³ The quantity supplied increases discontinuously, while price decreases discontinuously. However, the rise in quantity dominates the decline in price, and revenues increase discontinuously. Finally, when an entrepreneur increases the number of layers in her firm, the change in costs dominates the change in prices, and markups increase discontinuously.

For a given α , let $q(\alpha)$, $p(\alpha)$, $\mu^{MC}(\alpha)$, $\mu^{AC}(\alpha)$, $r(\alpha)$ correspond to the global solution to the entrepreneur's maximization problem. When the number of layers is constant, $q(\alpha)$, $p(\alpha)$, $\mu^{MC}(\alpha)$, $\mu^{AC}(\alpha)$, $r(\alpha)$ have the properties of their corresponding local solutions. When the entrepreneur increases the number of layers in her firm, marginal costs decrease discontinuously and $q(\alpha)$, $p(\alpha)$, $\mu^{MC}(\alpha)$, $\mu^{AC}(\alpha)$, $r(\alpha)$ adjust discontinuously. The profit function $\pi(\alpha)$ is the upper envelope of the local profit functions $\pi_L(\alpha)$. It is continuous, strictly increasing and convex in α . These results are summarized below:

Proposition 3 *Holding the number of layers constant, $q(\alpha)$, $p(\alpha)$, $\mu^{MC}(\alpha)$, $\mu^{AC}(\alpha)$, and $r(\alpha)$, are increasing with respect to α . When a firm increases the number of layers, $q(\alpha)$, $\mu^{MC}(\alpha)$, $\mu^{AC}(\alpha)$, and $r(\alpha)$ increase discontinuously, while $p(\alpha)$ decreases discontinuously. Moreover, $\pi(\alpha)$ is continuous, strictly increasing and convex with respect to α .*

¹¹An increase in α has two opposing effects on markups. First, prices rise with α , increasing both $\mu_L^{MC}(\alpha)$ and $\mu_L^{AC}(\alpha)$. Second, firms produce a greater quantity of the differentiated variety causing marginal costs to increase, and average costs to either rise or decline. In all cases, the first effect dominates, and both markups increase with α .

¹²Such an entrepreneur always exists, because for a given α , the slope of the profit function is increasing with the number of layers. Formally, in applying the envelope theorem to the profit function, it follows that: $\frac{\partial \pi_L(\alpha)}{\partial \alpha} = q_L(\alpha)$. Since for any α , $q_L(\alpha)$ is increasing with L it follows that the profit function has a steeper slope for a greater number of layers.

¹³Consider such a firm with demand draw $\alpha_{L,L+1}$. The optimal quantity is chosen so that marginal revenue equals marginal cost. Since for a given quantity, an organization with L layers has a greater marginal cost than an organization with $L + 1$ layers, it follows that $q_{L+1}(\alpha_{L,L+1}) > q_L(\alpha_{L,L+1})$, and at the optimal quantities $MC_L(q_L(\alpha_{L,L+1})) > MC_{L+1}(q_{L+1}(\alpha_{L,L+1}))$. In addition, since revenues increase discontinuously, total costs are greater when producing with $L + 1$ layers, i.e., $C_L(q_L(\alpha_{L,L+1})) < C_{L+1}(q_{L+1}(\alpha_{L,L+1}))$.

3.2 Equilibrium

Prior to entering the differentiated goods sector, entrepreneurs develop a product that yields a demand draw α from a known cumulative distribution $G(\alpha)$ with support $[\alpha_M, \infty]$. This process, however, costs f_E units of labor. Under the assumption that α_M is relatively small, there will always be demand draws that yield negative profits from operating a firm. Therefore, given the mass of entrants in the sector, M_E , a demand cutoff α_D exists such that all entrepreneurs with draw $\alpha < \alpha_D$ choose to exit, and the marginal entrepreneur's profits are zero: $\pi(\alpha_D, M) = 0$. Moreover, the optimal quantity supplied by a firm created by the marginal entrepreneur is the solution to:

$$q_D = \frac{N}{2\gamma} \left[\alpha_D - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC(q_D) \right]. \quad (6)$$

Isolating $\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})$ and substituting it back into equation (5), yields an expression for the profits of firms, as a function of α , and the endogenously determined demand and quantity cutoffs:

$$\pi(\alpha, \alpha_D, q_D) = \left[\frac{2\gamma}{N} q_D + MC(q_D) - \alpha_D + \alpha - \frac{\gamma}{N} q(\alpha) \right] q(\alpha) - C(q(\alpha)). \quad (7)$$

Consequently, the profits of a firm created by the marginal entrepreneur are equal to:

$$\pi(q_D) = \frac{\gamma}{N} q_D^2 + MC(q_D) q_D - C(q_D), \quad (8)$$

which only depend on the quantity cutoff, q_D .

The expected profits from entry, Π^e , is equal to the expected profits before entrepreneurs know their demand schedule minus the fixed entry cost. Unrestricted free entry implies that expected profits, Π^e , are zero, which yields the equilibrium free entry condition:

$$\int_{\alpha_D} \pi(\alpha, \alpha_D, q_D) dG(\alpha) = f_E. \quad (9)$$

Finally, the mass of firms operating in equilibrium is the mass of entrants with successful entry, and is therefore equal to: $M_E = M/(1 - G(\alpha_D))$.

The equilibrium is a set of values, q_D , α_D , and M , that solve equations (6), (8) and (9), given the parameters of the model and the distribution of demand draws $G(\alpha)$. Additionally, a condition on the parameter η ensures that consumers also demand the homogeneous good. I prove that:

Proposition 4 *If $\eta > \underline{\eta}$, then there exists a unique equilibrium.*

3.3 Comparative Statics with Respect to Market Size

Consider an increase in the size of a market. From the implicit function theorem it follows that:

$$\frac{\partial \alpha_D}{\partial N} = - \frac{\partial \Pi^e / \partial N + \partial \Pi^e / \partial q_D \partial q_D / \partial N}{\partial \Pi^e / \partial \alpha_D}. \quad (10)$$

The numerator is positive while the denominator is negative, implying that an increase in market size raises the demand cutoff, α_D . Intuitively, there are two effects behind this result. First, a direct effect: holding the mass of entrants fixed, an increase in the size of a market raises sales and profits. Second, an indirect effect: increased profits for entrants implies that potential profits rise as well, increasing the expected profits from entry, Π^e . To maintain the equilibrium condition that Π^e is zero, the term $\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p})$ rises, which intensifies the competition between firms, lowering expected profits and raising the bar for successful entry. Under a restriction on the parameter η , we can ensure that M rises as well.¹⁴ This result is summarized below:

Proposition 5 *An increase in N induces an increase in q_D , α_D and $\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p})$. In addition, if $\bar{\eta} > \eta$, then a rise in N induces an increase in M .*

Consider an entrepreneur with a demand draw $\alpha_{L,L+1}$ and indifferent between organizing production with L or $L + 1$ layers. The distance between $\alpha_{L,L+1}$ and α_D decreases with N . Intuitively, there are two effects behind this result. First, an increase in market size raises the bar for successful entry, increasing α_D . Second, greater competition between firms lowers markups and profits. However, lower markups also change the relative tradeoff of the fixed and variable costs from adding a layer to a firm. For firms producing on a sufficiently large scale, the benefit of adding a layer and increasing markups, exceeds the costs. Some entrepreneurs therefore reorganize their firm and add additional layers of managers so as to increase the profits of their firms.¹⁵ This result is summarized below:

Proposition 6 *The distance between the marginal entrepreneur with demand draw, α_D , and the entrepreneur who is indifferent between two organizational forms, $\alpha_{L,L+1}$, is decreasing with respect to N . In addition, the change in the distance between α_D and $\alpha_{L,L+1}$, increases with L .*

¹⁴Alternatively, it may be that $(\bar{\alpha} - \bar{p})$ rises and M decreases with N . The proof to this proposition shows there always exists an $\eta \in [\underline{\eta}, \bar{\eta}]$. Moreover, whether the number of varieties increases or decreases with market size does not affect the results below. From equation (8), it also follows that a rise in N increases q_D . However, this does not imply that firms with demand draw α_D increase their quantity in larger markets. The equilibrium consists of a mapping between the set of demand draws, α 's, and the set of quantities, q 's, that depends on N and the other parameters. How this mapping changes along its support determines how a firm's production changes with the size of the market. I observe, in fact, that the response to operating in a larger market is heterogeneous across firms (Appendix H contains simulations).

¹⁵There is evidence that prices and markups are lower in larger markets (see [Handbury and Weinstein \(2014\)](#) and [Bellone et al. \(2014\)](#)). Moreover, both effects do not imply that $\alpha_{L,L+1}$ decreases with N . When N increases, the mapping between demand draws and quantities changes. What matters, however, is the distance between $\alpha_{L,L+1}$ and α_D . The proof of Proposition 6 provides a lower and upper bound to the change in $\alpha_{L,L+1}$ with respect to N , and shows that they may be positive or negative. Additionally, the proof of Proposition 6 shows that the quantity produced by the firm with L layers, $q_L(\alpha_{L,L+1})$ decreases with N , while the quantity produced by the firm with $L + 1$ layers, $q_{L+1}(\alpha_{L,L+1})$, increases with N .

Let Λ_N denote the discrete cumulative distribution of firm organization in the economy. Given the model parameters, the probability mass of firms producing with at most L layers is equal to:

$$\Lambda_N(L) = \frac{G(\alpha_{L,L+1}) - G(\alpha_D)}{1 - G(\alpha_D)}. \quad (11)$$

Now consider how the distribution changes with N . When the size of the market changes, both the numerator and denominator in equation (11) are affected. On its own, Proposition 6 does not provide enough information to stochastically order the distribution of firm organization across markets of different sizes. However, if the distribution of demand draws has a non-increasing hazard rate, the cumulative distribution of firm organization in larger markets first-order stochastically dominates the distribution in smaller markets.¹⁶ This result is summarized in the following proposition:

Proposition 7 *Suppose that $N' > N$. If the distribution of demand draws $G(\alpha)$ has a non-increasing hazard rate, then the cumulative distribution of firm organization $\Lambda_{N'}$ first-order stochastically dominates Λ_N .*

Proposition 7 contains the main theoretical result of the paper. In larger markets, there is a greater mass of firms producing with a higher number of layers. This proposition also has implications for how the distributions of productivity and knowledge differ across markets of different sizes. The next subsection demonstrates the results for the productivity of firms.

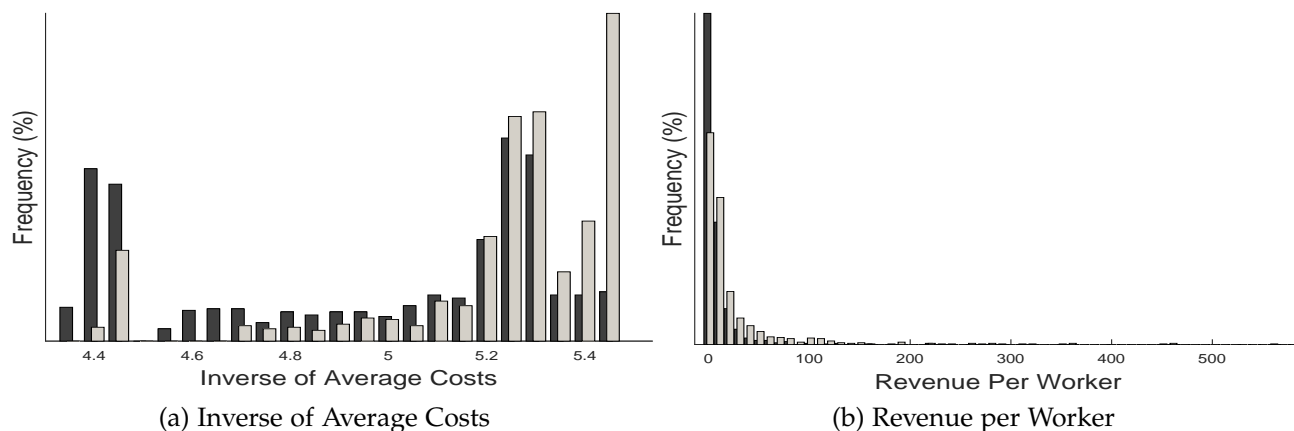
3.4 Numerical Simulations

This section uses numerical simulations to illustrate the model's implications on how the distributions of productivity vary across markets of different sizes. The figures presented in this section are from two simulations of the model. In both simulations, demand is drawn from a Pareto distribution, and the model parameters are identical in both simulations except for the size of the markets. Distributions with dark gray bars are from the smaller market. A complete description of the simulations and additional results are presented in the Supplementary Appendix H.

Figure 2 illustrates how the distribution of firm productivity varies across markets. In Figure 2a, productivity is defined as the inverse of average costs, which is a measure of quantity productivity, while in Figure 2b productivity is equal to revenue per worker, a measure of revenue productivity. Both figures show there are different shares of small and high productivity firms across markets.

¹⁶The Pareto distribution, regularly used in the heterogeneous firm literature, satisfies the non-increasing hazard rate property. The intuition behind the non-increasing hazard rate assumption is the following: holding L fixed, it ensures that the mass lost from firms changing their organizational hierarchy and producing with $L' \neq L$ is greater than the mass gained from firms changing their organizational hierarchy and now producing with L layers.

Figure 2: Distributions of Different Measures of Firm Productivity



Notes: Panel (a) illustrates the distribution of inverse of average costs across two markets of different sizes. Panel (b) illustrates the distribution of revenue per worker across two markets of different sizes. Light gray represents the larger market, while dark gray represents the smaller market. The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

The larger market contains a smaller share of low productivity firms and a greater share of high productivity firms. The Supplementary Appendix H illustrates results using other measures of quantity and revenue productivity and the conclusions stay the same: firms in larger markets organize into a greater number of layers and, as a result, are on average more productive.¹⁷

This study therefore establishes a new mechanism to explain differences in the productivity of firms across markets. Namely, firms in larger markets are more productive because they organize production into a greater number of layers. This is a direct result of the production framework of the model, and the effect that a tougher competitive environment has on firm organization. Firms producing with a greater number of layers, organize their workforce more efficiently, reduce average costs, hire more skilled workers, and are more productive.¹⁸ The sections that follow examine the model's predictions and assess how important this new mechanism is for understanding firm productivity differences across locations of varying sizes.

Before moving on to the empirical analysis, however, it is important to note that the model has specific assumptions, and as a result, only directly applies to particular market structures: monopolistically competitive industries with exclusively local demand. The majority of the analysis takes the model's assumptions seriously, and examines its implications on firms that operate in industries where demand is determined at the local level. In Section 5.3, I assess whether the

¹⁷The figures in the Supplementary Appendix H also illustrate how measures of quantity and revenue productivity change with firm re-organization. The results are similar to Caliendo and Rossi-Hansberg (2012) and further discussed in the Online Appendix G.

¹⁸This mechanism also explains differences in worker skills and incomes across locations. In larger locations, workers are more skilled and earn a higher income because they are employed in firms with a greater number of layers.

model's predictions are valid in other market structures.

4 Data and Identification

The main empirical analysis of the paper examines French firms operating in 2004 in industries where demand is determined at the local level. In what follows, I describe the data and discuss several identification concerns. Details on the construction of the data and the main variables used are provided in the Data Appendix.

4.1 Data Sources

Information was obtained from three French sources, all provided by the French National Statistical Institute for Statistics and Economic Studies (INSEE). The first is the Déclarations Annuelles des Données Sociales (DADS), which is built from mandatory tax reports. The DADS is an annually matched employer-employee dataset, is nearly exhaustive, and contains information on all workers who earn a salary in France. For each worker, the data report the establishment and firm of employment, as well as their salary, number of hours, and occupation within the firm. In addition, the data report, for every firm, its industry, the establishments it operates, and the municipality (*commune*) of each establishment.

The second source is the Fichier Complet Unifié de Suse (FICUS), which is built from mandatory fiscal declarations. The FICUS data is also nearly exhaustive, and contains annual balance sheet information on firms, such as total sales, material assets, and value added, as well as information on their legal structure (e.g., legal status and whether the firm belongs to a business group).

Finally, the third source is the Recensement de la Population (RP), a septennial census of the French population. Up until the year 1999, the RP is exhaustive and contains demographic information on all individuals and households in France.

4.2 Data Restrictions

In order to properly examine the model's implications, I impose several restrictions on the data. First, the model is silent on whether a firm can have one or many establishments. To be consistent with the model, I aggregate the data at the firm-level. Second, while owners of very small firms commonly also work in their firm, they often do not disclose a salary, and thus do not appear in the DADS (Bacheré (2015)). This creates measurement error in my baseline estimates, biasing results. I consequently only retain firms with at least 8 employees.¹⁹

¹⁹More precisely, when I examine the second implication of the model, there is a downward bias in my estimates as firms with unreported workers have higher productivity and less layers. This problem is more acute in small firms and

Finally, the model concerns single product firms operating in a monopolistically competitive industry and competing in a local market. I thus restrict the main analysis to firms that operate in one local market and in one of the following industries: Clothing and Shoe Retail, Traditional Restaurants, and Hair and Beauty Salons. The model's general assumptions are applicable to these sectors: they are non-tradeable industries with relatively local demand. The assumption of monopolistic competition also applies reasonably well. Even in small markets there are several firms operating in these industries, each firm provides a differentiated product, and because of their location and quality of service, each firm retains some market power.

4.3 Firm Organization & Productivity

To construct the organization of firms I use the method developed by [Caliendo et al. \(2015b\)](#) and classify employees into organizational layers using the first digit of the occupational codes from the DADS.²⁰ Indeed, a unique feature of the French occupational codes is that they rank workers within firms. Occupational code 2 refers to workers with the highest level of authority: owners who receive a salary from the firm and corporate officers. Occupational code 3 indicates senior managers and directors, occupational code 4 supervisors and employees with a higher level of responsibility than ordinary workers, and occupational codes 5 and 6 (grouped together here to form one category), refer to ordinary workers. Using these occupational categories, a firm can thus be divided into as many as four distinct layers. To further measure the total number of layers in firms I follow [Caliendo et al. \(2015b\)](#) and focus specifically on the relative ranking of occupations. In other words, a firm reporting ℓ occupational categories will have ℓ layers. For example, a firm with workers in occupations 2,3, and 4 is a three-layer firm. Not all firms, however, have an employee in every occupational category, such that there are four types of organizational structures in the data: one-layer, two-layer, three-layer, and four-layer firms.

To measure the productivity of firms, I combine balance sheet data from the FICUS with measures on the size and wage bill of firms from the DADS. Using this information, I construct different estimates of firm productivity. My baseline analysis uses a measure of revenue-labor productivity, defined as value-added per worker. This measure is used in the empirical literature examining

cannot be corrected using information from the FICUS. In addition, in the French data, value-added per worker is higher in firms with a single employee than in other small firms, decreases as the number of employees increases, and tends to stabilize at 4 employees (see [Bacheré \(2015\)](#) for details). In the industries examined in this study, value-added per worker tends to stabilize at about 5 employees, while value-added per hour exhibits the same pattern and tends to stabilize at about 8 employees.

²⁰In the knowledge-based management theory of firms, a layer is composed of a group of workers with similar wages, who have similar skills, and perform tasks at a similar level of authority ([Caliendo et al. \(2015b\)](#)). [Caliendo et al. \(2015b\)](#) also provide evidence for the manufacturing sector that this classification is consistent with the concept of layer in the knowledge-based management theory of firms. I provide similar evidence below.

the productivity of service firms and is closely related to Figure 2b from the model simulations.²¹ To assess the robustness of my results, I also adopt a more structural approach and measure productivity as Total Factor Productivity (TFP). I use the methods proposed by Levinsohn and Petrin (2003) and Wooldridge (2009), to address concerns about endogenous factor inputs.²² As a final robustness check, I also estimate TFP using a method adapted from Caliendo et al. (2015a), which directly accounts for the organization of firms (details are provided in the Online Appendix B).

Moreover, it is important to note that the balance sheet data does not contain information on prices or units produced, such that my analysis is limited to measures of revenue productivity. Firms may thus appear more productive, because they set higher prices. While this is a common concern with measures of revenue productivity, addressing the latter is beyond the scope of this paper (see, however, Greenstone et al. (2010) for a solution in a different context).

4.4 Definition of Local Markets

The model relates the size of markets to firm organization. One challenge in examining its implications is defining local markets. In this paper, I apply two geographical decompositions of mainland France. My baseline analysis reports results from employment areas (*zones d'emploi*), composed of geographical spaces in which most of the inhabitants reside and work within the area. More specifically, employment areas reflect local labor markets, they correspond to cities and their surrounding areas, or to metropolitan areas, and each municipality (*commune*) belongs to a single employment area. I also report results from urban areas (*unité urbaines*), or geographical regions with at least 2,000 inhabitants in spaces separated by no more than 200 meters. Urban areas are contiguously built-up spaces, they correspond to small cities and their suburban areas, and do not cover all of France.

Both geographical decompositions provide a reasonable definition of local markets. It is sensible to assume that most individuals eat, shop, and visit salons within their area of residence and employment, such that the relevant definition of a local market is an employment area. If this definition is too broad, and people perform these activities within the city they reside in, defining

²¹There is a long-standing debate on the appropriate measure of firm productivity in the service sector, with many researchers advocating the use of value-added per worker (see Haskel and Sadun (2009) for a discussion on retail firms). Yet, one disadvantage of this measure is that it ignores the other inputs used in production (see Syverson (2011) for a discussion). Additionally, here I use the number of workers as a measure of the size of firms, as I suspect errors in the number of hours reported in service firms. The Online Appendix reports results with the number of hours.

²²In terms of the model, TFP implies that the parameter A is also heterogeneous across firms. A change in A would have two effects on firms. First, for a given level of quantity, because they require less labor to produce a unit of output, firms with a greater A will have fewer layers. Second, for a given demand curve, firms with a greater A will produce a larger quantity, which will force them to have a greater number of layers. Which effect dominates depends on the assumptions in the model. Furthermore, the data indicate that firms in denser markets produce with a greater number of layers and that these firms tend to be more productive, implying that the second effect dominates. Moreover, all measures of TFP are estimated over the period 2002-2006. Similar results are also obtained when estimating TFP using the methods proposed by Olley and Pakes (1996) and Akerberg et al. (2015).

markets at the urban level is more appropriate. As shown below, using either decomposition leads to the same conclusions. An alternative concern is that both definitions are too broad and encompass several submarkets. This problem should be of minor concern however, since in my analysis both the dependent and independent variables are averages, and are thus affected in the same way from defining markets too broadly (see Briant et al. (2010)).

A second challenge is measuring the size of local markets. Within a geographical area, the model abstracts from any spatial dimensions and assumes varieties are accessible to all consumers. In the model, however, market size determines the level of competition between firms. For this reason I measure the size of markets using population density in the year 1999, defined as the total population residing in an area divided by the area's surface, measured in hectares (1 hectare = 0.01 km²). Density is the more suitable measure in this setting, as it accounts for the concentration of economic activity and better reflects the level of competition in local markets.

4.5 Empirical Strategy

The model is static and involves comparisons of firms across markets of different sizes. To examine its implications, my empirical analysis focuses on the year 2004 and relies mainly on cross-sectional variation in the size of different geographical areas. More specifically, I estimate equations of the following form, at some level ℓ , in the year 2004:

$$Y_{\ell,a} = \alpha + \gamma \log density_a + X_{\ell,a}\beta + \epsilon_{\ell,a}, \quad (12)$$

where $Y_{\ell,a}$ is a variable based on the characteristics of firms, $density_a$ is the local density of an area a in the year 1999, and $X_{\ell,a}$ contains additional controls. Depending on the implication being examined, equation (12) is estimated at the geographical area level or at the level of firms.

The model has several restrictive assumptions that must be taken into consideration when examining its implications. First, the presence of a homogeneous good sector that determines the wage of a unit of labor. This assumption implies that the industries included in the analysis need to be relatively small, in the sense that the firms do not employ a large fraction of the local labor force. Unlike most industries, those in this study satisfy this condition: on average they employ 3.1% (3.3%) of the labor of an employment area (urban area).²³ This condition also has implications for my identification strategy: any controls of the characteristics of areas are unlikely to be determined by the industries in this study.

A second assumption is that the wage of a unit of labor is the same across markets. In the model,

²³Alternatively, the different areas can be ranked by the share of the local labor employed in these industries. At the 50th percentile, the share of an employment area's (urban area's) labor employed in these industries is 2.9% (2.4%). At the 75th percentile it is 3.6% (3.8%), and at the 99th percentile it is 7.7% (16.3%).

firms do not determine local wages, yet changes in the latter have an effect on firm organization. Local wages determine the cost of hiring workers and thereby the cost schedule faced by firms.²⁴ For this reason, my analysis also controls for the local wage of a unit of labor. To measure this variable, I use the remaining industries in the economy and estimate area fixed effects from a Mincerian hourly wage regression that controls for skills and occupations (see Appendix B.1 for details). These fixed effects do not contain the returns to skills, nor do they reflect information on workers' positions within firms.

Moreover, in the model wages also determine incomes, which govern consumer expenditures and thus have an effect on the demand schedule of firms.²⁵ Thus, in several specifications I also control for local incomes. To construct a measure of incomes in an area, I calculate the median annual salary from a sample of individuals who reside in the same area and who are not employed in the industries under study.

Another assumption of the model is that preferences are identical across markets. This assumption is important when trying to identify the relationship between the size of markets and the organizational structure of firms. In the model, preferences shape the demand schedule of firms. Following this observation, Appendix A.3 demonstrates that adjusting preferences can yield the same economic outcomes as an increase in market size. Moreover, in urban economics it is well-known that agents sort into locations in part because of their preferences (see, for example, [Holmes and Sieg \(2015\)](#) and [Rosenthal and Ross \(2015\)](#)). It is also well-known that both young individuals and immigrants tend to live in denser areas, and to the extent that their consumption habits are different, differences in density may correlate with differences in local preferences. To account for the latter, I include the following demographic variables, constructed from the RP data for the year 1999: share of the local population between the ages of 25 and 59, unemployment rate among local active individuals, and share of the local population born outside France.

Furthermore, a frequent concern with specifications similar to equation (12) is that the size of markets is endogenous. For instance, there may be local shocks that are not accounted for in the estimation, determining both density and the outcomes of firms. To address this issue, I follow [Ciccone and Hall \(1996\)](#), [Combes et al. \(2008\)](#), and [Combes et al. \(2010\)](#) and instrument for density using historical measures dating as far back as 1831. This strategy is defensible as long as there is some persistence in the spatial distribution of the French population, and the local determinants of contemporary firm organization are not related to the size of markets in the past ([Combes et al. \(2010\)](#)). To the extent that technology, methods of production, and the theory of management have

²⁴If the cost of a unit of labor increases in denser markets, some firms will reduce costs by simultaneously decreasing the knowledge of employees in the lower layers of the organization and increasing the number of managerial layers.

²⁵Holding the cost of labor fixed, a model with a higher income per consumer is isomorphic to a model with a larger market and a unit income.

over the years changed the structure of firms, using very long lags makes this hypothesis plausible (see [Combes et al. \(2010\)](#) for a discussion on change in the French economy since 1831).²⁶

Moreover, the same endogeneity concern also applies to the variables that control for local incomes and the demographic characteristics of areas. I consequently measure these additional variables in the year 1999. Insofar as the local shocks are of short duration and not correlated over time, these variables can be treated as exogenous. However, because the nature of the shocks and their persistence is unknown, I also instrument for the additional variables using lagged values.

To summarize, the empirical analysis is conducted for the year 2004 and focuses on firms with at least 8 workers, operating in a monopolistically competitive industry, and competing in one local market. I rely on cross-sectional variation in the size of geographical areas, include a set of controls to account for any other local factors that may also determine the number of layers and the productivity of firms, and use historical data to instrument for density and the different characteristics of areas. Furthermore, to avoid confusion over the different possible ways of describing a local market and since the implications of the model are examined on geographical areas, the remainder of the paper refers to local markets as areas. In the remainder of this section, I present descriptive statistics.

4.6 Descriptive Statistics

In the Online Appendix B, Table B2 presents some general statistics of the data. The data cover 341 employment areas and 1,289 urban areas in France. Across all employment areas (urban areas) there are 27,508 (24,197) firms, and the correlation between density and number of firms is 0.863 (0.421). The correlations between density and the characteristics of areas are also positive. In both samples, the majority of firms operate in the Traditional Restaurants industry, which accounts for 75% (73%) of firms across employment areas (urban areas). The samples in the data have a different number of firms because while all municipalities in mainland France belong to an employment area, urban areas do not contain every municipality within the country. For this reason, and since the conclusions are generally the same, I largely report results using the employment area geographical decomposition of mainland France.²⁷

²⁶This identification strategy is valid to the extent that firms do not sort into locations. The ideal empirical design to address this issue would require quasi-random assignment of firms to locations, as in [Greenstone et al. \(2010\)](#) (see [Duranton and Turner \(2017\)](#) for a detailed explanation). Such an approach is, however, beyond the scope of this paper. Rather, I assess the extent of bias caused from sorting, using the parametric test developed by [Oster \(2017\)](#).

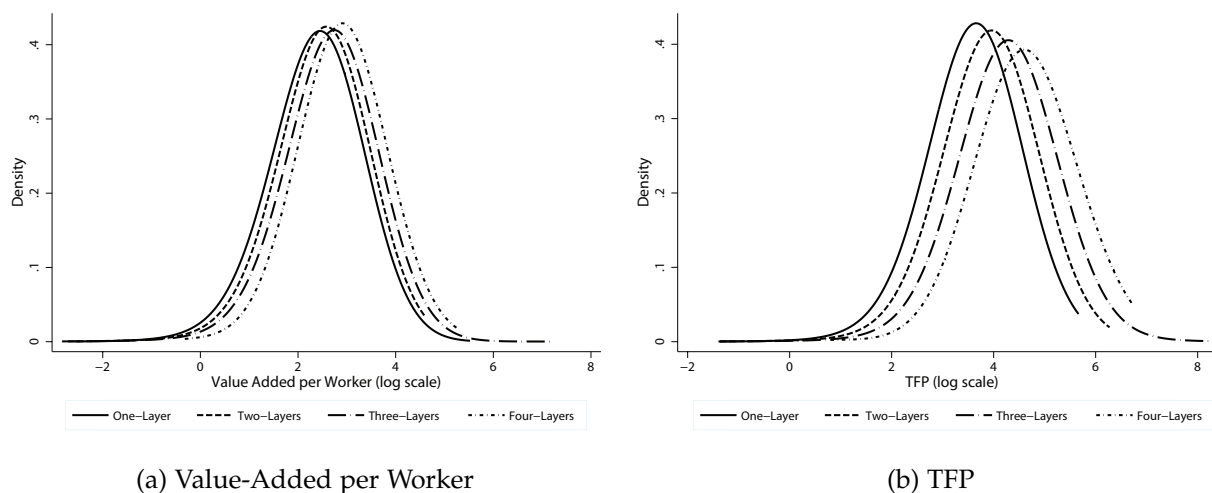
²⁷Online Appendices B.9 and B.10 provide descriptive statistics on the characteristics of firms across areas. The tables and figures illustrate that firms in denser areas are larger, more productive, and pay higher wages. These patterns are consistent with the theoretical model and the literature examining the productivity advantages of larger areas (see [Combes and Gobillon \(2015\)](#) for a review).

Table 1: Summary Statistics by Organization

Total Number of Layers	Number of Firms	Average Value Added	Average Number of Workers	Average Number of Hours	Average Hourly Wage
One-Layers	10,051	149.74	12.53	8,609	6.86
Two-Layers	12,509	232.33	17.28	12,922	7.79
Three-Layers	4,808	486.61	29.54	22,504	9.27
Four-Layers	590	888.29	45.35	36,678	10.66

Notes: Descriptive statistics of the observable characteristics of firms with the same number of layers. For each firm, the hourly wage is measured as: the total wage bill/total number of hours.

Figure 3: Kernel Density Productivity Distributions by Organization



Notes: Panel (a) illustrates the kernel density distribution of value-added per worker for one-layer, two-layer, three-layer, and four-layer firms. Panel (b) illustrates the kernel density distribution of estimates of firm TFP for one-layer, two-layer, three-layer, and four-layer firms. TFP is estimated using a Cobb-Douglas production function with capital and labor (number of workers), and the identifying assumptions of [Levinsohn and Petrin \(2003\)](#).

Descriptive Statistics: Firm Organization

The method developed by [Caliendo et al. \(2015b\)](#) to separate firms into layers of management has to date only been applied to firms operating in the manufacturing sector. Before proceeding to the main analysis, this subsection shows that using occupational categories to separate service firms into layers is also economically meaningful.

Table 1 first groups firms by their number of layers, and compares observable characteristics. The main takeaway is that firms with different organizational structures have distinct characteristics. Those with a greater number of layers are larger, in terms of value-added, number of workers, and number of hours, and pay higher hourly wages. These differences are also evident in Figures 3a and 3b, which plot the kernel density of value-added per worker and TFP estimated using the method from [Levinsohn and Petrin \(2003\)](#). Both figures show there is a ranking of distributions,

Table 2: Summary Statistics of Organizations

Total Number of Layers	$N_L^l \leq N_L^{l-1}, \forall l$	$N_L^2 \leq N_L^1$	$N_L^3 \leq N_L^2$	$N_L^4 \leq N_L^3$
Two-Layers	0.989	0.989	–	–
Three-Layers	0.857	0.962	0.894	–
Four-Layers	0.728	0.922	0.828	0.967
Total Number of Layers	$w_L^{l-1} \leq w_L^l, \forall l$	$w_L^1 \leq w_L^2$	$w_L^2 \leq w_L^3$	$w_L^3 \leq w_L^4$
Two-Layers	0.919	0.919	–	–
Three-Layers	0.798	0.937	0.859	–
Four-Layers	0.600	0.944	0.876	0.755

Notes: The share of firms that satisfy the different conditions. N_L^l denotes the number of workers in layer l in a firm with L total layers. w_L^l denotes the average hourly wage in layer l in a firm with L total layers, which is measured as the total wage bill in layer l /total number of hours in layer l .

with those firms operating with a greater number of layers being more productive.²⁸

Moreover, the classification of workers into layers is economically meaningful in that it reflects the hierarchical composition of firms. Table 2 reports that the fundamental property of hierarchies is present in the data: higher layers contain fewer workers who command greater wages. For instance, the top panel reports that nearly 85.7% of three-layer firms satisfy the condition that higher layers contains fewer workers, and the bottom panel indicates that nearly 79.8% of three-layer firms satisfy the condition that workers in higher layers earn greater wages.

