The Anatomy of French Production Hierarchies*

Lorenzo Caliendo
Yale University

Ferdinando Monte
Johns Hopkins University

Esteban Rossi-Hansberg
Princeton University

February 5, 2014

Abstract

We use a comprehensive dataset of French manufacturing firms to study their internal organization. We first divide the employees of each firm into 'layers' using occupational categories. Layers are hierarchical in that the typical worker in a higher layer earns more, and the typical firm occupies less of them. In addition, the probability of adding/dropping a layer is very positively/negatively correlated with value added. We then explore the changes in the wages and number of employees that accompany expansions in layers, or output. The empirical results indicate that reorganization, through changes in layers, is key to understand how firms expand and contract. For example, we find that firms that expand substantially add layers and pay *lower* average wages, in part by hiring less experienced employees, in all pre-existing layers. In contrast, firms that expand little and do not reorganize pay *higher* average wages in all pre-existing layers, partly by hiring more educated employees.

*We thank Ariel Burstein, Pierangelo De Pace, Luis Garicano, Bob Gibbons, Gene Grossman, Maria Guadalupe, Gordon Hanson, Oleg Itskhoki, Wolfgang Keller, Sam Kortum, Francis Kramarz, Claire Lelarge, Marc Melitz, Derek Neal, Steve Redding, Peter Schott, Chris Sims, John van Reenen, and Julie Wulf for useful conversations and comments. The computations in this paper were done at a secure data center located at CREST, Paris. Caliendo acknowledges the support from the Yale MacMillan Center. Caliendo: lorenzo.caliendo@yale.edu, Monte: ferdinando.monte@jhu.edu, and Rossi-Hansberg: erossi@princeton.edu.

1 Introduction

Labor is not a homogeneous input. Employees are distinct in their levels of skill, knowledge, experience, and a vast variety of other dimensions. So an important decision made by firms is to determine not only the number but also the characteristics of their employees, as well as the role that each of them plays in the firm. We refer to these decisions as the organization of a firm. In this paper we aim to describe empirically the organization of firms and how this organization is related to other firm characteristics. We are particularly interested in understanding if firms actively manage their organization –and therefore the number and characteristics of their employees– and how they do this. Understanding these decisions is important in order to understand the behavior of firms and therefore that of the aggregate economy. As far as we know, this is the first empirical study of the internal organization of firms that uses a comprehensive dataset with a large number of firms.

We use a sample of the large majority of French manufacturing firms during the period 2002-2007.[1] To organize the data in a practical and meaningful way, we first introduce the concept of a 'layer' of employees. The concept of a layer is adopted from the theory of management hierarchies proposed initially by Rosen (1982) and Garicano (2000) and used in the context of heterogeneous firms in Caliendo and Rossi-Hansberg (2012). In this theory a layer is a group of employees, with similar characteristics summarized in their knowledge, who perform similar tasks within the organization. Conceptually, these layers are hierarchical in the sense that higher layers of management are smaller and include more knowledgeable employees who have as subordinates employees in lower layers. Dividing the employees in real firms into layers requires some mapping between these concepts and the data.

The first part of the map involves using wages as a one-dimensional measure of the marketable characteristics of employees, that is, to view wages as a measure of the 'knowledge' of workers. In this vein, if one individual is more knowledgeable than another (in terms of practical knowledge used in production) he will obtain a higher wage. Later in the paper we discuss the particular way in which this average 'knowledge' in a layer is modified and how it is created using a combination of education and experience.

The second part of the map is to group workers into layers. To do so we use information on occupational characteristics. Fortunately, the French data we use provide hierarchical occupational categories. The top occupation includes owners who receive a wage. The next one below includes senior staff and top management positions; the next one employees at a supervisory level. The lowest two occupations include clerks and blue collar workers. We document that they earn similar wages and so we pull them together into one layer. This gives us a maximum of four hierarchical layers of employees. Of course, many firms (in fact, most of them) do not employ agents in all of these layers, something we exploit extensively in this paper.

We then investigate whether this division of the employees in a firm into layers is an economically meaningful classification. We cut the data in a variety of ways that indicate that it is. Firms with

[1] A detailed description of the data is relegated to Online Appendix B.

more layers are larger in terms of value added and employment, and, in general, pay higher wages. Around 50% of the variation in wages within firms is variation across layers. A large majority of firms have adjacent layers that start at the bottom. When they add or drop a layer, it is mostly an adjacent layer and they add or drop only one. Layers within firms are also different from each other. Lower layers are larger in the number of hours of work and employ agents who earn lower wages. Thus, most firms are hierarchical in their layers both in terms of wages and in terms of time employed at each layer. In addition, the probability of adding a layer is increasing in value added and the probability of dropping one is decreasing in value added. Finally, the firms that we observe adding/dropping a layer in a given period tend to grow/shrink in the previous periods, indicating that they progressively get closer to a size treshold that triggers the change. All these facts are very significant in the data and robust to accounting for industry and time fixed effects. So we conclude that the layers we identify using occupations are not arbitrary names but have an economic meaning in terms of the characteristics of the employees they group and the tasks they perform.

The next step is to understand how firms change their organization —the knowledge and number of employees at each layer— as they grow. It is useful to go back to the theory in order to guide our exploration. We rely on Caliendo and Rossi-Hansberg (2012, from now on CRH) as our guide, since their general equilibrium theory of production hierarchies allows for firm heterogeneity, which is important in the data.[2] In this theory firms organize production to economize on their use of knowledge: a costly input. Production requires time and knowledge. Workers in layer one work on the production floor. To produce, they need to solve the problems they face in production. Their knowledge allows them to solve some, but not all, the problems they face. If they can solve a problem output is immediately realized. Otherwise, they can ask agents in higher layers how to solve them. Since they do not know anything about these problems, they first ask the managers in layer two. These managers spend their time communicating with the workers and understanding their problems. They in turn solve some of them and pass the rest to the third-layer and so on. The problem of the firm is to decide the number of hours of work and the level of knowledge of employees in each layer and how many layers to have in the firm. The number of hours of the top manager is fixed and common across firms.

A firm with higher demand, or higher exogenous idiosyncratic productivity, optimally decides to have more layers. Its larger scale allows it to economize on the total cost of knowledge by having many layers of management with very knowledgeable managers at the top, but much less knowledgeable employees in the bottom layers. The theory in CRH implies that some firms that expand value added will add layers. However, some others might expand without adding layers since the expansion is not large enough to make the added cost of an extra layer (namely, the wage of the new top manager) worth paying. The trade-off is simple: lower 'marginal' cost from having less knowledgeable employees in the existing layers (because the new top manager can solve the less

[2] Throughout, we refer mostly to the theory of CRH, although some of the arguments we advance can be traced back to Garicano (2000) or Garicano and Rossi-Hansberg (2006).

frequent questions) versus higher 'fixed' cost from having to pay an extra, and large, wage of the top manager. So it is worth paying the 'fixed' cost only if the cumulative expansion since the last change is large enough. Thus, in the theory, firms that expand by adding layers reduce wages and increase the number of hours at all layers, while firms that expand but do not add layers increase both hours and wages at all layers.

This theory has implications on firm and layer-level outcomes. That is, it has implications for the number of workers and their average knowledge, and therefore the average wage, for each layer of employees. We go to the data guided by the implications of the theory. We look at firms that expand and add layers and firms that contract and drop layers. In particular, we estimate changes in log average wages by layer and changes in log average hours of work by layer, normalized by hours in the top layer, for firms that add or drop layers from one year to the next. We then look at firms that expand but do not add or drop layers.

In the data, firms that expand substantially in a given year tend to reorganize by changing layers. They contribute almost 40% of the total change in value added in the manufacturing sector. We find that wages in firms that expand and add layers behave differently than firms that expand but do not reorganize. If firms expand by adding layers, average wages in preexisting layers fall, while if firms expand without reorganizing, average wages in all layers rise. All these results reverse when we focus on firms that contract and either drop layers or do not reorganize. These results hold for each layer in firms with any number of layers. We do not find any instance in which they are contradicted by the data. Importantly, our findings are not simply the result of regrouping workers with the same wages across layers. We document that the distribution of wages in preexisting layers shifts down for all percentiles when firms grow by adding layers, while it shifts up for all percentiles when firms grow without reorganizing.

The results above document how firms affect the average 'knowledge' in each layer when they grow or decline. They do not, however, explain how firms manage to modify the characteristics of their employees to achieve these average changes.[3] The data available to explore these questions is not as well suited as the one we used for the results above; nevertheless, in Section 5 we ask two questions pertinent to understand how firms adjust average knowledge in a layer.

First, do firms affect the average wage in a layer by changing the composition of the workers in a layer or by changing the knowledge and wages of current employees? We find that when firms grow by adding a layer, average wages in a layer are reduced by changing the composition of employees in a layer and not by reducing individual wages. The extensive margin is also dominant when we look at firms that drop layers.

Second, if knowledge can be created with either formal education or experience, which of the two

[3]The frictions in the French labor market, and therefore the implications of firm reorganization on particular individuals, are certainly worth exploring but do not constitute the main focus of our paper. Instead, our aim is to understand if and how firms actively manage and reorganize their labor force. Our focus on firms decisions and their performance parallels the work of Bloom and van Reenen (2007), Bloom, et al. (2012), and Bloom, et al. (2011) all of which study the effect of management, not organization, on firm characteristics. Our approach also relates to Baker, et al. (1994), and Baker and Holmstrom (1995) who study how a particular firm organizes its internal labor force.

is affected when firms decide to change the average level of knowledge in a layer? Using estimated measures of labor market experience and years of formal education for each employee[4], we show that firms use these two forms of acquiring knowledge in systematic but distinct ways. Firms that grow without adding layers increase knowledge by hiring workers with more formal education, particularly at the bottom of the hierarchy. In contrast, firms that expand by adding layers tend to reduce knowledge by hiring less experience workers.[5] The behavior we uncover is consistent with the view that formal years of education provide the knowledge to solve the most common problems in an organization, the tasks handled at the lower layers of the hierarchy. In contrast, labor market experience provides the knowledge required to solve more infrequent problems, the tasks handled at the higher level of the hierarchy.

Several other papers have studied the internal organization of firms using small samples of producers (a few hundred). For instance, Caroli and van Reenen (2001) use surveys from England and France to find that the wage bill share of different skill levels change as firms delayer. These results support our finding that delayering is associated with systematic occupational shifts.[6] Garicano and Hubbard (2007) study the role of hierarchies as a means of organizing production in law firms. They use confidential data from law offices from the 1992 Census of Services. They find that as market size increases, the ratio of associates to partners increases. We document a similar finding for manufacturing firms in France. We find that as firms expand, either by adding layers or not, the number of hours worked by lower-level employees relative to higher-level ones expands. Rajan and Wulf (2006), using a sample of 300 large U.S. firms for the period 1986 to 1998, analyze how hierarchies of top-level managers have changed over time. The study shows that the CEOs' span of control has increased, while the number of layers between division heads and CEOs has gone down during the sample period. Thus, they find evidence that such hierarchies have 'flattened' over time and have decentralized their decision making. Using a large comprehensive dataset for France, we document that firms have also become 'flatter' during the period 2002 to 2007.

The objective of our paper is to study the organization of firms and how this organization changes as firms grow or decline. The theory in CRH emphasizes the notion of reorganization, a concept that we find helpful when looking at the data. In fact, conditioning on whether a firm reorganizes or not, allows us to uncover what we view as robust characteristics of firm behavior. For example, that they reduce average wages in pre-existing layers when they expand and reorganize, but they increase them when they expand without adding layers. Independently of the theory that provides the true explanation for these facts, they show that the concept of reorganization through changes in layers, measured by changes in the occupations employed in a firm, is useful to

[4] Direct measures of formal education and experience are not directly available in a way that can be matched to the firm data we are using. So we need to estimate worker level education and experience using a methodology that we describe in detail in Section 5.

[5] We also find that the reverse patterns hold when we look at firms that contract with and without changing layers.

[6] Caroli and van Reenen (2001) has the advantage of using a measure of delayering directly reported by managers rather than indirectly inferred from the occupational structure, like ours. The advantage of our approach is that it relies less on the subjective views of managers and more on their observed actions. Our measure also allows us to use the universe of manufacturing firms instead of specialized surveys.

characterize the behavior of firms.

The rest of the paper is organized as follows. The next section describes in more detail the essential features of the theory in CRH that guides our empirical exploration. Section 3 describes the data and our construction of layers and shows the basic characteristics of firms and layers. Section 4 presents our findings on organizational changes as a result of changes in layers and expansions in value added. Section 5 discusses how firms change average wages in a layer and Section 6 concludes. Online Appendix A presents a variety of robustness checks and extensions of the results in the main text. Online Appendix B describes in detail the data, its manipulation, and our empirical methodology.

2 A Theory of Organization with Heterogeneous Firms

In this section we discuss briefly the framework in CRH. Given that the purpose of the current paper is to describe and understand the data, we present the theory in its simplest form and do not discuss all the details fully. The interested reader is directed to CRH for the more technical discussions and all proofs of the results.

We consider an economy with $\tilde{N}$ identical agents with preferences that lead to a demand for variety α given by $x(p, \alpha; R, P)$ where p denotes the price, R revenue, and P the price index. We assume that agents like varieties with higher α better, so $\partial x(p, \alpha; R, P)/\partial\alpha > 0$ and, as usual, $\partial x(p, \alpha; R, P)/\partial p < 0$. Agents are endowed with one unit of time that they supply inelastically and obtain an equilibrium wage $\bar{w}$ for their unit of time. Agents acquire knowledge in order to solve the problems they encounter during production. Learning how to solve problems in an interval of knowledge of length z costs $\bar{w}cz$ (c teachers per unit of knowledge at cost $\bar{w}$ per teacher). Since the cost of knowledge is linear, agents receive it back as compensation for their work. Hence, the total wage of an employee with knowledge z is given by $w = \bar{w}[cz + 1]$.

