Technology, Information and the Decentralization of the Firm*

Daron Acemoglu[†], Philippe Aghion[‡], Claire Lelarge[§],

John Van Reenen[¶], and Fabrizio Zilibotti[‖]

August 21, 2006

Abstract

This paper develops a framework to analyze the relationship between the diffusion of new technologies and the decentralization of firms. Centralized control relies on the information of the principal, which we equate with publicly available information. Decentralized control, on the other hand, delegates authority to a manager with superior information. However, the manager can use her informational advantage to make choices that are not in the best interest of the principal. As the available public information about the specific technology increases, the trade-off shifts in favor of centralization. We show that firms closer to the technological frontier, firms in more heterogeneous environments and younger firms are more likely to choose decentralization. Using three datasets of French and British firms in the 1990s, we report robust correlations consistent with these predictions.

JEL No. O31, O32, O33, F23

Keywords: decentralization, heterogeneity, learning, the theory of the firm.

*We thank Erik Brynjolfsson, participants at seminars in CREST, CEPR (Paris), CIAR, Harvard-MIT, NBER, Toronto, three anonymous referees and Lawrence Katz for helpful comments and suggestions. Van Reenen would like to thank the ESRC for help with financing this research at the Centre for Economic Performance.

[†]Massachusetts Institute of Technology

[‡]Harvard University

[§]INSEE-CREST, Paris

[¶]London School of Economics and Centre for Economic Performance

[‖]University of Zürich

1 Introduction

Recent years have witnessed increasing interest in the determinants of firms' organizational choices. This interest is partly motivated by the belief that new technologies are inducing firms to become less hierarchical and more decentralized. Despite this interest, there is limited work on the determinants of the decentralization decisions of firms. This paper undertakes a theoretical and empirical investigation of how the allocation of authority within firms changes as the information structure in an industry evolves.

We develop a simple model where firms make choices on how to implement new technologies. Different organizational forms are distinguished by the amount of information they use in these decisions. As in Aghion and Tirole (1997), centralized control relies more on the information of the principal,[1] which we equate with publicly available information about past implementations of similar technologies. Decentralized control delegates authority to a manager, who potentially possesses more information than available in the public history. Nevertheless, because the interests of the principal and the manager are not perfectly aligned, the manager can use his informational advantage to make choices that are not in the best interest of the principal. This trade-off between the superior knowledge of the manager and the agency costs of managerial delegation determines the optimal degree of decentralization. The main focus of our analysis is on how learning from the experiences of other firms changes this trade-off. Typically, the more a principal learns from other firms regarding the implementation of new technologies, the less she needs to delegate control to the manager.

Using this basic framework we derive three sets of empirical predictions:

1. Firms closer to the technological frontier are more likely to choose decentralization, because they are dealing with new technologies about which there is only limited information in the public history.

2. Firms in more heterogeneous environments are more likely to be decentralized because greater heterogeneity makes learning from the experiences of others more difficult.

3. Young firms, that have had a limited history to learn about their own specific needs, are also more likely to be decentralized than older firms.

The bulk of the paper investigates these predictions using two large datasets of French firms and establishments and one smaller set of British establishments in the 1990s. We document

[1] Throughout the paper the principal could be thought of as either the owner or the chief operating officer of the firm.

a range of empirical patterns consistent with these predictions; firms closer to the technology frontier of their industry, firms operating in more heterogeneous environments and younger firms are more likely to choose decentralization.

In addition, since our theoretical approach emphasizes the importance of learning about the implementation of new technologies, we also separately look at high-tech industries (defined as those using information technology intensively). Consistent with our theoretical approach, we find that the relationship between heterogeneity or distance to frontier and decentralization is significantly stronger in high-tech than in low-tech industries.

Our main measure of decentralization is whether different units of the firm are organized into "profit centers". We also show that our main results are robust to proxying decentralization by the extent of delayering or measures of managerial autonomy over investment decisions. The results are also robust to the inclusion of a large number of controls, to using various different measures of heterogeneity and to different estimation strategies.

On the theoretical side, our paper is most closely related to the literature on the costs and benefits of delegation or decentralization in firms. A first strand of this literature, for example Baron and Besanko (1992) and Melumad, Mookherjee and Reichelstein (1995), investigate the conditions for delegated contracting to replicate the constrained efficient centralized contracting. As emphasized by Mookherjee (2006), however, the presence of complete contracts in these models implies that delegation can at best replicate the constrained efficient allocation, which is also achievable by centralized contracting. A second strand emphasizes information processing and communication costs as determinants of centralization or decentralization in firms.[2] Although we also stress the importance of learning, our focus is different, namely on how public information affects how much autonomy the principal would like to grant to the agent. Closer to our paper are the recent models emphasizing the trade-off between loss of control and better information under decentralization—in particular, Aghion and Tirole (1997), Baker, Gibbons and Murphy (1999), Rajan and Zingales (2001), Dessein (2002), and Hart and Moore (2005).[3] The main differences between these papers and ours are twofold: first, because there are no incentive effects of the form of the organization, our framework is significantly

[2] See, among others, Sah and Stiglitz (1986), Geanakoplos and Milgrom (1991), Radner (1992, 1993), Radner and Van Zandt (1992), Bolton and Dewatripont (1994), and Garicano (2000).

[3] The possibility that the implementation of new technologies could encourage delegation was first raised by Jensen and Meckling (1992). Aghion and Tirole (1997) emphasize the trade-off between loss of control and the agent's ex ante incentives to acquire information under decentralization. Hart and Moore (2005) show how the trade-off between loss of control and information can explain why in many hierarchies generalists command specialists. Dessein (2002) develops a model in which decentralization to a specialized agent entails a loss of control for the principal, but at the same time reduces the agent's incentive to miscommunicate her information to the principal.

simpler and allows us to focus on the basic trade-off between information and loss of control; second; we allow the principal to learn from other firms' or from their own past experience, which is the source of all the comparative static results we investigate in the empirical work.[4]

The main contribution of our paper is the empirical evidence we provide on the determinants of decentralization. Previous work in the literature focuses on the general move towards "flatter" organizations.[5] Rajan and Wulf (2005) provide the most systematic statistical description of recent organizational trends, showing a strong movement towards flatter corporations in the United States between 1986 and 1999. Bresnahan, Brynjolfsson and Hitt (2002) and Caroli and Van Reenen (2001) report a positive association between various measures of decentralization and organizational change on the one hand and information technology (and human capital) on the other. Baker and Hubbard (2003, 2004) document the effect of new technologies (on-board computers) on ownership patterns in the US trucking industry. Other related papers include Colombo and Delmastro (2004), who present empirical models of decentralization in Italian manufacturing plants, Lerner and Merges (1998), who examine the allocation of control rights in biotechnology alliances, and the papers by Black and Lynch (2001), Ichinowski et al (1997) and Janod and Saint-Martin (2004), which examine the impact of human resource practices and firm reorganization on productivity. None of these papers investigate the relationship between decentralization (or organizational change) and distance to frontier or heterogeneity.

The remaining part of the paper is organized as follows. Section 2 presents some preliminary data description to motivate the basic theoretical framework, which is developed in Section 3. Section 4 describes the data and our main econometric specification. Section 5 presents the empirical results. Section 6 concludes. Appendix A, which contains a more detailed exposition of the theory and the proofs from Section 3, and Appendix B, which contains additional data description and robustness checks, are available upon request and on the Web.[6]

[4]Acemoglu and Zilibotti (1999) present a different model where endogenous accumulation of information affects the internal organization of firms. In their model, a larger number of firms in the economy enables better relative performance of evaluation and creates a shift from direct to indirect monitoring. The number of firms in the economy is, in turn, determined endogenously as a function of the stage of development and the level of the capital stock.

[5]This phenomenon is described by different terms in different contexts, including decentralization, delayering and delegation. In the theory, consistent with the principal-agent literature, we use the term "delegation", while in the empirical analysis, we adopt the terms used in the datasets, namely "decentralization" in the first dataset and "delayering" in the second.

[6]http://econ-www.mit.edu/faculty/index.htm?prof_id=acemoglu&type=paper.

2 Basic Patterns

To motivate our focus in the paper, we first present some salient patterns from a database of several thousand French manufacturing firms, the "Changements Organisationnels et Informatisation" (COI). This dataset, as well as our two other datasets are described below. Our key indicator for decentralization from the COI is whether a firm is organized into profit centers or whether it is more centrally controlled with divisions organized as cost or production centers. A manager of a profit center is concerned with all aspects of the business that contribute to profitability, while managers in charge of production centers focus on output targets and managers of cost centers target costs. When a firm organizes its divisions into profit centers, it typically delegates substantially more authority to its managers (see the discussion in Section 4).

Figures 1-3 show the proportion of over 3,570 firms that are decentralized into profit centers broken down by various firm characteristics. Figure 1 divides firms into deciles depending on the "heterogeneity" of the firm's environment. Heterogeneity is measured as the difference between log productivity (value-added per hour) growth at the 90th and the 10th percentiles of the relevant four-digit industry. This variable is a natural measure of technological heterogeneity among firms within a four-digit industry; it will be greater in industries where some firms are experiencing much faster productivity growth than others. The construction of this variable is discussed in greater detail in Section 4.

Figure 1 shows a general increase in the probability of decentralization as we move from the firms in the least heterogeneous industries to the most heterogeneous industries; while 24% of the firms are decentralized in the second decile, this number is 41% in the tenth decile. The first decile is somewhat anomalous, but closer investigation shows that there is a disproportionately large number of less productive and older firms in these sectors, aspects which we now turn to.

Figure 2 plots the fraction of firms decentralized into profit centers against the "proximity to the frontier" (measured as the ratio of the firm's value added per hour to the value added per hour of the firm at the 99th percentile of the distribution in the same four-digit industry). While 27% of the firms in the bottom quintile of the proximity distribution are decentralized, of the firms in the top quintile, which are closer to the technology frontier, 38% are decentralized.

Finally, Figure 3 shows that younger firms are, on average, more decentralized than older firms: about 45% of the firms under the age of five years are decentralized compared to a rate of 30% for the older firms.

In the rest of the paper, we document that the patterns shown in Figures 1-3 are robust to a variety of controls, different estimation techniques and different measures approximating our theoretical concepts. For example, we show that the same broad patterns apply when we use the variation in productivity levels within four-digit industries (rather than productivity growth) and also a firm-specific index of heterogeneity, measuring the "distance" between the product mix of a firm and those by other firms in the population of French firms. We also show that our main results are robust to using alternative measures of decentralization, including measures of delayering and managerial autonomy over investment decisions (the age results are somewhat weaker with some of these alternative measures of decentralization).

Overall, the patterns in Figures 1-3, and our more detailed results below, suggest that firms that operate in more heterogeneous environments, that are closer to the technological frontier, and that are younger are more likely to be decentralized. These correlations, especially the relationship between decentralization and heterogeneity, indicate that information acquisition and learning may the important factors in the decentralization decisions of firms. In particular, firms in more heterogeneous environments, those closer to the technology frontier and younger firms naturally face greater uncertainty regarding their business decisions in general and the optimal implementation of new technologies in particular compared to firms that are in more homogeneous environments, farther from the frontier and more experienced. This motivates our theoretical approach emphasizing the relationship between learning and decentralization, which is presented in the next section. Although there may be alternative explanations for some of our findings, we are not aware of other approaches that can explain the evidence as satisfactorily as our theoretical framework.

3 Theory

In this section, we describe a theoretical environment linking information and technology choices to decentralization decisions. Our purpose is to highlight a number of implications that will be investigated in the empirical work below. More details on the theoretical framework, as well as the proofs of all the results stated here, are contained in the working paper version, Acemoglu et al. (2006), as well as in Appendix A.

Suppose that there is a ladder of technologies, $k = 1, 2, ...$ At each point in time, $t = 1, 2, ..$, each firm i has previously implemented up to some technology, say $k - 1$. The next technology in the ladder, k, becomes available to this firm with probability $p_i \in (0, 1]$. The parameter p_i thus measures the speed at which firm i climbs the technology ladder. The realizations of technological opportunities are independent across firms and over time. When a new technology becomes available to a firm, it decides how to implement it. In particular, the firm chooses between two actions, L and R, which correspond to two different choices in the implementation of the new technology. Dropping the time index, the choice of the firm is denoted by $x_{i,k} \in \{L, R\}$. One of these choices, $x^*_{i,k} \in \{L, R\}$, leads to successful implementation, while the other leads to an unsuccessful outcome. We will refer to $x^*_{i,k}$ as the *correct action.* Successful implementation of a technology increases the firm's productivity by a factor $\gamma > 1$, while unsuccessful implementation leaves the productivity of the firm unchanged.

We assume that the successful action for firm i in the implementation of technology k is given by

$$x^*_{i,k} = \begin{cases} x^*_k & \text{with probability } 1 - \varepsilon \\ \sim x^*_k & \text{with probability } \varepsilon \end{cases}, \tag{1}$$

where $x^*_k \in \{L, R\}$ is the *reference action* for technology k, $\sim x^*_k$ denotes "not x^*_k" (i.e., if $x^*_k = L$, then $\sim x^*_k = R$) and $0 < \varepsilon < 1/2$. Conditional on x^*_k, the realizations of $x^*_{i,k}$ and $x^*_{i',k}$ for any $i \neq i'$ are independent.[7] We assume that, for each technology, the prior probability that L (or R) is the reference action is equal to $1/2$.

This specification implies that there is a generally correct ("conventional") way of implementing each technology, given by the reference action, but differences in firms' specific needs and competencies imply that some firms need to take a different action for successful implementation. Equation (1) thus makes it clear that ε is a measure of the *heterogeneity* among firms: when ε is equal to zero, the reference action is the correct action for all firms; when ε is equal to $1/2$, the correct action is unrelated across firms.

Each firm is owned by a principal who maximizes its value conditional on the public information available. Successful implementation and hence profits depend on the organization of the firm. The two alternative organizational forms available to each firm are: *centralization* and *delegation.* With centralization (denoted by $d_{i,k} = 0$), the principal manages the firm and chooses $x_{i,k}$; with delegation ($d_{i,k} = 1$), the choice of $x_{i,k}$ is delegated to a manager.

The principal in firm i has no special skills in identifying the right action. Therefore, under centralization she bases her decision on the history of publicly available information relevant for technology k at the time of its decision, denoted by h^i_k. In contrast, the manager of firm i observes the correct action $x^*_{i,k}$, so that he knows exactly which action will lead to successful implementation. However, his interests may not be aligned with those of the owner. Following Aghion and Tirole (1997), we model this in a reduced form way, and assume that the preferred action of the manager for technology k is given by

$$z^*_{i,k} = \begin{cases} x^*_{i,k} & \text{with probability } \delta \\ \sim x^*_{i,k} & \text{with probability } 1 - \delta \end{cases}. \tag{2}$$

[7] This implies that $\{x^*_{i,k}\}_i$ is a Bernoulli sequence with a parameter of $1 - \varepsilon$ or ε (depending on whether $x^*_k = L$ or R).

This specification implies that δ is a measure of *congruence* between the firm's and the manager's objectives. Notice that equation (2) implies that the manager is informed about the right action for this particular firm (not only about the right reference action).

We adopt a number of simplifying assumptions to focus on the main implications of this framework. First, we assume that the relationship between the firm and each manager is short-term. Second, when $x_{i,k} = z^*_{i,k}$, the manager obtains a private benefit. We assume that managers are credit constrained and cannot compensate principals for these private benefits and that these private benefits are sufficiently large that it is not profitable for the principal to use an incentive contract to induce the manager to take the right action. These assumptions imply that delegation will lead to the implementation of the action that is preferred by the manager; thus, when there is delegation, $x_{i,k} = z^*_{i,k}$.[8]

Finally, we assume that $\delta \in (1/2, 1-\varepsilon)$, which implies that the manager's interests are more likely to be aligned with those of the principal than otherwise ($\delta > 1/2$) and that the conflict of interest between the principal and the manager is sufficiently severe that a principal who knows the reference action is more likely to make the correct choice if she, rather than the manager, decides ($\delta < 1-\varepsilon$).

The organizational form and implementation decisions by the principal of firm i for technology k depend on the history of public information h^i_k, which includes the outcomes of all previous attempts with technology k (in particular, which actions were chosen and whether they led to successful implementation). Since conditional on x^*_k the success or failure of different firms in the past are independent, all payoff-relevant information can be summarized by $h^i_k = \{n^i_k, \tilde{n}^i_k\}$: n^i_k is the number of firms that have attempted to implement this technology before firm i, and $\tilde{n}^i_k \leq n^i_k$ is the number of firms for whom L turned out to be the profitable action.[9] Note also that n^i_k is a direct measure of *distance to the technology frontier*. If n^i_k is high, it means that many other firms have implemented technology k before firm i. Therefore, comparative statics with respect to n^i_k will be informative about the impact of the distance to the technology frontier on decentralization decisions.

Let $\pi\left(d_{i,k}; h^i_k\right)$ denote the probability that firm i chooses the correct action conditional on

[8]Put differently, in this model the choice between centralization and delegation simply corresponds to whether or not the "advice" of the manager is followed by the principal. In particular, all the results would be identical if we considered a different game form in which the manager reports his recommendation and then the principal decides which action to take. In this alternative game form, "delegation" would correspond to the principal following the recommendation of the manager. See Acemoglu et al. (2006) and Appendix A for the results in the case where the principal can use incentive contracts.

[9]Note that $\tilde{n}^i_k$ is equal to the number of firms that have adopted technology k before i, chose $x_{i',k} = L$ and were successful, plus the number of firms that chose $x_{i',k} = R$ and were unsuccessful. The public information set also includes the organizational form chosen by firms that have previously adopted technology k, but equation (2) implies that $\tilde{n}^i_k$ is a sufficient statistic for this public information.

history h_k^i and the organizational form $d_{i,k}$. It can be shown that profit maximization in this context is equivalent to maximizing $\pi\left(d_{i,k}; h_k^i\right)$ in every period (see Acemoglu et al., 2006). Hence, the principal will choose $d_{i,k} = 1$ (delegation) if $\pi\left(d_{i,k} = 1; h_k^i\right) > \pi\left(d_{i,k} = 0; h_k^i\right)$.