Overall, these results are consistent with the management hierarchy models of [Garicano \(2000\)](#), and are comparable to the findings of [Caliendo et al. \(2015b\)](#). They also suggest that using occupational categories to separate firms into layers provides a reasonable representation of the hierarchical structure of service firms.

5 Empirical Analysis

5.1 Firm Organization Across Locations

I begin by examining the first main prediction of the model: firms in denser areas organize into a greater number of layers. The analysis proceeds in stages. I first group areas by their density and non-parametrically compare the distribution of firm organization in denser and less dense areas. I then turn to a regression analysis to examine the distribution of firm organization. Here and in the remaining subsections, I only report regression results in tables from my preferred specifications and include additional results in the Online Appendix.

²⁸In the Online Appendix B, additional figures illustrate a similar ranking of distributions using other characteristics of firms.

Table 3: Distribution of Firm Organization across Geographical Areas

	Number of Firms	One Layer	Two Layer	Three Layer	Four Layer
Employment Areas					
All	27,508	0.365	0.438	0.174	0.021
Below Median	4,428	0.464	0.412	0.114	0.009
Above Median	23,080	0.346	0.443	0.186	0.023
Urban Areas					
All	24,197	0.347	0.443	0.184	0.023
Below Median	2,054	0.474	0.422	0.097	0.005
Above Median	22,143	0.335	0.445	0.192	0.025

Notes: The share of firms with a given number of layers, across areas with below-median and above-median density.

Table 4: Mann-Whitney Distribution Tests

	Employment Areas		Urban Areas	
	Null Hypothesis: Distributions are Equal	Probability Above Median > Below Median	Null Hypothesis: Distributions are Equal	Probability Above Median > Below Median
ALL Industries	0.000	0.575	0.000	0.593
Clothing and Shoe Retail	0.001	0.547	0.264	0.527
Traditional Restaurants	0.000	0.578	0.000	0.598
Hair and Beauty Salons	0.000	0.585	0.000	0.595

Notes: Results of Mann-Whitney stochastic dominance test. The null hypothesis is that both distributions are equal. Columns (1) and (3) report the p-values of the test. Columns (2) and (4) report the probability that a random draw of a firm organization from areas with below-median density is greater than a random draw of a firm organization from areas with above-median density. Industries are grouped together based on their 3-digit NAF Rev 1 classification code.

Ordinal Stochastic Dominance Tests

Table 3 compares the distribution of firm organization across the different halves of the density distribution. In denser areas, a larger share of firms produce with a greater number of layers. For example, the top half of the employment area density distribution has a 0.014 percentage point higher share of four-layer firms, a 0.072 higher share of three-layer firms, and a 0.031 higher share of two-layer firms, relative to the bottom half of the distribution.

Table 4 further examines whether the cumulative distribution of firm organization in denser areas first-order stochastically dominates the distribution in less dense areas, using the Mann-Whitney test.²⁹ The underlying hypothesis of the test is that both distributions are identical, while the alternative is that one distribution has systematically larger values. Columns 1-2 in Table 4 compare the cumulative distribution of organizations across employment areas with above-median

²⁹The Mann-Whitney test relies on three assumptions: both samples are random, across and within sample observations are independent of one another, and the response variable is ordinal.

and below-median density, while columns 3-4 perform the same comparison across urban areas. Columns 1 and 3 report the p-values of the tests. In nearly every case the null hypothesis is rejected at conventional levels of significance.³⁰ Columns 2 and 4 report the probability that a randomly selected firm from an area with above-median density has a greater number of layers than a random firm from an area with below-median density. The results indicate that the cumulative distribution of firm organization in denser areas first-order stochastically dominates the distribution in less dense areas. For example when grouping all industries together, the Mann-Whitney test indicates that a randomly selected firm from an employment area with above-median density, is 57.5% more likely to have a greater number of layers than a random firm from an area with below-median density.

Regression Results: Distribution of Organizations

The Mann-Whitney test is reliable if, apart from density, all areas have the same characteristics. In reality, however, areas are different along many dimensions and these may be correlated with density and firm organization. Another concern, common in the urban economics literature, is that local shocks simultaneously determine firm organization and density. Both issues imply that the independence assumption of the Mann-Whitney test (observations are independent both across and within samples) may be violated. In what follows, I consequently examine the model's prediction using a regression analysis and address these additional concerns.

The simplest way to compare the distribution of firm organization is to estimate the relationship between density and the probability of firms having a given number of layers. More precisely, for firm i operating in industry k and area a , I estimate the following equation:

$$p_{i,k,a}^L = \alpha + \gamma \log density_a + X_{k,a}\beta + \epsilon_{i,k,a}, \quad (13)$$

where $p_{i,k,a}^L$ is an indicator variable that equals one if firm i has L layers, $density_a$ is the local density of an area in the year 1999, and $X_{k,a}$ contains the industry and area controls discussed above. According to the model, we should expect the estimates of γ to be negative for low values of L and positive for higher values.

Table 5 reports linear regression results with standard errors clustered at the employment area level. The unit of observation is a firm in an area, and every entry in the table reports a result from a separate regression.

³⁰Across urban areas, the Mann-Whitney test fails to reject the null hypothesis for the Clothing and Shoe Retail industry. Because there are fewer of these firms and their number in both groups is very unequal, the power of the Mann-Whitney test is lower for this industry (see Zimmerman (1987)). The probabilities reported, however, are comparable to the other industries, suggesting that a random firm from a denser area is likely to organize into a greater number of layers.

Column 1 reports OLS estimates from the most basic specification, which only controls for firm industry. The estimates have the expected sign and are significant at conventional levels when the dependent variable measures the probability of producing with one, three, and four layers. The magnitudes further imply that a firm operating in an employment area with a 100% higher density is 2.8% ($\approx -\ln(2) * 0.041$) less likely to operate with only one layer, 2.0% ($\approx \ln(2) * 0.029$) more likely to have three layers, and 0.5% ($\approx \ln(2) * 0.007$) more likely to be a four-layer firm. Put differently, increasing the density of an area the size of Lyon, the third most populated in city France, to an area the size of Paris, the most populated in city France, corresponds to a 0.147 percentage point decrease in the share of one-layer firms, a 0.104 point increase in the share of three-layer firms, and a 0.025 point increase in the share of four-layer firms.

Columns 2-3 include local area controls. Column 2 only controls for the cost of a unit of labor, while column 3 also controls for income and the demographic composition of employment areas. Relative to column 1, the estimated coefficients on density in column 2 are almost identical, implying that differences in the local cost of hiring workers do not play a prominent role in explaining the distribution of firm organization across locations. In contrast, in column 3 the magnitude of the estimated coefficients on density are lower than in column 1. This is not surprising because both income and preferences can affect the level of demand, and consequently the organization of firms. Nonetheless, the estimates have the expected sign and density continues to have an effect on the distribution of firm organization. The magnitudes in column 3 indicate that a firm in an area with a 100% higher density is 0.7% more likely to operate with three layers and 0.3% more likely to be a four-layer firm. In other words, increasing the density of an area the size of Lyon to an area the size of Paris corresponds to a 0.036 percentage point increase in the share of three-layer firms and a 0.014 point increase in the share of four-layer firms.

Columns 4-5 instrument for density and the local area controls. Because sample sizes are slightly different from columns 1 and 3, the first two columns of Table C1, in the Online Appendix C, report OLS results from the IV sample of firms. The magnitudes on density are similar. Column 4 returns to the specification in column 1 and instruments for density using long-lagged values measured in 1831, 1851, 1881, and 1901 with two-stage least squares (2SLS). When the dependent variable is the probability of having one and two layers, the magnitude of density is slightly lower than its OLS counterpart, implying that unobservable local shocks may be correlated with both the size of areas and the organization of firms. Density, however, continues to explain the distribution of firm organization.³¹ Column 5 further instruments for density and the local area controls except

³¹The instruments are not weak according to the test developed by Kleibergen and Paap (2006), and they all pass the Sargan-Hansen over-identification test at the 1% level. Because standard errors are clustered at the area level, to assess whether the instruments are weak, I rely on the following rule of thumb: the KP Wald F-statistic, from Kleibergen and Paap (2006), should be greater than 10 (see Staiger and Stock (1997), Stock et al. (2002), Stock and Yogo (2005), and

Table 5: Regression Results across Employment Areas

	(1)	(2)	(3)	(4)	(5)
Probability One-Layer Firm					
log density	-0.041 (0.002) ^a	-0.040 (0.002) ^a	-0.008 (0.006)	-0.040 (0.001) ^a	-0.013 (0.007) ^c
Probability Two-Layer Firm					
log density	0.004 (0.004)	0.004 (0.004)	-0.006 (0.006)	0.004 (0.004)	0.007 (0.007)
Probability Three-Layer Firm					
log density	0.029 (0.001) ^a	0.029 (0.002) ^a	0.010 (0.003) ^a	0.029 (0.002) ^a	0.005 (0.005)
Probability Four-Layer Firm					
log density	0.007 (0.000) ^a	0.006 (0.000) ^a	0.004 (0.000) ^a	0.007 (0.000) ^a	0.001 (0.001)
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	1463	10.01
R ² : One-Layer Firm	0.050	0.050	0.054	-	-
R ² : Two-Layer Firm	0.003	0.003	0.004	-	-
R ² : Three-Layer Firm	0.036	0.036	0.038	-	-
R ² : Four-Layer Firm	0.014	0.014	0.014	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	27,508	27,508	27,508	25,637	25,637

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (4) instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.098, 0.027, 0.369, and 0.067, when the dependent variable is the probability of a firm having respectively one, two, three, or four layers. Column (5) instruments for density and local characteristics using: density in 1831, 1881, and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, and the share of the population in 1968 residing in buildings built before 1949. The p-value of the over-id test is respectively 0.458, 0.177, 0.484, and 0.148, when the dependent variable is the probability of a firm having one, two, three or four layers.

for the cost of a unit of labor, as the hypothesis that it is exogenous cannot be rejected at conventional levels. The coefficients on density have the expected sign and are statistical indistinguishable from their OLS counterparts.³²

Overall, the evidence is consistent with the model. In denser areas, firms are less likely to be

[Murray \(2006\)](#) for details).

³²Although they pass the Sargan-Hansen over-identification test, the instruments just satisfy the rule of thumb advocated by [Staiger and Stock \(1997\)](#). In unreported results, I have also estimated the model with continuously updating GMM (CUE), which simulation evidence suggests is less sensitive to weak instruments than is 2SLS (see [Hahn et al. \(2004\)](#) and [Baum et al. \(2007\)](#)). The coefficients on density are less precisely estimated, but they remain consistent with the model.

composed of only one layer and more likely to be three and four-layer firms. These findings suggest that the cumulative distribution of firm organization in denser areas first-order stochastically dominates the distribution in less dense areas.

Table C1, in the Online Appendix C, further reports a set of robustness results. First, columns 3-4 estimate equation (13) using probit and logit models and report the marginal effect of density evaluated at the sample means. Second, to account for the fact that tasks are more specialized in denser areas (see [Stigler \(1951\)](#), [Garicano and Hubbard \(2007\)](#) and [Duranton and Jayet \(2011\)](#)) column 5 controls for the local degree of occupational concentration using a Herfindahl-Hirschman Occupational Index (HHI), whereas column 6 controls for the number of additional occupations in firms.³³ Third, to account for the fact that tasks may be more specialized in larger firms, columns 7-8 control for the size of firms. Finally, column 9 further simultaneously controls for several characteristics of firms in the data.³⁴ In all cases, the conclusions do not change. Density continues to affect the distribution of firm organization and the findings remain consistent with the model.

Moreover, the point estimates on density reported in column 3 of Table 5 are stable. This is reassuring in light of the concern that there may be omitted variable bias or that firms select into areas based on unobservable local factors. To further assess the extent to which the estimates may be biased from unobservable variables, I follow [Duranton and Turner \(2017\)](#) and implement the test proposed by [Oster \(2017\)](#) on the specifications of columns 3 in Table 5 and column 9 in Table C1. The results suggest that the influence of unobservable variables would need to be at least 4 times as important as the included set of controls for the coefficient of density to equal zero.³⁵

To summarize, the evidence is consistent with the model.³⁶ Even when controlling for the characteristics of areas and of firms, there is strong evidence that firms in denser areas operate with a greater number of layers.

³³One implication of the theoretical model is that there is greater vertical task specialization in denser areas, i.e., tasks within a firm that have different levels of authority. However, the model abstracts from horizontal task specialization, i.e., tasks within a firm with the same level of authority. Further, the correlation between the HHI index and the density is -0.261, implying that there is less occupational concentration in denser areas.

³⁴Along with firm size (the number of workers), column 9 controls for the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether they have a single establishment and whether they belong to a business group. In unreported results, I use measures of the endowments of employment areas as in [Combes et al. \(2016\)](#). The findings are similar.

³⁵This is greater than 1, the value recommended by [Oster \(2017\)](#). To conduct this test, I follow [Oster \(2017\)](#) and assume that $R_{\max}^2 = 1.3R^2$, and restrict the specification in column 3 to the same group of firms as in column 9.

³⁶In the Online Appendix C, the remaining tables consider different specifications. First, Tables C2-C3 report results using urban areas. The findings are similar and indicate that firms in denser areas organize into a greater number of layers. Second, Tables C4-C7 report regression results aggregated at the area-industry level. The conclusions remain the same. Finally, the remaining tables contain firm-level regression results with the dependent variable being the number of layers. In all cases, findings are consistent with the model and suggest that firms in denser areas organize into a greater number of layers. For instance, controlling for the characteristics of areas, the magnitudes in column 3 of Table C8 indicate that increasing the density of an employment area the size of Lyon to that of Paris, corresponds to an additional 0.093 layers in firms.

5.2 Organization & Productivity Gains from Denser Areas

I now turn to the second implication of the model and examine the extent to which differences in firm organization account for the productivity differences observed across locations.

To conduct my analysis, I adopt a three-step approach. In the first step, I obtain an estimate of the average productivity of areas while accounting for differences in the composition of firms. More precisely, for firm i operating in industry k and area a , I estimate the following equation:

$$\log \phi_{i,k,a} = \alpha + Z_{i,k,a}\delta + \theta_a + \epsilon_{i,k,a}, \quad (14)$$

where $\phi_{i,k,a}$ is a measure of firm productivity, θ_a are area fixed effects, and $Z_{i,k,a}$ contains controls of the characteristics of firms, including their industry. In the second step, I regress the area fixed effects on density as in the following equation:

$$\hat{\theta}_a = \alpha + \gamma \log \text{density}_a + X_a\beta + \epsilon_a, \quad (15)$$

where X_a contains local area controls. The parameter of interest, γ , measures the relationship between the average productivity of firms and the density of areas. To further determine the extent that firm organization contributes to this relationship, I compare the magnitudes of γ from two specifications, labeled Models 1 and 2. Model 1 never controls for the number of firm layers in equation (14), while Model 2 does. The difference in the point estimates of γ between Models 1 and 2 allows to assess how important firm organization is to the productivity differences across locations.³⁷

Before moving on to a comparison of Models 1 and 2, it is important to emphasize an additional prediction of theoretical model. In particular, the latter predicts that firms operating with a greater number of layers are on average more productive. Results from the first step of Model 2 are reported in the Online Appendix D. The findings are consistent with the model and imply that firms with an additional layer are on average 11.6% to 29.3% more productive.³⁸

I now turn to an examination of the point estimates of density from equation (15). Results are reported in Table 6. The unit of observation is an employment area and every entry in the table

³⁷Note that in equation (14), firm organization may be endogenous. More precisely, it may be the case that firm productivity and organization are simultaneously determined or that contemporaneous firm level shocks affect both variables. Although a common concern with many micro-level studies conducted at the firm-level, this issue is beyond the scope of this paper.

³⁸For a sense of the magnitudes, consider firms located in the third and fourth quartiles of the value-added per worker distribution. Average value-added per worker in the third quartile is 15.437, while in the fourth quartile it is equal to 26.193. In addition, firms on average organize with 1.869 layers in the third quartile, and 2.083 in the fourth. Using the relevant point estimates from Tables D1-D2, implies that differences in firm organization accounts for roughly 3.5% $\approx (\exp^{(0.116*(2.083-0.116)} - 1) * 15.437 / (26.193 - 15.437)$ to 3.8% $\approx (\exp^{(0.124*(2.083-0.116)} - 1) * 15.437 / (26.193 - 15.437)$ of the difference in average value-added per worker across the two quartiles of the productivity distribution. A similar exercise using the corresponding values for TFP yields magnitudes of 12.5% and 12.8%.

Table 6: Second-Step Regression Results across Employment Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.058 (0.008) ^a	0.045 (0.013) ^a	0.052 (0.008) ^a	0.042 (0.016) ^a	0.074 (0.013) ^a	0.075 (0.013) ^a
Model 2: With Org						
log density	0.045 (0.008) ^a	0.041 (0.013) ^a	0.041 (0.008) ^a	0.036 (0.015) ^b	0.062 (0.012) ^a	0.063 (0.012) ^a
Method	OLS	OLS	2SLS	2SLS	OLS	OLS
KP Wald F-Statistic	-	-	335	6.98	-	-
% Decrease	22.4	8.88	21.1	14.2	16.2	16.0
Sobel-ML Test (p-value)	0.000	0.000	0.000	0.010	0.001	0.001
R ² : Model 1	0.124	0.210	-	-	0.311	0.314
R ² : Model 2	0.080	0.162	-	-	0.206	0.208
2 nd -Step: Local Controls	No	Yes	No	Yes	Yes	Yes
1 st -Step: Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	341	341	304	304	340	340

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the estimate from a separate regression from equation (15). Sobel-ML Test is the multilevel mediation test from [Krull and MacKinnon \(1999\)](#) and [Krull and MacKinnon \(2001\)](#). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (3) instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.439 and 0.391 for Models 1 and 2. Column (4) instruments for density and local characteristics, except for the cost of a unit of labor, using: density in 1831, 1851, 1881, and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, the share of the population in 1968 residing in buildings built before 1949, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence. The p-value of the over-id test is 0.440 and 0.424 for Models 1 and 2.

reports a result from a separate regression.

In columns 1-4 productivity is measured as value-added per worker. Column 1 reports OLS estimates from the most basic specification, which only controls for the industry of firms in the first step. In both models, the estimated coefficient on density is positive and significant at conventional levels, implying that there are productivity gains from operating in denser areas. In Model 1, a 100% increase in density is associated with a 0.040 ($\approx 0.058 * \ln(2)$) increase in the average value-added per worker. Model 2 controls the number of layers in firms. Relative to Model 1, the magnitude of density is smaller, suggesting that part of the gains are accounted for by differences in firm organization. In Model 2, a 100% increase in density is associated with a 0.031 ($\approx 0.045 * \ln(2)$) increase in the average productivity of firms. This corresponds to a 22.4%

($\approx 100 * (0.058 - 0.045) / 0.058$) decrease in the magnitude of density estimated in Model 1.³⁹ In other words, differences in firm organization account for 22.4% of the productivity gains from operating in denser areas.

Column 2 further controls for the characteristics of areas in the second stage. Relative to column 1, the coefficients on density are lower, yet they remain positive and significant at conventional levels. A comparison of the point estimates from Models 1 and 2 indicates that 8.8% of the productivity gains are accounted for by differences in firm organization.

Columns 3-4 instrument for density and the characteristics of areas. Because the sample sizes are different, the first two columns of Table E1 report OLS results using the IV sample of firms. The magnitudes on density are lower than in columns 1-2, yet the conclusions remain the same. Column 3 returns to the specification in column 1 and instruments for density using long-lagged values. The magnitudes of density are statistically identical to their OLS counterparts, and suggest that differences in the organization of firms account for 21.1% of the gains from denser areas. Column 4 further instruments for density and the local area controls except for the cost of a unit of labor. The coefficients on density are again statistically indistinguishable from their OLS counterparts, and indicate that 14.2% of the productivity gains are accounted for by differences in firm organization.⁴⁰

Up to this point, the evidence is consistent with the model. The last two columns further examine whether the findings are robust across different measures of firm productivity. Column 5 measures productivity as the residual of a Cobb-Douglas production function with capital and labor, and estimated with the method from Levinsohn and Petrin (2003), which accounts for endogenous factor inputs. In column 6 the dependent variable is TFP estimated using the approach proposed by Wooldridge (2009), which also allows to identify the coefficient on labor in the estimation of firm TFP. The magnitudes on the estimated coefficients of density are similar across both methods, and the results indicate that differences in firm organization account for 16% of the productivity gains from operating in denser areas.

The remaining robustness checks are reported in the Online Appendix E.⁴¹ In all cases, the findings are consistent with the model, and confirm that differences in firm organization account for a meaningful share of the productivity gains from operating in denser areas.

³⁹To assess whether firm organization has a statistically significant effect on the productivity gains from denser areas, the tables also report a mediation test advocated by Krull and MacKinnon (1999) and Krull and MacKinnon (2001). While this test is similar to that of Sobel (Sobel (1982), Sobel (1986)), it accounts for the multilevel structure of the regressions. In the Online Appendix E, whenever the analysis is at the firm-level, I directly report results from Sobel's test.

⁴⁰In unreported results, the models have also been estimated with continuously updating GMM (CUE). The conclusions are similar.

⁴¹Table E1 controls for the additional characteristics of firms. Tables E2-E3 report results using urban areas. Tables E4-E7 report results at the firm-level. Tables E8-E9 report results using additional measures of firm productivity. Finally, Tables E10-E11 report separate results with firm organization fixed effects in equation (14), as well as, firm organization-industry fixed effects. In this last set of results, the estimated magnitudes on density are almost identical to those in Table 6.

Table 7: Extra Regression Results across Employment Areas

	CMORH LP TFP (1)	CMORH WD TFP (2)	VA Worker (3)	LP TFP (4)	WD TFP (5)
log density	0.048 (0.009) ^a	0.043 (0.009) ^a	0.027 (0.008) ^a	0.028 (0.008) ^a	0.028 (0.008) ^a
log density*1 Layer	-	-	0.015 (0.009)	0.013 (0.011)	0.013 (0.011)
log density*2 Layer	-	-	0.022 (0.011) ^b	0.014 (0.011)	0.014 (0.011)
log density*3 Layer	-	-	0.023 (0.035)	0.011 (0.034)	0.011 (0.034)
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	35.1	42.6	-	-	-
Test Interactions Terms					
on log density					
1 Layer = 2 Layer	-	-	0.540	0.944	0.959
1 Layer = 3 Layer	-	-	0.825	0.941	0.957
2 Layer = 3 Layer	-	-	0.969	0.913	0.936
R-Squared	0.168	0.135	0.237	0.343	0.337
2 nd -Step: Local Controls	Yes	Yes	-	-	-
1 st -Step: Industry FE	Yes	Yes	-	-	-
Industry & Local Controls	-	-	Yes	Yes	Yes
Firm Controls	-	-	Yes	Yes	Yes
Sample Size	340	340	26,791	23,459	23,459

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Entries display an estimate from a separate regression. Columns (1)-(2) report estimates from a separate regression from equation (15) with robust standard errors, with productivity estimated using the production function from [Caliendo et al. \(2015a\)](#) and the approaches from [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#). In columns (4)-(5) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Columns (1)-(2) compare the estimates with Model 1 from columns (5)-(6) of Table 6. Columns (3)-(5) estimate models in one step at the level of firms, and standard errors are clustered at the local area level. Along with the local area controls, columns (3)-(5) also control for firm size (the number of workers), the number of additional occupations, the amount of capital, the legal status of firms, whether firms have a single establishment, and whether they belong to a business group. All variables are interacted with the number of layers in firms.

5.3 Additional Results

I now turn to some additional results, as well as examine whether the findings are robust across different empirical specifications and market structures.⁴²

First, columns 1 and 2 of Table 7 report estimates of density using measures of firm TFP, based on the production function from [Caliendo et al. \(2015a\)](#). Consistent with the theoretical model, this production function takes the view that firm organization determines the efficiency of labor, and

⁴²Because the findings are similar across urban areas, here and in the remaining section I focus on employment areas.

directly controls for the number of layers in firms. These measures of firm TFP are also informative of the actual productivity gains from operating in denser areas, because they reflect productivity net of firm organization.⁴³ In column 1, TFP is estimated using the method from [Levinsohn and Petrin \(2003\)](#) while column 2 uses that from [Wooldridge \(2009\)](#) (details provided in the Online Appendix B). In both cases, the estimated coefficients on density are positive and significant, and are statistically smaller than the corresponding estimates in columns 5 and 6 of Table 6. The magnitudes imply that a 100% increase in density is associated with a 0.029 ($\approx 0.043 * \ln(2)$) to 0.033 ($\approx 0.048 * \ln(2)$) increase in the productivity of firms, which is 35.1% to 42.6% smaller than the relationships reported in Model 1, in columns 5 and 6 of Table 6.

Second, to assess whether the productivity gains from denser areas are the same across differently organized firms, columns 3-5 report results with all variables, including the characteristic of firms, interacted with their organization. Regressions are reported at the firm-level, and in column 3 the dependent variable is value-added per worker, while in columns 4-5 it is TFP. In all cases, the estimated coefficients on density are positive, yet most of the interaction terms are not statistically different from the zero. We also cannot reject the hypothesis that the interaction terms are different from one another, implying that the benefits from operating in denser areas are the same across differently organized firms. Put differently, using measures of revenue productivity, there is insufficient evidence to conclude that there are complementarities between the size of areas and the organization of service firms.

Third, an important assumption of the model is that demand is determined at the local level. Thus far, the empirical analysis has taken this assumption seriously and examined the model's implications on firms that operate in service industries with local demand. To assess whether the implications also hold in other market structures, the Online Appendix F reports results from the same empirical analysis conducted on manufacturing firms. Even though manufacturing demand is determined at a global level, it is reasonable to expect that due to local input-output linkages and firms' demand for local intermediate inputs, there is greater competition between manufacturing firms operating in denser areas.⁴⁴ The implications of the model should consequently also hold for the manufacturing sector. Table F1 reports regression results examining how manufacturing firm organization varies with the density of areas. The findings are consistent with the model and indicate that firms in denser areas organize into more layers. Table F2, which has the same structure

⁴³To be precise, for firm i in period t , TFP is equal to: $TFP_{it} = \exp(\ln va_{it} - \hat{\alpha}_k \ln k_{it} - \hat{\alpha}_C C(O_{it}, w) - \hat{\alpha}_O ORG_{it})$, where $\hat{\alpha}_k$, $\hat{\alpha}_C$, and $\hat{\alpha}_O$ are the estimated coefficients of capital, the wage bill, and the number of layers in firms. See Appendix B.5 for details.

⁴⁴There are, in addition to locating close to consumers, other factors that determine the agglomeration of manufacturing firms (see [Marshall \(1890\)](#), [Duranton and Puga \(2004\)](#), and [Combes et al. \(2012\)](#)). Nonetheless, the evidence suggests that input-output linkages are particularly important ([Ellison et al. \(2010\)](#)). Furthermore, the manufacturing data are constructed using the same approach as that used for the main industries of this study.

as Table 6, further reports estimates of density from Models 1 and 2. The estimated coefficients are positive and significant, yet they are smaller than the coefficients on density reported in Table 6. Nonetheless, the conclusions remain the same: 9.6% to 28.0% of the productivity gains from operating in denser areas are accounted for by the organizational differences in manufacturing firms.

Fourth, thus far the analysis has focused on the static gains from operating in denser areas. The Online Appendix G examines the model's implications for the short-term gains, estimated from specifications using short-term variation within firms. Consistent with Caliendo et al. (2015b) and Caliendo et al. (2015a), the analysis shows that different measures of firm productivity respond differently to a re-organization of a firm. For instance, while the inverse of average costs increases discontinuously when a firm adds a layer, revenue per worker decreases discontinuously.⁴⁵ To the extent that firm organization is determined by the density of areas, it then follows that the short-term gains may be greater or smaller once we account for the re-organization of firms. Using a panel dataset that covers the years 2000-2007, the Online Appendix G provides evidence consistent with this result. For instance, the analysis suggests that the short-term value-added per worker gains may be 7.1% to 11.1% greater once we account for changes in the number of layers in firms.⁴⁶

6 Conclusion

This paper documents a novel channel through which the size of local markets has an effect on the productivity of firms: differences in the internal organization of firms. To this end, I develop a theoretical model that yields two implications. First, the distribution of firm organization in larger markets first-order stochastically dominates the distribution in smaller markets. Second, firms in larger markets are more productive because they operate with a greater number of layers. Using administrative data, I also examine how the organization and the productivity of firms, operating in service industries with locally determined demand, vary across employment areas. I find that firms in denser areas operate with a greater number of layers and that firm organization is an important component of firm productivity. Finally, I observe that 8.8% to 22.4% of the productivity gains from operating in denser areas are explained by differences in firm organization. A separate analysis confirms that these results also extend to firms in the manufacturing sector.

These magnitudes are economically important. For a sense of their size, consider another im-

⁴⁵More precisely, let $n(\alpha)$ denote the total number of workers in a firm with demand draw α . One implication of the model is then that when a firm increases its number of layers: (i) the inverse of average costs, $1/AC(\alpha)$, increases discontinuously, (ii) output per worker, $q(\alpha)/n(\alpha)$, decreases discontinuously, (iii) revenue per worker, $r(\alpha)/n(\alpha)$, decreases discontinuously, and (iv) revenue per labor costs, $r(\alpha)/C(\alpha)$, decreases discontinuously. These relationships have implications for specifications estimating the gains from denser areas using short-term variation within firms.

⁴⁶The analysis also shows that the findings from the previous sections are not specific to the year 2004 and that they are robust to specifications that control for time-invariant characteristics of areas and of firms.

portant channel known to increase the productivity of areas: the spatial sorting of firms. While one mechanism cannot account for all the productivity advantages of larger areas, recent work has established that this channel separately accounts for approximately 50% of the latter (Gaubert (2018)). Although smaller, the effect of firm organization is roughly 17.6% ($\approx 8.8\%/50\%$) to 44.8% ($\approx 22.4\%/50\%$) of the size of the effect of sorting, which remains a tangible contribution.⁴⁷

Moreover, while my analysis focuses on the productivity advantages of larger areas, the results provide additional economic insights. More specifically, they imply that the division of labor, the share of managerial occupations, and the degree of task specialization - all of which are greater in larger areas - are related to the organization of knowledge in firms. These findings also have important policy implications. In particular, they suggest there is less scope for local intervention to increase the productivity of areas, because a share of the gains are already internalized and priced into firms.

More fundamentally, this paper emphasizes the importance of internal firm organization in shaping economic outcomes and provides a number of directions for future research. One reasonable conjecture is that firms' personnel policies and their wage structures are also related to their hierarchical organization. Consequently, the theory presented here could also have interesting implications with respect to the distribution of wages, as well as workers' learning and career trajectories across different areas. Furthermore, the framework of Garicano (2000) is based on a complementarity in the production of knowledge, which shapes economic outcomes and the productivity of firms. Similar to Combes et al. (2012), Behrens et al. (2014), and Gaubert (2018), future work might investigate how this complementarity interacts with other productive sources of areas to determine the size of cities, and the organization and the productivity of firms.⁴⁸ Finally, there is a need for empirical work to better understand the effect that firm organization has on the other outcomes of firms. My analysis has largely left untouched econometric issues related to the fact that most firm level variables are determined simultaneously. Future work might create new research designs and use exogenous variation to causally identify the effect of firm organization on the different outcomes of firms.

⁴⁷Note that this exercise simply compares the magnitude of the effects. It does not, however, imply that sorting along with firm organization account for 58.8% to 72.4% of the productivity advantages of denser areas. To make such a statement, my analysis would require a more general theoretical model, in which firms with different characteristics sort into different areas.

⁴⁸Such an analysis might also endeavor to separate the productivity advantages that arise from differences in firm organization from the other productive sources of larger areas. Such a framework could also generalizable to other sectors and provide additional channels through which the size of cities determines firm organization. Note that this paper's model directly applies to non-tradeable service industries, whose location patterns are predominantly determined by the location of consumers. These industries account, however, for a relatively small share of the economic activity of an area. It would consequently be of interest to investigate whether there are other channels that determine firm organization.