An entrepreneur pays a fixed entry cost f^E in units of labor to design her product. After doing so, she obtains a demand draw α from a known distribution $G(\alpha)$. The draw α determines the level of demand of the firm. If the entrepreneur decides to produce she pays a fixed cost f in units of labor. Production requires labor and knowledge. Agents employed in a firm act as production workers (layer $\ell = 1$) or managers (layers $\ell \geq 2$). We denote by n_L^ℓ, z_L^ℓ, and w_L^ℓ, the number, knowledge, and total wage of employees at layer $\ell = 1, 2, 3...$ of an organization with L layers.[7] Production workers use their unit of time to generate a production possibility that can yield one unit of output. For output to be realized the worker needs to solve a problem drawn from a distribution $F(z)$ with $F''(z) < 0$. Production workers learn how to solve the most frequent problems. The ones in the interval $\left[0, z_L^1\right]$. If the problem they face falls in $\left[0, z_L^1\right]$, production is realized; otherwise, they can ask a manager one layer above how to solve the problem. Managers spend h units of their time on each problem that gets to them. A manager at layer $\ell = 2$ tries to solve the problems

[7] Note that we label the lowest layer of the organization, the layer of production workers, as layer 1 while in Garicano (2000), Garicano and Rossi-Hansberg (2006), and Caliendo and Rossi-Hansberg (2012), the lowest layer is denoted by layer 0.

workers could not solve. Hence, they learn how to solve problems in $\left[z_L^1, z_L^1 + z_L^2\right]$. In general, the firm needs $n_L^\ell = h n_L^1 (1 - F(Z_L^{\ell-1}))$ managers of layer ℓ, where $Z_L^\ell = \sum_{l=1}^{\ell} z_L^l$.[8]

We characterize the problem using the variable cost function. Let $C(q;w)$ denote the minimum variable cost of producing q units, and $C_L(q;w)$ the same cost if we restrict the organization to producing with L layers of management. Then,

$$C(q;w) = \min_{L\geq 1} \{C_L(q;w)\} = \min_{L\geq 1,\ \{n_L^\ell, z_L^\ell\}_{l=1}^L \geq 0} \sum\nolimits_{\ell=1}^{L} n_L^\ell w_L^\ell \tag{1}$$

subject to

$$\begin{aligned} q &\leq F(Z_L^L) n_L^1, & (2)\\ w_L^\ell &= \bar{w}[c z_L^\ell + 1] \text{ for all } \ell \leq L, & (3)\\ n_L^\ell &= h n_L^1 [1 - F(Z_L^{\ell-1})] \text{ for } L \geq \ell > 1, & (4)\\ n_L^L &= 1. & (5) \end{aligned}$$

So one entrepreneur, $n_L^L = 1$, chooses the number of layers, L, employees at each layer, n_L^ℓ, and the interval of knowledge that they acquire, z_L^ℓ, subject to the output constraint and the time constraints of employees at each layer. Figure 1 illustrates the resulting average cost function $C(q;w)/q$ as a function of q. It is the lower envelope of the average cost functions restricted to have a given number of layers, $C_L(q;w)/q$. The minimum of these average cost functions decreases with the number of layers and is reached for higher output levels the higher the layer. Each point in that curve is associated with a particular organization. Namely, it is associated with a number of layers, and a number of employees and their knowledge at each layer.

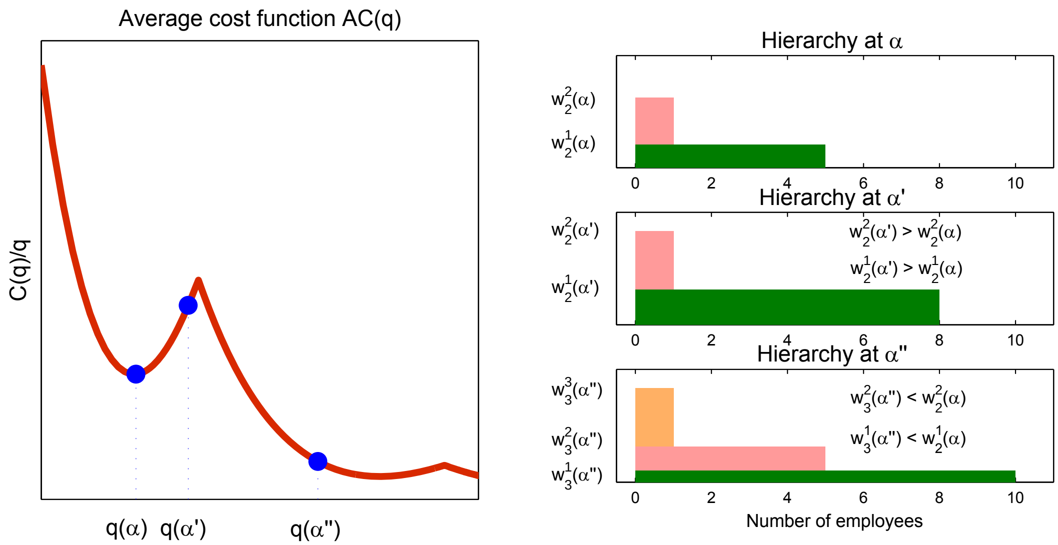
Importantly, as proven in CRH, as firms increase the number of layers by one in order to produce more, the number of agents in each layer increases and the knowledge in all pre-existing layers, and therefore the wage, decreases. In Figure 1 this is illustrated as a change from $q(\alpha)$ to $q(\alpha'')$. The logic is straightforward. Firms add layers to economize on the knowledge of their workers. So when they add a new top layer, they make the new manager deal with the rare problems and make lower level employees know less. The lower knowledge in all pre-existing layers reduces, by equation (4), the span of control of each manager in the organization (namely the number of subordinates per manager). However, the number of employees in all layers still goes up since the span of control of the new top manager is larger than one.

Figure 1 also illustrates how firms grow when they do not add layers (a change from $q(\alpha)$ to $q(\alpha')$ in the figure). In this case we see that the number of workers in all layers increases, as do the knowledge and wages of all workers. This is the only way the firm can expand given the number of layers. Because the span of managers is given by the knowledge of their subordinates, the only way the firm has to increase output is to increase the knowledge of its employees. Since the knowledge of agents at different layers is complementary, the firm does so at all layers. So the implications of

[8] To derive some of the implications of the theory, CRH specify the distribution of problems as an exponential, so $F(z) = 1 - e^{-\lambda z}$.

the model on how firms grow and how the wages and knowledge of their employees vary depends crucially on whether firms add layers or not. Clearly, firms that want to expand substantially change the number of layers, while firms that do not want to expand that much, in general, keep the number of layers fixed. We investigate all of these implications in our empirical analysis below.

Figure 1: The average cost function $C(q;w)/q$ as a function of q



So far we have not said anything about how the quantity produced is determined. To do so we need to turn to the profit maximization and entry decision of the firm. CRH embed the cost function discussed above into a standard Melitz (2003) type framework with heterogeneity in demand. Given that we exploit the general equilibrium of the model only in a limited way, we direct the reader to CRH for details. Here we only state that the model yields an optimal quantity produced, which is increasing as a function of the demand draw α. So a higher demand draw leads to a higher quantity and, as described above, a new organization.

To sum up, the model has the following implications:

1. Firms are hierarchical, $n_L^1 \geq ...n_L^\ell ... \geq n_L^L$ for all L.

2. Layers, L, sales pq, and total number of employees, $\sum_{\ell=1}^{L} n_L^\ell$, *increase* with α.

3. Given L, w_L^ℓ and n_L^ℓ *increase* with α at all ℓ.

4. Given α, w_L^ℓ *decreases* and n_L^ℓ *increases* with an increase in L at all ℓ.

Armed with these implications, and the way of organizing the data dictated by the theory, we now turn to our empirical analysis of the anatomy of French production hierarchies.

3 The Data

We use confidential data collected by the French National Statistical Institute (INSEE) for the period 2002 to 2007. We do not use the data before 2002 because the occupational categories, which we use to determine layers below, changed that year. To construct our unique dataset, we merge two different sources of mandatory reports. First is the BRN dataset, which includes the balance-sheet data of private firms. It includes 553,125 firm-year observations in the manufacturing sector. Second is the DADS dataset, which includes occupation, hours of work, and earning reports of salaried employees. In matching the datasets, we lose 6.9% of the observations and we lose another 11.5% from cleaning the data. The resulting sample covers on average over time 90.3% of total value added in manufacturing. We should note that small firms can choose not to report in the BRN. However, firms that choose not to report add up to a small share of value added, since they are included in the 9.7% of value added not included in our sample. A detailed description of the construction and characteristics of the dataset is included in Online Appendix B.

In order to dissect this large dataset in a way that we can understand and analyze through the lens of the theory described in the previous section, we first need to determine what constitutes a layer of workers or management in the data. To do so, we use the PCS-ESE classification codes for workers in the manufacturing sector. Remember that the notion of a layer is a group of employees who have similar knowledge levels and wages and who perform tasks at a similar level of authority. That is, our purpose is not to separate employees in a firm according to the functional characteristics of the tasks they perform (e.g., whether they are accountants or lawyers) but rather based on their hierarchical level in the organization, that is, based on the number of layers of subordinates that they have below them. The PCS-ESE is, we believe, ideal for this purpose. For manufacturing it includes five occupational categories classified with numbers from 2 to 6.[9] The classes we use, together with their class number, are:

2. Firm owners receiving a wage (which includes the CEO or firm directors).

3. Senior staff or top management positions (which includes chief financial officers, heads of human resources, and logistics and purchasing managers).

4. Employees at the supervisor level (which includes quality control technicians, technical, accounting, and sales supervisors).

5. Qualified and non-qualified clerical employees (secretaries, human resources or accounting employees, telephone operators, and sales employees).

6. Blue collar qualified and non-qualified workers (welders, assemblers, machine operators and maintenance workers).

Throughout the paper we merge classes 5 and 6, since the distribution of wages of workers in these two classes is extremely similar, indicating similar levels of knowledge. Table 1 shows

[9] Class 1 is only used for farmers and so is never present in our data.

percentiles of the distribution of wages in the different classes of workers, all expressed in 2005 euros.[10]

Table 1: Distribution of average hourly wage by occupation in 2005 euros

	CEO, directors	Senior staff	Supervisors	Clerks	Blue collars
Mean	81.39	47.83	26.58	19.01	20.70
p5	23.68	21.45	14.35	10.63	10.64
p10	28.60	25.01	16.21	11.79	11.82
p25	41.51	31.00	19.36	13.84	13.65
p50	58.06	38.28	23.11	16.49	15.97
p75	80.48	47.26	27.76	19.95	19.07
p90	114.51	59.91	34.15	24.66	23.40
p95	142.29	72.08	40.45	29.37	27.87

The distributions are clearly ranked. CEOs make the most money, and wages decrease as we reach classes 5 and 6, which are practically identical. In order to match the numbers of these occupational classes with the theory, we order them from the bottom up. So classes 5 and 6 will form the layer of production workers, namely, layer one. Class 4 of supervisors will form the second layer, layer 2. Senior staff will be included in layer 3, and CEOs and firm directors will form layer 4. So firms can have a maximum of 4 layers, starting with layer one and moving all the way up to layer 4.

3.1 Firms with a different number of layers are different

We aim to establish that this classification of employees into layers is a meaningful economic classification. Of course, this occupational classification could just constitute some arbitrary names given to particular workers in an organization that are not systematic across firms. The evidence in Table 1 suggests otherwise. Wages across these occupations are evidently ranked. Clearly, much more is needed. We dedicate the rest of this section to convincing the reader that this classification is useful. In Tables 2 and 3 we present some basic statistics of our dataset.

Table 2 presents the number of firms by year as well as average value added, hours of work, wages, and layers. There is little variation by year in the data, as is evident from the table. We classify a firm as having a particular layer if it reports employing a positive number of hours in that layer. On average, firms in our sample employ a positive number of hours in about 2.5 layers. It is important for our purposes that firms do not tend to employ workers in all layers, since we will analyze how firms change as they add or drop layers of management. Given that on average firms have only slightly more than 2.5 layers and that they can have a maximum of four layers, there are ample opportunities for firms to add new layers.

Table 3 presents the average characteristics of firms across layers. Clearly, firms with more layers are larger in terms of value added and hours. They also tend to pay higher wages. The

[10] Throughout the paper all nominal variables are expressed in 2005 euros.

Table 2: Data description by year

Year	Firms	Average VA	Average Hours	Average Wage	Average # of layers
2002	78,494	2,929	78,775	22.46	2.60
2003	76,927	2,922	77,813	22.69	2.58
2004	75,555	2,957	76,574	23.47	2.59
2005	74,806	2,799	73,078	23.64	2.55
2006	73,834	2,847	72,770	23.50	2.53
2007	71,859	2,709	68,908	24.10	2.51

Value added in 000s of 2005 euros. Wage = average hourly wage in 2005 euros.

last fact is more evident when we look at the median than at the mean given that there are some outliers for firms with only layer zero. Figures 2, 3, and 4 document the same facts using the whole distribution.

