The Anatomy of French Production Hierarchies*

Lorenzo Caliendo
Yale University

Ferdinando Monte
Johns Hopkins University

Esteban Rossi-Hansberg
Princeton University

July 19, 2012

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, output, or markets (by becoming exporters). 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 all pre-existing layers. In contrast, firms that expand little and do not reorganize pay *higher* average wages in all pre-existing layers.

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, including export status. We are particularly interested in understanding if firms actively manage their organization –and therefore the number and knowledge of their employees– and how they do this. Understanding these decisions is important in order to

*We thank Ariel Burstein, Luis Garicano, Bob Gibbons, Gene Grossman, Maria Guadalupe, Gordon Hanson, Oleg Itskhoki, Wolfgang Keller, Sam Kortum, Francis Kramarz, Claire Lelarge, Marc Melitz, Steve Redding, Peter Schott, Chris Sims, and John van Reenen 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.

understand the characteristics of firms and therefore the behavior 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 Garicano (2000) and used in the context of heterogeneous firms in Caliendo and Rossi-Hansberg (2012). In this theory a layer of employees is a group of managers or workers, with similar characteristics and 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. 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 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. 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

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

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 zero 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 one. 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 second-layer managers 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 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 expansion 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.

We go to the data guided by these implications. We look at firms that expand and add layers and firms that contract and drop layers. In particular, we estimate changes in log wages and changes in log hours of work normalized by hours in the top layer, for firms that add or drop layers. The implications discussed above hold for each layer in firms with any number of layers. We do not find any instance in which these implications are falsified by the data. We then look at firms that expand but do not add or drop layers. We find that these firms expand hours and wages at all layers. Note the key finding: *If firms expand by adding layers, wages in preexisting layers fall, while if firms expand without reorganizing, wages rise.* 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

[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).

it shifts up for all percentiles when firms grow without reorganizing. All these results reverse when we focus on firms that contract.

These findings are consistent with the theory in CRH but contradict theories of efficiency wages, or Nash bargaining, in which firms that expand always pay higher wages, independently of organizational changes. Several studies have documented that average wages at firms increase with firm size, or as firms expand: the 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 findings are not inconsistent with these findings. What we document, among other things, is that the size-wage premium is reversed when we condition on the firm reorganizing by adding layers and we focus on workers in preexisting layers.

Next we explore organizational change in firms that become exporters. Exporters are larger, employ more hours of labor, pay higher wages, and have more layers. Firms with more layers are much more likely to be exporters. For example, among firms with three layers of management (the highest number of layers given that they also have a layer of workers) 90.2% of the value added is generated by firms that also export. All of these facts are consistent with the standard finding in the literature that exporters are larger and are also consistent with CRH where larger firms have weakly more layers. Hence, it is perhaps more interesting to turn our attention to new exporters.

We find that new exporters are more likely to add layers than non-exporters (and symmetrically firms that exit exporting are more likely to drop layers). In addition, new exporters that add layers decrease average wages in existing layers while exporters that do not add layers increase them. The well-known finding (see Bernard and Jensen 1995, 1997, and Verhoogen 2008) that firms that become exporters pay higher wages is the result of a composition effect. In fact, the firms that expand significantly as a result of exporting, namely, the ones that add layers, reduce average wages. Furthermore, they do so at all pre-existing layers. In contrast, new exporters that do not change layers barely expand but do increase wages. Since there are more new exporters that do not change layers than there are exporters that do change layers, the average effect on wages is positive but small. The result is relevant for the conceptualization of new exporters. The notion that new exporters expand and increase the wages of their employees either because they upgrade their technology (and so the marginal product of labor is higher) or because profits are higher and they share them with workers (via a wage sharing or bargaining mechanism) is at odds with our data.[3] The data are consistent with a view in which new exporters that expand significantly change their organizational design and economize on knowledge by employing less knowledgeable employees who are paid less.

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.[4] Garicano

[3] Felbermayr, et al. (2008), Egger and Kreickemeier (2009), Helpman, et al. (2010) and Eaton, et al. (2011) propose models were the exporter-wage premium is the outcome of a bargaining mechanism.

[4] Caroli and van Reenen (2001) has the advantage of using a measure of delayering directly reported by managers

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. In a similar study, Guadalupe and Wulf (2010), using the same data as Rajan and Wulf (2006), find evidence that exposure to the export market makes firms reorganize their production. We find that firms that enter the export market are more likely to reorganize production than non-exporters.

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 presents the results for new exporters and firms that exit the export market. Section 6 discusses how firms change average wages in a layer and Section 7 concludes. Appendix A presents a variety of robustness checks and extensions of the results in the main text. 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

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.

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 $\ell = 0$) or managers (layers $\ell \geq 1$). We denote by n_L^ℓ, z_L^ℓ, and w_L^ℓ, the number, knowledge, and total wage of employees at layer $\ell = 0, 1, 2...$ of an organization with L layers of management (or $L + 1$ layers of employees, given that we call the layer of workers layer zero). 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 $F''(z) < 0$. Workers learn how to solve the most frequent problems. The ones in the interval $\left[0, z_L^0\right]$. If the problem they face falls in $\left[0, z_L^0\right]$, 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 $\ell = 1$ tries to solve the problems workers could not solve. Hence, they learn how to solve problems in $\left[z_L^0, z_L^0 + z_L^1\right]$. In general, the firm needs $n_L^\ell = hn_L^0(1 - F(Z_L^{\ell-1}))$ managers of layer ℓ, where $Z_L^\ell = \sum_{l=0}^{\ell} z_L^l$.[5]

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,

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

subject to

$$\begin{aligned} q &\leq F(Z_L^L) n_L^0, && (2)\\ w_L^\ell &= \bar{w}[cz_L^\ell + 1] \text{ for all } \ell \leq L, && (3)\\ n_L^\ell &= hn_L^0[1 - F(Z_L^{\ell-1})] \text{ for } L \geq \ell > 0, && (4)\\ n_L^L &= 1. && (5) \end{aligned}$$

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.

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

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. 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.

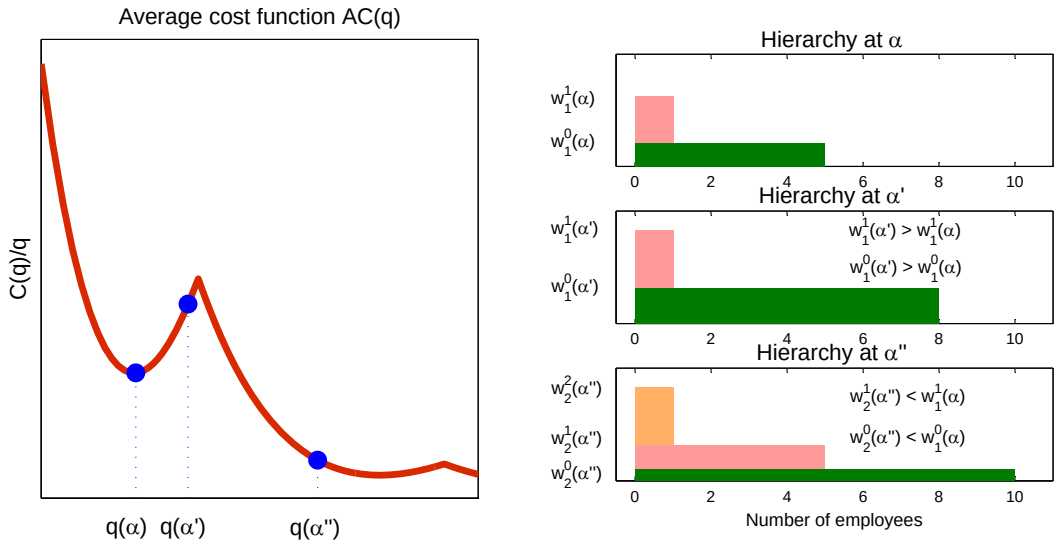


Figure 1: The average cost function $C(q;w)/q$ as a function of q

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 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.

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.

The model in CRH also allows us to study the effect of a new opportunity to export on the organization of firms. Essentially, the possibility to export has effects similar to those of a larger demand draw α. So exporters are larger and therefore have more layers. Furthermore, new exporters that add layers reduce the knowledge and wages of pre-existing layers and increase the number of employees in those layers. New exporters that do not add layers increase wages, knowledge, and the number of employees at all layers.

To sum up, the model has the following implications:

1. Firms are hierarchical, $n_L^0 \geq ...n_L^\ell... \geq n_L^L$ for all L.

2. Layers, L, sales pq, and total number of employees, $\sum_{\ell=0}^{L} n_L^\ell$, *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.

5. Exporters sell more and have more layers than non-exporters.

6. When firms start exporting:
   a) L increases weakly.
   b) Exporters that do not change L *increase* w_L^ℓ and n_L^ℓ at all ℓ.
   c) Exporters that increase L *decrease* w_L^ℓ and *increase* n_L^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 5.9% of the observations and we lose another 11.5% from cleaning the data. The resulting sample covers on average over time 90.7% 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 9.3% of value added not included in our sample. A detailed description of the construction and characteristics of the dataset is included in 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.[6] 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 percentiles of the distribution of wages in the different classes of workers, all expressed in 2005 euros.[7]

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

	CEO, directors	Senior staff	Supervisors	Clerks	Blue collars
Mean	75.60	47.91	26.30	19.06	20.83
p5	22.33	19.56	13.14	9.74	9.67
p10	26.99	23.07	15.01	11.00	10.94
p25	38.86	28.93	18.14	13.03	12.88
p50	54.62	35.96	21.87	15.63	15.24
p75	75.07	44.62	26.44	19.01	18.34
p90	106.04	56.95	32.76	23.73	22.58
p95	132.17	69.01	38.94	28.33	26.94

[6] Class 1 is only used for farmers and so is never present in our data.
[7] Throughout the paper all nominal variables are expressed in 2005 euros.

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 workers, namely, layer zero. Class 4 of supervisors will form the first layer of managers, layer 1. Senior staff will be included in layer 2, and CEOs and firm directors will form layer 3. So firms can have a maximum of 4 layers, starting with layer zero and moving all the way up to layer 3. Throughout the rest of the paper we refer to the number of layers in the firm by the number of management layers. So a firm that has a layer of workers and two layers of managers is referred to as a firm with two layers.

3.1 Firms with a different number of layers are different

We aim to establish that this classification of workers 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: Data description by year

		Average			
Year	Firms	VA	Hours	Wage	# of layers
2002	79,260	2,909	83,749	22.07	1.59
2003	77,768	2,900	82,704	22.51	1.58
2004	76,448	2,934	81,445	23.33	1.58
2005	75,426	2,786	78,090	24.07	1.55
2006	74,818	2,823	77,697	23.29	1.53
2007	72,918	2,685	73,515	23.98	1.50

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

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 1.5 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 1.5 layers of management and that they can have a maximum of three layers of management (plus one layer of workers), there are ample opportunities for firms to add new layers.

Table 3: Data description by number of layers in the firm					
# of		Average			Median
layers	Firm-years	VA	Hours	Wage*	wage
0	81,909	205	7,946	29.44	17.08
1	126,069	403	16,450	20.96	17.99
2	161,449	2,821	85,674	20.99	19.45
3	87,211	8,879	227,070	22.23	20.78

Value added in 000s of 2005 euros. Wage* = average hourly wage in 2005 euros from trimmed sample (0.05% top and bottom in each cell).

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 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.

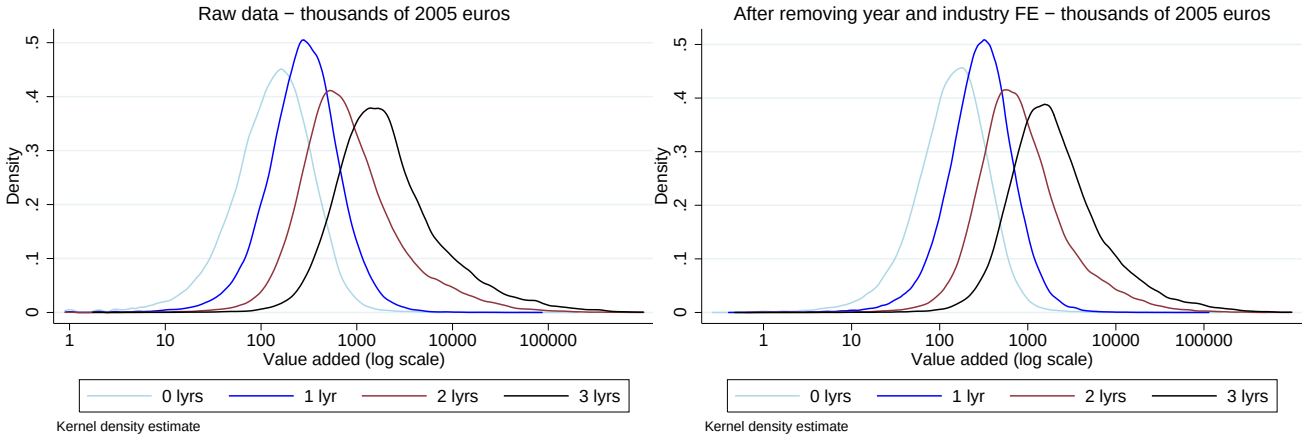


Figure 2: Value added distribution by number of layers

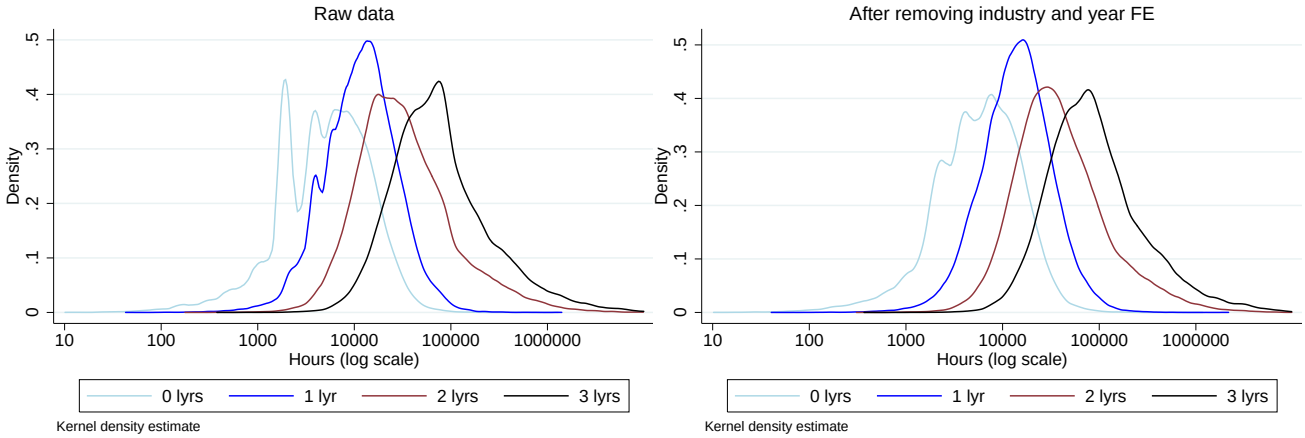


Figure 3: Hours distribution by number of layers

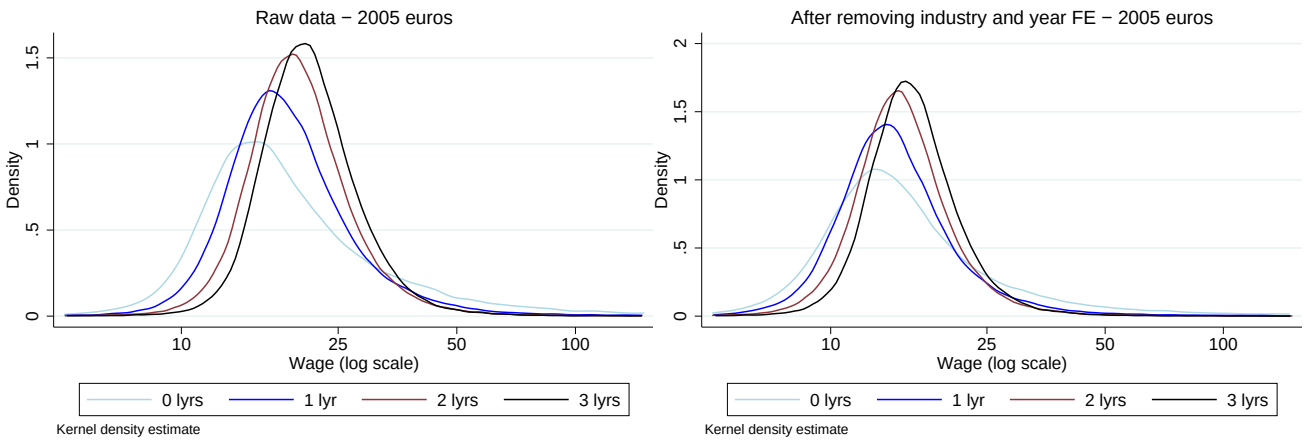


Figure 4: Firm average hourly wage distribution by number of layers

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 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 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 zero. That is, 87.17% of the firms in our sample that have zero layers of management actually have layer zero. These are also the largest firms in terms of employment, as they account for 95.86% of the total employment of firms with zero layers. As another example, consider the firms with two management layers. About 80% of them include layers zero to two. The remaining 20% are missing one of these layers and include layer three. 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 zero are the smallest firms in the sample. Clearly, all firms with three management layers have adjacent layers starting from zero since they have all layers. Overall, 81.57% of firms have adjacent layers starting at zero and they account for 96.65% 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 Appendix A.

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 first layer of management, which is larger than the second layer of management, 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 A39 in Appendix A presents the averages weighted by value added.

Table 4: Percentage of firms that have adjacent layers

Among firms with:	0 layers	1 layer	2 layers	3 layers	All firms
Unweighted	87.17	67.22	79.98	100	81.57
Weighted by VA	85.84	68.01	94.54	100	96.65
Weighted by hours	95.86	72.38	93.15	100	95.74

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 one layer 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 three layers of management are hierarchical in all the rankings in about 80% of cases. Table A40 in Appendix A presents the averages weighted by value added.

Table 5: Firms that satisfy a hierarchy in hours

# of layers	$N_L^\ell \geq N_L^{\ell+1}$ all ℓ	$N_L^0 \geq N_L^1$	$N_L^1 \geq N_L^2$	$N_L^2 \geq N_L^3$
1	85.3%	85.3%	-	-
2	62.0%	85.2%	74.0%	-
3	54.3%	85.8%	76.4%	86.6%

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

Table 6: Firms that satisfy a hierarchy in wages

# of layers	$w_L^{\ell+1} \geq w_L^\ell$ all ℓ	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	92.1%	92.1%	-	-
2	86.2%	93.6%	92.5%	-
3	79.7%	96.5%	94.4%	87.8%

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 firms with different numbers of layers. Each layer is represented using a square. The length of the square represents the average number of hours employed in the layer by firms with a given number of layers. The height of the square 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.

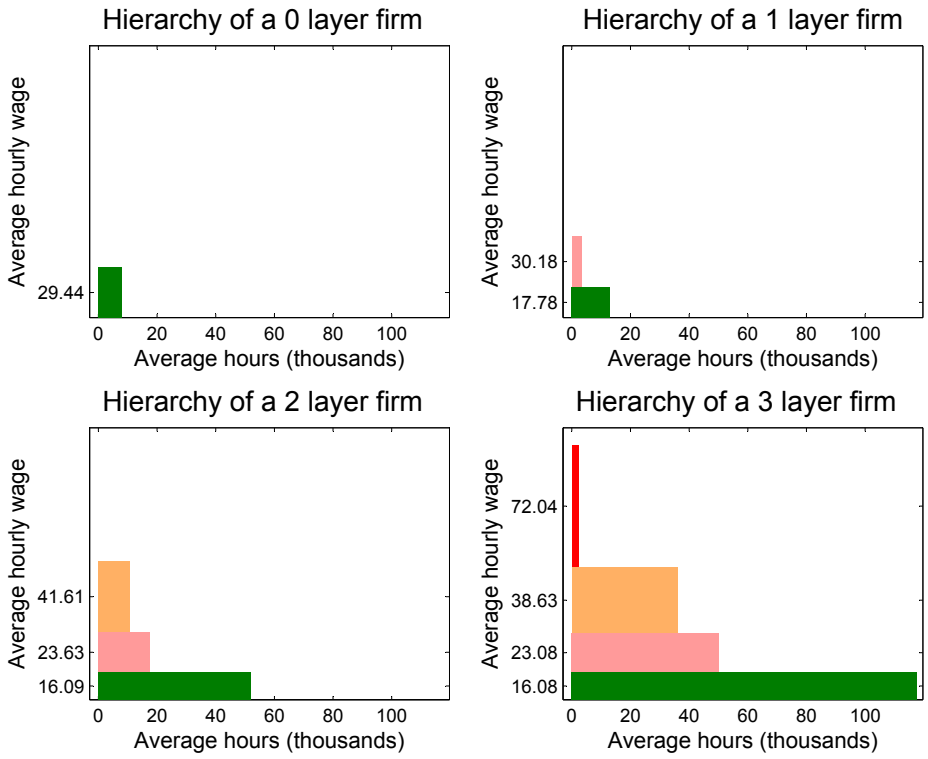


Figure 5: Representative hierarchies

Figure 6 presents the same representative hierarchies when 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 in Figure 6 after normalizing.

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 zero layers of management (since, by definition, for these firms there is no cross-layer variation in log wages) and grows

to 66% for firms with three layers of management. The table reassures us that variation in wages across layers is essential to understand the distribution of wages within firms.

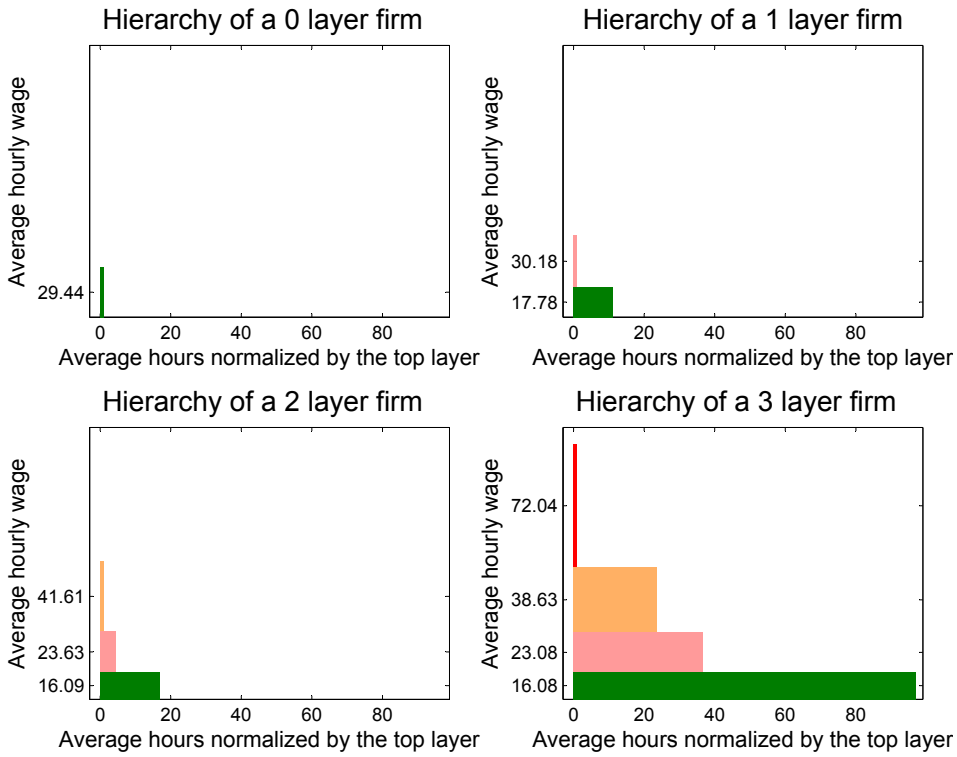


Figure 6: Representative hierarchies normalized by hours in the top layer

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

			Weighted by	
	Firm-years	Unweighted	Hours	VA
All firms	440,443	0.50	0.51	0.50
Firms with more than 0 layers	374,568	0.59	0.52	0.50
Firms with 0 layers	65,875	0.00	0.00	0.00
Firms with 1 layer	125,911	0.50	0.41	0.43
Firms with 2 layers	161,447	0.62	0.52	0.50
Firms with 3 layers	87,210	0.66	0.63	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 (75.4% for zero-layer firms and 82.9% for firms with one layer of management; see Table A2 in 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 Appendix A shows that the same pattern as in Table 8 is observed even if we weight the firms by their value added.