References

- Akerberg, Daniel A., Kevin Caves, and Garth Frazer. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83, 6: (2015) 2411–2451.
- Antras, Paul, Luis Garicano, and Esteban Rossi-Hansberg. "Offshoring in a Knowledge Economy." *Quarterly Journal of Economics* 121, 1: (2006) 31–77.
- . *Organizing Offshoring: Middle Managers and Communication Costs*. Cambridge, US: Harvard University Press, 2008, e. helpman, d. marin, t. vierdier eds. edition.
- Antras, Paul, and Esteban Rossi-Hansberg. "Organizations and Trade." *Annual Review of Economics* 43–64.
- Bacheré, Hervé. "Les Très Petites Entreprises, 2 Millions d' Unités Très Diverses." *Insee Focus* 24.
- Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. "Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing." *Stata Journal* 7, 4.
- Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. "Productive Cities: Sorting, Selection, and Agglomeration." *Journal of Political Economy* 122, 3: (2014) 507–553.
- Bellone, Flora, Patrick Musso, Lionel Nesta, and Frederic Warzynski. "International Trade and Firm-Level Markups when Location and Quality Matter." *Journal of Economic Geography* forthcoming.
- Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen. "The Distinct Effects of Information Technology and Communication Technology on Firm Organization." *Management Science* 60, 12: (2014) 2859–2885.
- Briant, Anthony, Pierre-Philippe Combes, and Miren Lafourcade. "Dots to Boxes: Do the Size and Shape of Spatial Units Jeopardize Economic Geography Estimations?" *Journal of Urban Economics* 67, 3: (2010) 287–302.
- Caicedo, Santiago, Jr. Lucas, Robert E., and Esteban Rossi-Hansberg. "Learning, Career Paths, and the Distribution of Wages." *American Economic Journal: Macroeconomics* 11, 1: (2019) 49–88.
- Caliendo, Lorenzo, Luca David Mion, Giordano Opromolla, and Esteban Rossi-Hansberg. "Productivity and Organization in Portuguese Firms." Working Paper 21811, National Bureau of Economic Research, 2015a.
- Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. "Anatomy of French Production Hierarchies." *Journal of Political Economy* 123, 4: (2015b) 809–852.
- Caliendo, Lorenzo, and Esteban Rossi-Hansberg. "The Impact of Trade on Organization and Productivity." *Quarterly Journal of Economic* 127, 3: (2012) 1393–1467.
- Charnoz, Pauline, Claire Lelarge, and Corentin Trevien. "Communication Costs and the Internal Organisation of Multi-plant Businesses: Evidence from the Impact of the French High-speed Rail." *The Economic Journal* 128, 610: (2018) 949–994.
- Ciccone, Antonio, and Robert E. Hall. "Productivity and the Density of Economic Activity." *American Economic Review* 86, 1: (1996) 54–70.

- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. "Spatial Wage Disparities: Sorting Matters." *Journal of Urban Economics* 63, 2: (2008) 723–742.
- . "The Costs of agglomeration: House and land prices in French cities.", 2016. Working Paper.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sebastien Roux. "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection." *Econometrica* 80, 6: (2012) 2543–2594.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sebastien Roux. "Estimating Agglomeration Economies with History, Geology, and Worker Effects." In *Agglomeration Economics*, edited by Edward L. Glaeser, University of Chicago Press, 2010, chapter 1, 15–66.
- Combes, Pierre-Philippe, and Laurent Gobillon. "The Empirics of Agglomeration Economies." In *Handbook of Regional and Urban Economics*, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, North-Holland, 2015, volume 5, chapter 5, 247–348.
- Dasgupta, Kunal. "Learning and Knowledge Diffusion in a Global Economy." *Journal of International Economics* 87, 2: (2012) 323–336.
- Duranton, Gilles, and Hubert Jayet. "Is the Division of Labour Limited by the Extent of the Market? Evidence from French Cities." *Journal of Urban Economics* 69, 1: (2011) 56–71.
- Duranton, Gilles, and Diego Puga. "Micro-Foundations of Urban Agglomeration Economies." In *Handbook of Regional and Urban Economics*, edited by J. Vernon Henderson, and Jacques-Francois Thisse, North-Holland, 2004, volume 4, chapter 10, 2063–2117.
- Duranton, Gilles, and Matthew A. Turner. "Urban Form and Driving: Evidence from U.S. Cities.", 2017. Working Paper.
- Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review* 100, 3: (2010) 1195–1213.
- Foster, Lucia, John Haltiwanger, and Chad Syverson. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98, 1: (2008) 394–425.
- Friedrich, Benjamin. "Trade Shocks, Firm Hierarchies and Wage Inequality.", 2015. Working Paper.
- Garicano, Luis. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economics* 108, 5: (2000) 874–904.
- Garicano, Luis, and Paul Heaton. "Information Technology, Organization, and Productivity in the Public Sector: Evidence from Police Departments." *Journal of Labor Economics* 28, 1: (2010) 167–201.
- Garicano, Luis, and Thomas N. Hubbard. "Managerial Leverage is Limited by the Extent of the Market: Hierarchies, Specialization, and the Utilization of Lawyer's Human Capital." *Journal of Law and Economics* 50, 1: (2007) 1–43.
- Garicano, Luis, and Esteban Rossi-Hansberg. "Inequality and the Organization of Knowledge." *American Economic Review Papers and Proceedings* 50, 1: (2004) 197–202.
- . "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics* 121, 4: (2006) 1383–1435.

- . “Knowledge-based Hierarchies: Using Organizations to Understand the Economy.” *Working Paper* .
- Gaubert, Cecile. “Firm Sorting and Agglomeration.” *American Economic Review* 108, 11: (2018) 3117–53.
- Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings.” *Journal of Political Economy* 118, 3: (2010) 536–598.
- Guadalupe, Maria, and Julie M. Wulf. “The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies.” *American Economic Journal: Applied Economics* 2, 4: (2010) 105–127.
- Gumpert, Anna. “The Organization of Knowledge in Multinational Firms.” *Working Paper* .
- Hahn, Jinyong, Jerry Hausman, and Guido Kuersteiner. “Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations.” *Econometrics Journal* 7, 1: (2004) 272–306.
- Handbury, Jessie, and David E. Weinstein. “Goods Prices and Availability in Cities.” *Review of Economic Studies* forthcoming.
- Haskel, Jonathan, and Raffaella Sadun. “Regulation and UK Retailing Productivity: Evidence from Micro Data.” IZA DP No. 4028.
- Holmes, Thomas J., and Holger Sieg. “Structural Estimation in Urban Economics.” In *Handbook of Regional and Urban Economics*, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, North-Holland, 2015, volume 5, chapter 2, 69–114.
- Kleibergen, Frank, and Richard Paap. “Generalized Reduced Rank Tests Using the Singular Value Decomposition.” *Journal of Econometrics* 133, 1: (2006) 97–126.
- Krull, Jennifer L., and David P. MacKinnon. “Multilevel Mediation Modeling in Group-Based Intervention Studies.” *Evaluation Review* 23, 4: (1999) 418–444.
- . “Multilevel Modeling of Individual and Group Level Mediated Effects.” *Multivariate Behavioral Research* 36, 2: (2001) 249–277.
- Levinsohn, James, and Kamil Petrin. “Estimating Production Functions Using Inputs to Control for Unobservables.” *Review of Economic Studies* 70, 2: (2003) 317–342.
- Marshall, Alfred. *Principles of Economics*. MacMillan, 1890.
- Melitz, Marc J., and Giancarlo I.P. Ottaviano. “Market Size, Trade and Productivity.” *Review of Economic Studies* 75, 1: (2008) 295–316.
- Murray, Michael P. “Avoiding Invalid Instruments and Coping with Weak Instruments.” *Journal of Economic Perspectives* 20, 4: (2006) 111–132.
- Olley, G. Steven, and Ariel Pakes. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica* 64, 6: (1996) 1263–1297.
- Oster, Emily. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 1–18.

- Ottaviano, Gianmarco P.I., Takatoshi Tabuchi, and Jacques-Francois Thisse. "Agglomeration and Trade Revisited." *International Economic Review* 43, 2: (2002) 409–435.
- Petrin, Amil, Brian P. Poi, and James Levinsohn. "Production function estimation in Stata using inputs to control for unobservables." *Stata Journal* 4, 2: (2004) 113–123.
- Raghuram, Rajan G., and Julie Wulf. "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies." *The Review of Economics and Statistics* 88, 4: (2006) 759–773.
- Rosen, Sherwin. "Specialization and Human Capital." *Journal of Labor Economics* 1, 1: (1983) 43–49.
- Rosenthal, Stuart S., and Stephen L. Ross. "Change and Persistence in the Economic Status of Neighborhoods and Cities." In *Handbook of Regional and Urban Economics*, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, North-Holland, 2015, volume 5, chapter 16, 1047–1120.
- Rosenthal, Stuart S., and William C. Strange. "Evidence on the Nature and Sources of Agglomeration Economics." In *Handbook of Regional and Urban Economics*, edited by J. Vernon Henderson, and Jacques-Francois Thisse, North-Holland, 2004, volume 4, chapter 49, 2119–2171.
- Rovigatti, Gabriele, and Vincenzo Mollisi. "PRODEST: Stata Module for Production Function Estimation Based on the Control Function Approach." Statistical Software Components, Boston College Department of Economics, 2016.
- Smith, Adam. *An Inquiry Into the Nature and Causes of the Wealth of Nations*. London, UK: W. Strahan and T. Cadell, 1776.
- Sobel, Michael E. "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models." *Sociological Methodology* 290–312.
- . "Some New Results on Indirect Effects and their Standard Errors in Covariance Structure Models." *Sociological Methodology* 16: (1986) 159–186.
- Spanos, Grigorios. "Organization and Export Performance." *Economics Letters* 146: (2016) 130 – 134.
- Staiger, Douglas, and James H. Stock. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65, 3: (1997) 557–586.
- Stigler, George J. "The Division of Labor is Limited by the Extent of the Market." *Journal of Political Economy* 59, 3: (1951) 185–193.
- Stock, James H., Jonathan H. Wright, and Motohiro Yogo. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20, 4: (2002) 518–529.
- Stock, James H., and Motohiro Yogo. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by Donald W. K. Andrews, and James H. Stock, Cambridge University Press, 2005, chapter 5, 80–108.
- Syverson, Chad. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy* 112, 6: (2004) 1181–1222.
- . "What Determines Productivity?" *Journal of Economic Literature* 49, 2: (2011) 326–365.

Wooldridge, Jeffrey M. "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables." *Economics Letters* 104, 3: (2009) 112–114.

Zimmerman, Donald W. "Comparative Power of Student t Test and Mann-Whitney U Test for Unequal Sample Sizes and Variances." *The Journal of Experimental Education* 55, 3: (1987) 171–174.

ONLINE APPENDIX - For Online Publication

A Appendix: Proofs and Additional Theoretical Results

This section contains proofs and additional theoretical results, presented in the following order:

1. Cost Minimization Problem: describes the cost minimization problem of firms.
2. Proofs of Propositions: contains the proofs from the different propositions in the paper.
 - (a) Proof of Proposition 2
 - (b) Proof of Proposition 3
 - (c) Proof of Proposition 4
 - (d) Proof of Proposition 5
 - (e) Proof of Proposition 6
 - (f) Proof of Proposition 7
3. Isomorphism Propositions: shows that changes in preferences can yield the same outcomes as changes in market size.

A.1 Cost Minimization Problem

Garicano (2000) characterizes the optimal and cost efficient way to organize production, in a framework where production requires the resolution of problems and it is costly to match problems with agents who know the solution. In this framework, Garicano (2000) shows that the optimal firm organization will have the following properties. First, production workers spend their time generating problems and trying to solve them, while managers spend all of their time solving other workers' problems. Second, the frequency of problems that agents solve is decreasing with their position in the firm. In other words, agents in lower layers acquire knowledge to solve the most frequent problems. Third, firm organizations never contain layers with overlapping intervals of knowledge. In other words, production workers learn to solve problems from 0 to z_L^1 , and managers in layer l learn to solve problems from z_L^{l-1} to z_L^l . And fourth, the size of each layer is optimally determined by the knowledge of agents in the layers below, the size of layer 1, and the communication technology. This is further described below. The intuition for these results is the following. The objective of firm organization is to efficiently match problems with agents who know the solution. With this structure, a firm ensures that managers with the knowledge to solve infrequent problems avoid common problems, and are able to leverage their knowledge over more problems.⁴⁹ In other words, this structure ensures that agents spend their time dealing with problems they specialize in solving. Furthermore, by not allowing agents' knowledge to overlap across the different layers, a firm eliminates redundancies and reduces costs. A firm's cost minimization problem takes this structure as given.

Firms compensate agents for their one unit of time and the knowledge they acquire, and wages are normalized to 1. Consider a one-layer firm and producing output q . In such a case, the only worker in the firm is the entrepreneur. Because she is alone, she optimally spends her time generating production problems. To increase her productivity, she acquires knowledge that allows her to solve the commonest problems. Therefore, her knowledge set ranges from 0 to z_1^1 , she is able to solve $F(z) = 1 - \exp^{-\lambda z_1^1}$ fraction of problems, and the expected output of the organization is: $A \left[1 - \exp^{-\lambda z_1^1} \right]$. It then follows that z_1^{*1} is equal to: $\frac{1}{\lambda} \ln \left(\frac{A-q}{A} \right)$. The firm's total cost is also equal to: $(cz_1^{*1} + 1)$.

Consider a firm with L layers, 1 layer of production workers and $L - 1$ layers of managers, and producing output q . In a firm with L layers, production workers are the only agents to generate problems. Thus, the number of available problems is equal to the number of agents in layer 1, n_1^1 . Additionally, the number of managers in each layer is determined by three factors: the fraction of time managers devote to listening to a problem, h , the number of problems generated in the firm, and the fraction of unsolved problems from agents in the layers below. In other words, the size, n_L^l , of managers in layer l is given by the following equation: $hn_L^l \exp^{-\lambda Z_L^{l-1}} = n_L^l$, where $Z_L^{l-1} = \sum_{i=1}^{l-1} z_L^i$ is the cumulative knowledge at layer l . The total output of a firm with L layers is determined by three components: the number of production problems generated in the firm, the fraction of problems solved, and the units of output generated from a solved problem, A . In other words, the total expected output of a L layer firm is equal to: $An_L^1 \left[1 - \exp^{-\lambda Z_L^1} \right]$. The firm's total cost is equal to: $\sum_{l=1}^L n_L^l [cz_L^l + 1]$.

Agents are compensated for their one unit of time and knowledge, and wages are normalized to 1. For a given number of layers, L , and a given quantity, q , the entrepreneur organizes her firm optimally, and decides on the number of agents to hire in each layer, n_L^l , and their knowledge, z_L^l ,

⁴⁹Intuitively, the efficient organizational structure is consistent with Rosen (1983): "the incentives for specialization ... arise from increasing returns to the utilization of human capital." The cost of acquiring knowledge is independent of how knowledge is used in the firm. To be efficient, a firm organizes production in such a way that maximizes the utilization rate of each agent's knowledge.

with the objective to minimize costs. The cost minimization problem of a firm with $L \geq 1$ layers is the following:

$$C_L(q) = \min_{\{n_l^l, z_l^l\} \geq 0} \sum_{l=1}^L n_l^l [cz_l^l + 1] \quad (16)$$

subject to

$$A \left[1 - \exp^{-\lambda Z_l^l} \right] n_l^l \geq q, \quad (17)$$

$$n_l^l = n_l^1 h \exp^{-\lambda Z_l^{l-1}} \text{ for } L \geq l \geq 2, \quad (18)$$

$$n_L^L = 1, \quad (19)$$

where $Z_l^l = \sum_{i=1}^l z_i^l$ is the cumulative knowledge at layer l . The first constraint indicates that the total output produced by the firm is at least q , the second constraint determines the size of each layer, while the last constraint ensures that the entrepreneur supplies all of her time to the firm.

The cost functions are derived in [Caliendo and Rossi-Hansberg \(2012\)](#). Here I simply provide their expressions. For a given number of layers, L , and output q , a firm's marginal cost is equal to:

$$MC_L(q) = \frac{ch}{\lambda A} \exp^{\lambda z_L^{*L}}. \quad (20)$$

where z_L^{*L} is the solution to the cost minimization problem. The total cost of a two-layer firm, producing output q , is equal to:

$$C_2(q) = \frac{c}{\lambda} \left(\frac{h}{A} \exp^{\lambda z_2^{*2}} q + \left(1 - \frac{\exp^{\lambda z_2^{*1}}}{h} \right) + \lambda z_2^{*2} + \frac{\lambda}{c} \right), \quad (21)$$

where z_2^{*2} and z_2^{*1} are the solutions to the cost minimization problem, and the total cost of a firm with $L \geq 3$ layers, producing output q , is equal to:

$$C_L(q) = \frac{c}{\lambda} \left(\frac{h}{A} \exp^{\lambda z_L^{*L}} q + \left(1 - \exp^{\lambda z_L^{*L-1}} \right) + \lambda z_L^{*L} + \frac{\lambda}{c} \right), \quad (22)$$

where z_L^{*L} and z_L^{*L-1} are the solutions to the cost minimization problem. Furthermore, the average costs of a firm with L layers and producing output q , is equal to: $AV_L(q) = \frac{C_L(q)}{q}$.

A.2 Proofs of Propositions

Proof of Proposition 2

We first derive the results for quantity $q(\alpha)$. We have:

$$\frac{\partial q(\alpha)}{\partial \alpha} = \frac{N}{2\gamma} \left[1 - \frac{\partial MC(q(\alpha))}{\partial q(\alpha)} \frac{\partial q(\alpha)}{\partial \alpha} \right].$$

Rearranging yields:

$$\frac{\partial q(\alpha)}{\partial \alpha} = \frac{\frac{N}{2\gamma}}{1 + \frac{\partial MC(q(\alpha))}{\partial q(\alpha)} \frac{N}{2\gamma}}.$$

Since within layers $\frac{\partial MC(q(\alpha))}{\partial q(\alpha)} > 0$, quantity is increasing with α .

Now moving onto prices. Since

$$\frac{\partial p(\alpha)}{\partial \alpha} = \frac{1}{2} \left[1 + \frac{\partial MC(q(\alpha))}{\partial q(\alpha)} \frac{\partial q(\alpha)}{\partial \alpha} \right],$$

and within layers, marginal costs are increasing with quantity, and quantity is increasing with α , $p(\alpha)$ is increasing with respect to α .

Now moving onto markups over marginal costs. Since

$$\frac{\partial \mu^{MC}(\alpha)}{\partial \alpha} = \frac{1}{2} \left[1 - \frac{\partial MC(q(\alpha))}{\partial q(\alpha)} \frac{\partial q(\alpha)}{\partial \alpha} \right],$$

and by substituting the expression for $\frac{\partial q(\alpha)}{\partial \alpha}$, within layers $\mu(\alpha)$ is increasing with respect to α .

Now moving onto markups over average costs. By definition markups over marginal costs, and markups over average costs are equal to:

$$\mu^{MC}(\alpha) = p(\alpha) - MC(q(\alpha)),$$

$$\mu^{AC}(\alpha) = p(\alpha) - AC(q(\alpha)).$$

It therefore follows that:

$$MC(q(\alpha)) - AC(q(\alpha)) = \mu^{AC}(\alpha) - \mu^{MC}(\alpha).$$

Marginal costs are increasing with quantity, while average cost curves are convex and attain their minimum when they intersect their associated marginal cost curve. It therefore follows that:

$$\frac{\partial [MC(q(\alpha)) - AC(q(\alpha))]}{\partial \alpha} = \frac{\partial [MC(q(\alpha)) - AC(q(\alpha))]}{\partial q(\alpha)} \frac{\partial q(\alpha)}{\partial \alpha} > 0,$$

which in turn implies that:

$$\frac{\partial \mu^{AC}(\alpha)}{\partial \alpha} > \frac{\partial \mu^{MC}(\alpha)}{\partial \alpha}.$$

Since $\frac{\partial \mu^{MC}(\alpha)}{\partial \alpha} > 0$ it follows that $\frac{\partial \mu^{AC}(\alpha)}{\partial \alpha} > 0$.

Now moving onto revenues. Since

$$\frac{\partial r(\alpha)}{\partial \alpha} = q(\alpha) \frac{\partial p(\alpha)}{\partial \alpha} + p(\alpha) \frac{\partial q(\alpha)}{\partial \alpha},$$

it follows that within layers, revenues are increasing with respect to α .

Now moving onto profits. From the maximization problem we know that,

$$\pi(\alpha) = \left[\alpha - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - \frac{\gamma}{N} q^*(\alpha) \right] q^*(\alpha) - C(q^*(\alpha)),$$

where * denotes the optimal quantities chosen. By the envelope theorem,

$$\frac{\partial \pi(\alpha)}{\partial \alpha} = q^*(\alpha).$$

Thus profits are increasing with respect to α . Since the optimal quantity produced, $q^*(\alpha)$ is increasing with the number of layers L , it follows that the slope of the profit function $\pi(\alpha)$ is increasing with L . And since,

$$\frac{\partial^2 \pi(\alpha)}{\partial^2 \alpha} = \frac{\partial q^*(\alpha)}{\partial \alpha}.$$

$\frac{\partial q^*(\alpha)}{\partial \alpha}$ is positive, profits are convex.

I now show that holding the number of layer fixed, profits are concave with respect to q . We know that

$$\pi(\alpha) = p(\alpha)q(\alpha) - C(q(\alpha)).$$

Substituting in for $p(\alpha) = \alpha - \frac{2\gamma}{N}q - \frac{\eta M}{\eta M + \gamma}(\bar{\alpha} - \bar{p})$, and taking the second derivative with respect to q yields:

$$\frac{\partial^2 \pi(\alpha)}{\partial^2 q} = -\frac{2\gamma}{N} - \frac{\partial MC(q(\alpha))}{\partial q},$$

which is negative. Thus profits are concave in q .

Proof of Proposition 3

I first derive the results with respect to quantity $q(\alpha)$. Since

$$\frac{\partial q(\alpha)}{\partial MC} = -\frac{N}{2\gamma},$$

when the firm increases the number of layers, marginal costs decrease discontinuously and quantity increases discontinuously.

Now onto prices. Since

$$\frac{\partial p(\alpha)}{\partial MC} = \frac{1}{2},$$

when the firm increases the number of layers, marginal costs decrease discontinuously, and thus prices decrease discontinuously as well.

Now moving on to markups over marginal costs. Since

$$\frac{\partial \mu^{MC}(\alpha)}{\partial MC} = -\frac{1}{2},$$

when the firm increases the number of layers, marginal costs decrease discontinuously, and thus markups over marginal costs increase discontinuously as well.

Now moving onto markups over average costs. Markups over average costs are also equal to: $\mu^{AC}(\alpha) = \pi(\alpha)/q(\alpha)$. The numerator of the derivative of this expression with respect to marginal costs is equal to:

$$q(\alpha) \frac{\partial \pi(\alpha)}{\partial MC} + \pi(\alpha) \frac{\partial q(\alpha)}{\partial MC}.$$

The first term is equal to: $q(\alpha) \frac{\partial \pi(\alpha)}{\partial MC} = q(\alpha) \left[-\frac{N}{2\gamma} MC(q(\alpha)) - q(\alpha) \right]$, where I have used the expression for $\frac{\partial r(\alpha)}{\partial MC}$ derived below. The second term is equal to $\pi(\alpha) \frac{\partial q(\alpha)}{\partial MC} = -\pi(\alpha) \frac{N}{2\gamma}$. Using the expression for profits, $\pi(\alpha) = \frac{\gamma}{N} q(\alpha)^2 + q(\alpha) MC(q(\alpha)) - C(q(\alpha))$, and eliminating common terms

implies that the numerator is equal to:

$$-\frac{q(\alpha)^2}{2} - \frac{N}{2\gamma}C(q(\alpha)),$$

which is negative. The denominator of the derivative of $\mu^{AC}(\alpha) = \pi(\alpha)/q(\alpha)$ with respect to marginal costs is equal to $q(\alpha)^2$ which is always positive. It therefore follows that when the firm increases the number of layers, marginal costs decrease discontinuously, and thus markups over average costs increase discontinuously as well.

Now moving onto revenues. Since

$$\frac{\partial r(\alpha)}{\partial MC} = -\frac{N}{2\gamma}MC(q(\alpha)),$$

as the firm increases the number of layers, marginal costs decrease discontinuously, revenues increase discontinuously as well.

Now moving onto profits. To show that profits are strictly increasing with respect to α and convex the arguments are similar to the proof of Proposition 2. Here I simply show that profits are continuous. Consider an entrepreneur who is indifferent between producing with layers L and $L + 1$. Then it follows:

$$\pi_L(\alpha) = \pi_{L+1}(\alpha).$$

Since, within layers profits are continuous and when an entrepreneur is indifferent between layers L and $L + 1$ profits are equal, thus profits are globally continuous with respect to α .

Proof of Proposition 4

The equilibrium of the model is determined from the zero-profit condition and the free-entry condition:

$$\pi(\alpha_D, M) = 0, \tag{23}$$

$$\int_{\alpha_D} \pi(\alpha, M) dG(\alpha) = f_E, \tag{24}$$

where M denotes the mass of firms operating in equilibrium and α_D is the demand draw of the entrepreneur who is indifferent between entering and exiting the market.

First I transform the equilibrium to be a function of q_D and α_D . From the first order condition of the firm's maximization problem, for a given α , quantity is determined by the equation:

$$q(\alpha) = \frac{N}{2\gamma} \left[\alpha - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC(q(\alpha)) \right].$$

For a given α there exists a unique quantity $q(\alpha)$, that is a solution to the expression of above. Rewriting this equation yields an expression of the term $\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})$ as a function the demand draw, α , and the optimal quantity produced, $q(\alpha)$:

$$\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) = \alpha - MC(q(\alpha)) - \frac{2\gamma}{N} q(\alpha).$$

Substituting this expression in the profit of the firm yields:

$$\pi(\alpha, q(\alpha)) = \frac{\gamma}{N} q(\alpha)^2 + MC(q(\alpha))q(\alpha) - C(q(\alpha)).$$

Doing the same for the marginal firm yields the expression for profit:

$$\pi(q_D) = \frac{\gamma}{N} q_D^2 + MC(q_D)q_D - C(q_D).$$

For a firm with demand draw α , I rewrite quantities, prices, markups and revenues as a function of q_D , p_D , μ_D and r_D , and the parameters of the model. This yields the following expressions:

$$\begin{aligned} q(\alpha) &= q_D + \frac{N}{2\gamma} [\alpha - \alpha_D + MC(q_D) - MC(q(\alpha))], \\ p(\alpha) &= p_D + \frac{1}{2} [\alpha - \alpha_D - MC(q_D) + MC(q(\alpha))], \\ \mu(\alpha) &= \mu_D + \frac{1}{2} [\alpha - \alpha_D + MC(q_D) - MC(q(\alpha))], \\ r(\alpha) &= r_D + q_D \frac{1}{2} [\alpha - \alpha_D - MC(q_D) + MC(q(\alpha))] + p_D \frac{N}{2\gamma} [\alpha - \alpha_D + MC(q_D) - MC(q(\alpha))] \\ &\quad + \frac{N}{4\gamma} [(\alpha - \alpha_D)^2 - (MC(q_D) - MC(q(\alpha)))^2]. \end{aligned}$$

The equilibrium is now determined by the solution to the following three equations:

$$\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) = \alpha_D - MC(q_D) - \frac{2\gamma}{N} q_D. \quad (25)$$

$$ZCP = \pi(\alpha_D, q_D) = 0, \quad (26)$$

$$FE = \int_{\alpha_D} \pi(\alpha, \alpha_D, q_D) dG(\alpha) - f_E = 0, \quad (27)$$

Here M , q_D and α_D are variables that are to be determined. Note that the solution to equation (26) depends only on q_D and the parameters of the model. Given the solution to equation (26), equation (27) is only a function of α_D and the parameters of the model. Finally, once q_D and α_D are both determined, M is determined equation (25). Therefore to prove that a solution exists, I need to show that there exists a q_D and α_D such that equations (26) and (27) are satisfied, and that a unique solution exists to (25).

First, I show that a solution to equation (26) exists. First, consider the slope of the profit function:

$$\frac{\partial \pi(q_D)}{\partial q_D} = \frac{2\gamma}{N} q_D + q_D \frac{\partial MC(q_D)}{\partial q_D} > 0.$$

When q_D is sufficiently large, $MC(q_D) > AC(q_D)$ and it follows that $\pi(q_D) = \frac{2\gamma}{N} q_D^2 + q_D MC(q_D) - C(q_D) > 0$. Since $\lim_{q_D \rightarrow 0} C(q_D) = w$ and $\lim_{q_D \rightarrow 0} MC(q_D) = 0$ it follows that $\lim_{q_D \rightarrow 0} \pi(q_D) < 0$. Therefore, there exists a unique q_D exists such that $\pi(q_D) = 0$.

Second, consider the equation (27). Using Leibniz's integral rule, the slope of the free-entry condition is:

$$\frac{\partial FE}{\partial \alpha_D} = -\pi(\alpha, \alpha_D, q_D) dG(\alpha_D) + \int_{\alpha_D} \frac{\partial \pi(\alpha, \alpha_D, q_D)}{\partial \alpha_D} dG(\alpha).$$

The first term by definition is equal to zero. I now show that the second term is positive. Using the expression for profits, and after eliminating common terms, it follows that:

$$\frac{\partial \pi(\alpha, \alpha_D, q_D)}{\partial \alpha_D} = \left[\frac{2\gamma}{N} q(\alpha) + q(\alpha) \frac{\partial MC(q(\alpha))}{\partial q(\alpha)} \right] \frac{\partial q(\alpha)}{\partial \alpha_D},$$

where from the expression of quantity it follows that:

$$\frac{\partial q(\alpha)}{\partial \alpha_D} = -\frac{\frac{N}{2\gamma}}{1 + \frac{N}{2\gamma} \frac{\partial MC(q(\alpha))}{\partial q(\alpha)}} < 0.$$

Therefore FE is downward sloping. Further, when $\alpha_D = \alpha_M$, $FE > 0$, and in the limit, when α_D approaches infinity $\lim_{\alpha_D \rightarrow \infty} FE < 0$. Hence, there exists a unique α_D such that $FE = 0$.

For a given number of layers L , there exists a solution to $\pi_L(\alpha_D^L, q_D^L) = 0$, and thus there are a discrete set of potential solutions. I now show that from this set, there is only one combination of α_D^L, q_D^L that satisfies the equilibrium.

Suppose not. Consider two possible solutions α_D^L, q_D^L and $\alpha_D^{L+1}, q_D^{L+1}$, associated with organizations with L and $L+1$ layers respectively. Without loss of generality, assume that $\alpha_D^L < \alpha_D^{L+1}$. In this case it follows that $\pi_L(\alpha_D^L, q_D^L) = \pi_{L+1}(\alpha_D^{L+1}, q_D^{L+1}) = 0$. By Proposition 2 it follows that all firms with demand draws in the interval $[\alpha_D^L, \alpha_D^{L+1}]$, will earn positive profits producing with an organization with L layers. By Proposition 2 it also follows that for the entrepreneur with demand draw α_D^{L+1} , $\pi_L(\alpha_D^{L+1}) > \pi_{L+1}(\alpha_D^{L+1}) = 0$, and so he will earn positive profits producing with an organization with L layers. Therefore $\alpha_D^{L+1}, q_D^{L+1}$ is not an equilibrium solution.

I now show that a unique M exists that satisfies equation (25). Since prices can be written as a function of α_D and q_D , it follows that \bar{p} is solely a function of α_D and q_D . Hence because the right-hand side of equation (25) is constant, while the left-hand side is increasing with respect to M , a unique solution exists and M is equal to:

$$M = \frac{\gamma}{\eta} \frac{\alpha_D - MC(q_D) - \frac{2\gamma}{N} q_D}{(\bar{\alpha} - \bar{p}) - \alpha_D + MC(q_D) + \frac{2\gamma}{N} q_D}$$

In equilibrium labor markets also clear. Labor is used for several purposes, as workers in the homogeneous sectors, as workers and managers in the differentiated good sector, as teachers, and to design new products. Let H be the mass of workers in the homogeneous good sector. As the total mass of agents in the economy is given by N , the labor market clearing condition is given by:

$$H + \frac{M}{1 - G(\alpha_D)} \left[f_E + \int_{\alpha_D} C(q(\alpha)) dG(\alpha) \right] = N. \quad (28)$$

I now show that if $\eta > \underline{\eta}$ both the homogeneous and differentiated goods will be produced in equilibrium. For simplicity, I define $K = \alpha_D - MC(q_D) - \frac{2\gamma}{N} q_D$ and $B = \int_{\alpha_D} \frac{\gamma}{N} q(\alpha) \frac{g(\alpha)}{1 - G(\alpha_D)}$. First consider the term $\bar{\alpha} - \bar{p}$. This can be rewritten as:

$$\begin{aligned}\bar{\alpha} - \bar{p} &= \int_{\alpha_D} \left[\alpha - \left[\alpha - K - \frac{\gamma}{N} q(\alpha) \right] \right] \frac{g(\alpha)}{1 - G(\alpha_D)} \\ &= K + B.\end{aligned}$$

It then follows that:

$$\begin{aligned}\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) &= \frac{\eta M}{\gamma + \eta M} [K + B] \\ &= K.\end{aligned}$$

And by isolating terms it follows that:

$$M = \frac{\gamma K}{\eta B}. \quad (29)$$

Next consider the equilibrium condition:

$$N - \frac{M}{1 - G(\alpha_D)} \left[[1 - G(\alpha_D)] f_E + \int_{\alpha_D} C(q(\alpha)) dG(\alpha) \right] > 0.$$

which can simply be rewritten as $N > M\bar{r}$. By substituting in the expression for M from above, it follows that:

$$\eta > \frac{\gamma K \bar{r}}{NB}. \quad (30)$$

Thus if $\eta > \underline{\eta} = \frac{\gamma K \bar{r}}{NB}$ both the homogeneous and differentiated goods will be produced in equilibrium.