Table 3: Data description by number of layers in the firm

Number of layers	Firm-years	Average VA	Average Hours	Average Wage	Median wage
1	80,326	201	7,656	26.90	17.50
2	124,448	401	15,706	21.82	18.64
3	160,030	2,834	80,488	22.31	20.41
4	86,671	8,916	211,098	23.89	22.04

Value added in 000s of 2005 euros. Wage = average hourly wage in 2005 euros.

The figures also show the distributions after we control for time and industry fixed effects (which as can be seen in the figure do very little). The distributions provide a picture very similar to the one we obtained from just looking at the means. Firms with more layers are larger and tend to pay higher average wages. As a robustness check, Table A1 and Figure A1 in Online Appendix A show the same findings using an alternative measure of wages (from DADS) that does not include payroll taxes and other expenses. A description of the data is presented in Online Appendix B. The evidence in this section has documented not only that layers do not simply group workers in arbitrary ways but that firms with different numbers of layers are different in economically meaningful ways.

3.2 Firms have adjacent layers and form hierarchies

So far we have studied firms with different total numbers of layers but not which layers firms actually do include in their organization. Table 4 shows that the vast majority of firms have adjacent layers starting from layer one. That is, 87.42% of the firms in our sample that have one layer actually have occupations 5 and 6. These are also the largest firms in terms of employment, as they account for 99.17% of the total employment of firms with one layer. As another example, consider the firms with three layers. About 80% of them include occupations from 3 to 5 and 6. The remaining 20%

Figure 2: Value added distribution by number of layers

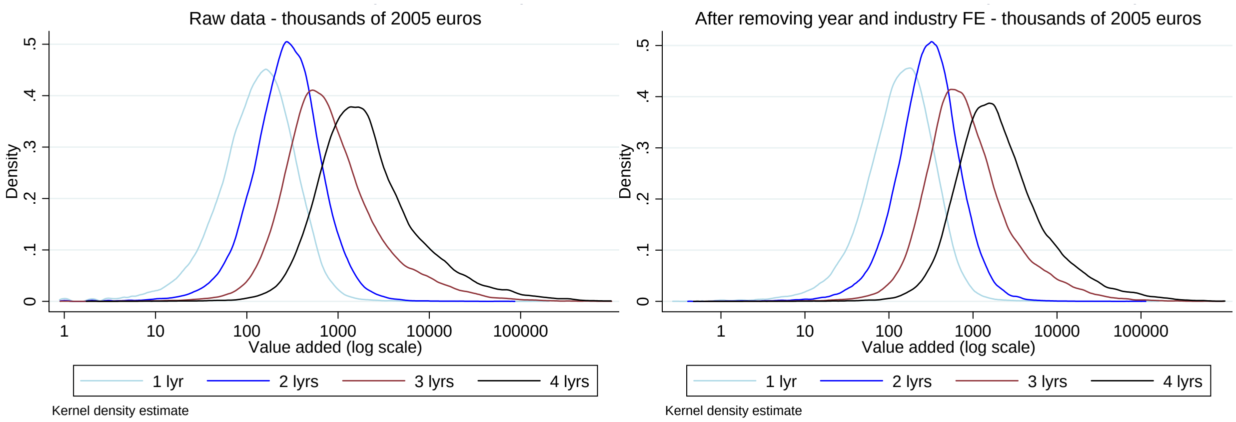


Figure 3: Hours distribution by number of layers

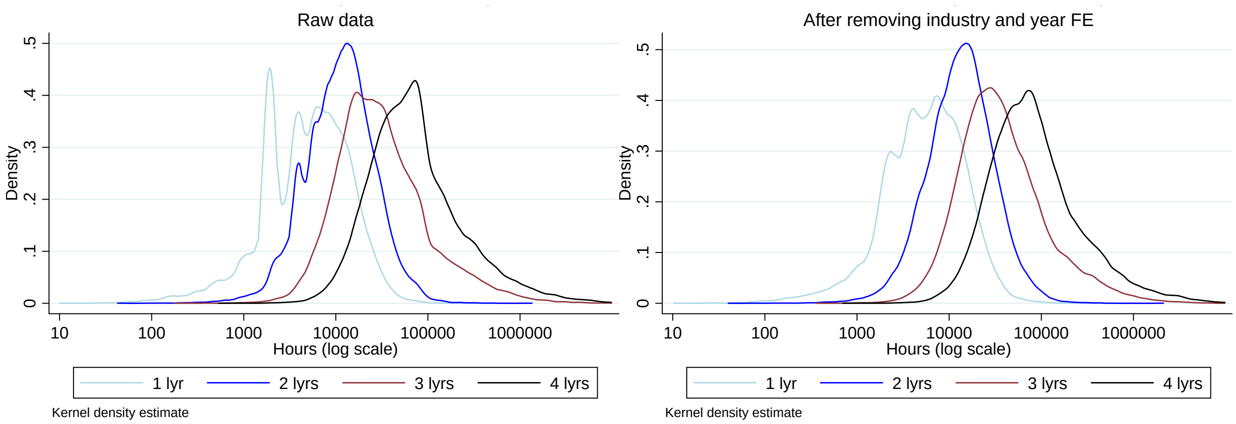
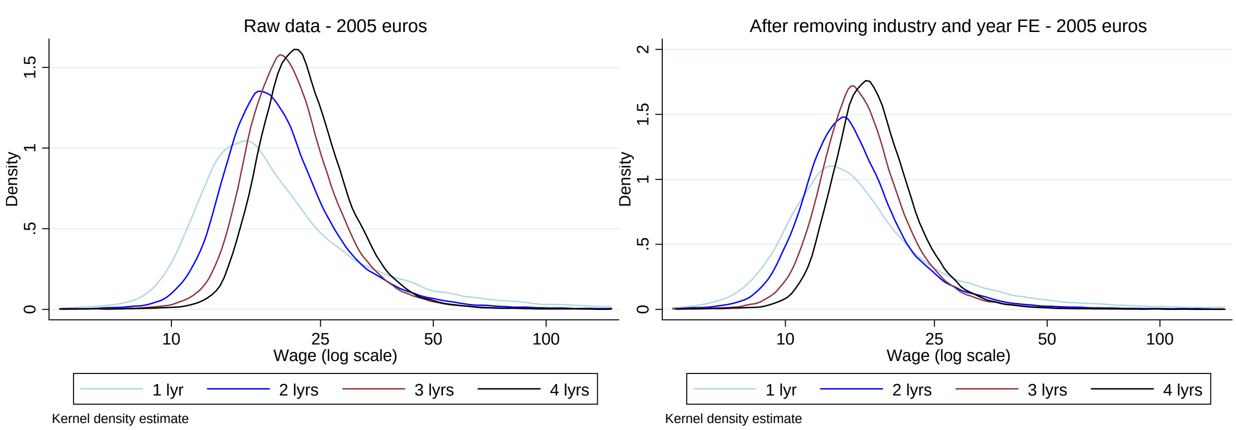


Figure 4: Firm average hourly wage distribution by number of layers



are missing one of these occupations and include occupation 2. Note that, weighted by employment or value added, these firms account for less than 7%. So the producers that do not have adjacent layers starting from layer one are the smallest firms in the sample. Clearly, all firms with four layers have adjacent layers starting from one since they have all layers. Overall, 81.69% of firms have adjacent layers starting at one and they account for 96.73% of value added. The results that follow in general do not depend on whether we restrict the sample to firms with adjacent layers (what we label 'selected sample'). These robustness checks are included in Online Appendix A.

Table 4: Percentage of firms that have adjacent layers

Among firms with:	1 layer	2 layers	3 layers	4 layers	All firms
Unweighted	87.42	67.39	80.01	100	81.69
Weighted by VA	87.69	68.40	94.60	100	96.73
Weighted by hours	99.17	72.56	93.07	100	95.69

Not only do firms have adjacent layers but they form hierarchies. That is, the number of hours employed in the lowest layer is, in most of them, larger than in the second layer, which is larger than the third layer, which is larger than the top layer (if the firm has all these layers). Table 5 presents the fraction of firms that satisfy this hierarchical criterion for hours in all layers and in each of them individually. Table A33 in Online Appendix A presents the averages weighted by value added.

Table 5: Firms that satisfy a hierarchy in hours

Number of layers	$N_L^\ell \geq N_L^{\ell+1}$ all ℓ	$N_L^1 \geq N_L^2$	$N_L^2 \geq N_L^3$	$N_L^3 \geq N_L^4$
2	85.6	85.6	-	-
3	63.4	85.9	74.8	-
4	56.5	86.9	77.5	86.9

N_L^ℓ = hours at layer ℓ of a firm with L layers.

A particular ranking of layers is hierarchical in the sense that the upper layer is smaller than the lower one in at least 74% of cases in the data. Almost all firms with two layers satisfy the ranking. However, only slightly more than half of the firms with all layers satisfy the ranking in all layers. So most firms are hierarchical in terms of hours, but we have a relatively large number of exceptions in at least one layer. In contrast, when we look at the hierarchy in wages –namely, whether workers in higher layers earn more than workers in lower layers– the hierarchy is satisfied in the vast majority of cases. We present this evidence in Table 6. All individual rankings are hierarchical in more than 87% of cases, and even firms with four layers are hierarchical in all the rankings in about 80% of cases. Table A34 in Online Appendix A presents the averages weighted by value added.

In sum, we conclude from this evidence that it is accurate to think of the representative firm as hierarchical, with more hours of work in lower layers, but workers that are paid less.

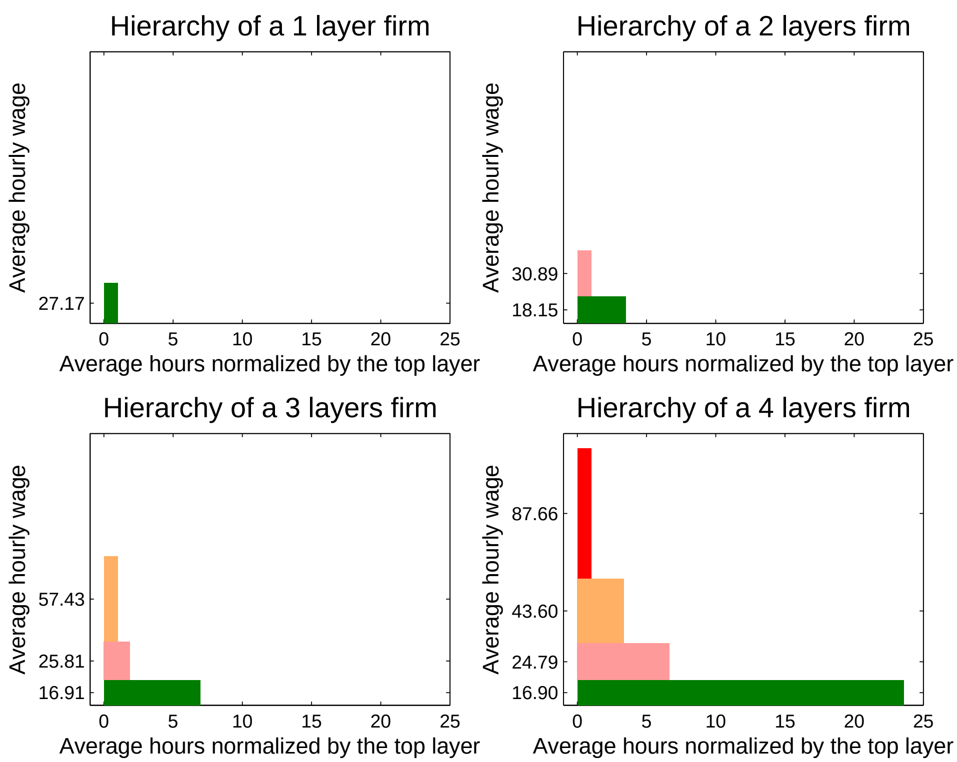
Figure 5 represents graphically the firms in our sample. Each panel in the graph represents

Table 6: Firms that satisfy a hierarchy in wages

Number of layers	$w_L^{\ell+1} \geq w_L^{\ell}$ all ℓ	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$	$w_L^4 \geq w_L^3$
2	92.1	92.1	-	-
3	86.3	93.7	92.5	-
4	80.1	96.6	94.5	87.9

firms with different numbers of layers. Each layer is represented using a rectangle. The length of the rectangle represents the average number of hours employed in the layer by firms with a given number of layers. The height of the rectangle represents the average hourly wage of employees in that layer (so the area is the total wage bill of the layer). The hierarchical organization of labor is evident. Also evident is the way in which firms with more layers organize differently. In the next section we study the particular changes in wages and hours by layer as firms expand. In the graph we normalize the number of hours of each layer by the number of hours in the top layer. Our model keeps the number of hours at the top layer fixed, so this normalization is desirable when we contrast the implications of the theory with our data. All the characteristics of the representative hierarchies that we discussed in Figure 5 are also present if we do not normalize the top layer.

Figure 5: Representative hierarchies normalized by hours in the top layer



The theory in CRH as well as our empirical analysis underscores *average* wages at each layer. Clearly this is a simplification since workers within a layer are bound to be heterogenous in their knowledge; for example, due to the individual histories of workers and the frictions faced by firms to hire and fire employees with particular levels of knowledge. Still, we have not said anything about how much of the variation in wages within the firm is explained by variation across layers rather than by variation within layers. This is relevant since if the fraction explained by cross-layer variation was negligible, our focus on layers would be clearly misguided, at least when it comes to analyzing the distribution of wages within firms. Table 7 shows that this is not the case.

The mean share of variation in log wages explained by variation in layers for all firms is about half, independently of how we weight firms. The share is zero for firms with one layer (since, by definition, for these firms there is no cross-layer variation in log wages) and grows to 66% for firms with four layers. The table reassures us that variation in wages across layers is essential to understand the distribution of wages within firms.