The above discussion establishes that when the principal chooses delegation, $\pi\left(1; h_k^i\right) = \delta$. On the other hand, under centralization, i.e., $d_{i,k} = 1$, the principal makes the optimal implementation decision given the publicly available information. Consequently, the probability of success when the principal chooses centralization, $\pi\left(0; h_k^i\right)$, is a stochastic variable that depends on history h_k^i, thus both on the firm's distance to the frontier, n_k^i, and on the experiences of firms that have previously implemented the technology, $\tilde{n}_k^i$. As the distance to frontier, n_k^i, increases, the history available to the principal expands and she learns the reference action x_k^* with greater precision. More specifically, when firm i is at the technology frontier, so that $n_k^i = 0$, the principal has no useful information and $\pi\left(0; h_k^i\right) = 1/2$. In contrast, as the principal observes the experiences of sufficiently many other firms, the probability that she chooses the correct action under centralization increases. In particular, it can be shown that $p\lim_{n_k^i \longrightarrow \infty} \pi\left(0; h_k^i\right) = 1 - \varepsilon$.[10] This implies that when n_k^i is small, $\pi\left(0; h_k^i\right)$ will be less than $\pi\left(1; h_k^i\right) = \delta > 1/2$, but as n_k^i increases, it will approach $1 - \varepsilon$ and thus exceed $\pi\left(1; h_k^i\right) = \delta$ (since $\delta < 1 - \varepsilon$). This argument establishes that delegation will be chosen by firms closer to the technology frontier, but not by those that are sufficiently behind. Denoting the optimal organizational choice given history h_k^i by $d_{i,k}^*\left(h_k^i\right)$, we can therefore establish the following result.

Proposition 1 ***(Distance to Frontier)*** Consider the adoption decision of technology k by firm i, and suppose that $\delta \in (1/2, 1 - \varepsilon)$. Then:
(i) For a firm at the technology frontier, i.e., $n_k^i = 0$, the principal chooses delegation, $d_{i,k}^*\left(h_k^i\right) = 1$.
(ii) For a firm sufficiently far from the technology frontier, i.e., $n_k^i \rightarrow \infty$, the principal (almost surely) chooses centralization. That is, $p\lim_{n_k^i \longrightarrow \infty} d_{i,k}^*\left(h_k^i = \left\{n_k^i, \tilde{n}_k^i\right\}\right) = 0$.

In the empirical analysis, we proxy distance to the technology frontier with the gap between the productivity of a particular firm and the highest productivity (or the highest percentile productivity) in the same industry. Firms that are further behind the frontier (i.e., those with higher n_k^i's) will be less productive because they have been unlucky in the past and have had

[10]The statements here and in Proposition 1 show that as $n_k^i \rightarrow \infty$, $\pi\left(0; h_k^i\right)$ will increase towards $1 - \varepsilon$. One might also conjecture that $\pi\left(0; h_k^i\right)$ and hence the probability of centralization should be monotonically increasing in n_k^i. In Acemoglu et al. (2006), we show that when n_k^i is low, integer issues may cause $\pi\left(0; h_k^i\right)$ to be nonmonotonic, but it is increasing in n_k^i "on average," i.e., when we average over neighboring values of n_k^i to smooth out integer issues.

fewer opportunities to adopt technologies, and also because these are typically the firms with lower p_i's that are slower in climbing the technology ladder. Using this proxy, we test the prediction that centralization increases with the distance to the frontier.[11]

Our next result links the parameter of heterogeneity, ε, to firms' decentralization decisions. Let $\Pr\left(d^*_{i,k}\left(h^i_k\right)=1\right)$ be the unconditional probability that firm i will choose delegation when implementing technology k.

Proposition 2 ***(Heterogeneity)*** Consider the adoption decision of technology k by firm i. Given the distance to frontier n^i_k, an increase in heterogeneity, ε, makes delegation more likely. That is, $\partial \Pr\left(d^*_{i,k}\left(h^i_k\right)=1\right)/\partial\varepsilon \geq 0$.

Intuitively, when ε is small, there is less heterogeneity in the environment and the performance of firms that have implemented the same technology in the past reveals more information about the reference action. Consequently, when ε is small, firms' posterior beliefs are more responsive to public information. In other words, given a history h^i_k, $\partial\pi\left(0;h^i_k\right)/\partial\varepsilon \leq 0$, so that as ε increases, delegation becomes more attractive at each history h^i_k.[12] The complication in the proof comes from the fact that a change in ε also affects the likelihood of different histories. Nevertheless, it can be proved that a greater ε changes the ex ante distribution of different histories in a direction that also increases the probability of delegation.

Proposition 2 provides the most interesting testable implication of our approach; it suggests that there should be more decentralization in industries with greater dispersion of performance across firms and also for firms that are more dissimilar to others. In the empirical section, we proxy heterogeneity using three different measures. First, we use the dispersion of productivity growth within a four-digit industry. This is a natural measure since a higher ε corresponds to greater variability in the successful implementation of a *given* technology and thus to greater variability in productivity growth. Second, we check these results using the dispersion in levels of productivity within an industry. Finally, we use a firm-level proxy for heterogeneity, the (IT-weighted) distance between the product mix of a particular firm and those of other firms in the same industry, the idea being that firms with a product mix that is more similar to others should be able to learn more from past experiences of other firms.

In Acemoglu et al. (2006), we extended this framework to derive a relationship between firm age and organizational structure. Firms learn not only from other firms, but also from

[11] Although in this section we state the results in terms of "distance to the technology frontier," in the empirical work it will be more convenient to use the inverse of this, "proximity to the frontier".

[12] Interestingly, this applies to both "correct" and "incorrect" beliefs. For instance, suppose that $x^*_k = L$, but R has been successful more than half of the time; when ε is small, the firm will assign higher probability to R being the correct action.

their own past experiences. The implication of this extension is that younger firms that have accumulated less "firm-specific" information are more likely choose delegation. Motivated by this observation, in our empirical analysis we also investigate the relationship between firm age and delegation.

4 Econometric specification and data

4.1 Empirical strategy

In our empirical analysis, we will document a number of correlations motivated by the theory presented in the previous section. Recall that the main predictions of our approach are:

1. Delegation should be more common for firms closer to the technological frontier.
2. Delegation should be more prevalent in environments with greater heterogeneity.
3. Young firms should be more likely to choose delegation.

We investigate these predictions by studying the relationship between various explanatory factors and decentralization decisions of several thousand French and British firms. Consider the following econometric model for delegation:

$$d^*_{ilt} = \alpha H_{ilt-1} + \beta PF_{ilt-1} + \gamma age_{ilt-1} + \mathbf{w}'_{ilt-1}\boldsymbol{\zeta} + u_{ilt}, \tag{3}$$

where i denotes firm, l denotes industry and t denotes time. d^*_{ilt} is a latent variable indicating the propensity to delegate authority to managers. H_{ilt-1} is a measure of heterogeneity, PF_{ilt-1} is a measure of "proximity to the frontier" (inverse measure of "distance to the frontier"), age_{ilt-1} denotes the age of the firm and $\mathbf{w}_{ilt-1}$ is a vector of other controls. All right-hand side variables refer to $t-1$, while the dependent variable is for t, which is an attempt to prevent the most obvious form of reverse causality. Nevertheless, we do not view estimates from equation (3) as corresponding to causal effects, since there may be other omitted factors, simultaneously affecting both the (lagged) right-hand side variables and the delegation decisions. All omitted factors are captured by the error term u_{ilt}, which we assume to be normally distributed.

In all of our specifications, we observe an indicator of decentralization, $d_{ilt} \in \{0, 1\}$, and in our baseline specifications, we assume that

$$d_{ilt} = \begin{cases} 1 & \text{if } d^*_{ilt} > 0 \\ 0 & \text{if } d^*_{ilt} \leq 0 \end{cases}, \tag{4}$$

where d^*_{ilt} is given by (3). Equation (4), combined with the fact that u_{it} is normally distributed, leads to the standard probit model (Wooldridge, 2002). We therefore estimate (3) by maximum likelihood probit. We check the robustness of our results by using logit and linear probability specifications. Throughout, we report standard errors that are robust and allow for an arbitrary variance-covariance structure at the four-digit industry level.

4.2 Data and measurement

We use two French and one British datasets. The use of multiple datasets is an important cross-validation of the robustness of our results. Our first and main dataset, "Changements Organisationnels et Informatisation" (COI), covers just over 4,000 manufacturing firms.[13] Using unique identifiers, firms in this dataset are matched to the dataset FUTE, which contains the entire population of French firms with more than 20 employees.[14] Many of our right-hand side variables are constructed from the FUTE and thus refer to this entire population. Since the COI contains some firms with less than 20 employees, the match leaves us with a total of 3,570 firms for our basic analysis.

Using the COI, we build a measure of decentralization based on the organization of a firm's business units into profit centers (see Appendix B for a more detailed description). In practice, once a firm grows beyond a certain size it faces the choice of retaining centralized control or allowing some decentralization. Firms are generally organized into business units, with different degrees of responsibility delegated to the managers of these units. While some firms retain complete command and control at the center, most create some form of "responsibility centers" for business unit managers.[15] These responsibility centers (from the most to the least decentralized) are profit centers, cost centers and revenue centers. Our key indicator for decentralization is whether the firm is organized primarily into profit centers. When a firm organizes into profit centers, the manager keeps track of both revenues and costs with the aim of maximizing profits. He is given considerable autonomy in the purchase of assets, hiring of personnel, the management of inventories and determination of bonuses and promotions.[16] In

[13]For previous uses of this dataset, see Aubert, Caroli and Roger (2004), Janod (2002), Janod and Saint-Martin (2004), Crépon, Heckel and Riedinger (2004), and Greenan and Mairesse (1999).

[14]FUTE also contains the population of non-manufacturing firms with more than 10 employees. These data are not published in the French National Accounts, so we worked directly with the underlying micro data located in the French statistical agencies in order to construct the appropriate variables. Similarly the information on the demographic structure of each firm (skills, worker age, hours, gender, etc.) had to be built up from the employee level datasources aggregated to the firm level. See Appendix B for details.

[15]For the meaning of the terms responsibility centers and profit centers in the business literature and in management, see, for example: http://smccd.net/accounts/nurre/online/chtr12a.htm. In addition, http://www.aloa.co.uk/members/downloads/PDF%20Output/costcentres.pdf, provides a standard discussion of autonomy of profit centers. Janod (2002) and Janod and Saint-Martin (2004) have previously used these data on profit centers as a measure of decentralization.

[16]Merchant's (1989, p. 10) book on profit centers explains: "The profit center managers frequently know their business better than top management does because they can devote much more of their time to following up developments in their specialized areas. Hence, top level managers usually do not have detailed knowledge of the actions they want particular profit center managers to take, and even direct monitoring of the actions taken, if it were feasible would not ensure profit center managers were acting appropriately."

contrast, a cost (revenue) center manager is responsible only for costs (revenue). Milgrom and Roberts (1992, pp. 229-230) contrast cost and profit centers managers as follows: “Managers who are given responsibility for profits, for example, are commonly given broader decision authority than those responsible just for costs or sales.” Overall about 30% of French firms in our sample are organized into profit centers.

Our second dataset, the “Enquête Reponse” (ER), is a survey of just under 3,000 French establishments covering all sectors of the economy conducted in 1998. This dataset is also matched with the FUTE to construct the right-hand side variables, which leaves us with a dataset of around 2,200 establishments. In this dataset, delegation can be measured in two ways. First, there is a direct question asked to plant managers regarding the degree of autonomy they enjoy in investment decisions relative to headquarters. Since this question only makes sense for firms that are part of a larger group, the analysis is restricted to this sub-sample (of 1,258 establishments). Second, there is a question related to delayering, which indicates whether there was any reduction in the number of hierarchical layers between 1996 and 1998. Although, a priori, delayering may be associated with more or less delegation (for example, because it may make the chief executive more informed about lower layers), existing evidence shows that delayering tends to involve delegating more power to lower layers of the managerial hierarchy (Rajan and Wulf, 2005, Caroli and Van Reenen, 2001).

Finally, we draw on a UK dataset, the 1998 Workplace Employee Relations Survey (WERS), which is similar in structure to ER. WERS does not have a question on plant managers’ autonomy over investment decisions, but contains a question on their autonomy from headquarters in making employment decisions. We use this question to measure the degree of decentralization. Unlike with the French data, for confidentiality reasons we are not allowed to legally match WERS with productivity at the firm level, though we can match productivity information at the four-digit industry level. Details on all three datasets are in the Appendix B.

Our indicator of proximity to the frontier is the gap between the log labor productivity of a firm (measured as value-added per hour) and the frontier (log) labor productivity in the primary four-digit industry of the firm, $\ln y_{ilt} - \ln y_{Flt}$, where F denotes the frontier, measured in a number of alternative ways. In addition to average labor productivity, we report robustness checks using Total Factor Productivity (TFP). We also construct several alternative indicators of “frontier” productivity. Our main measure is the highest productivity in the firm’s primary four-digit industry (defined as the 99th percentile to mitigate any measurement error from outliers that might arise had we used the maximum) again calculated from the FUTE dataset. We repeat the same exercise using other percentiles (90th and 95th), and we consider alternative measures based on the firm's productivity *rank* in the four-digit industry.

In addition to our main specification, we also allow $\ln y_{ilt-1}$ and $\ln y_{Flt-1}$ to have different coefficients in the regression equation, by estimating

$$d^*_{ilt} = \alpha H_{ilt-1} + \beta_1 \ln y_{ilt-1} + \beta_2 \ln y_{Flt-1} + \gamma age_{ilt-1} + \mathbf{w}'_{ilt-1}\boldsymbol{\zeta} + u_{ilt}. \quad (5)$$

This specification allows us to test whether $\beta_2 < 0$ (that is, whether, as predicted by our theory, delegation is negatively correlated with lagged frontier productivity) and also whether $\beta_1 = -\beta_2$. This robustness check is particularly important, since a positive correlation between distance to frontier and decentralization may reflect a positive effect of decentralization on the firm's own productivity. If this were the case, in equation (5) we would estimate $\beta_2 = 0$.

For heterogeneity H_{il}, we use three measures. All three measures are constructed from the FUTE dataset for the entire covered population of firms (in the UK we use the ABI Census data). Our benchmark measure of heterogeneity, H_l^G, is the dispersion of firm productivity growth within a four-digit sector. This measure captures the effect of the parameter ε in the model of Section 3, since high values of ε correspond to greater heterogeneity in the performance of firms implementing the same technology, thus to greater variability in productivity growth within a sector. We measure productivity growth by the average annual growth in value added per hour over the 1994 to 1997 period, and our main measure of dispersion is the difference in productivity growth rates between the 90th percentile and the 10th percentile in the four-digit industry. Thus we have

$$H_l^G \equiv (\Delta \ln y_{il})^{90} - (\Delta \ln y_{il})^{10} \quad (6)$$

where $(\Delta \ln y_{il})^P$ denotes the P th percentile of the distribution of productivity growth across all firms in industry l. We also consider several alternatives such as the difference between the 95th and the 5th percentiles (instead of the 90-10), the standard deviation of firm productivity growth rates and the standard deviation of the trimmed productivity growth distribution.

We also present results with an alternative measure of a heterogeneity, H_l^L, constructed similarly to H_l^G, but using productivity levels instead of growth rates (i.e., $H_l^L \equiv \ln y_{il}{}^{90} - \ln y_{il}{}^{10}$). This measure has two empirical disadvantages relative to our benchmark. First, it is likely to be correlated with the distance to the frontier, so the heterogeneity and proximity terms may be hard to identify separately. Second, the growth-based measure, H_l^G, is likely to be a better proxy for ε since it differences out time-invariant omitted variables affecting the level of productivity that are observable to firms but not the econometrician (such as management quality, brand differences, etc.).

Both of these measures of heterogeneity do not vary within a four-digit industry. Our third measure, H_i^F, is a firm-specific index of heterogeneity and quantifies (the inverse of) how many other firms are close "neighbors" of the firm in question in the product space. When there are more similar firms (neighbors), the firm will have greater opportunity to learn from the experiences of others and this will correspond to a lower value of ε in terms of our theoretical model.[17] To calculate this measure, for each firm i we compute the distribution of production across all four-digit sectors. We define $s_i \equiv (s_{i1}, \ldots, s_{il}, \ldots, s_{iL})$ as firm i's shares of production in each industry $l = 1, ..., L$ (by definition $\sum_{l=1}^{L} s_{il} = 1$). An element of the vector s_i will be equal to zero if a firm produces nothing in industry l and unity if a firm produces all its output in that particular industry. We then calculate the "closeness" of any two firms, i and i' in the FUTE as the uncentered correlation coefficient,

$$c_{ii'} \equiv \frac{\sum_{l=1}^{L} s_{il}.s_{i'l}}{\left(\sum_{l=1}^{L} s_{il}^2\right)^{\frac{1}{2}} \cdot \left(\sum_{l=1}^{L} s_{i'l}^2\right)^{\frac{1}{2}}},$$

which takes greater values when the production profiles of two firms are more similar and is equal to unity when the two profiles are identical. Since our theoretical approach emphasizes the importance of similarity in the context of experimenting with new technologies, our preferred measure of firm-specific heterogeneity is constructed with information technology (IT) weights:

$$H_i^F \equiv \log \left(\frac{\sum_{i', i' \neq i} c_{ii'} \cdot IT_{i'}}{\sum_{i', i' \neq i} IT_{i'}} \right)^{-1} \tag{7}$$

where $IT_{i'}$ is the level of investment in IT by firm i'. We also check the robustness of our results by looking at an alternative unweighted measure. The "inverse" in equation (7) makes sure that high levels of H_i^F correspond to high values of ε in terms of our theory.

An important concern with this firm-level heterogeneity measure is that it may be related to the level of product market competition. If there are many other firms "close" to a company in the product market space, then this company may be facing tougher competition.[18] To alleviate this concern, we control for various measures of product market competition, in particular the Lerner index (a proxy for price-cost margins), calculated as gross profits—value added minus labor costs—divided by sales, from the FUTE dataset. Moreover, we document below that there is a robust positive relationship between product market competition and delegation,[19] so the possible negative correlation between product market competition and

[17] This measure of closeness is inspired by Jaffe's (1986) approach in the context of patent spillovers, but uses the proportion of production in a four-digit industry. Jaffe originally used patent technology class, which has the potential disadvantage that many firms do not patent, especially in service sectors.

[18] See Bloom, Schankerman and Van Reenen (2004) for a discussion.

[19] See Nickell, Nicolitsas and Patterson (2001) and McKinsey Global Institute (2002).

H_i^F will, if anything, bias the results towards finding a negative effect of heterogeneity on delegation, which is the *opposite* of the prediction in Section 3.

In addition, since our theory emphasizes the implementation of new technologies, we estimate (3) and (5) separately in high-tech and low-tech subsamples (as measured by industry IT intensity). We expect the patterns suggested by our model to be more pronounced for high-tech firms.

To measure age, age_{ilt-1}, we use four dummies; age less than 5 years, between 5 and 9 years, between 10 and 19 years, and the reference category, greater than or equal to 20 years.

Means, medians and standard deviations for all the main variables are presented in Table 1. Appendix B gives greater detail on the data used. The average firm in our data has 323 employees, was born 22 years ago and has 3 plants.