Proof of Proposition 5

Consider an increase in N . The proof proceeds in steps. I first show that q_D and α_D increase with N . I then show that the term $\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})$ increases with N . And finally, I show under what conditions M increases with N .

From the zero-profit equation, it follows that:

$$0 = -\frac{\gamma}{N^2} q_D^2 + \frac{2\gamma}{N} q_D \frac{\partial q_D}{\partial N} + q_D \frac{\partial MC(q_D)}{\partial q_D} \frac{\partial q_D}{\partial N},$$

which after rearranging terms yields the result:

$$\frac{\partial q_D}{\partial N} = \frac{\frac{\gamma}{N^2} q_D}{\frac{2\gamma}{N} + \frac{\partial MC(q_D)}{\partial q_D}} > 0, \quad (31)$$

Since the denominator and numerator are both positive.

Now, consider the equation characterizing the expected profits of entry V^e :

$$\int_{\alpha_D} \pi(\alpha, \alpha_D, q_D) dG(\alpha) = f_E.$$

From this equation, it follows that:

$$\frac{\partial \alpha_D}{\partial N} = - \frac{\partial V^e / \partial N + \partial V^e / \partial q_D \partial q_D / \partial N}{\partial V^e / \partial \alpha_D}. \quad (32)$$

In the proof of Proposition 4, I showed that the denominator in equation (32) is negative. I now show that the numerator is positive. The profit of a firm with demand draw α is:

$$\pi(\alpha, \alpha_D, q_D) = \left[\frac{2\gamma}{N} q_D + MC(q_D) - \alpha_D + \alpha - \frac{\gamma}{N} q(\alpha) \right] q(\alpha) - C(q(\alpha)). \quad (33)$$

Using the envelope theorem and taking the derivative with respect to N and q_D yields:

$$\begin{aligned} \frac{\partial \pi(\alpha, \alpha_D, q_D)}{\partial N} + \frac{\partial \pi(\alpha, \alpha_D, q_D)}{\partial q_D} \frac{\partial q_D}{\partial N} &= \left[\frac{2\gamma}{N} q_D + MC(q_D) - \alpha_D + \alpha - \frac{\gamma}{N} q(\alpha) \right] q(\alpha) - C(q(\alpha)) \\ &= \left[-\frac{2\gamma}{N^2} q_D + \frac{2\gamma}{N} \frac{\partial q_D}{\partial N} + \frac{\partial MC(q_D)}{\partial N} \frac{\partial q_D}{\partial N} + \frac{\gamma}{N^2} q(\alpha) \right] q(\alpha) \\ &= \left[\frac{\gamma}{N^2} (q(\alpha) - q_D) \right] q(\alpha), \end{aligned} \quad (34)$$

where here α_D is held fixed and I substituted for $\frac{\partial q_D}{\partial N}$ using equation (31). Therefore since

$$\frac{\partial V^e}{\partial N} + \frac{\partial V^e}{\partial q_D} \frac{\partial q_D}{\partial N} = \int_{\alpha_D} \left[\frac{\partial \pi(\alpha, \alpha_D, q_D)}{\partial N} + \frac{\partial \pi(\alpha, \alpha_D, q_D)}{\partial q_D} \frac{\partial q_D}{\partial N} \right] dG(\alpha)$$

the numerator in equation (32) is positive and α_D is increasing with respect to N .

Rearranging the first order condition of the firm's maximization problem gives:

$$\alpha_D - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) = \frac{2\gamma}{N} q_D + MC(q_D),$$

and taking the derivative of this expression with respect to N yields:

$$\frac{\partial \alpha_D}{\partial N} - \frac{\partial \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})}{\partial N} = -\frac{\gamma}{N^2} q_D.$$

Hence it follows that:

$$\frac{\partial \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})}{\partial N} = \frac{\partial \alpha_D}{\partial N} + \frac{\gamma}{N^2} q_D > 0.$$

I now show that if $\bar{\eta} > \eta$ the mass of firms in the differentiated goods sector, M , will increase with N . In the proof of Proposition 4, I showed that M can be rewritten as:

$$M = \frac{\gamma K}{\eta B}.$$

where $K = \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p})$ and $B = \int_{\alpha_D} \frac{\gamma}{N} q(\alpha) \frac{g(\alpha)}{1-G(\alpha_D)}$. Taking the derivative of this expression with respect to N yields:

$$\frac{\partial M}{\partial N} = \frac{\gamma}{\eta B^2} \left[B \frac{\partial K}{\partial N} - K \frac{\partial B}{\partial N} \right].$$

Hence $\frac{\partial M}{\partial N}$ is positive if and only if $B\frac{\partial K}{\partial N} - K\frac{\partial B}{\partial N}$ is positive. I do not know the sign of $\frac{\partial B}{\partial N}$. If it is negative then it automatically follows that $\frac{\partial M}{\partial N}$ is positive, and I do not have to make an assumption on η . However assume that $\frac{\partial B}{\partial N}$ is positive. Then using the expression for M from above yields:

$$\frac{\frac{\partial K}{\partial N}}{\frac{\partial B}{\partial N}} > \frac{K}{B} = \frac{M\eta}{\gamma},$$

and by isolating η , one obtains following inequality:

$$\frac{\gamma\frac{\partial K}{\partial N}}{M\frac{\partial B}{\partial N}} > \eta. \quad (35)$$

Hence if $\bar{\eta} = \frac{\gamma\frac{\partial K}{\partial N}}{M\frac{\partial B}{\partial N}} > \eta$, the mass of firms in the differentiated goods sector, M , will increase with N .

I now show that there always exists an η such that $\bar{\eta} > \eta > \underline{\eta}$. Substituting the expressions for both terms yields:

$$\frac{\gamma K\bar{r}}{NB} < \frac{\gamma\frac{\partial K}{\partial N}}{M\frac{\partial B}{\partial N}}. \quad (36)$$

After rearranging terms it follows that:

$$\frac{M\bar{r}}{N} < \frac{B\frac{\partial K}{\partial N}}{K\frac{\partial B}{\partial N}}.$$

By assumption, for η to be in the set $[\underline{\eta}, \bar{\eta}]$ the following two conditions must simultaneously hold:

$$M\bar{r} < N \quad \& \quad K\frac{\partial B}{\partial N} < B\frac{\partial K}{\partial N}.$$

where by assumption $\frac{\partial B}{\partial N}$ is positive. Hence from these two conditions it follows that:

$$\frac{M\bar{r}}{N} < 1 < \frac{B\frac{\partial K}{\partial N}}{K\frac{\partial B}{\partial N}}.$$

Therefore there always exists an $\eta \in [\underline{\eta}, \bar{\eta}]$.

Proof of Proposition 6

Consider an entrepreneur with demand draw $\alpha_{L,L+1}$ that is indifferent between two organizational forms L and $L+1$. Then it follows that her profits are the same and:

$$\pi_L(\alpha_{L,L+1}, N) = \pi_{L+1}(\alpha_{L,L+1}, N).$$

In this section, I first show how q_L and q_{L+1} change with respect to N . Then I examine how $\alpha_{L,L+1}$ changes with N . Finally, I examine how $\alpha_{L,L+1}$ changes relative to $\alpha_{L+1,L+2}$ with respect to N .

Substituting the expression for profits, implies that:

$$\frac{\gamma}{N}q_L^2 + MC_L(q_L)q_L - C_L(q_L) = \frac{\gamma}{N}q_{L+1}^2 + MC_{L+1}(q_{L+1})q_{L+1} - C_{L+1}(q_{L+1}),$$

where q_L and q_{L+1} are the quantities produced by the entrepreneur with demand draw $\alpha_{L,L+1}$ when she is producing with L or $L + 1$ layers. Taking the derivative of this expression with respect to N and eliminating any common terms yields:

$$-\frac{\gamma}{N^2}q_L^2 + \frac{2\gamma}{N}q_L \frac{\partial q_L}{\partial N} + q_L \frac{\partial MC_L}{\partial q_L} \frac{\partial q_L}{\partial N} = -\frac{\gamma}{N^2}q_{L+1}^2 + \frac{2\gamma}{N}q_{L+1} \frac{\partial q_{L+1}}{\partial N} + q_{L+1} \frac{\partial MC_{L+1}}{\partial q_{L+1}} \frac{\partial q_{L+1}}{\partial N}. \quad (37)$$

Since q_L is the optimal quantity supplied by the entrepreneur, it satisfies the first order condition of the firm's maximization problem:

$$q_L = \frac{N}{2\gamma} \left[\alpha_{L,L+1} - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC_L(q_L) \right].$$

Similarly because q_{L+1} is the optimal quantity supplied by the entrepreneur, it satisfies the equation:

$$q_{L+1} = \frac{N}{2\gamma} \left[\alpha_{L,L+1} - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC_{L+1}(q_{L+1}) \right].$$

From these two expressions it follows that:

$$MC_L(q_L) + \frac{2\gamma}{N}q_L = MC_{L+1}(q_{L+1}) + \frac{2\gamma}{N}q_{L+1}. \quad (38)$$

Taking the derivative of equation (38) with respect to N yields:

$$\frac{\partial MC_L(q_L)}{\partial q_L} \frac{\partial q_L}{\partial N} - \frac{2\gamma}{N^2}q_L + \frac{2\gamma}{N} \frac{\partial q_L}{\partial N} = \frac{\partial MC_{L+1}(q_{L+1})}{\partial q_{L+1}} \frac{\partial q_{L+1}}{\partial N} - \frac{2\gamma}{N^2}q_{L+1} + \frac{2\gamma}{N} \frac{\partial q_{L+1}}{\partial N}. \quad (39)$$

Multiplying equation (39) by q_{L+1} , substituting this expression into equation (37), and rearranging yields:

$$\frac{\partial q_L}{\partial N} \left[\frac{2\gamma}{N} (q_L - q_{L+1}) + (q_L - q_{L+1}) \frac{\partial MC_L}{\partial q_L} \right] = \frac{\gamma}{N^2} (q_{L+1} - q_L)^2. \quad (40)$$

Since q_L is less than q_{L+1} the term on the right hand-side is positive, while the expression in brackets on the left-hand side is negative. Hence from equation (40) we have:

$$\frac{\partial q_L}{\partial N} < 0. \quad (41)$$

Performing the same steps as above, but multiplying equation (39) by q_L yields:

$$\frac{\partial q_{L+1}}{\partial N} \left[\frac{2\gamma}{N} (q_{L+1} - q_L) + (q_{L+1} - q_L) \frac{\partial MC_{L+1}}{\partial q_{L+1}} \right] = \frac{\gamma}{N^2} (q_{L+1} - q_L)^2. \quad (42)$$

which implies:

$$\frac{\partial q_{L+1}}{\partial N} > 0. \quad (43)$$

Hence for the two quantities q_L and q_{L+1} such that an entrepreneur is indifferent between two organizational forms, q_L is decreasing with N while q_{L+1} is increasing with respect to N . Therefore when controlling for market size, larger firms will have more layers.

I now examine how $\alpha_{L,L+1}$ changes with respect to N . The first order condition of the firm's maximization problem can be rewritten as:

$$q_L = \frac{N}{2\gamma} \left[\alpha_{L,L+1} - \alpha_D + \alpha_D - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC_L(q_L) \right].$$

Taking the derivative of this expression with respect to N and isolating common terms yields:

$$\frac{\partial q_L}{\partial N} \left[1 + \frac{N}{2\gamma} \frac{\partial MC_L(q_L)}{\partial q_L} \right] = \frac{q_L}{N} + \frac{N}{2\gamma} \frac{\partial (\alpha_{L,L+1} - \alpha_D)}{\partial N} + \frac{N}{2\gamma} \frac{\partial \left(\alpha_D - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) \right)}{\partial N}. \quad (44)$$

The expression in brackets is positive and $\frac{\partial q_L}{\partial N}$ is negative, so the left-hand side in equation (44) is negative. Also since:

$$\frac{q_L}{N} + \frac{N}{2\gamma} \frac{\partial \left(\alpha_D - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) \right)}{\partial N} = \frac{q_L}{N} - \frac{q_D}{2N} > 0,$$

the distance between $\alpha_{L,L+1}$ and α_D decreases with N . Namely,

$$\frac{\partial (\alpha_{L,L+1} - \alpha_D)}{\partial N} < 0.$$

I now proceed to analyze how $\alpha_{L,L+1}$ changes with N . Returning to the first order condition of the firm's maximization problem, taking the derivative of this expression with respect to N , and isolating common terms yields:

$$\frac{\partial q_L}{\partial N} \left[1 + \frac{N}{2\gamma} \frac{\partial MC_L(q_L)}{\partial q_L} \right] = \frac{q_L}{N} + \frac{N}{2\gamma} \frac{\partial \alpha_{L,L+1}}{\partial N} - \frac{N}{2\gamma} \frac{\partial \left(\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) \right)}{\partial N}.$$

Since the term on the left-hand side is negative, we have:

$$\frac{\partial \left(\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) \right)}{\partial N} - \frac{2\gamma}{N^2} q_L > \frac{\partial \alpha_{L,L+1}}{\partial N}.$$

which provides an upper bound to $\frac{\partial \alpha_{L,L+1}}{\partial N}$. Performing the same exercise with respect to q_{L+1} implies:

$$\frac{\partial \alpha_{L,L+1}}{\partial N} > \frac{\partial \left(\frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) \right)}{\partial N} - \frac{2\gamma}{N^2} q_{L+1}.$$

which provides a lower bound to $\frac{\partial \alpha_{L,L+1}}{\partial N}$.

Consider two entrepreneurs who are indifferent producing with two types of organizations. Denote the demand draw of the entrepreneur who is indifferent between $L, L+1$ layers as $\alpha_{L,L+1}$ and the demand draw of the entrepreneur who is indifferent between $L+1, L+2$ layers as $\alpha_{L+1,L+2}$. I now examine how $\alpha_{L,L+1}$ and $\alpha_{L+1,L+2}$ change relative to each other with respect to N . From the first order condition of the firm's maximization problem it follows that:

$$q_L(\alpha_{L,L+1}) = \frac{N}{2\gamma} \left[\alpha_{L,L+1} - \frac{\eta M}{\gamma + \eta M} (\bar{\alpha} - \bar{p}) - MC_L(q_L(\alpha_{L,L+1})) \right],$$

where $q_L(\alpha_{L,L+1})$ is the quantity supplied by the firm with demand draw $\alpha_{L,L+1}$ using an organization with L layers. This implies that:

$$\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p}) = \alpha_{L,L+1} - MC_L(q_L(\alpha_{L,L+1})) - \frac{2\gamma}{N}q_L(\alpha_{L,L+1}).$$

Returning to the first order condition of the firm's maximization problem with demand draw $\alpha_{L+1,L+2}$ and substituting in the expression for $\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p})$ from above, it follows that:

$$q_{L+1}(\alpha_{L+1,L+2}) = q_{L+1}(\alpha_{L,L+1}) + \frac{N}{2\gamma} [\alpha_{L+1,L+2} - \alpha_{L,L+1} + MC_{L+1}(q_{L+1}(\alpha_{L,L+1})) - MC_{L+1}(q_{L+1}(\alpha_{L+1,L+2}))],$$

where $q_{L+1}(\alpha_{L+1,L+2})$ is the quantity supplied by the firm with demand draw $\alpha_{L+1,L+2}$ using an organization with $L + 2$ layers. Taking the derivative of this expression with respect to N and isolating common terms yields:

$$\begin{aligned} \frac{\partial q_{L+1}(\alpha_{L+1,L+2})}{\partial N} \left[\frac{2\gamma}{N} + \frac{\partial MC_{L+1}(q_{L+1}(\alpha_{L+1,L+2}))}{\partial q_{L+1}(\alpha_{L+1,L+2})} \right] - \frac{\partial q_{L+1}(\alpha_{L,L+1})}{\partial N} \left[\frac{2\gamma}{N} + \frac{\partial MC_{L+1}(q_{L+1}(\alpha_{L,L+1}))}{\partial q_{L+1}(\alpha_{L,L+1})} \right] = \\ \frac{2\gamma}{N^2}(q_{L+1}(\alpha_{L+1,L+2}) - q_{L+1}(\alpha_{L,L+1})) + \frac{\partial \alpha_{L+1,L+2}}{\partial N} - \frac{\partial \alpha_{L,L+1}}{\partial N}. \end{aligned}$$

Since $\frac{\partial q_{L+1}(\alpha_{L+1,L+2})}{\partial N} < 0$ and $\frac{\partial q_{L+1}(\alpha_{L,L+1})}{\partial N} > 0$, the term on the left-hand side is negative, and therefore:

$$-\frac{2\gamma}{N^2}(q_{L+1}(\alpha_{L+1,L+2}) - q_{L+1}(\alpha_{L,L+1})) > \frac{\partial \alpha_{L+1,L+2}}{\partial N} - \frac{\partial \alpha_{L,L+1}}{\partial N}.$$

This provides an upper bound to $\frac{\partial \alpha_{L+1,L+2}}{\partial N} - \frac{\partial \alpha_{L,L+1}}{\partial N}$. Since $q_{L+1}(\alpha_{L+1,L+2}) - q_{L+1}(\alpha_{L,L+1}) > 0$, it follows that:

$$\frac{\partial \alpha_{L+1,L+2}}{\partial N} - \frac{\partial \alpha_{L,L+1}}{\partial N} < 0. \quad (45)$$

Therefore the distance between the demand draws $\alpha_{L,L+1}$ and $\alpha_{L+1,L+2}$ decreases with N . Adding and subtracting $\frac{\partial \alpha_D}{\partial N}$ on the left-hand side, isolating terms and taking into account that $\frac{\partial \alpha_{L+1,L+2}}{\partial N} - \frac{\partial \alpha_D}{\partial N}$ and $\frac{\partial \alpha_{L,L+1}}{\partial N} - \frac{\partial \alpha_D}{\partial N}$ are both negative, implies that the absolute change of $\alpha_{L+1,L+2}$ relative to α_D is greater than the change of $\alpha_{L,L+1}$ relative to α_D .

I can provide a lower bound for this expression as well. Using the same argument as above, but replacing $q_{L+1}(\alpha_{L+1,L+2})$ with the quantities produced by the entrepreneur with demand draw $\alpha_{L+1,L+2}$ using an organization of $L + 2$ layers, $q_{L+2}(\alpha_{L+1,L+2})$, and replacing $q_{L+1}(\alpha_{L,L+1})$ with the quantities produced by the entrepreneur with demand draw $\alpha_{L,L+1}$ using an organization of L layers, $q_L(\alpha_{L,L+1})$ yields:

$$\frac{\partial \alpha_{L+1,L+2}}{\partial N} - \frac{\partial \alpha_{L,L+1}}{\partial N} > -\frac{2\gamma}{N^2}(q_{L+2}(\alpha_{L+1,L+2}) - q_L(\alpha_{L,L+1})).$$

Proof of Proposition 7

The probability mass of firms producing with at most L layers is:

$$\Lambda_N(L) = \frac{[G(\alpha_{L,L+1}) - G(\alpha_D)]}{1 - G(\alpha_D)}, \quad (46)$$

with $\Lambda_N(L_{max}) = 1$ for some L_{max} which is the maximum number of layers in a firm. Taking the derivative of this expression with respect to N yields:

$$\frac{\partial \Lambda_N(L)}{\partial N} = \frac{[1 - G(\alpha_D)] \left[g(\alpha_{L,L+1}) \frac{\partial \alpha_{L,L+1}}{\partial N} - g(\alpha_D) \frac{\partial \alpha_D}{\partial N} \right] + [G(\alpha_{L,L+1}) - G(\alpha_D)] g(\alpha_D) \frac{\partial \alpha_D}{\partial N}}{[1 - G(\alpha_D)]^2}.$$

Eliminating common terms, adding and subtracting $[1 - G(\alpha_D)] g(\alpha_{L,L+1}) \frac{\partial \alpha_D}{\partial N}$, yields:

$$\begin{aligned} \frac{\partial \Lambda_N(L)}{\partial N} &= \frac{[1 - G(\alpha_D)] g(\alpha_{L,L+1}) \frac{\partial \alpha_{L,L+1}}{\partial N} - [1 - G(\alpha_{L,L+1})] g(\alpha_D) \frac{\partial \alpha_D}{\partial N}}{[1 - G(\alpha_D)]^2} \\ &= \frac{[1 - G(\alpha_D)] g(\alpha_{L,L+1}) \left[\frac{\partial \alpha_{L,L+1}}{\partial N} - \frac{\partial \alpha_D}{\partial N} \right] + [[1 - G(\alpha_D)] g(\alpha_{L,L+1}) - [1 - G(\alpha_{L,L+1})] g(\alpha_D)] \frac{\partial \alpha_D}{\partial N}}{[1 - G(\alpha_D)]^2}. \end{aligned} \quad (47)$$

The denominator is always positive, so the sign of $\frac{\partial \Lambda_N(L)}{\partial N}$ depends on the numerator. Previous sections have shown that $\frac{\partial \alpha_{L,L+1}}{\partial N} - \frac{\partial \alpha_D}{\partial N}$ is negative, and that $\frac{\partial \alpha_D}{\partial N}$ is positive. Hence the numerator in equation (47) will be negative if the following condition holds:

$$[1 - G(\alpha_D)] g(\alpha_{L,L+1}) \leq [1 - G(\alpha_{L,L+1})] g(\alpha_D),$$

which can be rewritten as

$$\frac{g(\alpha_{L,L+1})}{[1 - G(\alpha_{L,L+1})]} \leq \frac{g(\alpha_D)}{[1 - G(\alpha_D)]}. \quad (48)$$

Equation (48) is the hazard rate of the distribution of demand draws, $G(\alpha)$. Thus as long as $G(\alpha)$ has a non-increasing hazard rate, it follows that the probability mass of firms producing with at most L layers, $\Lambda_N(L)$ will be decreasing with respect to N . Therefore, if $N' > N$, it follows that the distribution of layers in economy N' , $\Lambda_{N'}$, will first order stochastically dominate the distribution of layers in economy N , Λ_N .

A.3 Isomorphisms Propositions

In this section I prove a simple theorem that relates to models that use Ottaviano-Tabuchi-Thisse preferences (Ottaviano et al. (2002)). I demonstrate that there is a correspondence between many market equilibria in this class of models, which makes it difficult to empirically test the model's prediction. In particular, the theorem implies that for any proportional change in the size of the market, there exists an equilibrium with the exact same outcomes, but that is derived from different parameters of the utility functions. This results stresses the importance of including demographic controls in regressions so as to proxy for consumers' preferences. In this section, I restrict my analysis to the utility function described in the paper with heterogeneous α , however analogous results can be derived from a model with constant α , as in Melitz and Ottaviano (2008).

Theorem 8 Consider the utility function (1). Let α , η and γ be parameters that govern agents' utility, and let N denote the size of the market. Then for any constant $k > 0$ the following parameters yield exactly same equilibrium outcomes: α , ηk and γk and Nk .

Proof. The proof relies on showing that the demand curve is the same in both frameworks. Consider the model with the parameters α, η, γ and N . Then a firm's demand curve is equal to:

$$p = \alpha + \frac{\eta M}{\eta M + \gamma} (\bar{\alpha} - \bar{p}) - \frac{\gamma}{N} q \quad (49)$$

where M is the mass of varieties in the economy. Replacing the parameters η, γ and N by the parameters $\eta k, \gamma k$ and Nk , yields the exact same demand curve. ■

Theorem 8 implies that any equilibrium derived from a model with Ottaviano-Tabuchi-Thisse utility and parameters $\{\alpha, \eta, \gamma, N\}$ is isomorphic to the set of equilibria from models with parameters $\{\alpha, \eta k, \gamma k, Nk\}$. This theorem has implications when researchers conduct comparative statics and take their model to the data. In particular consider the implications of an increase N in a closed economy. Theorem 8 implies the following:

Corollary 9 Consider the utility function (1). Let α, η and γ be parameters that govern agents' utility, and let N denote the size of the market. Consider the case of an increase in the size of the market of the sort $N^* = Nk$ with $k > 0$. Then the new equilibrium is identical to the following equilibrium in a market of size N and with parameters $\alpha, \frac{\eta}{k}$ and $\frac{\gamma}{k}$.

Corollary 9 implies that for any equilibrium generated from an increase in a market's size is isomorphic to an equilibrium where the market's size does not change, but where consumers have different preferences. Without controlling for agents' preferences, any estimates on how market size, N , affects economic outcomes across different regions, will be misidentified.

B Data Appendix

This section contains information on the construction of the data and reports additional descriptive statistics, presented in the following order:

1. Subsection B.1 explains the construction of the data composed of firms operating in the Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons industries, in the year 2004.
2. Subsection B.2 explains the construction of the data composed of firms operating in the manufacturing sector in the year 2004.
3. Subsection B.3 explains the construction of the panel data composed of firms operating in the Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons industries, in the years 2000-2007.
4. Subsection B.4 provides a description of the occupations in the Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons industries.
5. Subsection B.5 explains the different measures of firm TFP used in the paper.
6. Subsection B.6 list of main variables in the paper.
7. Subsection B.7 presents some general descriptive statistics of the data.
8. Subsection B.8 presents additional descriptive statistics on firm organization.
9. Subsection B.9 presents descriptive statistics across employment areas.
10. Subsection B.10 presents descriptive statistics across urban areas.

B.1 Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons: Year 2004.

The following French data sources that are used in this paper:

- Exhaustive Cross-Sections of the Déclarations Annuelles des Données Sociales: DADS
- Fichier Complet Unifié de Suse: FICUS
- Recensement de la Population: RP
- Restricted panel dimension of the DADS

Exhaustive Cross-Sections of the DADS

The DADS reports annual information on all almost every job (*poste de travail*) in mainland France. This study uses the years 2000-2007. The main empirical analysis of the paper is based on the year 2004. Apart from the years 2000-2001, where the identity of workers is unknown, for every job there is information on the identity of the worker, the firm, and the establishment within the firm. For each job, the DADS reports the number of hours, number of days, and salary, as well as the occupation within the firm. The DADS also contains demographic information on workers, such as their age, gender and municipality of residence and employment, and it also reports the main industry of the firm.

To clean the data, I first remove observations with missing worker or firm identifiers, and remove observations that have missing information for one of the following variables: number of days, number of hours, net and gross salary, occupation and municipality. I also remove public firms, firms with agricultural workers, and firms that contain errors in their occupational codes. Furthermore, in the DADS firms are classified into industries using the French NAF Rev 1 codes. At the 3-digit level, these industrial codes are equivalent to the NACE Rev 1.1 codes, however, at the 4-digit level the French industrial codes are more detailed: NACE Rev 1.1 has 514 classes while NAF Rev 1 has 712 classes. Finally, using the 1999 geographical definitions from INSEE, I match each municipality to an employment area and to an urban area.

The analysis is conducted at the firm-level. Using the French industrial codes, I further restrict the data to the different samples of firms in the study. The main sample is composed of firms that operate in one of the following industries: Clothing and Shoe Retail (NAF Rev 1 codes 524C and 524E), Traditional Restaurants (NAF Rev 1 code 553A) and Hair and Beauty Salons (NAF Rev 1 codes 930D and 930E). The manufacturing sample contains firms that operate in industries with a 2-digit level NAF Rev 1 code between 15 to 37. Note that from these samples, I only retain firms that operate in one single area and with at least 8 employees.

Furthermore, the method that I use to construct the number of layers in firms is based on [Caliendo et al. \(2015b\)](#). I use the 1st-digit of the cs-occupational codes, which, after cleaning, range from 2 to 6, to create the different hierarchical layers in firms. I also follow [Caliendo et al. \(2015b\)](#) and group ordinary administrative and production workers (occupations 5 and 6) together into the lowest layer of firms. With this method there are at most four different layers in a firm. Not every occupation, however, is observed in all firms. To measure the total number of layers in a firm, I simply count the number of layers present in a firm.

Fichier Complet Unifié de Suse: FICUS

The FICUS data is nearly exhaustive, and contains annual balance sheet information, such as sales, value-added and the value of capital, and some information on the structure of firms, such as whether firms belong to a French or international business group, and the legal status of firms. The years that are used in this study are from 2000-2007, and the main empirical analysis of the

paper is based on the year 2004. From this data I remove duplicate observations and drop firms with missing or negative values for one the following variables: value added, sales, total number of workers and salaries paid. I also remove observations in which value-added is larger than sales.

Recensement de la Population: RP

The RP data is a septennial census of the French population. In this study, I construct variables from the year 1999, and I use the year 1968 to construct instruments for the corresponding variables. For the years 1968 and 1999, the RP is exhaustive and contains demographic information on all individuals and households in France. For each individual, there is information on the household he belongs to, his demographic characteristics such as age, gender, education, employment status, and nationality, the characteristics of the household, and the location of the household at the municipal level.

Restricted panel dimension of the DADS

For my analysis, I also use the restricted panel dimension of the DADS to create controls for the local cost of a unit of labor and local incomes. From 1976-1993 the panel dimension of the DADS contains information on all workers born in October in even numbered years, and from 1993-2007 contains information on all workers born in October. For each individual there is information on their municipality of employment, and from 1993 onwards, their municipality of residence. The variables in both the exhaustive cross-section and the restricted panel of the DADS are generally the same.⁵⁰

From this restricted panel of the DADS, I first remove observations with missing or imputed worker, firm, or establishment identifiers, and remove observations that have missing information for variables of interest. I further retain male full-time workers between the ages of 18 and 65, aggregate the information to the worker-firm-occupation-year level, and for every worker and year, retain the firm-occupation with the highest salary.

From this data, I construct a measure of the local cost of a unit of labor. Using the years 1998-2000 and male workers that are not employed in one of the industries in the analysis, my measure of the local cost of a unit of labor, are the area fixed effects, g_a , estimated from the following equation:

$$\ln wage_{it} = \alpha + x_{it}\beta + occ_i + ind_j + g_a + \epsilon_i,$$

where $wage_{it}$ is the hourly wage of worker i in year t , x_{it} contains a quintic polynomial of a worker's age and time fixed effects, ind_j are industry fixed effects (3-digit level), and occ_i are cs-occupation fixed effects (1-digit level).⁵¹

From this data, I also construct a measure of local incomes. To construct a measure of local incomes, in the year 1999, I first group workers by their area of residence. I then calculate the median annual salary in each area, from the set of workers that are not employed in one of the industries in the analysis. I also use the same measure constructed from the year 1993, as an instrument.

Note that the local cost of a unit of labor and local incomes are constructed separately for employment and urban areas. Also, note that these variables are constructed separately for the different group of industries of the study. That is, in the main analysis, the sample used to construct these variables excludes workers employed in Clothing and Shoe Retail, Traditional Restaurants,

⁵⁰One difference between the panel and the exhaustive cross-sections of the DADS, however, is that in the panel individuals can be tracked over time.

⁵¹In the main analysis, these measures exclude workers employed in Clothing and Shoe Retail, Traditional Restaurants, and Hair and Beauty Salons. In the analysis on manufacturing firms, they exclude workers from the manufacturing sector.

and Hair and Beauty Salons. In the analysis of manufacturing firms, the sample used to construct these variables excludes workers employed in the manufacturing sector.

B.2 Manufacturing Dataset: Year 2004

The manufacturing dataset covers the year 2004. It is constructed using the same data sources as for the Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons, and all variables are created in the same way.

B.3 Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons: Panel Data, Years 2000-2007

The unbalanced panel dataset covers the years 2000-2007. It is constructed using the same data sources above, and all variables except for firm size are created the same. From the data, I retain firms that: (i) operate in a single area throughout the sample, (ii) that do not change industries, and (iii) firms that only exist for consecutive years. I also re-estimate firm productivity using the same methods described below for the main dataset of the paper. There are two differences, however, between the panel data and the main data of the paper. In the years 2000 and 2001, the DADS does not report the identity of individual workers. As a result, in all years, labor and the number of workers are approximated using jobs. And second, the size of areas is measured using the total number of private sector jobs in an area.

B.4 Description of Occupations in Clothing & Shoe Retail, Traditional Restaurants, Hair & Beauty Salons

For a sense of the hierarchical structure of firms in the Clothing and Shoe Retail, Traditional Restaurants, and Hair and Beauty Salons industries, Table B1 presents a description of the occupations that are associated with each occupational category in firms. It is important to note, however, that although the classification of occupational codes is detailed, it does not permit one to observe the tasks workers perform nor the chain of command within the firm.

Table B1: Description of Occupations

	Job Description	Examples
Layer Four	This occupational category includes the actual shopkeepers and owners when they are employees of their own trade, and corporate officers of a business.	Owners, CEO, CFO.
Layer Three	This occupational category includes employees occupying an executive or senior managerial position within a business.	Store Director, Head Chef, Department Head, General Manager.
Layer Two	This occupational category includes employees occupying a supervisory position within a business.	Bar Manager, Second Chef, Warehouse Manager, Sales Manager.
Layer One	This occupational category contains employees performing manual or administrative work who are either skilled and unskilled. Workers with degrees in the same field as their profession are considered skilled occupations. The rest are unskilled occupations.	Servers, Cooks, Dishwashers Barmen, Kitchen Helpers, Bus Boys, Receptionists, Cashiers, Estheticians, Hairdressers, Merchandisers, Warehouse Workers, Clerk, Sales Personnel.

Notes: List of occupations in the different layers in firms.

B.5 Estimating Productivity

In my analysis, I use several approaches to measure firm productivity. First, I use value-added per worker, which is a measure of revenue-labor productivity and is commonly used in the empirical literature examining the productivity of service firms.⁵² For this measure, I combine balance sheet information from the FICUS with measures on the size and wage bill of firms from the DADS.