Table 7: Mean share of variation in wages explained by cross-layer variation

			Weighted by	
	Firm-years	Unweighted	Hours	VA
All firms	434,872	0.50	0.51	0.49
Firms with more than 1 layer	370,997	0.59	0.51	0.50
Firms with 1 layer	63,875	0.00	0.00	0.00
Firms with 2 layers	124,299	0.50	0.41	0.43
Firms with 3 layers	160,028	0.62	0.51	0.50
Firms with 4 layers	86,670	0.66	0.53	0.50

3.3 Layer transitions depend on size and firms add or drop adjacent layers

Let us now investigate how many producers add or drop layers in a given period. Table 8 shows that between 60 and 70% of firms in a given period maintain their number of layers. From the remaining, some firms exit, with the exit rate decreasing with the number of layers. Clearly, of the firms that change layers, the majority adds or drops only one of them. In fact, out of the firms with adjacent layers, most of the firms that add one add an adjacent layer (75.5% for one-layer firms and 82.3% for firms with two layers; see Table A2 in Online Appendix A). Hence, when firms add or drop layers they tend to drop or add an adjacent layer, and only one of them. This is all consistent with the view, provided by the theory, that firms add layers to expand and drop layers to contract, and do so in a systematic way. Since very large expansions are rare, we see few transitions that add or drop more than one layer. Table A3 in Online Appendix A shows that the same pattern as in Table 8 is observed even if we weight the firms by their value added.

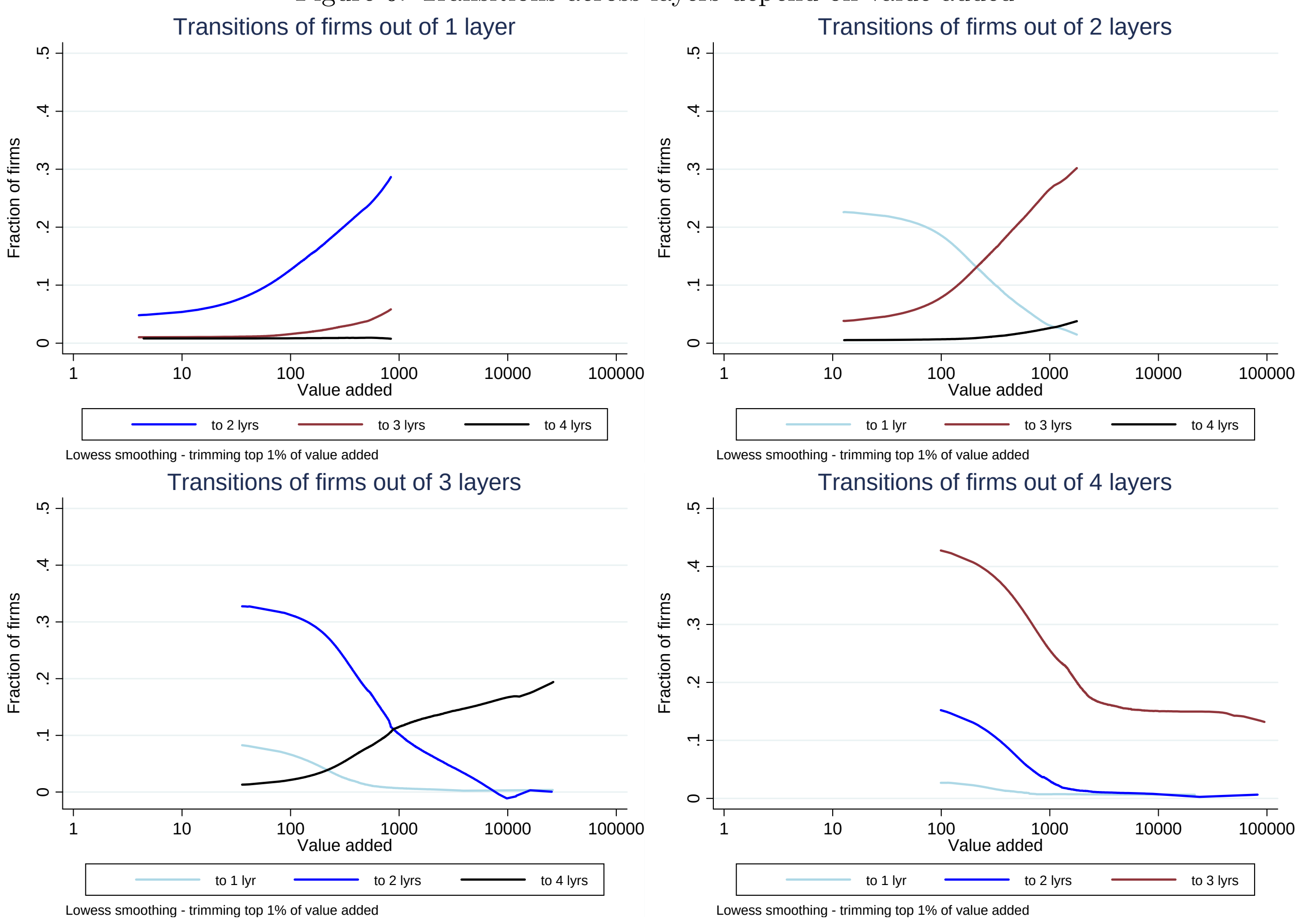
We can also study the probability of adding or dropping one or more layers as a function of the size of the firm in terms of its value added. Figure 6 shows a lowess smoothing interpolation of the probability of changing the number of layers to any count, as a function of the value added of the firm, for firms with different initial numbers of layers. If, for example, firms receive shocks to their demand parameter α over time, and these shocks are drawn from a common arbitrary distribution, the model predicts that the probability of adding a layer should increase with value added. In

Table 8: Distribution of layers at t+1 conditional on layers at t

		Number of layers at $t+1$					
		Exit	1	2	3	4	Total
Number	1	15.3	67.5	15.2	1.9	0.2	100
of	2	9.8	10.7	62.2	16.2	1.1	100
layers	3	7.7	1.2	13.1	67.6	10.5	100
at t	4	6.2	0.2	2.0	20.5	71.3	100

contrast, the probability of dropping a layer should decrease in value added. Furthermore, the probability of adding one layer should be larger than the probability of adding two, which should also be increasing in value added. This is exactly what happens in Figure 6.

Figure 6: Transitions across layers depend on value added

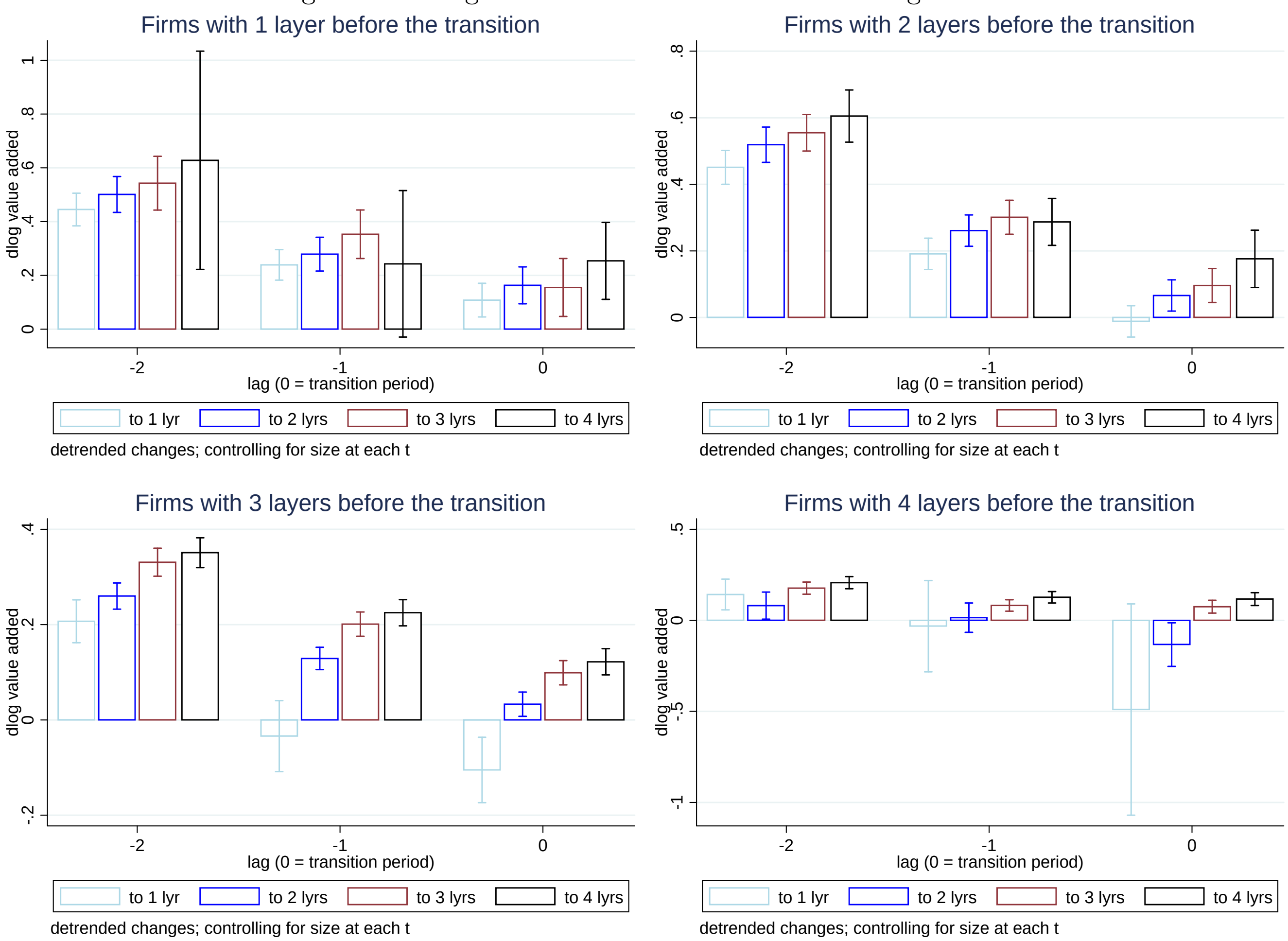


The probability of adding layers is always increasing in value added and of dropping is always decreasing, and the ranking of probabilities is always consistent with the predictions of our theory, augmented with some simple stochastic process for the fundamentals. Figure 6 does not include confidence bands in order to enhance the visibility of the curves. However, in Figure A2 in Online Appendix A we present a graph with all the individual observations and show that they line up fairly tight around the interpolation estimates. Online Appendix B describes all the details to construct Figure 6.

3.4 Trends before adding or dropping layers

The model we sketched in Section 2 suggests that the firm's decision of how many layers to have is characterized by a series of size-thresholds. Given a distribution of shocks to a firm's revenues (e.g. a demand or a productivity shock) a firm will add a new layer if the cumulative set of shocks since its last change is large enough. This logic implies that, on average and conditional on a number of layers, firms that add layers in t should grow faster in the previous couple of years than other firms with the same number of layers. Similarly, firms that drop layers at t should tend to grow slower in $t-1$, and $t-2$ than other firm with the same number of layers. Figure 7 presents evidence on this hypothesis.

Figure 7: Changes in value added before and during transition



We take all firms i in our dataset who have the sequence of layers (L, L, L, L') over time for any

L and $L' = \{1, 2, 3, 4\}$. We then estimate, for $k = 0, 1, 2$,

$$d \ln \widetilde{VA}_{it-k} = \sum_{L'=1}^{4} \gamma^1_{LL't-k} D_{LL'} + \gamma^2_{Lt-k} \ln VA_{it-k} + \epsilon_{it-k},$$

where $d \ln \widetilde{VA}_{it-k}$ is the detrended log change in value added from period $t-k$ for firm i, and $D_{LL'}$ is a dummy that takes the value of one if a firm with L layers at the end of the sequence has L' layers. Figure 7 presents the value of $\gamma^1_{LL't-k}$; namely, the effect of the LL' transition on the mean percentage growth rate in detrended value added, conditional on firm size, two periods before, one period before, and at the moment of the transition to other layers. We also present the 95% confidence interval.

It is clear from Figure 7 that firms that add layers in period t grow faster in $t-2$, $t-1$ and t, than firms that keep or reduce the number of layers. Furthermore, the more layers they add or drop in t, the faster they grow or shrink (although transitions of several layers are estimated less precisely due to small sample sizes). Note that these growth rates are conditional on the size of the firm, and so they are consistent with the view that firms that grow/shrink faster since the last transition are closer to the firm-specific threshold that makes them add/drop a layer.

The evidence in this section has documented that layers do not just group workers in arbitrary ways but that firms with different numbers of layers are different in economically meaningful ways. In addition, we have documented that changes in the number of layers are also systematic and are determined by the size and growth of the firm. We now turn to analyze how firms change their organization, layer by layer, when they decide to expand or contract.

4 How Do Firms Expand?

In this section we analyze how firms change when they expand or contract. The main body of evidence we present tracks firms over time and, therefore, controls for a variety of individual firm characteristics, such as industry. This, we believe, is the ideal way of analyzing the predictions of our theory in the data and we do so below.