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

		# of layers at $t+1$					
		Exit	0	1	2	3	Total
# of layers at t	0	15.4	67.1	15.3	2.0	0.2	100
	1	9.9	10.8	62.0	16.2	1.1	100
	2	7.6	1.2	13.2	67.5	10.5	100
	3	6.1	0.2	2.0	20.5	71.2	100

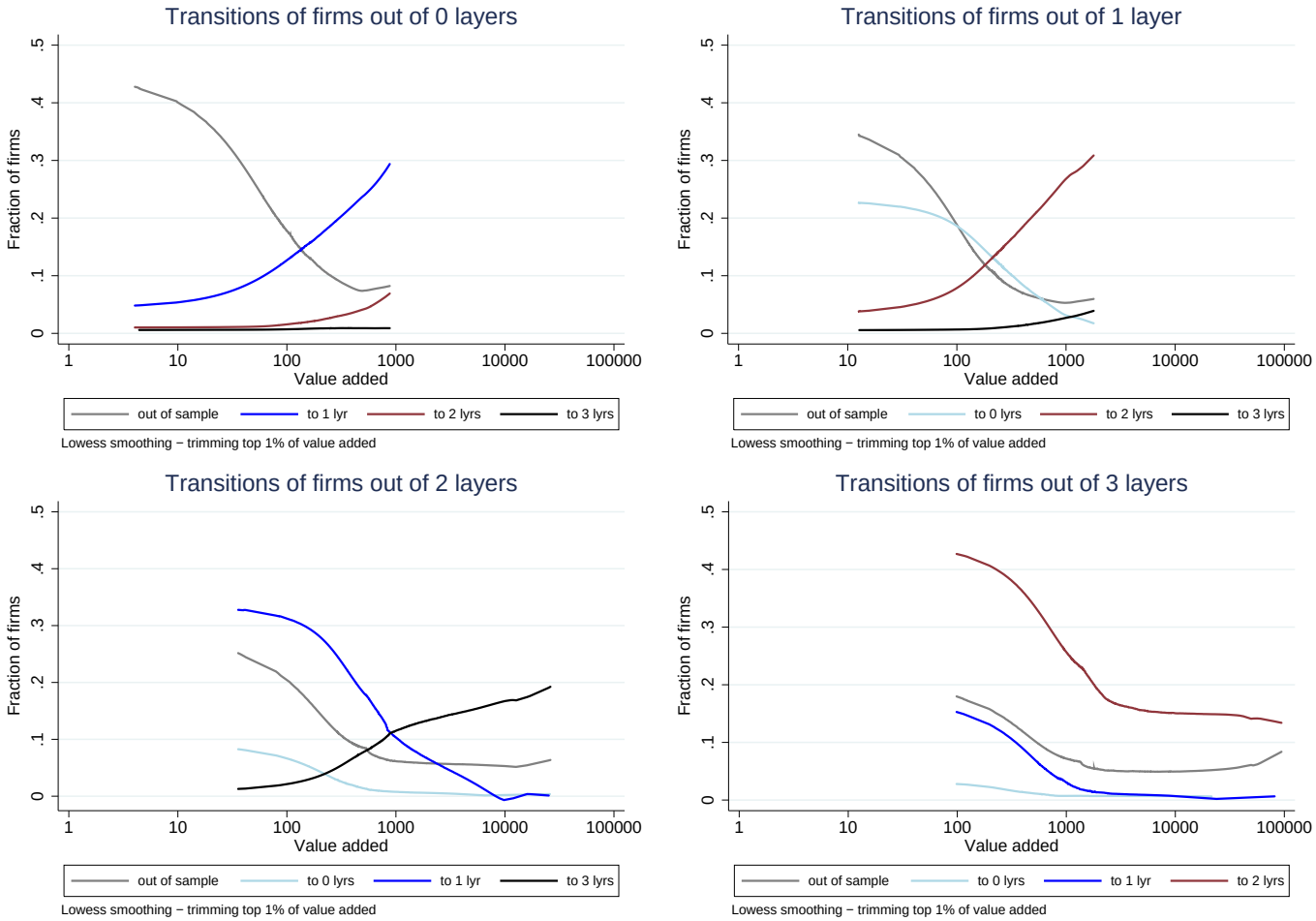


Figure 7: Transitions across layers depend on value added

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 7 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 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 7.

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 7 does not include confidence bands in order to enhance the visibility of the curves. However, in Figure A1 in Appendix A we present a graph with all the individual observations and show that they line up fairly tight around the interpolation estimates. Appendix B describes all the details and choices made to construct Figure 7.

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 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. However, before we present our results over time, it is informative to analyze briefly the cross-section of firms. In particular, we study the relationship between value added and normalized hours in a given layer, ℓ, for firms with L layers, n_L^ℓ, or the wage of the workers in a given layer ℓ, w_L^ℓ. The theory predicts that if we focus on firms with a given number of layers, L, n_L^ℓ, and w_L^ℓ should increase with value added absent any other source of heterogeneity across producers. Of course, firms are different in many dimensions and so instead of a one-to-one relationship we find substantial heterogeneity in n_L^ℓ and w_L^ℓ for any given level of value added. Nevertheless, Figures A3 to A9 in Appendix A show a positive and significant relationship between log value added and the log of n_L^ℓ and w_L^ℓ for every $\ell \in 0, 1, ..., L$ and every L.

The theory also has a more subtle prediction. Namely, that the slope of the relationship between the log of value added and normalized hours should decrease as we consider higher layers of management. 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. Figures A3 to A5 in Appendix A, together with Table 9, show that the ranking of the slopes is the one predicted by the model for firms with two management layers but not for firms with three layers (where the ranking is reversed). The differences between the elasticities are significant at the 1% level for all comparisons. The table also shows the estimates when we add time and industry fixed effects (again see Appendix B for details). The failure of the implication of the model for firms with three layers is the result of cross-sectional heterogeneity, as we show below.

Table 9: Normalized hours within layers

	Without FE		With FE	
Normalized hours	$\ln VA$	s.e.	$\ln VA$	s.e.
$\ln n_1^0$	.411	.005	.349	.005
$\ln n_2^0$	.186	.003	.158	.003
$\ln n_2^1$	.100	.002	.096	.003
$\ln n_3^0$	.663	.004	.669	.003
$\ln n_3^1$	.734	.003	.719	.003
$\ln n_3^2$	.773	.004	.765	.004

All p-values < 0.01. Excluding top and bottom 0.05%

Table 10: Wages within layers

	Without FE		With FE	
Wages	$\ln VA$	s.e.	$\ln VA$	s.e.
$\ln w_0^0$	.039	.002	.072	.004
$\ln w_1^0$	.051	.002	.059	.002
$\ln w_1^1$	.083	.002	.086	.002
$\ln w_2^0$	.040	.001	.045	.001
$\ln w_2^1$	.039	.001	.043	.001
$\ln w_2^2$	.068	.001	.073	.001
$\ln w_3^0$	.045	.001	.048	.001
$\ln w_3^1$	.043	.001	.046	.001
$\ln w_3^2$	.072	.001	.072	.001
$\ln w_3^3$	.192	.002	.192	.002

All p-values < 0.01. Excluding top and bottom 0.05%

Similarly, we can look at the ranking of the slopes of the relationship between log wages and log value added. Now the model predicts that the elasticity should be higher, the higher the management layer. The firm economizes by expanding the knowledge of higher layers more, since there are fewer of those managers. This is in fact the case in the majority of comparisons, except for a couple that are not significant, as is evident in Table 10. Thus, among firms with the same number

of layers, the larger ones pay higher wages, and proportionally higher at the top of the hierarchy. Note that these comparisons are treating all the sources of heterogeneity across firms that are not accounted for by value added as classical measurement error. This is certainly a stretch, so it is somewhat surprising that these predictions still emerge so clearly in our data.

The evidence above is informative about the differences across firms with similar characteristics, in particular, the same number of layers. However, it is confounded by other dimensions of heterogeneity across firms. We now move to time series evidence that tracks firms over time. Let us start by looking at the same prediction we analyzed in the cross-section. That is, we study how n_L^ℓ and w_L^ℓ change as the firm increases its value added. We estimate a regression of the form

$$d \ln n_{Lit}^\ell = \alpha_L^\ell + \beta_L^\ell d \ln VA_{it} + \varepsilon_{it}, \tag{6}$$

where i refers to a particular firm with L management layers, t is the time period, and d is a yearly time difference (e.g., $d \ln n_{Lit}^\ell = \ln n_{Lit+1}^\ell - \ln n_{Lit}^\ell$).

Table 11: Elasticity of n_L^ℓ with VA for firms that do not change L

# of layers	Layer	β_L^ℓ	s.e.	p-value	obs
1	0	0.044	0.012	0.00	65,114
2	0	0.046	0.009	0.00	91,833
2	1	0.019	0.010	0.07	91,833
3	0	0.109	0.014	0.00	53,053
3	1	0.048	0.013	0.00	53,053
3	2	0.037	0.013	0.01	53,053

Table 11 presents the estimates of β_L^ℓ for all $\ell \in 0, 1, ..., L$ and every L. Note that since we are 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, now 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. 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 A42 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. In particular, we estimate

$$d \ln w_{Lit}^\ell = \delta_L^\ell + \gamma_L^\ell d \ln VA_{it} + \varepsilon_{it}, \tag{7}$$

where now $d \ln w_{Lit}^\ell$ denotes the log difference in wages. The results are presented in Table 12 and are, as were the cross-sectional results, for the most part consistent with the theory. Namely, γ_L^ℓ is

positive and significant for all $\ell \in 0, 1, ..., L$ and every L. Furthermore, $\gamma_L^\ell < \gamma_L^{\ell'}$ for $\ell < \ell'$, in most cases, apart from a couple of instances where the difference is not significant. Hence, we reinforce the conclusion that 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. Table A37 in Appendix A presents several robustness checks.

Table 12: Elasticity of w_L^ℓ with VA for firms that do not change L

# of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
0	0	0.077	0.007	0.00	45,606
1	0	0.098	0.006	0.00	65,114
1	1	0.116	0.006	0.00	65,114
2	0	0.145	0.006	0.00	91,833
2	1	0.156	0.006	0.00	91,833
2	2	0.172	0.006	0.00	91,833
3	0	0.173	0.009	0.00	53,053
3	1	0.187	0.009	0.00	53,053
3	2	0.189	0.010	0.00	53,053
3	3	0.218	0.011	0.00	53,053

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 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 13 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 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 13.7%. The results are reversed when we select only firms that drop layers. Now wages rise by 13.6%.

These estimates demonstrate that in firms that expand by adding layers, wages in pre-existing layers fall. As far as we know, this is inconsistent with most theories in the labor and firm dynamics literatures. In these theories firms that expand always increase the wages they pay. 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 layers should decrease.

Table 13: Change in firm-level outcomes

	All	Increase L	No change in L	Decrease L
$d\ln$ total hours	-0.014***	0.056***	-0.011***	-0.093***
- detrended	-	0.070***	0.003***	-0.079***
$d\ln\sum_{\ell=0}^{L} n_L^\ell$	-0.011***	1.366***	0.012***	-1.408***
- detrended	-	1.377***	0.023***	-1.396***
$d\ln VA$	-0.008***	0.032***	-0.007***	-0.049***
- detrended	-	0.039***	0.001	-0.040***
$d\ln$ avg wage	0.018***	-0.001	0.018***	0.038***
- detrended	-	-0.020***	-0.000	0.020***
- common layers	0.020***	-0.117***	0.018***	0.156***
- - detrended	-	-0.137***	-0.002***	0.136***
% firms	100	12.75	73.48	13.78
% VA change	100	39.21	65.65	-4.87

*** significant at 1%.

The results in Table 13 suggest that wages in firms that add or drop layers behave differently than previously thought. Table 13 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 12.75% of the total. Furthermore, as they are on average larger than their counterparts that do not add layers, they contribute 39.2% of the total change in value added. Inversely, firms that drop layers represent 13.78% of firms and contribute -4.87% 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.

Table 14: $d \ln n^{\ell}_{Lit}$ for firms that transition						
# of layers		Layer	$d \ln n^{\ell}_{Lit}$	s.e.	p-value	obs
Before	After					
0	1	0	1.520	0.017	0.00	10432
0	2	0	1.745	0.053	0.00	1350
0	3	0	2.312	0.193	0.00	111
1	0	0	-1.585	0.017	0.00	11356
1	2	0	0.710	0.012	0.00	17052
1	2	1	0.533	0.012	0.00	17052
1	3	0	1.218	0.048	0.00	1168
1	3	1	1.018	0.047	0.00	1168
2	0	0	-1.801	0.046	0.00	1698
2	1	0	-0.696	0.012	0.00	17927
2	1	1	-0.537	0.012	0.00	17927
2	3	0	1.338	0.014	0.00	14228
2	3	1	1.277	0.016	0.00	14228
2	3	2	1.167	0.016	0.00	14228
3	0	0	-2.203	0.157	0.00	142
3	1	0	-1.112	0.041	0.00	1493
3	1	1	-0.948	0.039	0.00	1493
3	2	0	-1.427	0.014	0.00	15303
3	2	1	-1.359	0.015	0.00	15303
3	2	2	-1.274	0.015	0.00	15303

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 14 computes average log changes in 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 $d \ln n^{\ell}_{Lit}$. 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. So Table 14 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.[8] 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 15: $d \ln w_{Lit}^{\ell}$ for firms that transition

# of layers		Layer	$d \ln w_{Lit}^{\ell}$	s.e.	p-value	obs
Before	After					
0	1	0	-0.131	0.005	0.00	10432
0	2	0	-0.432	0.024	0.00	1350
0	3	0	-0.943	0.131	0.00	111
1	0	0	0.201	0.005	0.00	11356
1	2	0	-0.041	0.003	0.00	17052
1	2	1	-0.245	0.004	0.00	17052
1	3	0	-0.165	0.018	0.00	1168
1	3	1	-0.416	0.02	0.00	1168
2	0	0	0.489	0.022	0.00	1698
2	1	0	0.085	0.003	0.00	17927
2	1	1	0.275	0.004	0.00	17927
2	3	0	-0.008	0.002	0.00	14228
2	3	1	-0.054	0.003	0.00	14228
2	3	2	-0.185	0.004	0.00	14228
3	0	0	1.102	0.12	0.00	142
3	1	0	0.188	0.014	0.00	1493
3	1	1	0.417	0.017	0.00	1493
3	2	0	0.029	0.002	0.00	15303
3	2	1	0.060	0.003	0.00	15303
3	2	2	0.153	0.004	0.00	15303

We do the same analysis layer by layer for changes in wages. Namely, we compute the average (over i and t) of $d \ln w_{Lit}^{\ell}$ for firms that add or drop layers. We present the results in Table 15. Again, the table confirms that the results we obtained for firm-level outcomes hold layer by layer,

[8] 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 14 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.

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 Appendix A and the description of the tables are presented in 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 15.

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 zero to one layer of management, the new manager in layer one makes 53.1% more than the average wage in the firm before reorganization. This number can be much higher for larger firms. Firms that go from 2 to 3 layers pay the new top manager 335.7% more than the average worker in the firm before adding the layer. All of these results are presented in Table 16. We decompose the 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 16 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 15. 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.

Table 16: Decomposition of total log change in average wages

$\bar{w}^{\ell\leq L}_{L'it+1}/\bar{w}_{Lit}$				$w^{L'}_{L'it+1}/\bar{w}_{Lit}$			
From/to	1	2	3	From/to	1	2	3
0	0.975** (10,422)	0.838** (1,348)	0.679** (111)	0	1.531** (10,421)	1.435** (1,349)	1.461** (110)
1		0.940** (17,036)	0.886** (1,167)	1		2.067 ** (17,035)	2.034** (1,167)
2			0.974 (14,214)	2			4.357** (14,213)
s				$d\ln\bar{w}_{Lit}$			
From/to	1	2	3	From/to	1	2	3
0	0.741 ** (10,422)	0.621** (1,350)	0.572** (111)	0	−0.008* (10,421)	−0.195** (1,350)	−0.589** (111)
1		0.853 ** (17,036)	0.775** (1,167)	1		0.014 ** (17,035)	−0.050** (1,167)
2			0.947 ** (14,214)	2			0.013 ** (14,212)

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

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.

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 zero. 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.

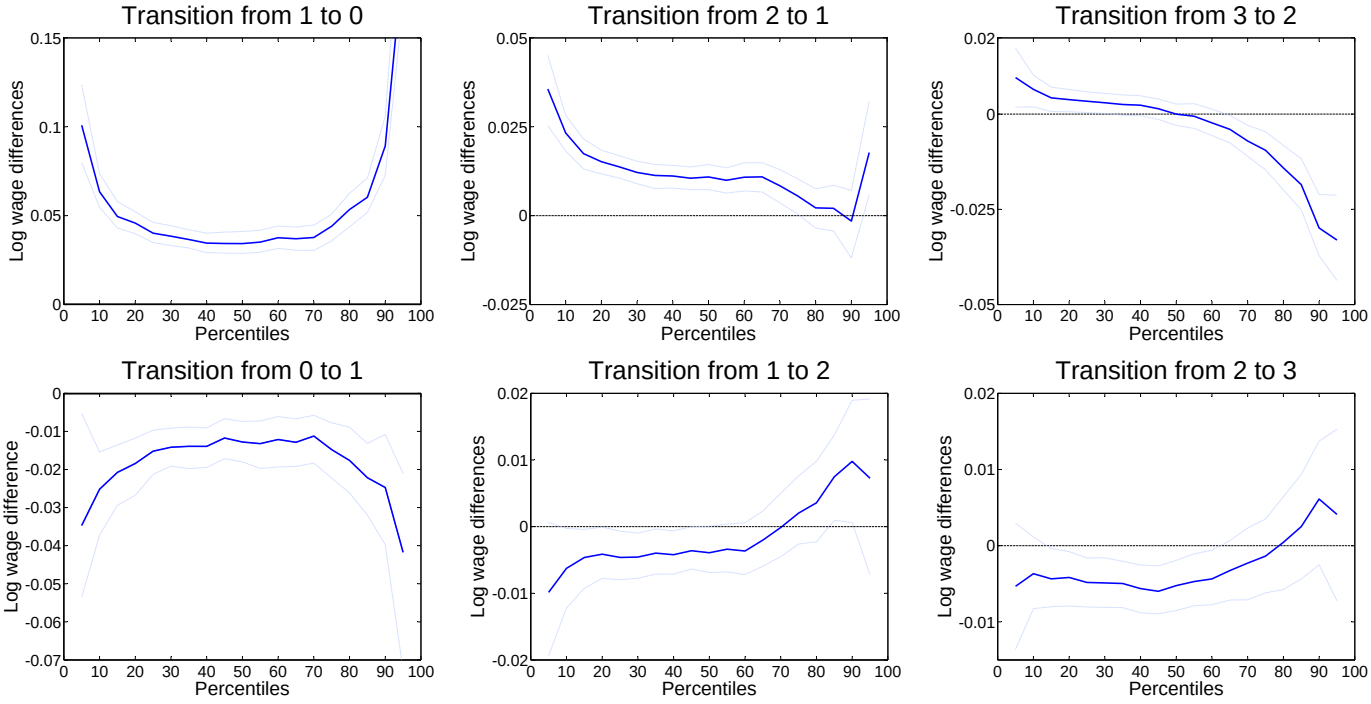


Figure 8: Difference in the distribution of wages after minus before the transition

Figure 8 presents the difference in the log wages at each percentile of the distribution for the different one-layer transitions.[9] The figure also shows a bootstrapped 95% confidence interval constructed as we explain in 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 16 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 15, that indicate that average wages in the upper pre-existing layers are the ones that fall the most when firms add a layer.

