Sorting Within and Across French Production Hierarchies*

Grigorios Spanos†

April 14, 2015

Abstract

The objective of this paper is to examine the assignment of workers to layers and firms. In particular, I use an administrative dataset of French workers to study the organization of manufacturing firms. First, I test whether higher ability workers are employed in the higher layers of firms. Second, I test whether there is positive assortative matching between workers in the different layers of firms. Third, I test whether abler managers supervise more workers. Finally, I test whether higher ability workers allow their managers to increase their span of control and employ larger teams. To do this, I first classify employees as residing in different organizational layers such as production and administrative workers, supervisors, senior managers, and owners and CEOs, using occupational codes. From a panel wage regression I then obtain estimates of workers' ability as in Abowd et al. (1999). I then study how workers sort into layers and across layers with other workers. I emphasize four results. First, higher ability workers are employed in the higher layers of firms. Second, I find evidence of positive assortative matching between workers in the different layers of firms. Third, I find evidence that abler managers supervise less workers. Finally, I also find weak evidence that higher ability workers allow their managers to increase their span of control and employ larger teams.

*I wish to thank Kunal Dasgupta, Gilles Duranton and Peter Morrow for excellent supervision. I am indebted to Pierre-Philippe Combes for helping me access and use the data. This paper also benefited from insightful discussions with Bernardo Blum, April Franco, Luis Garicano, Camille Hemet, Francis Kramarz, Nicholas Li, Matthieu Solignac, Marc Sagnier, Michel Serafinelli, Brian Tavares, Daniel Trefler, Ronald Wolthoff and Tanguy van Ypersele, and seminar participants at Brock University, the International Trade Workshop at the University of Toronto, and the AMSE Eco Lunch Seminar at Marseilles. Finally, I would like to thank the staff of the French National Statistical Institute, INSEE, in Paris, for their kindness and support. This project was conducted during my stay at the Aix-Marseilles School of Economics, and funded by the AMSE Pre-Doctoral Research Grant. All remaining errors are my own.

†Aix-Marseille School of Economics. E-mail: grigorios.spanos@utoronto.ca

1 Introduction

Ever since Coase (1937) economists have known that one of the most important problems a firm faces is how to organize inputs efficiently. However, classical economic models often abstract from firms' organizational decisions. A firm is like a black box, whereby inputs are directly mapped into a final good. However, an understanding of how firms organize is essential, because firms determine the allocation of productive resources in the economy.[1] One important organizational decision of firms is what types and how many workers they should hire, as well as what tasks should be assigned to which workers.

Despite much theoretical interest, very little is known empirically about how workers sort together in firms.[2] Several researchers have investigated whether good workers are employed in productive firms. While most empirical studies are concerned with how workers match with firms, far fewer studies have examined the different tasks workers perform, and whether better workers are employed with better workers in the other positions of a firm.

This paper fills the gap by examining how workers sort together in firms. My empirical strategy relies on the idea that firms can be thought of as hierarchical teams, composed of layers that perform different tasks. The lowest layer of a firm, for example, contains workers who focus on production, while higher layers contain individuals that perform managerial tasks. With this in mind, I examine how workers sort into teams and layers within each team. More precisely, within a team I first test whether higher ability workers are employed in higher layers. Second, across teams, I test whether there is positive assortative matching, in which the ability of an individual in one layer is positively correlated with the ability of a worker in another layer. Third, I investigate whether this sorting pattern is caused by higher ability workers allowing their managers to increase their span of control and employ larger teams, as suggested by Antras et al. (2006).

I use an administrative dataset of French workers, the Declarations Annuelles des Donnees Sociales (DADS) to test these predictions. I begin by classifying employees as residing in the different organizational layers of firms. With my dataset, I observe four distinct layers, production and administrative workers, supervisors, senior managers, and owners and CEOs, by using occupational codes. The concept of a layer that I use is from the management hierarchy theory of the firm that was introduced by Garicano (2000) and used empirically by Caliendo et al. (2014). In theory a layer corresponds to a set employees who earn similar wages, are of similar ability and perform tasks at a similar level of authority. Since firms are hierarchical teams, layers have the added property that, within a firm, higher layers contain fewer workers who are of greater

[1] As noted by Rosen (1982): 'The firm cannot be analyzed in isolation from other production units in the economy. Rather, each person must be placed in his proper niche, and the marriage of personnel to positions and to firms must be addressed directly.'

[2] For example, several studies have used models of firm organization to investigate earnings inequality (Garicano and Rossi-Hansberg (2006)), offshoring (Antras et al. (2006),Antras et al. (2008)) and knowledge diffusion (Dasgupta (2012)).

ability.

For every firm in the dataset, I calculate the total number of layers in the firm, and the size of each layer, in terms of labor hours worked. With my dataset I can observe four different types of organizations, one-layer firms, two-layer firms, three-layer firms, and four-layer firms. I show that this classification of employees into layers is meaningful and consistent with the concept of a layer discussed above.

Then for the years 1993 to 2004, I use the panel dimension of my dataset to obtain estimates of workers' ability. I estimate a Mincerian wage regression with individual fixed effects, as in Abowd et al. (1999). I use the individual fixed effects from my regression as my measure of worker ability.

Using these measures of the size and number of layers of firms, along with measures of worker ability in these layers, I test my main predictions. First, I conclude that higher ability workers are employed in the higher layers of firms. For example for four-layer firms, I find that an individual with a one hundred percent increase in his ability will on average reside 0.511 layers higher. Second, I find evidence of positive assortative matching between workers in the different layers of firms. For example, in four-layer firms, a one hundred percent increase in the average ability of workers in layer one is associated with a 0.320 increase in the average ability of workers in layer two. Third, I find evidence that abler managers supervise less workers. For example, in three-layer firms a one unit increase in the average ability of managers in layer three is associated with a 23.1 percent decrease in their span of control. Finally, I find only weak evidence that higher ability workers allow their managers to increase their span of control. For example, in four-layer firms, a one unit increase in the average ability of workers in layer three is associated with a 33.4 percent increase in their span of control.

In the last part of the paper, I address robustness of my results by assessing several potential threats to my empirical strategy. First, one concern with the empirical analysis is that my worker fixed effects are inconsistent. Because the worker fixed effects are incidental parameters from a wage regression, they can only be measured consistently as the number of years an individual is observed in the panel grows large. To resolve this issue, I conduct my analysis on a restricted sample of worker fixed effects for workers that I observe for at least 10 periods. Second, as discussed in Andrews et al. (2008) my measures of workers' ability may be misestimated and any positive correlation between the individual fixed effects is the result of a positive correlation between the estimated error of the individual fixed effects. To address this issue, I conduct my empirical analysis outside of the sample, for the year 2008, and only on the set of workers who have moved to a firm that they have never been employed in before. Taking both potential threats into account, I continue to find that higher ability workers are employed in the higher layers of firms, that the ability of individuals in one layer of a firm is positively correlated with the ability of workers in another layer, and only weak evidence that higher ability workers allow their managers to increase their span of control, and evidence that higher ability managers supervise less workers.

This paper is related to the broad literature on the theory of the firm allowing for management hierarchies. With the aim of explaining the distributions of firm size and earnings in the economy, a long-standing literature has examined how productive factors are allocated to managers with different abilities (for example Lucas (1977) and Rosen (1982)). To motivate my empirical strategy, I use a model by Garicano and Rossi-Hansberg (2006) in which agents with different cognitive abilities sort into occupations, layers and teams. Regardless of the distribution of knowledge in the economy, the equilibrium displays skill stratification, in the sense that agents with similar levels of cognitive ability sort into the same occupations and layers across firms. Agents with the least amount of knowledge become production workers, while agents with high levels of ability sort into managerial layers which correspond to higher layers of firms. The equilibrium also displays positive assortative matching, in the sense that higher ability managers organize into firms with higher ability subordinates. The mechanism behind this result is the following: in a given layer, agents of greater ability can solve a greater proportion of problems, and thus render their subordinates more productive. In turn, because they can solve a greater proportion of problems alone, higher ability subordinates require less of their superiors' time. This frees up the latter's time and allows managers to supervise more workers.

This paper is most closely related to Garicano and Hubbard (2005) who examine positive assortative matching between partners and associates in law firms in Texas.[3] Using data on lawyers' school of education and firm of employment, they find that associates are more likely to work at the same firm as partners who went to a similarly ranked school, consistent with positive assortative matching. The nature of their data, however, does not permit them to obtain a measure of workers' ability that varies across individuals who graduated from the same school. In addition their analysis is limited to two-layer firms: partners and associates. My dataset and classification strategy allow me to make progress on this issue, since I can identify up to four layers in firms. Finally, another distinction is that I examine the mechanism that is causing this sorting pattern: higher ability managers supervising larger teams.

More generally, this paper is also related to a large empirical literature examining sorting in labor markets. In particular, this literature is concerned with whether productive workers are matched with productive firms (see for example Abowd et al. (2003), Abowd et al. (2004), Martins (2008), Andrews et al. (2008)). My paper examines a related but different question: how workers sort with other workers in layers and firms.

My paper is most closely related to de Melo (2013), who uses matched employer-employee data from Brazil to examine whether workers of similar ability sort into the same firms. While Lopes de Melo focuses exclusively on whether workers of similar ability sort into the same firm, I also examine how workers sort into the different layers of firms, and the mechanism

[3]Iranzo et al. (2008) investigate whether production and non-production workers are complements or substitutes. Although in their analysis, managers are contained in their non-production worker classification, they do not focus on the relationship between managers and production workers.

that is causing this pattern. In addition I make progress on an econometric issue associated with testing whether higher ability workers sort into the same firms. Specifically, a positive correlation between workers' ability may be due to standard estimation error. I address this issue by examining whether my findings hold in the year 2008, for individuals who have moved to a firm that they have never been employed in before.

The paper proceeds as follows: Section 2 presents a brief description of the management hierarchy theory of a firm and its predictions. Section 3 introduces the data. Section 4 discusses the estimation strategy. Section 5 presents the descriptive statistics and summary results, Section 6 tests the model's predictions, and Section 7 presents robustness checks. Finally, section 8 concludes.

2 Model

In this section I briefly present the knowledge-based management hierarchy models from Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) and discuss their main implications. To fix ideas, I first present the model where teams have three layers, and then describe the general setting in which teams have any number of layers. For a complete exposition and for all proofs, I refer the reader to Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) and Antras et al. (2008).

2.1 Setup of the Model

In the model, a unit of output is produced only when a problem is solved. Problems are differentiated by their difficulty and for simplicity, assume that the difficulty of problems is drawn from a uniform distribution with support $[0, 1]$. Agents have one unit of time and are heterogeneous in their level of knowledge.[4] Agents' knowledge determine their ability to solve problems, and assume that knowledge is cumulative: an individual with knowledge z can solve all problems in the interval $[0, z]$.[5]

Production occurs in teams. Teams are composed of one layer of production workers, who spend their one unit of time drawing problems and attempt to solve them, and layers of manager, who do not draw problems but instead spend all of their time solving problems that their production workers cannot solve.[6] For managers to receive problems that other agents cannot solve,

[4]I abstract from the decision to acquire knowledge. For a model where agents acquire knowledge see Garicano and Rossi-Hansberg (2006).

[5]The output of such an individual from working alone is therefore:

$$y(z) = z. \tag{1}$$

[6]In other words, production workers specialize in routine tasks (i.e. production), while managers specialize in nonrountine tasks (i.e. problem solving).

communication is possible between managers and production workers within a team. Communication, however, entails a cost to the managers of a team.

In a team, production workers draw one problem per unit of time and if they can solve the problem a unit of output is produced. Otherwise, they ask their manager in the immediate layer above who in turn spends a fraction of her time communicating with the worker. If the manager knows the solution to the problem, then she conveys the solution to her worker who immediately produces a unit of output. If the manager does not know the solution, the production worker asks the manager two layers above. This process continues, until the problem is solved, or the production worker has seen a manager in every layer of the firm, at which point the problem remains unsolved and nothing is produced.