We first study the relationship between firm real value added (VA) and normalized hours n^ℓ_L in a given layer, ℓ, for firms with L layers. That is, we study how n^ℓ_L changes as the firm increases its value added. We first detrend all variables using aggregate trends. Namely, if i refers to a particular firm, $\tilde{n}^\ell_{Lit} = n^\ell_{Lit}/\bar{n}_t$ and $\widetilde{VA}_{it} = VA_{it}/\overline{VA}_t$ where $\bar{n}_t$ and $\overline{VA}_t$ are the average normalized hours and value added. We then estimate a regression of the form

$$d \ln \tilde{n}^\ell_{Lit} = \beta^\ell_L d \ln \widetilde{VA}_{it} + \varepsilon_{it}, \qquad (6)$$

where d denotes a yearly time difference (e.g., $d \ln n^\ell_{Lit} = \ln n^\ell_{Lit+1} - \ln n^\ell_{Lit}$).[11]

Table 9 presents the estimates of β^ℓ_L for all $\ell \in 1, 2, ..., L$ and every L. Note that since we are

[11] In the main text we use a specification that controls only for aggregate trends in normalized hours. However, all our results are robust to adding layer-specific time trends.

Table 9: Elasticity of hours with VA for firms that do not change L

Number of layers	Layer	β_L^ℓ	s.e.	p-value	obs
2	1	0.042	0.012	0.00	64,536
3	1	0.039	0.009	0.00	91,253
3	2	0.013	0.010	0.20	91,253
4	1	0.107	0.014	0.00	52,799
4	2	0.051	0.013	0.00	52,799
4	3	0.037	0.013	0.00	52,799

normalizing hours by the number of hours in the top layer, we can look at the value of β_L^ℓ for $\ell = 1, ..., L-1$ only. First note that as predicted by Implication 3 of the theory in Section 2, given L, firms grow by increasing the number of hours at all layers. Furthermore, the ranking of the values of β_L^ℓ always satisfies that $\beta_L^\ell > \beta_L^{\ell'}$ for $\ell < \ell'$, although in one instance the difference is not significant.[12] Hence, these results show that once we control for firm fixed effects, the predictions of the theory are in line with our findings. As a robustness check we present in Table A35 the results with the selected sample. We conclude that when firms grow but keep the same number of layers, they employ more hours of work at all layers but proportionally more in the lower layers. So firms become flatter, with a wider base.

We do the same analysis for wages. Namely, we study the relationship between value added and wages in a given layer for all firms. As we did for hours above, we detrend wages and value added by removing the yearly mean across all layers and firms. So we run

$$d \ln \tilde{w}_{Lit}^\ell = \gamma_L^\ell d \ln \widetilde{VA}_{it} + \varepsilon_{it}, \quad (7)$$

where $d \ln \tilde{w}_{Lit}^\ell$ is the log difference in detrended wages, $\tilde{w}_{Lit}^\ell = w_{Lit}^\ell / \bar{w}_t$, and $\bar{w}_t$ is the mean hourly wage across all firms in year t.[13]

The results are presented in Table 10 and are all consistent with the theory. Namely, γ_L^ℓ is positive and significant for all $\ell \in 1, 2, ..., L$ and every L. Furthermore, $\gamma_L^\ell < \gamma_L^{\ell'}$ for $\ell < \ell'$, in all cases. Hence, when firms grow without changing the number of layers, they increase wages (or knowledge according to the theory) in all layers, but they increase wages proportionally more at the top of the firm as the model predicts. Table A32 in Online Appendix A presents several robustness checks.

The above analysis paints a familiar picture of the way firms expand. Firms expand by adding more workers of all types, by hiring more knowledgeable workers, and by paying them more. Most models of firm dynamics (like Lentz and Mortensen 2008, among many others) share these features

[12] The theory in fact has a more subtle prediction. Namely, the slope of the relationship between the log of value added and normalized hours should decrease as we consider higher layers. The reason is that a larger firm with the same L has more knowledgeable workers in all layers, as discussed in Section 2, and so larger spans of control at all layers.

[13] Again, all our results are robust to adding layer specific time trends.

Table 10: Elasticity of wages with VA for firms that do not change L

Number of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
1	1	0.077	0.007	0.00	45,045
2	1	0.100	0.006	0.00	64,536
2	2	0.118	0.006	0.00	64,536
3	1	0.145	0.006	0.00	91,253
3	2	0.155	0.006	0.00	91,253
3	3	0.170	0.006	0.00	91,253
4	1	0.171	0.009	0.00	52,799
4	2	0.185	0.009	0.00	52,799
4	3	0.186	0.010	0.00	52,799
4	4	0.217	0.011	0.00	52,799

with the theory outlined in Section 2 when firms keep the number of layers constant. The next subsection shows that when a firm's expansion leads to a change in the number of layers –a reorganization of the firm– many of these findings are altered in a significant way, specifically, the one predicted by Implication 4 in Section 2 above.

4.1 Expansions that add layers

We first look at how firm-level outcomes change depending on whether firms add or drop layers of management. Table 11 shows the average log changes in total hours, total normalized hours, value added, and average wages (including and excluding the new top manager in the case of adding layers) for all firms, the ones that add layers, the ones that do not change layers, and the ones that drop layers. As one can see in the first column of the table, most of these variables exhibit some trend over time, and so the average log change is significantly different from zero. To account for this, we also present average changes after we control for time trends (see Online Appendix B for details). Clearly, adding layers is related to increasing hours, normalized hours, and value added. In contrast, firms that add layers decrease average wages once we take out the common time trend. Furthermore, if we look at wages in the pre-existing layers only, wages fall significantly, by 12.2%. The results are reversed when we select only firms that drop layers. Now wages rise by 12.2%.

These estimates demonstrate that in firms that expand by adding layers, average wages in pre-existing layers fall. This is inconsistent with many conceptualizations of firm dynamics where firms that expand always increase the wage of all their employees. Note also that since overall wages increase (without de-trending), as do wages of firms that do not exhibit changes in layers, the fall in wages cannot be the result of reverse causality in which drops in wages cause expansions. If that were the case, we would see drops in wages associated with expansions in all firms, not only the ones that add layers. Furthermore, we would not obtain the opposite result when we select only firms that drop layers. Of course, the theory in Section 2 is exactly consistent with this evidence on wages. Implication 4 says that, as firms add layers, the knowledge and therefore wages at all

Table 11: Change in firm-level outcomes

	All	Increase L	No change in L	Decrease L
$d\ln$ total hours	-0.015***	0.040***	-0.012***	-0.081***
- detrended	-	0.055***	0.003***	-0.066***
$d\ln\sum_{\ell=0}^{L} n_L^{\ell}$	-0.011***	1.362***	0.012***	-1.404***
- detrended	-	1.373***	0.023***	-1.392***
$d\ln VA$	-0.008***	0.032***	-0.007***	-0.050***
- detrended	-	0.040***	0.001	-0.041***
$d\ln$ avg wage	0.019***	0.015***	0.019***	0.025***
- detrended	-	-0.005***	-0.000	0.006***
- common layers	0.021***	-0.101***	0.019***	0.143***
- - detrended	-	-0.122***	-0.002***	0.122***
% firms	100	12.65	73.66	13.68
% VA change	100	40.12	65.08	-5.19

*** significant at 1%.

layers should decrease.

The results in Table 11 suggest that wages in firms that add or drop layers behave differently than previously thought. Table 11 also shows that the firms that add or drop layers represent an important fraction of firms in the economy, as well as an important fraction of value added. Firms that add layers represent 12.65% of the total. Furthermore, as they are on average larger than their counterparts that do not add layers, they contribute 40.12% of the total change in value added. Conversely, firms that drop layers represent 13.68% of firms and contribute -5.19% to the change in value added. Together, the firms that reorganize by changing layers and that therefore change wages in the new way we uncover represent more than a quarter of the firms in the economy and contribute more than 40% of the absolute changes in value added. So the firms that change their organization to expand and contract do not represent a fringe of the firms in the economy. They are essential to understanding firm dynamics and the associated labor market outcomes. The theory of organization with heterogeneous firms in CRH can rationalize the behavior of these firms.

We now proceed to analyze in more detail firms that change their layers of management. In particular, we are interested in whether firms that add layers add hours of work to all layers and decrease wages in all layers. The results above tell us that this is the case on average, but they do not imply that this happens layer by layer. Table 12 computes average log changes in detrended normalized hours for firms that transition between layers. Each line in the table represents a particular type of transition (e.g., from two to three layers) and a particular layer in firms that undergo that transition. The first column in the table indicates the number of management layers in the initial period and the second column the number of layers in the second period. The third column indicates the layer, ℓ, for which we are calculating the average (over i and t) of $d\ln\tilde{n}_{Lit}^{\ell}$. The fourth column indicates the coefficient of interest. Note first the sign of the average change. It is positive and significant for all firms that increase the number of layers (by one or more layers). Symmetrically, it is negative in all layers for all firms that drop one or more layers, exactly what we would expect from the theory in Section 2.

Table 12 indicates that the firm-level outcomes on normalized hours from Table 13 not only hold for the firm as a whole, but they hold layer by layer too. All our estimates are significant at the 1% level.[14] As we did above, we detrend all variables using aggregate trends. Online Appendix A presents a variety of robustness checks. In particular, it presents the results when we only use firms with adjacent layers and we condition on firm-level outcomes, such as expansions in hours or value added.

Table 12: Average log change in hours for firms that transition

# of layers Before	# of layers After	Layer	$d\ln\tilde{n}^{\ell}_{Lit}$	s.e.	p-value	obs
1	2	1	1.537	0.018	0.00	10,177
1	3	1	1.762	0.056	0.00	1,263
1	4	1	2.266	0.212	0.00	97
2	1	1	-1.582	0.017	0.00	11,106
2	3	1	0.716	0.012	0.00	16,800
2	3	2	0.539	0.012	0.00	16,800
2	4	1	1.205	0.049	0.00	1,129
2	4	2	1.004	0.048	0.00	1,129
3	1	1	-1.795	0.048	0.00	1,584
3	2	1	-0.682	0.012	0.00	17,666
3	2	2	-0.518	0.012	0.00	17,666
3	4	1	1.352	0.014	0.00	14,113
3	4	2	1.289	0.016	0.00	14,113
3	4	3	1.174	0.016	0.00	14,113
4	1	1	-2.119	0.173	0.00	123
4	2	1	-1.059	0.041	0.00	1,456
4	2	2	-0.918	0.040	0.00	1,456
4	3	1	-1.411	0.014	0.00	15,160
4	3	2	-1.345	0.015	0.00	15,160
4	3	3	-1.260	0.015	0.00	15,160

We do the same analysis layer by layer for changes in wages. Namely, we compute the average (over i and t) of $d\ln\tilde{w}^{\ell}_{Lit}$ for firms that add or drop layers. We present the results in Table 13.

[14] The theory also predicts that the proportional change in the hours of employees in higher layers should be larger than the proportional change in the hours of employees in lower layers. The reason is that the knowledge of all employees falls and so does the span of each manager. This results in positive but smaller proportional changes in hours at the bottom of the firm, when the firm adds a layer. In Table 12 we see, in most cases, exactly the opposite. As in the case of firms that keep the number of layers constant, the lower layers expand proportionally more. Several forces can be responsible for this mismatch between the theory and the data. First, during a year, firms that switch might also have grown without further changes in the number of layers. Since according to the theory the rank of the log changes is different depending on whether layers are added or not, what we see could be the result of one effect dominating the other. Second, there could be frictions in hiring that make lower layers easier to expand than higher ones. This would be the case if hiring more knowledgeable employees is more costly and takes more time. Finally, the theory suggests that changes in communication costs, h, as the firm adds layers could also reverse the implications of the theory on this ranking.

Again, the table confirms that the results we obtained for firm-level outcomes hold layer by layer, as the theory predicts. Firms that add layers reduce wages in all pre-existing layers and firms that drop layers increase wages in all the layers of the reorganized firm. Again, these results are robust to conditioning on large firm expansions in value added, normalized hours, or both, as well as to restricting the sample of firms with adjacent layers. Furthermore, we corroborate our results using a different source of wage data. Tables with the robustness checks are presented in Online Appendix A and the description of the tables are presented in Online Appendix B. The conclusion is that many firms expand by adding layers of management and these firms reduce the salary of workers in all pre-existing layers (or, accordingly, their knowledge as the theory would suggest). The theory also predicts that, as a firm adds layers, the wages of higher level managers should fall proportionally more than those of lower level ones (since their knowledge is more sustitutable with that of the top manager), a prediction also corroborated by Table 13.

Table 13: Average log change in wages for firms that transition

# of layers Before	# of layers After	Layer	$d\ln\tilde{w}^{\ell}_{Lit}$	s.e.	p-value	obs
1	2	1	-0.129	0.005	0.00	10,177
1	3	1	-0.332	0.020	0.00	1,263
1	4	1	-0.678	0.117	0.00	97
2	1	1	0.167	0.005	0.00	11,106
2	3	1	-0.050	0.002	0.00	16,800
2	3	2	-0.255	0.004	0.00	16,800
2	4	1	-0.150	0.015	0.00	1,129
2	4	2	-0.409	0.019	0.00	1,129
3	1	1	0.356	0.018	0.00	1,584
3	2	1	0.059	0.002	0.00	17,666
3	2	2	0.249	0.004	0.00	17,666
3	4	1	-0.021	0.002	0.00	14,113
3	4	2	-0.067	0.003	0.00	14,113
3	4	3	-0.199	0.004	0.00	14,113
4	1	1	0.804	0.109	0.00	123
4	2	1	0.139	0.012	0.00	1,456
4	2	2	0.372	0.016	0.00	1,456
4	3	1	0.009	0.002	0.00	15,160
4	3	2	0.040	0.003	0.00	15,160
4	3	3	0.134	0.004	0.00	15,160

The finding is surprising in light of theories of firm growth that do not put organization at center stage. Expansions associated with any revenue-enhancing shock (like a demand or an exogenous productivity shock) are, in virtually all of these theories, associated with increases in employee wages. The channels vary, but the result is the same. Sometimes wages rise because the marginal product of workers increases, sometimes because they share some of the rents, and sometimes because the bargaining position of the worker improves. In contrast, in the theory of CRH, wages go down because by adding an extra layer of management the firm can have top managers deal with the exceptional problems and make workers learn only how to solve the most common problems. The firm economizes on knowledge by having more experts together with less knowledgeable workers and lower-level managers who earn less. Of course, this new organization is only optimal if the firm expands enough, since it involves a higher fixed cost in terms of the salaries of its managers.