5 Results

5.1 Decentralization

Table 2 presents our basic findings using the decentralization measure from COI. Throughout, all regressions are estimated by maximum likelihood (ML) probit and we report marginal effects evaluated at the sample mean. All standard errors are computed using the Huber formula, where we allow for heteroskedasticity and clustering at the four-digit industry level. All regressions also include a full set of three-digit industry dummies.[20]

The first column includes only our key variables: heterogeneity (measured by the 90-10 of firm productivity growth), frontier productivity (measured as the 99th percentile of the productivity distribution in the firm's primary four-digit industry), own productivity, age dummies and the three-digit industry dummies. The results are consistent with the predictions in Section 3—all key variables take their expected signs and are statistically significant at the 5% level. The marginal effects of heterogeneity and own productivity are positive, whereas the marginal effects of frontier productivity and age are negative.[21] Firms in more heterogeneous environments are significantly more likely to be decentralized (the marginal effect is 0.211, while the standard error is 0.107). The youngest firms (under 5 years old) are 15% more likely to be decentralized than the oldest firms (those over 20 years old) and this difference is

[20] Since the frontier productivity term and the heterogeneity measure H_l^G are defined at the four-digit level, we could not identify their effects if we included four-digit industry dummies. Instead, with a full set of four-digit dummies, we can only identify the marginal effects of age and firm-specific productivity. These remain correctly signed and significant at conventional levels with a full set of four digit dummies.

[21] When included individually each variable is also significant. For example, when we drop all other variables except the industry dummies, the marginal effect of heterogeneity is -0.212, which is significant at the 5% level (see Acemoglu et al., 2006, for details).

significant at the 5% level. In column (2), we combine the frontier productivity and the own productivity terms in a single "proximity to frontier" term as in equation (3). The marginal effect of proximity to frontier is 0.167 (standard error = 0.024), while the marginal effect of heterogeneity is 0.252 (standard error = 0.102).[22] Overall, these patterns suggest that firms that are in more heterogeneous environments and closer to the technology frontier of their industry are more likely to be decentralized.[23]

The remaining columns in Table 2 include a large number of additional controls to check whether the correlations we report are driven by omitted variables. These additional covariates are: the Lerner index, a foreign ownership dummy, the log number of plants of the firm, the (log of) capital stock divided by value added, log firm size, the fraction of employees working with computers, the fraction of high skilled workers, the average age of workers, the firm's market share, a specialization/inverse diversification index, as well as a number of industry-level variables, in particular, (the log of) capital stock divided by employment, IT expenditures divided by employment and the Herfindahl index. The fixed capital stock and computer use variables are included both as potential controls and also to bring the measure of labor productivity closer to TFP by controlling for the contribution of various components of the capital stock. Firm-level worker characteristics are included since these may affect organizational choices; for example, firms with more skilled workers and/or younger workers might be more likely to decentralize control. The additional controls improve the fit of the model, but the heterogeneity, age and productivity terms all remain individually significant at the 5% level or less. Also notable is that in the specification of column (3), which includes all the additional covariates, we do not reject the hypothesis that $\beta_1 = -\beta_2$, that is, the hypothesis that frontier and own labor productivity terms have equal and opposite-signed coefficients (p-value>0.10).

The estimated effects of these other variables are consistent with the existing literature. Firms that are more skill-intensive (Caroli and Van Reenen, 2001), that employ younger workers (Aubert et al, 2006), that have more workers using computers and/or that are more in more IT-intensive industries (Bresnahan et al, 2002) appear significantly more likely to be decentralized. Furthermore, firms that are large, multi-plant, foreign owned and/or less specialized (more diversified) are also more likely to be decentralized, possibly because their production processes are more complex. Firm-level capital stock or industry-level capital stock do not

[22]The Wald test rejects the restriction that $\beta_1 = -\beta_2$ at the 5% level, though when additional covariates are included in columns (3) and (4), this restriction is no longer rejected.

[23]Since our main models are cross-sectional, we cannot distinguish age and cohort effects. Consequently the positive coefficient on age may be driven by firms founded in more recent years being more likely to adopt "best practice" organizational forms (see, e.g., Ichinowski et al, 1997).

appear to have a major effect on decentralization. There is also a robust negative relationship between the Lerner index, our (inverse) proxy for product market competition, and the probability of decentralization, which implies that more concentrated industries are associated with less decentralization. We discuss this association further in the concluding section.

Since the theory in Section 3 relates decentralization decisions to the adoption of new technologies, we expect a stronger relationship between decentralization and heterogeneity in the high-tech industries than in the low-tech sectors. We define "high-tech" sectors to be those with an average ratio of IT investment per worker greater than the sample median and re-estimate the equations on these two sectors separately. Consistent with our expectations, the marginal effects and significance of all the key variables are greater in the high-tech sectors than in the low-tech sectors. For example, the heterogeneity index, H_l^G, is positive and significant in the high-tech sectors (column (5)), but negative and insignificant in the low-tech sectors (column (6)). The marginal effects of proximity to frontier and of the youngest age dummy are twice as large in the high-tech sectors as in the low-tech sectors. Wald tests show that these differences are significant at the 1% level for heterogeneity and at the 5% level for proximity to frontier (but insignificantly different for age).

Overall, the results in Table 2 suggest that, consistent with our theory and the relationships shown in Figures 1-3, firms that operate in more heterogeneous environments, that are closer to the technology frontier, and that are younger are significantly more likely to be decentralized.

5.2 Magnitudes

To gauge the quantitative magnitudes of the estimates in Table 2, we look at the impact of doubling each variable starting from its sample mean.

Using the estimate of the marginal effect of heterogeneity in column (4) of Table 2, 0.251, we find that doubling the mean value of heterogeneity (the 90-10 of firm productivity growth in the industry) from 0.275 to 0.550 increases the predicted probability of a firm being decentralized into profit centers by approximately 7 percentage points ($0.251 \times 0.275 \approx 0.069$) starting from a base of 30 percent, which is a sizeable effect. Thus, in "elasticity" terms, a doubling of heterogeneity is associated with a 23% increase in the probability of decentralization (a 6.9 percentage point increase on a base of 30 percent).[24]

Again using the estimate from column (4) of Table 2, doubling the proximity measure leads to a substantial increase in the probability of decentralization of about 11 percentage points which represents a 37% increase on the base of about 30 percent ($0.164 \times \ln 2/0.3$). Also

[24] A one standard deviation increase in heterogeneity (0.087) results in a smaller increase in decentralization probabilities: a 2.2 percentage point, or 7.3% increase.

using the estimates from column (4) of Table 2, doubling the age of a firm from four years to eight years reduces the probability of decentralization by a third (11 percentage points on a 30 percent base). These calculations suggest that the statistical associations documented in Table 2 appear to be economically as well as statistically significant.

5.3 Alternative measures of heterogeneity

Table 3 contrasts our basic measure of heterogeneity (the inter-decile range of firm productivity growth rates in the industry) with several alternative indicators of heterogeneity. The first column of Table 3 replicates the specification from the last column of Table 2 for comparison. The next three columns (2)-(4) use alternative measures of heterogeneity, H_l^G still based on the dispersion of productivity growth rates across firms within the four-digit industry. Column (2) shows a similar result to column (1) using the difference between the productivity growth rates at the 95th and 5th percentiles (instead of the 90th and 10th percentiles). The marginal effect is 0.142 with a standard error of 0.069. In column (3) we use the standard deviation of the growth rate, which also has a positive marginal effect, but is only significant at the 10% level. This lack of precision may be due to a number of outliers in the firm-level productivity growth distribution. In column (4), we use the standard deviation calculated after trimming the top and bottom 5% of the firm-level productivity growth distribution and obtain a much larger and much more significant marginal effect.

Column (5) includes the heterogeneity term based on firm productivity levels, H_l^L. The marginal effect of this variable is positive, but is not statistically significant at the 5% level. The estimated magnitude is comparable to that in column (1), however; a doubling of H_l^L is associated with a 27% increase in decentralization (a 8.1 percentage point increase on a base of 30 percent) compared to 23% for our benchmark measure, H_l^G. Furthermore, as with our benchmark results in Table 2, the level-based measure of heterogeneity, H_l^L, has a large and statistically significant marginal effect of 0.271 in the high-tech sample (column (6)). In contrast, its marginal effect is insignificant (and negative) in the low-tech sample (column (7)) and is also significantly different from the estimate in the high-tech sample (p-value=0.009).

Columns (8)-(11) report results using the firm-level measure of heterogeneity, H_i^F.[25] Recall that this index measure is the (inverse) IT-weighted distance of a firm to all other firms. This is an entirely different source of variation in heterogeneity and thus constitutes a useful cross-validation of the main results. In column (8) of Table 3 H_i^F has a marginal effect of 0.063 and a standard error of 0.031. The next two columns show that, as with the other

[25] All of the results in Table 2 are similar if we use this measure. For example, the most parsimonious specification in column (1) of Table 2 gives a marginal effect of H_i^F of 0.112 with a standard error of 0.034.

measures, the effect of heterogeneity is stronger in the high-tech subsample than among the low-tech industries (0.098 with a standard error of 0.048 versus 0.019 with a standard error of 0.037). Finally, in column (11) we look at the simpler unweighted measure of the firm-level heterogeneity measure. This is also useful as another check to see whether this measure is capturing some competition-related factors. If that were the case, we would expect the unweighted measure to be stronger. The unweighted measure also has a positive effect, but with a smaller coefficient that is only statistically significant at the 10% level. This suggests that, consistent with our theoretical approach, the IT weights increase the explanatory power of the firm-level heterogeneity index.[26]

Overall, the results in this table show that there is a robust positive association between heterogeneity and decentralization, particularly in high-tech industries.[27]

5.4 Further robustness checks

In addition, we conducted a large number of robustness checks (see Acemoglu et al, 2006, and Appendix B for details). These checks show that our main results do not depend on the precise functional form, control variables or the exact sample.

First, estimating the marginal effects by OLS or logit gave very similar marginal effects to the probit baseline of column (3) Table 2.

Second, alternative measures of productivity and distance the frontier also gave similar results. For example, results using total factor productivity were very close to those using labor productivity. We also experimented with alternative definitions of the distance to frontier using an ordinal measure (the rank of the firm's labor productivity in the four-digit industry) or lower percentiles of the productivity distribution to the 99th percentile in order to measure the frontier (e.g., the 95th and the 90th). Again, the results were qualitatively similar, but the marginal effects of the frontier became progressively weaker as we used the 95th and the 90th percentiles. This pattern is not surprising, since we expect 95th and 90th percentiles to be are poorer measures of the technology frontier than the 99th percentile.

Third, although a single firm can be organized into divisions with each division being decentralized as a profit center, the measure of profit centers may be more natural for firms that

[26] If we include both the weighted and the unweighted measures together with all covariates, the weighted measure is positive and significant at the 5% level (marginal effect=0.184, standard error=0.096), while the unweighted measure is negative (marginal effect=-0.142, standard error=0.109).

[27] One concern with any measure of heterogeneity is that, since it is correlated with uncertainty in firm's environment, it may affect the extent of the moral hazard problem between the firm and the manager (an issue we have abstracted from in the model). Nevertheless, everything else equal, this effect would bias the results against finding a positive association between heterogeneity and decentralization, since greater uncertainty should increase agency costs, making decentralization less attractive.

are part of larger groups. To investigate this issue further and also to exclude potentially owner-managed firms, we re-estimated our basic specification on the sub-sample of 1,793 firms that are part of a larger corporate group. Reassuringly, the effects of heterogeneity and proximity to frontier are now considerably larger and more significant (0.461 with a standard error of 0.140 for heterogeneity and -0.303 with a standard error of 0.056 for the frontier).

Another concern is that we have allocated a single "frontier" to each firm, whereas firms that operate across multiple industries will have multiple "frontiers". To address this concern, we limited the sample to firms that have at least 80% of their sales in their primary four-digit industry, since the multiple industry issue should not be a serious concern for these firms. In this limited sample of 2,555 firms, both heterogeneity and the frontier terms remain highly significant, but the marginal effect of the frontier term is somewhat smaller; -0.179 instead of -0.225 in the baseline specification.

We also estimated instrumental-variable models to address the issue of endogeneity of our main right-hand side variables. Our strategy was to use the UK counterparts of our variables as instruments. Although this approach does not solve all possible endogeneity problems, it is a useful check against reverse causality concerns. We constructed heterogeneity variables identical to H_l^G based on the dispersion of productivity growth among British firms for the same time period to instrument French industry-level heterogeneity. We also constructed the 99th percentile of the productivity distribution in each four-digit British industry as a potential instrument for the French proximity to frontier. The details are provided in Acemoglu et al. (2006). Briefly, these instruments are highly significant in the first stages. Using instrumental-variables probit (see Lee, 1981), we estimated positive and significant effects of both heterogeneity and proximity to frontier in the second stage. The marginal effects in this case are 1.572 for heterogeneity and 0.456 for the proximity to frontier (compared to 0.230 and 0.167 when treating these variables as exogenous). These instrumental-variable results therefore suggest that, if anything, treating heterogeneity and proximity as exogenous may be causing some attenuation due to measurement error and making us underestimate the impact of heterogeneity and proximity to the frontier on decentralization.

5.5 Alternative measures of decentralization

Two alternative measures of decentralization are control over investment decisions and delayering. Whether an establishment's senior managers can make investment decisions without consulting headquarters is clearly directly related to delegation of authority. Case studies and econometric evidence suggest that reducing the layers of the managerial hierarchy tends to be associated with decentralized decision making.[28] There are questions on delayering and the autonomy of investment decision making in our second French dataset, the Enquête Reponse (ER).[29] Delayering is defined as the removal of one or more layers of the managerial hierarchy between 1996 and 1998. We defined our indicator of investment autonomy/decentralization to be equal to unity if the plant manager had "full" or "important" authority in making investment decisions independently of central headquarters and zero if she had "limited" or "no" autonomy in making such decisions.

Table 4 shows the results of estimating equation (5) for these alternative measures both for the full sample and also separately for the high-tech and the low-tech sub-samples (constructed using industry IT intensity as in Table 2). In addition, we limit the sample to firms that are part of a larger group, because the question on delegation of investment decisions from headquarters is only relevant for these firms. In columns (1)-(3) the dependent variable is an indicator of whether the firm allows autonomy over investment decisions to its plant managers. In columns (4)-(9), the dependent variable is an indicator for whether there was a reduction in the number of layers in the managerial hierarchy between 1996 and 1998.

In column (1) of Table 4 frontier productivity is negatively and significantly related to the probability of allowing managers to make investment decisions without consulting headquarters (decentralization). Heterogeneity is positively related to decentralization, but (like age) is insignificant. When we distinguish between high-tech (column (2)) and low-tech sectors (column (3)), the results are stronger. In the high-tech sectors, the marginal effect of heterogeneity is positive and significant at 5%, whereas in the low-tech sectors it is insignificant. Similarly, the marginal effect of proximity to the frontier is negatively and significantly related to decentralization in the high-tech sample, but is positive and insignificant in the low-tech sample. Own productivity and age are insignificant in both subsamples.

The next six columns use the measure of delayering as the dependent variable. In column (4) the productivity terms are both correctly signed and significant at the 5% level, suggesting that the closer is a firm to the technology frontier the more likely it is to choose delayering. Younger firms are significantly more likely to delayer than older firms.[30] The heterogeneity

[28] See, for example, Rajan and Wulf (2005) or Caroli and Van Reenen (2001).

[29] In the COI dataset there is an indicator of the number of hierarchical levels, but as discussed in Appendix B, a better datasource to measure delayering is the ER. Briefly, this is because the Enquête Reponse question on delayering refers explicitly to changes in management, which is more in line with the theory. The COI question, by contrast, refers to the number of "hierarchical levels" and is thus likely to be more informative on hierarchies involving production workers.

[30] Although the ER data is at the establishment level, the regressions in Table 4 use firm age to make the results comparable to those in Tables 2 and 3. The young firm dummy remains positive and significant if we also condition on establishment age.

term is positive and significant at the 10% level. When we split the sample into high-tech (column (5)) and low-tech (column (6)) sectors, the marginal effects of heterogeneity and proximity are again much larger in the high-tech sectors than in the low-tech sectors, but the standard errors are also much larger in both samples. In contrast, the age effects are larger in the low-tech sample, which is the opposite of the prediction of our theory.

Since the delayering variable measures "organizational change" (rather than the "level" of decentralization as in our previous dependent variable), we also considered regressions where the productivity terms are in differences rather than in levels. Since we do not have reliable time-series information on the heterogeneity term and some of the other covariates, they are still included in levels. The results, presented in columns (7)-(9), are similar to the benchmark estimates, but somewhat weaker. The frontier growth term is correctly signed, but no longer significant and the own productivity term is also insignificant.[31] The heterogeneity measure remains positive and significant in the full sample. With the sample split, it is no longer significant in either sample (presumably because of the smaller number of observations), though, as expected, the marginal effect is substantially larger in the high-tech sample.

In summary, the results from using delayering and autonomy over investment decisions as alternative indicators of decentralization broadly support our earlier conclusions. Decentralization appears to be more likely when the environment is more heterogeneous and firms are closer to the technology frontier, particularly in high-tech sectors, though the age results appear to be somewhat less robust.

5.6 Decentralization in Britain

We complement our evidence from the French micro datasets with an analysis of the British Workplace Employee Relations Survey (WERS98). The French Enquête Reponse was modeled on WERS and we use the 1998 wave to match the year used in ER. The WERS cross-section does not have a question on autonomy over investment decisions, but there is a similar question on the establishment manager's autonomy over employment decisions. Senior managers were asked whether they were able to take decisions on staff recruiting without consulting company headquarters. Our WERS sample is further restricted because we are only able to match manufacturing establishments to industry-level information (unfortunately census information on

[31] The weakness of the frontier growth term in this case is related to the higher correlation between productivity growth and heterogeneity variables (recall that heterogeneity is defined here as the decile ratio of productivity growth rates in the firm's four-digit industry). In column (7) if we drop the heterogeneity and firm productivity terms, the marginal effect of frontier productivity growth increases to -0.064 with a standard error of 0.031. If we use the full specification of column (7), but just include two-digit (instead of the usual three-digit) industry dummies, the marginal effect of frontier growth becomes -0.074 with a standard error of 0.038.

non-manufacturing is not available over this time period). Finally, we are unable to condition on the rich set of firm level covariates as in the French data, because confidentiality restrictions limit the data that can be matched at the firm level (such as firm-level output, capital or age).