Consider a team composed of three layers, one manager in layer three with knowledge z_m^3, n_m^2 managers in layer two with knowledge z_m^2 and n_p^1 production workers in layer one with knowledge z_p^1.[7] Let h the time cost per problem that a manager incurs communicating with a production worker and assume that this cost is the same across all managers. The number of managers in layer two, n_m^2, is determined by the number of problems in the team, the fraction of problems production workers cannot solve, and the communication costs, and is equal to:

$$n_p^1 h[1 - z_p^1] = n_m^2, \tag{2}$$

where $h[1 - z_p^1]$ is the total cost per unit of time that a manager in layer two incurs while working in a team composed of production workers with knowledge z_p^1. Similarly, since managers in layer two can solve z_m^2 fraction of the problems, the time constraint of a manager in layer three is equal to:

$$n_p^1 h[1 - z_m^2] = 1, \tag{3}$$

where $h[1 - z_m^2]$ is the total cost per unit of time that the top manager incurs while working in a team composed of managers in layer two with knowledge z_m^2. The communication technology therefore limits the amount of interactions managers can have with their subordinates, and this in turn determines the number of production workers a manager in layer two can supervise, n_p^1, and the number of managers in layer two, n_m^2. Since the manager in layer three receives all problems that her production workers and her managers in layer two cannot solve, and she can only solve a fraction z_m^3 of them, the output of the team is equal to:

$$y(z_m^3, z_m^2, z_p^1) = n_p^1 z_m^3. \tag{4}$$

This production framework has several important properties that determine the equilibrium allocation of managers and production workers to teams. First, regardless of their occupation,

[7] I slightly depart from the notation in Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) and refer to firms by the total number of layers instead of the number of management layers.

agents are not perfect substitutes in production. Second, in this framework matching is many to one and because they share their knowledge with other agents in team, managers increase the value of their knowledge and concentrate on problems that only they can solve. Third, production is asymmetrically sensitive to skill. Since they can leverage their knowledge over many workers, managers are more important to the output of a team. And fourth, managers in layer three, managers in layer two, and production workers are complements. The mechanism behind this result is the following. Managers of greater knowledge can solve a greater proportion of problems, and thus render subordinates more productive. In turn, more knowledgeable subordinates increase the productivity of their managers. Since all individuals have one unit of time, managers are constrained in the number of agents that they can supervise and because they can solve a greater proportion of problems on their own, more knowledgeable subordinates spend less time communicating with managers which allows the latter to supervise larger teams.

In equilibrium the assignment of agents into occupations and teams has the following properties. First because production is asymmetrically sensitive to skill, there is skill stratification in the sense that agents with greater knowledge sort into managerial occupations while agents with lesser knowledge become production workers. Second because managers in a given layer of a team are complements in production with managers in the other layers and with production workers, there is positive assortative matching between agents in the different layers of teams. Third, managers of greater knowledge supervise more subordinates and employ larger teams, and fourth subordinates of greater knowledge are employed in larger teams. These two properties follow from positive assortative matching between agents in the different layers of a firm and managers' time constraint.

Furthermore, these results are generalizable to an economy where firms can have any number of layers (see Garicano and Rossi-Hansberg (2006) for details). An important point to note is that any model with a production function that exhibits similar interactions between managers and production workers will, in a general equilibrium, yield similar results (Garicano and Hubbard (2008)). More specifically, as long as high skill agents raise the productivity of their subordinates, and better individuals require less supervision, then in equilibrium, high skill individuals will form firms with more and better subordinates in the layers below. These properties summarized in the proposition below:

Proposition 1 *With L layer firms, the equilibrium assignment of individuals to occupations and teams has the following properties:*

- *Individuals with the greatest knowledge sort into the top managerial occupations, individuals with intermediate knowledge sort into managerial occupations, while individuals with least knowledge become production workers.*
- *There is positive assortative matching between managers in the different layers of a firm and between managers and production workers.*

Figure 1: Equilibrium in Economy with Two-Layer Firms

(a) Assignment of Production Workers to Managers (b) Managers' Span of Control on Managers' Knowledge (c) Managers' Span of Control on Subordinates' Knowledge

- *For all layers, managers with greater knowledge supervise more individuals in the layers below.*

2.2 Discussion: Taking the Model to the Data

In this section thus far, I have presented the intuition of the model, and its predictions. The sections further below are concerned with estimating a measure of worker's ability, testing whether there is positive assortative matching between workers in the different layers of firm, and testing the mechanism that is driving this sorting pattern: able workers require less supervision, and thus allow their superiors to supervise more of them. Given my measure of ability, testing for positive assortative matching is straightforward. As done in the literature that investigates matching between workers and firms, to test whether there is positive assortative matching between workers in the different layers of firms, I estimate a correlation between the average ability of workers in the different layers of firms. To test whether abler managers supervise more subordinates, I rely on the equations that characterize managers' time constraint. Equations (2) and (3) can be generalized to a firm with L layers. Rearranging, taking logs, and defining the span of control of workers in layer $l+1$ as the ratio of the size of layer 1 to the size of layer $l+1$ one obtains the following expression:

$$\begin{aligned} \ln span^{l+1} &= \ln \frac{n^1}{n^{l+1}} \\ &= \ln h - \ln[1 - z^l]. \end{aligned} \tag{5}$$

where n^1 denote the size of layer 1, n^{l+1} denotes the size of layer $l+1$ and z^l denotes the ability of agents in layer l. As managers occupy all other layers except layer 1, I drop the subscripts p and m from the notation. These equations have the following characteristics in common: they are a function of the number of production workers, and depend on the knowledge of the individuals in the layer below. According to the model, therefore, managers' span of control should be increasing with their subordinates ability. To test this prediction, I approximate equation (5) with the following equation:

$$\ln span^{l+1} = \gamma_0 + \gamma_1 \widehat{z^l} + u, \tag{6}$$

where $\widehat{z^l}$ represents my measure of the ability of agents in layer l. If the mechanism described by the model, that determines how agents sort together into firms, holds in the data then the estimated coefficient γ_1 should be positive and significant.

Furthermore, since there is positive assortative matching, there is a one-to-one correspondence between the ability of individuals in the different layers of a firm. Therefore, equation (5) can also be rewritten as:

$$\ln span^{l+1} = \ln h - \ln[1 - f(z^{l+1})]. \tag{7}$$

where $f()$ is a function that maps the ability of workers in layer l to layer $l+1$; i.e. $z^l = f(z^{l+1})$. I approximate equation (7) with the following equation:

$$\ln span^{l+1} = \gamma_0 + \gamma_1 \widehat{z_{l+1}} + u, \tag{8}$$

where $\widehat{z_{l+1}}$ represents my measure of the ability of agents in layer $l+1$. If the mechanism described by the model holds in the data then the estimated coefficient γ_1 should be positive and significant.

3 Data Description & Classification of Layers

The data are extracted from the Déclarations Annuelles des Données Sociales (DADS), which are provided and maintained by the French National Statistical Institute for Statistics and Economic Studies (INSEE). The DADS are matched employer-employee datasets and are constructed from administrative records that must be completed by all employers in France. A report must be filled by each establishment for every one of its employees, so there is a unique record for each employee-establishment-year combination. The DADS contains two datasets: a panel of workers born in October and that runs from 1976 to 2008, and from 1993 to 2008, exhaustive cross-sections of all workers in mainland France.[8]

In both the panel and cross-section datasets, for each observation, there is information on employees' characteristics, such as age, gender, and occupation, basic information on the establishment, such as location, industry and the parent firm, and basic firm level information, such as the firm's industry. For each observation there is also information on annual earnings,

[8]Until 1993 the DADS only contained information on individuals born in October in an even numbered year. From 1993 onwards, the DADS contains information on all individuals born in October.

denominated in 2007 euros, number of days worked, and number of hours worked.[9]

As discussed further below, I use the panel dataset to obtain measures of workers' ability. For computational tractability, I restrict the sample to the years 1993 to 2004, and to all full-time workers who are born in October in an even numbered year, are between the ages 18 and 65 and work in mainland France. For the years 1993 to 2004, there are 4,999,728 observations, corresponding to 753,092 workers in 399,676 firms. Appendix A and Abowd et al. (1999) provide further details on the data and information on how wages are determined in France.

For the year 2004, I use information from the cross-section to measure the total number of layers in firms and the size of each layer.[10] In the management hierarchy theory of the firm by Garicano (2000) a layer corresponds to a set employees who earn similar wages, are of similar ability and perform tasks at a similar level of authority (Caliendo et al. (2014)).[11] To construct the different layers of firms, I adopt the strategy put forth by Caliendo et al. (2014), and use one-digit occupational codes, which range from 2 to 6, to classify employees into layers.[12] In total, I can classify employees into four distinct layers. Layer 1 corresponds to qualified and non-qualified administrative workers and blue-collar workers. It contains all workers with occupational codes 5 and 6. Layer 2 is composed of supervisors and individuals with a higher level of responsibility than ordinary workers, and contains all workers with an occupational code 4. Layer 3 is composed of senior directors and top management staff and contains all workers with an occupational code 3. And, layer 4 corresponds to owners who receive a wage and CEOs. It contains all workers with occupational code 2. I consider a firm to have a layer if there is at least one employee in the exhaustive cross-section employed there.

Finally, I merge the information from both the panel and exhaustive cross-section datasets together, and retain only firms that operate in the manufacturing sector, and firms that operate in only one industry and location.[13] In all, the matched dataset contains $23,916$ firms that operate in 17 industries, of which 2,160 are one-layer firms, 3,322 are two-layer firms, 7,860 are three-layer firms and 5,450 are four-layer firms.

Table 1 reports summary statistics of firms in the matched sample. In the table firms are grouped by their organizational structure, their total number of layers. The average firm in the sample has an organizational structure that is consistent with the knowledge-based management

[9] Information on the total number of hours worked is only available after 1993.

[10] Since the panel of the DADS is only a 5 percent sample of the population, it is not suitable to properly measure the total number of layers and the size of each layer in a firm. Appendix A and Caliendo et al. (2014) provide further details on the exhaustive cross-section data.

[11] The concept of a layer that I use is therefore independent of the actual occupations of employees, such as whether they are lawyers, engineers or computer programmers. Instead it depends on their knowledge, productive ability, and their relative position in the organizational hierarchy of firms. In addition, since firms are hierarchical, layers have the added property that within a firm higher layers contain less workers who are of greater ability.

[12] The occupational codes range from 1 to 6. I have removed all firms operating in the agricultural and fishing industries, which correspond to occupational code 1.

[13] Unlike the exhaustive cross-section, since the panel data is based on a 5 percent sample of the French population, it contains information on a sample of all firms operating in mainland France. Approximately 1 percent of firms in the panel dataset are not matched.

Table 1: Description of Manufacturing Firms by Total Number of Layers

Total Number of Layers	Average Number of employees	Average Number of hours	Average Wage	Average Ability	Median Wage	Median Ability	Standard Deviation Wage	Standard Deviation Ability
ONE								
1st layer	11.00	11,844.12	2.13	0.433	2.10	0.442	0.195	0.435
TWO								
1st layer	17.16	21,874.82	2.18	0.347	2.14	0.358	0.231	0.411
2nd layer	2.71	4,115.16	2.40	0.428	2.38	0.419	0.288	0.409
THREE								
1st layer	49.16	71,431.91	2.33	0.352	2.27	0.357	0.321	0.375
2nd layer	11.72	19,900.35	2.61	0.473	2.57	0.496	0.352	0.365
3rd layer	6.25	10,350.94	2.96	0.580	2.96	0.594	0.349	0.402
FOUR								
1st layer	58.57	85,769.69	2.36	0.359	2.29	0.365	0.335	0.381
2nd layer	14.90	25,643.13	2.64	0.501	2.61	0.519	0.361	0.372
3rd layer	8.91	14,515.65	3.00	0.574	2.98	0.582	0.362	0.389
4th layer	1.22	2,165.00	3.27	0.772	3.39	0.765	0.378	0.398

Notes: Descriptive statistics of manufacturing firms that are in both the exhaustive cross-section and panel datasets of the DADS, for the year 2004. *These statistics are reported separately for firms with different number of layers. Column* 1 *refers to the layer within a firm. Columns* 2 *report the average number of employees in a given layer, while column* 3 *reports the average number of hours worked by employees in a given layer. These measures are obtained from the exhaustive cross-section of the DADS. Columns* 4, 6 *and* 7 *report the average log hourly wages, median log hourly hourly wages and standard deviation of log hourly wages within a layer. Columns* 5, 7 *and* 9 *report the average ability, median ability and standard deviation of ability of workers in a given layer. Measures of wages and ability values are obtained from the panel dataset of the DADS. Ability is estimated from equation (9).*

hierarchical theory of a firm. On average lower layers of firms are larger than the layers above, and contain workers that earn lower wages and are of lower ability. For example for the average four-layer firm, layer 1 has 58.57 employees, layer 2 has 14.90 employees, layer 3 has 8.91 employees and layer 4 has 1.22 employees. The findings are similar if one measures the size of layers by the number of hours worked. Returning again to four-layer firms, the average log-hourly wages of workers in layer 4 are 3.27, the average log-hourly wages of workers in layer 3 are 3.00, and the mean log-hourly wages of workers in layers 2 and 1 are 2.64 and 2.36 respectively. Therefore there is a clear ranking in wages. The same ranking also holds for ability, where for example in four-layer firms workers who reside in layer 4 have the greatest average ability, and are succeeded by layers 3, 2, and 1. The classification of workers into layers, therefore has economic meaning. The evidence is consistent with the view that firms are hierarchies, in the sense that higher layers of a firm are smaller and contain workers of earning higher wages and are of greater ability.