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

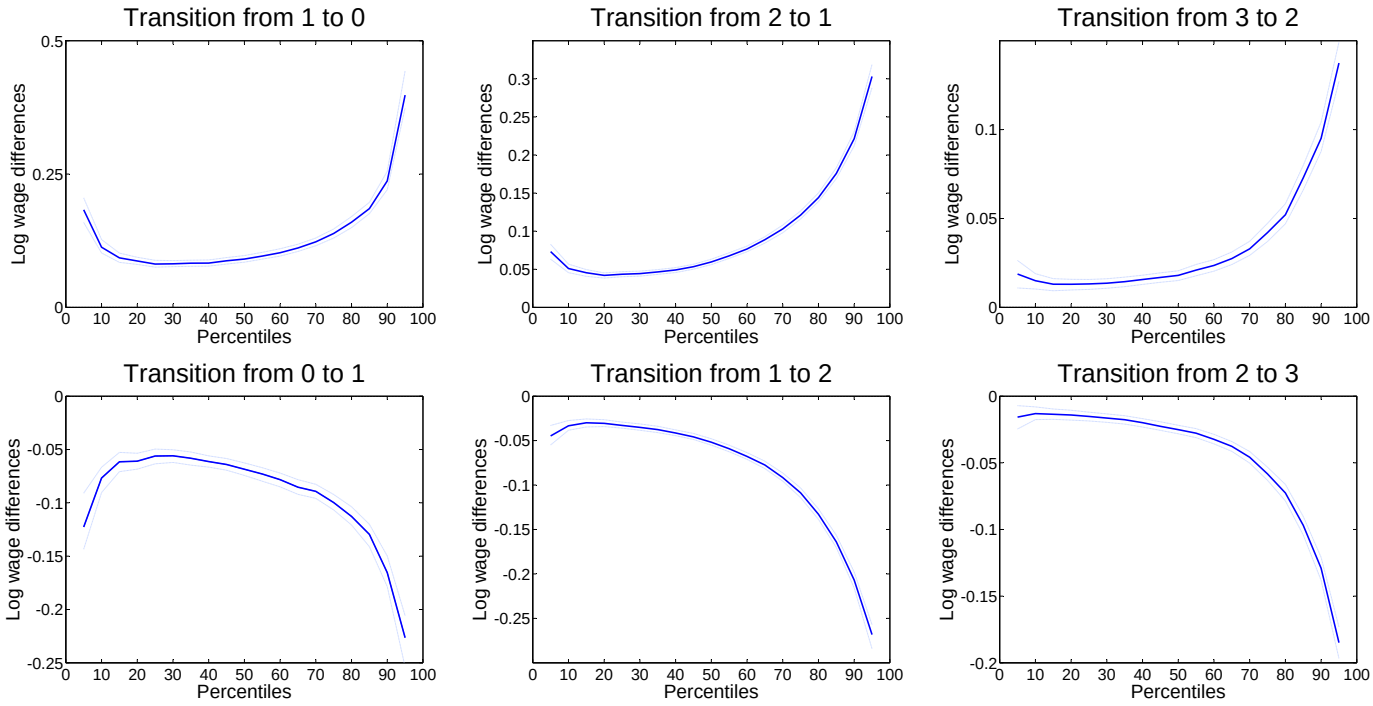


Figure 9: Difference in the distribution of wages in pre-existing layers after minus before the transition

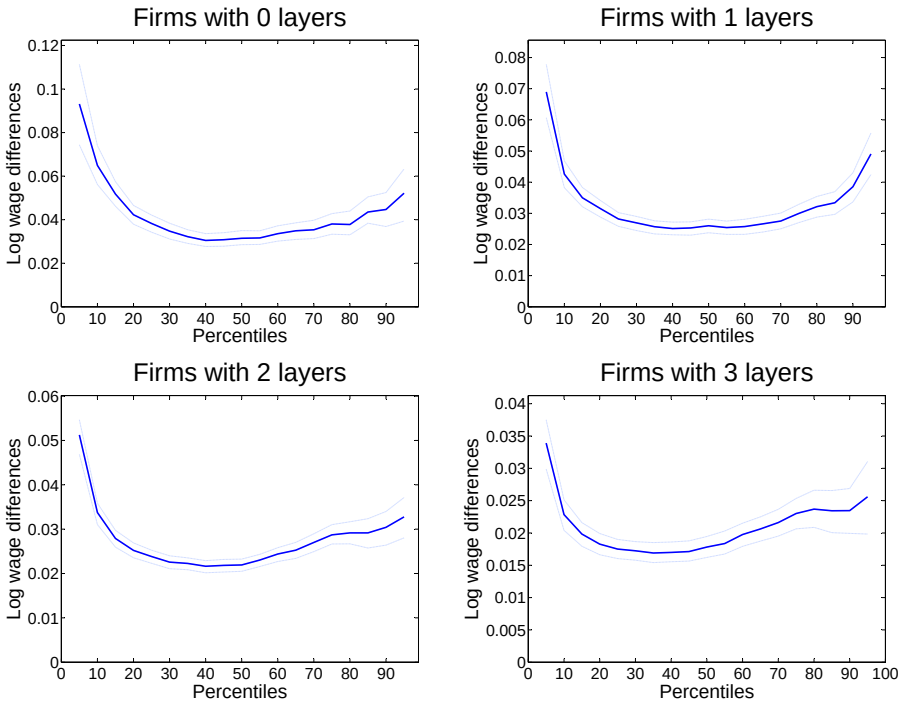


Figure 10: Difference in the distribution of wages for firms that do not transition conditioning on $d \ln VA > 0$

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.[10] 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.[11] 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.

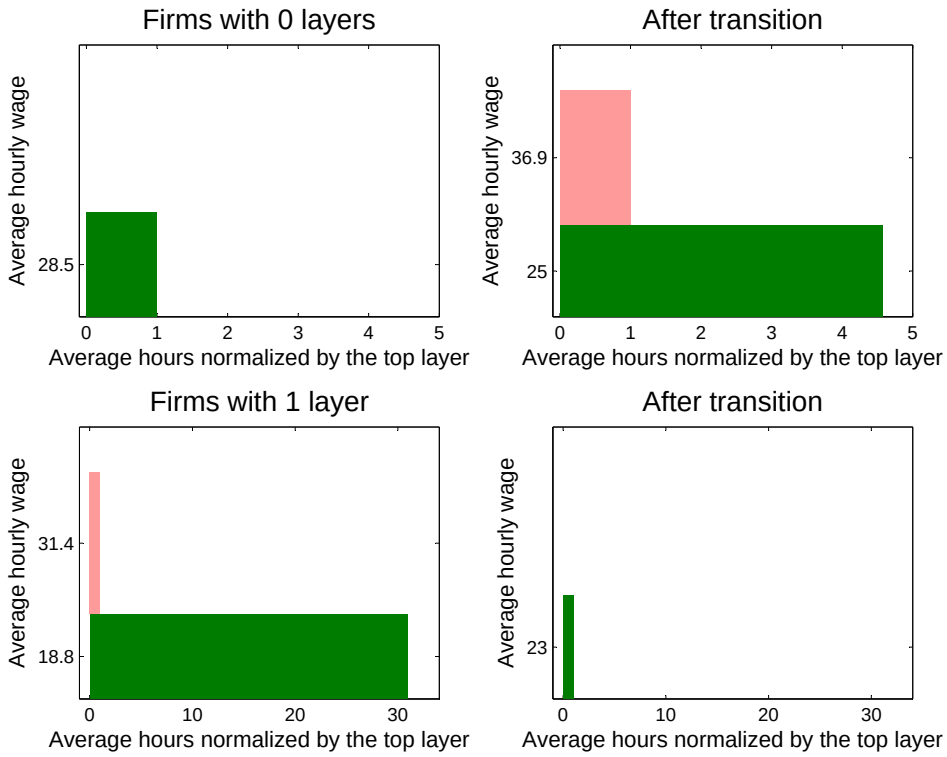


Figure 11: Representative transitions between firms with 0 and 1 layers

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 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.

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

[11] In Figure A12 in 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 2 and 3 layers and for most percentiles in firms with 0 and 1 layers. In the case of 0 and 1 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.

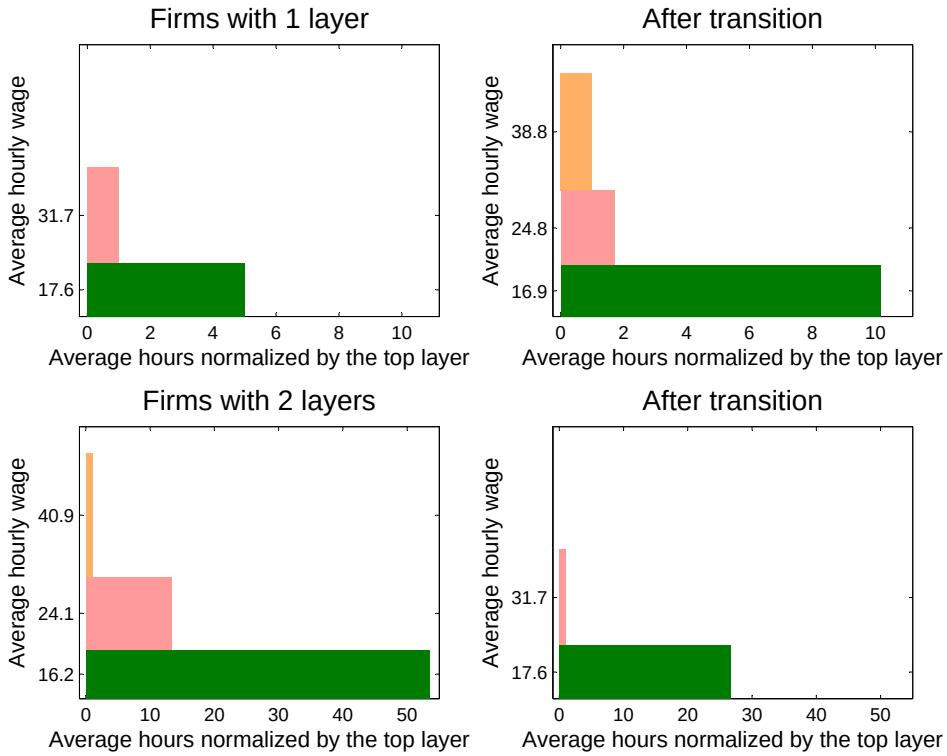


Figure 12: Representative transitions between firms with 1 and 2 layers

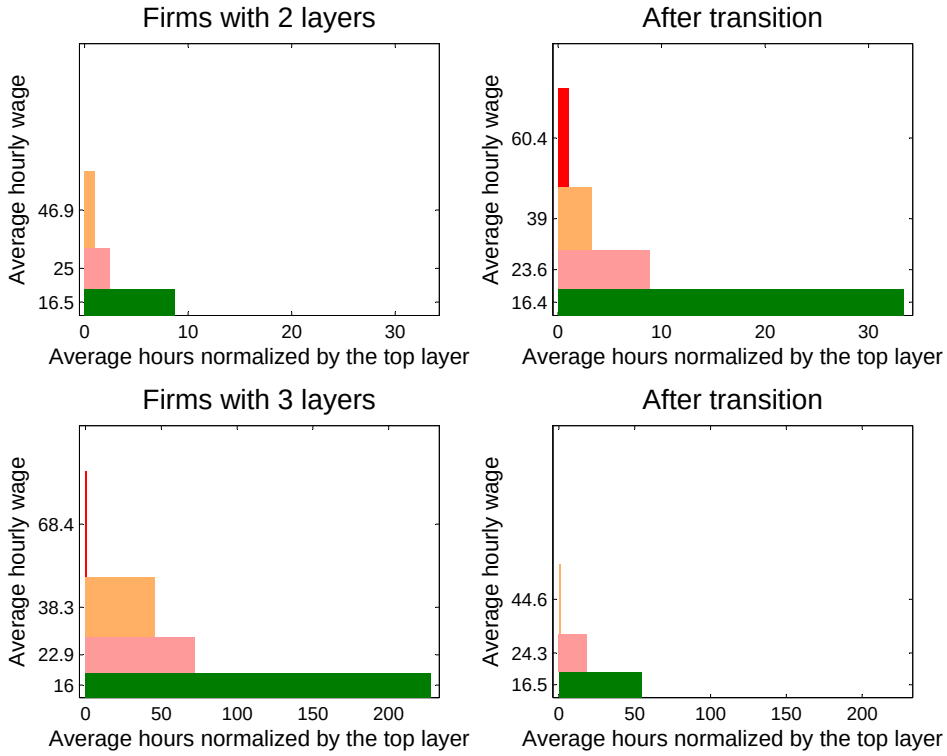


Figure 13: Representative transitions between firms with 2 and 3 layers

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 A10 and A11 in Appendix A. In the next section we explore a particular form of a firm's expansion and its resulting reorganization. Namely, we study firms that start to export goods to serve foreign markets.

5 Reorganizing to Export

It is well known by now that exporters are larger in terms of value added and employment (see Bernard and Jensen, 1999, and Bernard, et al., 2007, among others). This is clearly the case in our data as well.[12] They also pay slightly higher wages. As Table 17 shows, they have more layers of management as well. The average number of layers of management among non-exporters is 1.23, meaning that the average exporter has a layer of workers, a layer of management and a fraction of a second layer of management. If we look at exporters, they have 1.96 layers of management on average. Hence, Implication 5 of the theory in Section 2 is confirmed by the data.

Table 17: Description of exporters

	Average			
	VA	Hours	Wage	# of layers
Non-exporters	800	29,026	23.03	1.23
Exporters	5,343	141,615	23.39	1.96

See the footnote in Table 2. The difference in wages is significant at 1%.

About 44% of the firms in our data export, and they account for slightly more than 83% of value added, with some variation across years. More relevant for our purposes is that the firms that export tend to have more layers. As Table 18 shows, of the firms with three layers of management, 71.6% of them export, while for firms with only workers, only 18.9% of them export.

Table 18: Share of exporters by number of layers

# of layers	Unweighted	Weighted by VA
0	18.9%	25.2%
1	31.7%	38.3%
2	54.1%	82.1%
3	71.6%	90.2%

These numbers are expected given that from Section 3 we know that larger firms have more layers, and from Table 17 and the vast literature on exporting, we know that larger firms export.

[12] Part of our data is used in Eaton, et al. (2011) to study the exporting behavior of firms. As a result, some of these facts for France are known from their paper. However, they have no results on layers or firm reorganization conditional on changing or keeping constant the number of layers.

Table 19 presents the composition of firms by number of layers. Out of all exporters, only 7.5% have only a layer of workers, while 42.6% have two layers of management, and 30.4% have three. So, as shown in Table 19, there is substantial heterogeneity in the number of layers of exporters and non-exporters. Furthermore, most exporters have many layers, while most non-exporters have only one or two.

Table 19: Composition of firms by number of layers

# of layers	Non-exporters	Exporters
0	26.4%	7.5%
1	34.3%	19.5%
2	29.4%	42.6%
3	9.9%	30.4%
Total	100%	100%

We now focus on firms that become exporters during the period in our sample: new exporters. New exporters are more likely to add layers than non-exporters. Table 20 shows that the probability of adding one or more layers for new exporters is significantly higher than for non-exporters, regardless of the initial count of layers. The probability of keeping the same number of layers goes down if the firm has zero or one layer of management, while the probability of keeping the same number of layers increases for firms with two or three layers. Given that firms with three layers cannot add layers, this is natural. We conclude from this evidence that new exporters tend to add layers, consistent with Implication 6a of the theory. Of course, there are some that also drop layers, but there are fewer of those firms than among firms that do not start to export. Table A5 in Appendix A shows that firms that exit the export market are more likely to drop layers than exporters that do not exit.

Table 20: Layer transitions for new exporters relative to non-exporters

		# of layers at $t+1$			
		0	1	2	3
# of layers at t	0	-9.43	6.61	2.31	0.51
	1	-2.57	-3.49	5.29	0.77
	2	-0.87	-4.83	2.84	2.87
	3	-0.18	-2.20	-2.45	4.83

All significant at 1%.

In fact, the new exporters that add layers expand on average much more than the ones that do not reorganize. Table 21 shows the changes in hours, normalized hours, value added and average wages for all new exporters, the ones that add layers, and the ones that do not change L. The table is similar to Table 13, which we analyzed in the previous section, but it limits the sample

to new exporters only. Table A4 in Appendix A presents the results for firms that exit the export market. As in Table 13 we present results when we detrend using trends for all firms in the economy (not only new exporters, of course). Firms that start exporting increase value added on average by 3.8%. The ones that add layers increase value added by much more, 11.6%, while the ones that do not change layers increase value added by only 3.3%. We find similar numbers for hours and normalized hours. Namely, new exporters that add layers expand much more than firms that do not add layers.Now let's look at wages. After detrending, new exporters pay wages similar to those paid before; so do firms that do not change layers. In contrast, firms that change layers decrease wages by a significant 1.8%. Perhaps more relevant is that this average change masks a composition effect between the new top manager and pre-existing layers. When we focus on wages of employees in pre-existing layers, we find that wages fall by 11.9% in firms that add layers (13.9% if we detrend), while they increase by 2.1% in firms that do not change layers (although the change is insignificant when we detrend).

Table 21: Behavior of firms that enter the export market

	All	Increase L	No change in L
$d\ln$ total hours	0.021***	0.126***	0.015***
- detrended	0.035***	0.141***	0.029***
$d\ln\sum_{\ell=0}^{L} n_L^\ell$	0.008	1.237***	0.024***
- detrended	0.019**	1.248***	0.035***
$d\ln VA$	0.038***	0.116***	0.033***
- detrended	0.046***	0.125***	0.041***
$d\ln$ avg wage	0.018***	0.000	0.021***
- detrended	-0.000	-0.018**	0.003
- common layers	0.018***	-0.119***	0.021***
- - detrended	-0.002	-0.139***	0.001
% firms	100	14.62	70.61
% VA change	100	18.62	73.66

** significant at 5%, *** significant at 1%.

The results paint a picture consistent with the one presented in the previous section. New exporters that reorganize reduce wages in pre-existing layers. Furthermore, these are the new exporters that actually expand significantly. The firms that add layers account for 14.62% of new exporters and 18.62% of the total increase in value added by new exporters. In sum, many firms expand little when they become exporters; these firms increase the salaries of all their employees. Some firms expand a lot when they start to export. They reorganize, add layers, and pay lower wages to employees in the pre-existing layers and higher than average wages to the new top manager.

We now proceed to verify that these results hold layer by layer. We look first at firms that do not add layers. As in the previous section, we estimate the regression in equation (6) for the change

in normalized hours. The only difference is that now we use only firms that either start to export or stop exporting in the year in which we measure the change in normalized hours. The results for β_L^ℓ are presented in Table 22. Many of these estimates are not significant. The ones that are are positive as predicted by the theory. As we showed using Table 21, the firms that do not add layers expand very little, so it is hard to estimate β_L^ℓ precisely enough to have significant results.

Table 22: Firms that change export status and do not change L

# of layers	Layer	β_L^ℓ	s.e.	p-value	obs
1	0	-0.011	0.035	0.76	6,968
2	0	0.017	0.024	0.47	10,507
2	1	-0.015	0.027	0.58	10,507
3	0	0.200	0.053	0.00	4,896
3	1	0.073	0.038	0.06	4,896
3	2	0.084	0.042	0.05	4,896

The results for wages using equation (7) for the sample of firms that change export status are presented in Table 23. Now the estimates are much more significant and robust. All the values of γ_L^ℓ are positive and significant and they tend to increase with ℓ given L. The ranking of the elasticities is not always significant, but it is in most cases, and when it is, it corresponds to the one predicted by the theory. Namely, the wage of the higher-level managers expands proportionally more than that of the lower-level ones. Hence, Implication 6b is also corroborated by the evidence.

Table 23: Firms that change export status and do not change L

# of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
0	0	0.108	0.022	0.00	3,263
1	0	0.110	0.016	0.00	6,968
1	1	0.119	0.018	0.00	6,968
2	0	0.169	0.017	0.00	10,507
2	1	0.186	0.018	0.00	10,507
2	2	0.193	0.019	0.00	10,507
3	0	0.199	0.033	0.00	4,896
3	1	0.219	0.034	0.00	4,896
3	2	0.218	0.034	0.00	4,896
3	3	0.219	0.035	0.00	4,896

The final prediction of the theory to contrast with our data is Implication 6c, which states that new exporters that add layers decrease w_L^ℓ and increase n_L^ℓ at all ℓ. We have already argued in the previous section that firms that add layers decrease wages and increase hours at all layers, and

Table 20 shows that new exporters tend to add layers. So it is natural to expect that in fact the predictions of the theory will be corroborated by the data.

Table 24: $d \ln n_{Lit}^{\ell}$ for $i's$ that transition and change export status

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.482	0.074	0.00	528
0	2	0	1.536	0.195	0.00	95
0	3	0	2.990	0.289	0.00	15
1	0	0	-1.482	0.084	0.00	520
1	2	0	0.670	0.046	0.00	1132
1	2	1	0.584	0.045	0.00	1132
1	3	0	0.936	0.175	0.00	91
1	3	1	0.907	0.149	0.00	91
2	0	0	-1.561	0.213	0.00	100
2	1	0	-0.600	0.046	0.00	1119
2	1	1	-0.438	0.048	0.00	1119
2	3	0	1.070	0.049	0.00	861
2	3	1	1.006	0.057	0.00	861
2	3	2	0.877	0.056	0.00	861
3	0	0	-2.900	0.304	0.00	16
3	1	0	-1.162	0.161	0.00	105
3	1	1	-0.880	0.156	0.00	105
3	2	0	-1.228	0.056	0.00	872
3	2	1	-1.159	0.061	0.00	872
3	2	2	-1.045	0.059	0.00	872

Table 24 presents the average log change in the number of hours for all transitions and all layers. The table is similar to Table 14, but it uses the sample of firms that enter the export market and add layers and firms that exit the export market and drop layers. The results establish that, for all transitions and layers, firms that add layers increase the number of hours, while firms that drop layers decrease them. The numbers are also extremely large. In the data there are 1,132 firms with one layer of management that add a second layer and start to export. Our results indicate that the average percentage change in the number of hours in layer zero is 95% (exp(0.67)) and 74% in layer one. So new exporters with one layer of management that add a new layer almost double in employment size. Some transitions are smaller and some even larger. However, all of them tend to be fairly large and the effects are extremely significant and robust.

The next step is to look at wages. Again, we study the change in average log wages for all transitions and layers for the sample of new exporters that add layers and firms that exit the export

market and drop layers. The results are displayed in Table 25. The prediction in Implication 6c is broadly corroborated by the data (although the change in log wages in $\ell = 0$ for transitions from 2 to 3 layers is not significant). New exporters that add layers decrease wages and the firms that exit the export market and drop layers increase wages. As we show in Table 20, new exporters tend to add more layers than non-exporters. Similarly firms that exit the export market tend to drop more layers than exporters. Again, Appendix A tables A34 to A36, A38 and A41 show several robustness checks for firms that transition and for firms that do not transition, using only firms with adjacent layers and DADS wages.

Table 25: $d \ln w_{Lit}^{\ell}$ for $i's$ that transition and change export status

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.144	0.022	0.00	528
0	2	0	-0.593	0.108	0.00	95
0	3	0	-1.031	0.353	0.01	15
1	0	0	0.219	0.026	0.00	520
1	2	0	-0.025	0.010	0.01	1132
1	2	1	-0.232	0.015	0.00	1132
1	3	0	-0.158	0.043	0.00	91
1	3	1	-0.334	0.056	0.00	91
2	0	0	0.524	0.088	0.00	100
2	1	0	0.074	0.010	0.00	1119
2	1	1	0.247	0.015	0.00	1119
2	3	0	0.004	0.011	0.67	861
2	3	1	-0.043	0.013	0.00	861
2	3	2	-0.165	0.017	0.00	861
3	0	0	0.769	0.346	0.04	16
3	1	0	0.126	0.049	0.01	105
3	1	1	0.465	0.073	0.00	105
3	2	0	0.023	0.009	0.01	872
3	2	1	0.051	0.012	0.00	872
3	2	2	0.169	0.016	0.00	872

So exporters tend to reduce wages as a result of adding layers. To continue with our previous example, a new exporter that had one layer of management and added another as a result of its decision to start exporting *reduces* the wages of its workers in layer zero by 2.5%, and the wage of managers in layer one *declines* by 26%. In contrast, as Table 26 (discussed below) shows, the newly hired second layer manager earns 97.5% more than the average wage in the firm before the change. Appendix A shows robustness checks with the alternative wage database.

The result should change our view on the distribution of the gains from exporting. The view that new exporters pay higher wages is flawed. Most new exporters expand little and do not change their organization. They hire more hours and pay higher wages. The new exporters that expand substantially add layers of management. They hire substantially more workers but pay these workers less (since according to the theory they also know less). The new exporters that expand and add layers exhibit more dispersion in wages within the firm.