The results are presented in Table 5. Column (1) includes the first measure of heterogeneity (the difference between the 95th percentile and the 5th percentile of the productivity growth rates in the four-digit industry) with only a full set of three-digit industry dummies as extra controls. Heterogeneity is positively and significantly associated with decentralization at the 5% level. The next column performs the same exercise for the 90-10, the relationship is still positive and significant at the 10% level. Column (3) includes the frontier growth term which is negatively signed as we would expect from the theory, but insignificant. The fourth column includes the age dummies. These are insignificant and show no clear pattern (possibly because in this dataset we only have establishment age rather than firm age).

The fifth and sixth columns include all the covariates. There appears to be some evidence that firms facing less competition are significantly less likely to decentralize. More importantly, the heterogeneity terms measured either as the 90-10 (column (5)) or the 95-5 (column (6)) percentile differences remain positive and significant. The frontier term enters negatively in the regressions in both columns and is significant at the 5% level. Both of these findings are consistent with our theory and with the results we presented from the French datasets, even though they are taken from a different dataset from a different country.

6 Conclusions

Despite considerable academic and popular interest in the changes in the internal organization of the firm, we are far from a theoretical or an empirical consensus on the determinants of the organizational decisions of firms and on the reasons why there has recently been a significant move towards greater decentralization. In this paper we presented a simple model of the relationship between technology, information and decentralization and empirically investigated the main predictions of this model using three micro-level datasets. In our model, firms delegate authority to managers, i.e., "decentralize", in order to use the manager's superior information about the implementation of new technologies. Because the interests of the manager and the principal are not perfectly aligned, such delegation entails a costly loss of control for the principal. The model predicts that as available public information about the implementation of new technologies increases, firms should become less likely to decentralize, whereas firms dealing with new (frontier) technologies should be more likely to decentralize. We also showed that firms in more heterogeneous environments and young firms are more likely to choose decentralization. These are intuitive, but quite novel, predictions, and have, to the best of our knowledge, never been investigated empirically.

We documented that in all three datasets the correlations are broadly consistent with the predictions of our model. Firms in more heterogeneous environments and those that are closer to the frontier of their industry are more likely to choose decentralization. Moreover, consistent with the predictions of the theory, these results are stronger for firms in high-tech sectors. The results are robust to using a variety of alternative measures of decentralization and heterogeneity. We also found that younger firms tended to be more likely to decentralize, though this result was less robust when we looked at alternative measures of decentralization. These results suggest that the recent move towards more decentralized organizations may be driven, in part, by the rapid diffusion of new technologies and the entry of young firms.

The theory and empirical results taken together suggest that learning and information accumulation may have important effects on the internal organization of firms and may be especially important for decentralization decisions. Our analysis also highlights a number of avenues for future research. First, the same forces pushing towards decentralization may also encourage spinoffs and reduce the incentives for vertical integration. An important area for future research is to study vertical integration and decentralization decisions jointly.

Second, our empirical results showed a robust positive association of competition and decentralization. An interesting question is to investigate the channels through which competition may affect decentralization. One possibility is that competition may increase the value of information, because falling behind competitors may be costly to firms, thus encouraging delegation to the manager who has superior information. Yet another effect of a more competitive environment may be through disciplining the manager; faced with greater competition, managers may be forced to take profit-maximizing decisions more often, thus reducing the conflict of interest between the principal and the manager. This would naturally increase delegation, since delegation becomes more attractive to the principal.

Another interesting area of future research is to investigate whether the statistical associations between proximity to frontier or heterogeneity and decentralization correspond to the causal effects of these variables on the internal organization of the firm. Specifying and estimating a more structural model would be a fruitful approach for this purpose. Finally, our approach suggests a natural reason for cross-country differences in the internal organization of firms; we may expect less decentralization in developing countries where most firms use well-established (rather than frontier) technologies. Theoretical and empirical analysis of cross-country patterns of organizational forms is another promising area for future research.

References

[1] Acemoglu Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti (2006) "Technology, Information and the Decentralization of the Firm" NBER Working Paper No. 12206.

[2] Acemoglu, Daron, and Zilibotti, Fabrizio (1999), "Information Accumulation in Development," *Journal of Economic Growth*, 4, 5-38.

[3] Aghion, Philippe, and Tirole, Jean (1997), "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1-29.

[4] Aubert, Patrick, Caroli, Eve, and Roger, Muriel (2006), "New Technologies, Workplace Organization and the Structure of the Workforce: Firm-Level Evidence," *Economic Journal*, 116, 73-93.

[5] Baker, George, Gibbons, Robert, and Murphy, Kevin (1999), "Informal Authority in Organizations," *Journal of Law, Economics, and Organization*, 15, 56-73.

[6] Baker, George, and Hubbard, Thomas (2003), "Make Versus Buy In Trucking: Asset Ownership, Job Design and Information", *American Economic Review*, 93, 551-72.

[7] Baker, George, and Hubbard, Thomas (2004), "Contractibility and Asset Ownership: On Board Computers and Governance in US Trucking," *Quarterly Journal of Economics*, 119, 1443-79.

[8] Baron, David, and Besanko, David (1992), "Information, Control, and Organizational Structure", *Journal of Economics and Management Strategy*, 1, 237-275.

[9] Black, Sandra, and Lynch, Lisa (2001), "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity," *Review of Economics and Statistics*, 83, 434-445.

[10] Bloom, Nicholas, Schankerman, Mark, and Van Reenen, John (2004), "Technological Spillovers and Product Market Rivalry", *Centre for Economic Performance Discussion Paper*, 675.

[11] Bolton, Patrick, and Dewatripont, Mathias (1994), "The Firm as a Communication Network," *Quarterly Journal of Economics*, 109, 809-839.

[12] Bouwens, Jan, and Van Lent, Laurence (2004), "Assessing the Performance of Profit Center Managers," mimeo, University of Tilburg.

[13] Bresnahan, Timothy, Brynjolfsson, Erik, and Hitt, Lorin (2002), "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence," *Quarterly Journal of Economics*, 117, 339-376.

[14] Caroli, Eve, and Van Reenen, John, (2001), "Skill Biased Organizational Change," *Quarterly Journal of Economics*, 116, 1448-1492.

[15] Colombo, Massimo, and Delmastro, Marco (2004), "Delegation of Authority in business organizations: An empirical test," *Journal of Industrial Economics*, 52, 53-80.

[16] Crépon, Bruno, Heckel, Thomas, and Riedinger, Nicolas (2004), "Computerization, Organizational Change and Productivity: Microeconometric Evidence for France," mimeo, INSEE.

[17] Dearden, John (1987), "Measuring profit center managers," *Harvard Business Review*, 65, 84-88.

[18] Dessein, Wouter (2002), "Authority and Communication in Organizations," *Review of Economic Studies*, 69, 811-838.

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

[20] Geanakoplos, John, and Milgrom, Paul (1991), "A Theory of Hierarchies Based on Limited Managerial Attention," *Journal of the Japanese and International Economies*, 5, 205-225.

[21] Greenan, Nathalie, and Mairesse, Jacques, (1999), "Organizational Change in French Manufacturing: What Do We Learn From Firm Representatives and From Their Employees," *NBER Working Paper*, 7285.

[22] Hart, Oliver, and Moore, John (2005), "On the Design of Hierarchies: Coordination versus Specialization," *Journal of Political Economy*, 113, 675-702.

[23] Ichniowski, Casey, Prenushi, Giovanna, and Shaw, Kathryn (1997), "The effects of Human Resource Management: A Study of Steel Finishing Lines," *American Economic Review*, 87, 291-313.

[24] Jaffe, Adam (1986), "Technological Opportunity and Spillovers from Firms' Patents, Profit, and Market Value," *American Economic Review*, 76, 984-1001.

[25] Janod, Véronique (2002), "Changements Organisationnel, Qualification et Croissance," PhD Dissertation, Université Paris 1 Panthéon-Sorbonne.

[26] Janod, Véronique, and Saint-Martin, Anne (2004), "Measuring Work Reorganization and its Impact on Firm Performance: an Estimate on French Manufacturing Firms over 1995-1999," *Labour Economics*, 11, 785-798.

[27] Jensen, Michael C. and Meckling, William H. (1992) "Specific and General Knowledge, and Organizational Structure," in Lars Werin and Hans Wijkander (eds) *Contract Economics*, Blackwell: Oxford

[28] Lee, Lung-Fei (1981), "Simultaneous equation models with discrete and censored dependent variables," in Manski, Charles, and McFadden, Daniel (eds), *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge MA: MIT Press, 346-364.

[29] Lerner, Josh, and Merges, Robert (1998), "The Control of Biotechnological Alliances: An Empirical Examination of the Biotechnology Industry," *Journal of Industrial Economics*, 46, 125-156

[30] McKinsey Global Institute (2002), "Reaching Higher Productivity Growth in France and Germany," *MGI Reports*, http://www.mckinsey.com/knowledge/mgi/europe/index.asp

[31] Melumad, Nahum, Mookherjee, Dilip, and Reichelstein, Stefan (1995), "Hierarchical Decentralization of Incentive Contracts," *Rand Journal of Economics*, 26, 654-672.

[32] Merchant, Kenneth (1989), *Rewarding Results: Motivating Profit Center Managers*, Cambridge: Harvard Business School Series in Accounting and Control.

[33] Milgrom, Paul, and Roberts, John (1992), *Economics, Organization and Management*, New Jersey: Prentice Hall.

[34] Mookherjee, Dilip (2006), "Decentralization, Hierarchies and Incentives: A Mechanism Design Approach," *Journal of Economic Literature*, 64, 367-390.

[35] Nickell, Stephen (1996), "Competition and Corporate Performance," *Journal of Political Economy*, 104, 724-746.

[36] Nickell, Stephen, Nicolitsas, Daphne, and Patterson, Malcolm (2001), "Does Doing Badly Encourage Managerial Innovation?," *Oxford Bulletin of Economics and Statistics*, 3, 5-28.

[37] Radner, Roy (1993), "The Organization of Decentralized Information Processing," *Econometrica*, 61, 1109-1146.

[38] Radner, Roy, and Van Zandt, Timothy (1992), "Information Processing in Firms and Returns to Scale," *Annales d'Economie et de Statistique*, 25/26, 265-298.

[39] Rajan, Raghuram, and Zingales, Luigi (2001), "The Firm as a Dedicated Hierarchy: A Theory of the Origin and Growth of Firms," *Quarterly Journal of Economics*, 116, 805-851.

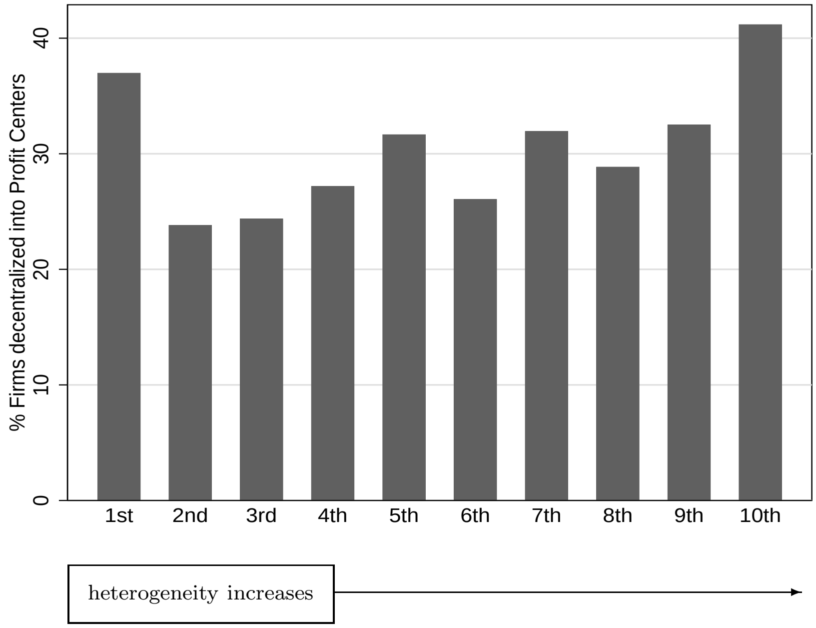
[40] Rajan, Raghuram, and Wulf, Julie (2005), "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies," *Review of Economics and Statistics*, forthcoming.

[41] Sah, Raaj, and Stiglitz, Joseph (1986), "The Architecture of Economic Systems: Hierarchies and Polyarchies," *American Economic Review*, 76, 716-727.

[42] Wooldridge, Jeffrey (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press.

Figure 1: Heterogeneity and Decentralization

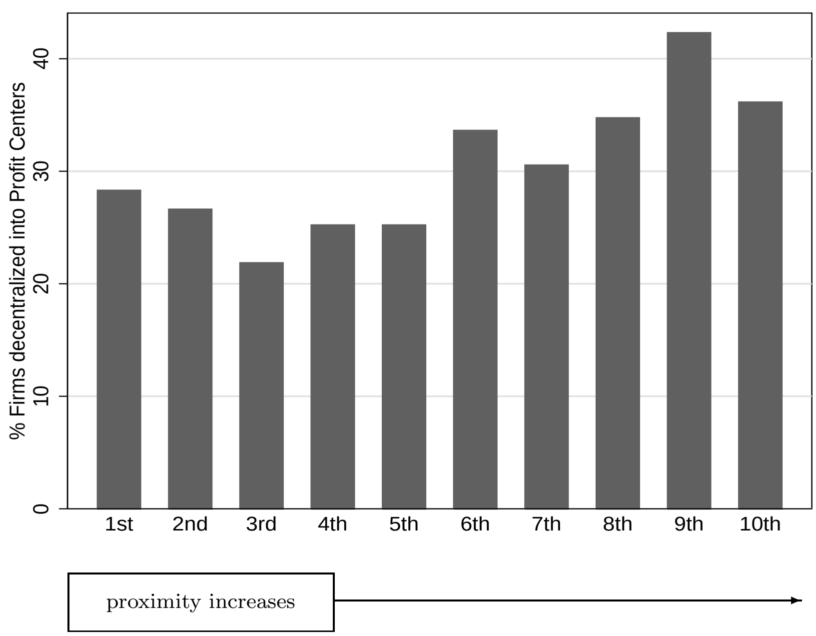
Decentralization to Profit Centers (COI)



NOTES: The X-axis divides all firms into deciles of heterogeneity from the first decile (low heterogeneity) to the tenth decile (high heterogeneity). Heterogeneity is measured by an index of dispersion (the difference of the 90^{th} minus the 10^{th} percentile) of productivity growth between firms in a four digit industry (see text). The Y-axis indicates the proportion of firms that are decentralized into profit centers in the relevant decile group. The sample is the COI sample (3,570 French firms in 1997).

Figure 2: Proximity to Frontier and Decentralization

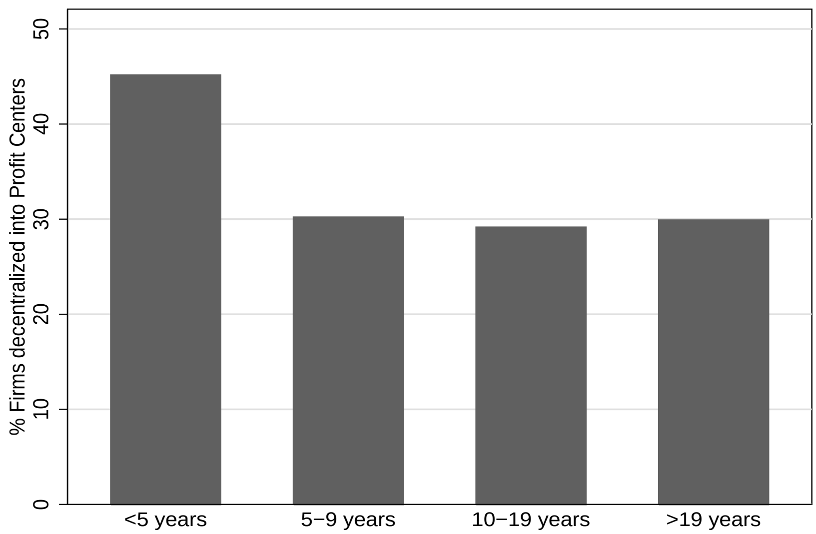
Decentralization to Profit Centers (COI)



NOTES: The X-axis divides all firms into deciles of proximity to frontier (in terms of value added per hour) from the first decile (low proximity to frontier) to the tenth decile (high proximity to frontier). The Y-axis indicates the proportion of firms that are decentralized into profit centers in the relevant decile group. The sample is the COI sample (3,570 French firms in 1997).

Figure 3: Age and Decentralization

Decentralization to Profit Centers (COI)



NOTES: Firms are grouped into age bands (dated from the birth of the firm). The Y-axis indicates the proportion of firms that are decentralized into profit centers in the relevant age group. The sample is the COI sample (3,570 French firms in 1997).

Table 1: Descriptive Statistics
(Enquête COI, 1994-1997)

Variable	Source	Mean	Median	St dev
Decentralization measures				
Organized into profit centers	COI	0.304	0	0.216
Decentralization of investment decisions	ER	0.484	0	0.500
Delayering	ER	0.436	0	0.496
Heterogeneity measures: Industry level heterogeneity of labour productivity growth				
$90^{\text{th}} - 10^{\text{th}}$ percentiles	DADS/FUTE	0.275	0.263	0.087
$95^{\text{th}} - 5^{\text{th}}$ percentiles	DADS/FUTE	0.443	0.406	0.160
Standard deviation	DADS/FUTE	0.177	0.165	0.066
Standard deviation after trimming	DADS/FUTE	0.088	0.082	0.033
Heterogeneity measures: Industry level heterogeneity of labour productivity level				
$90^{\text{th}} - 10^{\text{th}}$ percentiles	DADS/FUTE	0.897	0.861	0.229
Heterogeneity measure: Firm level inverse share of close firms in the product space				
Share of close (IT weighted) firms, %	FUTE	0.343	0.138	0.049
log Firm specific heterogeneity		7.111	6.587	2.381
Share of close (unweighted) firms, %	FUTE	0.216	0.096	0.308
Distance to technological frontier				
Firm labour productivity	DADS/FUTE	0.163	0.143	0.089
Sectoral 99^{th} perc. labour productivity	DADS/FUTE	0.508	0.397	0.315
Proximity to frontier	DADS/FUTE	0.358	0.334	0.159
log Proximity to frontier		-1.125	-1.096	0.457
Other firm level variables				
Foreign ownership	LIFI	0.173	0	0.143
Firms belonging to a larger group	LIFI	0.502	1	0.500
Number of plants	DADS	3.092	1	8.510
Firm age	SIRENE	21.658	18	12.740
Capital intensity (/value added)	FUTE	1.143	0.907	1.036
Number of workers	FUTE	323.463	88.375	677.080
% workers working with computers	COI	59.669	71.846	26.300
Unskilled workers	DADS	27.004	22.623	20.202
Age of workers	DADS	38.870	39.010	3.403
Lerner index	FUTE	0.075	0.068	0.077
Market share	FUTE	1.732	0.404	4.171
Herfindahl index	FUTE	0.049	0.031	0.057
Specialization	FUTE	0.831	0.931	0.203
Other industry level variables				
Capital intensity (per worker)	BRN	404.987	289.242	369.064
IT investment (per worker)	EAE/FUTE	0.849	0.600	0.725

NOTES: These descriptive statistics are based on the COI sample (3,570 observations), except for "Decentralization of investment decisions" and "Delayering" (ER sample: 1,258 observations). The COI dataset is a firm level survey providing information on organization and other firm characteristics in 1997; it covers manufacturing sectors only. The ER dataset is an establishment survey containing information about organizational change between 1996 and 1998; it covers both manufacturing and non-manufacturing sectors. The FUTE files contain the firms balance sheet and further accounting information; it refers to the entire population of French firms having more than 20 employees in manufacturing industries and more than 10 employees in other industries. The DADS files consist of yearly mandatory employer reports of each worker's hours (and gross earnings subject to payroll taxes); they cover the entire population of French firms. The LIFI files describe the structure of ownership of large French firms; it also includes information about their main interests in other companies. SIRENE is the index of all French firms (and administrations). Units of currency are thousands of French Francs at the 1995 prices. See text for variable definitions.