4 Estimating Ability

To obtain measures of workers' ability from the data, I use the empirical approach of Abowd et al. (1999) which has been developed further by Card et al. (2013). I model log hourly wages,

w_{it}, of worker i in time t, as a linear function of a time-invariant worker component θ_i, a time-invariant firm-layer component $\psi_{J(i,l,t)}$, time varying worker characteristics x_{it}, and a mean-zero error term ϵ_{it}. The equation to be estimated is:

$$\ln w_{it} = x_{it}\beta + \theta_i + \psi_{J(i,l,t)} + \epsilon_{it}. \tag{9}$$

The term θ_i captures the portable part of a worker's wages that remain with him as he moves across firms, or layers within firms. The variation of this term reflects a worker's productivity, bargaining ability and labor market discrimination. In the subsequent analysis, I use θ_i as my measure of workers' ability. The terms x_{it} captures how workers' earnings evolve with changes in their observable attributes, such as labor market experience. In my estimation, I use age as a proxy for experience.[14] Although in theory, workers only form firms with other workers, I include in equation (9) firm-layer fixed effects, $\psi_{J(i,l,t)}$, which are meant to identify firm attributes that affect every worker's earnings in a given layer in a firm equally, such as compensation policies, bargaining strength in the labor market, and productivity. Alternatively, since not all workers are employed in the same firm throughout their career, one can interpret the firm-layer fixed effects as partially accounting for any permanent influences past employees may have on the current organization, or any influences that affect individuals' earnings in a given layer of a firm that are the result of workers in the other layers.

To identify all of the econometric parameters in equation (9), I assume, as in Abowd et al. (1999), that the error term ϵ_{it} is strictly exogenous. Under this assumption, the parameter β can be consistently estimated as the number of workers, N, the number of firm-layers, J, and the number of years, T, increases. The parameters θ_i and $\psi_{J(i,l,t)}$ can only be separately identified by workers who switch employers, or layers within employers in the panel. In the dataset, there are in total of 1,156,816 worker displacements. Since θ_i is an incidental parameter, consistent estimates for it can only be obtained as the number of years a worker is observed grows large. Table 1 in the Appendix A presents the distribution of the number of years a worker in observed in the panel. Over 50 percent of workers are observed for 6 years or more. Similarly, $\psi_{J(i,l,t)}$ can only be consistently estimated if the number of workers in a layer in a firm, or the number of years grows large. Table 2 in the Appendix A presents the distribution of the number of workers observed in a layer in a firm in a given year, as well as the number of years firms' layers are observed in the panel. The average number of workers in a layer in a firm is 2.67, and over 50 percent of firms' layers are observed for 2 years or more.

To estimate equation (9) I focus on the largest connected group, that is the largest group of layers within firms that, over the years, have had at least one employee in common with another layer in the same or a different organization. In the panel, the largest connected group contains

[14] Since in equation (9) age cannot be separately identified from worker and time fixed effects, I exclude any time trends from the analysis. Indeed one can show that age can be written as a linear combination of the time and worker fixed effects.

Table 2: Summary Statistics

Sample Year	$1993 - 2004$
Worker and Firm-Layer Parameters	
Number of Worker Effects	$753,092$
Number of Firm-Layer Effects	$569,198$
Summary of Parameter Estimates	
St. Dev. of Wages	0.4417
St. Dev. of Worker Effects	0.3940
St. Dev. of Firm Effects	0.2508
St. Dev. of $X\beta$	
RMSE of AKM Residual	0.1717
Adjusted R-Squared	0.8489
Correlations	
Wages & Worker Effects	0.2509
Wages & Firm Effects	0.5073
Worker Effects & Firm Effects	-0.1636
Comparison with the Match Effects Model	
RMSE of Match Model	0.1490
Adjusted R-Squared	0.8862
ADDENDUM	
Sample Size	$4,999,728$

Notes: Results from OLS estimation of equation (9). $X\beta$ includes age and age squared interacted with gender. The match model includes $X\beta$ the region effects and a separate dummy for each worker-firm pair, corresponding to a job.

753,092 workers and 569,198 layers within firms. To estimate equation (9), I use the algorithm put forth by Guimaraes and Portugal (2010), which builds on the algorithm of Abowd et al. (2003).[15]

5 Results & Descriptive Statistics

Table 2 summarizes the estimation results from regression (9). To summarize my findings I report the standard deviation of log hourly wages, of the worker and firm-layer effects, as well as of the time-varying observables. I also report the root mean squared error (RMSE) of the residuals and the adjusted R-squared of the estimation. One important point to note is that the standard deviation of the worker effects is less than the standard deviation of wages. In the model because workers of different abilities are more productive from working in firms rather than alone, individuals' wages are amplified relative to their ability, and hence the standard deviation of wages is greater than the standard deviation of abilities, consistent with the data.

In Table 2, I also report correlations. The correlation between the worker and firm-layer fixed effects is -0.1636. This finding is similar to the empirical literature that investigates how workers

[15] The output of the algorithm provides a non-unique set of solutions for the worker and firm fixed effects. To make the effects unique, the algorithm sets the average of the firm fixed effects to zero.

sort into firms. As many researchers report there is a negative correlation between worker and firm fixed effects, estimated from a log-linear wage equation. In my analysis, I abstract from this correlation, since I am concerned with how employees in each layer of a firm match, rather than how workers match with firms. Furthermore, the correlation between the individual fixed effects and log-hourly wages is 0.2509. Therefore, individuals of higher ability earn more.

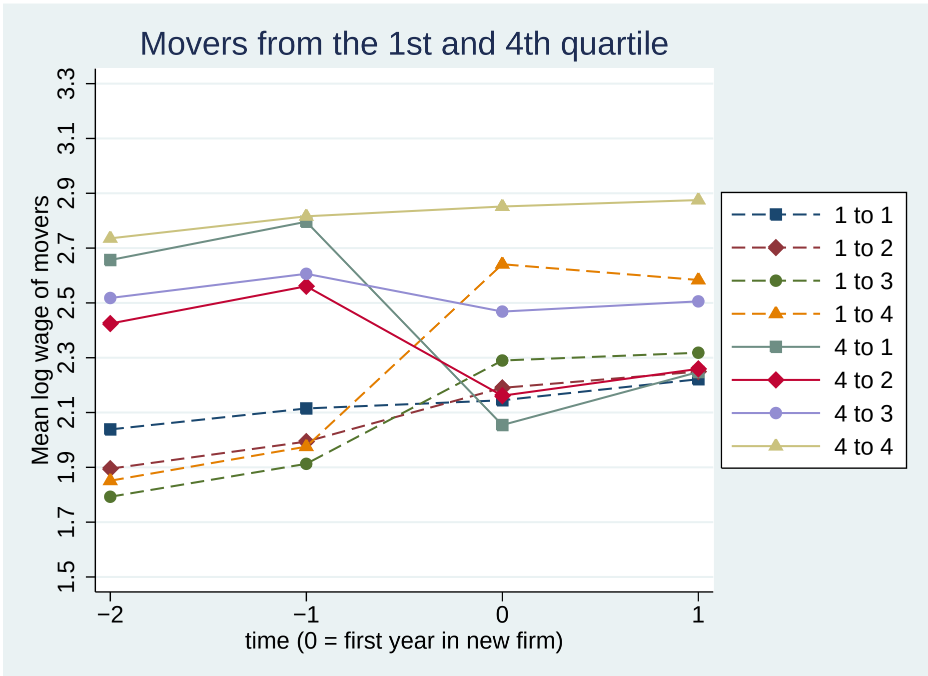
Table 2 also contains the adjusted R-squared and RMSE of a model with unrestricted match effects, that is a separate dummy for each worker-firm-layer job spell. If match effects are an important determinant of workers' wages, then a model with worker-firm-layer match effects should provide a markedly better fit to the data. The match effects model has an adjusted R-squared of 0.8862 and a RMSE of 0.1490, while the adjusted R-squared from the estimation of equation (9) is 0.8489 and the RMSE is 0.1717. The match effects model, therefore, fits the data slightly better than a specification with separate worker and firm-layer effects. Although this indicates that a match component is present in wages, the improvement in fit is modest.

As in Card et al. (2013), I further examine the importance of a match component to wages. In particular, I examine the wage dynamics of all individuals who changed firms, or layers within firms, in the years 1993 to 2004 with at least two consecutive years in the new and old position. I classify the origin and destination positions by the quartile of the estimated firm-layer effects and calculate the average hourly wages of agents in each cell two years before and after the move. I report the results in Table 3. If the error term in equation (9) is exogenous, changes in the wages of individuals who transition from one quartile to the other should be relatively symmetric, and individuals who move to new firms, or layer within firms, within the same quartile should not experience a wage gain. In addition, the increase in wages of workers who transition to different quartiles should be monotonically increasing with the distance of the quartiles. These conditions hold in Table 3. For visual aide Figure 2 panel (a) illustrates the wage profiles of workers in the first and fourth quartiles. The gains or losses of individuals who transition to quartiles is monotonically increasing with the distance between the quartiles, and the gains or losses are relatively symmetric. Panel (b) illustrates the wage profiles of workers that remain within the same quartile. These profiles are relatively flat. Therefore, at a minimum, the model in equation (9) is a relatively decent first approximation to wages.

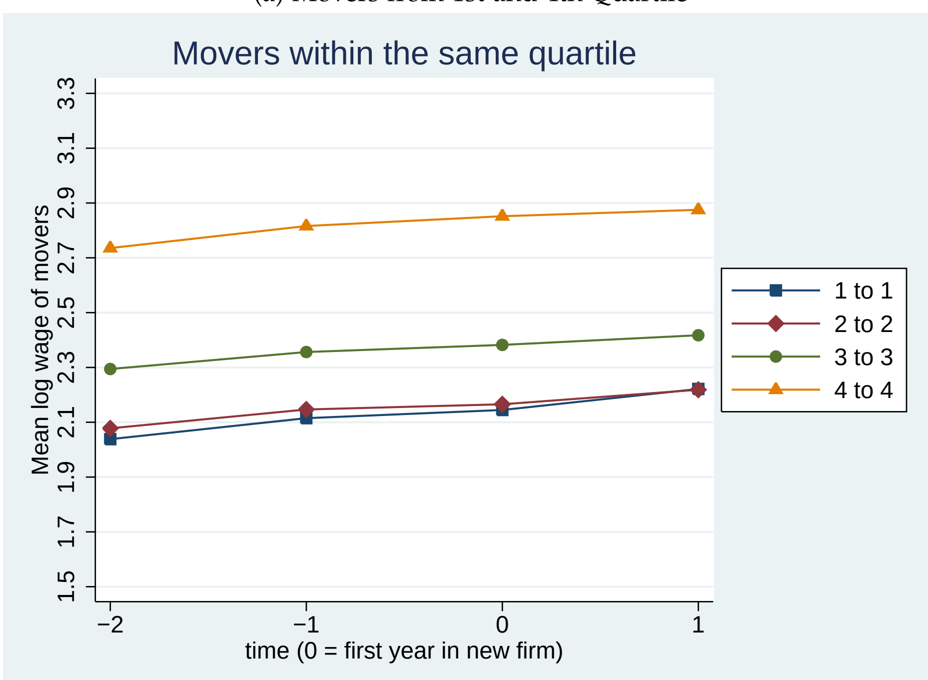
6 Tests

6.1 Testing Skill Stratification

I first test for skill stratification, that abler individuals occupy the upper layers of organizations. Testing whether this sorting pattern holds in my sample is important because it confirms that workers in a layer are abler than their subordinates. If this is not the case, then the model's prediction of positive assortative matching and the mechanism that determines how individuals sort



(a) Movers from 1st and 4th Quartile



(b) Movers within the Same Quartile

Figure 2: Wages of Movers

Table 3: Mean Log Wages by Transitions and Years

Origin-Destination Quartile	Number of Observations	Two Years Before	One Year Before	One Year After	Two Years After	Change from Two Years Before and After
1 to 1	25,775	2.03	2.11	2.14	2.22	0.19
1 to 2	15,654	1.89	1.99	2.19	2.24	0.35
1 to 3	11,759	1.79	1.91	2.28	2.31	0.52
1 to 4	3,410	1.85	1.97	2.64	2.58	0.73
2 to 1	16,427	2.09	2.20	2.03	2.12	0.03
2 to 2	51,732	2.07	2.14	2.16	2.21	0.14
2 to 3	44,670	2.07	2.16	2.28	2.32	0.25
2 to 4	10,489	2.08	2.19	2.55	2.55	0.47
3 to 1	11,468	2.25	2.37	1.98	2.11	−0.14
3 to 2	40,717	2.17	2.25	2.14	2.21	0.04
3 to 3	109,545	2.29	2.35	2.38	2.41	0.12
3 to 4	42,056	2.42	2.50	2.66	2.69	0.27
4 to 1	3,445	2.65	2.79	2.05	2.24	−0.41
4 to 2	8,550	2.42	2.56	2.16	2.25	−0.17
4 to 3	30,478	2.51	2.60	2.46	2.50	−0.01
4 to 4	72,529	2.73	2.81	2.85	2.87	0.14

Notes: Descriptive statistics of job transitions from the estimation of equation (9).

into organizations is clearly false.[16] For individual i employed in firm $j(i)$, I therefore estimate the following equation:

$$layer^L_{j(i)} = \mu_0 + \mu_1 ability_i + X_{j(i)} + u_{j(i)}, \tag{10}$$

where $layer^L_{j(i)}$ is the layer in firm j with a total number of L layers that worker i occupies, $ability_i$ is the estimated ability of worker i, and $X_{j(i)}$ are industry and location controls. Equation (10) is estimated across firms with the same total number of layers, separately. In equation (10) the interest is in how agents sort into layers, so the coefficient of interest is μ_1, and if abler individuals occupy the upper layers of organizations then μ_1 will be positive and significant.