Table 26: Decomposition of total log change in average wages

$\bar{w}^{\ell \leq L}_{L'it}/\bar{w}_{Lit}$

From/to	1	2	3
0	0.956** (527)	0.814** (94)	0.629** (15)
1		0.963** (1,132)	0.872** (91)
2			0.988** (861)

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

From/to	1	2	3
0	1.445** (527)	1.510** (94)	1.309** (15)
1		1.975** (1,132)	1.854** (90)
2			5.244** (861)

s

From/to	1	2	3
0	0.740** (528)	0.586** (95)	0.649** (14)
1		0.848** (1,132)	0.749** (90)
2			0.941** (861)

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

From/to	1	2	3
0	−0.024 (528)	−0.300** (95)	−0.779* (15)
1		0.027** (1,132)	0.002 (91)
2			0.027** (861)

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

Table 26 shows the same decomposition we showed in Table 16 in the previous section. It separates the change in wages in the firm in the contribution to the average of the new top manager and the change in the wage of the pre-existing layers. The top left panel shows that the average wage of all existing layers decreases as firms add layers (and we know from Table 25 that it decreased for each layer individually). The top right panel shows the wage of the new top manager relative to the mean wage of the firm before the change. Clearly, wage dispersion in the firm increases substantially when it starts to export and adds layers.

We end this section as we did the last one, with a graphical representation of the change in firms as they become exporters. Figures 14 to 16 show the organization of firms that had zero, one, or two layers of management before they became exporters. The first row in each graph represents the old and new organization when the firm adds layers and starts to export. The second row represents the organization when the firm starts to export but keeps constant the number of layers.

Perhaps the most striking observation coming out of Figures 14 to 16 is how little firms change when they do not add layers compared to when they do. Hopefully, these figures are convincing in showing that new exporters expand by adding layers, adding employment, and reducing wages. The reduction in wages challenges, as far as we know, all theories of trade that do not add explicit organizational choices.

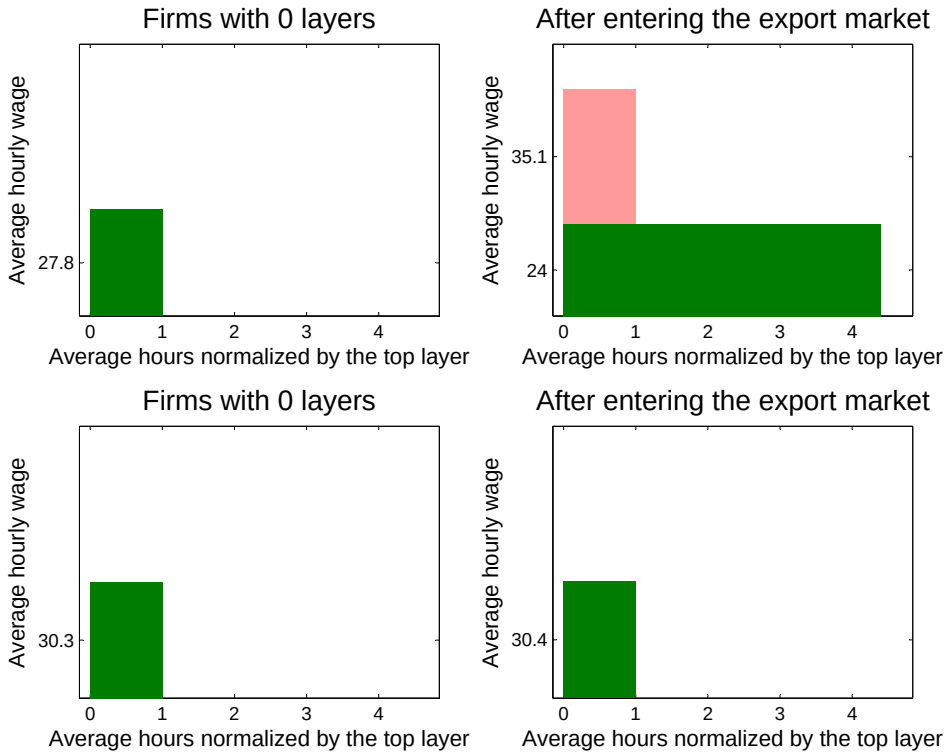


Figure 14: Representative new exporters with $L = 0$

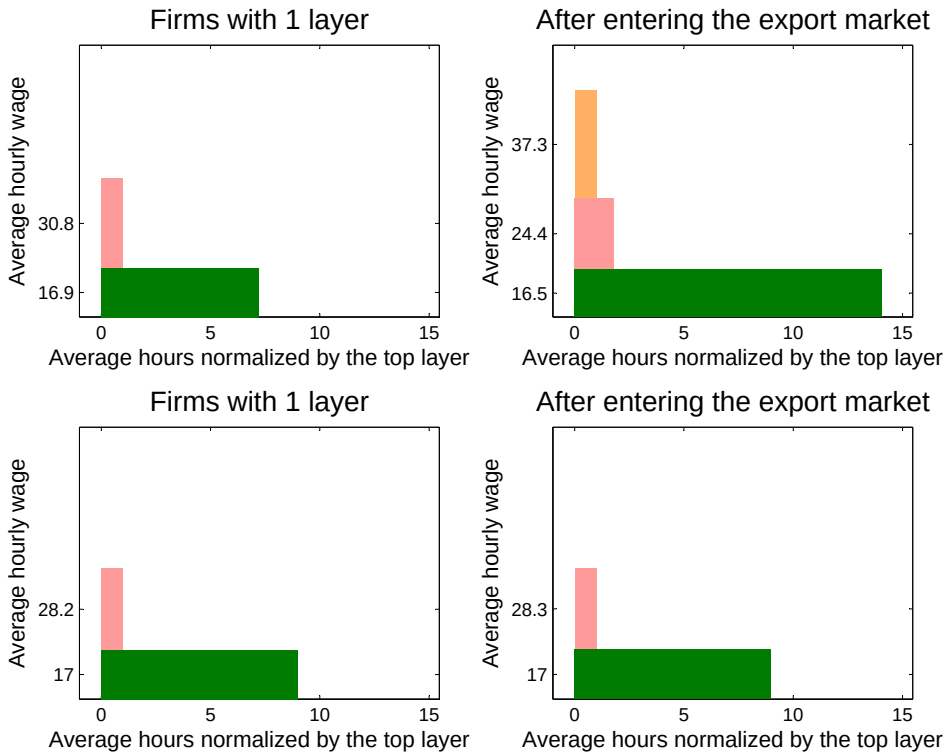


Figure 15: Representative new exporters with $L = 1$

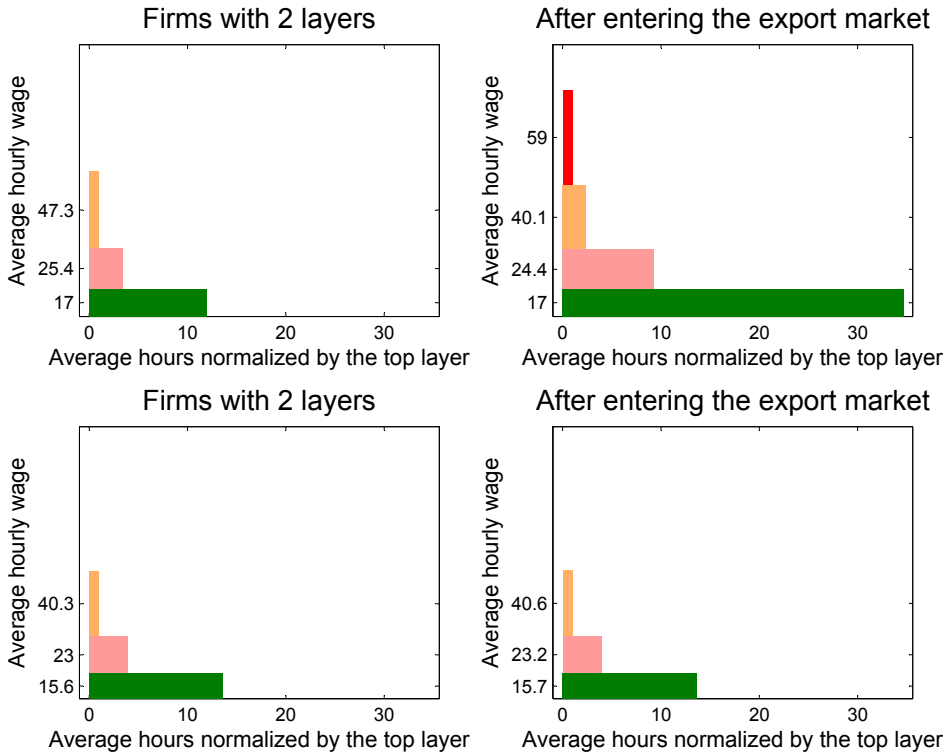


Figure 16: Representative new exporters with $L = 2$

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

The theory in CRH is useful as a guide into the optimal behavior of firms at the level of a layer. However, it is silent about how firms achieve any change in hours, average knowledge, or average wages in a layer when they face costs or restrictions to adjust freely. For example, how do they manage to lower the average wages in pre-existing layers conditional on the institutional environment in which they operate? The answer to this question will surely be specific to the particular characteristics of the French labor market. As such, it is interesting for learning something about France, but not necessarily about firms in general. This is why this question is not a central concern of our paper. Our conclusions and the interpretation of our findings is not affected by the answer to this question. Still, we want to make sure that it is not impossible for firms to do what we are arguing they are, in fact, doing. Of course, we know that labor market frictions, institutions, and regulations make reducing the wage of a particular employee complicated, if not impossible. So how can the prediction that firms that add layers reduce average wages in pre-existing layers be so clearly present in our data?

The data we use is very detailed in that they allow us to observe hierarchical occupations together with wages at the occupation level for most manufacturing firms. To study how reorganization affects layer-level average wages, we track hours of work, and their average hourly wage, as they enter, stay, or leave the layer during a transition.[13] Studying these transitions we find that

[13] See description of Tables A43 to A46 in the appendix for a detailed account of this excercise and findings.

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.[14] This reduces the average wage in the layer. That is, the adjustment happens via an extensive margin. Firms adjust the composition of the employees in the layer and not the individual wages that they command.

This way of adjusting the level of knowledge and wages in a layer clearly reflects the presence of some form of downward wage rigidities. Hence, 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 importance of this argument.

7 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 enormously 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. Furthermore, exporters that expand substantially do so by adding layers.

We find that the organization of labor in the firm depends dramatically 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). 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 (for example, because they start to export) 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.[15] The evidence we provide is extremely robust and needs to be confronted by all

[14] Tables A43 to A46 in the appendix present these results for all transitions. Table A43 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 A44 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 0 in firms that go from 3 to 2 layers of management).

[15] Calvo and Wellisz (1979) and Qian (1994) present theories were efficiency wages decrease as the span of control is reduced. However, these theories assume a fixed organizational form and so are silent about the way in which

existing and future theories of firm dynamics and export behavior.

Our results also have implications for the distribution of wages within the firm and how this distribution changes as firms expand or start to export. Since most large expansions are done through reorganization via new layers, firm growth and export behavior lead to increases in the range of wages within the firm. This is the case because the new managers in firms that add layers make much more than the average employee before the change.

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 a trade liberalization 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.

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.

---

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] Bernard, A. B. and J. B. Jensen, (1995). "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987," *Brookings Papers in Economic Activity. Microeconomics*, 67-112.

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

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

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

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

[7] Caliendo, L. and E. Rossi-Hansberg, (2012). "The Impact of Trade on Organization and Productivity," *Quarterly Journal of Economics*, forthcoming.

[8] 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.

[9] 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.

[10] Eaton, J., S. Kortum, and F. Kramarz (2011). "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79:5, 1453-1498.

[11] Eaton, J., S. Kortum, F. Kramarz, and R. Sampognaro, (2011). "Dissecting the French Export Wage Premium," Working Paper.

[12] Egger, H., and U. Kreickemeier, (2009). "Firm Heterogeneity and the Labor Market Effects of Trade Liberalisation," *International Economic Review*, 50: 187-216.

[13] Felbermayr, G., J. Prat, and H. Schmerer, (2008). "Globalization and Labor Market Outcomes: Bargaining, Search Frictions, and Firm Heterogeneity," *IZA discussion paper no.* 3363.

[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] Guadalupe, M., and J. Wulf, (2010). "The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies," *American Economic Journal*: Applied Economics, 2:4, 105-27.

[19] Helpman, E., O. Itskhoki, and S. Redding, (2010). "Inequality and Unemployment in a Global Economy," *Econometrica*, 78: 1239-1283.

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

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

[22] 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.

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

[24] 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.

[25] Verhoogen, E. A., (2008). "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 123:2, 489-530.