We can decompose the total log change in average wages in the firm into two parts. The first part is the change in wages of workers in existing layers, which, as we know, is negative by the results discussed above. The second part is the change induced by adding a new agent at the top of the hierarchy. Agents in the new added layer earn more than the average worker in the firm since they are added, in the vast majority of cases, to the top layer. For example, for firms that transition from one to two layers, the new manager in layer two makes 50.7% more than the average wage in the firm before reorganization. This number can be much higher for larger firms. Firms that go from 3 to 4 layers pay the new top manager 338.5% more than the average worker in the firm before adding the layer. All of these results are presented in Table 14. We decompose the detrended average wage in the firm, $d \ln \bar{w}_{Lit}$, as

$$d \ln \bar{w}_{Lit} = \ln \bar{w}_{L'it+1} - \ln \bar{w}_{Lit} = \ln \left(\left(\bar{w}^{\ell \leq L}_{L'it+1} / \bar{w}_{Lit} \right) s + \left(w^{L'}_{L'it+1} / \bar{w}_{Lit} \right) (1-s) \right)$$

where $\bar{w}^{\ell \leq L}_{L'it+1}$ is the average wage in all pre-existing layers in the reorganized firms with $L' > L$ layers, $w^{L'}_{L'it+1}$ is the wage of the new top manager, and s is the fraction of hours of work done by employees in pre-existing layers. Table 14 presents each of these components. The fact that $\bar{w}^{\ell \leq L}_{L'it+1} / \bar{w}_{Lit}$ is below one for all transitions is, for practical purposes, just a re-expression of the results in Table 13. The upper-right panel shows the earnings of the top managers as a fraction of the average wage in the firm before transition. Clearly, since workers in pre-existing layers earn less, but the new top manager makes more, the overall effect of adding a layer is ambiguous and not particularly robust. The relevant finding is that new managers are the only ones in the reorganized firm who earn more after adding layers.

The results in this section suggest that in order to understand the behavior of firms that expand, it is essential to condition on whether the expansion requires a reorganization. On average, the firms that expand the most tend to reorganize. So to understand expansions we need to understand reorganization. The salient fact in the data is that when firms expand and reorganize they pay workers on pre-existing layers on average less. This is consistent with the view, borrowed from the theory in CRH, that the firm wants less knowledgeable workers on pre-existing layers after the reorganization. Note that this finding does not challenge the many empirical studies (Abowd, et al. 1999, Brown and Medoff 1989, Oi and Idson 1999, Bernard and Jensen 1997, 1999, Frias, et al. 2009) that have found that average firm wages (or the wages of a particular class of workers) increase with firm size or as firms expand. In fact, we find some evidence that this is true in our sample too. What our finding says is that this is not the case when we condition on the firm reorganizing by adding layers. Furthermore, we find that a substantial fraction of expansions are in fact paired with this type of reorganization.

Table 14: Decomposition of total log change in average wages

$\bar{w}^{\ell\le L}_{L'it+1}/\bar{w}_{Lit}$

From/to	2	3	4
1	0.963*** (10,167)	0.865*** (1,262)	0.733*** (96)
2		0.926*** (16,783)	0.876*** (1,128)
3			0.958*** (14,099)

$w^{L'}_{L'it+1}/\bar{w}_{Lit}$

From/to	2	3	4
1	1.507*** (10,166)	1.501*** (1,263)	1.602*** (97)
2		2.040 *** (16,783)	2.021*** (1,129)
3			4.385*** (14,099)

s

From/to	2	3	4
1	0.741 *** (10,166)	0.620*** (1,262)	0.563*** (97)
2		0.853 *** (16,784)	0.775*** (1,128)
3			0.948 *** (14,099)

$d\ln \bar{w}_{Lit}$

From/to	2	3	4
1	−0.007* (10,166)	−0.094*** (1,263)	−0.305** (97)
2		0.005 ** (16,784)	−0.033** (1,129)
3			−0.001 (14,098)

All results from trimmed sample at 0.05%. *significant at 10%, ** at 5%, *** at 1%. Number of obs. in parenthesis.

One might be concerned that the results presented above are just driven by a reclassification of agents across occupations. For example, as the firm expands the oldest and best paid worker might be now called a floor supervisor and since she has the highest wage, this relabeling results in a lowering of the average wage of the agents that remain as workers in layer one. This relabeling of jobs would show up in our results as a firm that added a layer and reduced the average wage in pre-existing layers, even though the actual distribution of wages within the firm has not changed. To address this potential concern we look at the change in the distribution of wages within the firm after a reorganization.

Figure 8 presents the difference in the log wages at each percentile of the distribution for the different one-layer transitions.[15] The figure also shows a bootstrapped 95% confidence interval constructed as we explain in Online Appendix B. First note that the distribution of wages in firms that transition actually changes significantly. This eliminates the concern that our results are just the result of meaningless relabeling of employees. Furthermore, note how wages in the lower part of the distribution fall for all transitions where firms add layers and they increase when firms drop them, exactly as we have been arguing.

Figure 8 also shows that, in some transitions, at the upper tail of the distribution wages rise when firms add layers and fall when they drop them. This was expected given the results in Table 14 where we show that agents in the newly added layer make more money than the average employee in the firm before the transition. Still, to confirm that this is in fact driven by the agents hired at the new layer, in Figure 9 we present a parallel figure without including the hours of work in the new layer (or the dropped layer). The results are very stark and consistent with our interpretation. In all transitions where the firms add (drop) a layer we observe a shift down (up) in the distribution of wages after the transition. In fact, the largest changes in the distribution are now observed at the upper tail of the distribution. This was expected from the results in Table 13, that indicate

[15] Transitions of more than one layer look similar and we omit them for brevity.

Figure 8: Difference in the distribution of wages after minus before the transition

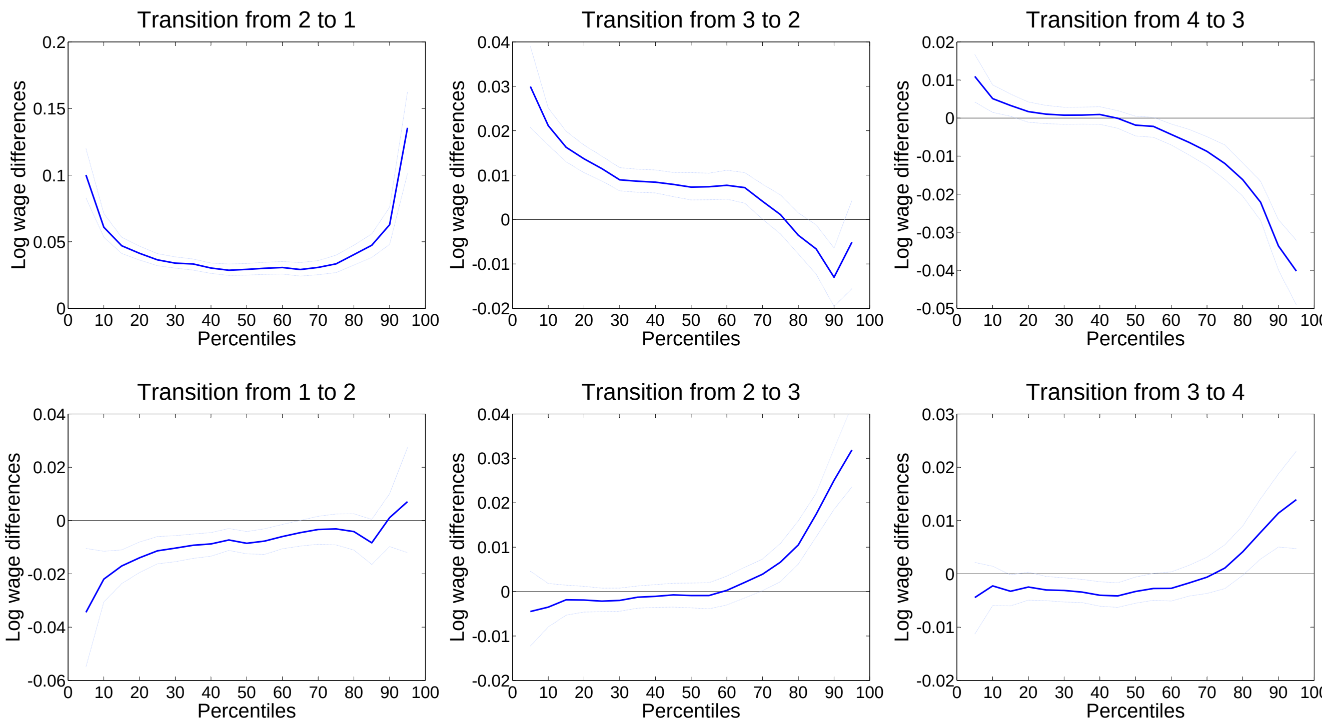
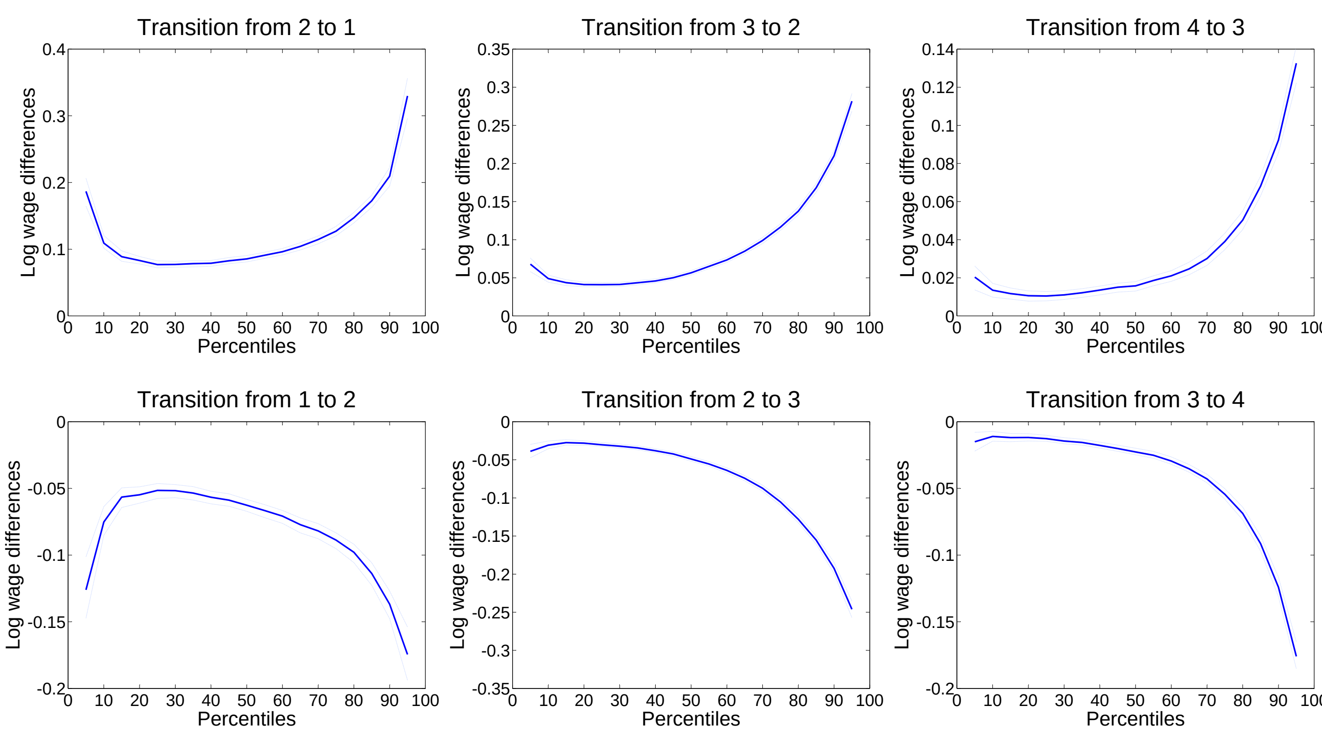


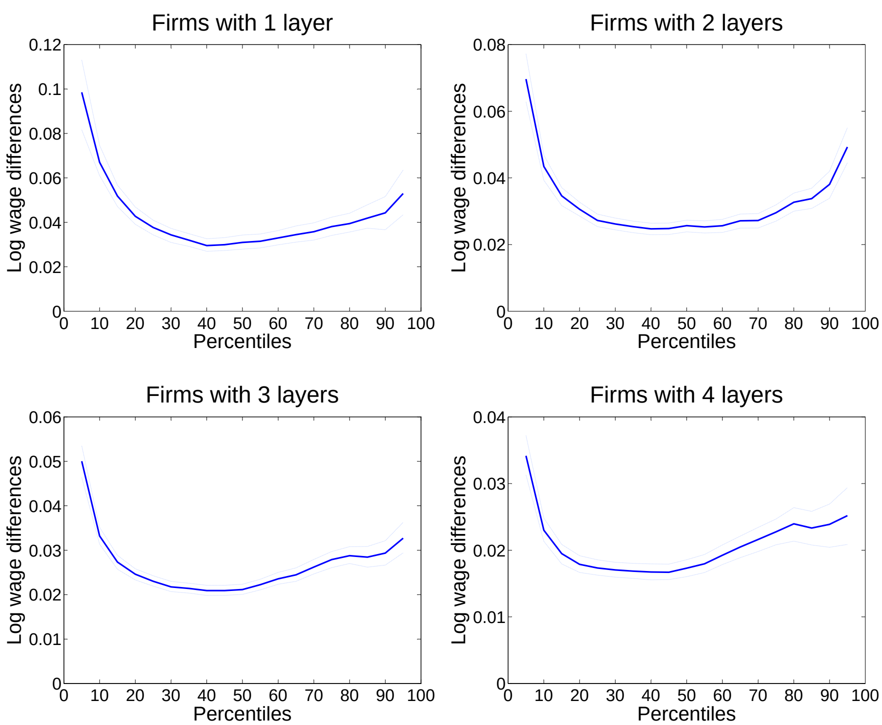
Figure 9: Difference in the distribution of wages in common layers after minus before the transition



that average wages in the upper pre-existing layers are the ones that fall the most when firms add a layer.