Table 4 reports the regression results. Each entry in the table reports the estimated coefficient of μ_1. Because of the large number of indicator variables in the regressions, I estimate equation (10) using OLS. Rows 1 to 3 contain the results for firms with 2, 3 and 4 layers in their organization, respectively. The first column in Table 4 indicates the total number of layers in firms, the second column contains the sample size of the regressions, and the third to seventh columns report the estimated value of the coefficient, μ_1.

[16] The argument made in the model is that organizations exist to optimally utilize the knowledge of their workers. By shielding knowledgeable agents from easy tasks, a hierarchy allows abler individuals to focus on solving more complex or harder problems, while lower ability individuals focus on easier or commoner problems. This implies that within a firm, higher ability agents occupy the upper layers of organizations.

Table 4: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
TWO	4,432	0.053^{***}	0.044^{***}	0.047^{***}	0.043^{***}	0.041
		(0.013)	(0.012)	(0.013)	(0.013)	(0.031)
THREE	19,841	0.369^{***}	0.347^{***}	0.349^{***}	0.336^{***}	0.411^{***}
		(0.014)	(0.014)	(0.014)	(0.014)	(0.019)
FOUR	16,003	0.469^{***}	0.446^{***}	0.438^{***}	0.428^{***}	0.511^{***}
		(0.017)	(0.017)	(0.016)	(0.016)	(0.020)
Industry FE		No	Yes	No	Yes	No
Area FE		No	No	Yes	Yes	No
Firm FE		No	No	No	No	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *White-heteroskedastic standard errors (clustered at the firm level) in parentheses. OLS regressions for equation (10). Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient* μ_1 *from equation (10). The dependent variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the* 18 *manufacturing industries. Area fixed effects correspond to the* 341 *employment areas in mainland France.*

The regressions in column (3) report how agents sort into layers across firms, industries and locations. In all three regressions the coefficients are positive and significant at the one percent level. The column indicates that an individual with a one hundred percent increase in his ability and employed in a two-layer firm will on average reside 0.053 layers higher, while if he is employed in an organization with three-layers he will on average reside 0.369 layers higher and if he is employed in a four-layer firm he will on average reside 0.469 layers higher.

Even within industries and locations, Table 4 reports that higher ability agents occupy the upper layers of firms. To examine how agents sort into layers within industries, in column (4) I include industry fixed effects. The coefficients remain positive and significant at the one percent level. In column (5) I include location fixed effects so as to examine how agents sort into layers within locations. The coefficients in column (5) remain positive and significant. And finally, in column (6) I include both industry and location fixed effects. The findings indicate that an individual with a one hundred percent increase in his ability and employed in a two-layer firm will on average reside 0.043 layers higher, if he is employed in a three-layer firm will on average reside 0.336 layers higher, and if he is employed in an organization with four layers he will on average reside 0.428 layers higher. Therefore even within industries and within locations abler individuals occupy the upper layers of organizations.

The regressions in column (7) contain firm fixed effects and examine how agents sort into layers within firms. In two out of three regressions, the coefficient μ_1 is positive and significant at the one percent level. For four-layer firms, the result indicates that an individual with a one hundred percent increase in his ability will on average reside 0.511 layers higher. For organizations with three layers an agent with one hundred percent increase in his ability will on average reside

0.411 layers higher. Within three and four layer firms, therefore, abler individuals are employed in the upper layers of organizations.

In column (7), the value of μ_1 for two-layer firms is 0.041 and is not significant at the ten percent level. This would suggest that within two-layer firms, higher ability agents are not sorting into the upper layers of organizations. However, in light of the fact that the coefficient of μ_1 in column (7) is similar in magnitude to column (6), and that in the average two-layer firm there are 1.3 observations in the dataset, these findings are inconclusive.

To summarize from the evidence presented in Table 4 one can conclude that there is skill stratification, in the sense that abler agents occupy the higher layers of organizations. In other words, agents in higher layers of firms are of greater ability than their subordinates in the layers below. I now proceed to examine whether there is positive assortative matching between agents in the different layers of firms, and whether the mechanism that determines the sorting pattern is as suggested by the model.

6.2 Testing for Sorting

I now test for positive assortative matching between workers in the different layers of firms.[17] According to the knowledge-based management theory of firms layers are composed of workers who are of similar ability. A representative measure a layer's ability is the weighted average ability of workers occupying the layer.[18] More specifically, let N_j^l be the number of individuals in layer l at firm j, $H_{j(i)}^l$ be the number of hours performed by individual i in layer l in firm j and H_j^l be the total number of hours in layer l at firm j. The measure of the ability of layer l at firm j that I use, is the following:

$$ability_j^l = \sum_{i=1}^{N_j^l} \frac{H_{j(i)}^l}{H_j^l} ability_i. \tag{11}$$

where the summation is taken over all individuals in layer l at firm j.[19]

For firms with the same organizational structure, or the same total number of layers, positive assortative matching implies that the best workers team up with the best workers in other layers of firms. For example, when comparing two firms with the same total number of layers, say 2, the firm with the best production workers in layer 1 also employs the best managers in layers 2.

[17] Appendix B contains additional tests for positive assortative matching. Appendix B tests whether better workers sort into organizations with better co-workers. I adopt the approach of de Melo (2013), and investigate whether a worker's fixed effect is positively correlated with that of his co-workers. I conduct this analysis across several dimensions and find evidence in favor of positive assortative matching.

[18] I use a weighted average to account for the fact that some workers may be employed for the full year in a firm. In such a case, these workers cannot have the same impact on a firm, as workers who have been employed for the entire year.

[19] Note that this construction is only possible for layers where I there is at least one employee in the panel dataset of the DADS. As the panel is only a five percent sample of the French population, for many firms $ability_j^l$ remains undefined.

In other words, there should be a positive correlation between the ability of workers in the different layers of firms. To test for positive assortative matching, I therefore estimate the following equation:

$$ability_j^l = \alpha_0 + \alpha_1 ability_j^{l-g} + X_j\beta + u_j, \tag{12}$$

where $ability_j^l$ is the estimated weighted average ability of all workers in firm j who are in layer l, and $ability_j^{l-g}$ is the estimated weighted average ability of all workers in firm j who are in layer $l-g$, for $g = 1, .., l-1, l$. The firm controls X_j are firm observable variables such as firm age, an indicator for whether the firm already existed in the first year I have information, 1976, as well as indicator variables for industry and location. I include industry and location fixed effects because the assignment of workers to layers and firms may be different across industries and locations. I estimate equation (12), for firms with the same organizational structure and for the different values of l and g. In equation (12) the interest is how $ability_j^l$ varies with $ability_j^{l-g}$ across firms with the same total number of layers. If there is a positive assortative matching, then the coefficient α_1 will be positive and significant.

Table 5 reports the results. Each entry in the table illustrates the estimated coefficient of α_1 between two layers. The first column indicates the total number of layers in firms. The second column indicates the layer for which weighted average ability is the left-hand-side variable in equation (12), and the third column indicates the layer for which weighted average ability is the right-hand-side variable in equation (12). The fourth column reports the sample size of the regressions, while the fifth to tenth columns report estimated values of the coefficient. In Table 5 all the standard errors are White-heteroskedasticity consistent standard errors.

Table 5 reports that the sample size varies across regressions, even for firms with the same total number of layers. For example, in regressions with four-layer firms the sample can be as small as 15 observations or as large as $1,249$ observations. For firms with a given number of layers, the sample size increases when I estimate equation (12) with lower layers. The reasons are twofold. First given the nature of the data, I do not observe workers in all layers of firms. And second, in the data I am more likely to observe a worker in the lower layer of an organization.[20]

Almost all of the coefficients reported in Table 5 have a positive sign. There are three notable exceptions. In row six the coefficients between the ability of workers in layers four and layers two in four-layer firms are negative but not significant. In all three cases, however, the reported coefficients are imprecise. Because the sample size is small relative to the number of control variables there is not much independent variation in the data. This lack of independent variation in the data may also account for the reported negative coefficients.

Column (4) contains no controls and tests how agents sort together into firms across industries and locations. In these regressions, the majority of the estimated values of α_1 are positive and

[20]If firms are hierarchies then there are more workers in the lower layers of firms. Therefore in a 1/12 random sample of the population of workers one is more likely to observe individuals employed in the lower layers of firms.

Table 5: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	2	1	142	0.103	0.116	0.005	0.145	0.110	0.155
				(0.095)	(0.098)	(0.178)	(0.185)	(0.095)	(0.196)
THREE	3	2	457	0.240***	0.240***	0.223**	0.224**	0.241***	0.227**
				(0.068)	(0.072)	(0.089)	(0.094)	(0.068)	(0.094)
THREE	3	1	662	0.301***	0.314***	0.277***	0.295***	0.302***	0.290***
				(0.065)	(0.066)	(0.080)	(0.086)	(0.064)	(0.084)
THREE	2	1	1385	0.233***	0.217***	0.249***	0.235***	0.235***	0.236***
				(0.034)	(0.034)	(0.037)	(0.037)	(0.034)	(0.037)
FOUR	4	3	15	0.636*	0.692*			0.692	
				(0.332)	(0.268)			(0.406)	
FOUR	4	2	22	−0.180	−0.158			−0.117	
				(0.225)	(0.308)			(0.212)	
FOUR	4	1	37	0.324	0.354			0.408	
				(0.243)	(0.259)			(0.271)	
FOUR	3	2	452	0.195***	0.178***	0.215**	0.169	0.194***	0.167
				(0.063)	(0.067)	(0.090)	(0.103)	(0.063)	(0.105)
FOUR	3	1	687	0.077	0.072	0.115	0.117	0.075	0.118
				(0.056)	(0.058)	(0.070)	(0.073)	(0.056)	(0.073)
FOUR	2	1	1249	0.341***	0.332***	0.332***	0.322***	0.340***	0.320***
				(0.046)	(0.047)	(0.058)	(0.060)	(0.047)	(0.060)
Industry FE				No	Yes	No	Yes	No	Yes
Area FE				No	No	Yes	Yes	No	Yes
Firm Controls				No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (12). Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The dependent variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

significant, indicating that across industries and locations there is positive assortative matching. For example, in organizations with four layers, a one unit increase in the average ability of workers in layer one is associated with a 0.341 average increase in the average ability of workers in layer two.

Even within industries and locations, there is evidence of positive assortative matching. Column (6) contains industry fixed effects and examines how agents sort together within industries, while column (7) contains location fixed effects and examines how agents sort together within locations. In most cases the coefficients are positive and significant, suggesting that even within industries or within locations, the best workers team up with the best workers in other layers of firms. Column (8) reports regression results with both industry and location fixed effects. The findings indicate that there is positive assortative matching. For example, in organizations with four layers, a one hundred percent increase in the average ability of workers in layer one is associated with 0.322 average increase in the average ability of workers in layer two. In standardized units, this implies that a one standard deviation increase in the average ability of agents in layer one corresponds to a 0.301 standard deviation increase in the average ability of agents in layer two.

Column (9) in Table 5 reports results of regressions with firm observables as controls. Almost all of the coefficients are positive and half are significant. Finally column (10) reports results with the full set of controls, industry, location and firm observables. The coefficients remain positive and significant. For example, in organizations with three layers, a one hundred percent increase in the average ability of workers in layer one is associated with 0.236 average increase in the average ability of workers in layer two, and with a 0.290 average increase in the ability of workers in layer three. To obtain a sense of the strength of this relationship, a one standard deviation increase in the average ability of agents in layer two, corresponds to a 0.224 and 0.259 standard deviation increase in the average ability of agents in layers two and three, respectively.