8 Appendix A

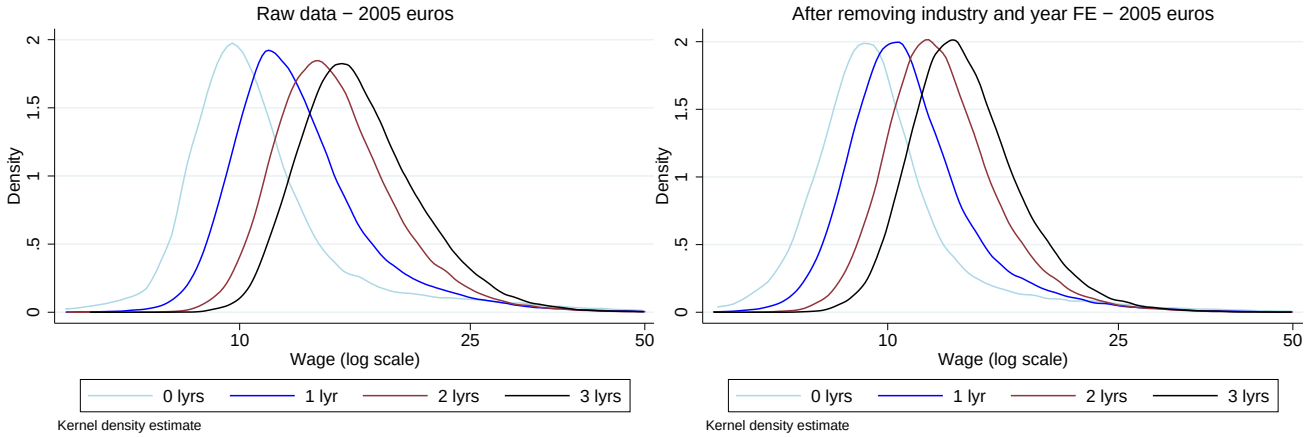


Figure A1: Firm average hourly wage distribution by number of layers, DADS wages

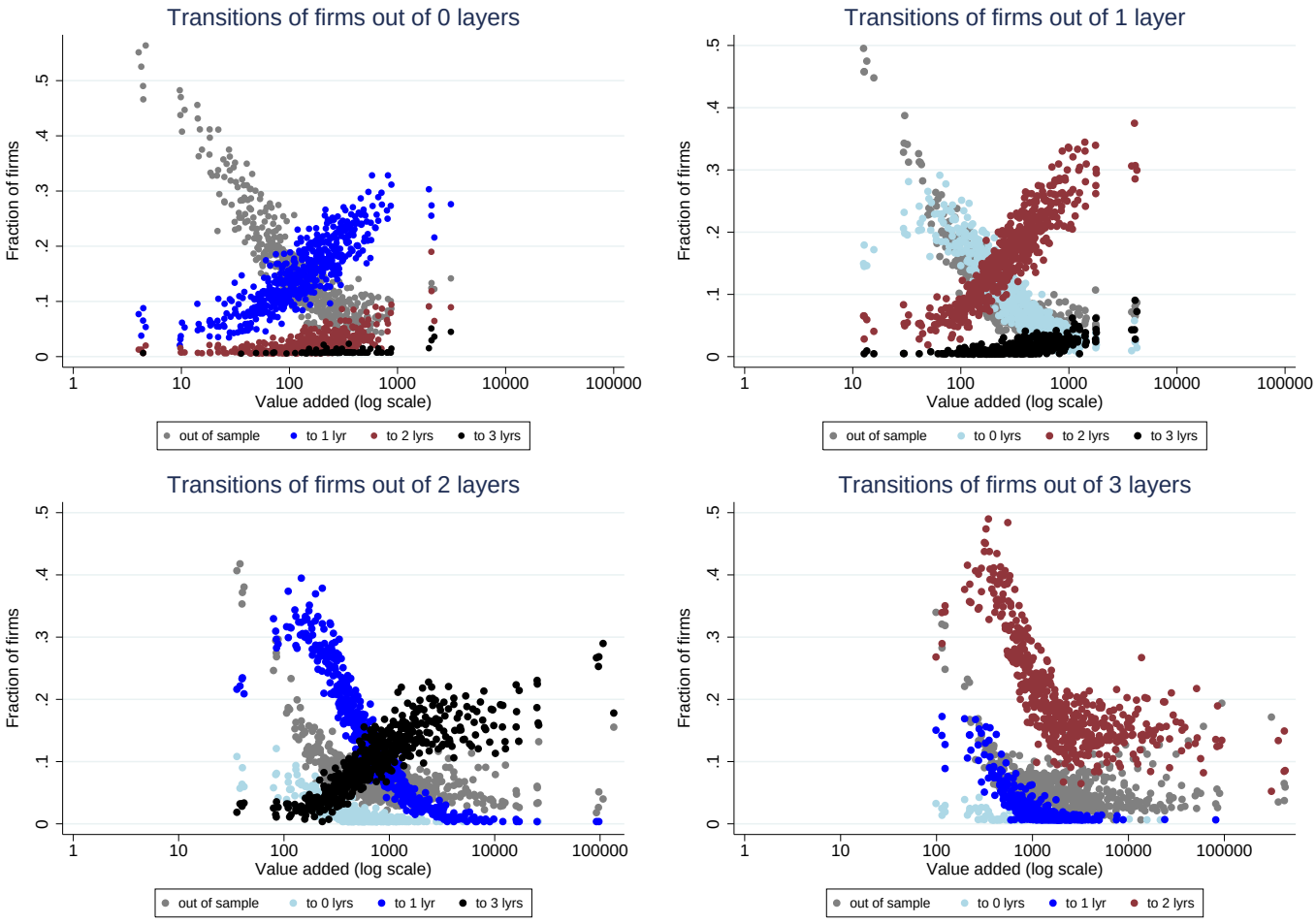


Figure A2: Transitions across layers depend on value added

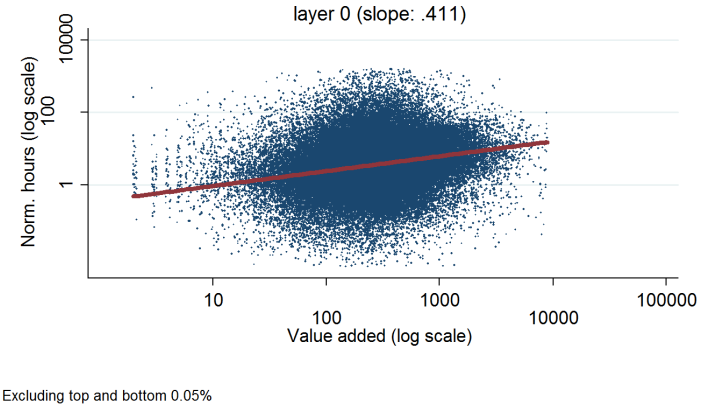


Figure A3: Normalized hours vs. firm value added among 1 layer firms

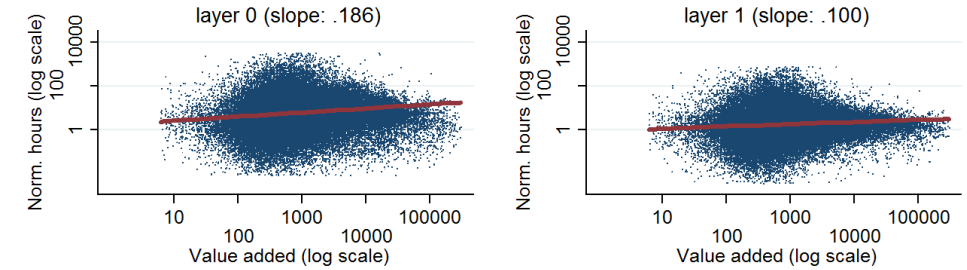


Figure A4: Normalized hours vs. firm value added among 2 layer firms

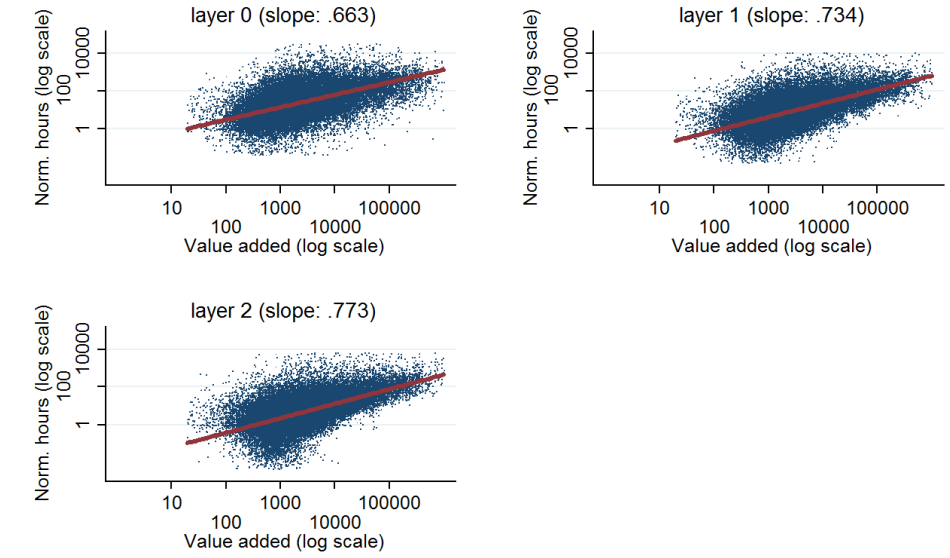


Figure A5: Normalized hours vs. firm value added among 3 layer firms

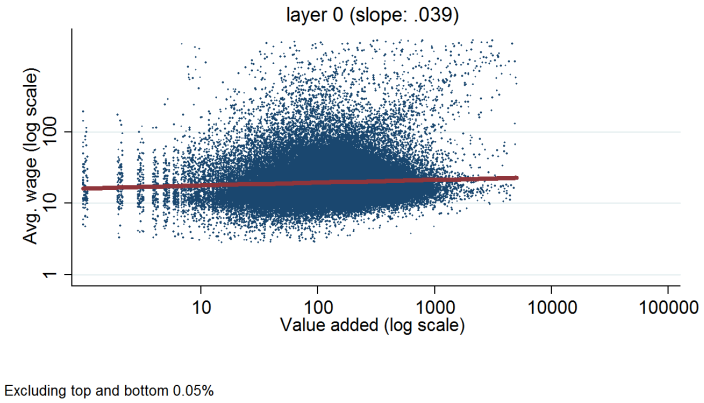


Figure A6: Wages vs. firm value added among 1 layer firms

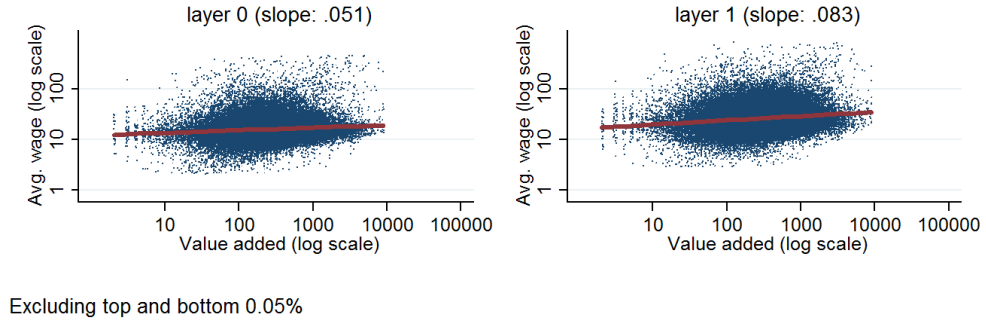


Figure A7: Wages vs. firm value added among 2 layer firms

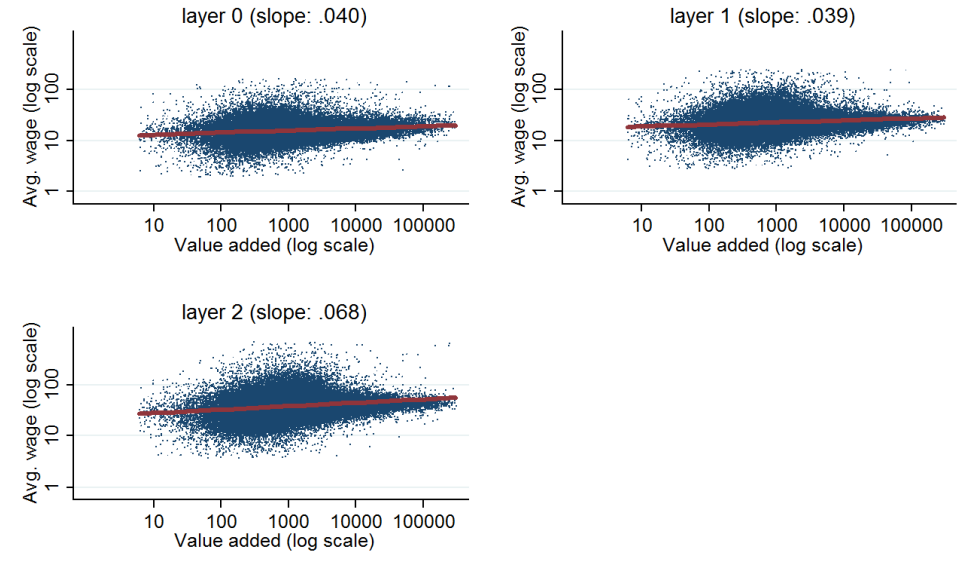


Figure A8: Wages vs. firm value added among 3 layer firms

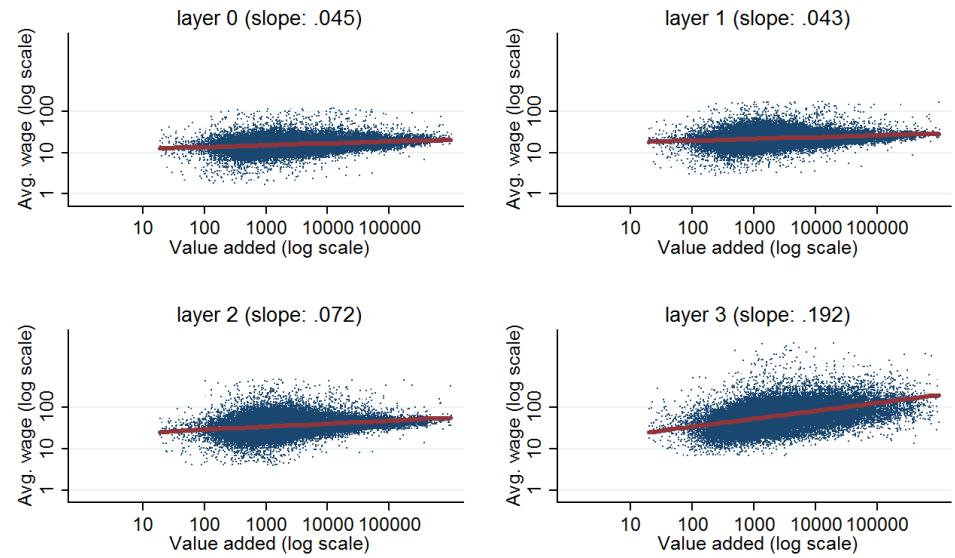


Figure A9: Wages vs. firm value added among 4 layer firms

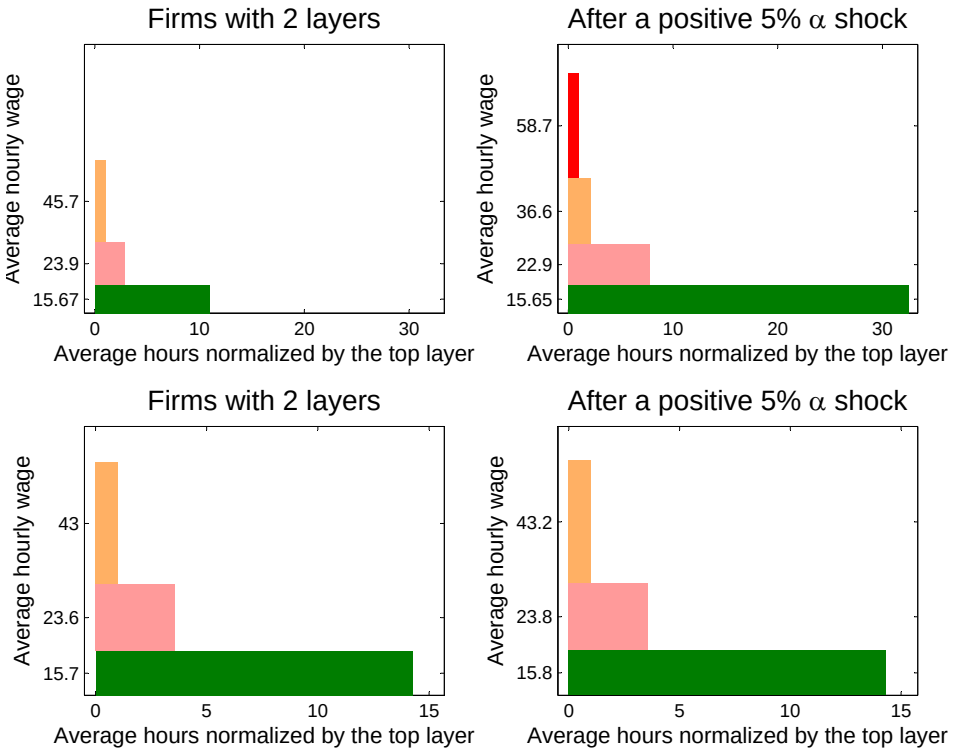


Figure A10: Transition after a 5% positive α shock

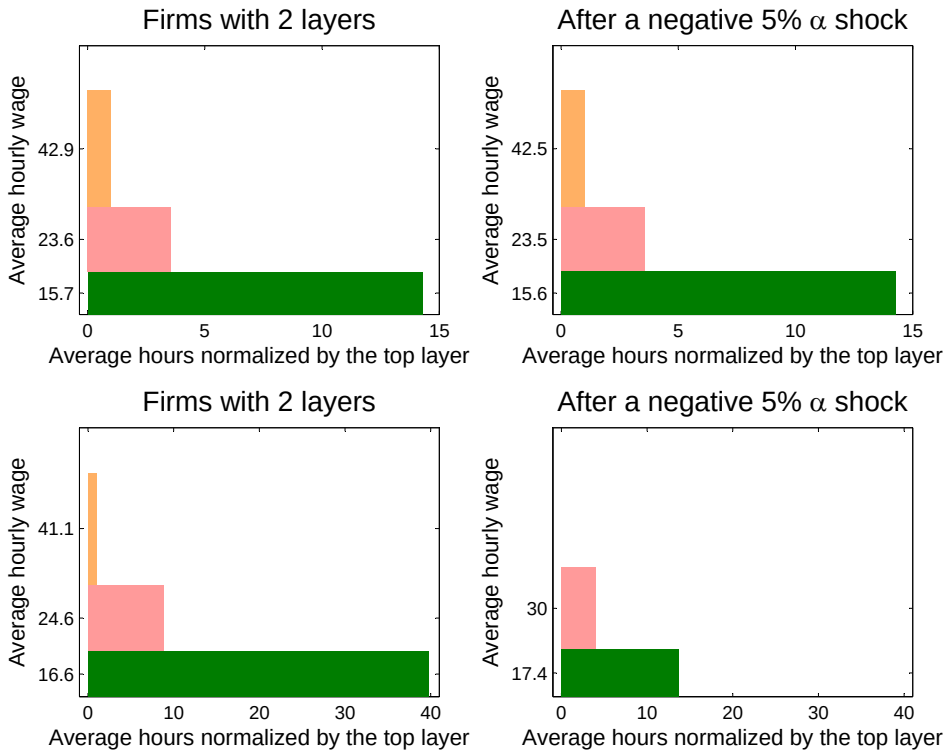


Figure A11: Transition after a 5% negative α shock

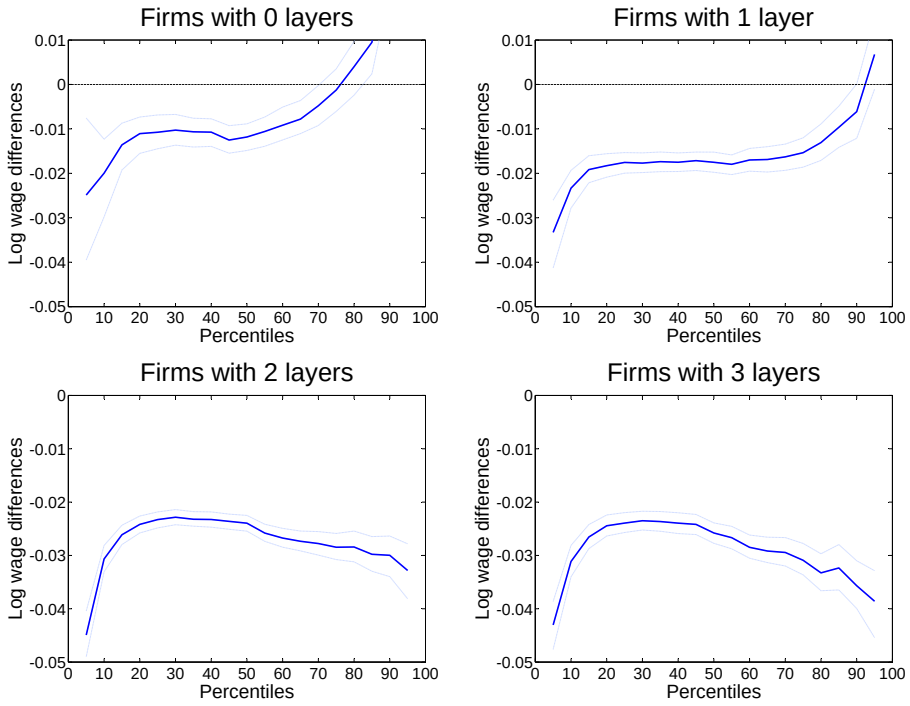


Figure A12: Difference in the distribution of wages for firms that do not transition conditioning on $d\ln VA < 0$

Table A1: Data description by number of layers in the firm, DADS wages					
# of		Average			Median
layers	Firm-years	VA	Hours	Wage*	wage
0	81,909	205	7,946	11.69	10.18
1	126,069	403	16,450	13.19	12.08
2	161,449	2,821	85,674	15.12	14.22
3	87,211	8,879	227,070	16.67	15.71

Value added in 000s of 2005 euros. Wage* = average hourly wage in 2005 euros from trimmed sample (0.05% top and bottom in each cell).

Table A2: % of firms that transition to an adjacent layer				
Conditioning on firms with adjacent layers				
	# of layers			
	0	1	2	3
Transition up	75.4	82.9	100	-
Transition down	-	91.3	60.3	75.9

Table A3: Distribution of layers at $t+1$ conditional on layers at t, weighted by VA							
		# of layers at $t+1$					
		Exit	0	1	2	3	Total
# of layers at t	0	11.9	64.1	19.6	3.7	0.7	100
	1	7.2	6.6	62.2	21.7	2.2	100
	2	5.8	0.2	2.5	72.6	19.0	100
	3	7.7	0.0	0.2	13.3	78.8	100

Table A4: Behavior of firms that exit the export market			
	All	Increase L	No change in L
$d\ln$ total hours	-0.027***	-0.147***	-0.017***
- detrended	-0.013***	-0.133***	-0.003
$d\ln\sum_{\ell=0}^{L} n_L^\ell$	-0.009	-1.251***	0.018***
- detrended	0.002	-1.239***	0.029***
$d\ln VA$	-0.046***	-0.124***	-0.035***
- detrended	-0.038***	-0.115***	-0.028***
$d\ln$ avg wage	0.006***	0.027***	0.008***
- detrended	-0.012***	-0.010	-0.010***
- common layers	-0.010***	0.154***	0.008***
- - detrended	-0.010***	0.134***	-0.012***
% firms	100	15.46	70.68
% VA change	100	20.28	76.01

*** significant at 1%.

Table A5: Layer transitions for firms that stop exporting relative to exporters					
		# of layers at $t+1$			
		0	1	2	3
# of layers at t	0	2.85***	-1.8*	-0.95**	-0.10
	1	1.29	1.09***	-2.00***	-0.38*
	2	0.66***	6.02***	-4.41***	-2.27***
	3	0.35***	1.87***	6.61***	-8.83***

* significant at 10%, ** significant at 5%, *** significant at 1%.

Change in normalized hours for firms that change layers (Robustness checks)

Table A6

Conditioning on $d \ln VA > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.514	0.024	0.00	5475
0	2	0	1.724	0.071	0.00	760
0	3	0	2.652	0.251	0.00	71
1	0	0	-1.531	0.023	0.00	6491
1	2	0	0.716	0.016	0.00	9211
1	2	1	0.542	0.016	0.00	9211
1	3	0	1.233	0.064	0.00	638
1	3	1	1.039	0.064	0.00	638
2	0	0	-1.756	0.057	0.00	1054
2	1	0	-0.693	0.016	0.00	9739
2	1	1	-0.529	0.017	0.00	9739
2	3	0	1.346	0.020	0.00	7482
2	3	1	1.279	0.021	0.00	7482
2	3	2	1.166	0.022	0.00	7482
3	0	0	-2.296	0.171	0.00	85
3	1	0	-1.155	0.054	0.00	858
3	1	1	-1.025	0.051	0.00	858
3	2	0	-1.390	0.019	0.00	8186
3	2	1	-1.305	0.021	0.00	8186
3	2	2	-1.239	0.020	0.00	8186

Table A7

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} N_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.499	0.030	0.00	3717
0	2	0	1.745	0.082	0.00	583
0	3	0	2.863	0.291	0.00	50
1	0	0	-1.518	0.027	0.00	4562
1	2	0	0.768	0.019	0.00	6231
1	2	1	0.546	0.020	0.00	6231
1	3	0	1.347	0.079	0.00	423
1	3	1	1.020	0.078	0.00	423
2	0	0	-1.709	0.062	0.00	806
2	1	0	-0.730	0.020	0.00	6642
2	1	1	-0.536	0.020	0.00	6642
2	3	0	1.335	0.024	0.00	4758
2	3	1	1.248	0.026	0.00	4758
2	3	2	1.118	0.026	0.00	4758
3	0	0	-2.210	0.181	0.00	69
3	1	0	-1.242	0.063	0.00	615
3	1	1	-1.052	0.060	0.00	615
3	2	0	-1.426	0.023	0.00	5401
3	2	1	-1.331	0.025	0.00	5401
3	2	2	-1.232	0.025	0.00	5401

Table A8

Conditioning on $d \ln \sum_{\ell=0}^{L} N_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.463	0.024	0.00	5847
0	2	0	1.690	0.066	0.00	875
0	3	0	2.575	0.230	0.00	72
1	0	0	-1.515	0.022	0.00	6775
1	2	0	0.754	0.015	0.00	9338
1	2	1	0.518	0.016	0.00	9338
1	3	0	1.373	0.066	0.00	650
1	3	1	1.026	0.062	0.00	650
2	0	0	-1.728	0.053	0.00	1146
2	1	0	-0.730	0.016	0.00	9998
2	1	1	-0.517	0.016	0.00	9998
2	3	0	1.311	0.019	0.00	7369
2	3	1	1.218	0.021	0.00	7369
2	3	2	1.099	0.021	0.00	7369
3	0	0	-2.213	0.183	0.00	104
3	1	0	-1.204	0.052	0.00	903
3	1	1	-0.978	0.050	0.00	903
3	2	0	-1.469	0.019	0.00	8416
3	2	1	-1.378	0.021	0.00	8416
3	2	2	-1.263	0.020	0.00	8416

Table A9

Conditioning on $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.520	0.017	0.00	10432
0	2	0	1.745	0.053	0.00	1350
0	3	0	2.312	0.193	0.00	111
1	0	0	-1.585	0.017	0.00	11356
1	2	0	1.105	0.012	0.00	13695
1	2	1	0.803	0.013	0.00	13695
1	3	0	1.436	0.047	0.00	1065
1	3	1	1.157	0.047	0.00	1065
2	0	0	-1.801	0.046	0.00	1698
2	1	0	-1.160	0.012	0.00	14006
2	1	1	-0.846	0.014	0.00	14006
2	3	0	1.677	0.015	0.00	11947
2	3	1	1.641	0.016	0.00	11947
2	3	2	1.463	0.017	0.00	11947
3	0	0	-2.203	0.157	0.00	142
3	1	0	-1.359	0.039	0.00	1342
3	1	1	-1.081	0.041	0.00	1342
3	2	0	-1.794	0.014	0.00	12864
3	2	1	-1.749	0.015	0.00	12864
3	2	2	-1.574	0.016	0.00	12864

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in normalized hours for firms that change layers (Robustness checks)

Table A10

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.514	0.024	0.00	5475
0	2	0	1.724	0.071	0.00	760
0	3	0	2.652	0.251	0.00	71
1	0	0	-1.531	0.023	0.00	6491
1	2	0	1.076	0.016	0.00	7589
1	2	1	0.767	0.018	0.00	7589
1	3	0	1.389	0.064	0.00	595
1	3	1	1.132	0.066	0.00	595
2	0	0	-1.756	0.057	0.00	1054
2	1	0	-1.139	0.016	0.00	7758
2	1	1	-0.805	0.018	0.00	7758
2	3	0	1.634	0.020	0.00	6459
2	3	1	1.588	0.022	0.00	6459
2	3	2	1.406	0.023	0.00	6459
3	0	0	-2.296	0.171	0.00	85
3	1	0	-1.421	0.051	0.00	770
3	1	1	-1.165	0.053	0.00	770
3	2	0	-1.730	0.019	0.00	6948
3	2	1	-1.668	0.021	0.00	6948
3	2	2	-1.511	0.022	0.00	6948

Selected sample with adjacent layers

Table A11

Adjacent layers

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.787	0.019	0.00	6788
0	2	0	2.364	0.061	0.00	796
0	3	0	2.650	0.200	0.00	75
1	0	0	-1.841	0.019	0.00	7016
1	2	0	1.244	0.017	0.00	8339
1	2	1	1.087	0.017	0.00	8339
1	3	0	1.654	0.062	0.00	594
1	3	1	1.529	0.061	0.00	594
2	0	0	-2.284	0.056	0.00	943
2	1	0	-1.286	0.018	0.00	8204
2	1	1	-1.158	0.018	0.00	8204
2	3	0	1.844	0.017	0.00	10233
2	3	1	1.838	0.018	0.00	10233
2	3	2	1.678	0.018	0.00	10233
3	0	0	-2.476	0.154	0.00	92
3	1	0	-1.507	0.053	0.00	780
3	1	1	-1.432	0.052	0.00	780
3	2	0	-1.892	0.016	0.00	11614
3	2	1	-1.866	0.017	0.00	11614
3	2	2	-1.743	0.016	0.00	11614

Table A12

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.824	0.025	0.00	3636
0	2	0	2.351	0.083	0.00	446
0	3	0	3.000	0.274	0.00	47
1	0	0	-1.822	0.025	0.00	3845
1	2	0	1.518	0.021	0.00	4134
1	2	1	1.261	0.023	0.00	4134
1	3	0	1.845	0.083	0.00	304
1	3	1	1.671	0.085	0.00	304
2	0	0	-2.214	0.072	0.00	567
2	1	0	-1.606	0.023	0.00	3847
2	1	1	-1.352	0.025	0.00	3847
2	3	0	1.999	0.022	0.00	5111
2	3	1	1.993	0.024	0.00	5111
2	3	2	1.781	0.025	0.00	5111
3	0	0	-2.434	0.198	0.00	56
3	1	0	-1.679	0.064	0.00	426
3	1	1	-1.593	0.065	0.00	426
3	2	0	-2.049	0.021	0.00	5699
3	2	1	-2.017	0.022	0.00	5699
3	2	2	-1.851	0.022	0.00	5699

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in normalized hours for firms that change layers (Robustness checks)

Selected sample with adjacent layers

Table A13

Conditioning on $d \ln VA > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.824	0.025	0.00	3636
0	2	0	2.351	0.083	0.00	446
0	3	0	3.000	0.274	0.00	47
1	0	0	-1.822	0.025	0.00	3845
1	2	0	1.248	0.023	0.00	4594
1	2	1	1.100	0.023	0.00	4594
1	3	0	1.709	0.087	0.00	319
1	3	1	1.578	0.086	0.00	319
2	0	0	-2.214	0.072	0.00	567
2	1	0	-1.287	0.025	0.00	4331
2	1	1	-1.157	0.024	0.00	4331
2	3	0	1.840	0.023	0.00	5434
2	3	1	1.824	0.024	0.00	5434
2	3	2	1.657	0.025	0.00	5434
3	0	0	-2.434	0.198	0.00	56
3	1	0	-1.522	0.069	0.00	449
3	1	1	-1.500	0.065	0.00	449
3	2	0	-1.858	0.022	0.00	6123
3	2	1	-1.821	0.023	0.00	6123
3	2	2	-1.711	0.022	0.00	6123

Table A14

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} N_L^\ell > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.854	0.030	0.00	2528
0	2	0	2.516	0.092	0.00	335
0	3	0	3.378	0.382	0.00	27
1	0	0	-1.855	0.030	0.00	2623
1	2	0	1.321	0.027	0.00	3099
1	2	1	1.121	0.027	0.00	3099
1	3	0	1.880	0.104	0.00	203
1	3	1	1.594	0.105	0.00	203
2	0	0	-2.205	0.080	0.00	407
2	1	0	-1.338	0.031	0.00	2873
2	1	1	-1.190	0.029	0.00	2873
2	3	0	1.806	0.028	0.00	3450
2	3	1	1.783	0.030	0.00	3450
2	3	2	1.574	0.030	0.00	3450
3	0	0	-2.341	0.207	0.00	40
3	1	0	-1.550	0.079	0.00	316
3	1	1	-1.492	0.073	0.00	316
3	2	0	-1.883	0.027	0.00	4064
3	2	1	-1.830	0.028	0.00	4064
3	2	2	-1.683	0.027	0.00	4064