These results are particularly meaningful when contrasted with Figure 10 where we present the difference in the log wages at each percentile of the distribution for firms that expand *without* reorganizing.[16] As before, the distribution of wages changes significantly. However, now the distribution of wages shifts up significantly at all percentiles of the distribution. So, as expected from the layer-level results above, the distribution of wages of firms that grow and reorganize shifts down for preexisting hours of work, while the distribution of wages of firms that grow without reorganizing shifts up.[17] These results underscore how, even if one is not interested in layer-level outcomes, conditioning on reorganization is essential to understand firm growth and its implications for wages and the characteristics of a firm's labor force.

Figure 10: Difference in the distribution of wages for firms that do not transition and $dlnVA > 0$



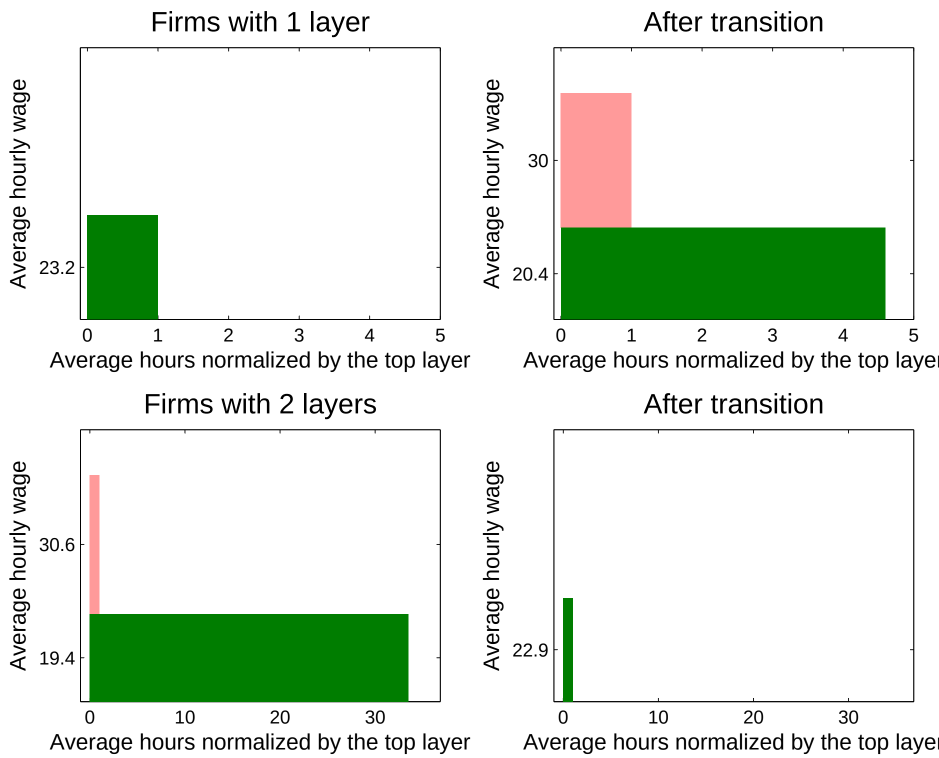
We finish this section with a graphical illustration of firm transitions and the representative changes in wages and normalized hours that result from those transitions. Figures 11 to 13 showcase the results. The main characteristics of these changes have been analyzed before and are consistent

[16] Again, bootstrapped 95% confidence intervals are included in the figure and an explanation of how we constructed the figure is presented in Online Appendix B.

[17] In Figure A5 in Online Appendix A we present the change in the distribution of wages conditioning on firms that contract. In this case the distribution shifts down, as expected from the theory and the layer-level outcomes, for all percentiles in firms with 3 and 4 layers and for most percentiles in firms with 1 and 2 layers. In the case of 1 and 2 layer firms wages at the very top of the distribution increase. This might be the result of some rent extraction by top managers and owners before firms exit.

with the first 4 implications of the theory in Section 2. The top row shows a representative firm that adds a layer and the bottom row a firm that drops a layer. The figure emphasizes the dramatic changes associated with reorganizations in the data.

Figure 11: Representative transitions between firms with 1 and 2 layers



We could also illustrate how firms respond to an expansion of, say, 5% in value added. In that case we can also illustrate what happens with the firms that do not add layers (which we know increase normalized hours and wages). We present these results in Figures A3 and A4 in Online Appendix A.

5 How Do Firms Change the Average Wage in a Layer?

We finish our analysis with an exploration of how firms manage to adjust the average wage in pre-existing layers. The answer to this question will be specific to the particularities of the French labor market. As such, it is interesting for learning something about France, but perhaps less so about firms in general. Still, we want to illuminate as much as possible how firms do what we are arguing they are, in fact, doing. Of course, we know well that labor market frictions, institutions, and regulations make reducing the wage of a particular employee complicated, if not impossible. So, for example, how can the prediction that firms that add layers reduce average wages in pre-existing layers be so clearly present in our data? We explore two margins of adjustment. First, we study if firms adjust average wages in a layer, which we interpret as the average knowledge in the layer,

Figure 12: Representative transitions between firms with 2 and 3 layers

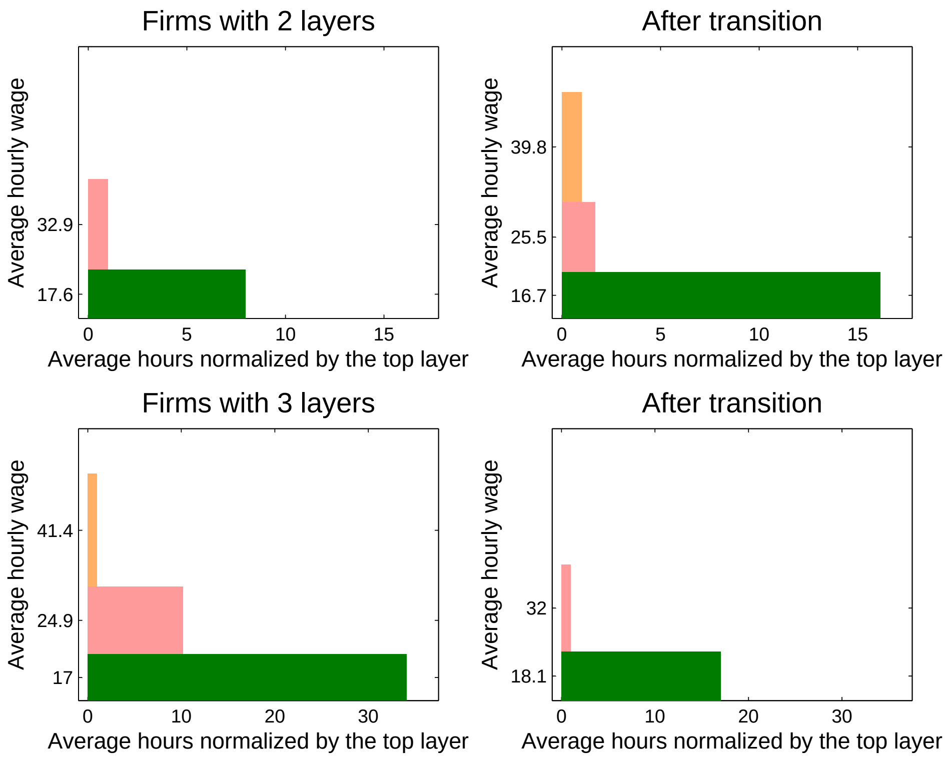
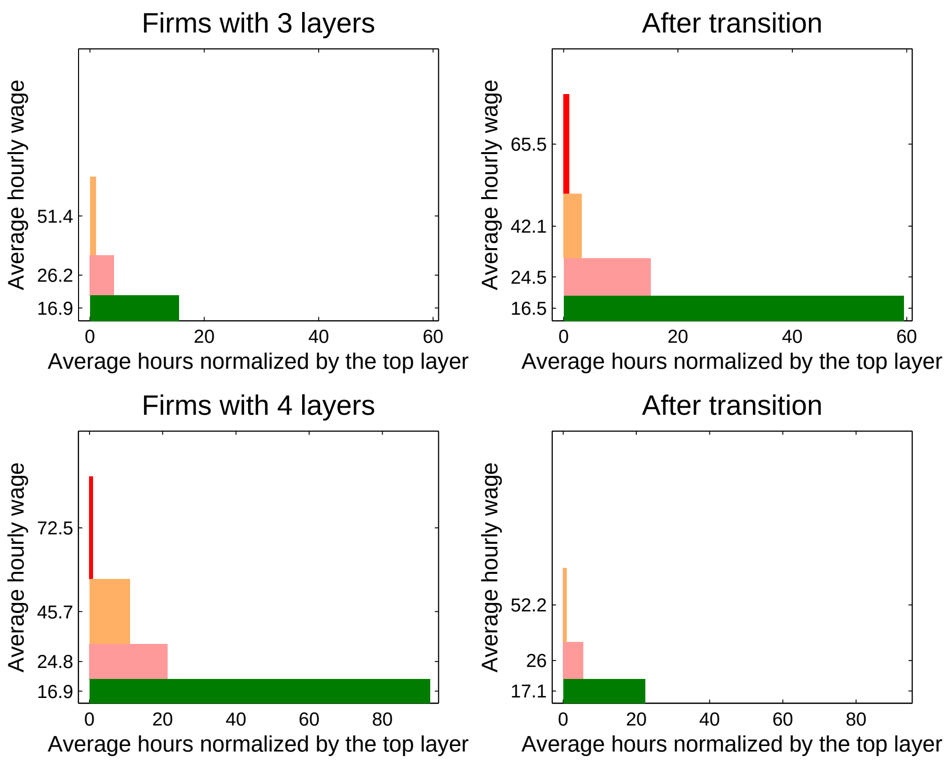


Figure 13: Representative transitions between firms with 3 and 4 layers



by adjusting the wages of existing employees or by adjusting their composition. We then explore which changes in the characteristics of employees lead to the changes in wages or knowledge that we observe. More specifically, we study to what extent the changes in wages are the result of changes in formal education or changes in labor market experience.

5.1 Extensive versus intensive margin

Our data, and any other data in France, are insufficient to track the universe of employees in a firm over time, as would be required to understand the individual agent effects of changes in organization.[18] The data only allows us to study how reorganization affects layer-level average wages by tracking, across two adjacent periods, hours of work, and average hourly wages, as employees enter, stay, or leave the layer during a transition.[19] Studying these transitions we find that firms reduce wages by promoting or firing the highest paid hours of work in a layer and by hiring new hours of work that are paid less. They keep the salaries of hours of work that stay in the layer essentially unchanged, although in some cases they raise them slightly.[20] This reduces the average wage in the layer. That is, the adjustment in the average wage happens mostly through the extensive margin. Firms adjust the composition of the employees in the layer and not the individual wages that they command.[21]

Note that there is nothing mechanical about this finding. Firms are actively deciding to use the extensive margin to change the average wage in a layer in a particular direction. This is particularly clear when we contrast this finding with the behavior of firms that expand without reorganizing. In those firms, average wages in each layer increase even though the normalized number of hours increases as well. That is, those firms choose to hire new workers in each layer that earn more than the average, not less.

This way of adjusting wages in a layer could be rationalized with the presence of some form of downward wage rigidities. In that case, the reductions we observe in the average wages of pre-existing layer in firms that add layers necessarily understate the reductions that we would observe absent these frictions. Similar empirical explorations in other, more flexible, labor markets would shed light on the validity of this argument.

[18] We cannot use a methodology like the one in Abowd, et al. (1999). The reason is that their data is just a random subsample of the workforce over time (all individuals employed in French enterprises that were born in October of even-numbered years) and therefore, it is impossible to reconstruct the whole organization of the firm to study individual workers' heterogeneity using such data.

[19] See description of Tables A36 to A39 in Online Appendix B for a detailed account of this exercise and findings.

[20] Tables A36 to A39 in Online Appendix A present these results for all transitions. Table A36 shows that the results for hours that stay in the layer are small and not particularly systematic, so they do not dominate the effect on the average wage in a layer. In contrast, Table A37 shows that the wages of hours that enter the layer are always significantly lower than the wages of hours that leave the layer (and vice-versa for firms that drop layers, except for layer 1 in firms that go from 4 to 3 layers of management).

[21] Several studies have documented that average wages at firms increase with firm size, or as firms expand: the firm size-wage premium (Abowd, et al., 1999, Brown and Medoff, 1989, Oi and Idson, 1999, Bernard and Jensen, 1997, 1999, and Frias, et al., 2009). Our results are not inconsistent with these findings: our theory does not address how individual-level outcomes change in response to a reorganization, and we cannot study this question with our data.