Second, I adopt a more structural approach and measure firm productivity as Total Factor Productivity (TFP). In this case, I combine balance sheet information from FICUS with measures on the size and wage bill of firms from the DADS. I then remove any firms with missing, zero, or negative reported values of intermediate materials and/or capital, that operate in more than one area, and with at most 7 employees. I also remove firms that have gaps in their years of operations, that is firms that exit and re-enter the sample. In the main analysis and the dataset with manufacturing firms, firm productivity is estimated using the years 2002-2006, while in the panel analysis, the productivity of firms is estimated using the years 2000-2007. Additionally, because not all firms belong to an urban area, firm productivity is estimated separately over both geographical decompositions of mainland France. In other words, in the main sample, firm productivity is estimated once across employments areas, and separately over the sample of firms operating in urban areas.

To deal with endogenous factor inputs, I estimate firm TFP using the methods put forth by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#). More precisely, for firm i in year t , I assume that value-added, va_{it} , is equal to:

$$\ln va_{it} = \alpha_k \ln k_{it} + \alpha_l \ln l_{it} + \psi_t + \epsilon_{it}, \quad (50)$$

where k_{it} is the capital of firm i at time t , l_{it} denotes the amount of labor used and ψ_t are year fixed effects. The residual, ϵ_{it} , is the sum of the productivity of firm i at time t , ω_{it} , and an independently and identically distributed error term, η_{it} . I assume that productivity follows a first-order Markov process of the form $\omega_{it} = E[\omega_{it}|\omega_{it-1}] + \zeta_{it}$, where ζ_{it} is an innovation term that is independently and identically distributed and uncorrelated with past values of labor and capital. Capital is assumed to be predetermined at time t , and labor is a free variable determined at time t and correlated with the innovation term, ζ_{it} .

The methods proposed by [Levinsohn and Petrin \(2003\)](#), and [Wooldridge \(2009\)](#) to measure TFP are well known. To have enough variation in the data, firm productivity is estimated at the 3-digit NAF Rev 1 level. To estimate the production function, I rely on the *levpet* ([Petrin et al. \(2004\)](#)) and *prodest* ([Rovigatti and Mollisi \(2016\)](#)) commands in Stata, and measure firm TFP as the residual. For example, TFP estimates from equation (50) are equal to:

$$TFP_{it} = \exp(\ln va_{it} - \hat{\alpha}_k \ln k_{it} - \hat{\alpha}_l \ln l_{it}) \quad (51)$$

where $\hat{\alpha}_k$ and $\hat{\alpha}_l$ are the estimated coefficients.

For simplicity, in the paper, these measures of firm TFP are labeled: LP TFP and WD TFP. It is important to note that in nearly all estimates of firm TFP l_{it} is measured using the number of workers. Only Tables E8 and E9, in the Online Appendix E, report firm TFP results with l_{it} measured using the number of hours.

And third, I estimate TFP using a method adapted from [Caliendo et al. \(2015a\)](#). The theory in this paper takes the view that organization is a choice made by firms and that firm organization determines the efficiency of labor. In other words, the number of units of output produced from

⁵²Indeed there has been a long-standing debate about the appropriate measure of firm productivity in service firms. See [Haskel and Sadun \(2009\)](#) for a discussion on retail firms. One issue with value-added per worker, however, is that it ignores the other inputs used in production. See [Syverson \(2011\)](#) for a discussion on the different measures of firm productivity and the references therein.

a unit of labor depends on the number of layers in firms. Firm organization, however, does not determine A , the number of units of output produced when a problem is solved, which most closely corresponds to measures of firm TFP. Under the assumption that entrepreneurs can devote less than their full unit of labor to a firm, for firm i at time t , [Caliendo et al. \(2015a\)](#) show that value-added is equal to:

$$\ln va_{it} = \alpha_k \ln k_{it} + \alpha_C C(O_{it}, w) + \alpha_O ORG_{it} + \psi_t + \epsilon_{it}, \quad (52)$$

where O_{it} is the number of problems solved by workers in firm i at time t , $C(O_{it}, w)$ is the firm's wage bill at time t , and remaining terms are defined above. Equation (52) is also estimated using the methods proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#) to deal with endogenous factor inputs. I assume that firm productivity follows a first-order Markov process of the form $\omega_{it} = E[\omega_{it}|\omega_{it-1}] + \zeta_{it}$, where ζ_{it} is an innovation term that is independently and identically distributed. To identify the parameters α_k , α_C and α_O using the structural methods I assume that firm organization and the wage bill are free variables, determined at time t , and correlated with the innovation term ζ_{it} . Additionally, I assume that the innovation term is uncorrelated with past values of the wage bill and the number of layers in firms. The benefit of using this last approach is that it explicitly takes into consideration the number of layers in firms and controls for the quality of labor in the estimation of firm TFP.⁵³

Equation (52) is again estimated using the methods proposed by [Levinsohn and Petrin \(2003\)](#), and [Wooldridge \(2009\)](#). Finally, once the parameters are estimated, I define firm TFP as:

$$TFP_{it} = \exp(\ln va_{it} - \hat{\alpha}_k \ln k_{it} - \hat{\alpha}_C C(O_{it}, w) - \hat{\alpha}_O ORG_{it}) \quad (53)$$

where $\hat{\alpha}_k$, $\hat{\alpha}_C$ and $\hat{\alpha}_O$ are the estimated coefficients. For simplicity, in the paper, these measures of firm TFP are labeled: CMORH LP TFP and CMORH WD TFP. Finally, I also define TFP as in [Caliendo et al. \(2015a\)](#). In this case, it is equal to:

$$TFP_{it} = \exp(\ln va_{it} - \hat{\alpha}_k \ln k_{it} - \hat{\alpha}_C C(O_{it}, w)) \quad (54)$$

where $\hat{\alpha}_k$ and $\hat{\alpha}_C$ are the estimated coefficients from equation (52). For simplicity, in the paper, these measures of firm TFP are labeled: CMORH LP TFP + Org and CMORH WD TFP + Org.

⁵³Although equation (52) is very similar to [Caliendo et al. \(2015a\)](#) there is a difference. In their paper [Caliendo et al. \(2015a\)](#) take the position that organization not only determines the efficiency of labor but also firm TFP, which maps into the parameter A of the model. More precisely, [Caliendo et al. \(2015a\)](#) assume that the productivity of firms, ω_{it} follows the following process: $\omega_{it} = \rho\omega_{it-1} + \beta ORG_{it} + \zeta_{it}$. The number of layers in a firm do not only have a contemporaneous effect on TFP, but also past organizational structures can also affect current TFP. See [Caliendo et al. \(2015a\)](#) for details.

B.6 List of Variables

Main Variable of the Size of Markets:

- **Density:** there are two measures of density used in this paper. In the static analysis over the year 2004, density is equal to the total population residing in an area in the year 1999 divided by the surface area, measured in hectares squared. In the panel analysis over the years 2000-2007, density is defined as the total number of private sector jobs in an area divided by the surface area, measured in hectares squared.

Measures of Firm Productivity:

- **Value-Added per Worker:** total value-added in a firm divided by the total number of workers, constructed from the FICUS and the DADS.
- **LP TFP:** firm TFP estimated using a Cobb-Douglas production function, with labor and capital, and the structural assumptions from [Levinsohn and Petrin \(2003\)](#). This measure of firm TFP is defined in equation (51). Apart from Tables E8 and E9 in the Online Appendix E, all measures of firm TFP are estimated with labor measured using the number of workers. In Tables E8 and E9, firm TFP is estimated with labor measured using the number of hours.
- **WD TFP:** firm TFP estimated using a Cobb-Douglas production function, with labor and capital, and the structural assumptions from [Wooldridge \(2009\)](#). This measure of firm TFP is defined in equation (51). Apart from Tables E8 and E9 in the Online Appendix E, all measures of firm TFP are estimated with labor measured using the number of workers. In Tables E8 and E9, firm TFP is estimated with labor measured using the number of hours.
- **CMORH LP TPF:** firm TFP estimated using a production function adapted from [Caliendo et al. \(2015a\)](#), and the structural assumptions from [Levinsohn and Petrin \(2003\)](#). This measure of firm TFP is defined in equation (53).
- **CMORH WD TPF:** firm TFP estimated using a production function adapted from [Caliendo et al. \(2015a\)](#), and the structural assumptions from [Wooldridge \(2009\)](#). This measure of firm TFP is defined in equation (53).
- **CMORH LP TPF + Org:** firm TFP and the estimated effect of organization, both obtained from a production function adapted from [Caliendo et al. \(2015a\)](#), and the structural assumptions from [Levinsohn and Petrin \(2003\)](#). This measure of firm TFP is defined in equation (54).
- **CMORH WD TPF + Org:** firm TFP and the estimated effect of organization, both obtained from a production function adapted from [Caliendo et al. \(2015a\)](#), and the structural assumptions from [Wooldridge \(2009\)](#). This measure of firm TFP is defined in equation (54).
- **Value-Added per Hour:** total value-added in a firm divided by the total number of hours, constructed from the FICUS and the DADS.

Firm Controls:

- **Organization:** the total number of layers in a firm, constructed from the DADS.
- **Total number of workers:** the total number of workers in the firm, constructed from the DADS.
- **Total number of hours:** the total number of hours in the firm, constructed from the DADS.

- **Capital:** value of capital in firms, obtained from the FICUS.
- **Number of additional occupations:** the number of additional occupations in firms, constructed from the DADS. This variable is measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$.
- **Single Establishment:** indicator variable for whether a firm has a single establishment, constructed from the DADS.
- **Business Group:** information on whether firms belong to a business group, obtained from the FICUS.
- **Legal status:** information on the legal status of firms, obtained from the FICUS.

Income and Wage Controls:

- **Cost of a unit of labor:** constructed using the Restricted panel dimension of the DADS.
- **Median annual salary of individuals:** for the years 1993 and 1999, constructed using the restricted panel dimension of the DADS.

Demographic Controls:

The following variables are constructed from the RP data from the year 1999 and 1968:

- the share of the local population between the ages of 25 and 59 in 1999, the share of the local population born outside France in 1999, and the share of unemployed active workers in 1999, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, the average person per household in 1968, the share of the population in 1968 residing in buildings built before 1949, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence.

Additional Controls:

- **Herfindahl-Hirschman Occupational Index (HHI):** occupational index measured at the 2-digit cs-occupational level, in the year 1999, constructed from the DADS.
- **Average Hourly Wage:** total wage bill in a firm divided by the number of hours in the firm. The total wage bill is calculated using net wages and constructed from the DADS.
- **Average Hourly Wage in layer l :** total wage bill in a layer l divided by the number of hours in layer l . The total wage bill in layer l is calculated using net wages, and constructed from the DADS.

Historical Measures of Density:

Historical measures of density are from [Combes et al. \(2008\)](#).

B.7 General Descriptive Statistics

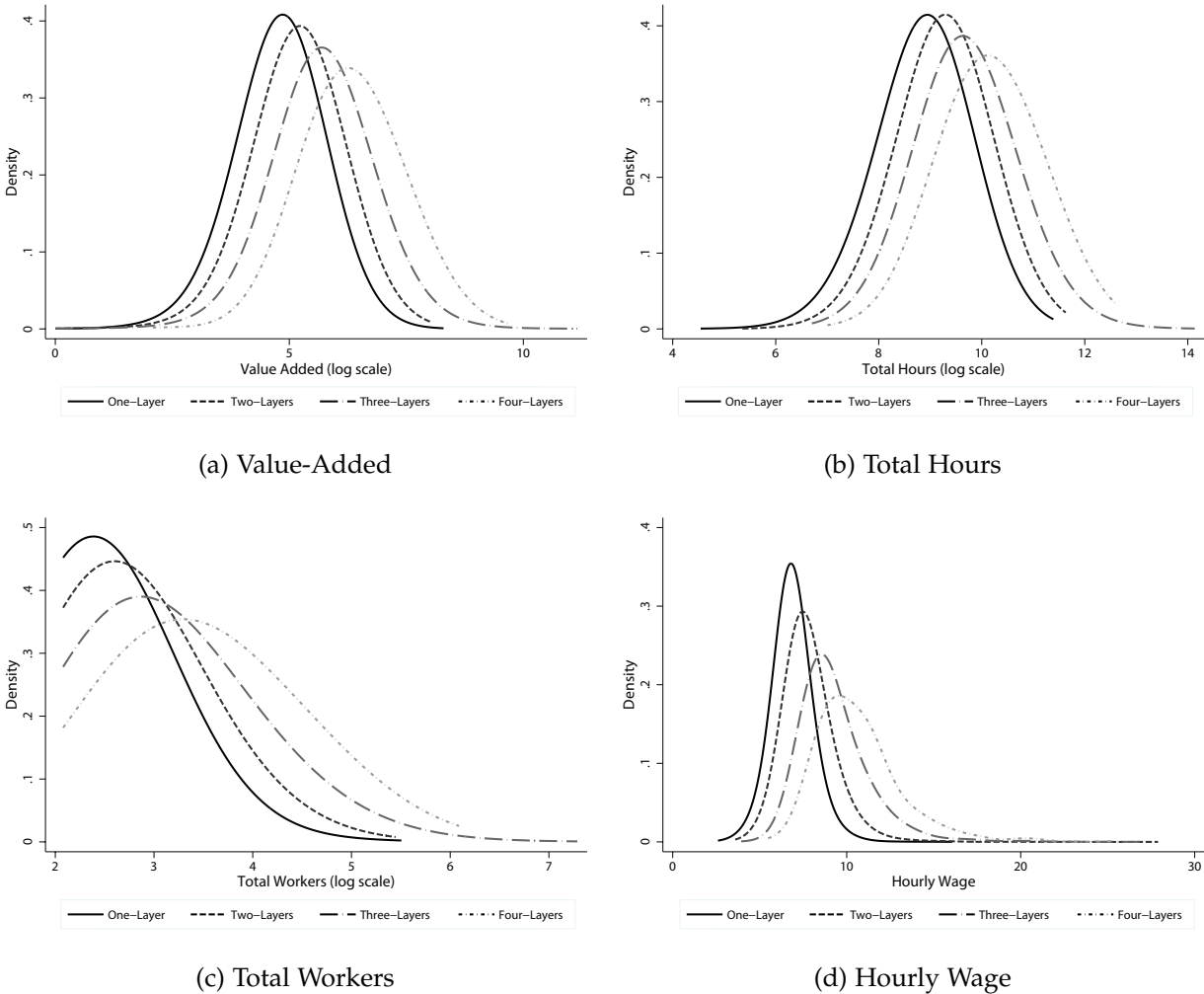
Table B2: General Summary Statistics

	Employment Areas	Urban Areas
Number of Areas	341	1,289
Average Density	2.63	3.02
Median Density	0.743	2.13
St. Dev. of Density	12.47	2.93
Total Number of Firms	27,508	24,197
Clothing and Shoe Retail	2,879	2,819
Traditional Restaurants	20,620	17,605
Hair and Beauty Salons	4,009	3,773
Average Number of Firms per Area	80.66	18.77
Median Number of Firms per Area	33	3
St. Dev. of Number of Firms per Area	221.60	186.46
Corr. B/T Density and Number of Firms	0.863	0.421
Average Number of Establishments Per Firm	1.09	1.09
Median Number of Establishments Per Firm	1	1
St. Dev. of Number of Establishments Per Firm	0.410	0.434
Corr. B/T Density and Local Market Controls		
Cost of a Unit of Labor	0.043	0.055
Median Annual Salary	0.288	0.062
Share of Population Between 25 and 59	0.336	0.161
Share of Population Born Outside France	0.362	0.205
Share of Population Unemployed	0.030	0.197

Notes: Summary statistics across local markets.

B.8 Additional Descriptive Statistics: Firm Organization

Figure B.4: Kernel Density Distributions across Organizations



Notes: Panel (a) illustrates the kernel density distribution of value-added for one-layer, two-layer, three-layer and four-layer firms. Panel (b) illustrates the kernel density distribution of the total number of hours for one-layer, two-layer, three-layer and four-layer firms. Panel (c) illustrates the kernel density distribution of the total number of workers for one-layer, two-layer, three-layer and four-layer firms. Panel (d) illustrates the kernel density distribution of the hourly wage for one-layer, two-layer, three-layer and four-layer firms. For each firm, the hourly wage is measured as: the total wage bill/total number of hours.

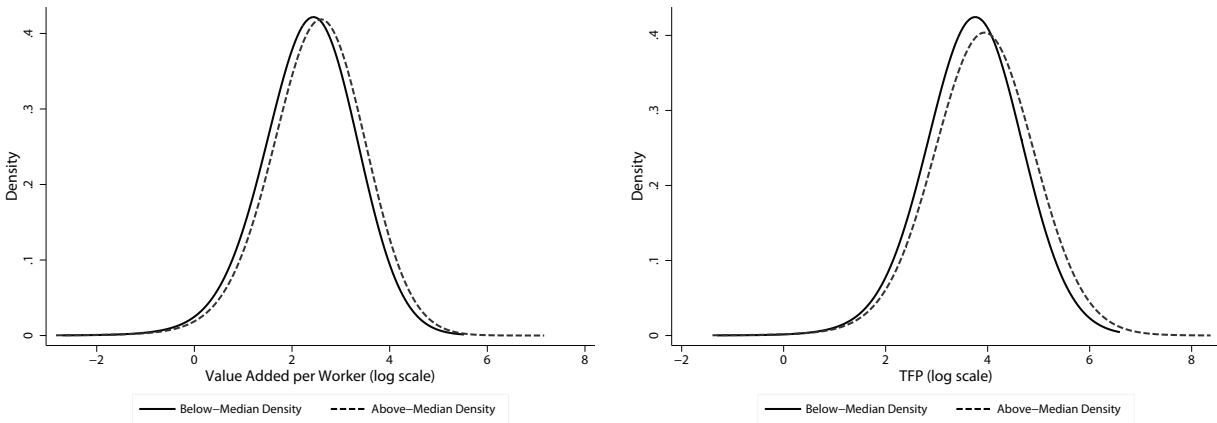
B.9 Descriptive Statistics: Firms across Employment Areas

Table B3: Summary Statistics across Employment Areas

	Number of Firms	Average Number of Layers	Average Value Added	Average Number of Workers	Average Number of Hours	Average Hourly Wage
All	27,508	1.85	260.66	18.29	13,531	7.77
Below Median	4,428	1.66	186.20	15.54	10,762	7.25
Above Median	23,080	1.88	274.95	18.82	14,062	7.87

Notes: Summary statistics of firms across employment areas. For each firm, the hourly wage is measured as: the total wage bill/total number of hours.

Figure B.5: Kernel Density Distributions across Employment Areas

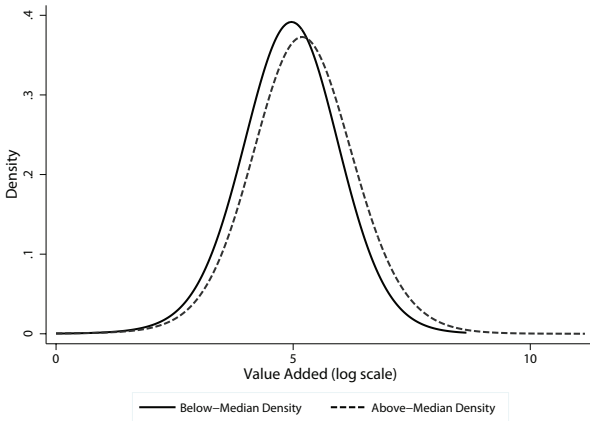


(a) Value-Added per Worker

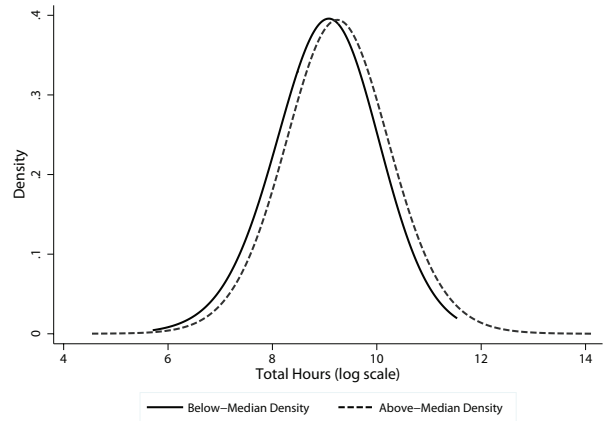
(b) TFP

Notes: Panel (a) illustrates the kernel density distribution of value-added per worker across firms operating employment areas with above and below median density. Panel (b) illustrates the kernel density distribution of estimates of firm TFP across firms operating employment areas with above and below median density. TFP is estimated using a Cobb-Douglas production function with capital and labor (number of workers), and the identifying assumptions of [Levinsohn and Petrin \(2003\)](#).

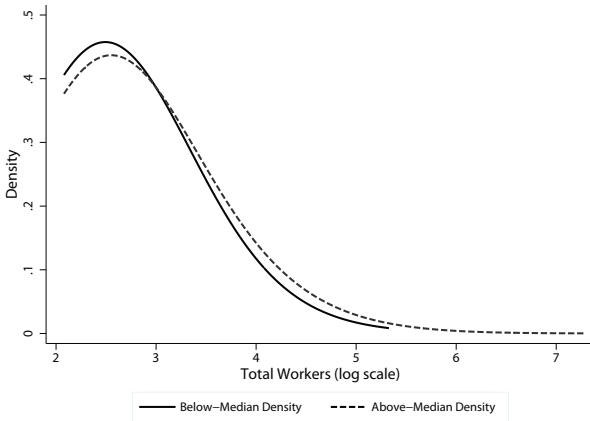
Figure B.6: Kernel Density Distributions across Employment Areas



(a) Value-Added



(b) Total Hours



(c) Total Workers



(d) Hourly Wage

Notes: Panel (a) illustrates the kernel density distribution of value-added across firms operating employment areas with above and below median density. Panel (b) illustrates the kernel density distribution of the total number of hours across firms operating employment areas with above and below median density. Panel (c) illustrates the kernel density distribution of the total number of workers across firms operating employment areas with above and below median density. Panel (d) illustrates the kernel density distribution of the hourly wage across firms operating employment areas with above and below median density. For each firm, the hourly wage is measured as: the total wage bill/total number of hours.

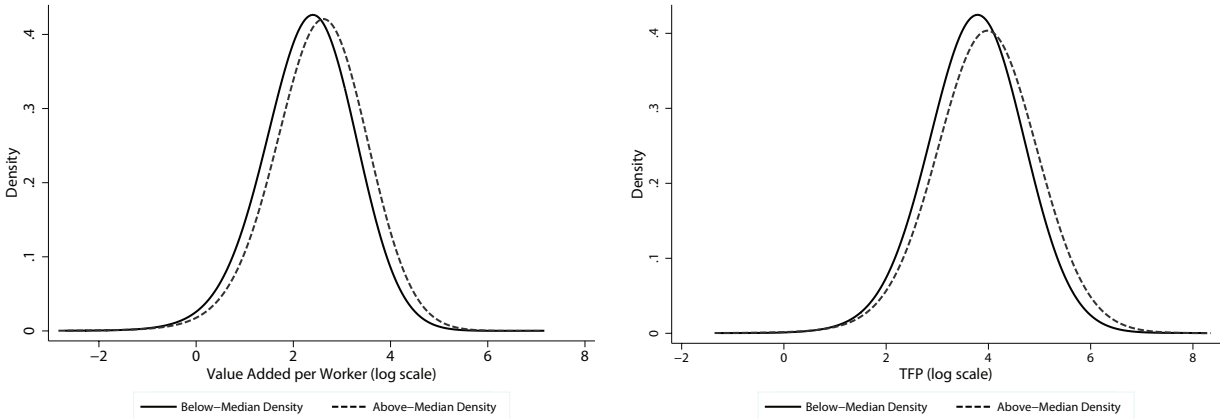
B.10 Descriptive Statistics: Firms across Urban Areas

Table B4: Summary Statistics across Urban Areas

Urban Areas	Number of Firms	Average Number of Layers	Average Value Added	Average Number of Workers	Average Number of Hours	Average Hourly Wage
All	24,197	1.88	274.62	18.76	14,198	7.84
Below Median	2,054	1.63	215.30	14.89	9,706	7.40
Above Median	22,143	1.90	280.12	19.12	14,615	7.88

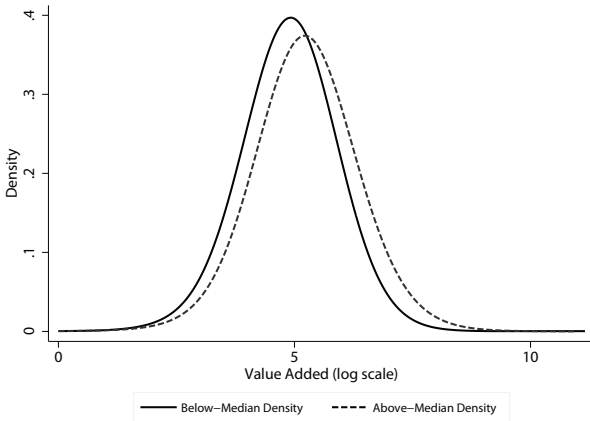
Notes: Summary statistics of firms across urban areas. For each firm, the hourly wage is measured as: the total wage bill/total number of hours.

Figure B.7: Kernel Density Distributions across Urban Areas

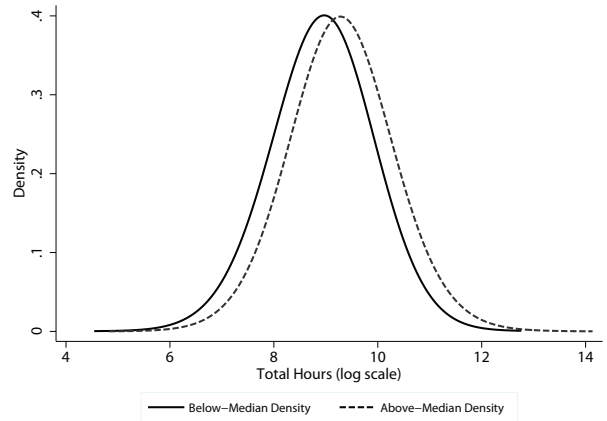


Notes: Panel (a) illustrates the kernel density distribution of value-added per worker across firms operating urban areas with above and below median density. Panel (b) illustrates the kernel density distribution of estimates of firm TFP across firms operating urban areas with above and below median density. TFP is estimated using a Cobb-Douglas production function with capital and labor (number of workers), and the identifying assumptions of [Levinsohn and Petrin \(2003\)](#).

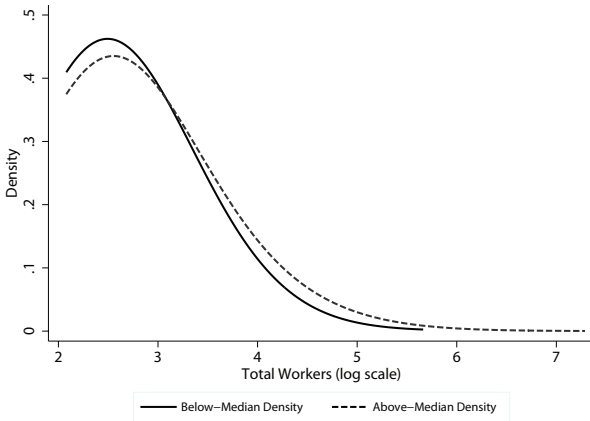
Figure B.8: Kernel Density Distributions across Urban Areas



(a) Value-Added



(b) Total Hours



(c) Total Workers



(d) Hourly Wage

Notes: Panel (a) illustrates the kernel density distribution of value-added across firms operating urban areas with above and below median density. Panel (b) illustrates the kernel density distribution of the total number of hours across firms operating urban areas with above and below median density. Panel (c) illustrates the kernel density distribution of the total number of workers across firms operating urban areas with above and below median density. Panel (d) illustrates the kernel density distribution of the hourly wage across firms operating urban areas with above and below median density. For each firm, the hourly wage is measured as: the total wage bill/total number of hours.

C Regression Results: Distribution of Firm Organization

This section contains additional results that examine firm organization across locations. The tables in this section are the following:

1. Table C1: Additional Regression Results across Employment Areas
2. Table C2: Regression Results across Urban Areas
3. Table C3: Additional Regression Results across Urban Areas
4. Table C4: Aggregate Regression Results across Employment Areas with Dependent Variable Share of Firms with a given Number of Layers
5. Table C5: Additional Aggregate Regression Results across Employment Areas with Dependent Variable Share of Firms with a given Number of Layers
6. Table C6: Aggregate Regression Results across Urban Areas with Dependent Variable Share of Firms with a given Number of Layers
7. Table C7: Additional Aggregate Regression Results across Urban Areas with Dependent Variable Share of Firms with a given Number of Layers
8. Table C8: Regression Results across Employment Areas with Dependent Variable Total Number of Layers in Firms
9. Table C9: Additional Regression Results across Employment Areas with Dependent Variable Total Number of Layers in Firms
10. Table C10: Regression Results across Urban Areas with Dependent Variable Total Number of Layers in Firms
11. Table C11: Additional Regression Results across Urban Areas with Dependent Variable Total Number of Layers in Firms

Table C1: Additional Regression Results across Employment Areas

	IV Sample (1)	IV Sample (2)	Probit (3)	Logit (4)	Herf.-Hir. Occ. Index (5)	Add. Occ. (6)	Number Workers (7)	Number Hours (8)	Add. Controls (9)
Prob. One-Layer Firm									
log density	-0.042 (0.002) ^a	-0.013 (0.004) ^a	-0.011 (0.006) ^c	-0.012 (0.006) ^b	-0.009 (0.006)	-0.005 (0.005)	-0.002 (0.005)	-0.000 (0.007)	-0.005 (0.005)
Prob. Two-Layer Firm									
log density	0.005 (0.004)	-0.000 (0.005)	0.006 (0.006)	0.007 (0.006)	-0.005 (0.006)	-0.007 (0.006)	-0.007 (0.006)	-0.008 (0.006)	-0.005 (0.005)
Prob. Three-Layer Firm									
log density	0.029 (0.001) ^a	0.009 (0.003) ^a	0.006 (0.002) ^b	0.006 (0.002) ^b	0.011 (0.003) ^a	0.009 (0.002) ^a	0.006 (0.002) ^b	0.005 (0.003) ^c	0.007 (0.003) ^b
Prob. Four-Layer Firm									
log density	0.007 (0.000) ^a	0.003 (0.000) ^a	0.002 (0.000) ^a	0.001 (0.000) ^b	0.004 (0.000) ^a	0.004 (0.000) ^a	0.003 (0.000) ^a	0.003 (0.000) ^a	0.003 (0.000) ^a
Method	OLS	OLS	MLE	MLE	OLS	OLS	OLS	OLS	OLS
R ² : One-Layer Firms	0.053	0.057	-	-	0.055	0.111	0.142	0.164	0.206
R ² : Two-Layer Firms	0.003	0.005	-	-	0.005	0.006	0.004	0.008	0.029
R ² : Three-Layer Firms	0.037	0.039	-	-	0.038	0.077	0.118	0.110	0.153
R ² : Four-Layer Firms	0.014	0.014	-	-	0.014	0.031	0.052	0.046	0.087
Wage Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	25,637	25,637	27,508	27,508	27,508	27,508	27,508	27,508	26,791

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (3) and (4) report the marginal effects of density evaluated at the mean, estimated using probit and logit models. Column (5) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (6) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Columns (7) and (8) control for the number of workers and hours in firms, respectively. Column (9) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table C2: Regression Results across Urban Areas

	(1)	(2)	(3)	(4)	(5)
Probability One-Layer Firm					
log density	-0.072 (0.006) ^a	-0.073 (0.006) ^a	-0.043 (0.005) ^a	-0.075 (0.009) ^a	-0.043 (0.008) ^a
Probability Two-Layer Firm					
log density	0.010 (0.003) ^a	0.010 (0.003) ^a	0.012 (0.005) ^b	0.008 (0.003) ^b	0.022 (0.008) ^a
Probability Three-Layer Firm					
log density	0.050 (0.007) ^a	0.050 (0.007) ^a	0.023 (0.003) ^a	0.054 (0.009) ^a	0.018 (0.006) ^a
Probability Four-Layer Firm					
log density	0.011 (0.001) ^a	0.011 (0.001) ^a	0.007 (0.000) ^a	0.012 (0.002) ^a	0.003 (0.001) ^c
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	13.28	6.67
R ² : One-Layer Firm	0.051	0.051	0.055	-	-
R ² : Two-Layer Firm	0.004	0.004	0.004	-	-
R ² : Three-Layer Firm	0.032	0.033	0.037	-	-
R ² : Four-Layer Firm	0.011	0.011	0.012	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	24,197	24,192	24,192	20,537	20,537

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an urban area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (4) only instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.157, 0.377, 0.845, and 0.493, when the dependent variable is the probability of a firm having respectively one, two, three or four layers. Column (5) instruments for density and local characteristics using the following variables: density in 1831, 1851, 1881, and 1901, the median annual salary of individuals residing in an urban area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, and the share of the population in 1968 residing in buildings built before 1949. The p-value of the over-id test is respectively 0.804, 0.596, 0.271, and 0.401, when the dependent variable is the probability of a firm having one, two, three or four layers.