Table A15

Conditioning on $d \ln \sum_{\ell=0}^{L} N_L^\ell > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.831	0.024	0.00	3888
0	2	0	2.481	0.076	0.00	476
0	3	0	3.127	0.288	0.00	38
1	0	0	-1.846	0.024	0.00	3966
1	2	0	1.309	0.022	0.00	4506
1	2	1	1.092	0.022	0.00	4506
1	3	0	1.832	0.083	0.00	307
1	3	1	1.520	0.083	0.00	307
2	0	0	-2.269	0.069	0.00	567
2	1	0	-1.325	0.024	0.00	4322
2	1	1	-1.163	0.024	0.00	4322
2	3	0	1.784	0.022	0.00	5292
2	3	1	1.759	0.024	0.00	5292
2	3	2	1.550	0.024	0.00	5292
3	0	0	-2.467	0.191	0.00	57
3	1	0	-1.562	0.067	0.00	456
3	1	1	-1.446	0.064	0.00	456
3	2	0	-1.935	0.021	0.00	6356
3	2	1	-1.887	0.022	0.00	6356
3	2	2	-1.728	0.022	0.00	6356

Table A16

Conditioning on $d \ln \sum_{\ell=0}^{L} n_L^\ell > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.787	0.019	0.00	6788
0	2	0	2.364	0.061	0.00	796
0	3	0	2.650	0.200	0.00	75
1	0	0	-1.841	0.019	0.00	7016
1	2	0	1.527	0.016	0.00	7445
1	2	1	1.265	0.018	0.00	7445
1	3	0	1.804	0.059	0.00	563
1	3	1	1.628	0.061	0.00	563
2	0	0	-2.284	0.056	0.00	943
2	1	0	-1.610	0.017	0.00	7245
2	1	1	-1.364	0.018	0.00	7245
2	3	0	2.024	0.016	0.00	9549
2	3	1	2.028	0.017	0.00	9549
2	3	2	1.822	0.018	0.00	9549
3	0	0	-2.476	0.154	0.00	92
3	1	0	-1.674	0.049	0.00	736
3	1	1	-1.539	0.051	0.00	736
3	2	0	-2.080	0.015	0.00	10831
3	2	1	-2.059	0.016	0.00	10831
3	2	2	-1.883	0.016	0.00	10831

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in average wages for firms that change layers (Robustness checks)

Table A17

Conditioning on $d \ln VA > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.113	0.007	0.00	5475
0	2	0	-0.411	0.031	0.00	760
0	3	0	-0.848	0.161	0.00	71
1	0	0	0.199	0.007	0.00	6491
1	2	0	-0.014	0.003	0.00	9211
1	2	1	-0.211	0.005	0.00	9211
1	3	0	-0.130	0.022	0.00	638
1	3	1	-0.374	0.026	0.00	638
2	0	0	0.481	0.028	0.00	1054
2	1	0	0.061	0.004	0.00	9739
2	1	1	0.246	0.005	0.00	9739
2	3	0	0.019	0.003	0.00	7482
2	3	1	-0.022	0.004	0.00	7482
2	3	2	-0.152	0.006	0.00	7482
3	0	0	1.018	0.156	0.00	85
3	1	0	0.174	0.020	0.00	858
3	1	1	0.397	0.023	0.00	858
3	2	0	0.000	0.003	0.96	8186
3	2	1	0.029	0.004	0.00	8186
3	2	2	0.119	0.005	0.00	8186

Table A18

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} N_L^\ell > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.210	0.009	0.00	3717
0	2	0	-0.545	0.037	0.00	583
0	3	0	-1.191	0.205	0.00	50
1	0	0	0.313	0.009	0.00	4562
1	2	0	-0.066	0.004	0.00	6231
1	2	1	-0.263	0.006	0.00	6231
1	3	0	-0.233	0.030	0.00	423
1	3	1	-0.502	0.033	0.00	423
2	0	0	0.639	0.034	0.00	806
2	1	0	0.128	0.005	0.00	6642
2	1	1	0.309	0.007	0.00	6642
2	3	0	-0.019	0.004	0.00	4758
2	3	1	-0.065	0.005	0.00	4758
2	3	2	-0.200	0.007	0.00	4758
3	0	0	1.265	0.179	0.00	69
3	1	0	0.259	0.027	0.00	615
3	1	1	0.478	0.028	0.00	615
3	2	0	0.048	0.004	0.00	5401
3	2	1	0.076	0.005	0.00	5401
3	2	2	0.164	0.006	0.00	5401

Table A19

Conditioning on $d \ln \sum_{\ell=0}^{L} N_L^\ell > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.276	0.007	0.00	5847
0	2	0	-0.670	0.033	0.00	875
0	3	0	-1.443	0.170	0.00	72
1	0	0	0.338	0.007	0.00	6775
1	2	0	-0.116	0.004	0.00	9338
1	2	1	-0.320	0.005	0.00	9338
1	3	0	-0.329	0.028	0.00	650
1	3	1	-0.591	0.029	0.00	650
2	0	0	0.708	0.029	0.00	1146
2	1	0	0.158	0.004	0.00	9998
2	1	1	0.342	0.005	0.00	9998
2	3	0	-0.059	0.004	0.00	7369
2	3	1	-0.109	0.004	0.00	7369
2	3	2	-0.244	0.006	0.00	7369
3	0	0	1.454	0.146	0.00	104
3	1	0	0.300	0.022	0.00	903
3	1	1	0.529	0.023	0.00	903
3	2	0	0.079	0.003	0.00	8416
3	2	1	0.108	0.004	0.00	8416
3	2	2	0.197	0.005	0.00	8416

Table A20

Conditioning on $d \ln \sum_{\ell=0}^{L} n_L^\ell > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.131	0.005	0.00	10432
0	2	0	-0.432	0.024	0.00	1350
0	3	0	-0.943	0.131	0.00	111
1	0	0	0.201	0.005	0.00	11356
1	2	0	-0.039	0.003	0.00	13695
1	2	1	-0.202	0.004	0.00	13695
1	3	0	-0.179	0.019	0.00	1065
1	3	1	-0.411	0.021	0.00	1065
2	0	0	0.489	0.022	0.00	1698
2	1	0	0.080	0.003	0.00	14006
2	1	1	0.213	0.004	0.00	14006
2	3	0	-0.011	0.003	0.00	11947
2	3	1	-0.046	0.003	0.00	11947
2	3	2	-0.149	0.004	0.00	11947
3	0	0	1.102	0.120	0.00	142
3	1	0	0.192	0.016	0.00	1342
3	1	1	0.395	0.018	0.00	1342
3	2	0	0.032	0.003	0.00	12864
3	2	1	0.048	0.003	0.00	12864
3	2	2	0.112	0.004	0.00	12864

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in average wages for firms that change layers (Robustness checks)

Table A21

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.113	0.007	0.00	5475
0	2	0	-0.411	0.031	0.00	760
0	3	0	-0.848	0.161	0.00	71
1	0	0	0.199	0.007	0.00	6491
1	2	0	-0.011	0.004	0.01	7589
1	2	1	-0.177	0.005	0.00	7589
1	3	0	-0.139	0.023	0.00	595
1	3	1	-0.375	0.027	0.00	595
2	0	0	0.481	0.028	0.00	1054
2	1	0	0.057	0.004	0.00	7758
2	1	1	0.191	0.006	0.00	7758
2	3	0	0.017	0.003	0.00	6459
2	3	1	-0.015	0.004	0.00	6459
2	3	2	-0.121	0.006	0.00	6459
3	0	0	1.018	0.156	0.00	85
3	1	0	0.181	0.022	0.00	770
3	1	1	0.376	0.024	0.00	770
3	2	0	0.003	0.004	0.42	6948
3	2	1	0.020	0.004	0.00	6948
3	2	2	0.083	0.005	0.00	6948

Selected sample with adjacent layers

Table A22

Adjacent layers

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.054	0.005	0.00	6788
0	2	0	-0.188	0.024	0.00	796
0	3	0	-0.414	0.120	0.00	75
1	0	0	0.125	0.005	0.00	7016
1	2	0	-0.013	0.003	0.00	8339
1	2	1	-0.062	0.004	0.00	8339
1	3	0	-0.098	0.021	0.00	594
1	3	1	-0.147	0.023	0.00	594
2	0	0	0.282	0.024	0.00	943
2	1	0	0.061	0.004	0.00	8204
2	1	1	0.103	0.004	0.00	8204
2	3	0	-0.003	0.003	0.34	10233
2	3	1	-0.016	0.003	0.00	10233
2	3	2	-0.088	0.004	0.00	10233
3	0	0	0.535	0.121	0.00	92
3	1	0	0.161	0.019	0.00	780
3	1	1	0.183	0.020	0.00	780
3	2	0	0.026	0.003	0.00	11614
3	2	1	0.026	0.003	0.00	11614
3	2	2	0.072	0.004	0.00	11614

Table A23

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.028	0.006	0.00	3636
0	2	0	-0.161	0.030	0.00	446
0	3	0	-0.297	0.132	0.03	47
1	0	0	0.116	0.007	0.00	3845
1	2	0	0.019	0.005	0.00	4134
1	2	1	-0.029	0.005	0.00	4134
1	3	0	-0.076	0.028	0.01	304
1	3	1	-0.108	0.031	0.00	304
2	0	0	0.277	0.032	0.00	567
2	1	0	0.031	0.006	0.00	3847
2	1	1	0.063	0.006	0.00	3847
2	3	0	0.024	0.004	0.00	5111
2	3	1	0.017	0.004	0.00	5111
2	3	2	-0.048	0.005	0.00	5111
3	0	0	0.458	0.147	0.00	56
3	1	0	0.134	0.026	0.00	426
3	1	1	0.153	0.027	0.00	426
3	2	0	-0.005	0.004	0.24	5699
3	2	1	-0.009	0.004	0.05	5699
3	2	2	0.027	0.005	0.00	5699

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in average wages for firms that change layers (Robustness checks)

Selected sample with adjacent layers

Table A24

Conditioning on $d \ln VA > 0$

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.028	0.006	0.00	3636
0	2	0	-0.161	0.030	0.00	446
0	3	0	-0.297	0.132	0.03	47
1	0	0	0.116	0.007	0.00	3845
1	2	0	0.016	0.004	0.00	4594
1	2	1	-0.027	0.005	0.00	4594
1	3	0	-0.072	0.027	0.01	319
1	3	1	-0.099	0.030	0.00	319
2	0	0	0.277	0.032	0.00	567
2	1	0	0.039	0.005	0.00	4331
2	1	1	0.073	0.006	0.00	4331
2	3	0	0.025	0.004	0.00	5434
2	3	1	0.015	0.004	0.00	5434
2	3	2	-0.054	0.006	0.00	5434
3	0	0	0.458	0.147	0.00	56
3	1	0	0.132	0.025	0.00	449
3	1	1	0.142	0.026	0.00	449
3	2	0	-0.004	0.004	0.25	6123
3	2	1	-0.007	0.004	0.09	6123
3	2	2	0.031	0.005	0.00	6123

Table A25

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} N_L^{\ell} > 0$

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.099	0.007	0.00	2528
0	2	0	-0.244	0.035	0.00	335
0	3	0	-0.566	0.211	0.01	27
1	0	0	0.212	0.009	0.00	2623
1	2	0	-0.031	0.006	0.00	3099
1	2	1	-0.074	0.006	0.00	3099
1	3	0	-0.168	0.038	0.00	203
1	3	1	-0.219	0.040	0.00	203
2	0	0	0.403	0.041	0.00	407
2	1	0	0.104	0.007	0.00	2873
2	1	1	0.133	0.008	0.00	2873
2	3	0	-0.014	0.005	0.01	3450
2	3	1	-0.024	0.006	0.00	3450
2	3	2	-0.101	0.007	0.00	3450
3	0	0	0.661	0.196	0.00	40
3	1	0	0.217	0.033	0.00	316
3	1	1	0.233	0.033	0.00	316
3	2	0	0.045	0.005	0.00	4064
3	2	1	0.042	0.005	0.00	4064
3	2	2	0.082	0.006	0.00	4064

Table A26

Conditioning on $d \ln \sum_{\ell=0}^{L} N_L^{\ell} > 0$

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.158	0.006	0.00	3888
0	2	0	-0.349	0.035	0.00	476
0	3	0	-0.880	0.205	0.00	38
1	0	0	0.234	0.007	0.00	3966
1	2	0	-0.081	0.005	0.00	4506
1	2	1	-0.126	0.006	0.00	4506
1	3	0	-0.260	0.035	0.00	307
1	3	1	-0.328	0.036	0.00	307
2	0	0	0.456	0.036	0.00	567
2	1	0	0.133	0.006	0.00	4322
2	1	1	0.166	0.006	0.00	4322
2	3	0	-0.054	0.004	0.00	5292
2	3	1	-0.066	0.005	0.00	5292
2	3	2	-0.146	0.006	0.00	5292
3	0	0	0.854	0.181	0.00	57
3	1	0	0.281	0.029	0.00	456
3	1	1	0.302	0.030	0.00	456
3	2	0	0.078	0.004	0.00	6356
3	2	1	0.075	0.004	0.00	6356
3	2	2	0.118	0.005	0.00	6356

Table A27

Conditioning on $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.054	0.005	0.00	6788
0	2	0	-0.188	0.024	0.00	796
0	3	0	-0.414	0.120	0.00	75
1	0	0	0.125	0.005	0.00	7016
1	2	0	-0.010	0.004	0.01	7445
1	2	1	-0.059	0.004	0.00	7445
1	3	0	-0.101	0.022	0.00	563
1	3	1	-0.151	0.024	0.00	563
2	0	0	0.282	0.024	0.00	943
2	1	0	0.053	0.004	0.00	7245
2	1	1	0.093	0.004	0.00	7245
2	3	0	-0.003	0.003	0.26	9549
2	3	1	-0.013	0.003	0.00	9549
2	3	2	-0.081	0.004	0.00	9549
3	0	0	0.535	0.121	0.00	92
3	1	0	0.159	0.020	0.00	736
3	1	1	0.187	0.020	0.00	736
3	2	0	0.026	0.003	0.00	10831
3	2	1	0.023	0.003	0.00	10831
3	2	2	0.066	0.004	0.00	10831

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in average wages for firms that change layers - DADS (Robustness checks)

Table A28

DADS

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.073	0.003	0.00	10432
0	2	0	-0.172	0.011	0.00	1350
0	3	0	-0.320	0.061	0.00	111
1	0	0	0.100	0.003	0.00	11356
1	2	0	-0.016	0.002	0.00	17052
1	2	1	-0.220	0.003	0.00	17052
1	3	0	-0.044	0.007	0.00	1168
1	3	1	-0.295	0.014	0.00	1168
2	0	0	0.181	0.009	0.00	1698
2	1	0	0.044	0.001	0.00	17927
2	1	1	0.234	0.003	0.00	17927
2	3	0	0.009	0.001	0.00	14228
2	3	1	-0.037	0.002	0.00	14228
2	3	2	-0.168	0.004	0.00	14228
3	0	0	0.274	0.044	0.00	142
3	1	0	0.064	0.006	0.00	1493
3	1	1	0.293	0.012	0.00	1493
3	2	0	0.014	0.001	0.00	15303
3	2	1	0.045	0.002	0.00	15303
3	2	2	0.138	0.003	0.00	15303

Table A29

Conditioning on $d \ln VA > 0$

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.076	0.004	0.00	5475
0	2	0	-0.178	0.015	0.00	760
0	3	0	-0.370	0.081	0.00	71
1	0	0	0.101	0.003	0.00	6491
1	2	0	-0.017	0.002	0.00	9211
1	2	1	-0.214	0.004	0.00	9211
1	3	0	-0.037	0.010	0.00	638
1	3	1	-0.281	0.019	0.00	638
2	0	0	0.183	0.012	0.00	1054
2	1	0	0.041	0.002	0.00	9739
2	1	1	0.226	0.004	0.00	9739
2	3	0	0.009	0.002	0.00	7482
2	3	1	-0.032	0.003	0.00	7482
2	3	2	-0.162	0.005	0.00	7482
3	0	0	0.227	0.052	0.00	85
3	1	0	0.064	0.009	0.00	858
3	1	1	0.287	0.017	0.00	858
3	2	0	0.014	0.002	0.00	8186
3	2	1	0.043	0.003	0.00	8186
3	2	2	0.133	0.005	0.00	8186

Table A30

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} N_L^{\ell} > 0$

# of layers Before	# of layers After	Layer	Change	s.e.	p-value	obs
0	1	0	-0.105	0.005	0.00	3717
0	2	0	-0.225	0.017	0.00	583
0	3	0	-0.509	0.101	0.00	50
1	0	0	0.127	0.004	0.00	4562
1	2	0	-0.034	0.003	0.00	6231
1	2	1	-0.232	0.005	0.00	6231
1	3	0	-0.058	0.012	0.00	423
1	3	1	-0.328	0.022	0.00	423
2	0	0	0.218	0.014	0.00	806
2	1	0	0.058	0.002	0.00	6642
2	1	1	0.239	0.005	0.00	6642
2	3	0	0.001	0.002	0.50	4758
2	3	1	-0.045	0.004	0.00	4758
2	3	2	-0.180	0.007	0.00	4758
3	0	0	0.293	0.060	0.00	69
3	1	0	0.084	0.011	0.00	615
3	1	1	0.303	0.019	0.00	615
3	2	0	0.026	0.002	0.00	5401
3	2	1	0.054	0.003	0.00	5401
3	2	2	0.142	0.006	0.00	5401

Refer to Appendix B for a detailed description of how these tables were constructed.

Change in average wages for firms that change layers - DADS (Robustness checks)

Table A31

Conditioning on $d \ln \sum_{\ell=0}^{L} N_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.116	0.004	0.00	5847
0	2	0	-0.248	0.014	0.00	875
0	3	0	-0.493	0.079	0.00	72
1	0	0	0.135	0.004	0.00	6775
1	2	0	-0.040	0.002	0.00	9338
1	2	1	-0.244	0.004	0.00	9338
1	3	0	-0.080	0.011	0.00	650
1	3	1	-0.342	0.018	0.00	650
2	0	0	0.237	0.012	0.00	1146
2	1	0	0.064	0.002	0.00	9998
2	1	1	0.248	0.004	0.00	9998
2	3	0	-0.005	0.002	0.00	7369
2	3	1	-0.055	0.003	0.00	7369
2	3	2	-0.191	0.005	0.00	7369
3	0	0	0.341	0.050	0.00	104
3	1	0	0.088	0.009	0.00	903
3	1	1	0.317	0.016	0.00	903
3	2	0	0.028	0.002	0.00	8416
3	2	1	0.058	0.003	0.00	8416
3	2	2	0.147	0.004	0.00	8416

Table A32

Conditioning on $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.073	0.003	0.00	10432
0	2	0	-0.172	0.011	0.00	1350
0	3	0	-0.320	0.061	0.00	111
1	0	0	0.100	0.003	0.00	11356
1	2	0	-0.013	0.002	0.00	13695
1	2	1	-0.176	0.003	0.00	13695
1	3	0	-0.047	0.008	0.00	1065
1	3	1	-0.279	0.014	0.00	1065
2	0	0	0.181	0.009	0.00	1698
2	1	0	0.039	0.002	0.00	14006
2	1	1	0.172	0.003	0.00	14006
2	3	0	0.009	0.001	0.00	11947
2	3	1	-0.025	0.002	0.00	11947
2	3	2	-0.129	0.004	0.00	11947
3	0	0	0.274	0.044	0.00	142
3	1	0	0.061	0.007	0.00	1342
3	1	1	0.265	0.012	0.00	1342
3	2	0	0.015	0.001	0.00	12864
3	2	1	0.031	0.002	0.00	12864
3	2	2	0.095	0.003	0.00	12864

Table A33

Conditioning on $d \ln VA > 0$ & $d \ln \sum_{\ell=0}^{L} n_L^{\ell} > 0$

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.076	0.004	0.00	5475
0	2	0	-0.178	0.015	0.00	760
0	3	0	-0.370	0.081	0.00	71
1	0	0	0.101	0.003	0.00	6491
1	2	0	-0.013	0.002	0.00	7589
1	2	1	-0.179	0.004	0.00	7589
1	3	0	-0.039	0.010	0.00	595
1	3	1	-0.275	0.019	0.00	595
2	0	0	0.183	0.012	0.00	1054
2	1	0	0.037	0.002	0.00	7758
2	1	1	0.172	0.004	0.00	7758
2	3	0	0.009	0.002	0.00	6459
2	3	1	-0.022	0.003	0.00	6459
2	3	2	-0.129	0.005	0.00	6459
3	0	0	0.227	0.052	0.00	85
3	1	0	0.060	0.010	0.00	770
3	1	1	0.256	0.017	0.00	770
3	2	0	0.015	0.002	0.00	6948
3	2	1	0.033	0.003	0.00	6948
3	2	2	0.096	0.004	0.00	6948

Refer to Appendix B for a detailed description of how these tables were constructed.

Firms that change export status and transition

Table A34

$d \ln n^{\ell}_{Lit}$ conditioning on adjacent layers

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	1.701	0.083	0.00	332
0	2	0	2.159	0.207	0.00	49
0	3	0	2.209	0.421	0.00	6
1	0	0	-1.819	0.091	0.00	291
1	2	0	1.182	0.065	0.00	577
1	2	1	1.084	0.060	0.00	577
1	3	0	1.332	0.192	0.00	53
1	3	1	1.152	0.178	0.00	53
2	0	0	-2.155	0.293	0.00	46
2	1	0	-1.191	0.066	0.00	517
2	1	1	-1.053	0.066	0.00	517
2	3	0	1.508	0.062	0.00	590
2	3	1	1.534	0.066	0.00	590
2	3	2	1.297	0.065	0.00	590
3	0	0	-2.569	0.353	0.00	11
3	1	0	-1.692	0.254	0.00	45
3	1	1	-1.447	0.260	0.00	45
3	2	0	-1.695	0.063	0.00	637
3	2	1	-1.679	0.067	0.00	637
3	2	2	-1.488	0.064	0.00	637

Table A35

$d \ln w^{\ell}_{Lit}$ conditioning on adjacent layers

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.062	0.020	0.00	332
0	2	0	-0.283	0.115	0.02	49
0	3	0	-0.641	0.611	0.34	6
1	0	0	0.096	0.023	0.00	291
1	2	0	0.003	0.012	0.83	577
1	2	1	-0.062	0.015	0.00	577
1	3	0	-0.062	0.049	0.21	53
1	3	1	-0.148	0.053	0.01	53
2	0	0	0.216	0.072	0.00	46
2	1	0	0.040	0.012	0.00	517
2	1	1	0.071	0.016	0.00	517
2	3	0	0.012	0.013	0.37	590
2	3	1	0.008	0.014	0.00	590
2	3	2	-0.074	0.016	0.00	590
3	0	0	0.019	0.127	0.89	11
3	1	0	0.085	0.056	0.14	45
3	1	1	0.119	0.070	0.10	45
3	2	0	0.020	0.011	0.08	637
3	2	1	0.011	0.012	0.40	637
3	2	2	0.079	0.014	0.00	637

Table A36

$d \ln w^{\ell}_{Lit}$ conditioning on DADS wages

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.064	0.012	0.00	528
0	2	0	-0.256	0.051	0.00	95
0	3	0	-0.766	0.263	0.01	15
1	0	0	0.127	0.014	0.00	520
1	2	0	-0.012	0.006	0.03	1132
1	2	1	-0.219	0.013	0.00	1132
1	3	0	-0.096	0.030	0.00	91
1	3	1	-0.272	0.048	0.00	91
2	0	0	0.271	0.054	0.00	100
2	1	0	0.052	0.006	0.00	1119
2	1	1	0.226	0.013	0.00	1119
2	3	0	0.009	0.005	0.06	861
2	3	1	-0.038	0.010	0.00	861
2	3	2	-0.160	0.015	0.00	861
3	0	0	0.305	0.154	0.07	16
3	1	0	0.047	0.014	0.00	105
3	1	1	0.386	0.052	0.00	105
3	2	0	0.019	0.005	0.00	872
3	2	1	0.047	0.009	0.00	872
3	2	2	0.165	0.015	0.00	872

Refer to Appendix B for a detailed description of how these tables were constructed.