5.2 Education or experience to adjust knowledge and wages

So far we have interpreted changes in the average wage in a layer as a direct measure of changes in the average knowledge of agents employed in that layer. This is consistent with a long tradition in labor economics that interprets wage differentials as compensating for individual characteristics (Rosen, 1986). In our case, wage differentials are compensating for the marketable characteristics of individuals that we call knowledge and that individuals use to 'solve problems'. Of course, if we had a direct and reliable measure of average knowledge in a layer we could repeat our study using this measure instead of wages.

We face several difficulties when we embark in this alternative exercise. First, knowledge is a complex concept that is not easily observable in the data. We can only observe a subset of worker characteristics that in reality are combined together to form the 'knowledge' of an individual. Perhaps the two most important of these characteristics are formal education and experience. Still, it is not obvious that any of these characteristics on its own is a better measure of knowledge than a market-based measure like wages. In fact, it is likely that the knowledge of individuals is optimally created by combining them in a way that depends on the relative cost of acquiring knowledge in these different ways. Second, there is the issue of data availability. There is no dataset in France that provides employer-employee matched measures of formal education, experience and perhaps other worker characteristics for every individual employee in an organization. So, instead, we use the most comprehensive labor market survey available in France to estimate formal education and experience in our data set.

Table 15: Elasticity of 'knowledge' with VA for firms that do not change L

# of layers	Layer	Experience	p-value	Education	p-value	obs
1	1	0.0014	0.69	0.0015	0.03	45,009
2	1	-0.0101	0.01	0.0042	0.00	64,469
2	2	0.0094	0.03	0.0032	0.00	64,469
3	1	-0.0103	0.00	0.0038	0.00	91,161
3	2	-0.0011	0.97	0.0026	0.00	91,161
3	3	0.0077	0.00	0.0011	0.10	91,161
4	1	-0.0154	0.00	0.0027	0.00	52,730
4	2	-0.0036	0.28	0.0026	0.00	52,730
4	3	-0.0001	0.97	0.0002	0.79	52,730
4	4	0.0073	0.02	-0.0030	0.07	52,730

We estimate the average years of education and the potential labor market experience for each employee in DADS. We start by using the French Labor Force Survey (Enquête Emploi), a survey that contains information on wages, years of education, age and individual characteristics, among other things. This data allows us to estimate a relationship between log-years of education (or log-potential labor market experience[22]) and the subset of individual characteristics that we can also find in the DADS dataset, namely, the hourly wage, age, industry and time fixed effects, and differ-

[22] Defined as age minus the age of the worker at the end of her formal education.

ent measures of occupation. We then use the resulting coefficients to estimate the average education across all hours employed in each layer in DADS (and similarly for labor market experience).[23]

Table 16: Average change in 'knowledge' for firms that change L

# of layers Before	# of layers After	Layer	Experience	p-value	Education	p-value	obs
1	2	1	-0.1082	0.00	-0.0040	0.00	10,171
1	3	1	-0.1835	0.00	-0.0025	0.29	1,261
1	4	1	-0.3300	0.00	0.0252	0.03	97
2	1	1	0.0962	0.00	0.0053	0.00	11,088
2	3	1	-0.0437	0.00	0.0001	0.82	16,778
2	3	2	-0.1813	0.00	0.0021	0.01	16,778
2	4	1	-0.0644	0.00	0.0021	0.29	1,124
2	4	2	-0.2284	0.00	0.0083	0.01	1,124
3	1	1	0.1374	0.00	0.0056	0.00	1,584
3	2	1	0.0444	0.00	0.0022	0.53	17,626
3	2	2	0.1533	0.00	-0.0004	0.00	17,626
3	4	1	-0.0113	0.00	0.0010	0.10	14,098
3	4	2	-0.0377	0.00	-0.0009	0.00	14,098
3	4	3	-0.1764	0.00	0.0238	0.82	14,098
4	1	1	0.1974	0.00	-0.0019	0.95	123
4	2	1	0.0726	0.00	-0.0001	0.12	1,454
4	2	2	0.1720	0.00	-0.0045	0.00	1,454
4	3	1	0.0132	0.00	-0.0015	0.26	15,150
4	3	2	0.0248	0.00	-0.0006	0.00	15,150
4	3	3	0.1134	0.00	-0.0204	0.00	15,150

Table 15 presents the elasticity of potential labor market experience and education with respect to value added for firms that do not change layers. Table 16 presents the average change in potential labor market experience and education for firms that change layers. We detrended all measures as we did above for the results in Tables 10 and 13. Our findings are robust to several different specifications and forms of detrending the data (see again Online Appendix A).

Consider first the firms that grow without adding layers. As we showed in Table 10 these firms increase average wages in all layers. Table 15 shows that these firms increase knowledge by increasing the formal education in the layer. This is particularly true in the lower layers of the organization. At the highest level of the organization, firms also increase average knowledge by hiring more experienced workers. So all of them increase the average wage or knowledge, but they do so using formal education at the bottom of the organization, and using also experience in the top layer. This behavior is consistent with the view that formal years of education primarily provide the knowledge to solve the most routine problems in an organization, the problems that are handled at the lower layers of the hierarchy. In contrast, labor market experience provides the

[23] Online Appendix B provides details about the exact estimating equation we use, as well as other details and robustness checks on our procedure to estimate these formal education and potential experience measures.

knowledge required to solve more infrequent problems, the tasks handled in the higher levels of the hierarchy.

Table 16 presents results for firms that change layers. Consider the case of firms that add layers, the ones that, according to Table 13 reduce the average wage, or knowledge, in each pre-existing layer. These firms mostly use experience to reduce the average knowledge in a layer and leave average levels of education mostly unchanged. The result is evident in Table 16 where the measure of experience declines in all pre-existing layers for firms that add layers and they rise for all firms that drop layers. These changes are large and precisely estimated. In contrast, for education, the results are very small, not systematic, and in many cases not significant.

These results complete our description of the changes in organization as firms grow. A firm that grows and adds a layer hires more employees at the bottom that have a similar level of education than the ones it already employs. It also promotes the more experienced employees to the higher layers, which lowers experience in all layers. Of course, the firm could balance the implied reduction in knowledge by employing more knowledge employees at the bottom of the organization. However, in contrast to the case when they expand without adding layers, they choose not to do so. This is the active decision that leads to a lower average wage and level of knowledge.

6 Conclusion

This paper provides the first anatomy of organizations using a large scale dataset. Previous studies focused on particular details of the organization and included only a few hundred firms. In contrast, we use virtually all the manufacturing firms in France during the period 2002-2007. To study the organization of these producers and how they modify their organization, we use occupational data to classify workers in layers. The concept of a layer of employees and our empirical implementation using hierarchical occupations have proven useful in analyzing these data. First, firms actively manage the number of layers in the firm. Firms expand by adding them and contract by dropping them.

We find that the organization of labor in the firm depends on the number of layers. In our data, studying the behavior of hours of work or wages in a firm without classifying workers in layers and conditioning on changes in these layers results in ambiguous firm-level effects that are hard to understand and not robust. In contrast, once we classify workers in layers, the behavior of wages and hours of work for each layer is easy to understand through the lens of a theory of hierarchical organization, such as the one in CRH (which borrows the technology from Garicano, 2000, and Rosen, 1982). Firms that expand without adding layers increase hours at all layers as well as average wages at all layers. They essentially behave the way most theories would suggest expanding firms behave. In contrast, firms that add layers expand much more, add many hours of work to all layers, but reduce average wages in all pre-existing layers. The fact that the wages of all pre-existing layers go down is, as far as we know, a novel empirical finding. Furthermore, it is hard to rationalize with theories that do not have the organization of knowledge at their core.[24] The evidence we provide is extremely robust and we hope will be confronted by the different theories of firm dynamics.

Although somewhat more speculative given the nature of the French data, we also present evidence that the changes in layer-level outcomes are in general achieved by changing the composition of workers in a layer and not by changing the knowledge or wages of continuing employees. Furthermore, these changes are achieved by actively managing the experience and education of the labor force in a layer. Our results suggest that changes in experience are used to reduce and increase knowledge when firms change layers, while education is essentially kept constant. In contrast, when firms expand without adding layers, they hire more educated employees in virtually all layers of the hierarchy, but primarily more experienced employees at the top.

It is important to be clear about the way in which our results should be interpreted. Our study has identified patterns in the data and clear robust correlations between organizational change and a variety of firm characteristics, like the distribution of employees across layers and their wages. The patterns we uncover are useful to discriminate between theories of firm growth and their implications for labor market outcomes. They establish a set of facts that theories of firms and wages should relate to in the future. What we have not done in this paper is to identify exogenous shocks that lead to reorganization. So our findings cannot be used to compute, say, the size of the effect of an economy wide change in technology on the wages paid by firms, or the number and types of the workers they hire. This analysis could easily be done in a country where such a shock is clearly identifiable.

Moreover, our results have implications for firm and layer-level outcomes, but are silent about how re-organization affects individual workers. This explains our empirical approach, and justifies why we do not focus on individual worker characteristics. Understanding the experience of individual workers as firms grow is an interesting research agenda, but not one we embark on in this paper. Using the approach in this paper, such a study could be done in a country with a matched employer-employee dataset that tracks the universe of employees over time as they move across layers and firms.

Our empirical analysis of firms' organization can be replicated easily for any country that has data on wages by occupation at the firm level. This opens the possibility of studying how firm' organization affects a variety of economic phenomena. Hopefully, our paper will convince others of the importance and relevance of this endeavor.

[24] Calvo and Wellisz (1979) and Qian (1994) present theories where efficiency wages decrease as the span of control is reduced. However, these theories are silent about the way in which growing firms reorganize and, as a result, change the number of employees and wages at each layer.

References

[1] Abowd, J., F. Kramarz, and D. Margolis (1999). "High-Wage Workers and High-Wage Firms," *Econometrica*, 67:2, 251-333.

[2] Baker, G., M. Gibbs, and B. Holmstrom (1994). "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109:4, 921-955.

[3] Baker, G., and B. Holmstrom (1995). "Internal Labor Markets: Too Many Theories, Too Little Evidence," *American Economic Review*, 85:2, 255-259.

[4] Bernard, A. B. and J. B. Jensen, (1997). "Exporters, Skill Upgrading and the Wage Gap," *Journal of International Economics*, 42:1, 3-31.

[5] Bernard, A. B. and J. B. Jensen, (1999). "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, 47:1, 1-25.

[6] Bernard, A. B, J. B. Jensen, S. Redding and P. Schott, (2007). "Firms in International Trade," *Journal of Economic Perspectives*, 21:3, 105-130.

[7] Bloom, N. and J. van Reenen, (2007). "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 122:4, 1351-1408.

[8] Bloom, N., R. Sadun, J. van Reenen, (2012). "The Organization of Firms Across Countries," *Quarterly Journal of Economics*, forthcoming.

[9] Bloom, N., B. Eifert, D. McKenzie, A. Mahajan, and J. Roberts, (2011). "Does Management Matter: Evidence from India," *Quarterly Journal of Economics*, forthcoming.

[10] Brown, C. and J. Medoff (1989). "The Employer Size-Wage Effect," *Journal of Political Economy*, 97:5, 1027-1059.

[11] Caliendo, L. and E. Rossi-Hansberg, (2012). "The Impact of Trade on Organization and Productivity," *Quarterly Journal of Economics*, 127:3, 1393-1467.

[12] Calvo, G. A., and S. Wellisz, (1978). "Supervision, Loss of Control, and the Optimum Size of the Firm," *Journal of Political Economy*, 86:5, 943-952.

[13] Caroli E., and J. van Reenen, (2001)."Skill-Biased Organizational Change? Evidence From A Panel Of British And French Establishments," *Quarterly Journal of Economics*, 116:4, 1449-1492.

[14] Frías, J., D. S. Kaplan and E. A. Verhoogen, (2009). "Exports and Wage Premia: Evidence from Mexican Employer-Employee Data." Working paper.

[15] Garicano, L., (2000). "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108:5, 874-904.

[16] Garicano, L. and T.N. Hubbard, (2007). "Managerial Leverage Is Limited by the Size of the Market: Theory and Evidence from the Legal Services Industry," *Journal of Law and Economics*, 50: 1-44.

[17] Garicano, L. and E. Rossi-Hansberg, (2006). "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, 121:4, 1383-1435.

[18] Lentz, R. and D. T. Mortensen, (2008). "An Empirical Model of Growth Through Product Innovation," *Econometrica*, 76:6, 1317-1373.

[19] Melitz, M. J., (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71:6, 1695–1725.

[20] Oi, W. and Idson, T. (1999). "Firm Size and Wages," *Handbook of Labor Economics: Volume 3*, O. Ashenfelter and D. Card (eds), Elsevier Science: Amsterdam, 2165-2214.

[21] Qian, Y. (1994). "Incentives and Loss of Control in an Optimal Hierarchy," *Review of Economic Studies,* 61:3, 527-544.

[22] Rajan, R. G. and J. Wulf, (2006). "The Flattening Firm: Evidence on the Changing Nature of Firm Hierarchies from Panel Data," *Review of Economics and Statistics*, 88:4, 759-773.

[23] Rosen, S., (1982). "Authority, Control, and the Distribution of Earnings," *The Bell Journal of Economics*, 13:2, 311-323.

[24] Rosen, S., (1986). "The Theory of Equalizing Differences". In Ashenfelter, Orley; Layard, Richard. *The Handbook of Labor Economics 1*. New York: Elsevier. pp. 641-692.