Table C3: Additional Regression Results across Urban Areas

	IV Sample (1)	IV Sample (2)	Probit (3)	Logit (4)	Herf.-Hir. Occ. Index (5)	Add. Occ. (6)	Number Workers (7)	Number Hours (8)	Add. Controls (9)
Prob. One-Layer Firm									
log density	-0.078 (0.006) ^a	-0.042 (0.006) ^a	-0.043 (0.005) ^a	-0.043 (0.005) ^a	-0.041 (0.005) ^a	-0.037 (0.005) ^a	-0.025 (0.005) ^a	-0.018 (0.004) ^a	-0.019 (0.004) ^a
Prob. Two-Layer Firm									
log density	0.009 (0.003) ^b	-0.015 (0.007) ^b	0.012 (0.005) ^b	0.012 (0.005) ^b	0.008 (0.005)	0.011 (0.005) ^b	0.012 (0.005) ^b	0.008 (0.005)	0.010 (0.005) ^b
Prob. Three-Layer Firm									
log density	0.056 (0.007) ^a	0.021 (0.005) ^a	0.024 (0.004) ^a	0.024 (0.004) ^a	0.024 (0.004) ^a	0.019 (0.003) ^a	0.010 (0.001) ^a	0.006 (0.003) ^c	0.006 (0.003) ^c
Prob. Four-Layer Firm									
log density	0.013 (0.001) ^a	0.005 (0.001) ^a	0.007 (0.001) ^a	0.007 (0.001) ^a	0.007 (0.001) ^a	0.006 (0.000) ^a	0.003 (0.001) ^a	0.002 (0.001) ^b	0.002 (0.001) ^b
Method	OLS	OLS	MLE	MLE	OLS	OLS	OLS	OLS	OLS
R ² : One-Layer Firms	0.050	0.053	-	-	0.055	0.110	0.140	0.165	0.201
R ² : Two-Layer Firms	0.004	0.004	-	-	0.004	0.006	0.004	0.006	0.027
R ² : Three-Layer Firms	0.029	0.033	-	-	0.037	0.077	0.116	0.111	0.151
R ² : Four-Layer Firms	0.009	0.011	-	-	0.012	0.031	0.054	0.048	0.089
Wage Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	20,537	20,537	24,192	24,192	24,192	24,192	24,192	24,192	23,558

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an urban area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (3) and (4) report the marginal effects of density evaluated at the mean, estimated using probit and logit models. Column (5) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (6) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Columns (7) and (8) control for the number of workers and hours in firms, respectively. Column (9) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table C4: Aggregate Regression Results across Employment Areas with Dependent Variable Share of Firms with a given Number of Layers

	(1)	(2)	(3)	(4)	(5)
Share One-Layer Firm					
log density	-0.070 (0.006) ^a	-0.070 (0.006) ^a	-0.025 (0.007) ^a	-0.060 (0.006) ^a	-0.021 (0.012) ^c
Share Two-Layer Firm					
log density	0.039 (0.006) ^a	0.040 (0.006) ^a	0.015 (0.007) ^b	0.030 (0.006) ^a	0.000 (0.011)
Share Three-Layer Firm					
log density	0.025 (0.003) ^a	0.025 (0.003) ^a	0.007 (0.005)	0.025 (0.003) ^a	0.018 (0.007) ^b
Share Four-Layer Firm					
log density	0.004 (0.000) ^a	0.004 (0.000) ^a	0.003 (0.001) ^b	0.004 (0.000) ^a	0.002 (0.001)
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	335	6.98
R ² : One-Layer Firm	0.258	0.259	0.343	-	-
R ² : Two-Layer Firm	0.113	0.124	0.170	-	-
R ² : Three-Layer Firm	0.117	0.126	0.169	-	-
R ² : Four-Layer Firm	0.081	0.089	0.104	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	341	341	341	304	304

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Robust standard errors in parentheses. Entries display the second-stage estimates from a separate regression. The structure of regressions is as in section 5.2. Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (4) only instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.235, 0.437, 0.831, and 0.809, when the dependent variable is the share of respectively one-layer, two-layer, three-layer or four-layer firms. Column (5) instruments for density and local characteristics using the following variables: density in 1831, 1851, 1881, and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, the share of the population in 1968 residing in buildings built before 1949, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence. The p-value of the over-id test is respectively 0.398, 0.355, 0.387, and 0.716, when the dependent variable is the share of one-layer, two-layer, three-layer or four-layer firms.

Table C5: Additional Aggregate Regression Results across Employment Areas with Dependent Variable Share of Firms with a given Number of Layers

	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Number Workers (5)	Number Hours (6)	Add. Controls (7)
Share One-Layer Firm							
log density	-0.066 (0.007) ^a	-0.025 (0.007) ^a	-0.025 (0.007) ^a	-0.021 (0.007) ^a	-0.015 (0.006) ^b	-0.015 (0.007) ^b	-0.014 (0.006) ^b
Share Two-Layer Firm							
log density	0.036 (0.006) ^a	0.011 (0.007)	0.015 (0.007) ^b	0.014 (0.007) ^c	0.014 (0.007) ^c	0.013 (0.007) ^c	0.011 (0.007) ^c
Share Three-Layer Firm							
log density	0.025 (0.003) ^a	0.010 (0.005) ^b	0.007 (0.005)	0.004 (0.005)	-0.000 (0.005)	0.000 (0.005)	0.001 (0.005)
Share Four-Layer Firm							
log density	0.004 (0.000) ^a	0.003 (0.001) ^b	0.003 (0.001) ^a	0.002 (0.001) ^b	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
R ² : One-Layer Firms	0.236	0.313	0.347	0.314	0.301	0.214	0.204
R ² : Two-Layer Firms	0.096	0.158	0.171	0.163	0.168	0.146	0.128
R ² : Three-Layer Firms	0.122	0.162	0.175	0.140	0.108	0.060	0.066
R ² : Four-Layer Firms	0.080	0.096	0.107	0.078	0.028	0.017	0.029
Wage Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	304	304	341	341	341	341	341

Notes: *a, b, c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the second-stage estimates from a separate regression. The structure of regressions is as in section 5.2. Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Columns (5) and (6) control for the number of workers and hours in firms, respectively. Column (7) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table C6: Aggregate Regression Results across Urban Areas with Dependent Variable Share of Firms with a given Number of Layers

	(1)	(2)	(3)	(4)	(5)
Share One-Layer Firm					
log density	-0.059 (0.011) ^a	-0.060 (0.011) ^a	-0.043 (0.012) ^a	-0.050 (0.018) ^a	-0.050 (0.032)
Share Two-Layer Firm					
log density	0.019 (0.010) ^c	0.020 (0.011) ^c	0.016 (0.012)	0.021 (0.017)	0.021 (0.030)
Share Three-Layer Firm					
log density	0.037 (0.006) ^a	0.037 (0.006) ^a	0.023 (0.007) ^a	0.029 (0.010) ^a	0.040 (0.017) ^a
Share Four-Layer Firm					
log density	0.003 (0.001) ^c	0.003 (0.001)	0.003 (0.002)	-0.000 (0.005)	-0.011 (0.016)
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	256	5.19
R ² : One-Layer Firm	0.017	0.018	0.039	-	-
R ² : Two-Layer Firm	0.002	0.003	0.009	-	-
R ² : Three-Layer Firm	0.019	0.019	0.038	-	-
R ² : Four-Layer Firm	0.002	0.002	0.004	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	1,289	1,285	1,285	480	480

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Robust standard errors in parentheses. Entries display the second-stage estimates from a separate regression. The structure of regressions is as in section 5.2. Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an urban area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (4) only instruments for density using density in 1831 and 1901. The p-value of the over-id test is 0.033, 0.027, 0.680, and 0.767, when the dependent variable is the share of respectively one-layer, two-layer, three-layer or four-layer firms. Column (5) instruments for density and local characteristics using the following variables: density in 1831 and 1901, the median annual salary of individuals residing in an urban area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, the share of the population in 1968 residing in buildings built before 1949, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence. The p-value of the over-id test is respectively 0.530, 0.539, 0.809, and 0.338, when the dependent variable is the share of one-layer, two-layer, three-layer or four-layer firms.

Table C7: Additional Aggregate Regression Results across Urban Areas with Dependent Variable Share of Firms with a given Number of Layers

	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Number Workers (5)	Number Hours (6)	Add. Controls (7)
Share One-Layer Firm							
log density	-0.069 (0.015) ^a	-0.045 (0.017) ^b	-0.040 (0.012) ^a	-0.040 (0.012) ^a	-0.029 (0.011) ^b	-0.013 (0.011)	-0.017 (0.011)
Share Two-Layer Firm							
log density	0.037 (0.015) ^b	0.027 (0.017)	0.014 (0.012)	0.015 (0.012)	0.016 (0.012)	0.012 (0.012)	0.010 (0.012)
Share Three-Layer Firm							
log density	0.032 (0.008) ^a	0.021 (0.010) ^b	0.022 (0.007) ^a	0.022 (0.007) ^a	0.012 (0.007) ^c	0.002 (0.007)	0.007 (0.007)
Share Four-Layer Firm							
log density	-0.000 (0.005)	-0.003 (0.007)	0.003 (0.002) ^c	0.003 (0.002)	0.000 (0.002)	-0.002 (0.002)	-0.000 (0.002)
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
R ² : One-Layer Firms	0.042	0.089	0.040	0.038	0.027	0.014	0.013
R ² : Two-Layer Firms	0.013	0.037	0.010	0.009	0.010	0.008	0.006
R ² : Three-Layer Firms	0.025	0.045	0.038	0.036	0.024	0.014	0.015
R ² : Four-Layer Firms	0.000	0.018	0.004	0.003	0.001	0.005	0.002
Wage Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	480	480	1,285	1,285	1,285	1,285	1,277

Notes: *a, b, c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the second-stage estimates from a separate regression. The structure of regressions is as in section 5.2. Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an urban area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Columns (5) and (6) control for the number of workers and hours in firms, respectively. Column (7) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table C8: Regression Results across Employment Areas with Dependent Variable Total Number of Layers in Firms

Dependent Variable	(1)	(2)	(3)	(4)	(5)
Total Number of Layers					
log density	0.084 (0.003) ^a	0.084 (0.004) ^a	0.028 (0.007) ^a	0.084 (0.003) ^a	0.021 (0.012) ^c
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	1463	10.01
R-squared	0.075	0.075	0.079	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	27,508	27,508	27,508	25,637	25,637

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (4) only instruments for density using density in 1831, 1851, 1881 and 1901. The p-value of the over-id test is 0.236. Column (5) instruments for density and local characteristics using the following variables: density in 1831, 1881 and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, and the share of the population in 1968 residing in buildings built before 1949. The p-value of the over-id test is 0.426.

Table C9: Additional Regression Results across Employment Areas with Dependent Variable Total Number of Layers in Firms

Dependent Variable	IV	IV	Herf.-Hir.	Add.	Number	Number	Add.
Total Number of Layers	Sample	Sample	Occ. Index	Occ.	Workers	Hours	Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log density	0.086 (0.003) ^a	0.031 (0.007) ^a	0.029 (0.008) ^a	0.023 (0.006) ^a	0.015 (0.006) ^b	0.011 (0.009)	0.020 (0.006) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
R-squared	0.077	0.081	0.079	0.165	0.234	0.241	0.316
Wage Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	25,637	25,637	27,508	27,508	27,508	27,508	26,791

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Columns (5) and (6) control for the number of workers and hours in firms, respectively. Column (7) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table C10: Regression Results across Urban Areas with Dependent Variable Total Number of Layers in Firms

Dependent Variable					
Total Number of Layers	(1)	(2)	(3)	(4)	(5)
log density	0.146 (0.016) ^a	0.147 (0.016) ^a	0.080 (0.008) ^a	0.155 (0.023) ^a	0.068 (0.013) ^a
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	13.28	6.67
R-squared	0.070	0.070	0.077	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	24,197	24,192	24,192	20,537	20,537

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an urban area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (4) only instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.388. Column (5) instruments for density and local characteristics using the following variables: density in 1831, 1851, 1881, and 1901, the median annual salary of individuals residing in an urban area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, and the share of the population in 1968 residing in buildings built before 1949. The p-value of the over-id test is 0.553.

Table C11: Additional Regression Results across Urban Areas with Dependent Variable Total Number of Layers in Firms

Dependent Variable	IV	IV	Herf.-Hir.	Add.	Number	Number	Add.
Total Number of Layers	Sample	Sample	Occ. Index	Occ.	Workers	Hours	Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log density	0.161 (0.017) ^a	0.074 (0.010) ^a	0.080 (0.009) ^a	0.068 (0.007) ^a	0.042 (0.007) ^a	0.030 (0.007) ^a	0.030 (0.007) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
R-squared	0.065	0.072	0.077	0.164	0.233	0.245	0.314
Wage Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Income Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	20,537	20,537	24,192	24,192	24,192	24,192	23,558

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an urban area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of active workers unemployed. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Columns (5) and (6) control for the number of workers and hours in firms, respectively. Column (7) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

D Regression Results: Firm Productivity, First Step

This section contains additional results that examine firm productivity across locations. The tables in this section are the following:

1. Table **D1**: First-Step Regression Results across Employment Areas
2. Table **D2**: Additional Second-Step Regression Results across Employment Areas
3. Table **D3**: First-Step Regression Results across Urban Areas
4. Table **D4**: Additional Second-Step Regression Results across Urban Areas

Table D1: First-Step Regression Results across Employment Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Organization						
log density	0.122 (0.012) ^a	0.122 (0.012) ^a	0.124 (0.013) ^a	0.124 (0.013) ^a	0.286 (0.015) ^a	0.293 (0.015) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
R-Squared	0.134	0.134	0.133	0.133	0.212	0.194
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	27,508	27,508	25,637	25,637	23,459	23,459

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (14). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step.

Table D2: Additional First-Step Regression Results across Employment Areas

Dependent Variable	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Add. Controls (5)
VA Worker					
Organization					
log density	0.124 (0.013) ^a	0.124 (0.013) ^a	0.122 (0.012) ^a	0.117 (0.013) ^a	0.116 (0.015) ^a
Method	OLS	OLS	OLS	OLS	OLS
R-Squared	0.133	0.133	0.134	0.135	0.256
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	25,637	25,637	27,508	27,508	26,791

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (14). Firm productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Column (5) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table D3: First-Step Regression Results across Urban Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Organization						
log density	0.124 (0.016) ^a	0.124 (0.016) ^a	0.126 (0.017) ^a	0.126 (0.017) ^a	0.288 (0.023) ^a	0.295 (0.024) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
R-Squared	0.156	0.156	0.122	0.122	0.241	0.224
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	24,197	24,197	20,537	20,537	20,649	20,649

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (14). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step.

Table D4: Additional First-Step Regression Results across Urban Areas

Dependent Variable	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Add. Controls (5)
VA Worker					
Organization					
log density	0.126 (0.017) ^a	0.126 (0.017) ^a	0.124 (0.016) ^a	0.119 (0.016) ^a	0.122 (0.020) ^a
Method	OLS	OLS	OLS	OLS	OLS
R-Squared	0.122	0.122	0.156	0.157	0.270
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	20,537	20,537	24,197	24,197	23,563

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (14). Firm productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Column (5) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

E Regression Results: Firm Productivity, Second Step

This section contains additional results that examine firm productivity across locations. The tables in this section are the following:

1. Table E1: Additional Second-Step Regression Results across Employment Areas
2. Table E2: Second-Step Regression Results across Urban Areas
3. Table E3: Additional Second-Step Regression Results across Urban Areas
4. Table E4: Firm-Level Regression Results across Employment Areas
5. Table E5: Additional Firm-Level Second-Step Regression Results across Employment Areas
6. Table E6: Firm-Level Regression Results across Urban Areas
7. Table E7: Additional Firm-Level Second-Step Regression Results across Urban Areas
8. Table E8: Regression Results across Employment Areas with Other Measures Productivity
9. Table E9: Regression Results across Urban Areas with Other Measures Productivity
10. Table E10: Regression Results across Employment Areas with Other Measures of Firm Organization
11. Table E11: Regression Results across Urban Areas with Other Measures of Firm Organization

Table E1: Additional Second-Step Regression Results across Employment Areas

Dependent Variable	IV Sample	IV Sample	Herf.-Hir. Occ. Index	Add. Occ.	Add. Controls
VA Worker	(1)	(2)	(3)	(4)	(5)
Model 1: Without Org					
log density	0.053 (0.008) ^a	0.036 (0.012) ^a	0.046 (0.013) ^a	0.044 (0.013) ^a	0.044 (0.012) ^a
Model 2: With Org					
log density	0.040 (0.008) ^a	0.031 (0.012) ^b	0.041 (0.013) ^a	0.041 (0.013) ^a	0.042 (0.011) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	32.5	13.8	10.8	6.8	4.7
Sobel-ML Test (p-value)	0.000	0.000	0.000	0.004	0.033
R ² : Model 1	0.108	0.228	0.210	0.204	0.202
R ² : Model 2	0.067	0.190	0.162	0.162	0.171
2 nd -Step: Local Controls	No	Yes	Yes	Yes	Yes
1 st -Step: Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	304	304	341	341	341

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the estimate from a separate regression from equation (15). Sobel-ML Test is the multilevel mediation test from Krull and MacKinnon (1999) and Krull and MacKinnon (2001). Firm productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Column (5) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table E2: Second-Step Regression Results across Urban Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.067 (0.014) ^a	0.080 (0.015) ^a	0.085 (0.020) ^a	0.115 (0.032) ^a	0.114 (0.017) ^a	0.116 (0.017) ^a
Model 2: With Org						
log density	0.055 (0.014) ^a	0.071 (0.014) ^a	0.075 (0.020) ^a	0.107 (0.031) ^a	0.089 (0.016) ^a	0.091 (0.016) ^a
Method	OLS	OLS	2SLS	2SLS	OLS	OLS
KP Wald F-Statistic	-	-	256	5.19	-	-
% Decrease	17.9	11.2	11.7	6.9	21.9	21.5
Sobel-ML Test (p-value)	0.000	0.000	0.003	0.128	0.000	0.000
R ² : Model 1	0.014	0.034	-	-	0.057	0.058
R ² : Model 2	0.010	0.028	-	-	0.037	0.038
2 nd -Step: Local Controls	No	Yes	No	Yes	Yes	Yes
1 st -Step: Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	1,289	1,285	480	480	1,214	1,214

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the estimate from a separate regression from equation (15). Sobel-ML Test is the multilevel mediation test from [Krull and MacKinnon \(1999\)](#) and [Krull and MacKinnon \(2001\)](#). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (3) instruments for density using density in 1831 and 1901. The p-value of the over-id test is 0.166 and 0.269 for Models 1 and 2. Column (4) instruments for density and local characteristics, except for the cost of a unit of labor, using: density in 1831 and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, the share of the population in 1968 residing in buildings built before 1949, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence. The p-value of the over-id test is 0.171 and 0.206 for Models 1 and 2.

Table E3: Additional Second-Step Regression Results across Urban Areas

Dependent Variable VA Worker	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Add. Controls (5)
Model 1: Without Org					
log density	0.090 (0.017) ^a	0.088 (0.019) ^a	0.081 (0.015) ^a	0.079 (0.015) ^a	0.067 (0.014) ^a
Model 2: With Org					
log density	0.077 (0.017) ^a	0.081 (0.019) ^a	0.073 (0.015) ^a	0.071 (0.014) ^a	0.064 (0.014) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	14.4	7.9	9.8	10.1	4.4
Sobel-ML Test (p-value)	0.000	0.025	0.000	0.000	0.066
R ² : Model 1	0.047	0.090	0.034	0.034	0.028
R ² : Model 2	0.036	0.077	0.029	0.028	0.026
2 nd -Step: Local Controls	No	Yes	Yes	Yes	Yes
1 st -Step: Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	480	480	1,285	1,285	1,277

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the estimate from a separate regression from equation (15). Sobel-ML Test is the multilevel mediation test from Krull and MacKinnon (1999) and Krull and MacKinnon (2001). Firm productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Column (5) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table E4: Firm-Level Regression Results across Employment Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.049 (0.004) ^a	0.039 (0.009) ^a	0.049 (0.004) ^a	0.042 (0.009) ^a	0.052 (0.008) ^a	0.053 (0.008) ^a
Model 2: With Org						
log density	0.038 (0.003) ^a	0.036 (0.009) ^a	0.038 (0.003) ^a	0.040 (0.009) ^a	0.045 (0.008) ^a	0.045 (0.008) ^a
Method	OLS	OLS	2SLS	2SLS	OLS	OLS
KP Wald F-Statistic	-	-	2701	9.10	-	-
KP Wald F-Stat: M2	-	-	2460	9.10	-	-
% Decrease	22.4	7.6	22.4	4.7	13.4	15.0
Sobel Test (p-value)	0.000	0.000	0.000	0.050	0.001	0.001
R ² : Model 1	0.088	0.092	-	-	0.098	0.074
R ² : Model 2	0.109	0.113	-	-	0.193	0.174
Local Controls	No	Yes	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	27,508	27,508	26,531	26,531	23,459	23,459

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression estimated in one step. Sobel Test is the mediation test from [Sobel \(1982\)](#) and [Sobel \(1986\)](#). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (3) instruments for density using density in 1881 and 1901. The p-value of the over-id test is 0.681 and 0.735 for Models 1 and 2. Column (4) instruments for density and local characteristics, except for the cost of a unit of labor, using: density in 1881 and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, and the share of the population in 1968 residing in buildings built before 1949. The p-value of the over-id test is 0.117 and 0.180 for Models 1 and 2.

Table E5: Additional Firm-Level Regression Results across Employment Areas

Dependent Variable	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Add. Controls (5)
VA Worker					
Model 1: Without Org					
log density	0.048 (0.004) ^a	0.037 (0.009) ^a	0.039 (0.009) ^a	0.038 (0.009) ^a	0.040 (0.008) ^a
Model 2: With Org					
log density	0.037 (0.003) ^a	0.034 (0.009) ^a	0.036 (0.009) ^a	0.035 (0.009) ^a	0.038 (0.008) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	22.9	8.1	7.6	7.8	5.0
Sobel Test (p-value)	0.000	0.001	0.000	0.000	0.001
R ² : Model 1	0.087	0.091	0.092	0.096	0.226
R ² : Model 2	0.108	0.112	0.113	0.113	0.239
Local Controls	No	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	26,531	26,531	27,508	27,508	26,791

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression estimated in one step. Sobel Test is the mediation test from Sobel (1982) and Sobel (1986). Firm productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Column (5) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table E6: Firm-Level Regression Results across Urban Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.073 (0.012) ^a	0.060 (0.012) ^a	0.086 (0.015) ^a	0.051 (0.022) ^b	0.095 (0.009) ^a	0.097 (0.009) ^a
Model 2: With Org						
log density	0.054 (0.009) ^a	0.050 (0.011) ^a	0.067 (0.012) ^a	0.042 (0.021) ^c	0.074 (0.009) ^a	0.074 (0.009) ^a
Method	OLS	OLS	2SLS	2SLS	OLS	OLS
KP Wald F-Stat: M1	-	-	13.28	6.67	-	-
KP Wald F-Stat: M2	-	-	12.83	6.67	-	-
% Decrease	26.0	16.6	22.0	17.6	22.1	23.7
Sobel Test (p-value)	0.000	0.000	0.000	0.000	0.000	0.000
R ² : Model 1	0.074	0.080	-	-	0.097	0.074
R ² : Model 2	0.096	0.102	-	-	0.193	0.175
Local Controls	No	Yes	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	24,197	24,192	20,537	20,537	20,644	20,644

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression estimated in one step. Sobel Test is the mediation test from [Sobel \(1982\)](#) and [Sobel \(1986\)](#). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (3) instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.421 and 0.359 for Models 1 and 2. Column (4) instruments for density and local characteristics, except for the cost of a unit of labor, using: density in 1831, 1851, 1881, and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, and the share of the population in 1968 residing in buildings built before 1949. The p-value of the over-id test is 0.185 and 0.203 for Models 1 and 2.

Table E7: Additional Firm-Level Regression Results across Urban Areas

Dependent Variable VA Worker	IV Sample (1)	IV Sample (2)	Herf.-Hir. Occ. Index (3)	Add. Occ. (4)	Add. Controls (5)
Model 1: Without Org					
log density	0.079 (0.013) ^a	0.044 (0.018) ^b	0.065 (0.013) ^a	0.058 (0.012) ^a	0.054 (0.010) ^a
Model 2: With Org					
log density	0.058 (0.009) ^a	0.034 (0.018) ^c	0.055 (0.012) ^a	0.050 (0.011) ^a	0.051 (0.010) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	26.5	22.7	15.3	13.7	5.5
Sobel Test (p-value)	0.000	0.000	0.000	0.000	0.000
R ² : Model 1	0.068	0.074	0.080	0.084	0.209
R ² : Model 2	0.090	0.095	0.102	0.102	0.224
Local Controls	No	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	20,537	20,537	24,192	24,192	23,558

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression estimated in one step. Sobel Test is the mediation test from Sobel (1982) and Sobel (1986). Firm productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Column (3) controls for the Herfindahl-Hirschman Occupational Index (HHI), measured in the year 1999, which is constructed at the 2-digit occupational level. Column (4) controls for the number of additional occupations in firms, measured as: $\ln(\# \text{ of occupations} - \# \text{ of layers} + 1)$. Column (5) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether firms have only a single establishment and whether they belong to a business group.

Table E8: Regression Results across Employment Areas with Other Measures of Firm Productivity

	Area Level			Firm Level		
	VA Hour (1)	LP TFP (2)	WD TFP (3)	VA Hour (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org log density	0.041 (0.010) ^a	0.059 (0.010) ^a	0.058 (0.009) ^a	0.025 (0.008) ^a	0.040 (0.006) ^a	0.039 (0.006) ^a
Model 2: With Org log density	0.040 (0.010) ^a	0.052 (0.009) ^a	0.051 (0.009) ^a	0.024 (0.007) ^a	0.036 (0.006) ^a	0.035 (0.006) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
% Decrease	2.4	11.8	12.0	4.0	10.0	10.5
Sobel-ML Test (p-value)	0.053	0.001	0.001	-	-	-
Sobel Test (p-value)	-	-	-	0.036	0.001	0.001
R ² : Model 1	0.103	0.241	0.227	0.037	0.811	0.809
R ² : Model 2	0.103	0.168	0.161	0.039	0.823	0.821
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	341	340	340	27,508	23,459	23,459

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses in columns (1)-(3) and clustered standard errors at the area level in columns (4)-(6). Sobel-ML Test is the multilevel mediation test from [Krull and MacKinnon \(1999\)](#) and [Krull and MacKinnon \(2001\)](#). Sobel Test is the mediation test from [Sobel \(1982\)](#) and [Sobel \(1986\)](#). Columns (1)-(3) reports results from a separate regression from equation (15), while columns (4)-(6) report regression results from a single step. In columns (1) and (4) firm productivity is value-added per hour. In columns (2)-(3) and (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of hours. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

Table E9: Regression Results across Urban Areas with Other Measures of Firm Productivity

	Area Level			Firm Level		
	VA Hour (1)	LP TFP (2)	WD TFP (3)	VA Hour (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.005 (0.012)	0.062 (0.013) ^a	0.058 (0.013) ^a	0.019 (0.007) ^a	0.061 (0.007) ^a	0.057 (0.007) ^a
Model 2: With Org						
log density	0.002 (0.012)	0.048 (0.013) ^a	0.044 (0.013) ^a	0.016 (0.007) ^b	0.047 (0.007) ^a	0.044 (0.007) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
% Decrease	40.0	22.5	24.1	14.6	22.9	22.8
Sobel-ML Test (p-value)	0.050	0.000	0.000	-	-	-
Sobel Test (p-value)	-	-	-	0.027	0.000	0.000
R ² : Model 1	0.027	0.036	0.035	0.034	0.831	0.832
R ² : Model 2	0.029	0.028	0.028	0.037	0.843	0.843
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	1,285	1,214	1,214	24,192	20,644	20,644

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses in columns (1)-(3) and clustered standard errors at the area level in columns (4)-(6). Sobel-ML Test is the multilevel mediation test from [Krull and MacKinnon \(1999\)](#) and [Krull and MacKinnon \(2001\)](#). Sobel Test is the mediation test from [Sobel \(1982\)](#) and [Sobel \(1986\)](#). Columns (1)-(3) reports results from a separate regression from equation (15), while columns (4)-(6) report regression results from a single step. In columns (1) and (4) firm productivity is value-added per hour. In columns (2)-(3) and (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of hours. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

Table E10: Regression Results across Employment Areas with Other Measures of Firm Organization

	Area Level			Firm Level		
	VA Worker (1)	LP TFP (2)	WD TFP (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org log density	0.045 (0.013) ^a	0.074 (0.013) ^a	0.075 (0.013) ^a	0.039 (0.009) ^a	0.052 (0.008) ^a	0.053 (0.008) ^a
Model 3: With Org log density	0.041 (0.013) ^a	0.063 (0.012) ^a	0.064 (0.012) ^a	0.035 (0.009) ^a	0.045 (0.008) ^a	0.045 (0.008) ^a
Model 4: With Org log density	0.041 (0.013) ^a	0.061 (0.012) ^a	0.062 (0.012) ^a	0.036 (0.009) ^a	0.045 (0.008) ^a	0.046 (0.008) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
Model 1 vs Model 3 % Decrease	8.8	14.8	14.6	10.2	13.4	15.0
Model 1 vs Model 4 % Decrease	8.8	17.5	17.3	7.6	13.4	13.2
R ² : Model 1	0.210	0.311	0.314	0.092	0.098	0.074
R ² : Model 2	0.163	0.211	0.212	0.113	0.194	0.175
R ² : Model 3	0.162	0.202	0.203	0.115	0.204	0.185
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	341	340	340	27,508	23,459	23,459

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses in columns (1)-(3) and clustered standard errors at the area level in columns (4)-(6). Columns (1)-(3) reports results from a separate regression from equation (15), while columns (4)-(6) report regression results from a single step. In columns (1) and (4) firm productivity is value-added per worker. In columns (2)-(3) and (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Model 3 uses organization fixed effects to account for the number of layers in firms. Model 4 uses organization-industry fixed effects. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

Table E11: Regression Results across Urban Areas with Other Measures of Firm Organization

	Area Level			Firm Level		
	VA Worker (1)	LP TFP (2)	WD TFP (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org log density	0.080 (0.015) ^a	0.114 (0.017) ^a	0.116 (0.017) ^a	0.060 (0.012) ^a	0.095 (0.009) ^a	0.097 (0.009) ^a
Model 3: With Org log density	0.071 (0.014) ^a	0.090 (0.016) ^a	0.092 (0.016) ^a	0.050 (0.011) ^a	0.074 (0.009) ^a	0.075 (0.009) ^a
Model 4: With Org log density	0.071 (0.014) ^a	0.085 (0.016) ^a	0.087 (0.016) ^a	0.050 (0.011) ^a	0.072 (0.009) ^a	0.073 (0.009) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
Model 1 vs Model 3 % Decrease	11.2	21.0	20.6	16.6	16.6	22.6
Model 1 vs Model 4 % Decrease	11.2	25.4	25.0	16.6	16.6	24.7
R ² : Model 1	0.034	0.057	0.058	0.084	0.097	0.074
R ² : Model 2	0.029	0.038	0.038	0.102	0.194	0.176
R ² : Model 3	0.028	0.033	0.034	0.104	0.203	0.186
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	1,285	1,214	1,214	24,192	20,649	20,649

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses in columns (1)-(3) and clustered standard errors at the area level in columns (4)-(6). Columns (1)-(3) reports results from a separate regression from equation (15), while columns (4)-(6) report regression results from a single step. In columns (1) and (4) firm productivity is value-added per worker. In columns (2)-(3) and (5)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Model 3 uses organization fixed effects to account for the number of layers in firms. Model 4 uses organization-industry fixed effects. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

F Additional Results: The Manufacturing Sector

This section contains additional results that examine firm organization and productivity across locations. Tables in this section are the following:

1. Table **F1**: Manufacturing Sector - Regression Results across Employment Areas
2. Table **F2**: Manufacturing Sector - Second-Step Regression Results across Employment Areas

Table F1: Manufacturing Sector - Regression Results across Employment Areas

	(1)	(2)	(3)	(4)	(5)
Probability One-Layer Firm					
log density	-0.013 (0.000) ^a	-0.013 (0.000) ^a	-0.005 (0.001) ^a	-0.012 (0.001) ^a	-0.000 (0.002)
Probability Two-Layer Firm					
log density	-0.007 (0.002) ^a	-0.008 (0.002) ^a	-0.006 (0.002) ^b	-0.006 (0.002) ^a	-0.005 (0.003)
Probability Three-Layer Firm					
log density	0.009 (0.001) ^a	0.010 (0.001) ^a	0.003 (0.002)	0.009 (0.001) ^a	-0.002 (0.003)
Probability Four-Layer Firm					
log density	0.010 (0.001) ^a	0.011 (0.001) ^a	0.008 (0.002) ^a	0.009 (0.001) ^a	0.008 (0.002) ^a
Method	OLS	OLS	OLS	2SLS	2SLS
KP Wald F-Statistic	-	-	-	987	5.53
R ² : One-Layer Firm	0.407	0.407	0.408	-	-
R ² : Two-Layer Firm	0.033	0.033	0.034	-	-
R ² : Three-Layer Firm	0.122	0.122	0.123	-	-
R ² : Four-Layer Firm	0.089	0.089	0.090	-	-
Wage Controls*	No	Yes	Yes	No	Yes
Income Controls	No	No	Yes	No	Yes
Demographic Controls	No	No	Yes	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Sample Size	53,601	53,601	53,601	50,108	50,108

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. *: variables always treated as exogenous. Clustered standard errors at the area level in parentheses. Entries display the estimate from a separate regression of equation (13). Industry fixed effects are at the 4-digit NAF Rev 1. level. Wage controls include the cost of a unit of labor. Income controls include the median annual salary of individuals residing in an area. Demographic controls include the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (4) instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.423, 0.154, 0.492, and 0.556, when the dependent variable is the probability of a firm having respectively one, two, three or four layers. Column (5) instruments for density and local characteristics using: density in 1831, 1881, and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, the share of the population in 1968 living in their principal residence, the share of the population in 1968 that own their housing, the share of the population in 1968 with a kitchen their residence, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence. The p-value of the over-id test is respectively 0.055, 0.045, 0.686, and 0.614, when the dependent variable is the probability of a firm having one, two, three or four layers.