Table A37: Elasticity of w_L^ℓ with VA for firms that do not change L
Robustness checks

Conditioning on selected sample					
# of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
0	0	0.067	0.009	0.00	39,914
1	0	0.106	0.009	0.00	42,071
1	1	0.118	0.009	0.00	42,071
2	0	0.145	0.007	0.00	71,424
2	1	0.155	0.007	0.00	71,424
2	2	0.170	0.007	0.00	71,424
3	0	0.173	0.010	0.00	53,053
3	1	0.187	0.010	0.00	53,053
3	2	0.189	0.011	0.00	53,053
3	3	0.218	0.011	0.00	53,053
With DADS wages					
# of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
0	0	0.000	0.003	0.94	45,606
1	0	0.001	0.002	0.63	65,114
1	1	0.019	0.003	0.00	65,114
2	0	-0.005	0.002	0.01	91,833
2	1	0.006	0.002	0.01	91,833
2	2	0.022	0.003	0.00	91,833
3	0	-0.007	0.002	0.05	53,053
3	1	0.008	0.002	0.00	53,053
3	2	0.009	0.003	0.01	53,053
3	3	0.038	0.006	0.00	53,053

Table A38: Elasticity of w_L^ℓ with VA for firms that change export status and do not change L
Robustness checks

Conditioning on selected sample					
# of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
0	0	0.105	0.024	0.00	2,720
1	0	0.134	0.023	0.00	4,550
1	1	0.120	0.024	0.00	4,550
2	0	0.165	0.019	0.00	8,031
2	1	0.177	0.019	0.00	8,031
2	2	0.182	0.021	0.00	8,031
3	0	0.199	0.033	0.00	4,896
3	1	0.219	0.034	0.00	4,896
3	2	0.218	0.034	0.00	4,896
3	3	0.219	0.035	0.00	4,896
With DADS wages					
# of layers	Layer	γ_L^ℓ	s.e.	p-value	obs
0	0	0.030	0.009	0.00	3,263
1	0	0.006	0.005	0.25	6,968
1	1	0.016	0.009	0.08	6,968
2	0	-0.006	0.005	0.19	10,507
2	1	0.010	0.006	0.09	10,507
2	2	0.017	0.008	0.03	10,507
3	0	-0.010	0.005	0.07	4,896
3	1	0.011	0.008	0.17	4,896
3	2	0.010	0.009	0.30	4,896
3	3	0.010	0.016	0.53	4,896

Table A39: Firms that satisfy a hierarchy in hours, weighted by VA				
# of layers	$N_L^\ell \geq N_L^{\ell+1}$ all ℓ	$N_L^0 \geq N_L^1$	$N_L^1 \geq N_L^2$	$N_L^2 \geq N_L^3$
1	88.8%	88.8%	-	-
2	63.2%	79.2%	76.3%	-
3	57.1%	77.7%	73.4%	98.1%

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

Table A40: Firms that satisfy a hierarchy in wages, weighted by VA				
# of layers	$w_L^{\ell+1} \geq w_L^\ell$ all ℓ	$w_L^1 \geq w_L^0$	$w_L^2 \geq w_L^1$	$w_L^3 \geq w_L^2$
1	94.1%	94.1%	-	-
2	96.2%	98.1%	98.1%	-
3	87.6%	99.2%	99.1%	89.2%

Robustness checks, conditioning on selected sample

Table A41: Firms that change export status and do not change L					
# of layers	Layer	β_L^ℓ	s.e.	p-value	obs
1	0	-0.045	0.044	0.31	4,550
2	0	0.009	0.026	0.73	8,031
2	1	-0.013	0.028	0.64	8,031
3	0	0.200	0.053	0.00	4,896
3	1	0.073	0.038	0.06	4,896
3	2	0.084	0.042	0.05	4,896

Table A42: Elasticity of n_L^ℓ with VA for firms that do not change L					
# of layers	Layer	β_L^ℓ	s.e.	p-value	obs
1	0	0.027	0.014	0.06	42,071
2	0	0.036	0.009	0.00	71,424
2	1	0.013	0.011	0.26	71,424
3	0	0.109	0.014	0.00	53,053
3	1	0.048	0.013	0.00	53,053
3	2	0.037	0.013	0.01	53,053

Table A43

Log diff. in hourly wage (after minus before the transition) for hours staying in the layer

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	0.001	0.00	0.79	8779
0	2	0	-0.072	0.02	0.00	953
0	3	0	-0.338	0.13	0.01	68
1	0	0	0.114	0.00	0.00	9645
1	2	0	0.022	0.00	0.00	15118
1	2	1	0.022	0.00	0.00	9358
1	3	0	-0.034	0.01	0.02	981
1	3	1	-0.034	0.02	0.10	536
2	0	0	0.243	0.02	0.00	1264
2	1	0	0.059	0.00	0.00	16048
2	1	1	0.086	0.00	0.00	10055
2	3	0	0.020	0.00	0.00	13455
2	3	1	0.028	0.00	0.00	11975
2	3	2	0.037	0.00	0.00	8912
3	0	0	0.557	0.13	0.00	80
3	1	0	0.111	0.01	0.00	1276
3	1	1	0.165	0.02	0.00	723
3	2	0	0.039	0.00	0.00	14508
3	2	1	0.046	0.00	0.00	12948
3	2	2	0.049	0.00	0.00	10348

Table A44

Log diff. in hourly wage of hours entering the layer (after transition) versus hours leaving the layer (before transition)

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.268	0.01	0.00	7564
0	2	0	-0.571	0.03	0.00	1133
0	3	0	-0.954	0.13	0.00	94
1	0	0	0.233	0.01	0.00	7848
1	2	0	-0.130	0.00	0.00	13375
1	2	1	-0.391	0.01	0.00	11406
1	3	0	-0.246	0.02	0.00	982
1	3	1	-0.515	0.02	0.00	929
2	0	0	0.527	0.02	0.00	1321
2	1	0	0.072	0.00	0.00	13707
2	1	1	0.378	0.01	0.00	11530
2	3	0	-0.087	0.00	0.00	12604
2	3	1	-0.154	0.00	0.00	10045
2	3	2	-0.339	0.01	0.00	10329
3	0	0	1.059	0.12	0.00	123
3	1	0	0.199	0.02	0.00	1226
3	1	1	0.497	0.02	0.00	1137
3	2	0	-0.033	0.00	0.00	13584
3	2	1	0.021	0.00	0.00	10771
3	2	2	0.188	0.01	0.00	10450

Table A45

Log diff. in hourly wage of new hours entering the layer versus hours staying in the layer (after transition)

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	-0.139	0.00	0.00	6216
0	2	0	-0.104	0.01	0.00	763
0	3	0	-0.099	0.04	0.02	61
1	0	0	0.029	0.00	0.00	8294
1	2	0	-0.097	0.00	0.00	12269
1	2	1	-0.156	0.01	0.00	4681
1	3	0	-0.086	0.01	0.00	843
1	3	1	-0.123	0.02	0.00	352
2	0	0	0.069	0.01	0.00	1131
2	1	0	-0.015	0.00	0.00	13854
2	1	1	0.037	0.00	0.00	6821
2	3	0	-0.073	0.00	0.00	12360
2	3	1	-0.106	0.00	0.00	8930
2	3	2	-0.168	0.01	0.00	5710
3	0	0	0.037	0.03	0.21	69
3	1	0	0.030	0.01	0.00	1157
3	1	1	0.026	0.02	0.14	552
3	2	0	-0.056	0.00	0.00	13451
3	2	1	-0.060	0.00	0.00	10241
3	2	2	0.019	0.01	0.00	7980

Table A46

Log diff. in hourly wage of hours leaving the layer versus hours who stayed in the layer (before the transition)

# of layers		Layer	Change	s.e.	p-value	obs
Before	After					
0	1	0	0.091	0.00	0.00	8148
0	2	0	0.140	0.01	0.00	912
0	3	0	0.185	0.03	0.00	58
1	0	0	-0.052	0.00	0.00	6738
1	2	0	0.049	0.00	0.00	13614
1	2	1	0.114	0.00	0.00	6929
1	3	0	0.091	0.01	0.00	919
1	3	1	0.176	0.02	0.00	448
2	0	0	-0.038	0.01	0.00	981
2	1	0	-0.013	0.00	0.00	13097
2	1	1	-0.070	0.01	0.00	4900
2	3	0	0.033	0.00	0.00	12646
2	3	1	0.055	0.00	0.00	9729
2	3	2	0.174	0.01	0.00	7313
3	0	0	-0.065	0.03	0.03	71
3	1	0	-0.020	0.01	0.03	1090
3	1	1	-0.046	0.02	0.02	468
3	2	0	0.018	0.00	0.00	13547
3	2	1	0.011	0.00	0.00	9805
3	2	2	-0.011	0.01	0.07	6464

9 Appendix B - Data Description

Our dataset is built from two data sources, a firm-level source and a worker-level source, both collected from the French National Statistical Institute (INSEE). We cover the manufacturing sector of metropolitan France for the years 2002-2007. The firm-level source contains balance-sheet information for all the firms reporting their income under the Bénéfice Réel Normal (BRN) fiscal regime. This regime is compulsory for firms above a certain revenue threshold, but it can still be adopted by smaller firms. The ratio between the value added of manufacturing firms in our original, uncleaned BRN dataset and value added in manufacturing as reported by the French National Statistical Institute is 96.4% on average. Each row in this dataset contains, among other things, a firm identifier, total employment, total wages and employer-paid payroll taxes, total value added, domestic, export and total sales, and an industry classification. The worker-level data source is the Déclarations Annuel des Données Sociales (DADS). This dataset is built on mandatory employer filing of the earnings of each salaried employee in France subject to the French payroll taxes in a given year. Each row of this dataset is an employment spell and contains, among other things, a worker identifier, his or her occupation, a firm identifier (which will be matched to the BRN dataset), the number of hours worked, and the total gross wage received by the worker from the firm.

9.1 Data processing

We start with the firm-level dataset, keeping only firms in the manufacturing sector. There are in total 553,125 firm-year observations. We drop some existing duplicated firm-year identifiers and then all firms with non-positive value added, total employment, total labor cost (i.e. total wages plus payroll taxes) or total sales. This leaves us with 11.5% fewer observations. We then move to the worker-level dataset. Starting from the universe of all observations, we keep the observations referring to all employees in any firm in the French private sector. We drop observations with missing or non-positive hours or total wage or with missing occupations. At this point, we match this dataset with the firm-level dataset based on the firm identifier and year and keep the observations that are matched (workers for which we find the firm and vice versa). This gives us a dataset of about 25.6 million observations for 6 years.[16]

To recover the occupational structure at the firm level, we work with the occupational code reported in the worker-level data. The occupational classification used in the DADS is the PCS-ESE 2003, and its first digit identifies 5 occupational categories relevant for manufacturing firms: firm owners and CEO (code 2), senior staff or top management positions (code 3), supervisors (code 4), white collar workers (code 5), blue collar workers (code 6).[17] We relabel code 6 into code

[16] Note that the number of observations in any given year does not correspond to total employment. A worker can have more than one row, for example, because of a job change during the year.

[17] In this process, we find 599 observations of workers categorized as farmers (code 1). Since the firms in our sample are only in the manufacturing sectors, we exclude from the analysis the 127 firm-year observations associated to these workers.

5, in order to create a unique category of blue and white collar workers, since their hourly wage distribution coincides in the data (see Table 1). We are left with 4 occupational categories: for each firm-year, we sum total hours and total wages of all the observations with the same occupation to recover the occupational structure at the firm level. During the matching we lose about 5.9% of the original dataset. Finally, we trim away firm-year observations with average hourly wages (see below for details) above the 99.95^{th} percentile, which would otherwise cause abnormal swings in the average wage by year. Our final dataset is composed of 456,638 firm-year observations. These observations represent on average 90.7% of the value added in the manufacturing sector in France.

To compute the average hourly wage at each occupation, we follow two alternative approaches. The first simply divides the total wage resulting from the DADS by the total number of hours in the same source. This approach doesn't include payroll taxes and some other worker-related expenditures paid by the firm. To approximate more closely the effective labor cost of each occupation type for the firm, we compute the share of wages paid to a given occupation in the total wages paid by the firm in the DADS source; we then apply this share to the total labor cost as in the BRN dataset to compute the labor cost of the occupation considered. We finally divide this total cost by the number of hours reported in the DADS to obtain the hourly wage from the BRN source. This latter estimate is our preferred measure of hourly wages; we will use the former measure for purposes of comparison and robustness checks.

We finally recode the name of the occupations into layers of management. A firm reporting c occupational categories will be said to have $L = c - 1$ layers of management: hence, in our data we will have firms that have from 0 to 3 layers of management. We drop the name of the occupation and assume that firms always grow from the lowest (occupations 5+6) to the highest skill (occupation 2), irrespective of the name of the occupation itself. Hence a firm with occupational categories 3 and 5 will have 1 layer of management, and its organization will consist of a layer 0 corresponding to blue and white collar workers, and a layer 1 corresponding to senior staff. This choice raises the question, what is the fraction of firms with the 'correct' ordering of layers? Define a firm with adjacent layers as a firm that only 1) reports occupations in adjacent categories, 2) starting from occupation 5+6. In this definition, firms with occupations 2 and 4 or 2 and 3 are not firms with adjacent layers; a firm with occupational categories 4 and 5 is. Table 4 shows that on average, 81.6% of the firm-years in the sample (representing 96.7% of value added and 95.7% of hours worked) meet this requirement. After this recoding, we can compute normalized hours. We define normalized hours, n_L^ℓ, in layer $\ell = 0, ..., 3$ in a firm with L layers of management as the number of hours in layer l per unit of hour worked in the top layer.

In results involving export behavior, we also exclude from the analysis firms that in some year have negative domestic or export sales; the sample in this case is of 456,559 firm-year observations.

All monetary values are deflated to 2005 euros using a CPI deflator for France. Value added is always expressed in thousands of euros; all hourly wages are in euros.

9.2 Description of tables and figures

9.2.1 Definitions

Some concepts are recurring in the explanation of a majority of the tables and figures. We define them here and consider them understood in what follows.

Average hourly wage from BRN: the total labor cost resulting from the balance sheet divided by the number of hours in the DADS source.

Average hourly wage in layer ℓ from BRN: the total labor cost for layer ℓ in the BRN divided by the number of hours reported in the DADS; the total labor cost for layer ℓ from the BRN is the share of wages paid to layer ℓ in the total wages paid by the firm as in the DADS source, multiplied by the total labor cost in the BRN.

Average hourly wage from DADS: the total wage payments in the DADS to all occupations divided by the number of hours, always from the DADS source.

Firms with adjacent layers or firms in the 'selected sample': it is a firm that 1) reports occupations in adjacent 1 digit PCS-ESE occupational categories, 2) starting from occupation 5+6 (blue and white collar workers). For examle, firms with occupations 2 and 4 or 2 and 3 do not have adjacent layers; a firm with occupational categories 4 and 5 does and has 1 layer of management.

Layer number: is the position of the workers in the hierarchy of the firm, starting from 0 (lowest layer, present in all firms) to 3 (highest layer, only present in firms with 3 layers of management).

Normalized hours: normalized hours in layer $\ell = 0, ..., L$ in a firm with L layers of management are the number of hours in layer ℓ per unit of hour worked in the top layer L.

Number of layers of management: the total number of 1 digit occupations in the firm from the DADS source minus 1.

Representative hierarchy of an L layers of management firm: for each layer $\ell = 0, ..., L$, we compute the average number of hours and the average hourly wage in the layer from the BRN among firms with L layers of management. To mitigate the impact of outliers, the set of firms used to compute each average is all the firms with the given number of layers, trimming observations below the 0.05^{th} and above the 99.95^{th} percentile of the outcome variable. In a figure portraying a firm with L layers of management, $L + 1$ rectangles are shown, one for each layer. The horizontal length of each rectangle is proportional to the number of hours in the layer (hours are reported along the $x-$axis) while its height is proportional to the average hourly wage (the value of the wage is reported to the left of each rectangle).

9.2.2 Descriptions

Table 1: This table reports, for each 1-digit occupational code present in the PCS-ESE 2003, mean and percentiles of the hourly wage distribution across all firms and years in the data. One observation in an occupation is the average hourly wage in a given firm-year from the BRN source, conditional on the firm reporting the occupation. The average hourly wage is the total labor

cost from the BRN dataset for an occupation, divided by the number of hours reported in this occupation. The total labor cost for an occupation is computed multiplying the total labor cost from the firm balance sheet times the share of wages paid to the occupation as resulting from the DADS source. Occupation 6 excludes one outlier, which would have driven its mean to 28.94.

Table 2: This table reports, for each year, the number of firms in the dataset and corresponding averages across all firms for selected variables. *VA* is the value added in the firm's balance sheet. *Hours* is the average number of hours from the DADS source. *Wage* is the average hourly wage from the BRN. *# of layers* is the average number of layers of management across firms in each year.

Table 3 and Table A1: Table 3 reports summary statistics on firm-level outcomes, grouping firm-year observations according to the number of layers of management reported (*# of layers*). *Firm-years* is the number of firm-years observations in the data with the given number of layers of management. *VA* is the average value added from the firm's balance sheet. *Hours* is the average number of total hours from the DADS source. *Wage* is the average hourly wage from the BRN; to mitigate the impact of outliers, the set of firms used to compute the average is all the firms with the given number of layers, trimming observations below the 0.05^{th} and above the 99.95^{th} percentile of the hourly wage distribution; the untrimmed averages across all firms in the cell are 30.48, 21.54, 21.21, and 22.41 for layers 0 to 3, respectively. *Median Wage* is the median across all firms in the cell of the average hourly wage from the BRN source. Table A1 was constructed in the same way using wages from DADS.

Table 4: This table reports the fraction of firms with adjacent layers conditioning on the number of layers of management in the firm (first four columns) and overall (fifth column). The first row reports the simple fraction of firms; the second and third rows assign a weight that is proportional to the total value added in the balance sheet, and to the total hours in the DADS, respectively.

Table 5, Table 6, Table A39 and Table A40: Table 5 reports the fraction of firms that satisfy a hierarchy in hours, grouping firms by their number of layers of management (*# number of layers*). *Hours* N_L^ℓ is the number of hours reported in layer ℓ in an L layers of management firm from the DADS source. For $L = 1, 2, 3$, and $\ell = 0, ..., L-1$, we say that a firm satisfies a hierarchy in hours between layers number ℓ and $\ell + 1$ in a given year if $N_L^\ell \geq N_L^{\ell+1}$, i.e. if the number of hours worked in layer ℓ is at least as large as the number of hours worked in layer $\ell + 1$; moreover, we say that a firm satisfies a hierarchy at all layers if $N_L^\ell \geq N_L^{\ell+1}$ $\forall l = 0, ..., L-1$, i.e. if the number of hours worked in layer ℓ is at least as large as the number of hours in layer $\ell + 1$, for all layers in the firm. Following these definitions, the top panel reports, among all firms with $L = 1, 2, 3$ layers of management, the fraction of those that satisfy a hierarchy in hours at all layers (first column), and the fraction of those that satisfy a hierarchy in hours between layer ℓ and $\ell + 1$, with $\ell = 0, ..., L-1$ (second to fourth column). Table A39 reports the same information when weighting each firm according to the share of its value added among all firms with L layers. Table 6 is the same as Table 5 for the case of wages, where w_L^ℓ is the average hourly wage in layer ℓ

from the BRN in an L layers of management firm. Table A40 reports the same information when weighting each firm according to the share of its value added among all firms with L layers.

Table 7: Table 7 reports statistics on the mean share of variation in log hourly wages within a firm associated to variation in average wages between layers. For each firm-year, we compute the R^2 of a regression of log hourly wages of workers within a firm on a constant and dummies for layers (all except one), weighted by the number of hours each worker provides to the firm. For each row, the column *unweighted* reports the average R^2 across all firm-years, while the remaining two columns to the right report the same average when weighting firms by their total number of hours or total value added. The column firm-years reports the number of firm-years used to compute the statistics in the corresponding row (note that for some firms - for example firms with only one worker - the R^2 cannot be computed, and hence the total number of firm-year in the dataset does not correspond to the total number of firm-years used). Each row differs from the others according to the subsample of firm-years used in computing the average R^2.

Table 8 and Table A3: Table 8 reports the distribution of the number of layers of management at time t+1, grouping firms according to the number of layers of management at time t. Among all firms with L layers of management ($L = 0, ..., 3$) in any year from 2002 to 2006, the columns report the fraction of firms that have layers $0, ..., 3$ the following year (from 2003 to 2007), or are not present in the dataset, *Exit*. The elements in the table sum to 100% by row. Table A3 reports the same information when weighting each firm according to the share of its value added.

Table 9 and Table 10: Table 9 shows the slope and robust standard errors of a set of regressions of log number of normalized hours worked at a given layer on log value added of the firm across firm-year observations. Specifically, each row $\ln n_L^\ell / n_L^L$ reports the results of two regressions. The columns with heading *Without FE* (*With FE*) report the slope and the robust standard error of a regression of log normalized hours in layer ℓ on log value added among firms with L layers of management, in the raw data (controlling for year and 2 digit industry fixed effects). To mitigate the impact of outliers, the set of firms used to compute each regression is all the firms with the given number of layers, trimming observations below the 0.05^{th} and above the 99.95^{th} percentile of each variable. Table 10 was constructed in the same way for the case of hourly wages.

Table 11 and Table A42: These two tables report the results of regressions of log change in normalized hours by layer on log change in value added for firms that do not change their number of layers of management L across two adjacent periods. Specifically, we run a regression of log change in normalized hours at layer ℓ (*layer*) in a firm with L (*# of layers in the firm*) layers of management on a constant and log change in value added across all the firms that stay at L layers of management across two adjacent years. β_L^ℓ is the coefficient on log change in value added, *s.e.* and *p-value* are its robust standard error and p-value, and *obs* is the number of observations in the regression. The first table (Table 11) uses all the firms that are observed to stay at the same layer, while the second one (Table A42) replicates it considering only firms with adjacent layers before and after the transition.