Table 2: Determinants of Decentralization
(Enquête COI)

Dependent variable	Firm decentralized into Profit Centers					
Industries	Full sample				High-tech	Low-tech
	(1)	(2)	(3)	(4)	(5)	(6)
Heterogeneity	0.211	0.252	0.296	0.251	0.679	-0.062
	(0.107)	(0.102)	(0.127)	(0.115)	(0.189)	(0.135)
Frontier, 99th percentile	-0.101	-	-0.225	-	-	-
($\ln y_{Fl}$)	(0.039)		(0.045)			
Labour productivity, firm level	0.182	-	0.141	-	-	-
($\ln y_{il}$)	(0.026)		(0.033)			
Proximity to frontier	-	0.167	-	0.164	0.224	0.103
(constrained term $\ln y_{il} - \ln y_{Fl}$)		(0.024)		(0.028)	(0.040)	(0.044)
Firm age<5 years	0.151	0.151	0.172	0.174	0.215	0.122
	(0.040)	(0.040)	(0.041)	(0.041)	(0.059)	(0.055)
5≤ Firm age<10 years	0.012	0.012	0.066	0.066	0.069	0.049
	(0.021)	(0.021)	(0.022)	(0.022)	(0.032)	(0.029)
10≤Firm age<20 years	-0.007	-0.007	0.039	0.040	-0.008	0.083
	(0.019)	(0.019)	(0.019)	(0.019)	(0.027)	(0.028)
Other firm and industry controls (cont.)	no	no	yes	yes	yes	yes
Industry dummies	yes (73)	yes (73)	yes (73)	yes (73)	yes (52)	yes (42)
Mean of dependent variable	0.304	0.304	0.304	0.304	0.378	0.232
Observations	3,570	3,570	3,570	3,570	1,767	1,803

NOTES: All coefficients are marginal effects from probit maximum likelihood estimation. Robust standard errors corrected for arbitrary variance-covariance matrix at the four-digit industry level in parentheses. Industry variables are defined at the four-digit level (except industry dummies at the three-digit level). All right hand side variables are lagged and averaged over three years (1994-1997). Labor productivity is the log of value added per hour, frontier is defined as the 99th percentile of the productivity distribution in the firm's four-digit industry. Heterogeneity is defined as the dispersion of productivity growth rate within a four digit industry (the 90th percentile less the 10th percentile). The omitted category for firm age is "firm age greater than twenty years". High-tech subsample includes all firms in industries with greater than median IT investment per worker. Low-tech subsample includes all firms in industries with less than median IT investment per worker. See text for variable definitions.

Table 2: (-cont.) Determinants of Decentralization
(Enquête COI)

Dependent variable	Firm decentralized into Profit Centers					
Industries	Full sample				High-tech	Low-tech
	(1)	(2)	(3)	(4)	(5)	(6)
Lerner index	-	-	-0.660	-0.733	-0.947	-0.547
			(0.144)	(0.136)	(0.168)	(0.218)
log Number of plants	-	-	0.041	0.041	0.027	0.052
			(0.015)	(0.015)	(0.021)	(0.017)
log Firm size	-	-	0.110	0.110	0.134	0.098
			(0.009)	(0.009)	(0.015)	(0.014)
% Workers working with computers	-	-	0.220	0.217	0.238	0.189
			(0.034)	(0.034)	(0.071)	(0.034)
Firm (high) skill %	-	-	0.169	0.153	0.206	0.090
			(0.049)	(0.048)	(0.078)	(0.059)
Firm log capital / value added	-	-	0.008	0.009	-0.003	0.022
			(0.012)	(0.012)	(0.017)	(0.016)
Foreign ownership	-	-	0.047	0.045	0.064	0.015
			(0.025)	(0.025)	(0.038)	(0.032)
Firm age of workers /10	-	-	-0.057	-0.060	-0.155	0.008
			(0.025)	(0.025)	(0.046)	(0.027)
Firm market share	-	-	-0.577	-0.574	-0.821	-0.520
			(0.260)	(0.257)	(0.382)	(0.419)
log Herfindahl index	-	-	-0.015	-0.011	0.031	-0.024
			(0.017)	(0.017)	(0.027)	(0.022)
log Firm specialization	-	-	-0.071	-0.070	-0.119	-0.027
			(0.030)	(0.030)	(0.047)	(0.039)
Sector log capital stock	-	-	-0.064	-0.074	-0.115	-0.033
per worker			(0.040)	(0.040)	(0.035)	(0.043)
Sector log IT investment	-	-	0.116	0.102	0.059	0.111
per worker			(0.019)	(0.019)	(0.044)	(0.029)
Industry dummies	yes (73)	yes (73)	yes (73)	yes (73)	yes (52)	yes (42)
Mean of dependent variable	0.304	0.304	0.304	0.304	0.378	0.232
Observations	3,570	3,570	3,570	3,570	1,767	1,803

Table 3: Determinants of Decentralization (Alternative Measures of Heterogeneity)
(Enquête COI)

Dependent variable (mean=0.304)	Firm decentralized into Profit Centers										
Measure of heterogeneity	Heterogeneity: dispersion of productivity *growth* in four-digit industry				Heterogeneity: dispersion of productivity *levels* in four-digit industry			Heterogeneity: log of the inverse share of "close" firms in the product space (/10)			
	$90^{th} - 10^{th}$ percentiles	$95^{th} - 5^{th}$ percentiles	Standard deviation	Std dev after trimming	$90^{th} - 10^{th}$ percentiles			IT weighted			Un-weighted
Industry	Full sample				Full sample	High-tech	Low-tech	Full sample	High-tech	Low-tech	Full sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Heterogeneity	0.251	0.142	0.231	0.870	0.090	0.271	-0.064	0.063	0.098	0.019	0.055
	(0.115)	(0.069)	(0.132)	(0.316)	(0.062)	(0.092)	(0.090)	(0.031)	(0.048)	(0.037)	(0.035)
Proximity to frontier	0.164	0.168	0.152	0.165	0.153	0.222	0.100	0.159	0.208	0.104	0.146
	(0.028)	(0.029)	(0.028)	(0.028)	(0.028)	(0.041)	(0.044)	(0.028)	(0.039)	(0.043)	(0.027)
Firm age<5 years	0.174	0.173	0.171	0.173	0.171	0.211	0.121	0.177	0.214	0.123	0.173
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.059)	(0.056)	(0.041)	(0.060)	(0.056)	(0.041)
5≤ Firm age<10 years	0.066	0.067	0.067	0.066	0.068	0.067	0.049	0.067	0.068	0.049	0.067
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.032)	(0.029)	(0.022)	(0.032)	(0.029)	(0.022)
10≤Firm age<20 years	0.040	0.041	0.040	0.040	0.039	-0.009	0.083	0.041	-0.005	0.082	0.040
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.027)	(0.028)	(0.019)	(0.027)	(0.028)	(0.019)
Lerner index	-0.733	-0.750	-0.708	-0.743	-0.725	-0.952	-0.534	-0.721	-0.895	-0.544	-0.697
	(0.136)	(0.137)	(0.134)	(0.136)	(0.132)	(0.172)	(0.217)	(0.136)	(0.168)	(0.218)	(0.133)
Industry dummies	yes (73)	yes (73)	yes (73)	yes (73)	yes (73)	yes (52)	yes (42)	yes (73)	yes (52)	yes (42)	yes (73)
Observations	3,570	3,570	3,570	3,570	3,570	1,767	1,803	3,570	1,767	1,803	3,570

NOTES: All coefficients are marginal effects from probit maximum likelihood estimation. Robust standard errors corrected for arbitrary variance-covariance matrix at the four-digit industry level in parentheses. In column (1) the measure of heterogeneity is the difference between the productivity *growth rate* for the firm at the 90^{th} percentile less the 10^{th} percentile; in column (2) it is the 95^{th} percentile less the 5^{th} percentile; in column (3) it is the standard deviation of productivity growth rates and in column (4) it is the same as column (2) except we trim the bottom and top 5% of productivity growth distribution in each four-digit industry. In columns (5) to (7) the measure of heterogeneity is the difference between the productivity *level* for the firm at the 90^{th} percentile less the 10^{th} percentile. In columns (8) to (10), the heterogeneity measure is the log of the inverse of the share of IT weighted "close" firms in the product space; whereas in column 11) these firms are not IT weighted. Full set of firm and industry level controls included; see text for variable definitions.

Table 4: Determinants of Decentralization (Alternative Measures of Decentralization)
(Enquête Reponse)

Dependent variable	Decentralization of investment decisions, 1998			Some delayering between 1996 and 1998					
Specification of proximity to frontier	Level			Level			Growth rate		
Industries	Full sample (1)	High - tech (2)	Low- tech (3)	Full sample (4)	High - tech (5)	Low- tech (6)	Full sample (7)	High - tech (8)	Low- tech (9)
Heterogeneity	0.108	0.455	-0.127	0.355	0.398	0.075	0.434	0.396	0.102
	(0.165)	(0.235)	(0.415)	(0.185)	(0.283)	(0.301)	(0.200)	(0.304)	(0.294)
Frontier (99^{th} percentile)	-0.110	-0.249	0.021	-0.115	-0.152	0.011	-0.058	-0.078	-0.043
	(0.054)	(0.079)	(0.103)	(0.051)	(0.080)	(0.086)	(0.048)	(0.064)	(0.067)
Labour productivity (firm)	0.072	0.064	0.104	0.126	0.118	0.066	-0.056	0.164	-0.137
	(0.052)	(0.077)	(0.075)	(0.056)	(0.072)	(0.081)	(0.101)	(0.157)	(0.108)
Firm age$<$5 years	-0.036	-0.016	-0.027	0.183	0.099	0.275	0.054	0.102	0.017
	(0.056)	(0.080)	(0.092)	(0.088)	(0.113)	(0.129)	(0.107)	(0.171)	(0.114)
5$\leq$Firm age$<$10 years	-0.007	-0.044	0.022	0.047	-0.007	0.076	0.034	-0.031	0.095
	(0.048)	(0.074)	(0.061)	(0.056)	(0.064)	(0.084)	(0.059)	(0.068)	(0.090)
10$\leq$Firm age$<$20 years	0.004	-0.015	0.026	0.069	0.059	0.068	0.072	0.056	0.099
	(0.040)	(0.059)	(0.054)	(0.039)	(0.054)	(0.055)	(0.040)	(0.053)	(0.054)
Lerner index	-0.034	-0.046	-0.029	-0.934	-0.166	-0.981	-0.666	0.048	-0.977
	(0.009)	(0.024)	(0.114)	(0.288)	(0.416)	(0.402)	(0.257)	(0.036)	(0.359)
Industry dummies	yes (81)	yes (49)	yes (31)	yes (70)	yes (47)	yes (43)	yes (69)	yes (47)	yes (42)
Mean of dependent variable	0.484	0.475	0.493	0.440	0.553	0.325	0.436	0.553	0.320
Observations	1,258	648	610	1,049	526	523	1,011	505	501

NOTES: All coefficients are marginal effects from probit maximum likelihood estimation. Robust standard errors corrected for arbitrary variance-covariance matrix at the four-digit industry level in parentheses. All establishments are part of a large (French or foreign) group but are not Head Quarters. Heterogeneity is defined as the dispersion of productivity growth rates within a four digit industry (the 90^{th} percentile less the 10^{th} percentile). Full set of firm and industry level controls included; see text for variable definitions.

Table 5: Determinants of Decentralization in Britain
(British WERS98)

Dependent variable (mean=0.805)	Decentralization of employment decisions					
	(1)	(2)	(3)	(4)	(5)	(6)
Heterogeneity	0.273	-	-	-	0.316	-
($95^{th}-5^{th}$ percentiles)	(0.130)				(0.129)	
Heterogeneity	-	0.540	-	-	-	0.659
($90^{th}-10^{th}$ percentiles)		(0.325)				(0.312)
Frontier	-	-	-0.051	-	-0.204	-0.156
(99^{th} percentile)			(0.073)		(0.073)	(0.073)
Establishment age<5 years	-	-	-	-0.076	-0.099	-0.123
				(0.115)	(0.116)	(0.114)
5≤Est. age<10 years	-	-	-	0.086	0.055	0.049
				(0.081)	(0.089)	(0.089)
10≤Est. age<20 years	-	-	-	-0.127	-0.164	-0.173
				(0.077)	(0.076)	(0.075)
Many competitors	-	-	-	-	0.127	0.150
					(0.082)	(0.078)
Few competitors	-	-	-	-	0.210	0.228
					(0.070)	(0.065)
No competitors	-	-	-	-	*ref*	*ref*
Other firm and industry controls	no	no	no	no	yes	yes
Industry dummies	yes (64)	yes (64)	yes (64)	yes (64)	yes (64)	yes (64)
Observations	236	236	236	236	236	236

NOTES: All coefficients are marginal effects from probit maximum likelihood estimation. Robust standard errors corrected for arbitrary variance-covariance matrix at the four-digit industry level in parentheses. Data are from the 1998 British Workplace Employee Relations Survey (WERS); they include manufacturing establishments only. Dependent variable is a dummy variable indicating whether "Establishment's manager is able to make decisions on which staff to recruit without consulting Head Office". Heterogeneity and frontier are averaged between 1994 and 1997. All regressions include a control for employment size (current, lagged one year and lagged five years), the proportion of young workers (under 20 years old), the proportion of older workers (aged over 50 years old), the proportion of unskilled manual workers and the proportion of part-time workers. See text for variable definitions.

APPENDICES NOT FOR PUBLICATION FOR

Daron Acemoglu, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti

"Technology, Information and the Decentralization of the Firm"

Appendix A: Theory

In this appendix, we provide a more detailed outline of the model presented in this theory section, and present proofs of the claims and propositions stated there.

The economy is assumed to be populated by a set $\mathcal{F}$ of firms. Time is discrete. New technologies $k = 1, 2, ...$ become available sequentially and randomly to firms. As a new opportunity of technology adoption materializes, firms must choose the form of its implementation. In each period, one (and only one) new technology becomes available to the firm $i \in \mathcal{F}$ with positive probability $p_i \in (0, 1]$. The speed at which new opportunities become available can differ across firms, and p_i measures the speed at which firm i climbs the technology ladder. The realizations of technological opportunities are independent across firms and over time.

Each technology can be implemented successfully or unsuccessfully. The successful implementation of a technology increases the firm's productivity by a factor $\gamma > 1$, while unsuccessful implementation leaves the productivity of the firm unchanged. So the law of motion of firm i's productivity is given by

$$y_{i,t} = \gamma^{S_i(t)} y_{i,t-1}, \tag{8}$$

where $y_{i,t}$ is the productivity (and revenue) of the firm at time t and $S_i(t)$ is an indicator function taking the value 1 if a technological opportunity arises for firm i at time t and is successfully implemented, and zero otherwise. Whether the technology is successful or not depends on an action taken by the firm, which we denote by $x_{i,k,t} \in \{L, R, \varnothing\}$, with $x_{i,k,t} = \varnothing$ standing for not attempting the technology, and L and R denoting two alternative ways of implementing the new technology, or two "actions". We will see below that the firm will never choose $x_{i,k,t} = \varnothing$, so the relevant choice is between L and R. One of these two actions, denoted by $x^*_{i,k} \in \{L, R\}$, leads to successful implementation, while the other leads to an unsuccessful outcome. As stated in the text, we refer to the action leading to successful implementation as the *correct action*, and assume that it is given by (1) as specified in the text. We also assume that conditional on x^*_k, $x^*_{i,k}$ and $x^*_{i',k}$ for any $i \neq i'$ are independent, or stated alternatively, the sequence $\left\{x^*_{i,k}\right\}_{i\in\mathcal{F}}$ is exchangeable in the sense that any permutation of this sequence is equally likely. This implies that $\left\{x^*_{i,k}\right\}_{i\in\mathcal{F}}$ is a Bernoulli sequence with a parameter of $1-\varepsilon$ or ε (depending on whether $x^*_k = L$ or R). We also assume that

$$\Pr\left(x^*_k = L\right) = q_0 = \frac{1}{2}.$$

As discussed in the text, each firm is owned by a principal, who maximizes the present discounted value (PDV) of profits, with discount factor $\beta \in (0, 1)$, conditional on information, h^t, and their initial productivity, $y_{i,t-1} < \infty$. Let us first ignore any costs on the side of firms, so profits are equal to revenue (productivity), and therefore the objective function of firm i is given by

$$V\left(y_{i,t-1}, h^t\right) = \mathbb{E}\left[\sum_{j=0}^{\infty} \beta^j y_{i,t+j} \mid h^t\right]. \tag{9}$$

We assume that $\beta < \gamma^{-1}$ so that the firm's value $V\left(y_{i,t-1}, h^t\right)$ always remains finite.

The above specification makes it clear that it is always advantageous for the firm to implement a technology whenever it becomes available, so we can restrict attention to $x_{i,k,t} \in \{L, R\}$. This is because, by delaying the implementation of technology k available at date t, the firm may obtain more information about this technology, but by assumption, it will never have an opportunity to adopt this technology again. Instead, at $t+1$, technology $k+1$ would become available to firm i with probability p_i. Consequently, a firm's productivity and value will depend on how many technologies have become available in the past to this firm and how successful the firm has been in implementing them. We can also note that none of the results will be affected if we allow firms to implement past unsuccessful new technologies again in the future. In this case, the cost of unsuccessful implementation will be simply the delay in realizing the increase in productivity.

Success in the implementation of technologies depends on the organization of the firm. The two alternative organizational forms available to each firm are: centralization and delegation. In the first, the principal (owner) manages the firm and chooses the action $x_{i,k}$, while in the second, the action choice is delegated to a manager.