One observation from Table 5 is that the magnitudes of α_1 are small. A small magnitude, however, is not necessarily inconsistent with the theory, since the assignment of agents into teams depends on the parameters of the model, and in particular on the distribution of abilities in the economy.[21] Furthermore, there are several rows in Table 5 where although the coefficients are positive, they are never significant. For example, in organizations with four layers, there appears to be no relationship between the average ability of agents in layers three and one. The same results hold for two-layer organizations. This suggests that there is no sorting between agents in these layers, however the fact that the coefficients are always positive indicates that there is a relationship in the data, albeit not strong.[22]

[21]The small magnitudes for α_1 are not problematic. If one were to assume a continuum of agents, as in Antras et al. (2006), then the mass of managers will be smaller than the mass of production workers. In this case, the matching function would have a slope that is less than 1.

[22]For two layer firms, this is consistent with the findings in table 4, which report that there is little evidence of sorting between agents and layers in firms.

To summarize the results, out of the possible 51 estimated coefficients, 28 are positive and significant at the five percent level, 2 are positive and significant at the ten percent level, 18 are positive but not significant, and 3 are negative and not significant. Therefore, apart for two-layer firms, these results provide evidence that there is positive assortative matching between workers in different layers of firms.

6.3 Testing the mechanism

Until now, I have found evidence that workers in a layer are of higher ability than their subordinates in the layers below, and that there is positive assortative matching between the workers in the different layers of firms. I now proceed to test the model's mechanism behind this sorting pattern. I proceed in two steps. First I test whether agents' span of control increases with their own ability. And second, I test whether agents' span of control increases with their subordinates' ability.

Let HR_j^l be the total number of hours worked by employees in layer l at firm j. I define the span of control of workers in layer l as:

$$span_j^l = \frac{HR_j^1}{HR_j^l}. \tag{13}$$

In other words, my measure of the span of control of workers in layer l is the ratio of the total number of hours in layer 1, to the number of hours in layer l.[23] The argument is that all workers in layer l supervise N_j^l individuals in layer 1, and these individuals spend a total of HR_j^1 hours at the firm. Dividing by the total number of hours worked by employees in layer l, HR_j^l, one obtains the number of hours a worker in layer l is expected to devote to supervising individuals in layer 1. This definition of span of control is closely related to the firms' maximization constraint in the model, discussed in the previous section, and it has the advantage of being invariant to the number of hours in the highest layer of the organization.[24]

In the model, the mechanism that is causing agents to sort together into firms is the following. Managers benefit from working with abler production workers because they take up less of their time, which allows managers to supervise more of them. Also, because abler managers can solve a greater number of problems, they increase the productivity of their workers. Therefore, in equilibrium, abler managers will be working with abler production workers and managers' span of control will be increasing with their ability. To test the mechanism of the model, therefore, I estimate the following equation:

[23] Since I cannot observe reporting relationships within organizations, this is the only measure available. I obtain HR_j^l from the exhaustive cross-section of the DADS.

[24] In the Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006) all workers have one unit of time available. Also, the number of workers in the top layer of a firm is normalized to one.

Table 6: Testing Mechanism - Managers' Ability

Total Number of Layers	layer l	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	2	601	-0.243	-0.161	-0.213	-0.109	-0.249	-0.115
			(0.158)	(0.137)	(0.232)	(0.197)	(0.160)	(0.195)
THREE	2	2412	-0.175^{**}	-0.084	-0.081	-0.037	-0.159^{**}	-0.027
			(0.036)	(0.072)	(0.083)	(0.076)	(0.078)	(0.076)
THREE	3	1183	-0.231^{**}	-0.116	-0.067	-0.040	-0.273^{**}	-0.067
			(0.111)	(0.097)	(0.121)	(0.110)	(0.108)	(0.109)
FOUR	2	1918	-0.120^{*}	-0.068	-0.062	-0.055	-0.117^{*}	-0.056
			(0.066)	(0.061)	(0.070)	(0.063)	(0.066)	(0.063)
FOUR	3	1042	-0.139	0.001	-0.216	-0.109	-0.116	-0.092
			(0.127)	(0.106)	(0.133)	(0.120)	(0.125)	(0.120)
FOUR	4	80	0.094	-0.001	0.311	-0.021	-0.074	0.486
			(0.295)	(0.339)	(0.457)	(0.971)	(0.296)	(1.274)
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (14). Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (14). The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, l. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

$$\ln span_j^l = \gamma_0 + \gamma_1 ability^l{}_j + X_j\beta + u_j, \tag{14}$$

where $ability^l{}_j$ is the estimated weighted average ability of all workers in firm j who are in layer l, and $span_j^l$ is defined above. The controls X_j are firm age, whether the firm was present in 1976, and indicator variables for industry and location. I estimate equation (14) for firms with the same number of layers and for different values of l, separately. In equation (14) I am interested in how the span of control of agents in layer l varies with their ability. If abler subordinates render their superiors more productive by allowing them to supervise more workers, then γ_1 should be positive and significant.

Table 6 reports regression results. The table has a similar structure to Table 5. Each entry in the table reports an estimated value of the coefficient γ_1. The first column reports the total number of layers in firms. The second column reports the layer for which weighted average ability is the right-hand-side variable in equation (14), the third column reports the sample size of the regressions, and the fourth to ninth columns report estimated values of γ_1. In Table 6 all the standard errors are White-heteroskedasticity consistent standard errors.

First note that the reported sample sizes in Table 6 vary across regressions. For example, in regressions with four-layer firms the sample size can be as small 80 observations or as large as 1,918 observations. The sample size also increases for regressions examining the mechanism in the lower layers of firms. As explained previously, this is not surprising given nature of the dataset and, because firms are hierarchies, in the dataset there are less employees in the higher layers of firms. Also the reported sample sizes are different from Table 5, since to estimate equation (14) only the ability of one employee in a layer has to be recorded in the dataset.

In Table 6 the vast majority of the estimates of γ_1 are negative, however only a handful of them are significant. The only consistent exception is the regressions of equation (14) reported in the last row, for agents in layer four in four-layer firms. In this case, the coefficients alternate sign, however they are never significant, and so they do not lead to a firm conclusion.

Column (1) contains no controls and examines how workers' ability varies with their span of control across industries and locations. The results indicate that workers' span of control is decreasing with their ability. For example in three-layer firms a one unit increase in the average ability of workers in layer two is on average associated with a 17.5 percent decrease in their span of control.

Within industries and locations, the relationship remains negative but not significant. Column (2) examines how workers' span of control varies with their ability within industries, while column (3) examines the relationship within locations. In both models the evidence suggests that there may be a negative relationship between agents' ability and their span of control, however this is not conclusive. Column (4) examines the relationship within industries and locations. The results remain the same: although the coefficients are negative, they are not significant.

Column (5) in Table 6 reports results with firm observables as controls. In column (5) two coefficients are negative and significant at the five percent level. In organizations with three layers, a one unit increase in the average ability of workers in layer three is associated with a 27.3 decrease in their span of control, while a one unit increase in the average ability of workers in layer two is associated with a 15.9 decrease in their span of control. Finally column (6) reports results with the full set of controls, industry, location and firm observables. The coefficients remain negative but not significant.

To summarize the results reported in Table 6, out of the 36 estimated coefficients, 4 are negative and significant at the five percent level, 2 are negative and significant at the ten percent level, 26 are negative but not significant, and 4 are positive but not significant. In light of these findings, the evidence suggests that although there is positive assortative matching between workers in the different layers of firms, the mechanism that is driving this sorting pattern is not present in the data. Indeed, the findings suggest that the opposite may be taking place. Abler managers form teams with abler production workers, however these workers take up more of the managers' time.

I now test whether agents' span of control increases with their subordinates' ability. I estimate the following equation:

Table 7: Testing Mechanism - Subordinates' Ability

Total Number of Layers	layer l-1	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	1	2863	-0.070	-0.098^{**}	-0.031	-0.068	-0.055	-0.061
			(0.053)	(0.051)	(0.057)	(0.054)	(0.053)	(0.055)
THREE	1	6430	-0.077^{**}	-0.035	-0.061	-0.032	-0.070^{*}	-0.028
			(0.036)	(0.035)	(0.038)	(0.036)	(0.036)	(0.036)
THREE	2	2413	-0.268^{**}	-0.169^{**}	-0.206^{**}	-0.158^{*}	-0.253^{***}	-0.150^{*}
			(0.082)	(0.076)	(0.085)	(0.081)	(0.081)	(0.081)
FOUR	1	4494	-0.143^{***}	-0.105^{**}	-0.115^{***}	-0.088^{**}	-0.140^{***}	-0.091^{**}
			(0.045)	(0.043)	(0.045)	(0.043)	(0.045)	(0.043)
FOUR	2	1918	-0.161^{*}	-0.089	-0.084	-0.067	-0.158^{**}	-0.067
			(0.082)	(0.074)	(0.086)	(0.080)	(0.082)	(0.080)
FOUR	3	1042	0.198	0.273	0.249	0.316^{**}	0.218^{*}	0.334^{**}
			(0.131)	(0.131)	(0.153)	(0.148)	(0.129)	(0.147)
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (15). Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (15). The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

$$\ln span_j^l = \gamma_0 + \gamma_1 ability_j^{l-1} + X_j\beta + u_j, \tag{15}$$

where $ability_j^{l-1}$ is the estimated weighted average ability of all workers in firm j who are in layer $l-1$, and $span_j^l$ is defined above. The controls X_j are the same as in equation (14). As before I estimate equation (15) for firms with the same number of layers and for different values of l, separately. In equation (15) the variable of interest is γ_1. If the mechanism behind the sorting of agents into organizations is correct, and abler subordinates allow their superiors to supervise more workers, then γ_1 should be positive and significant.

Table 7 presents the results from regression (15). The table has a similar structure to Table 6. As Table 7 reports in most of the regressions γ_1 is negative. The findings are more conclusive than the results reported in Table 6. For example, in three-layer firms, the relationship between the average ability of workers in layer 2 and the span of control of workers in layer 3 is negative and significant throughout the table. For workers in layer three in four-layer firms, however, there is evidence in favor of the mechanism suggested by the model: higher ability subordinates allow the superiors to increase their span of control. Because for the other regressions the findings are similar to Table 6, I only discuss the results reported in the last row of Table 7.

Column (4) examines how workers' ability varies with their superiors span of control within industries and locations. For four-layer firms the coefficient is positive and significant at the five percent level. A unit increase in the average ability of workers in layer three is associated with a 31.6 percent increase in the span of control of workers in layer four. Further column (5) reports results with firm observables as controls and column (6) reports the results for the full set of controls. In column (5) the coefficient of γ_1 is positive and significant at the ten percent level. In column (6) it is positive and significant at the five percent level and indicates that a one unit increase in the average ability of workers in layer three is associated with a 33.4 percent increase in the span of control of workers in layer four. In light of these results, for workers in layer four in organizations with four layers, there is some evidence to suggest that the mechanism is present in the data.

Therefore, the evidence is mixed. There is limited evidence in favor of the mechanism described by the model, and evidence to suggest that the opposite is taking place: abler managers form teams with abler production workers, however, these production workers take up more of the managers' time, which limits the amount of agents managers can supervise.

6.4 Additional Results

In this section, I examine the data more closely. I proceed in two steps. First, I examine whether the findings in the previous section are the same across firms with different sizes. Second, I examine whether the relationship holds for mono-establishment firms, that is firms that consist of only one plant. In this section I only report estimation results for equations (14) and (15) with the full set of controls.

It may be the case that large firms are using different production technologies, or that reporting relationships within large organizations are different. To account for this, within an industry I classify firms into quartiles by size, where I calculate firms' size from the number of workers in the organization. Tables 8 and 9 report the results.