Table 12 and Table A37: These tables report the results of regressions of log change in

hourly wage by layer on log change in value added for firms that do not change their number of layers of management L across two adjacent periods. Specifically, we run a regression of log change in average hourly wage at layer ℓ (*layer*) in a firm with L (*# of layers in the firm*) layers of management on a constant and log change in value added across all the firms that stay at L layers of management across two adjacent years. β is the coefficient on log change in value added, *s.e.* and *p-value* are its robust standard error and p-value, and *obs* is the number of observations used in the regression. Table 12 uses hourly wage at layer ℓ from the BRN. The first panel in Table A37 uses the selected sample, all the firms that are observed to stay at the same layer. The second panel uses hourly wage at layer ℓ from DADS.

Table 13: This table shows changes in firm-level outcomes between adjacent years for all firms (*All*), and for the subsets of those that increase (*Increase L*), don't change (*No change in L*) and decrease (*Decrease L*) layers. It reports changes in log hours, log normalized hours, log average wage from the BRN, and log average wage in common layers for the whole sample. The change in average wage for common layers in a firm that transitions from L to L' layers is the change in the average wage from the BRN computed using only the first $\min\{L, L'\}$ layers before and after the transition. To detrend a variable, we subtract from all the log changes in a given year the average change during the year across all firms. In the last two rows of the table, *% of firms* is the percentage of firms observed doing each type of transition; *% VA change* is the fraction of the total change in real value added observed in the dataset accounted for by firms making the given transition.

Table 14 and Tables A6-A16: This set of tables show estimates of the average log change in normalized hours at each layer ℓ (*Layer*) among firms that transition from L (*# of layers before*) to L' layers (*# of layers after*), with $L \neq L'$: for a transition from L to L', we can only evaluate changes for layer number $\ell = 0, ..., \min\{L, L'\}$. *Change* is the average log change in the transition, estimated as a regression of the log change in the number of normalized hours in layer ℓ in two adjacent years on a constant. *s.e.* is the robust standard error associated with the constant, *p-value* its p-value, and *obs* the number of observations used in the regression. Different tables differ according to the sample of observations used to compute the change. The first table (Table 14) uses all observed transitions in the sample; all the other tables condition on positive (negative) changes in the indicated variables when firms add (drop) layers; in particular, observations are selected on changes in value added (A6), in value added and total hours (A7), in total hours of the firm (A8), in the total normalized hours of the firm (A9), and in value added and total normalized hours (A10). The table series *Selected sample* (Tables A11-A16) replicates the same set of 6 regression tables when we consider only firms with adjacent layers before and after the transition.

Table 15 and Tables A17-A33: This set of tables show estimates of the average log change in hourly wage at each layer ℓ among firms that transition from L (*# of layers before*) to L' layers (*# of layers after*), with $L \neq L'$: for a transition from L to L', we can only evaluate changes for layer number $\ell = 0, ..., \min\{L, L'\}$. *Change* is the average log change in the transition, estimated as a regression of the log change in the average hourly wage at layer ℓ in two adjacent years on

a constant. *s.e.* is the robust standard error associated with the constant, *p-value* its p-value, and *obs* the number of observations used in the regression. We start by describing tables that use hourly wage at each layer ℓ from the BRN (a correspondent set of tables replicate the results using average hourly wage at each layer from DADS). Different tables differ according to the sample of observations used to compute the change. The first table (Table 15) uses all observed transitions in the sample; all the other tables condition on positive (negative) changes of the indicated variables when firms add (drop) layers; in particular, observations are selected on changes in value added (A17), in value added and total hours (A18), in total hours of the firm (A19), in the total normalized hours of the firm (A20), and in value added and total normalized hours (A21). The Table series *Selected sample* (Tables A22-A27) replicates the same set of 6 regression tables when we consider only firms with adjacent layers before and after the transition. The table series *DADS* (Tables A28-A33) shows the average log change in hourly wage at each layer ℓ from DADS, using all firms.

Table 16: This table shows the sources of change in the average hourly wage from the BRN, by type of transition. For a given firm transitioning from L to $L' > L$ layers, write the log change in average wage as $d\ln \bar{w}_{Lit} = \ln \bar{w}_{L'it+1} - \ln \bar{w}_{Lit} = \ln\left(\frac{\bar{w}^{\ell\leq L}_{L'it+1}}{\bar{w}_{Lit}} s + \frac{w^{L'}_{L'it+1}}{\bar{w}_{Lit}}(1-s)\right)$. In this notation, $\bar{w}^{\ell\leq L}_{L'it+1}$ is the average hourly wage after the transition in the common layers, $\bar{w}_{Lit}$ is the average hourly wage before the transition, s is the share of hours of the common layers after the transition, and $w^{L'}_{L'it+1}$ is the average wage in the layers added after the transition. In the data, we form the quantities $\frac{\bar{w}^{\ell\leq L}_{L'it+1}}{\bar{w}_{Lit}}$, s and $\frac{w^{L'}_{L'it+1}}{\bar{w}_{Lit}}$ for each firm and report in the cells the average of each of these quantities in the first three panels; the fourth panel shows the overall average log change $d\ln \bar{w}_{Lit}$ in hourly wage during the indicated transition. Each cell is computed excluding observations below the 0.05^{th} and above the 99.95^{th} percentile.

Table 17: This table shows summary statistics of firm-year variables broken down by export status of the firm. *VA* is the value added in the firm's balance sheet. *Hours* is the number of hours from the DADS source. *Wage* is the average hourly wage in the firm across all occupations. *# of layers* is the number of layers of management in the firm.

Table 18: This table shows the fraction of firms in the data that have positive export sales, grouping firms by their number of layers of management (*# of layers*). *Unweighted* reports the fraction of exporters in the raw data. *Weighted by VA* reports the fraction of firms that are exporters, where each firm receives a weight proportional to its value added in the year considered.

Table 19: This table shows the composition of firms by number of layers of management (*# of layers*), broken down by export status.

Table 20 and Table A5: Table 20 shows, for each transition from L to L', the difference between the fraction of domestic sellers entering the export market that make the transition and the fraction of domestic sellers not entering the export market that make the same transition. Each number is computed as follows. Consider all the firms that have L layers in a given period, and either do not export before and after, or do not export before but export after; among all these firms, let d be an indicator variable equal to 1 if the firm is transitioning from L to L' in a given

period, and zero otherwise; we run a regression of d on a constant and a dummy equal to 1 if the firm is entering the export market. Robust standard errors are used to compute the significance levels. Table A5 replicates the same analysis for the case of all firms that exit the export, the ones that drop layers and the ones that do not change layers.

Table 21: This table focuses on firms entering the export market and shows changes in firm-level outcomes between adjacent years for all firms (*All*), and for the subsets of those that increase (*Increase L*), and don't change (*No change in L*) layers. It reports changes in log hours, log normalized hours, log average wage from the BRN, and log average wage in common layers for the whole sample. It is parallel to Table 13 and shares the same content and methodology. The time trend is computed with respect to all firms in the sample, not just new exporters.

Table 22 and Table A41: Table 22 reports the results of regressions of log change in normalized hours by layer on log change in value added for firms that do not change their number of layers of management L across two adjacent periods and either enter or exit the export market. Specifically, we run a regression of log change in normalized hours at layer ℓ (*layer*) in a firm with L (*# of layers in the firm*) layers of management on a constant and log change in value added across all the firms that stay at L layers of management across two adjacent years. β is the coefficient on log change in value added, *s.e.* and *p-value* are its robust standard error and p-value, and *obs* is the number of observations used in the regression. This table shares with Table 11 the content and the methodology. Table A41 reports the same as Table 22 only using the selected sample.

Table 23 and Table A38: Table 23 reports the results of regressions of log change in hourly wage by layer on log change in value added for firms that do not change their number of layers of management L across two adjacent periods and either enter or exit the export market. Specifically, we run a regression of log change in average hourly wage at layer ℓ (*layer*) in a firm with L (*# of layers in the firm*) layers of management on a constant and log change in value added across all the firms that stay at L layers of management across two adjacent years. β is the coefficient on log change in value added, *s.e.* and *p-value* are its robust standard error and p-value, and *obs* is the number of observations used in the regression. This table shares with Table 12 the content and the methodology. The first panel in Table A38 uses the selected sample, all the firms that are observed to stay at the same layer. The second panel in Table A38 uses hourly wage at layer ℓ from DADS.

Table 24 and Table A34: Table 24 shows estimates of the average log change in normalized hours at each layer ℓ among firms that enter (exit) the export market and concurrently add (drop) layers of management. Specifically, the table reports estimates of the average log change in normalized hours at each layer ℓ (*Layer*) among firms that transition from L (*# of layers before*) to L' layers (*# of layers after*), with $L \neq L'$: for a transition from L to L', we can only evaluate changes for layer number $\ell = 0, ..., \min\{L, L'\}$. *Change* is the average log change in the transition. This table shares with table 14 the content and the methodology. Table A34 presents the results conditioning on the selected sample.

Table 25, Table A35, and Table A36: Table 25 shows estimates of the average log change in hourly wage at each layer ℓ from the BRN among firms that enter (exit) the export market

and concurrently add (drop) layers of management. Specifically, the table reports estimates of the average log change in hourly wage at each layer ℓ (*Layer*) among firms that transition from L (*# of layers before*) to L' layers (*# of layers after*), with $L \neq L'$: for a transition from L to L', we can only evaluate changes for layer number $\ell = 0, ..., \min\{L, L'\}$. *Change* is the average log change in the transition. This table shares with table 15 the content and the methodology. Table A35 presents the results using the selected sample, all the firms that are observed to stay at the same layer. Table A36 presents the results using hourly wage at layer ℓ from DADS.

Table 26: This table shows the sources of change in the average hourly wage from the BRN, by type of transition, focusing on firms entering the export market. The contents and methodology are the same as those reported in Table 16.

Table A4: This table reports the fraction of firms that have adjacent layers at time t+1, conditional on having adjacent layers at time t. *Transition Up* reports, among all firms with L adjacent layers of management in any year (from 2002 to 2006) that stay in the sample the year after, the fraction of those moving to $L+1$ adjacent layers of management, with $L = 0, 1, 2$. *Transition Down* reports, among all firms with L adjacent layers of management in any year (from 2002 to 2006) that stay in the sample the year after, the fraction of those moving to $L-1$ adjacent layers of management, with $L = 1, 2, 3$.

Table A43-A46: These tables report the sources of changes in the average wage by layer during a transition. Some introductory notation will aid clarity. For any given firm, denote with i an employee, with $h(i)$, $w(i)$ and $\ell(i)$ the hours worked, the total wage received, and his or her wage before a transition, and let us use primes to denote the same outcomes after a transition, $h'(i)$, $w'(i)$ and $\ell'(i)$. If an employee is not present in the firm before a transition then $\ell(i) = -1$, and $h(i) = w(i) = 0$ (and analogously if an employee is not present after a transition). In the employee level dataset for year t, a given employee's row reports his or her outcomes for year t (after a transition) and year $t-1$ (before a transition), provided the employee stays in the same firm. Given the way the data is reported, an employee may have $\ell(i) = -1$ even if he or she was in the same firm but switched to a different plant, or switched occupation in the middle of the year (rather than at the end of the preceding year): in fact, more rows of data may be present for the same employee. For these reasons, our processing implies an over-estimation of hours leaving the layer and entering the layer, as opposed to hours staying in the layer during a transition. To track the flow of hours across layers we reconstruct transitions observed from year $t-1$ to year t by using only the employee level dataset for year t (rather than using separately the datasets at time t and time $t-1$, which would lose such flow). In a number of cases (less than 1/30,000) we are not able to match perfectly the layer structure of the firm with the one recovered by only using information of employees at time t in each year. For a given firm transitioning from L to L' layers, fix a layer $\bar{\ell}$ which is common to both before and after the transition. Hours worked in $\bar{\ell}$ after the transition can be grouped in hours worked by employees coming 1) from the same layer or 2) from outside the layer. For each of these groups, we can compute an average hourly wage: denote these two hourly wages with $\bar{w}'\left(\bar{\ell}\right) = [\sum_{i:\ell'(i)=\ell(i)=\bar{\ell}} w'(i)] \,/\, [\sum_{i:\ell'(i)=\ell(i)=\bar{\ell}} h'(i)]$, and $\bar{w}'\left(-\bar{\ell}\right) = [\sum_{i:\ell'(i)=\bar{\ell}\neq\ell(i)} w'(i)]$

/ $[\sum_{i:\ell'(i)=\bar{\ell}\neq\ell(i)} h'(i)]$, respectively. Similarly, we can group hours worked in $\bar{\ell}$ before the transition in hours worked by employees who will 1) stay in the same layer or 2) leave the layer, and compute their average hourly wage: denote these two hourly wages with $\bar{w}\left(\bar{\ell}\right) = [\sum_{i:\ell'(i)=\ell(i)=\bar{\ell}} w(i)]$ / $[\sum_{i:\ell'(i)=\ell(i)=\bar{\ell}} h(i)]$, and $\bar{w}\left(-\bar{\ell}\right) = [\sum_{i:\ell(i)=\bar{\ell}\neq\ell'(i)} w(i)]$ / $[\sum_{i:\ell(i)=\bar{\ell}\neq\ell'(i)} h(i)]$, respectively. Tables A43- A46 report differences between these averages across all firms where both quantities can be computed, for each transition and each layer common to before and after the transition. Table A43 reports the average of $\ln\left(\bar{w}'\left(\bar{\ell}\right)/\bar{w}\left(\bar{\ell}\right)\right)$, i.e., the average change in the hourly wage for hours worked by employees who don't change layer. Table A44 reports the average of $\ln\left(\bar{w}'\left(-\bar{\ell}\right)/\bar{w}\left(-\bar{\ell}\right)\right)$, i.e., the average log difference between the hourly wage after the transition of hours which came from outside the layer and the hourly wage before the transition of hours which are going to leave the layer. Table A45 reports the average of $\ln\left(\bar{w}'\left(-\bar{\ell}\right)/\bar{w}'\left(\bar{\ell}\right)\right)$, i.e., the average log difference, after the transition, in the hourly wage of hours coming from outside the layer vs. hours who stayed in the layer. Table A46 reports the average of $\ln\left(\bar{w}\left(-\bar{\ell}\right)/\bar{w}\left(\bar{\ell}\right)\right)$, i.e., the average log difference, before the transition, in the hourly wage of hours who will leave the layer vs. hours who will stay in the layer.

Figure 2, Figure 3, and Figure 4: These figures report kernel density estimates of the distribution of log value added (Figure 2), log hours worked (Figure 3) and log hourly wage (Figure 4) by number of layers in the firm. We describe Figure 2; Figures 3 and 4 are constructed in the same way but instead of using value added, we use hours worked and hourly wage, respectively. The left panel reports the kernel density estimate of the distribution of log value added on the raw data: one density is estimated for each group of firms with the same number of layers. The right panel shows the same density estimated after removing year and industry fixed effects. To remove these effects, we run a regression of the form $\log v_i = \alpha + \sum_j \beta_j layers_{ij} + \sum_i \delta_i industry_{ij} + \sum \gamma_i year_{ij} + \varepsilon_i$, where v_i is value added for firm-year i, and $layers_{i.}$, $industry_{i.}$, and $year_{i.}$ are a set of layers, 2 digit industry and year dummies, respectively. The omitted dummy for layers is for firms with zero layers of management. The log value added for firm-year i without year and industry fixed effects is then $\log \hat{v}_i = \log m + \sum_j \hat{\beta}_j layers_{ij} + \hat{\varepsilon}_i$, where we set m to the median value added in 2002 for firms with zero layers of management. We then compute 4 kernel density estimates of the distribution of $\log \hat{v}_i$, grouping firms according to their number of layers.

Figure 7 and Figure A1: These figures show the probability of transition away from the current layer as a function of the initial value added of the firm. Figure A1 shows these probabilities in the raw data. Each panel reports transition probabilities starting from a different initial number of layers. To produce the panel of transitions out of layer $L = 0, ..., 3$, we take for each year (from 2002 to 2006) all the firms with L layers of management and group them into 100 bins according to their value added; for each bin, we compute the fraction of firms that will have any number of layers (or exit the dataset) in the following period and plot the average value added in the bin against this fraction. In Figure 7, for each transition series we then apply a lowess smoothing for all the probabilities estimated from the first to the 99^{th} bin.

Figures 8: This figure portrays, for each transition type, the estimated log change (on the

y-axis) in the percentiles (on the x-axis) of the wage distribution within firms, after the transition vs. before the transition, and associated 95% bootstrapped confidence intervals. To compute these panels, we first construct a employee-level dataset that contains log hourly wage of each employee, a firm identifier, year and current number of layers of the firm, and we remove year and firm fixed effects from the log hourly wage distribution. We then focus on all the hours worked in firm-years observations where the firm is making a transition from ℓ to ℓ' layers (both before and after the transition takes place) and compute the p^{th} percentile (for $p = 1, ..., 99$) in two distributions: the wage distribution after the transition and the wage distribution before the transition. Each distribution of hourly wages is computed making sure that each employee receives a weight proportional to his or her number of hours in the firm, but giving to each firm the same weight, regardless of the total number of hours worked in it (to be consistent with our estimates of changes in log wages in Table 15 in the paper). The y-axis in each panel reports the log difference between the wage at the p^{th} percentile after, less its correspondent wage before the transition (hence, a negative number implies a shift to the left of the distribution at the given percentile). To compute standard errors, we performed 500 bootstrap replications of this process, clustering the sampling at firm-transition level (i.e., one cluster contains all the employees present either before or after a transition in one firm) to preserve the within-firm and within-transition correlation in wages present in the data, and we report the 5^{th} and 95^{th} percentiles of these replications.

Figures 9: This figure portrays, for each transition type, the estimated log change (on the y-axis) in the percentiles (on the x-axis) of the wage distribution in common layers within firms, after the transition vs. before the transition, and associated 95% bootstrapped confidence intervals. To build it we follow the same process described in Figure 8, to which we refer, with the only difference that after removing year and firm fixed effects, we focus on the wage distribution implied by all hours worked in layers which are common before and after the transition.

Figure 10 and Figure A12: Figure 10 portrays, for firms staying at a given number of layers in two consecutive years and positive change in value added, the estimated log change (on the y-axis) in the percentiles (on the x-axis) of the wage distribution within firms, the year after less the year before, and associated 95% bootstrapped confidence intervals. To build it we follow the same process described in Figure 8. Figure A12 portrays the estimated log change (on the y-axis) in the percentiles (on the x-axis) of the wage distribution within firms, for firms staying at a given number of layers in two consecutive years and negative change in value added.

Figures 11-13: These figures depict hierarchies of firms before and after a transition, using normalized hours: each panel focuses on transitions from L to $L+1$ layers of management and from $L+1$ to L, for $L = 0, 1, 2$. In each panel, for fixed L, the left column portrays the representative hierarchy with L layers of management and normalized hours, using only firms that will make the indicated transition. The right column estimates the hierarchy after the transition using the average log changes in each quantity as resulting from our estimates reported in Tables 14 and 15: the wage (normalized hours) in layer ℓ after the transition from L to $L+1$ is computed as the average hourly wage (normalized hours) before the transition multiplied times $\exp(b)$, where

b is the average log change in wages (normalized hours) at layer ℓ in transitions from L to $L+1$. Normalized hours after a transition are always set to 1 for the top layer. An analogous procedure is followed for transitions from $L+1$ to L. For transitions one layer up, the hourly wage for the top layer after the transition is computed as follows: we estimate the average log change in the wage of the top layer ($\ln w_{L+1}^{L+1} - \ln w_L^L$) and multiply the wage at the top layer before the transition times the exponential of this change.

Figures 14-16: These figures portray hierarchies of firms before and after a transition, using normalized hours, for firms entering the export market: each panel of four graphs focuses on transitions from L to $L+1$ layers of management (first row), and from L to L (second row), for $L = 0, 1, 2$. To estimate the hierarchies before and after a transition (first row of graphs), we consider only the subset of firms with L layers that will enter the export market and have $L+1$ layers the following period. We use these firms to compute the representative hierarchy before the transition (top left graph). To estimate the representative hierarchy after the transition (top right graph), we follow the procedure indicated in Figure 11 and use the estimated log changes for firms entering the export market from Tables 24 and 25. To estimate the hierarchies for firms that do not switch while entering (second row of graphs) we consider only the subset of firms with L layers that will enter the export market and stay at L layers the following period. We use this set of firms to compute the representative hierarchy before the transition (bottom left graph). To compute the change in each layer level outcome, we regress the change in log outcome on change in log value added, with no constant; each layer level quantity after the transition is computed as the layer level quantity before the transition times the exponential of the predicted log change associated with the mean increase in value added observed for these firms.

Figures A3-A9: Figures A3-A5 show the relation between the number of normalized hours worked at each layer against the value added in the firm, grouping firms by number of layers of management across firm-year observations. Each of the 3 panels fixes the number of layers of management $L = 1, 2, 3$ and shows L scatterplots, showing the relation between the number of normalized hours worked in each layer $\ell = 1, .., L$ against the value added in the firm. To mitigate the impact of outliers, the set of firms used to produce each graph is all the firms with the given number of layers, trimming observations below the 0.05^{th} and above the 99.95^{th} percentile of each variable. The regression line fits a linear relation between log normalized hours and log value added and is computed using the same trimming procedure. Figures A6-A9 replicated the same for the case of average hourly wage from the BRN at each layer ℓ.

Figure A10 and Figure A11: These figures show the change in the hierarchical structure of a firm with 2 layers of management associated with (first panel) a 5% increase in value added when the firm adds a third layer (first row) or stays at 2 layers (second row), and (second panel) a 5% decrease in value added when the firm drops one layer (first row) or stays at 2 layers (second row). Consider the first panel (the second panel will follow an analogous procedure). To estimate the hierarchies before and after a transition (first row of graphs), we consider only the subset of firms with 2 layers that will become 3 layers the following period, increasing value added and total

hours. We group these firms in three terciles according to the size of their value added before the transition, and we use the middle tercile of these firms to compute the representative hierarchy before the transition (top left graph). To estimate the representative hierarchy after the transition, we regress the log change in layer level outcomes (normalized hours and hourly wage for each layer $\ell = 0, ..., 2$) on a set of three dummies (one for each tercile) and log change in value added, using the same subset of firms; again, to mitigate the impact of outliers, each of the six regressions is run on the subsample of firms trimmed of the observations below the 0.05^{th} and above the 99.95^{th} percentile of each variable. Each layer level quantity after the transition is computed as the layer level quantity before the transition times the exponential of the predicted log change for a firm in the middle tercile following a 5% change in value added. The average hourly wage of the third layer after the transition is computed using changes in the top layer (consistent with Figures 10 - 12), following the same procedure described here. To estimate the change in the hierarchy for firms that do not transition (second row of graphs) we proceed as follows. We compute the representative hierarchy before the transition using firms in the middle tercile of value added among those that do not transition. To compute the change in each layer level outcome we regress the change in log outcome on change in log value added, with no constant, using the set of all firms staying at 2 layers (trimming as above); each layer level quantity after the transition is computed as the layer level quantity before the transition times the exponential of the predicted log change associated with a 5% increase in value added.