The principal has no special skills in identifying the right action, so she bases her decision on the publicly available information h^t. Without loss of any generality, for the decision regarding technology k, we can restrict attention to the history that is payoff relevant to technology k. Moreover, throughout we will focus on a (representative) firm i that has access to technology k at time t. In this case, with a slight abuse of terminology, we will refer to the payoff-relevant history for the implementation of technology k as a *history* h_k^i. As discussed further below, under centralization the principal will choose $x_{i,t,k} = L$ when the posterior that $x_k^* = L$ given history h_k^i,

$$q\left(h_k^i\right) = \Pr\left(x_k^* = L \mid h_k^i\right)$$

is greater than 1/2.

Recall that $\pi\left(d_{i,k}; h_k^i\right)$ is the posterior probability that the firm chooses the correct action conditional on the history h_k^i and the organizational form $d_{i,k}$. We start our analysis by comparing $\pi\left(d_{i,k}; h_k^i\right)$ under the two organizational forms, delegation and centralization (Lemmas 1 and 2). Next we show that profit-maximizing firms always choose the organization form that maximizes $\pi\left(d_{i,k}; h_k^i\right)$ (Lemma 3). Finally, we turn to the main testable implications of the theory, linking distance to frontier (n_k^i) and heterogeneity (ε) to the probability that firms choose either delegation or centralization (Propositions 1 and 2).

As noted in the text, when the principal delegates the implementation of a new technology to an informed manager, the probability of success is constant and equal to δ, i.e.,

$$\pi\left(1; h_k^i\right) = \delta. \tag{10}$$

If, on the other hand, the principal retains authority, the probability of success is a stochastic variable that depends on h_k^i, thus both on the firm's distance to the frontier, n_k^i, and on the experiences of firms further ahead, captured by the number of firms $\tilde{n}_k^i$ out of n_k^i for whom L was the right action. The following lemma provides the expression for $\pi\left(0; h_k^i\right)$.

Lemma 1 Given a history h_k^i, the probability of success for a principal who retains authority is

$$\pi\left(0; h_k^i\right) = \begin{cases} \varepsilon + q\left(h_k^i\right)(1-2\varepsilon) & \text{if } q\left(h_k^i\right) \geq 1/2 \\ \varepsilon + \left(1 - q\left(h_k^i\right)\right)(1-2\varepsilon) & \text{if } q\left(h_k^i\right) < 1/2. \end{cases} \tag{11}$$

where

$$q\left(h_k^i\right) = \frac{(1-\varepsilon)^{\tilde{n}_k^i} \varepsilon^{n_k^i - \tilde{n}_k^i}}{(1-\varepsilon)^{\tilde{n}_k^i} \varepsilon^{n_k^i - \tilde{n}_k^i} + (1-\varepsilon)^{n_k^i - \tilde{n}_k^i} \varepsilon^{\tilde{n}_k^i}} = \frac{1}{1 + \left(\frac{1-\varepsilon}{\varepsilon}\right)^{n_k^i - 2\tilde{n}_k^i}} \tag{12}$$

is the posterior probability that the right choice is L.

Proof. The probability of success for a firm choosing centralization, conditional on action $x_{i,k}$ is:

$$\pi\left(0;h_k^i \mid x_{i,k}=L\right) = q\left(h_k^i\right)(1-\varepsilon)+(1-q\left(h_k^i\right))\varepsilon, \quad (13)$$
$$\pi\left(0;h_k^i \mid x_{i,k}=R\right) = (1-q\left(h_k^i\right))(1-\varepsilon)+q\left(h_k^i\right)\varepsilon, \quad (14)$$

where $q\left(h_k^i\right)$ is given by (12). In equation (13), $q\left(h_k^i\right)(1-\varepsilon)$ is the posterior probability that the reference action is L and that the correct action for the firm coincides with the reference action, whereas $(1-q\left(h_k^i\right))\varepsilon$ is the posterior probability that the reference action is R and that the correct action for the firm differs from the reference action. Equation (14) has a similar form, with $(1-q\left(h_k^i\right))(1-\varepsilon)$ as the posterior probability that the reference action is R and that the correct action for the firm coincides with the reference action and $q\left(h_k^i\right)\varepsilon$ as the posterior probability that the reference action is R and that the correct action for the firm differs from the reference action.

Thus

$$\begin{aligned}\pi\left(0;h_k^i\right) &= \max\left\langle \pi\left(0,h_k^i \mid x_{i,k}=L\right), \pi\left(0,h_k^i \mid x_{i,k}=R\right)\right\rangle \\ &= \begin{cases} \varepsilon+q\left(h_k^i\right)(1-2\varepsilon) & \text{if } q\left(h_k^i\right)\geq 1/2 \\ \varepsilon+\left(1-q\left(h_k^i\right)\right)(1-2\varepsilon) & \text{if } q\left(h_k^i\right)<1/2. \end{cases},\end{aligned}$$

establishing the result. ■

To understand the expression of $\pi\left(0;h_k^i\right)$, note that the principal chooses L if $q\left(h_k^i\right)\geq 1/2$ and R if $q\left(h_k^i\right)\leq 1/2$. The first row of (11) gives then the probability that, when $q\left(h_k^i\right)\geq 1/2$, the firm correctly infers that the reference action is L *and* the reference action coincides with the correct action. The second row gives the probability that, when $q\left(h_k^i\right)<1/2$, the firm incorrectly infers that the reference action is R *and* the reference action does not coincide with the correct action (hence, due to a double mistake, the firms makes the successful adoption).

The expression of $q\left(h_k^i\right)$ in (12) follows from the Bayes' rule, given the Bernoulli assumption and $q_0=1/2$. It may be useful to note that since $\varepsilon\in(0,1/2)$, $q\left(h_k^i\right)>1/2$ whenever $\tilde{n}_k^i>n_k^i/2$, and $q\left(h_k^i\right)<1/2$ whenever $\tilde{n}_k^i<n_k^i/2$. More importantly, $q\left(h_k^i\right)$ is a random variable that depends on the realization of the stochastic vector h_k^i. Since h_k^i consists of a deterministic component, n_k^i, and a stochastic one, $\tilde{n}_k^i$, we can determine the first moment of $q\left(h_k^i\right)$ conditional on the sample size n_k^i, $E\left(q_k^i|n_k^i\right)$, where expectation is taken over possible realizations of $\tilde{n}_k^i$.

The following lemma establishes how this conditional expectation changes with n_k^i and the limiting behavior of $q\left(h_k^i\right)$. In the rest of the analysis, without loss of any generality, we assume that $x_k^*=L$.

Lemma 2 *Let $E\left(q_k^i|n_k^i\right)$ denote the expectation of $q\left(h_k^i\right)$ conditional on sample size n_k^i and suppose that $x_k^*=L$. Then $E\left(q_k^i|n_k^i\right)$ is an increasing function of n_k^i. Moreover, $p\lim_{n_k^i\longrightarrow\infty} q\left(h_k^i=\left\{n_k^i,\tilde{n}_k^i\right\}\right)=1$.*

Proof. (1) To see that $E\left(q_k^i|n_k^i\right)$ is increasing in n_k^i, let $(q_n)_k^i$ denote the posterior probability of firm i that $x_k^*=L$ conditional on a history of length n, i.e., $\left(q_{n_k^i}\right)_k^i\equiv E\left(q_k^i|n_k^i\right)$. We need to prove that $(q_n)_k^i$ is increasing in n when $x_k^*=L$. First note that $(q_n)_k^i$ is a sufficient statistic for the history h_k^i. Thus, by the Law of Iterated Expectations, we have

$$(q_{n+1})_k^i = E\left((q_{n+1})_k^i \mid (q_n)_k^i\right) > (q_n)_k^i\,.$$

Therefore, it is sufficient to prove that

$$E\left((q_{n+1})_k^i \mid (q_n)_k^i\right) > (q_n)_k^i\,.$$

Consider a firm with a posterior $(q_n)_k^i$ conditional on a sample of n and suppose that it obtains an additional observation, i.e., observes the realization of a $n+1^{st}$ firm. Since, given $x_k^* = L$, L and R will be revealed to be the right choice for the $n+1^{st}$ firm with respective probabilities $1-\varepsilon$ and ε, Bayes' rule implies that

$$(q_{n+1})_k^i = \begin{cases} \frac{(q_n)_k^i \times (1-\varepsilon)}{(q_n)_k^i \times (1-\varepsilon) + \left(1-(q_n)_k^i\right) \times \varepsilon} & \text{with prob. } 1-\varepsilon \\ \\ \frac{(q_n)_k^i \times \varepsilon}{\left(1-(q_n)_k^i\right) \times (1-\varepsilon) + (q_n)_k^i \times \varepsilon} & \text{with prob. } \varepsilon \end{cases}$$

Therefore,

$$\begin{aligned} E\left((q_{n+1})_k^i \mid (q_n)_k^i\right) &= (1-\varepsilon) \times \frac{(q_n)_k^i \times (1-\varepsilon)}{(q_n)_k^i \times (1-\varepsilon) + \left(1-(q_n)_k^i\right) \times \varepsilon} \\ &\quad + \varepsilon \times \frac{(q_n)_k^i \times \varepsilon}{\left(1-(q_n)_k^i\right) \times (1-\varepsilon) + (q_n)_k^i \times \varepsilon}. \end{aligned}$$

This implies that

$$\begin{aligned} E\left((q_{n+1})_k^i \mid (q_n)_k^i\right) & - (q_n)_k^i \\ &= \frac{\left(1-(q_n)_k^i\right)^2 (1-2\varepsilon)^2 \times (q_n)_k^i}{\left(\varepsilon \times (q_n)_k^i + \left(1-(q_n)_k^i\right) \times (1-\varepsilon)\right)\left((q_n)_k^i \times (1-\varepsilon) + \varepsilon \times \left(1-(q_n)_k^i\right)\right)} \\ &> 0, \end{aligned}$$

establishing the desired result.

(2) Next, to prove $p\lim_{n_k^i \longrightarrow \infty} q_{i,k}\left(h_k^i\right) = 1$ note that from the viewpoint of $t = 0$, $q\left(h_k^i\right)$ is a random variable, since the history h_k^i is a random vector. We need to show that for almost all histories $q\left(h_k^i\right)$ will become arbitrarily close to 1. We will do this by using the Continuous Mapping Theorem (e.g., van der Vaart, 1998, Theorem 2.3). First, when $x_k^* = L$, by the strong law of large numbers $p\lim_{n_k^i \to \infty} \tilde{n}_k^i / n_k^i = 1-\varepsilon$. The Continuous Mapping Theorem then implies that any continuous function, $G\left(\tilde{n}_k^i, n_k^i\right)$, converges in probability to $\lim_{n_k^i \to \infty} G\left((1-\varepsilon)\, n_k^i, n_k^i\right)$. (12) is such a function, so

$$p \lim_{n_k^i \longrightarrow \infty} q_{i,k}\left(h_k^i\right) = \lim_{n_k^i \longrightarrow \infty} \frac{1}{1 + \left(\frac{\varepsilon}{1-\varepsilon}\right)^{n_k^i[2(1-\varepsilon)-1]}} = 1,$$

where the last equality follows from $\varepsilon < 1/2$. ■

Lemma 2 establishes the intuitive result that, as the history relevant to technology k expands, the principal learns the reference action x_k^* with increasing precision.

Next, we establish that profit-maximizing firms always choose the organization form which maximizes $\pi\left(d_{i,k}; h_k^i\right)$. Although firms have a dynamic objective function, given by (9), the maximization program is equivalent to a sequence of static problems. Intuitively, the current organization choice only affects the PDV of the firm via its effect on current productivity, $y_{i,t}$, so the optimal strategy simply maximizes the probability of successful implementation of new technologies in each period.

Lemma 3 A firm i maximizing (9) chooses, for all technologies k,

$$d_{i,k}^*\left(h_k^i\right) \in \{0,1\} = \arg\max_{d_{i,k}} \pi\left(d_{i,k}; h_k^i\right).$$

In particular, given history h_k^i, firm i will choose $d_{i,k}^*\left(h_k^i\right) = 1$ (delegation) if $\pi\left(0; h_k^i\right) < \delta$, and $d_{i,k}^*\left(h_k^i\right) = 0$ (centralization) if $\pi\left(0; h_k^i\right) > \delta$, where $\pi\left(0; h_k^i\right)$ is given by (11).

Proof. First, we note that the value of the firm admits a recursive representation, so that for $s^* \in \mathcal{S}$ that maximizes (9) with starting productivity y_i and corresponding history h_k^i, let this maximum value be $V_i\left(y_i, h_k^i\right)$. Then

$$V_i\left(y_i, h_k^i\right) = \max_{d_{i,k}\in\{0,1\}} \mathbb{E}\left[V_i\left(y_i, h_k^i \mid d_{i,k}\right) \mid h_k^i\right], \text{ and} \tag{15}$$

$$\begin{aligned} V_i\left(y_i, h_k^i \mid d_{i,k}\right) &= p_i\pi\left(d_{i,k}; h_k^i\right)\left(\gamma y_i + \beta\mathbb{E}\left[V_i\left(\gamma y_i, h_{k+1}^{i+1}\right) \mid h_k^i\right]\right) \\ &\quad + \left(1 - p_i\pi\left(d_{i,k}; h_k^i\right)\right)\left(y_i + \beta\mathbb{E}\left[V_i\left(y_i, h_{k+1}^i\right) \mid h_k^i\right]\right). \end{aligned}$$

In particular, with probability p_i, the opportunity to implement the next technology arrives and it is successfully implemented with probability $\pi\left(d_{i,k}; h_k^i\right)$. If it is successfully implemented, y_i increases to γy_i, and otherwise it stays at y_i. Future probabilities of opportunities and success are unaffected by these choices or realizations. The resulting optimal policy $d_{i,k}^*$ in this recursive representation is the organizational form induced by s^*.

To characterize the form of the value function in (15) and the optimal policy $d_{i,k}^*$, we guess and verify that $V_i\left(y_i, h_k^i \mid d_{i,k}\right)$ takes a linear form, namely

$$V_i\left(y_i, h_k^i \mid d_{i,k}\right) = A_i\left(h_k^i \mid d_{i,k}\right) y_i.$$

We then substitute the guess into the value function so as to solve for the unknown coefficient $A_i\left(h_k^i \mid d_{i,k}\right)$. This yields the following recursive equation for A_i:

$$\begin{aligned} A_i\left(h_k^i \mid d_{i,k}\right) &= p_i\pi\left(d_{i,k}; h_k^i\right)\gamma + \left(1 - p_i\pi\left(d_{i,k}; h_k^i\right)\right) + \\ &\quad \beta\left(p_i\pi\left(d_{i,k}; h_k^i\right)\gamma + \left(1 - p_i\pi\left(d_{i,k}; h_k^i\right)\right)\right) \max_{d_{i,k+1}\in\{0,1\}} \mathbb{E}\left[A_i\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right]. \end{aligned} \tag{16}$$

The optimal organizational form is simply that which maximizes $A_i\left(h_k^i \mid d_{i,k}\right)$. This is equivalent, in turn, to choose organization so as to maximize $\pi\left(d_{i,k}; h_k^i\right)$ (note that the term $\mathbb{E}\left[A\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right]$ is outside of the firm's control). Namely, the firm will choose organizational form $d_{i,k} = 1$ if $\pi\left(1; h_k^i\right) > \pi\left(0; h_k^i\right)$, and $d_{i,k} = 0$ otherwise. Since the program is recursive, we have established that choosing organization period-by-period so as to maximize $\pi\left(d_{i,k}; h_k^i\right)$ maximizes the value of the firm.

The second part of the Lemma follows immediately from the previous analysis. ■

Lemmas 1, 2 and 3 enable us to characterize how the organizational form changes with distance to frontier, n_k^i, and heterogeneity, ε. Consider, first, distance to frontier. Suppose, in particular, that a firm is at the technology frontier, so that it is the first firm implementing technology k. In this case, $n_k^i = 0$, and so $q\left(h_k^i\right) = 1/2$ and $\pi\left(0; h_k^i\right) = q_0 = 1/2 < \delta$. Thus, when no public information is available, the firm will choose delegation, $d_{i,k}^*\left(h_k^i\right) = 1$. Next, consider the other extreme where the firm is far behind the technology frontier, so that many other firms have implemented the same technology before. In this case, we have $n_k^i \to \infty$ and $p\lim_{n_k^i \longrightarrow \infty} q\left(h_k^i\right) = 1$ from Lemma 2. Equation (11) then implies $p\lim_{n_k^i \longrightarrow \infty} \pi\left(0, h_k^i\right) = 1 - \varepsilon > \delta$, where the inequality follows from the fact that $\delta \in (1/2, 1-\varepsilon)$. Hence, the firm will (almost surely) choose centralization, $d_{i,k}^*\left(h_k^i\right) = 0$. This discussion establishes the following proposition:

Proposition 1 (Distance to Frontier) Consider the adoption decision of technology k by firm i, and suppose that Assumption 1 holds and $x_k^* = L$. Then:
(i) For a firm at the frontier, i.e., $n_k^i = 0$, the principal chooses delegation, $d_{i,k}^*\left(h_k^i\right) = 1$.
(ii) For a firm sufficiently far from the frontier, i.e., $n_k^i \to \infty$, the principal chooses almost surely centralization, i.e., $p\lim_{n_k^i \longrightarrow \infty} d_{i,k}^*\left(h_k^i = \left\{n_k^i, \tilde{n}_k^i\right\}\right) = 0$.

From this proposition and the fact that, $E\left(q_k^i | n_k^i\right)$ is increasing in n_k^i (Lemma 2), one might expect a more general result, such that as distance to frontier n_k^i increases, decentralization becomes more likely. Unfortunately, though intuitive, this result is not correct because of integer issues, though it is true when integer issues are ignored, for example by smoothing the relationship between $E\left(q_k^i|n_k^i\right)$ and n_k^i. Therefore, in the empirical analysis, we disregard the integer problem and focus on the prediction that centralization increases with the distance to the frontier. We will proxy distance to the frontier with the gap between the productivity of a firm and the highest productivity (or more precisely the highest percentile productivity) in the same industry. It is clear that firms further from the frontier in terms of having high n_k^i's are less productive, since these are the firms that have been unlucky and have had fewer opportunities to adopt technologies, and they are also likely to be the ones with relatively low p_i's, that is, those that are slower in climbing their technology ladder.