From Tables 8 and 9 a weak pattern emerges. Even though the majority of the coefficients are negative, in higher quartiles, the magnitude of γ_1 decreases and even becomes positive. For example, in Table 9 for organizations with four layers, in the first quartile, a one unit increase in the average ability of workers in layer three is associated with a 175.0 percent average decrease in the span of control of workers in layer four, in the second quartile it is associated with a 151.0 percent average decrease, and in the third quartile the corresponding value is a 37.0 percent decrease. In the fourth quartile a one unit increase in the average ability of workers in layer three, however, is associated with a 100.1 percent average increase in the span of control of workers in layer four. This pattern does not hold uniformly throughout table. Although it less robust it is also present in Table 6. For example in organizations with three layers, in the first quartile, a unit a one unit increase in the average ability of workers in layer two is associated with a 56.0 percent average decrease in their span of control, while in the fourth quartile the corresponding value is a

Table 8: Testing Mechanism - Managers' Ability

Total Number of Layers	layer l	Sample Size	First Quartile	Second Quartile	Third Quartile	Fourth Quartile
TWO	2	601	-0.484^{**}	0.326	0.189	1.015^{*}
			(0.243)	(0.265)	(0.478)	(0.574)
THREE	2	2413	-0.560^{***}	-0.104	-0.103	0.296^{***}
			(0.157)	(0.124)	(0.110)	(0.090)
THREE	3	1183	-0.495^{**}	-0.221	0.042	0.143
			(0.207)	(0.190)	(0.169)	(0.125)
FOUR	2	1918	-0.238^{***}	-0.266^{*}	-0.106^{**}	0.071
			(0.289)	(0.142)	(0.079)	(0.072)
FOUR	3	1042	-0.095	-0.571^{**}	-0.209	0.015
			(0.437)	(0.272)	(0.151)	(0.125)
FOUR	4	82	-1.608	-0.619	1.086	1.405
			(0.913)	(1.711)	(0.976)	(1.210)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (14). Each row displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 interacted with the firm size quartile, for regressions with the full set of controls. The dependent variable is the span of control of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, l. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

29.6 percent increase. Therefore, there is some weak evidence that indicates that the mechanism described by the model holds in large organizations, while for small firms the evidence suggests that abler agents supervise less workers.

In addition, it may be the case that reporting relationships are only specific to a physical location. In particular if a firm is operating multiple plants, their organization may be different than what is suggested by the theory. To account for this, in Table 10 I report regression results for mono-establishment organizations.[25] The results are similar to those reported in the previous section. Apart for the span of control of agents in layer four, the estimated values of γ_1 are negative. Therefore, the conclusion remains the same. There is limited evidence in favor of the mechanism described by the model, and evidence to suggest that the agents' span of control is decreasing with ability.

7 Robustness Checks

7.1 Inconsistent Estimates

There are three threats to my estimates of workers' ability. All stem from my estimation of worker fixed effects. First, because the worker fixed effects are incidental parameters from regression (9), consistent estimates for them can only be obtained as the number of years an individual

[25] This removes 925 observations from the dataset.

Table 9: Testing Mechanism - Subordinates' Ability

Total Number of Layers	layer l-1	Sample Size	First Quartile	Second Quartile	Third Quartile	Fourth Quartile
TWO	1	2863	-0.341^{***}	-0.006	0.589^{***}	1.216^{***}
			(0.072)	(0.077)	(0.123)	(0.246)
THREE	1	6430	-0.374^{***}	0.131^{**}	0.042	-0.063
			(0.077)	(0.055)	(0.053)	(0.060)
THREE	2	2413	-0.867^{***}	-0.107	0.136	-0.122
			(0.158)	(0.126)	(0.124)	(0.105)
FOUR	1	4494	-0.189	-0.013	-0.057^{*}	-0.183^{***}
			(0.172)	(0.083)	(0.058)	(0.059)
FOUR	2	1918	-0.052	-0.267	0.106	-0.106
			(0.320)	(0.174)	(0.102)	(0.092)
FOUR	3	1042	-1.750^{***}	-1.510^{***}	-0.376^{**}	1.001^{***}
			(0.518)	(0.274)	(0.167)	(0.160)

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (15). Each row displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 interacted with the firm size quartile, for regressions with the full set of controls. The dependent variable is the span of control of workers in layer $l-1$. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

is observed in the panel grows large. Since for the years 1993 to 2004, the average worker is observed for 6 years, not all of the estimates of the time-invariant component, θ_i, identify a consistent measure of a worker's ability. Although the panel is short, to get a sense of how important is this issue, I conduct my analysis on workers that I observe for at least 10 periods. Tables 11 12, 13 and 14 present the regression results for this restricted sample.

Table 11 reports the tests for skill stratification. As in the previous table, higher ability agents occupy the upper layers of organizations. In addition in model 5, even within two-layer firms this relationship is now significant and indicates that an individual with a one hundred percent increase in his ability will on average reside 0.254 layer higher.

Table 12 reports regression results that test for positive assortative matching. First, note that in comparison to Table 12 not all of the reported estimates have a positive sign. In two-layer firms, within industries and locations, a one unit increase in the weighted average ability of agents in layer one is associated with a 2.018 decrease in the weighted average ability of agents in layer two. Although this would suggest that there is negative assortative matching between agents in layers one and two, the sample is small relative to the number of controls. In addition, in columns (5) and (9) the reported relationship is positive and significant at the five percent level. Therefore, these findings do not lead to firm conclusion. Second in organizations with four layers, in contrast to the results reported in Table 5 there is now weak evidence in favor of positive assortative matching between workers in layers three and one. In column (5), for example, a unit

Table 10: Testing Mechanism - Mono-Establishment Firms

Total Number of Layers	layer l	Sample Size	Model (1)	Total Number of Layers	layer l-1	Sample Size	Model (2)
TWO	2	591	-0.135	TWO	1	2723	-0.072
			(0.200)				(0.057)
THREE	2	2309	-0.046	THREE	1	6140	-0.033
			(0.077)				(0.037)
THREE	3	1120	-0.075	THREE	2	2309	-0.161^{**}
			(0.114)				(0.082)
FOUR	2	1783	-0.061	FOUR	1	4203	-0.100^{*}
			(0.066)				(0.045)
FOUR	3	952	-0.102	FOUR	2	1783	-0.058
			(0.126)				(0.084)
FOUR	4	82	0.486	FOUR	3	952	0.270^{*}
			(1.274)				(0.148)
Industry FE			Yes	Industry FE			Yes
Area FE			Yes	Area FE			Yes
Firm Controls			Yes	Firm Controls			Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equations (14) and (15) . Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (15) with the full set of controls. The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable in the first column is the estimated weighted average ability of workers in a lower layer, l. The right-hand side variable in the first column is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

increase in the weighted average ability of of workers in layer one corresponds to a 0.135 increase in the average ability of workers in layer three. Third, note that estimated magnitudes of several of the coefficients are larger than in Table 5. For example in column (8), in three-layer firms a unit increase in the average ability of workers in layer one is associated with a 0.370 increase in the average ability of workers in layer two. In standardized units, a one standard deviation in the average ability of agents in layer one corresponds to a 0.341 standard deviation increase in the ability of agents in layers two, which is greater than the 0.224 standard deviation increase reported from Table 5.

For the restricted sample, Table 13 tests whether abler workers have a greater span of control. In general the results are similar to those reported in Table 6. However, in three-layer organizations, there is now convincing evidence that higher ability agents in layer three supervise less workers. As reported in column (6) a one unit increase in the average ability of individuals in layer three is associated with a 40.7 decrease in their span of control.

Table 14 reports estimates of γ_1 from equation (15). Although the general conclusions are similar to Table 7, there are two differences. First, there is no longer any evidence to suggest that abler workers in layer three allow agents in layer four to increase their span of control. And

Table 11: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
TWO	1,746	0.118^{***}	0.093^{***}	0.110^{***}	0.092^{***}	0.254^{***}
		(0.024)	(0.023)	(0.027)	(0.026)	(0.078)
THREE	10,374	0.634^{***}	0.604^{***}	0.612^{***}	0.596^{***}	0.805^{***}
		(0.024)	(0.024)	(0.023)	(0.023)	(0.035)
FOUR	8,601	0.706^{***}	0.685^{***}	0.688^{***}	0.677^{***}	0.838^{***}
		(0.028)	(0.028)	(0.027)	(0.027)	(0.035)
Industry FE		No	Yes	No	Yes	No
Area FE		No	No	Yes	Yes	No
Firm FE		No	No	No	No	Yes

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. *White-heteroskedastic standard errors (clustered at the firm level) in parentheses. OLS regression results of equation (10) for workers with at least 10 years in the dataset. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient μ_1 from equation (10). The dependent variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the* 18 *manufacturing industries. Area fixed effects correspond to the* 341 *employment areas in mainland France.*

second, excluding the last row, there are now 9 estimated values of γ_1 that are positive but not significant.

Therefore, the conclusions remain the same. Although there is evidence that the best workers team up with the best workers in other layers of firms, there is limited evidence in favor of the mechanism described by the model, and evidence to suggest that the opposite is taking place. In particular that abler managers form teams with abler production workers, however, these production workers take up more of the managers' time, thereby limiting the amount of agents managers' can supervise.

7.2 Estimation Error

A positive correlation between the individual fixed effects may be the result of using standard econometric techniques. As discussed in Abowd et al. (2004) and Andrews et al. (2008), in equation (9) there is a negative correlation between the worker and firm-layer effects caused from standard estimation error. When the firm-layer fixed effects in equation (9) are on average underestimated, the individual fixed effects will be overestimated, and when the firm-layer fixed effects are on average overestimated, the individual fixed effects will be underestimated. Because in the panel workers transition between layers within firms, this implies that my regressions may suffer from non-classical measurement error, biasing results.

To resolve these issues I conduct my analysis only on workers who have moved to a new employer in the year 2008. For this sample of workers, any errors caused by miss-estimated firm-layer fixed effects will be uncorrelated with one another and uncorrelated with the workers' span

Table 12: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	2	1	38	0.354^{**}	0.326	0.127	-2.018^{**}	0.348^{**}	-1.670
				(0.136)	(0.208)	(1.653)	(0.365)	(0.139)	(1.628)
THREE	3	2	232	0.455^{***}	0.449^{***}	0.502^{***}	0.514^{***}	0.460^{***}	0.511^{***}
				(0.075)	(0.068)	(0.123)	(0.121)	(0.074)	(0.124)
THREE	3	1	380	0.419^{***}	0.427^{***}	0.281^{**}	0.274^{**}	0.411^{***}	0.262^{**}
				(0.086)	(0.089)	(0.121)	(0.130)	(0.087)	(0.120)
THREE	2	1	747	0.382^{***}	0.373^{***}	0.387^{***}	0.372^{***}	0.381^{***}	0.370^{***}
				(0.047)	(0.047)	(0.060)	(0.060)	(0.047)	(0.061)
FOUR	4	3	9	0.674				0.882	
				(0.469)				(0.585)	
FOUR	4	2	14	0.112	0.081			0.111	
				(0.183)	(0.969)			(0.193)	
FOUR	4	1	19	0.226	0.554			0.242	
				(0.346)	(0.334)			(0.409)	
FOUR	3	2	225	0.348^{***}	0.358^{***}	0.423^{***}	0.391^{**}	0.336^{***}	0.369^{***}
				(0.092)	(0.105)	(0.158)	(0.171)	(0.093)	(0.169)
FOUR	3	1	379	0.138^{***}	0.137^{***}	0.114	0.117	0.135^{**}	0.127
				(0.066)	(0.065)	(0.114)	(0.119)	(0.068)	(0.121)
FOUR	2	1	679	0.357^{***}	0.361^{***}	0.318^{**}	0.329^{**}	0.356^{***}	0.326^{**}
				(0.085)	(0.086)	(0.135)	(0.141)	(0.084)	(0.142)
Industry FE				No	Yes	No	Yes	No	Yes
Area FE				No	No	Yes	Yes	No	Yes
Firm Controls				No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regression results of equation (12) for workers with at least 10 years. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The dependent variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

of control. In these regressions, the coefficients will only suffer from attenuation bias, however the sign of the estimated coefficients will more properly reflect the relationships of interest. Because the sample sizes are small in these regressions, unlike in the previous sections Tables (15), (16), (17) and (18) do not report the results from all models.

For the year 2008, Table 15 reports how workers sort into layers and organizations. Because in the sample there are not many workers employed in the same firm, regression results that examine how workers sort into layers within organizations are omitted. Further, in all the regressions reported in Table 15 the coefficient μ_1 is positive. For firms with three and four layers, the reported coefficients are also significant at the one percent level. Therefore in three and four layer firms, higher ability agents occupy the upper the layers of organizations. In two-layer organizations μ_1 is no longer significant, however, it has a similar magnitude as the results reported

Table 13: Testing Mechanism - Managers' Ability

Total Number of Layers	layer l	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	2	263	-0.137	-0.024	-0.020	0.257	-0.206	0.162
			(0.242)	(0.189)	(0.398)	(0.403)	(0.252)	(0.401)
THREE	2	1471	-0.128^{*}	-0.089	0.000	0.006	-0.115	0.0140
			(0.090)	(0.083)	(0.099)	(0.090)	(0.090)	(0.090)
THREE	3	775	-0.359^{**}	-0.321^{***}	-0.305^{*}	-0.385^{**}	-0.385^{***}	-0.407^{***}
			(0.140)	(0.125)	(0.155)	(0.152)	(0.139)	(0.151)
FOUR	2	1198	-0.031	-0.016	-0.018	-0.019	-0.035	-0.025
			(0.080)	(0.073)	(0.097)	(0.086)	(0.080)	(0.086)
FOUR	3	700	-0.165	-0.037	-0.185	-0.096	-0.143	-0.085
			(0.144)	(0.118)	(0.168)	(0.149)	(0.143)	(0.151)
FOUR	4	54	0.089	-0.159	0.905	-0.518	-0.119	
			(0.381)	(0.563)	(1.010)	(3.235)	(0.376)	
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regression results of equation (14) for workers observed for at least 10 years. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (14). The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, l. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the coefficient of the estimated firm fixed effects from regression (9), the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

in Table 4. This suggests that there is no sorting between agents into layers, however the fact that the coefficients are always positive indicates that there is a relationship in the data, albeit not strong.