Table F2: Manufacturing Sector - Second-Step Regression Results across Employment Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	LP TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.035 (0.004) ^a	0.030 (0.005) ^a	0.025 (0.004) ^a	0.030 (0.006) ^a	0.052 (0.005) ^a	0.050 (0.005) ^a
Model 2: With Org						
log density	0.027 (0.004) ^a	0.026 (0.004) ^a	0.018 (0.004) ^a	0.025 (0.005) ^a	0.047 (0.005) ^a	0.043 (0.005) ^a
Method	OLS	OLS	2SLS	2SLS	OLS	OLS
KP Wald F-Statistic	-	-	335	9.09	-	-
% Decrease	22.8	13.3	28.0	16.6	9.6	14.0
Sobel-ML Test (p-value)	0.000	0.000	0.000	0.001	0.000	0.000
R ² : Model 1	0.182	0.452	-	-	0.467	0.445
R ² : Model 2	0.130	0.379	-	-	0.400	0.375
2 nd -Step: Local Controls	No	Yes	No	Yes	Yes	Yes
1 st -Step: Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	341	341	304	304	341	341

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Robust standard errors in parentheses. Entries display the estimate from a separate regression from equation (15). Sobel-ML Test is the multilevel mediation test from [Krull and MacKinnon \(1999\)](#) and [Krull and MacKinnon \(2001\)](#). In columns (1)-(4) firm productivity is value-added per worker. In columns (5)-(6) productivity is TFP estimated using the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level included in the first step. Local area controls refer to: the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Column (3) instruments for density using density in 1831, 1851, 1881, and 1901. The p-value of the over-id test is 0.092 and 0.063 for Models 1 and 2. Column (4) instruments for density and local characteristics, except for the cost of a unit of labor, using: density in 1831, 1851, 1881, and 1901, the median annual salary of individuals residing in an area in 1993, the share of the local population between the ages of 25 and 39 in 1968, the share of the local population born outside France in 1968, the share of active workers unemployed in 1968, average persons per household in 1968, the share of the population in 1968 living in their principal residence, the share of the population in 1968 that own their housing, the share of the population in 1968 with a kitchen their residence, the share of the population in 1968 with heating in their residence, the share of the population in 1968 with a toilet in their residence, the share of the population in 1968 with a bathtub or shower in their residence, and the share of the population in 1968 having access to a telephone in their residence. The p-value of the over-id test is 0.465 and 0.402 for Models 1 and 2.

G Short-Term Gains & Analysis Within Firms

G.1 Main Analysis

Up to here, the analysis has examined the static gains from operating in denser areas. In this final section, I turn to an analysis of the short-term gains. To motivate results, I first return to the theoretical model to illustrate how different measures of firm productivity change with a re-organization of a firm. I then use a panel dataset to examine the short-term productivity gains from operating in denser areas, while taking into account changes in firm organization. In the interest of space, this section only focuses on the main findings. The full set of results are reported in the Online Appendix G.2.

One result of the model that has not received any attention thus far, is how the different measures of firm productivity adjust to a re-organization of a firm. Figure G.1 uses the same simulations from Section 3.4 to illustrate these patterns. Figure G.1a measures productivity as the inverse of average costs, while in Figure G.1b productivity is equal to revenue per worker. Both figures show that the different measures of firm productivity respond differently to a re-organization of a firm. When a firm increases its number of layers, the inverse of average costs increases discontinuously, while revenue per worker decreases discontinuously. The Supplementary Appendix H also illustrates results with productivity defined as output per worker and revenue per labor costs. It shows that both measures decrease discontinuously when a firm increases its number of layers. These patterns are the same as in [Caliendo and Rossi-Hansberg \(2012\)](#), and they have also been examined in [Caliendo et al. \(2015a\)](#). They are summarized below:

Summary 10 *Let $n(\alpha)$ denote the total number of workers in a firm with demand draw α . When a firm increases its number of layers: (i) the inverse of average costs, $1/AC(\alpha)$, increases discontinuously, (ii) output per worker, $q(\alpha)/n(\alpha)$, decreases discontinuously, (iii) revenue per worker, $r(\alpha)/n(\alpha)$, decreases discontinuously, and (iv) revenue per labor costs, $r(\alpha)/C(\alpha)$, decreases discontinuously.*

These results are relevant for empirical specifications relying on short-term variation within firms. More precisely, they imply that depending on the measure of firm productivity, the short-term gains from operating in denser areas may be greater, or smaller, once we account for firm organization. For instance, consider the following first-difference equation:

$$\Delta \log \phi_{i,k,a,t} = \alpha + \gamma \Delta \log \text{density}_{a,t} + \lambda \Delta \text{ORG}_{i,k,a,t} + \Delta X_{i,k,a,t} \beta + \Delta \epsilon_{i,k,a,t}, \quad (55)$$

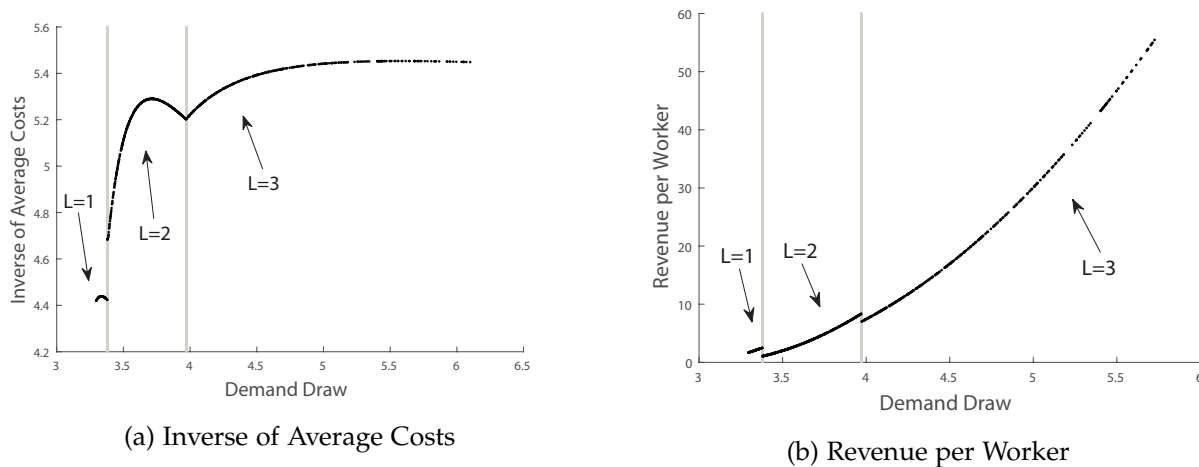
where $\phi_{i,k,a,t}$ is a measure of the productivity of firm i operating in area a , in industry k , and at time t , $\text{ORG}_{i,k,a,t}$ measures the number of layers in firms, and Δ denotes changes from period $t - 1$ to t .

Equation (55) identifies the coefficient on firm organization from changes in the number of layers in firms. Now consider the baseline measure of firm productivity: value-added per worker. It follows from the discussion above that we should expect the estimated coefficient on firm organization, λ , to be negative. And, to the extent that firm organization is determined by the density of areas, it follows that the short-term gains from operating in denser areas, estimated in γ , should be greater once we account for changes in the organization of firms.

To examine this claim, I return to the main industries of the study and use an unbalanced panel dataset that covers the years 2000-2007 (details on the construction of the data are provided in the Data Appendix). Results are reported in Table F3. All specifications contain industry-year fixed effects and standard errors clustered at the local area level.

Before moving to the results, however, it is important to emphasize two differences between the variables in the panel and the previous data. First, since the population of areas is not counted every year, the panel data measures the size of areas using employment density, defined as the

Figure G.1: Productivity and Firm Re-Organization



Notes: Panel (a) illustrates the optimal relationship between the inverse of average costs and the heterogeneous demand draw of a firm. Panel (b) illustrates the optimal relationship between the revenue per worker and the heterogeneous demand draw of a firm. Demand Draw denotes the demand parameter, α , and L , the total number of layers in a firm. The parameters used to simulate the model are reported in the second row of Table H1 in the Supplementary Appendix H.

total number of private sector jobs in an area divided by the area's surface, measured in hectares.⁵⁴ And second, because in the years 2000-2001, the DADS does not report information on the identity of workers, in all years, labor and the number of workers, are approximated using jobs. Using the panel data, columns 1-2 of Tables G1-G4 report results from the main firm level specifications from the previous sections. The point estimates are similar and lead to the same conclusions, suggesting that the difference in variables is not a cause for concern.⁵⁵

Columns 1-2 of Table F3 first examine the short-term relationship between the employment density of areas and the organization of firms. To save space, both columns report results with dependent variable the total number of layers in firms. Column 1 only controls for the employment density of areas, while column 2 also controls for local characteristics. In all cases, the estimated coefficient on employment density has the expected sign and the findings remain consistent with the theory: an increase in the density of an area is associated with an increase in the number of layers in firms.⁵⁶

The last four columns of Table F3 turn to firm productivity and report results from specifications with and without controls for the number of layers in firms. Columns 3-4 measure productivity as value-added per worker. In both cases, the estimated coefficient on employment density has the expected sign, yet it is not significant at conventional levels. This limits the ability of this section to draw strong conclusions, yet the magnitudes are consistent with the discussion above. In both

⁵⁴The correlation between population density, and employment density from the year 2004, is 0.961 across employment areas and 0.847 across urban areas.

⁵⁵To show that the findings are not specific to the year 2004, columns 3-4 of Tables G1-G4 report results grouping all years together. In all cases, the point estimates on density suggest that the conclusions from previous sections are not unique to a particular year. And, the last column of Tables G1-G4 also report results from specifications that control for the time-invariant characteristics of areas, using area fixed effects. The findings are weaker yet the conclusions are consistent with previous results.

⁵⁶In the Online Appendix G.2, Table G5 reports results on the distribution of firm organization. Even though point estimates are less precisely estimated, the findings suggest that an increase in the density of an area is associated with a decrease in the number of one-layer and two-layer firms, and an increase in the number of three-layer and four-layer firms.

Table F3: Organization and Productivity Panel Regression Results

	Number of Layers			VA	VA	LP	CMORH LP
	(1)	(2)		Worker	Worker	TFP	TFP + Org
				(3)	(4)	(5)	(6)
			M1: W/out Org				
log density	0.064 (0.026) ^b	0.058 (0.026) ^b	log density	0.028 (0.045)	0.018 (0.043)	0.052 (0.031) ^c	0.048 (0.026) ^c
			M2: With Org				
			log density	0.030 (0.046)	0.020 (0.043)	0.052 (0.031) ^c	0.049 (0.026) ^c
			organization	-0.036 (0.002) ^a	-0.036 (0.002) ^a	0.001 (0.001)	-0.014 (0.001) ^a
Method	OLS	OLS	Method	OLS	OLS	OLS	OLS
R ²	0.005	0.005	R ² : M1	0.001	0.001	0.003	0.003
			R ² : M2	0.004	0.004	0.003	0.004
Ind-Year FE	Yes	Yes	Ind-Year FE	Yes	Yes	Yes	Yes
Local Controls	No	Yes	Local Controls	No	Yes	Yes	Yes
Sample Size	110,477	110,477	Sample Size	110,477	110,477	104,024	104,024

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

columns, the coefficient on firm organization is statistically significant, and its magnitude implies that an extra layer is associated with a 3.6% decrease in the productivity of a firm. Comparing the point estimates from Models 1 and 2, also illustrates the main argument of this section: the short-term gains from operating in denser areas may be greater once we account for changes in firm organization. For instance, the magnitudes in column 1 suggest that accounting for firm organization increases the short-term productivity gains by roughly 7.1%.

Columns 5-6 further assess whether the findings are similar with measures of firm TFP. In column 5, firm TFP is estimated using a Cobb-Douglas production function with capital and labor, and the identifying assumptions from [Levinsohn and Petrin \(2003\)](#). In this case, the point estimate on firm organization is insignificant and its magnitude is nearly zero.⁵⁷ Column 6 further reports results based on measures of firm TFP from [Caliendo et al. \(2015a\)](#).⁵⁸ In this case, the coefficient on firm organization is negative and significant, and its magnitude implies that an extra layer is associated with a 1.4% decrease in the productivity of a firm. The point estimates further suggest that accounting for firm organization increases the short-term gains from operating in denser areas by 2.0%.

Additional results are reported in the Online Appendix [G.2](#).⁵⁹ Moreover, it is important to

⁵⁷This is a bit surprising, however, TFP is estimated using assumptions that are outside of the model. I also have estimated equation (55) separately for each year. The coefficient on firm organization is negative in some years while positive in others, when using TFP from a Cobb-Douglas production function with only capital and labor. The coefficient on organization remains negative across all years, when using the other measures of firm productivity.

⁵⁸This measure of TFP is not the same as in Table 7. Given the estimated production function, $\ln va_{it} = \hat{\alpha}_k \ln k_{it} + \hat{\alpha}_C C(O_{it}, w) + \hat{\alpha}_O ORG_{it} + \hat{\epsilon}_{it}$, TFP is equal to: $TFP_{it} = \exp(\ln va_{it} - \hat{\alpha}_k \ln k_{it} - \hat{\alpha}_C C(O_{it}, w))$.

⁵⁹Tables [G6-G11](#) reports results using different specifications and other measures of firm TFP. Table [G7](#) also illustrates that standard measures of firm TFP, estimated using Cobb-Douglas production function with capital and labor, and with the approaches of [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#) do not always decrease with a re-organization of

emphasize that the findings in this section are consistent with the conclusions from Section 5.2. Both sections focus on different type of gains from operating in denser areas - static and short-term - and as result, they use different variation from the data to identify the coefficient on density and firm organization. Put differently, despite the fact that value-added per worker decreases from an additional layer in firms, as firms move along their new production schedule, they ultimately become more productive.⁶⁰

firms. This may be due to the structural assumptions used to estimate firm TFP, which include assumptions about the production function that are not part of the model. This may also be due to the fact that value-added increases faster than a weighted average of capital and labor. However, it can also be a result of some measures of firm TFP increasing faster, as firms move along their new production schedule. Additionally, Table G8 also reports results using estimates of TFP from [Caliendo et al. \(2015a\)](#). The findings are consistent with their analysis, and indicate that an additional layer is associated with a 0.9% to a 1.7% drop in the productivity of a firm.

⁶⁰To illustrate this effect, Tables G10 and G11 report results from specifications with longer lags, using the sample to firms observed 2003 and 2006 and for at least 3 years. The findings still indicate that an extra layer is associated with a decrease in firm productivity. However, they also suggest that as equation (55) is estimated using longer lags the magnitude of the coefficient on firm organization decreases, suggesting that firms become more productive as they move along their new production schedule.

G.2 Additional Tables

This remaining section contains additional results that examine the outcomes of firms using the panel data from the years 2000-2007. The tables in this section are the following:

1. Table G1: Robustness Checks: Panel Regression Results on the Distribution of Firm Organization - Employment Areas
2. Table G2: Robustness Checks: Panel Value-Added per Worker Productivity Regression Results - Employment Areas
3. Table G3: Robustness Checks: Panel LP TFP Productivity Regression Results - Employment Areas
4. Table G4: Robustness Checks: Panel WD TFP Productivity Regression Results - Employment Areas
5. Table G5: Panel Regression Results on the Distribution of Firm Organization - Employment Areas
6. Table G6: Value-Added per Worker Panel Regression Results - Employment Areas
7. Table G7: TFP Panel Regression Results - Employment Areas
8. Table G8: CMORH TFP + Org Panel Regression Results - Employment Areas
9. Table G9: CMORH TFP Panel Regression Results - Employment Areas
10. Table G10: Value-Added per Worker & TFP Panel Regression Results with Different Lags - Employment Areas
11. Table G11: CMORH TFP + Org Panel Regression Results with Different Lags - Employment Areas

Table G1: Robustness Checks: Panel Regression Results on the Distribution of Firm Organization - Employment Areas

	(1)	(2)	(3)	(4)	(5)
Total Number of Layers					
log density	0.032 (0.008) ^a	0.028 (0.006) ^a	0.020 (0.008) ^b	0.018 (0.005) ^a	0.087 (0.053) ^c
Probability One-Layer Firm					
log density	-0.011 (0.007)	-0.009 (0.005) ^c	-0.003 (0.005)	-0.002 (0.004)	-0.027 (0.026)
Probability Two-Layer Firm					
log density	-0.004 (0.007)	-0.005 (0.005)	-0.008 (0.005)	-0.009 (0.003) ^b	-0.013 (0.023)
Probability Three-Layer Firm					
log density	0.010 (0.003) ^a	0.010 (0.002) ^a	0.008 (0.002) ^a	0.007 (0.001) ^a	0.021 (0.016)
Probability Four-Layer Firm					
log density	0.005 (0.001) ^a	0.004 (0.000) ^a	0.004 (0.000) ^a	0.003 (0.000) ^a	0.019 (0.013)
Method	OLS	OLS	OLS	OLS	OLS
R ² : Total Number of Layers	0.078	0.078	0.090	0.090	0.105
R ² : One-Layer Firm	0.052	0.052	0.063	0.063	0.080
R ² : Two-Layer Firm	0.004	0.004	0.006	0.007	0.016
R ² : Three-Layer Firm	0.037	0.037	0.044	0.044	0.051
R ² : Four-Layer Firm	0.016	0.016	0.016	0.016	0.020
Year 2004	Yes	Yes	No	No	No
All Years	No	No	Yes	Yes	Yes
Population Density	Yes	No	Yes	No	No
Employment Density	No	Yes	No	Yes	Yes
Wage Controls	Yes	Yes	Yes	Yes	No
Income Controls	Yes	Yes	Yes	Yes	No
Demographic Controls	No	Yes	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Area FE	No	No	No	No	Yes
Sample Size	20,411	20,411	164,822	164,822	164,822

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display an estimate from a separate regression. In columns (1) and (3) the size of areas is measured using population density. In columns (2), (4), and (5) the size of areas is measured using employment density. Columns (1)-(2) restrict the sample to the year 2004, while columns (3)-(5) use all years and estimates the models using year fixed effects. Column (5) additionally includes area fixed effects.

Table G2: Robustness Checks: Panel Value-Added per Worker Productivity Regression Results - Employment Areas

VA Per Worker	(1)	(2)	(3)	(4)	(5)
Model 1: Without Org					
log density	0.045 (0.010) ^a	0.035 (0.006) ^a	0.041 (0.007) ^a	0.033 (0.004) ^a	0.032 (0.034)
Model 2: With Org					
log density	0.040 (0.010) ^a	0.031 (0.006) ^a	0.039 (0.007) ^a	0.031 (0.004) ^a	0.021 (0.032)
organization	0.140 (0.013) ^a	0.140 (0.013) ^a	0.126 (0.008) ^a	0.125 (0.008) ^a	0.122 (0.009) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	11.1	11.4	4.8	6.0	34.3
Sobel Test (p-value)	0.000	0.000	0.006	0.000	0.051
R ² : Model 1	0.090	0.090	0.093	0.093	0.110
R ² : Model 2	0.115	0.115	0.113	0.113	0.129
Year 2004	Yes	Yes	No	No	No
All Years	No	No	Yes	Yes	Yes
Population Density	Yes	No	Yes	No	No
Employment Density	No	Yes	No	Yes	Yes
Wage Controls	Yes	Yes	Yes	Yes	No
Income Controls	Yes	Yes	Yes	Yes	No
Demographic Controls	No	Yes	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Area FE	No	No	No	No	Yes
Sample Size	20,411	20,411	164,822	164,822	164,822

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display an estimate from a separate regression. Sobel Test is the mediation test from Sobel (1982) and Sobel (1986). Productivity is value-added per worker. In columns (1) and (3) the size of areas is measured using population density. In columns (2), (4), and (5) the size of areas is measured using employment density. Columns (1)-(2) restrict the sample to the year 2004, while columns (3)-(5) use all years and estimates the models using year fixed effects. Column (5) additionally includes area fixed effects.

Table G3: Robustness Checks: Panel LP TFP Productivity Regression Results - Employment Areas

LP TFP	(1)	(2)	(3)	(4)	(5)
Model 1: Without Org					
log density	0.058 (0.008) ^a	0.047 (0.006) ^a	0.057 (0.008) ^a	0.048 (0.005) ^a	0.074 (0.040) ^c
Model 2: With Org					
log density	0.049 (0.008) ^a	0.040 (0.005) ^a	0.052 (0.007) ^a	0.044 (0.005) ^a	0.048 (0.033)
organization	0.289 (0.014) ^a	0.289 (0.014) ^a	0.283 (0.009) ^a	0.283 (0.010) ^a	0.278 (0.010) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	15.5	14.8	8.7	8.3	54.1
Sobel Test (p-value)	0.000	0.000	0.007	0.001	0.037
R ² : Model 1	0.089	0.090	0.102	0.103	0.123
R ² : Model 2	0.185	0.185	0.191	0.192	0.207
Year 2004	Yes	Yes	No	No	No
All Years	No	No	Yes	Yes	Yes
Population Density	Yes	No	Yes	No	No
Employment Density	No	Yes	No	Yes	Yes
Wage Controls	Yes	Yes	Yes	Yes	No
Income Controls	Yes	Yes	Yes	Yes	No
Demographic Controls	No	Yes	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Area FE	No	No	No	No	Yes
Sample Size	19,332	19,332	156,293	156,293	156,293

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display an estimate from a separate regression. Sobel Test is the mediation test from [Sobel \(1982\)](#) and [Sobel \(1986\)](#). Productivity is TFP estimated using a Cobb-Douglas production function and the approach proposed by [Levinsohn and Petrin \(2003\)](#), with labor measured using the number of workers. In columns (1) and (3) the size of areas is measured using population density. In columns (2), (4), and (5) the size of areas is measured using employment density. Columns (1)-(2) restrict the sample to the year 2004, while columns (3)-(5) use all years and estimates the models using year fixed effects. Column (5) additionally includes area fixed effects.

Table G4: Robustness Checks: Panel WD TFP Productivity Regression Results - Employment Areas

WD TFP	(1)	(2)	(3)	(4)	(5)
Model 1: Without Org					
log density	0.058 (0.008) ^a	0.048 (0.006) ^a	0.058 (0.008) ^a	0.049 (0.005) ^a	0.079 (0.042) ^c
Model 2: With Org					
log density	0.049 (0.008) ^a	0.039 (0.006) ^a	0.052 (0.007) ^a	0.044 (0.005) ^a	0.052 (0.034)
organization	0.310 (0.015) ^a	0.310 (0.015) ^a	0.305 (0.010) ^a	0.304 (0.010) ^a	0.299 (0.011) ^a
Method	OLS	OLS	OLS	OLS	OLS
% Decrease	15.5	18.7	10.3	10.2	34.1
Sobel Test (p-value)	0.000	0.000	0.007	0.001	0.037
R ² : Model 1	0.101	0.102	0.115	0.116	0.136
R ² : Model 2	0.204	0.205	0.212	0.213	0.227
Year 2004	Yes	Yes	No	No	No
All Years	No	No	Yes	Yes	Yes
Population Density	Yes	No	Yes	No	No
Employment Density	No	Yes	No	Yes	Yes
Wage Controls	Yes	Yes	Yes	Yes	No
Income Controls	Yes	Yes	Yes	Yes	No
Demographic Controls	No	Yes	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Area FE	No	No	No	No	Yes
Sample Size	19,332	19,332	156,293	156,293	156,293

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the area level in parentheses. Entries display an estimate from a separate regression. Sobel Test is the mediation test from [Sobel \(1982\)](#) and [Sobel \(1986\)](#). Productivity is TFP estimated using a Cobb-Douglas production function and the approach proposed by [Wooldridge \(2009\)](#), with labor measured using the number of workers. In columns (1) and (3) the size of areas is measured using population density. In columns (2), (4), and (5) the size of areas is measured using employment density. Columns (1)-(2) restrict the sample to the year 2004, while columns (3)-(5) use all years and estimates the models using year fixed effects. Column (5) additionally includes area fixed effects.

Table G5: Panel Regression Results on the Distribution of Firm Organization - Employment Areas

	(1)	(2)	(3)	(4)	(5)	(6)
Total Number of Layers						
log density	0.064 (0.026) ^b	0.058 (0.026) ^b	0.060 (0.026) ^b	0.152 (0.045) ^a	0.111 (0.038) ^a	0.081 (0.035) ^b
Probability One-Layer Firm						
log density	-0.015 (0.017)	-0.012 (0.017)	-0.012 (0.017)	-0.061 (0.023) ^a	-0.045 (0.023) ^b	-0.032 (0.021)
Probability Two-Layer Firm						
log density	-0.019 (0.024)	-0.019 (0.024)	-0.023 (0.024)	-0.007 (0.025)	-0.003 (0.028)	-0.006 (0.028)
Probability Three-Layer Firm						
log density	0.021 (0.022)	0.019 (0.022)	0.022 (0.022)	0.047 (0.019) ^a	0.033 (0.021)	0.027 (0.021)
Probability Four-Layer Firm						
log density	0.013 (0.010)	0.013 (0.010)	0.012 (0.010)	0.021 (0.013)	0.016 (0.011)	0.010 (0.011)
Method	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	FD	FD	Firm FE	Firm FE	Firm FE
R ² : Total Number of Layers	0.005	0.005	0.006	0.723	0.723	0.730
R ² : One-Layer Firm	0.002	0.002	0.003	0.687	0.687	0.692
R ² : Two-Layer Firm	0.001	0.001	0.001	0.521	0.521	0.524
R ² : Three-Layer Firm	0.002	0.002	0.002	0.533	0.533	0.539
R ² : Four-Layer Firm	0.001	0.001	0.002	0.488	0.489	0.493
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	No	Yes	Yes	No	Yes	Yes
Wage Controls	No	Yes	Yes	No	Yes	Yes
Income Controls	No	Yes	Yes	No	Yes	Yes
Firm Controls	No	No	Yes	No	No	Yes
Sample Size	110,477	110,477	107,789	143,324	143,324	140,064

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Columns (1)-(3) estimate the model using first differences. Columns (4)-(6) estimate the model using firm fixed effects. Column (3) controls for the size, the number of additional occupations and the amount of capital in firms. Column (6) controls for firm size (the number of workers), the number of additional occupations, the amount of capital, and the legal status of firms, as well as whether they belong to a business group. In specifications with firm fixed effects, the demographic, income and wage controls are interacted with a linear time trend.

Table G6: Value-Added per Worker Panel Regression Results - Employment Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	VA Worker (4)	VA Worker (5)	VA Worker (6)
Model 1: Without Org						
log density	0.028 (0.046)	0.018 (0.043)	0.050 (0.035)	0.026 (0.033)	0.032 (0.040)	0.007 (0.037)
Model 2: With Org						
log density	0.030 (0.046)	0.020 (0.043)	0.054 (0.036)	0.028 (0.033)	0.038 (0.040)	0.011 (0.036)
organization	-0.036 (0.002) ^a	-0.036 (0.002) ^a	-0.023 (0.003) ^a	-0.024 (0.003) ^a	-0.031 (0.002) ^a	-0.032 (0.002) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	FD	Firm FE	Firm FE	Firm FE	Firm FE
R ² : Model 1	0.002	0.002	0.762	0.762	0.818	0.818
R ² : Model 2	0.004	0.004	0.763	0.763	0.819	0.819
Lagged Productivity	No	No	No	No	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	No	Yes	No	Yes	No	Yes
Sample Size	110,477	110,477	143,324	143,324	110,477	110,477

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. Productivity is value-added per worker. Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Columns (1)-(2) estimate the model using first differences. Columns (3)-(4) estimate the model using firm fixed effects. Columns (5)-(6) estimate the model using firm fixed effects and using lagged productivity as an additional control. In columns (3)-(6), the sample is restricted to firms that are observed for at least 2 years. In specifications with firm fixed effects, the demographic, income and wage controls are interacted with a linear time trend.

Table G7: TFP Panel Regression Results - Employment Areas

	LP TFP (1)	LP TFP (2)	LP TFP (3)	WD TFP (4)	WD TFP (5)	WD TFP (6)
Model 1: Without Org						
log density	0.052 (0.031) ^c	0.109 (0.036) ^a	0.069 (0.029) ^b	0.053 (0.031) ^c	0.112 (0.037) ^a	0.072 (0.029) ^b
Model 2: With Org						
log density	0.052 (0.031) ^c	0.106 (0.036) ^a	0.068 (0.029) ^b	0.053 (0.031) ^c	0.109 (0.036) ^a	0.070 (0.029) ^b
organization	0.001 (0.001)	0.021 (0.002) ^a	0.007 (0.002) ^a	0.002 (0.001)	0.023 (0.002) ^a	0.009 (0.002) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	Firm FE	Firm FE	FD	Firm FE	Firm FE
R ² : Model 1	0.003	0.847	0.907	0.003	0.856	0.913
R ² : Model 2	0.003	0.847	0.907	0.003	0.856	0.913
Lagged Productivity	No	No	Yes	No	No	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	104,024	136,105	104,024	104,024	136,105	104,024

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. In columns (1)-(3) productivity is TFP estimated using a Cobb-Douglas production function and the approach proposed by [Levinsohn and Petrin \(2003\)](#), with labor measured using the number of workers. In columns (4)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approach proposed by [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Columns (1) and (4) estimate the model using first differences. Columns (2) and (5) estimate the model using firm fixed effects. Columns (3) and (6) estimate the model using firm fixed effects and using lagged firm productivity as an additional control. In columns (2),(4),(5) and (6) the sample is restricted to firms that are observed for at least 2 years. In columns (2)-(3) and (5)-(6) the demographic, income and wage controls are interacted with a linear time trend.

Table G8: CMORH TFP + Org Panel Regression Results - Employment Areas

	CMORH LP TFP + Org (1)	CMORH LP TFP + Org (2)	CMORH LP TFP + Org (3)	CMORH WD TFP + Org (4)	CMORH WD TFP + Org (5)	CMORH WD TFP + Org (6)
Model 1: Without Org						
log density	0.048 (0.026) ^c	0.079 (0.030) ^a	0.065 (0.025) ^a	0.046 (0.026) ^c	0.073 (0.029) ^b	0.062 (0.025) ^b
Model 2: With Org						
log density	0.049 (0.026) ^c	0.080 (0.030) ^a	0.067 (0.025) ^a	0.047 (0.026) ^c	0.074 (0.029) ^b	0.065 (0.025) ^a
organization	-0.014 (0.001) ^a	-0.008 (0.002) ^a	-0.013 (0.001) ^a	-0.017 (0.001) ^a	-0.013 (0.002) ^a	-0.017 (0.001) ^a
R ² : Model 1	0.003	0.966	0.980	0.003	0.968	0.980
R ² : Model 2	0.004	0.966	0.980	0.004	0.968	0.981
Method	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	Firm FE	Firm FE	FD	Firm FE	Firm FE
Lagged Productivity	No	No	Yes	No	No	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	104,024	136,105	104,024	104,024	136,105	104,024

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. Firm TFP is defined as the residual along with the effect of organization, both of which are estimated from the production function from [Caliendo et al. \(2015a\)](#) and using the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#). Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Columns (1) and (4) estimate the model using first differences. Columns (2) and (5) estimate the model using firm fixed effects. Columns (3) and (6) estimate the model using firm fixed effects and using lagged firm productivity as an additional control. In columns (2),(4),(5) and (6) the sample is restricted to firms that are observed for at least 2 years. In specifications with firm fixed effects, the demographic, income and wage controls are interacted with a linear time trend.

Table G9: CMORH TFP Panel Regression Results - Employment Areas

	CMORH LP TFP (1)	CMORH LP TFP (2)	CMORH LP TFP (3)	CMORH WD TFP (4)	CMORH WD TFP (5)	CMORH WD TFP (6)
log density	0.051 (0.026) ^c	0.085 (0.030) ^a	0.073 (0.024) ^a	0.049 (0.026) ^c	0.078 (0.030) ^a	0.069 (0.025) ^a
R^2	0.003	0.966	0.980	0.003	0.968	0.981
Method	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	Firm FE	Firm FE	FD	Firm FE	Firm FE
Lagged Productivity	No	No	Yes	No	No	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	104,024	136,105	104,024	104,024	136,105	104,024

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. Firm TFP is defined as the residual estimated from the production function from [Caliendo et al. \(2015a\)](#) and using the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#). Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers. Columns (1) and (4) estimate the model using first differences. Columns (2) and (5) estimate the model using firm fixed effects. Columns (3) and (6) estimate the model using firm fixed effects and using lagged firm productivity as an additional control. In columns (2),(4),(5) and (6) the sample is restricted to firms that are observed for at least 2 years. In specifications with firm fixed effects, the demographic, income and wage controls are interacted with a linear time trend.