We next turn to heterogeneity. Let us define $\Pr\left(d^*_{i,k}\left(h^i_t\right)=1\right)$ as the ex ante probability that firm i facing technology k at time t will choose decentralization. This probability is clearly a function of the parameters of the model, in particular ε, which measures the extent of heterogeneity, and the firm's distance to frontier. In particular, recall that greater ε translates into a greater heterogeneity in the firm's environment. The following proposition establishes that greater heterogeneity—higher ε—encourages decentralization (proof in Appendix A):

Proposition 2 (Heterogeneity) Consider the adoption decision of technology k by firm i. Given the distance to frontier, we have

$$\frac{\partial \Pr\left(d^*_{i,k}\left(h^i_t\right)=1\right)}{\partial\varepsilon}\geq 0,$$

so that an increase in ε makes delegation more likely.

This result is proved using the following three lemmas (which are themselves proved below).

Lemma 4 *For all $n_k^i\in\mathbb{N}$ and $\tilde{n}_k^i\in\mathbb{N}$ with $\tilde{n}_k^i\leq n_k^i$, we have*

$$\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)=\pi\left(0;\left(n_k^i,n_k^i-\tilde{n}_k^i\right)\right).$$

This lemma states that firms updates their beliefs *symmetrically* after signals suggesting either L or R to be the more likely correct action.

Lemma 5 *Either $\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)<\delta$ for all $\tilde{n}_k^i\in[0,n_k^i]$, or there exists a unique integer, $Q\left(\varepsilon,n_k^i\right)\in\mathbb{Z}_+$, such that*

$$\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)\geq\delta\Leftrightarrow\begin{cases}\text{either} & \tilde{n}_k^i\leq Q\left(\varepsilon,n_k^i\right),\\ \text{or} & \tilde{n}_k^i\geq n_k^i-Q\left(\varepsilon,n_k^i\right).\end{cases}$$

In the latter case, $Q\left(\varepsilon,n_k^i\right)\leq\left(n_k^i-1\right)/2$ and $Q\left(\varepsilon,n_k^i\right)$ is non-increasing in ε.

This lemma states that the posterior that the firm will choose the correct action will be greater than the threshold for decentralization, δ, if the number of successful L actions in the past are either smaller or greater than a specific threshold depending on the integer $Q\left(\varepsilon,n_k^i\right)$. Finally, we have the following technical result:

Lemma 6 *Let*

$$\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,\bar{Q}\right)\equiv\Pr\left[\tilde{n}_k^i\leq\bar{Q}\right]+\Pr\left[\tilde{n}_k^i\geq n_k^i-\bar{Q}\right]\tag{17}$$

Then,

$$\frac{\partial\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,\bar{Q}\right)}{\partial\bar{Q}}\geq 0,$$

and for any $\bar{Q}\leq\left(n_k^i-1\right)/2$,

$$\frac{\partial\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,\bar{Q}\right)}{\partial\varepsilon}\leq 0.$$

Lemmas 4, 5 and 6 prove Proposition 2. In particular, unless $\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)<\delta$ for all $\tilde{n}_k^i\in[0,n_k^i]$, we have that

$$\Pr\left(d_{i,k}^*\left(h_t^i\right)=1\right)=\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,Q\left(\varepsilon,n_k^i\right)\right).$$

By definition,

$$\frac{d\Pr\left(d_{i,k}^*\left(h_t^i\right)=1\right)}{d\varepsilon}=\left.\frac{\partial\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,Q\left(\varepsilon,n_k^i\right)\right)}{\partial\varepsilon}\right|_{Q\left(\varepsilon,n_k^i\right)=\bar{Q}}+\frac{\partial\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,Q\left(\varepsilon,n_k^i\right)\right)}{\partial Q\left(\varepsilon,n_k^i\right)}\frac{dQ\left(\varepsilon,n_k^i\right)}{d\varepsilon}.$$

Lemma 6 implies that the first term is non-positive and that $\partial\chi\left(n_k^i,\tilde{n}_k^i,\varepsilon,\bar{Q}\right)/\partial\bar{Q}\geq 0$, while Lemma 5 establishes that $dQ\left(\varepsilon,n_k^i\right)/d\varepsilon\leq 0$. Finally, if $\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)<\delta$ for all $\tilde{n}_k^i\in[0,n_k^i]$, a change in ε has no impact on $\Pr\left(d_{i,k}^*\left(h_t^i\right)=1\right)$, which is equal to 0. This establishes the proposition.

Proof of Lemma 4: The equality follows from the assumption that $q_0=1/2$. More formally, equations (11) and (12) imply that for all $\tilde{n}_k^i\leq\left(n_k^i-1\right)/2$,

$$\begin{aligned}\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)&=\varepsilon+\left(1-\frac{1}{1+\left(\frac{1-\varepsilon}{\varepsilon}\right)^{n_k^i-2\tilde{n}_k^i}}\right)(1-2\varepsilon)\\&=\varepsilon+\frac{1}{1+\left(\frac{1-\varepsilon}{\varepsilon}\right)^{-\left(n_k^i-2\tilde{n}_k^i\right)}}(1-2\varepsilon)=\pi\left(0;\left(n_k^i,n_k^i-\tilde{n}_k^i\right)\right).\end{aligned}$$

The same conclusion follows from applying (11) to the case in which $\tilde{n}_k^i>\left(n_k^i-1\right)/2$.■

Proof of Lemma 5: Suppose that there exists $\tilde{n}_k^i\in[0,n_k^i]$ such that $\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)\geq\delta$. Let $X\left(\varepsilon,n_k^i\right)\in\mathbb{R}_+$ be the unique value of X that solves the following equation:

$$\tilde{\pi}\left(0;\left(n_k^i,X\right)\right)\equiv\varepsilon+\left(1-\frac{1}{1+\left(\frac{1-\varepsilon}{\varepsilon}\right)^{n_k^i-2X}}\right)(1-2\varepsilon)=\delta.$$

This equation has a unique solution since, by hypothesis, $\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)\geq\delta$ for some $\tilde{n}_k^i\in[0,n_k^i]$; moreover, $\tilde{\pi}\left(0;\left(n_k^i,n_k^i/2\right)\right)=1-\varepsilon<\delta$, and the left-hand side is continuous and monotonically decreasing in both ε and X in the range where $X\in\left(0,n_k^i/2\right)$. [To see that $\tilde{\pi}\left(0;\left(n_k^i,X\right)\right)$ is decreasing in ε, let $\zeta\left(n_k^i,X\right)\equiv\left[1+\left(\frac{1-\varepsilon}{\varepsilon}\right)^{n_k^i-2X}\right]^{-1}$. In the range $X\in\left(0,n_k^i/2\right)$, we have $\frac{\partial}{\partial\varepsilon}\tilde{\pi}\left(0;\left(n_k^i,X\right)\right)=1-2\left(1-\zeta\left(n_k^i,X\right)\right)-(1-2\varepsilon)\frac{\partial\zeta\left(n_k^i,X\right)}{\partial\varepsilon}$. Since $\zeta\left(n_k^i,X\right)<1/2$ and $\partial\zeta\left(n_k^i,X\right)/\partial\varepsilon>0$, $\partial\tilde{\pi}\left(0;\left(n_k^i,X\right)\right)/\partial\varepsilon<0$ follows]. Therefore, there exists a unique $X\left(\varepsilon,n_k^i\right)$ such that $\tilde{\pi}\left(0;\left(n_k^i,X\left(\varepsilon,n_k^i\right)\right)\right)=\delta$. Lemma 4 implies that $\tilde{\pi}\left(0;\left(n_k^i,X\left(\varepsilon,n_k^i\right)\right)\right)=\delta$ if and only if $\tilde{\pi}\left(0;\left(n_k^i,n_k^i-X\left(\varepsilon,n_k^i\right)\right)\right)=\delta$. Let $Q\left(\varepsilon,n_k^i\right)\in\mathbb{Z}_+$ be the largest integer smaller than X, which will be the threshold number of realization of either L or R such that $\pi\left(0;\left(n_k^i,\tilde{n}_k^i\right)\right)>\delta$. More formally, $Q\left(\varepsilon,n_k^i\right)\equiv\max_{z\in\mathbb{Z}_+}\left\{z\leq X\left(\varepsilon,n_k^i\right)\right\}$. The properties of $X\left(\varepsilon,n_k^i\right)$ immediately imply that $Q\left(\varepsilon,n_k^i\right)$ is no larger than $\left(n_k^i-1\right)/2$ and is non-increasing in ε.■

Proof of Lemma 6: Since $\tilde{n}_k^i$ is the number of successes out of n_k^i in a Bernoulli trial, then holding

Q constant at $\bar{Q}$, we have that

$$\begin{aligned}
\Pr\left[\tilde{n}_k^i \leq \bar{Q}\right] &= \sum_{\tilde{n}_k^i=0}^{\bar{Q}} \binom{n_k^i}{\tilde{n}_k^i} (1-\varepsilon)^{\tilde{n}_k^i} \varepsilon^{n_k^i-\tilde{n}_k^i}, \text{ and} \\
\Pr\left[\tilde{n}_k^i \geq 1-\bar{Q}\right] &= \sum_{\tilde{n}_k^i=n_k^i-\bar{Q}}^{n_k^i} \binom{n_k^i}{\tilde{n}_k^i} (1-\varepsilon)^{\tilde{n}_k^i} \varepsilon^{n_k^i-\tilde{n}_k^i}, \\
&= \sum_{\tilde{n}_k^i=0}^{\bar{Q}} \binom{n_k^i}{\tilde{n}_k^i} \varepsilon^{\tilde{n}_k^i} (1-\varepsilon)^{n_k^i-\tilde{n}_k^i}.
\end{aligned}$$

The latter equation implies

$$\begin{aligned}
\frac{\partial \Pr\left[\tilde{n}_k^i \geq 1-\bar{Q}\right]}{\partial \varepsilon} &= \sum_{\tilde{n}_k^i=0}^{\bar{Q}} \binom{n_k^i}{\tilde{n}_k^i} \frac{\partial}{\partial \varepsilon}\left(\varepsilon^{\tilde{n}_k^i} (1-\varepsilon)^{n_k^i-\tilde{n}_k^i}\right) \\
&= \binom{n_k^i}{0}\left(-n_k^i (1-\varepsilon)^{n_k^i-1}\right) + \\
&\quad \binom{n_k^i}{1}\left((1-\varepsilon)^{n_k^i-1} - \left(n_k^i-1\right)\varepsilon(1-\varepsilon)^{n_k^i-2}\right) + \\
&\quad \binom{n_k^i}{2}\left(2\varepsilon(1-\varepsilon)^{n_k^i-2} - \left(n_k^i-2\right)\varepsilon^2(1-\varepsilon)^{n_k^i-3}\right) + \\
&\quad \ldots\ldots + \\
&\quad \binom{n_k^i}{\bar{Q}-1}\left(\left(\bar{Q}-1\right)\varepsilon^{\bar{Q}-2}(1-\varepsilon)^{n_k^i-\bar{Q}-1} - \left(n-\left(\bar{Q}-1\right)\right)\varepsilon^{\left(\bar{Q}-1\right)}(1-\varepsilon)^{n_k^i-\bar{Q}}\right) + \\
&\quad \binom{n_k^i}{\bar{Q}}\left(\bar{Q}\varepsilon^{\bar{Q}-1}(1-\varepsilon)^{n_k^i-\bar{Q}} - \left(n-\bar{Q}\right)\varepsilon^{\bar{Q}}(1-\varepsilon)^{n_k^i-\bar{Q}-1}\right).
\end{aligned}$$

Evaluating these terms and canceling them pairwise, we obtain

$$\frac{\partial \Pr\left[\tilde{n}_k^i \geq 1-\bar{Q}\right]}{\partial \varepsilon} = -\binom{n_k^i}{\bar{Q}}\left(n-\bar{Q}\right)\varepsilon^{\bar{Q}}(1-\varepsilon)^{n_k^i-\bar{Q}-1}.$$

A similar argument establishes:

$$\frac{\partial \Pr\left[\tilde{n}_k^i \leq \bar{Q}\right]}{\partial \varepsilon} = \binom{n_k^i}{\bar{Q}}\left(n-\bar{Q}\right)(1-\varepsilon)^{\bar{Q}}\varepsilon^{n_k^i-\bar{Q}-1}.$$

Combining these two expressions, we have

$$\begin{aligned}
\frac{\partial \chi\left(n_k^i, \tilde{n}_k^i, \varepsilon, \bar{Q}\right)}{\partial \varepsilon} &= \frac{\partial \Pr\left[\tilde{n}_k^i \geq 1-\bar{Q}\right]}{\partial \varepsilon} + \frac{\partial \Pr\left[\tilde{n}_k^i \leq \bar{Q}\right]}{\partial \varepsilon}, \\
&= \binom{n_k^i}{\bar{Q}}\left(n-\bar{Q}\right)\left((1-\varepsilon)^{\bar{Q}}\varepsilon^{n_k^i-\bar{Q}-1} - \varepsilon^{\bar{Q}}(1-\varepsilon)^{n_k^i-\bar{Q}-1}\right), \\
&= \binom{n_k^i}{\bar{Q}}\left(n-\bar{Q}\right)(1-\varepsilon)^{n_k^i-1}\left(\left(\frac{1-\varepsilon}{\varepsilon}\right)^{\bar{Q}} - \left(\frac{1-\varepsilon}{\varepsilon}\right)^{n_k^i-\bar{Q}-1}\right) < 0,
\end{aligned}$$

where the last inequality follows using the facts that $\varepsilon < 1/2$ and $\bar{Q} \leq \left(n_k^i - 1\right)/2$.■

Intuitively, when ε is small, the performance of firms that have implemented the same technology in the past reveals more information about the reference action. Thus, firms' posterior beliefs are more responsive to public information. Note that this applies to both "correct" and "incorrect" beliefs. For instance, suppose that $x_k^* = L$, but in the sample available to the firm R has been successful more than half of the time; then, when ε is small, the firm will assign higher probability to R being the correct action (i.e., $\pi\left(0; h_k^i \mid x_{i,k} = R\right)$ will take on a larger value). The complication in establishing Proposition 2 comes from the fact that a change in ε affects the likelihood of different histories. Nevertheless, the proof establishes that a greater ε changes the ex ante distribution of different histories so as to also increase $\Pr\left(d_{i,k}^*\left(h_t^i\right) = 1\right)$.

Proposition 2 provides the most interesting testable implication of our approach; it suggests that there should be more decentralization in industries with greater dispersion of performance across firms and also for firms that are more dissimilar to others.

Finally, we briefly extend our basic model to derive a relationship between firm age and organizational structure. In the model analyzed so far the deviation between the reference action and the correct action for each firm was independently and identically distributed across technologies, firms and time. Consequently, a firm's information on how to implement technology k was independent from that firm's actions and performance on previous technologies $k' < k$. More generally, one could assume that there is a positive correlation between the correct actions that a firm should take across successive technologies, for example, because the specific skills of the employees or the culture of the organization differ across firms. In this case, the firm could learn from its own past experience as well as from the experiences of other firms.

Since solving the signal extraction problem with multiple sources of uncertainty is complicated and not our main focus here, we assume in this subsection that firms cannot learn from other firms. This enables us to focus instead on firms' learning from their own performance. The analogue to equation (1) in the text is

$$x_{i,k}^* = \begin{cases} x_i^* & \text{with probability } 1-\varepsilon_i \\ \sim x_i^* & \text{with probability } \varepsilon_i \end{cases},$$

for any $I_i(k;t) = 1$. This equation implies that although the reference action for firm i is $x_i^* \in \{L, R\}$, the correct action for technology k may differ from this with some probability $\varepsilon_i < 1/2$.

This equation implies that the updating problem is now identical to that discussed previously, up to a reinterpretation of the information set. In particular, what used to be history $h_k^i = \left\{n_k^i, \tilde{n}_k^i\right\}$ is now replaced by $h_k^i = \left\{n_i^k, \tilde{n}_i^k\right\}$, where, given $I_i(k;t) = 1$, n_i^k denotes the number of technologies that firm i has implemented before technology k, and $\tilde{n}_i^k$ denotes the number of times in which action L turned out to be the correct choice in this firm's own experience in the past. Given this reinterpretation, our previous analysis implies (proof omitted):

Proposition 3 ***(Age)*** Consider the adoption decision of technology k by firm i, and suppose that Assumption 1 holds and $x_i^* = L$. Then:
(i) For the youngest firm with $n_i^k = 0$, we have $q\left(h_t^i = \{0,0\}\right) = \pi\left(0, h_t^i = \{0,0\}\right) = 1/2 < \delta$, and the principal chooses delegation, $d_{i,k}^*\left(h_t^i\right) = 1$;
(ii) For a sufficiently old firm,, i.e., $n_i^k \to \infty$, we have $p\lim_{n_i^k \longrightarrow \infty} q\left(h_t^i = \left\{n_i^k, \tilde{n}_i^k\right\}\right) = 1$, $p\lim_{n_k^i \longrightarrow \infty} \pi\left(0, h_t^i = \left\{n_i^k, \tilde{n}_i^k\right\}\right) = 1 - \varepsilon > \delta$, and the principal almost surely chooses centralization, i.e., $p\lim_{n_k^i \longrightarrow \infty} d_{i,k}^*\left(h_t^i = \left\{n_i^k, \tilde{n}_i^k\right\}\right) = 0$.

We next briefly discuss the possibility that firms may use incentive contracts to induce managers to choose the right action. Let the private benefit on the manager when the implements is preferred action be By_{t-1}. We will show that when B (the benefit accruing to the manager when she chooses her preferred action) is sufficiently large, such incentive contracts will not be optimal. The intuition is that because managers are credit constrained, incentive contracts give the right incentive to managers only by transferring rents to them. If B is large, this is not profitable for the principal.

Let us assume that the principal decides whether to hire the manager before knowing whether that $I_i(k,t) = 1$. Let us also normalize the outside option of the manager to zero, and recall that the manager is also risk neutral. Given the credit constraints of the manager , the optimal contract takes a simple form: the principal will pay the manager $By_{i,t-1}$ in case of success. Both when the manager is unsuccessful in the implementation of the new technology and when there is no new technology to be implemented, it is optimal for the principal to pay him zero. This contract will induce the manager to choose the right action. It will also meet his participation constraint, since the manager will receive $By_{i,t-1}$ with probability $p_i\delta > 0$.

The alternative is to pay the manager zero irrespective of success, and let him choose his preferred action. This contract also meets the manager's participation constraint, since he derives private benefits from the implementation of the project. This option was the one analyzed in the text. Since the issue of whether there is delegation or not is only interesting in the case when there is a technology to be implemented, let us focus on time t such . We then have:

Proposition 4 Suppose that $B > \frac{(\gamma-1)(1-\delta)}{1-\beta p_i \gamma}$ and that $I_i(k,t) = 1$. Then, for any history $h_k^i \in \mathcal{H}_k^t$, the optimal strategy for the principal of firm i is not to offer an incentive contract to the manager.