Table 16 reports the results of tests for positive assortative matching. Because there are not many workers employed in the same firm, Table 16 only reports results for three-layer and four-layer organizations. First note that in firms with three layers there is some evidence of negative assortative matching. The majority of the reported coefficients have a negative sign, and one coefficient is significant at the five percent level. Across industries and locations a one unit increase in the average ability of agents in layer one corresponds to a 0.130 decrease in the average ability of workers in layer two. As additional controls are added, however, this relationship remains negative but is no longer significant. Given that the size of the samples are small, however, these findings to not lead to a firm conclusion.

Second, in organizations with four layers, there is evidence to suggest that better workers are employed with better workers in the other layers of firms. The majority of the reported coefficients have a positive sign, and two are significant at the five percent level. For example, across industries and locations, a one unit increase in the average ability of agents in layer one corresponds to a 0.193 average increase in the average ability of workers in layer two. These

Table 14: Testing Mechanism - Subordinates' Ability

Total Number of Layers	layer l-1	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	1	1225	-0.022	-0.046	0.049	0.026	-0.018	0.033
			(0.080)	(0.076)	(0.102)	(0.093)	(0.080)	(0.093)
THREE	1	3919	-0.029	0.025	0.003	0.041	-0.030	0.040
			(0.047)	(0.046)	(0.048)	(0.047)	(0.047)	(0.047)
THREE	2	1471	-0.207^{**}	-0.158	-0.120	-0.100	-0.199^{**}	-0.096
			(0.100)	(0.096)	(0.105)	(0.103)	(0.101)	(0.103)
FOUR	1	3014	-0.121^{**}	-0.075	-0.113^{**}	-0.082	-0.122^{**}	-0.084
			(0.052)	(0.050)	(0.055)	(0.053)	(0.052)	(0.053)
FOUR	2	1198	-0.050	-0.012	0.043	0.040	-0.050	0.041
			(0.100)	(0.093)	(0.117)	(0.110)	(0.099)	(0.110)
FOUR	3	700	-0.061	-0.023	-0.008	0.025	-0.034	0.031
			(0.151)	(0.136)	(0.168)	(0.161)	(0.150)	(0.163)
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regression results of equation (15) for workers observed for at least 10 years. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (15). The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the coefficient of the estimated firm fixed effects from regression (9), the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

findings are consistent with previous results.

Table 17 and 18 report results that test for the mechanism. As in the previous section, there is no longer any evidence to suggest that abler workers in layer three allow agents in layer four to increase their span of control. And second, although the majority of the coefficients are negative few are significant at the five percent level.

Therefore, the conclusions remain the same. Even though there is evidence that the best workers team up with the best workers in other layers of firms, there is limited evidence in favor of the mechanism suggested by the model, and evidence to suggest that the opposite is taking place.

7.3 Biased Estimates

Third, the estimated worker fixed effects may be biased. If workers in a given layer render their subordinates more productive, and if their subordinates make them more productive, then this should be reflected in wages. If this is the case, then the worker fixed effect in equation (9) is not only identifying the productivity of a worker, but also the impact his co-workers have on his productivity. In other words, for worker i employed at time t in firm $J(i,t)$, the estimated individual fixed effect from equation (9), $\hat{\theta}_i$, is equal to: $\theta_i + \phi_J(i,t)$, where $\phi_J(i,t)$ is the impact

Table 15: Regression Results for Skill Stratification

Total Number of Layers	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)
TWO	444	0.028	0.031	0.044	0.042
		(0.024)	(0.024)	(0.037)	(0.039)
THREE	2,537	0.229***	0.200***	0.222***	0.196***
		(0.030)	(0.028)	(0.028)	(0.027)
FOUR	2,112	0.258***	0.227***	0.219***	0.195***
		(0.042)	(0.041)	(0.042)	(0.042)
Industry FE		No	Yes	No	Yes
Area FE		No	No	Yes	Yes
Firm FE		No	No	No	No

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors (clustered at the firm level) in parentheses. OLS regression results of equation (10) for workers with at least 10 years in the dataset. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient μ_1 from equation (10). The dependent variable is the layer that a worker occupies. The right-hand side variable is the ability of the worker. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France.*

Table 16: Regression Results for Sorting Tests

Total Number of Layers	layer l	layer l-g	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
THREE	3	2	18	−0.114	−0.407*			0.209
				(0.173)	(0.170)			(0.133)
THREE	3	1	47	−0.001	−0.018	0.693	0.921	−0.091
				(0.132)	(0.168)	(0.789)	(2.020)	(0.131)
THREE	2	1	95	−0.130**	−0.107	−0.159	−0.059	−0.122
				(0.077)	(0.090)	(0.252)	(0.461)	(0.112)
FOUR	3	2	43	0.068	0.015	1.242***		0.035
				(0.095)	(0.167)	(0.276)		(0.099)
FOUR	3	1	70	0.117	0.012	−0.006	−0.181	0.115
				(0.101)	(0.106)	(0.282)	(0.392)	(0.099)
FOUR	2	1	131	0.193**	0.177*	0.297	0.316	0.170*
				(0.085)	(0.101)	(0.369)	(0.459)	(0.091)
Industry FE				No	Yes	No	Yes	No
Area FE				No	No	Yes	Yes	No
Firm Controls				No	No	No	No	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regression results of equation (12) for workers with at least 10 years. Each cell displays the estimate of a separate regression for firms with the same total number of layers and across two layers of firms. The table only reports the value of the coefficient α_1 from equation (12). The dependent variable is the estimated weighted average ability of workers in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l - g$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976. Entries are omitted because the sample size was too small.*

Table 17: Testing Mechanism - Managers' Ability

Total Number of Layers	layer l	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	2	25	-0.344	-0.192	2.758		-0.352	
			(0.719)	(0.766)	(6.934)		(0.996)	
THREE	2	303	-0.381	-0.264	-0.056	-0.088	-0.410^{*}	-0.146
			(0.244)	(0.178)	(0.355)	(0.314)	(0.246)	(0.317)
THREE	3	123	-0.359^{**}	0.106	0.609	1.154	0.162	1.157
			(0.140)	(0.419)	(1.190)	(1.608)	(0.372)	(1.250)
FOUR	2	322	-0.259	-0.121	-0.296	-0.125	-0.246	-0.092
			(0.198)	(0.184)	(0.334)	(0.364)	(0.197)	(0.364)
FOUR	3	158	-0.498	-0.305	-0.884	-0.234	-0.517	-0.536
			(0.404)	(0.311)	(1.002)	(1.071)	(0.409)	(1.124)
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regression results of equation (14) for workers in new firms. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (14). The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, l. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the coefficient of the estimated firm fixed effects from regression (9), the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

worker i's co-workers have on his productivity. If there is positive sorting, then it is safe to assume that $cov(\theta_i, \phi_J(i,t))$ is positive. For another worker h employed at time t in the same firm as i, then it should be the case that $cov(\theta_h, \phi_J(i,t))$ is positive as well. Hence, this would imply that there is nonclassical measurement error in my subsequent regressions.

Moreover, if the model is an accurate description of the real world, since there is positive assortative matching between workers in the different layers of firms, for a given layer, $\phi_J(i,t)$ should be increasing with θ_i. Therefore, according to the model the bias should be increasing in the ability of an individual. For regressions (12), (14) and (15), this would further bias the coefficient of interests, α_1 and γ_1 in favor of finding a positive result. Therefore, one interpretation of my results is that they present an upper bound on the relationships of interest.

8 Conclusion

Understanding how workers sort together with other workers into layers and firms is crucial for understanding the organization of firms. Without knowledge of the precise nature of the interactions between workers in the different layers of firms, it is difficult to comprehend how firms organize production. Additionally, pinpointing the mechanism that is causing this sorting pattern is essential for determining why workers sort together in firms. Finally, better knowledge

Table 18: Testing Mechanism - Subordinates' Ability

Total Number of Layers	layer l-1	Sample Size	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
TWO	1	367	0.028	0.054	−0.358	−0.310	0.033	−0.331
			(0.150)	(0.157)	(0.245)	(0.278)	(0.153)	(0.284)
THREE	1	1462	−0.093	−0.133*	−0.124	−0.158*	−0.102	−0.170*
			(0.081)	(0.078)	(0.097)	(0.094)	(0.081)	(0.095)
THREE	2	303	−0.554**	−0.478**	−0.162	−0.234	−0.545**	−0.221
			(0.269)	(0.218)	(0.371)	(0.383)	(0.272)	(0.386)
FOUR	1	1010	−0.060	0.002	−0.167	−0.059	−0.060	−0.056
			(0.099)	(0.095)	(0.113)	(0.108)	(0.099)	(0.108)
FOUR	2	322	−0.440**	−0.289	−0.601	−.447	−0.420*	−0.391
			(0.223)	(0.207)	(0.379)	(0.389)	(0.219)	(0.399)
FOUR	3	158	−0.474	−0.004	−0.533	−0.381	−0.479	−0.673
			(0.419)	(0.406)	(1.079)	(1.231)	(0.428)	(1.306)
Industry FE			No	Yes	No	Yes	No	Yes
Area FE			No	No	Yes	Yes	No	Yes
Firm Controls			No	No	No	No	Yes	Yes

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regression results of equation (15) for workers in new firms. Each cell displays the estimate of a separate regression for firms with the same total number of layers. The table only reports the value of the coefficient γ_1 from equation (15). The dependent variable is the estimated span of control of agents in layer l. The right-hand side variable is the estimated weighted average ability of workers in a lower layer, $l-1$. Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the coefficient of the estimated firm fixed effects from regression (9), the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

of how workers sort into layers and firms is important for understanding earnings' inequality, and how firms respond to changes in their market environment.

This paper directly examined how workers sort together in firms. My empirical strategy relies on the idea that firms can be thought of as hierarchical teams, composed of layers that perform different tasks. Using French administrative data, I conclude that, within firms, higher ability workers are employed in the higher layers of firms, and across firms, there is positive assortative matching between workers in the different layers of firms. Third, I find only weak evidence for the mechanism, as suggested by Garicano and Rossi-Hansberg (2006), that is causing this sorting pattern: higher ability workers allow their managers to increase their span of control and employ larger teams. Finally, I also find evidence that higher ability managers supervise less workers.

An important question remains to be answered. The findings presented in this study indicate that although there is some evidence that higher ability workers allow their managers to increase their span of control and employ larger teams, there is also evidence that the opposite is taking place. An important question therefore remains to be answered: If better workers sort into firms with other better workers, what is causing this sorting pattern?

References

Abowd, John M., Creecy Robert H., and Francis Kramarz. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." *Working Paper* .

Abowd, John M., Francis Kramarz, Paul Lengerman, and Sebastien Perez-Duarte. "Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States." *Working Paper* .

Abowd, John M., Francis Kramarz, and David N. Margolis. "High Wage Workers and High Wage Firms." *Econometrica* 67, 3: (1999) 251–333.

Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Upward Richard. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Statistical Artefact?" *Journal of the Royal Statistical Society* 171, 3: (2008) 673–697.

Antras, Paul, Luis Garicano, and Esteban Rossi-Hansberg. "Offshoring in a Knowledge Economy." *Quarterly Journal of Economics* 121, 1: (2006) 31–77.

———. *Organizing Offshoring: Middle Managers and Communication Costs*. Cambridge, US: Harvard University Press, 2008, e. helpman, d. marin, t. vierdier eds. edition.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. "Anatomy of French Production Hierarchies." *Journal of Political Economy* Forthcoming.

Card, David, Jorg Heining, and Patrick Kline. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128, 3: (2013) 967–1015.

Coase, Ronald. "The Nature of the Firm." *Economica* 16, 4: (1937) 386–405.

Dasgupta, Kunal. "Learning and Knowledge Diffusion in a Global Economy." *Journal of International Economics* 87, 2: (2012) 323–336.

Garicano, Luis. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economics* 108, 5: (2000) 874–904.

Garicano, Luis, and Thomas N. Hubbard. "Hierarchical Sorting and Learning Costs: Theory and Evidence from the Law." *Journal of Economic Behavior and Organization* 58, 2: (2005) 349–369.