Table G10: Value-Added per worker & TFP Panel Regression Results with Different Lags - Employment Areas

	VA Worker (1)	VA Worker (2)	VA Worker (3)	LP TFP (4)	LP TFP (5)	LP TFP (6)	WD TFP (7)	WD TFP (8)	WD TFP (9)
Model 1: Without Org									
log density	-0.030 (0.072)	0.018 (0.061)	0.017 (0.034)	0.071 (0.060)	0.083 (0.052)	0.109 (0.043) ^b	0.075 (0.060)	0.086 (0.052) ^c	0.113 (0.043) ^b
Model 2: With Org									
log density	-0.022 (0.073)	0.023 (0.060)	0.019 (0.034)	0.072 (0.061)	0.082 (0.052)	0.108 (0.043) ^b	0.075 (0.061)	0.085 (0.052)	0.111 (0.043) ^b
organization	-0.048 (0.004) ^a	-0.031 (0.004) ^a	-0.024 (0.005) ^a	-0.007 (0.003) ^c	0.007 (0.003) ^b	0.023 (0.004) ^a	-0.005 (0.003)	0.009 (0.003) ^a	0.025 (0.004) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	FD	FD	FD	FD	FD	FD	FD	FD
Lag Difference Δ	1 year	2 year	3 year	1 year	2 year	3 year	1 year	2 year	3 year
R^2 : Model 1	0.001	0.003	0.002	0.003	0.004	0.005	0.003	0.004	0.005
R^2 : Model 2	0.007	0.006	0.004	0.003	0.004	0.006	0.003	0.004	0.007
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	21,111	21,111	21,111	19,694	19,694	19,694	19,694	19,694	19,694

Notes: *a, b, c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. The sample is restricted to firms observed in the year 2003 and 2006 and for at least 3 periods. In columns (1)-(3) productivity is value-added per worker. In columns (4)-(6) productivity is TFP estimated using a Cobb-Douglas production function and the approach proposed by [Levinsohn and Petrin \(2003\)](#), with labor measured using the number of workers. In columns (7)-(9) productivity is TFP estimated using a Cobb-Douglas production function and the approach proposed by [Wooldridge \(2009\)](#), with labor measured using the number of workers. Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

Table G11: CMORH TFP + Org Panel Regression Results with Different Lags - Employment Areas

	CMORH LP TFP + Org (1)	CMORH LP TFP + Org (2)	CMORH LP TFP + Org (3)	CMORH WD TFP + Org (4)	CMORH WD TFP + Org (5)	CMORH WD TFP + Org (6)
Model 1: Without Org						
log density	0.046 (0.049)	0.081 (0.047) ^c	0.087 (0.042) ^b	0.040 (0.048)	0.082 (0.046) ^c	0.084 (0.042) ^b
Model 2: With Org						
log density	0.049 (0.050)	0.083 (0.047) ^c	0.088 (0.042) ^b	0.043 (0.048)	0.084 (0.046) ^c	0.085 (0.042) ^b
organization	-0.021 (0.004) ^a	-0.013 (0.003) ^a	-0.009 (0.003) ^b	-0.023 (0.004) ^a	-0.016 (0.003) ^a	-0.014 (0.003) ^a
Method	OLS	OLS	OLS	OLS	OLS	OLS
Method	FD	FD	FD	FD	FD	FD
Lag Difference Δ	1 year	2 year	3 year	1 year	2 year	3 year
R^2 : Model 1	0.003	0.004	0.005	0.002	0.004	0.004
R^2 : Model 2	0.005	0.005	0.005	0.005	0.005	0.005
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	19,694	19,694	19,694	19,694	19,694	19,694

Notes: *a,b,c*: significant at the 1%, 5% and 10% level. Clustered standard errors at the local area level in parentheses. Entries display an estimate from a separate regression. The sample is restricted to firms observed in the year 2003 and 2006 and for at least 3 periods. Firm TFP is defined as the residual along with the effect of organization, both of which are estimated from the production function from [Caliendo et al. \(2015a\)](#) and using the approaches proposed by [Levinsohn and Petrin \(2003\)](#) and [Wooldridge \(2009\)](#). Industry fixed effects are at the 4-digit NAF Rev 1. level. Local controls include the cost of a unit of labor, the median annual salary of individuals residing in an area, the share of the local population between the ages of 25 and 59, the share of the local population born outside France, and the share of unemployed active workers.

Supplementary Appendix - For Online Publication

H Model Simulations

This section illustrates the outcomes from two simulations of the model. Section H.1 first provides an overview of the simulation parameters, describes the equilibrium values, and illustrates differences in the distribution of firm organization. Section H.2 illustrates the differences in the outcomes of firms and Section H.3 the differences in the various measures of firm productivity. Finally, Section H.4 illustrates the differences in the knowledge and income across both simulations.

H.1 Simulation Parameters & Distribution of Organizations

Table H1 lists the complete set of parameters and their simulation values. There are two economies of different sizes, represented in Models 1 and 2. In Model 1, the size of the market is $N = 500$, while in Model 2 market size is equal to: $N = 1000$. In both economies, demand is drawn from a Pareto distribution with coefficient $k = 3.95$ and support $[1, \infty]$. As a result, the cumulative distribution of demand draws is equal to: $G(\alpha) = 1 - \alpha^{-3.95}$.⁶¹ The remaining parameters are identical in both models.

Table H2 presents equilibrium values. An increase in market size raises the demand draw of the marginal firm, α_D , and the aggregate term, $\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p})$. These results are consistent with Proposition 5. Table H2 also reports the demands draws at which entrepreneurs are indifferent between two organizational structures, $\alpha_{L,L+1}$. It is not always the case that the cutoffs, $\alpha_{L,L+1}$, decrease in larger markets. However, consistent with Proposition 6, the distance between α_D and $\alpha_{L,L+1}$ decreases with N .

Figure H.1 illustrates the cumulative distribution of firm organization in both economies. The distribution from Model 1 is represented with dark gray bars, while the distribution from Model 2 is represented using light gray bars. Throughout this section, I maintain the same color scheme.

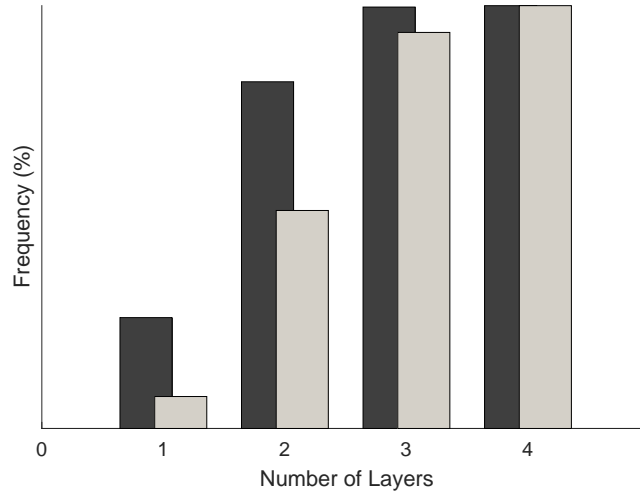
As illustrated in Figure H.1 there is a ranking of the cumulative distributions of firm organization. Intuitively, a larger market affects the distribution of firm organization through two channels. First, it induces tougher selection and raises the demand cutoff, α_D . And second, it lowers markups, and as a result induces a mass of firms to re-organize production in favor of more layers. The Pareto distribution always satisfies the non-decreasing hazard rate property. It thus follows from Proposition 7 that the cumulative distribution of firm organization in Model 2 first-order stochastically dominates the distribution in Model 1. Applying the Mann-Whitney test to the simulated data confirms the result. It rejects the hypothesis that both distributions are equal at the one percent level of statistical significance (the p-value of the test statistic is 0.000). The Mann-Whitney test indicates that a firm chosen at random from the larger market is 69.4% more likely to have a greater number of layers than a randomly selected firm from the smaller market

H.2 Firm-Level Simulation Results

Figure H.2 illustrates the optimal firm allocations in both economies. Panel (H.2a) focuses on firm quantity. It demonstrates that there is a heterogeneous response of $q(\alpha)$ to a rise in N . An increase in market size has two opposing effects on firms' demand schedule. First, a direct effect: a rise in N rotates the inverse demand curve away from the quantity-axis, increasing firm demand. And second, an indirect effect: a rise in N increases the aggregate term, $\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p})$, leading to a

⁶¹The parameter k is relatively large for computational tractability.

Figure H.1: Cumulative Distribution of Organizations.



Notes: The figure illustrates the cumulative distribution of the total number of layers across two markets of different size. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

Table H1: Parameter Values

	N	A	h	λ	c	γ	η	k	α_m	f_E
Model 1	500	10	0.42	28	14	5	3	3.95	1	1.75
Model 2	1000	10	0.42	28	14	5	3	3.95	1	1.75

Notes: Parameters used in simulations of Models 1 and 2.

Table H2: Equilibrium Values

	α_D	q_D	$\frac{\eta M}{\gamma + \eta M}(\bar{\alpha} - \bar{p})$	M	$\bar{\alpha}$	\bar{p}	$\alpha_{0,1}$	$\alpha_{1,2}$	$\alpha_{2,3}$
Model 1	2.368	6.893	2.069	3.248	3.145	0.013	2.557	3.759	9.432
Model 2	3.294	7.361	3.031	3.375	4.453	0.005	3.379	3.971	6.804

Notes: Equilibrium Values from simulations of Models 1 and 2.

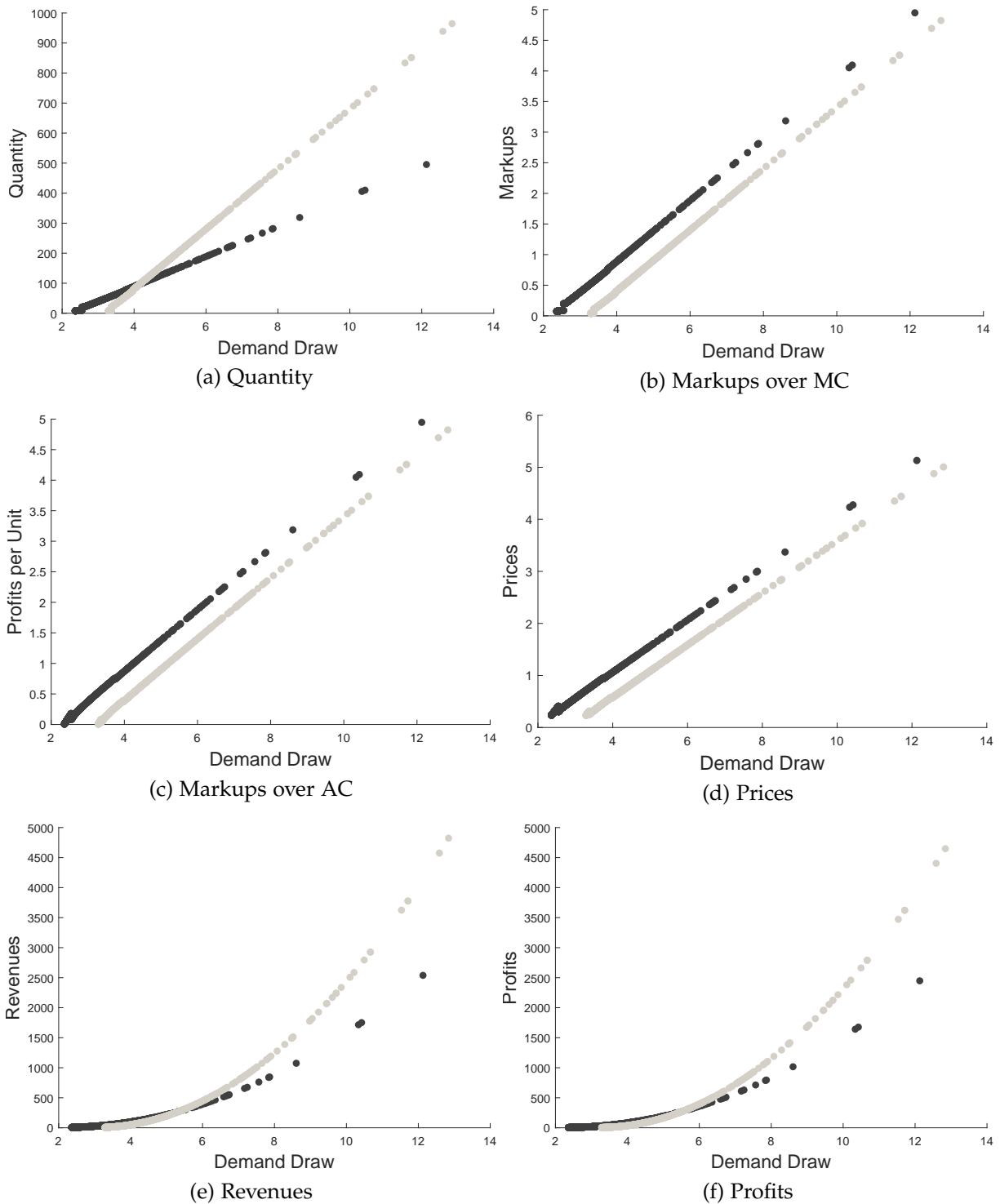
downward shift in firm demand. The direct effect dominates in firms with a sufficiently high α , who increase production in the larger market. The indirect effect, however, dominates in firms with low α , who produce less in the larger market.

Panels (H.2b) and (H.2c) focus on markups. The figures show that across all firms both $\mu^{MC}(\alpha)$ and $\mu^{AC}(\alpha)$ decrease with N . Note that markups also decrease in firms that change their organizational structure. Furthermore, lower markups over average costs imply that profit per unit declines with N . In turn, this implies that the equilibrium mapping from q to π changes as well: for any level of output, profits are lower in the larger market.

Panel (H.2d) shows that firm price decreases with N . There are two factors that determine the change in $p(\alpha)$. First, $q(\alpha)$ changes with N , which affects marginal costs.⁶² Second, since in the

⁶²Firms that do not change their organization and produce more output increase their marginal costs, while firms that produce less output and do not change their organizational structure, decrease their marginal costs. The opposite, however, takes place in firms that change their organization. Holding markups constant, an increase in marginal costs induces firms to charge higher prices, while a decrease in marginal costs induces firms to charge lower prices.

Figure H.2: Firm Outcomes across Markets



Notes: Panel (a) illustrates the optimal relationship between firm quantity and the demand draw, across two markets. Panel (b) illustrates the optimal relationship between firm markups over marginal costs and the demand draw, across two markets. Panel (c) illustrates the optimal relationship between firm markups over average costs and the demand draw, across two markets. Panel (d) illustrates the optimal relationship between firm price and the demand draw, across two markets. Panel (e) illustrates the optimal relationship between firm revenues and the demand draw, across two markets. Panel (f) illustrates the optimal relationship between firm profits and the demand draw, across two markets. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). Demand Draw denotes the demand parameter, α . The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

larger market demand is more elastic, $\mu^{MC}(\alpha)$ is lower, leading to a decrease in $p(\alpha)$. The change in markups dominates any change in marginal costs, and as a result, prices decline with N .

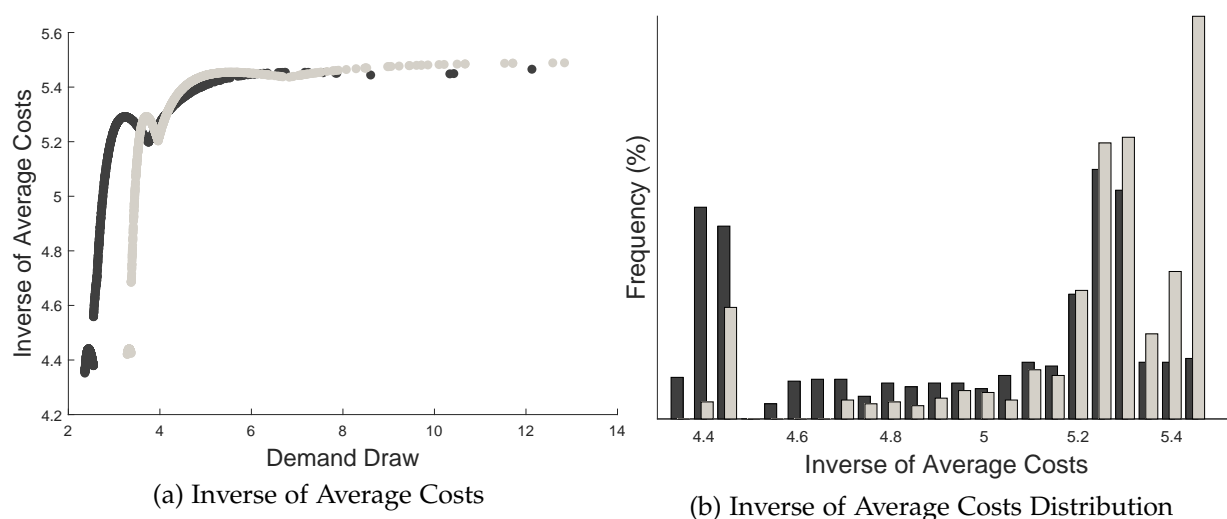
Panel (H.2f) further shows that firm profits may decrease or increase with N . There are two factors that again determine how $\pi(\alpha)$ changes with N . First, as demonstrated in panel (H.2c), $\mu^{AC}(\alpha)$ declines with N , decreasing profits. Second, as shown in panel (H.2a), $q(\alpha)$ changes as well. Holding markups constant, an increase in output raises profits. Consequently, firms that reduce their quantity in the larger market earn lower profits, because both effects work in the same direction. In contrast, in firms that increase their output, both effects work in the opposite direction. Beyond some quantity, the second effect dominates and $\pi(\alpha)$ increases in the larger market.

H.3 Productivity Simulation Results

I turn to firm productivity. As in [Caliendo and Rossi-Hansberg \(2012\)](#) I report results from several measures of firm productivity. The first two measures are based on quantity, while the last two measures are based on revenues.

Quantity-Based Productivity Measures

Figure H.3: The Impact of Market Size on Inverse of Average Costs.



Notes: Panel (a) illustrates the optimal relationship between inverse of average costs and the heterogeneous demand draw of a firm, across two markets. Panel (b) illustrates the distribution of inverse of average costs across two markets. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). Demand Draw denotes the demand parameter, α . The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

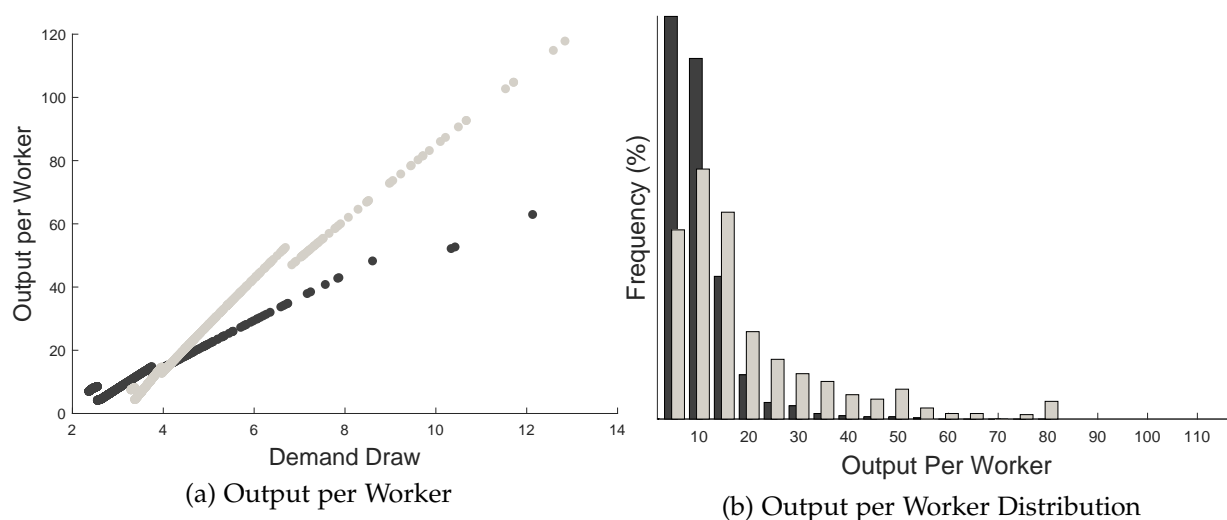
Figure H.3 illustrates simulation results from the first measure of quantity-based productivity, the inverse of average costs, which is equal to: $\phi_1(\alpha) = \frac{q(\alpha)}{C(q(\alpha))} = \frac{1}{AV(q(\alpha))}$.⁶³ Figure H.3a illustrates

⁶³A previous section characterized the cost function of firms. To summarize, the important points are the following. First, with the exception at the MES, firms' marginal costs are not equal to their average costs. Second, average cost are neither constant nor a monotonic function of quantity. This implies that $\phi_1(\alpha)$ will also be neither constant nor a monotonic function of quantity. Third, the minimum average cost is decreasing with the number of layers, and the level of output that attains the minimum average cost is increasing with the number of layers. This implies that firms producing with more layers can attain a greater productivity. And fourth, because the quantity produced by firms and their organization depend on the size of the market, $\phi_1(\alpha)$ will also depend on the size of the market. Together these

$\phi_1(\alpha)$ at the firm-level. The main takeaway from the figure is that there is a heterogeneous response of $\phi_1(\alpha)$ to a rise in N . The productivity of firms with relatively low α decreases in the larger market, while the productivity of firms with medium and high demand draws rises. Note that this is different from the findings of [Caliendo and Rossi-Hansberg \(2012\)](#), where in a closed-economy, an increase in market size only raises wages and the mass of firms, but does not affect output. Consequently in their framework, in a closed-economy, the number of layers and the productivity of firms do not vary with the size of markets.

Figure H.3a presents the distribution of $\phi_1(\alpha)$ in both economies. It shows different shares of small, medium and high productivity firms across both markets. The larger market has a smaller fraction of low and medium productivity firms, and a greater mass of high productivity firms. In Model 2, $\phi_1(\alpha)$ has a mean of 5.21 and a standard deviation of 0.275. In comparison, in Model 1 the mean of $\phi_1(\alpha)$ is equal to 4.94 and its standard deviation is 0.376.

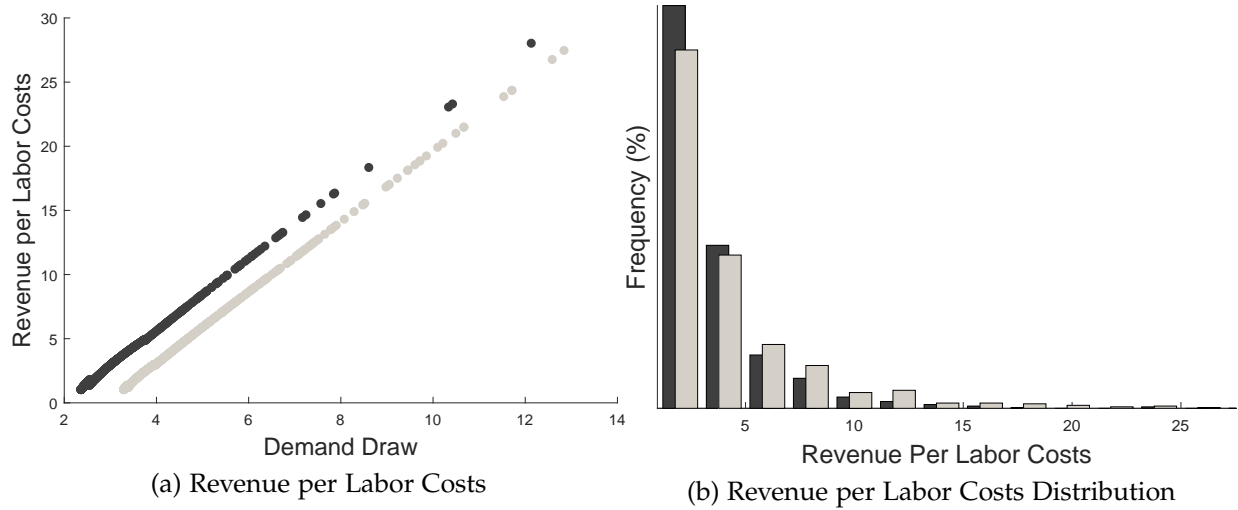
Figure H.4: The Impact of Market Size on Output per Worker.



Notes: Panel (a) illustrates the optimal relationship between output per worker and the heterogeneous demand draw of a firm, across two markets. Panel (b) illustrates the distribution of output per worker across two markets. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). Demand Draw denotes the demand parameter, α . The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

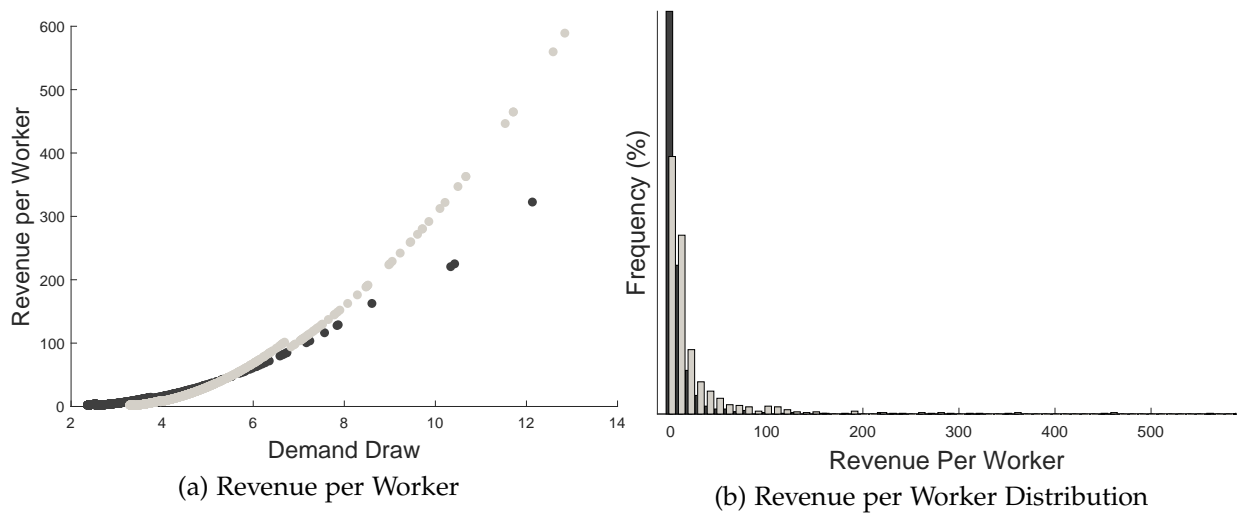
Figure H.4 presents simulation results from output per worker, or labor productivity, which is equal to: $\phi_2(\alpha) = \frac{q(\alpha)}{\sum_{j=1}^L n_L^j(\alpha)}$. Figure H.4a illustrates $\phi_2(\alpha)$ at the firm-level. It again shows a heterogeneous response of firm productivity to a rise in N . Labor productivity decreases in firms with very low demand draws, and rises for the remaining firms. The firms that experience an increase $\phi_2(\alpha)$, are the ones that increased their output in the larger market. Figure H.4b presents the distribution of $\phi_2(\alpha)$ in both markets. It shows that the larger market has a smaller share of low productivity firms, and a greater share of medium and high productivity firms. In Model 2, $\phi_2(\alpha)$ has a mean of 19.30 and a standard deviation of 18.85. In contrast, in Model 1 the mean of $\phi_2(\alpha)$ is equal to 9.94 and its standard deviation is 6.45.

Figure H.5: The Impact of Market Size on Revenue per Labor Costs.



Notes: Panel (a) illustrates the optimal relationship between revenue per labor costs and the heterogeneous demand draw of a firm, across two markets. Panel (b) illustrates the distribution of revenue per labor costs across two markets. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). Demand Draw denotes the demand parameter, α . The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

Figure H.6: The Impact of Market Size on Revenue per Worker.



Notes: Panel (a) illustrates the optimal relationship between revenue per worker and the heterogeneous demand draw of a firm, across two markets. Panel (b) illustrates the distribution of revenue per worker across two markets. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). Demand Draw denotes the demand parameter, α . The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

Revenue-Based Productivity Measures

Figure H.5 presents simulation results from $\psi_1(\alpha)$, which measures firm revenues per labor costs and is equal to: $\psi_1(\alpha) = \frac{r(\alpha)}{C(\alpha)}$. Figure H.5a illustrates $\psi_1(\alpha)$ at the firm-level. It shows that $\psi_1(\alpha)$ points suggest that the distribution of productivity will be different across locations.

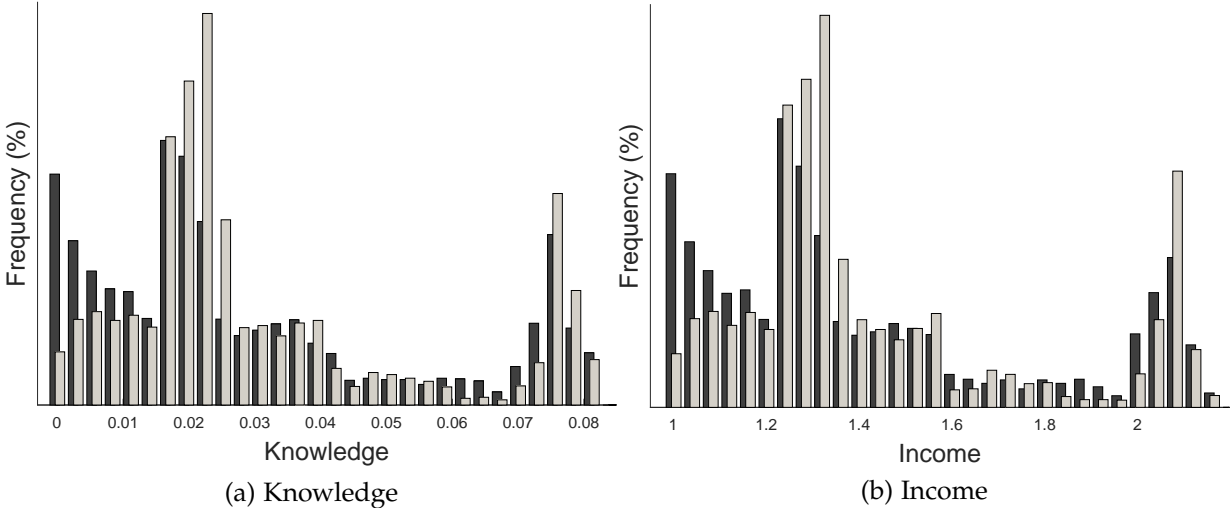
declines in the larger market. This is expected, since $p(\alpha)$ decreases with N . Figure H.5b further shows that the larger market has a smaller fraction of low and medium productivity firms, and a greater mass of high productivity firms. These differences are due to the tougher selection in the larger market. Furthermore, in Model 2, $\psi_1(\alpha)$ has a mean of 4.22 and a standard deviation of 4.28. In Model 1, the mean of $\psi_1(\alpha)$ is equal to 3.32 and its standard deviation is 2.77.

Figure H.6 presents simulation results from revenue per worker, or revenue-labor productivity, which is equal to: $\psi_2(\alpha) = \frac{r(\alpha)}{\sum_{j=1}^L n_j^L(\alpha)}$. Naturally, the outcomes of $\psi_2(\alpha)$ with respect to N resemble the outcomes of labor productivity, as $\psi_2(\alpha) = p(\alpha)\phi_2(\alpha)$. Figure H.6a illustrates $\psi_2(\alpha)$ at the firm-level. The main takeaway from the figure is that there is a heterogeneous response of $\psi_2(\alpha)$ with respect to N . The productivity of firms with relatively low demand draws decreases in the larger market, while the productivity of firms with medium and high draws increases with N . Figure H.6b presents the distribution of firm productivity in both markets. It shows that the larger market has a lower share of low productivity firms, and a greater mass of medium and high productivity firms. In Model 2, $\psi_2(\alpha)$ has a mean of 29.79 and a standard deviation of 128.74. In Model 1, the mean of $\psi_2(\alpha)$ is equal to 9.54 and its standard deviation is 20.83.

H.4 Knowledge and Income Simulation Results

I now turn to incomes and knowledge and examine how their distributions are different across Models 1 and 2.

Figure H.7: Distributions of Knowledge & Income.



Notes: Panel (a) illustrates the distribution of knowledge across two markets. Panel (b) illustrates the distribution of income across two markets. Light gray represents the larger market ($N = 1000$), while dark gray represents the smaller market ($N = 500$). The parameters used to simulate both models are reported in Table H1 in the Supplementary Appendix H.

Table H3: Qualitative Comparisons on the Distributions of Income

	Model 1	Model 2
Mean	1.4095	1.4421
Standard Deviation	0.3480	0.3264

Notes: Comparisons of distribution of income from Model 1 ($N = 500$) and Model 2 ($N = 1000$).

Knowledge

Panel (a) in Figure H.7 illustrates the distribution of knowledge in both economies. The larger market, Model 2, contains a greater share of agents with intermediate levels of knowledge, and a smaller share of agents with low levels of knowledge. This effect is simply due to a greater mass of firms organizing with more layers in the larger market. Because in the larger market firms produce with more layers, the knowledge of existing workers decreases. At the same time, because firms employ more intermediate managers, there are more agents with intermediate levels of knowledge in the economy. The second effect dominates and the mass of agents with low levels of knowledge is reduced in the larger market.

Income

Panel (b) presents the distribution of income in both economies. The distribution of income closely resembles the distribution of knowledge, because an agent's income is equal to: $(cz + 1)$. Table H3 reports the mean and standard deviation of the distribution of income from both models. In Table H3 the distribution of income has a higher mean but a lower standard deviation in the larger market relative to the smaller market. A higher mean is consistent with empirical studies that examine how wages differ across locations. A conclusion emerging from these studies is that workers earn higher wages in denser markets (for example, see [Combes and Gobillon \(2015\)](#)). The numerical simulations suggest that the model is able to qualitatively account for this fact. The simulations, however, are unable to account for the fact that wage inequality is greater in larger markets (for example, see [Combes and Gobillon \(2015\)](#)). This difference may be due to the fact that the simulations attribute a relatively large weight to firms with low α (i.e. the Pareto shape parameter, k , is relatively large).