———. "The Returns to Knowledge Hierarchies." *Working Paper* .

Garicano, Luis, and Esteban Rossi-Hansberg. "Inequality and the Organization of Knowledge." *American Economic Review Papers and Proceedings* 50, 1: (2004) 197–202.

———. "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics* 121, 4: (2006) 1383–1435.

Guimaraes, Paulo, and Pedro Portugal. "A Simple Feasible Procedure to Fit Models with High-Dimensional Fixed Effects." *The Stata Journal* 10, 4: (2010) 628–649.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti. "Skill Dispersion and Firm Productivity: An Analysis with Matched-Employer Employee Data." *Journal of Labor Economics* 26, 2: (2008) 247–285.

Lucas, Robert E. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9, 2: (1977) 508–523.

Martins, Pedro S. "Dispersion in Wage Premiums and Firm Performance." *Economic Letters* 101, 2: (2008) 63–65.

de Melo, Rafael Lopes. "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." *Working Paper* .

Rosen, Sherwin. "Authority, Control, and the Distribution of Earnings." *The Bell Journal of Economics* 13, 2: (1982) 311–323.

Appendix A: Data Appendix

The Panel Dataset of the DADS

To estimate worker and firm fixed effects, I use the years 1993 to 2004 from the panel dataset of the DADS. Initially, the dataset contains 24,882,933 total observations, 5,469,362 workers and 1,614,337 firms. I remove from the dataset any workers or firms that cannot be properly identified or that have missing values. For reasons of computational tractability, I restrict the sample to all workers who are born in an even numbered year, are between the ages 18 and 65 and work in continental France. I also eliminate from the sample all individuals I observe only once in the panel and who are not full-time workers. In a given year, an individual may hold multiple jobs. In case of multiple jobs, for a given year I keep the worker's employment with the highest salary. Finally, I also eliminate all firms in the agricultural and fishing industries and all industries in which there are some coding problems present. In all, for the years 1993 to 2004, there are X observations, 1,X workers and X firms. From this sample of workers and firms, to obtain an exact estimate of worker and firm-layer fixed effects I find the largest connected group. The largest connected group contains 4,999,728 observations, 753,092 workers, 399,676 firms and 569,198 firm-layer pairs.

For the years 1993 to 2004 Table 1 presents distribution of the number of years workers are observed in the panel dataset of the DADS.

Table 1: Distribution of the number of years workers are observed in the panel

variable	mean	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Years	6.63	2	2	2	2	4	6	10	11	11	12	12

For the years 1993 to 2004 Table 2 presents distribution of the number of years firms are observed in the panel dataset of the DADS as well as the distribution of the number of workers that are observed in a firm in a given year.

Table 2: Distribution of the number of years layers within firms are observed in the panel and the distribution of workers per layer in a firm

variable	mean	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Years	3.28	1	1	1	1	1	2	4	8	10	11	12
Workers	2.67	1	1	1	1	1	1	2	4	7	22	9,674

The Exhaustive Cross-Section of the DADS

The exhaustive cross-section of the DADS contains information on all workers who earn a positive wage in a french establishment. For a given year, the observations are at the worker-establishment level. Within a firm, a worker can have multiple jobs if he is employed in two different establishments. To clean the data, I first remove any observations that do not have a positive amount of hours, days, occupation or wage reported. I also remove any observations in which the firm and individual information is missing.

For every firm we have information on its industry of operation. I also have this information for every establishment that comprises a firm. I classify firms into industries using the industry of operation of the firm. Further I remove any industries in which there are classification errors. Since I focus on manufacturing firms only, I remove any industries that are not in manufacturing.

Table 19: Descriptive Statistics from the Exhaustive Cross-Section Dataset

Exhaustive Cross-Section:

Total		Average		Median		Standard Deviation	
Number of Layers	Number of Firms	Number of Employees	Number of Hours	Number of Employees	Number of Hours	Number of Employees	Number of Hours
1	160,904	3.45	3,382.28	2	2,028	4.73	4,980.97
2	74,676	8.36	9,437.35	5	6,084	10.80	11,971.35
3	52,949	23.90	31,701.27	11	13,342	53.59	80,918.29
4	14,434	59.92	82,872.88	33	47,106	96.12	119,415.10

Notes: Descriptive statistics from the exhaustive cross-section of the DADS for the year 2004.

I construct my measure of layers using the first-digit of the CS classification codes. I consider a firm as having a layer if an employee is present in that layer. I classify firms by the number of management layers present in their organization. In other words, a firm where layer 1, layer 3 and layer 3 are present is defined as a three-layer firm. To identify the layers in a firm, I use the first digit of the CS occupational codes which range from 2 to 6. Therefore in total I can identify up to four layers. Layer 1 corresponds to qualified and non-qualified administrative workers and blue-collar workers. It contains all workers with CS occupational codes 5 and 6, respectively. I group CS occupational codes 5 and 6 together because their distribution of ability are similar. Layer 2 is composed of supervisors and individuals with higher level of responsibility than ordinary workers, and contains all workers with an occupational code 4. Layer 3 is composed of senior directors and top management staff and contains all workers with an occupational code 3. Layer 4 corresponds to owners who receive a wage and CEOs. It contains all workers with a CS occupational code 2.

For every layer, I calculate the total number of employees in a layer. If a worker appears in two different establishments but in a different layer, I treat him as two separate observations. I also calculate the total number of hours per layer. If a worker appears in two different establishments, I keep both observations to calculate the total number of hours per layer. Further, I remove any firms that record a positive number of hours for workers in occupations, with codes different from 2 to 6.

Merged Datasets

For the year 2004 I merge the information from the panel and exhaustive cross-section datasets together. Unlike the exhaustive cross-section, since the panel data is based on a 5 percent sample of the French population, it contains information on a sample of all firms operating in mainland France. Approximately 1 percent of firms in the panel dataset are not matched. I keep only firms that operate in the manufacturing sector, and remove any firms that operate in more than one industry and location. In total the matched dataset contains 23,916 firms that operate in 17 industries, of which 2,160 are one-layer firms, 3,322 are two-layer firms, 7,860 are three-layer firms and 5,450 are four-layer firms.

Tables 19 and 20 contains descriptive statistics of firms in the exhaustive cross-section dataset and the matched dataset for the year 2004, respectively. As is evident from the average and the median number of workers and the average and the median number of hours worked in a firm, the matched dataset contains larger firms.[26] For both measures of firm size, number

[26] As explained further below, one reason for this result is that I obtain my measures of workers' ability from the largest connected set of workers and firms.

Table 20: Descriptive Statistics from the Matched Dataset

Matched Sample:

Total Number of Layers	Number of Firms	Average Number of Employees	Average Number of Hours	Median Number of Employees	Median Number of Hours	Standard Deviation Number of Employees	Standard Deviation Number of Hours
1	2,160	11.00	11,844.12	8	8,233.50	23.14	25,319.03
2	3,322	19.88	25,989.84	14	18,173.50	22.11	28,176.84
3	7,860	67.14	101,683.20	37	53,741	115.92	183,040.00
4	5,450	83.61	128,090.50	53	81,350	99.54	155,456.00

Notes: Descriptive statistics from the the matched sample dataset for the year 2004.

of employees and the number of hours, the standard deviation in the matched sample is also greater than in the population. The sample is therefore biased towards larger firms, and is not representative of the entire population of firms. To the extent that the theory applies to all firms the unrepresentativeness of the sample is not of concern.

Appendix B: Test for Positive Assortative Matching

In this section I test whether there is positive assortative matching between workers and their co-workers in a firm. I adopt the approach proposed by Lopes de Melo (2013). Building on the frictional matching model of Shimer and Smith (2003), Lopes de Melo (2013) shows that even though wages are not monotone with respect to firm productivity, wages will be monotone with respect to workers' skills. Therefore, if better workers are sorting together into firms, the correlation between a worker's fixed effect and the average fixed effect of his co-workers should be positive.

To test this prediction, I conduct three exercises. First I correlate the worker fixed effect with the average fixed effect of his co-workers for all firms in the economy, and for all firms with the same number of layers, separately. Second, I perform the same exercise, but within the layers of firms. And finally, I add additional structure and estimate the following equation:

$$\theta_i = \alpha_0 + \alpha_1 \overline{\theta}_{-i} + X_j \beta + u_j, \tag{16}$$

where $\overline{\theta}_{-i}$ is the average ability of workers i's co-workers. I include as controls indicators for industry, the age of the firm, whether the firm was present in 1976, and controls for industry and location. If workers of similar ability are employed in the same firms, then α_1 should be positive and significant. I conduct these tests only for the year 2004.

Tables 21 and 22 presents the correlation results. The first column of table 21 presents the correlation for all workers in the economy and for firms with the same number of layers, while the second column contains the size of the sample used to estimate the correlation. The third column presents the correlation between the worker fixed effects and the firm fixed effects, and the fourth columns presents the number of observations used to estimate the correlation. In column one, the correlations are all positive. For all firms in the economy, the correlation between the worker fixed effect and the average ability of his co-workers is 0.352. Table 22 presents the same correlation but across the layers within a firm, l. Again all correlations are positive.

Tables 23 and 24 report regression results. In the tables all standard errors are robust. In tables 23 and 24 not all coefficients are positive and significant. In table 23, apart for firms that organize with one layer, the coefficients are positive and significant at the one percent level. In table 24, the results are mixed. For the higher layers of firms, the coefficient of α_1 is negative and significant. For example, in a four-layer firm, for a worker in layer two, a one unit in the average ability of his co-workers is associated with -0.205 average decrease in his ability. In all, we can conclude that there is positive assortative matching between workers in the same organization, however within the same layer of an organization the evidence is mixed.

Table 21: Regression Results Sorting from Regression 16

Total Number of Layers	corr(θ_i; $\overline{\theta}_{-i}$)	N	corr(θ; ψ)	N
All	0.352	31,941	−0.277	31,941
ONE	0.374	481	−0.559	2,432
TWO	0.382	1,871	−0.521	4,432
THREE	0.362	15,773	−0.310	19,841
FOUR	0.335	13,816	−0.290	16,003

Notes: Correlations between the ability of workers and their co-workers.

Table 22: Regression Results Sorting from Regression 16

Total Number of Layers	layer	corr(θ_i; $\overline{\theta}_{-i}$)	N	corr(θ; ψ)	N
ONE	1	0.374	481	−0.503	481
TWO	1	0.361	1,819	−0.423	1,819
TWO	2	0.639	52	−0.805	52
THREE	1	0.355	15,026	−0.287	15,026
THREE	2	0.310	648	−0.293	648
THREE	3	0.539	99	−0.585	99
FOUR	1	0.325	13,127	−0.282	13,127
FOUR	2	0.300	550	−0.300	550
FOUR	3	0.429	139	−0.533	139
FOUR	4				

Notes: Correlations between the ability of workers and their co-workers within the layers of firms.

Table 23: Regression Results Sorting from Regression 16

Total Number of Layers	Model 1 α_1	Model 1 std. err.	Model 2 α_1	Model 2 std. err.	N
All	0.413***	0.009	0.412***	0.009	31,933
ONE	−0.149	0.106	−0.174	0.107	481
TWO	0.182***	0.031	0.171***	0.031	1,871
THREE	0.405***	0.013	0.402***	0.013	15,765
FOUR	0.374***	0.015	0.370***	0.015	13,816
Industry FE		Yes		Yes	
Area FE		Yes		Yes	
Firm Controls		No		Yes	

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (16). The table only reports the value of the coefficient α_1 from equation (16). Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*

Table 24: Regression Results Sorting from Regression 16

Total Number of Layers	layer	Model 1 α_1	Model 1 std. err.	Model 2 α_1	Model 2 std. err.	N
ONE	1	−0.149	0.106	−0.174	0.107	481
TWO	1	0.135***	0.032	0.119***	0.032	1,819
TWO	2	0.093	0.415	−1.014***	0.013	52
THREE	1	0.399***	0.014	0.399***	0.014	15,018
THREE	2	−0.057	0.033	−0.0639	0.068	648
THREE	3	−0.090	0.185	−0.144	0.177	99
FOUR	1	0.372***	0.016	0.368***	0.016	13,127
FOUR	2	−0.205***	0.066	−0.240***	0.065	550
FOUR	3	−0.591***	0.131	−0.595***	0.130	139
FOUR						
Industry FE		Yes		Yes		
Area FE		Yes		Yes		
Firm Controls		No		Yes		

*Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. White-heteroskedastic standard errors in parentheses. OLS regressions for equation (12). Industry fixed effects correspond to the 18 manufacturing industries. Area fixed effects correspond to the 341 employment areas in mainland France. Firm controls include the age of the firm, and whether the firm was present in the first year of the panel, 1976.*