Proof. Let $d_{i,k} = 2$ denote firm i's decision to delegate control with full compensation to the manager for choosing the profit-maximizing action at date t on technology k when $I_i(k,t) = 1$. The value of the decentralized firm offering the manager an incentive contract is

$$\begin{aligned} V_i\left(y_i, h_k^i \mid 2\right) &= p_i[(\gamma - B)\, y_i + \beta \mathbb{E}\left[V_i\left(\gamma y_i, h_{k+1}^i\right) \mid h_k^i\right]] \\ &\quad +(1-p_i)\{y_i + \beta \mathbb{E}\left[V_i\left(y_i, h_{k+1}^i\right) \mid h_k^i\right]\}. \end{aligned}$$

Solving this functional equation leads to $V_i\left(y_i, h_k^i \mid 2\right) = A_i\left(h_k^i \mid 2\right) y_i$, where:

$$\begin{aligned} A_i\left(h_k^i \mid 2\right) &= [p_i(\gamma - B) + 1 - p_i] + \\ &\quad \beta[p_i\gamma + 1 - p_i] \max_{d_{i,k+1}\in\{0,1,2\}} \mathbb{E}\left[A_i\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right]. \end{aligned} \tag{18}$$

Instead, the value of a decentralized firm offering the manager a flat wage is $V_i\left(y_i, h_k^t \mid 1\right) = A_i\left(h_k^t \mid 1\right) y_i$, where

$$\begin{aligned} A_i\left(h_k^i \mid 1\right) &= [p_i(\delta\gamma + (1-\delta)) + 1 - p_i] \\ &\quad +\beta[p_i(\delta\gamma + (1-\delta)) + 1 - p_i] \max_{d_{i,k+1}\in\{0,1,2\}} \mathbb{E}\left[A_i\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right]. \end{aligned} \tag{19}$$

Which of the two regimes yields a larger value to the firm depends on whether $A_i\left(h_k^i \mid 1\right)$ is larger or smaller than $A_i\left(h_k^i \mid 2\right)$. To establish when this is the case, note that

$$\begin{aligned} A_i\left(h_k^i \mid 1\right) &> A_i\left(h_k^i \mid 2\right) \Leftrightarrow \\ B &> (\gamma-1)(1-\delta)\left(1 + \beta \max_{d_{i,k+1}\in\{0,1,2\}} \mathbb{E}\left[A_i\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right]\right). \end{aligned} \tag{20}$$

An upper bound to the future value of the firm can be calculated by assuming that, from period $(t+1)$ onwards, the firm will innovate successfully whenever a new technology opportunity arises, which takes place with probability p_i, and will pay no managerial wage. This yields

$$\max_{d_{i,k+1}\in\{0,1,2\}} \mathbb{E}\left[A_i\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right] < \bar{A} \equiv \frac{p_i\gamma}{1-\beta p_i\gamma}.$$

Thus, substituting $\max_{d_{i,k+1}\in\{0,1,2\}} \mathbb{E}\left[A_i\left(h_{k+1}^i \mid d_{i,k+1}\right) \mid h_k^i\right]$ by $\bar{A}$ we obtain the sufficient condition

$$B > \frac{(\gamma-1)(1-\delta)}{1-\beta p_i\gamma} \Rightarrow A_i\left(h_k^i \mid 1\right) > A_i\left(h_k^i \mid 2\right)$$

for incentive contracts not to be profitable for the principal. ■

Appendix B

B.1 French Data

COI ("Changements Organisationnels et Informatisation," SESSI)

This is a firm level survey providing information on organization and other firm characteristics conducted in 1997. It covers manufacturing sectors only (4,153 firms). There are several questions on organizational design.

ER (Enquête Reponse 1998; "Relations Professionnelles et Négociations d'Entreprise," DARES)

The Enquête Reponse is an establishment level survey. This contains information about organizational change between 1996 and 1998. It covers both manufacturing and non-manufacturing sectors and is an updated version of the Reponse 1992 survey used by Caroli and Van Reenen (2001). 2,943 establishments of manufacturing and non-manufacturing sectors were surveyed with senior managers being asked questions about industrial relations, organization and other aspects of performance in 1998.

FUTE files ("Format Unifié Total d'Entreprises," INSEE)

The FUTE dataset is the key data we use to construct many of the variables used in the paper. FUTE is constructed from the merging of two datasets, the BRN ("Bénéfices Réels Normaux") and the EAE ("Enquêtes Annuelles d'Entreprises"), that are then checked rigorously for consistency at INSEE.

The BRN files consist of firms' balance sheets collected annually by the Direction Générale des Impôts (Fiscal Administration) and provides firm-level accounting information (value added, capital investment, wage bills,employment, etc.). This tax regime is mandatory for the companies that have a level of sales higher than 3.8 million Francs, but can also be also disclosed by smaller firms. These files include around 600,000 firms,[32] in the private non-financial, non-agricultural sectors each year and covers around 80% of total output in the French economy. The EAE survey is conducted by SESSI (production industries), INSEE (Services and Trade), the Ministry of Agriculture and Ministry of Equipment (Transportation and Construction). The annual survey is mandatory and exhaustive for firms hiring more than 20 workers. It includes a detailed sectoral description of the various activities of each firm surveyed (the amount of each kind of output).[33]

DADS files ("Déclarations Annuelles de Données Sociales")

The DADS files consists of yearly mandatory employer reports of each worker's gross earnings subject to payroll taxes. Hours are also reported since 1993 (but of good quality only since 1994). These files include around 27 million workers each year (27,535,562 in 1996 after some basic cleaning), which we aggregate at the plant (1,587,157 plants in 1996) or firm level (1,379,032 firms in 1996) to get information on the workforce structure (age, gender, skill group in terms of hours worked). We also use the total hours series necessary for the measures of productivity underlying the heterogeneity and proximity to the frontier measures (see below).

[32] 630,593 firms in 1996 of which 489,783 report a strictly positive number of workers.

[33] In French, the question was :
"Répartir le chiffre d'affaires net hors taxes et les exportations directes de votre entreprise selon les différentes activités conformément aux nomenclatures officielles d'activités et de produits. Le total du chiffre d'affaires net doit correspondre au montant du poste du compte de résultat. Les reventes en l'état de marchandises ou de produits doivent être déclarées dans une ou plusieurs rubriques *négoce.*"

LIFI Surveys ("Liaisons Financières," INSEE)

Yearly survey describing the structure of ownership of French firms of the private sector whose financial investments in other firms (participations) are higher than 8 million Francs or having sales above 400 million Francs or a number of workers above 500.

Even after keeping only firms who are in the COI, BRN, DADS and EAE we are still left with over 90% of the original COI sample (3,751 observations). The firms who we lose tend to be the smallest firms. We loose a few more observations in our regressions due to missing values on the some of the questions in COI (final sample for regressions is 3,570 observations). For the Reponse sample, we only keep firms that are part of a larger French or foreign group, but that are not the corporate head quarters (final sample for regressions is 1,258 observations).

B.2 Variable Definitions

The firm and industry level quantitative variables introduced in the regressions are averaged over four years (COI) if available (three years for Reponse, "Delayering"). Unless otherwise indicated all industry variables are at the four-digit NACE level.

B.2.1 Decentralization into Profit Centers

Our main measure of decentralization is from the COI. Managers were asked:

Is your firm organized into profit centers ?

The translation of the French definition used in COI is "*Organization in profit centers. A profit center is an enterprise unit that has a margin of budgetary manoeuvre, and therefore some relative autonomy in their choices (usually it has its own accounting system to measure their profit).*"

We coded the measure of decentralization to be unity if the manager answered "yes" to this question and zero if the answer was "no".

Using the COI we build a measure of decentralization based on the organization of its business units into profit centers.[34] In practice, once a firm gets beyond a minimal size it faces the choice of retaining central control or allowing some decentralization. Firms are generally organized into business units and different firms make decisions about what degree of responsibility to devolve to the managers of these units. Some firms retain complete command and control at the center, but most create some form of "responsibility centers" for business unit managers.[35] Business scholars delineate three broad types of responsibility centers (from the most to the least decentralized): profit centers, cost centers and revenue centers. Our key indicator for decentralization is whether the firm is organized primarily into profit centers. As its name suggests, when a firm organizes into profit centers a manager is responsible for the profits of the unit she manages. In general the profit center manager is given considerable autonomy to make decisions on the purchase of assets, hiring of personnel, setting salary and promotion schedules and managing inventories. A manager of a profit center is concerned with all aspects of the business that contribute to profitability. Such a manager keeps track of both revenues and costs with the aim of maximizing profit. As one management specialist puts it:

"The profit center managers frequently know their business better than top management does because they can devote much more of their time to following up developments in their specialized areas. Hence, top level managers usually do not have detailed knowledge of the actions they want particular profit center managers to take, and even direct monitoring of the actions taken, if it were feasible would not ensure profit center managers were acting appropriately." (*Motivating Profit Center Managers*, Merchant, 1989, p.10)

In contrast to a profit center manager, a cost center manager will have the quantity or quality of output set by someone higher up in the organization. The manager is delegated with some power,

[34] This follows Janod (2002) and Janod and Saint-Martin (2004).

[35] For an introduction to responsibility centers in general and profit centers in particular see, for example: http://smccd.net/accounts/nurre/online/chtr12a.htm.

however, in order to try and reduce costs. He will be able to decide on some short-run (but not long-term) asset purchases, hire temporary and contract staff (but not permanent employees) and manage inventories. A revenue center manager has the least autonomy of all.[36] She is told to spend a certain amount of resource and account for revenues but has no (or little) discretion to exceed spending limits. Inventories are managed but staff and investments are not acquired unless he is authorized explicitly to do so.

There are numerous examples from the business literature on the greater autonomy of profit centers. It is well recognized that organizing divisions into profit centers delegates more power to managers, and it is generally agreed that a characteristic of companies that organize divisions into profit centers is that it "allows decision making and power to be delegated effectively". Similarly, the first disadvantage of profit centers is viewed as "loss of overall central control of the company." [37]

Although it is possible in principle for a profit center manager to be monitored on profits and yet not be given any powers to affect these profits would seem sub-optimal for the firm (Dearden, 1987, Merchant, 1989, Bouwens and van Lent, 2004). A profit center manager would be held responsible for outcomes that he cannot affect, so this would de-motivate such managers. Some organizations like this probably exist - the only way to know more would be to have subjective questions on the degree to which different profit center managers have greater decision making powers. The advantage of our profit center variable is that it is an objective feature of the firm and does not rely on a manager's subjective statement of his power relative to a senior manager.

In short, we have an indicator equal to unity if the firm is organized into profit centers and a zero otherwise. So the base group contains firms who are organized primarily into responsibility centers with less autonomy (i.e. cost and revenue centers) and those firms who have no responsibility centers at all and maintain command and control. Unfortunately, the data does not allow us to distinguish the latter groups more finely.

B.2.2 Managerial Autonomy to Make Investment Decisions

In the Enquête Reponse 1998 the establishment's senior manager was asked how much autonomy from headquarters she had to make decisions over investment.[38] Answers were coded to be one if she answered that she had "full" autonomy or "important"' autonomy and coded to zero if she had "limited" or "no" autonomy. This is used as the dependent variable in the first three columns of Table 4. We consider only establishments that are part of a wider group (as a single site establishment will not have a separate headquarters).

B.2.3 Delayering

While COI provides data about the current organization of the firm, the ER dataset provides information on organizational "changes" (i.e., whether a firm became more or less hierarchical) rather than "levels".

Our preferred measure of delayering is from the Enquête Reponse 1998 where we use the following question:
For any of the following technologies and methods, would you tell us whether it is implemented in your establishment ?
- Shortening of the hierarchical line (delayering of an intermediate hierarchical level).
The indicator used is a dummy variable coded to one if the respondent answered "yes" to this question.

[36] In fact "revenue center" is rather a misnomer because a notional revenue is assigned by the organization's controller based on activities and transfer prices. "Expense center" is sometimes used as the manager accounts mainly for the expenses incurred.

[37] These quotes are taken from the educational web-site:
http://www.aloa.co.uk/members/downloads/PDF%20Output/costcentres.pdf
See also Janod (2002).

[38] In French: "Par rapport au siège ou à la maison mère de l'entreprise ou du groupe, quelle est l'autonomie de votre établissement en matière d'investissement?
Totale / Importante / Limitée / Nulle."

Case study and econometric evidence suggest that delayering is associated with decentralization (Rajan and Wulf, 2005, Caroli and Van Reenen, 2001).

B.2.4 Proximity to Technological Frontier

Value-added (VA_{ilt}, FUTE) is defined as sales minus purchases of materials. It is deflated with the value added price index at the two-digit level (NAF36) available from National Accounts (value added at 1995 prices)
Total hours ($HOURS_{ilt}$, DADS)
Capital Stock (K_{ilt}, FUTE) is computed from firm level fixed assets. This information is registered at historical cost in the balance sheets. We recover volumes by deflating the initial measure by the investment price index (National Accounts) at the date considered minus an estimated age of capital. This age is calculated as the ratio of depreciated assets over fixed assets multiplied by an average equipment length of life (16 years).

Labour Productivity and TFP are defined as:

$$\begin{aligned} y_{ilt} &= \ln(VA_{ilt}) - \ln(HOURS_{ilt}) \\ TFP_{ilt} &= \ln(VA_{ilt}) - \alpha_l \ \ln(HOURS_{ilt}) - (1-\alpha_l) \ \ \ln(K_{ilt}) \end{aligned}$$

Where α_l is the wage bill share of value added in the four-digit NACE (we also considered an economy-wide weight of 0.7). We drop firms reporting divergent values of total number of employees in the FUTE and in the DADS (values greater than double one way or the other). Industries represented by less than ten firms in the FUTE are also dropped. For each firm (and like other firm level variables introduced in the regressions) the labour productivity and TFP values are averaged over four years if available (three years respectively Delayering in Enquete Reponse).

The industry "frontier" ($y_{F_l t}$ or $TFP_{F_l t}$) is defined as the 99$^{\text{th}}$ percentile (or 95$^{\text{th}}$, or 90$^{\text{th}}$ when specified) of the obtained series at the NACE four-digit level. The constrained term defined as $GAP^y_{ilt} = \ln(y_{ilt}) - \ln(y_{F_l t})$ is a firm level measure of proximity to the technological frontier.

Another alternative measure of distance to frontier is the rank of firms in their industry (the firms are ranked according to their labour productivity in the regressions presented).

B.2.5 Heterogeneity Measures

We use the FUTE to construct $\Delta \ln y_{ilt}$, the firm specific annual productivity growth rate (value added per hour) for all firms. We average this growth rate for up to three years. We then construct the percentiles of the inter-firm productivity growth distribution within each four-digit NACE sector. The 90-10 is $(\Delta \ln y_{ilt})^{90} - (\Delta \ln y_{ilt})^{10}$ where $(\Delta \ln y_{ilt})^{90}$ is the productivity growth at the 90$^{\text{th}}$ percentile and $(\Delta \ln y_{ilt})^{10}$ is productivity growth at the 10$^{\text{th}}$ percentile. Alternative measures of heterogeneity are based on other indicators of dispersion of the same series of firm level labour productivity growth rates: the 95-5 is $(\Delta \ln y_{ilt})^{95} - (\Delta \ln y_{ilt})^{5}$, the standard deviation, the standard deviation after trimming bottom and top 5 % of values in each four-digit industry.

The levels-based measure of industry it originates is defined similarly as $(\ln y_{ilt})^{90} - (\ln y_{ilt})^{10}$.

In addition, we use the FUTE files to construct two additional firm - level measures of heterogeneity A firm i is characterized by its vector of kind of productions (sold, l being one of its markets):

$$S_i = (S_{i1}, \ldots, S_{il}, \ldots, S_{iL}), \quad \text{or in shares :} \quad s_i = \left(\underbrace{\frac{S_{i1}}{\sum_{h \in \mathcal{L}} S_{ih}}}_{s_{i1}}, \ldots, \underbrace{\frac{S_{il}}{\sum_{h \in \mathcal{L}} S_{ih}}}_{s_{il}}, \ldots, \underbrace{\frac{S_{iL}}{\sum_{h \in \mathcal{L}} S_{ih}}}_{s_{iL}} \right)$$

where $\mathcal{L}$ refers to the set of industries. Our main index of firm-level heterogeneity is constructed as

$$H_i^F = \log\left(\frac{\sum_{i'\in\mathcal{N},i'\neq i} c_{ii'}\cdot IT_{i'}}{\sum_{i'\in\mathcal{N},i'\neq i} IT_{i'}}\right)^{-1}$$

for firm i at time t, with N referring to the sample of firms in the FUTE, and IT_i refers to the level of IT investment of firm i,[39] and the closeness measure is

$$c_{ii'} = \frac{\sum_{l\in\mathcal{L}} s_{il}.s_{i'l}}{(\sum_{l\in\mathcal{L}} s_{il}^2)^{\frac{1}{2}}.(\sum_{l\in\mathcal{L}} s_{i'l}^2)^{\frac{1}{2}}},$$

The alternative measure is an unweighted version of our main measure calculated as:

$$H_i^{FA} = \log\left(\frac{\sum_{i'\in\mathcal{N},i'\neq i} c_{ii'}}{N-1}\right)^{-1},$$

where N is the total number of firms in the set $\mathcal{N}$.

B.2.6 Other Firm Level Variables

All firm level variables are averaged over four years if available.

Lerner Index: We calculate gross profits as value added minus labor costs and then divide by sales. This proxies for the average profit margin. All these variables are sourced from FUTE. We considered also making a control for total capital costs using the stock of capital and the average user-cost (this is also deducted from gross profits to allow a normal rate of return on fixed capital). This lead to very similar results.

Capital Intensity: Fixed capital stock divided by value added. Sourced from FUTE.

Firm / Plant age: Information available from the SIRENE dataset (reproduced in the DADS) or from files reporting the yearly creations of firms (Firm Demography Department). Plant age is available in the Reponse survey.

Joint Stock Firms: Indicator of a firm being a Joint Stock Company (as opposed to smaller and less anonymous structures, e.g. limited liability firms). Sourced from FUTE.

Foreign Ownership: Indicator of whether a firm is part of a larger foreign group. Sourced from LIFI.

Number of Plants: Number of plants belonging to each firm (and their region of localization). Sourced from DADS.

Size: The *number of workers* at the plant level for Reponse and at the firm level for COI. Sourced from DADS.

Skills: Share of hours worked by skilled workers at the firm level. We consider as unskilled: *Industrial blue collar workers* (CS 67, Ouvriers non qualifiés de type industriel); *Craftsmen* (CS 68, Ouvriers non qualifiés de type artisanal), *Foremen and Supervisors* (CS 53, Agents de surveillance), *Clerical* (CS 55, Employés de commerce), *Personnel of the direct services to the private individuals sectors* (CS56, Personnels des services directs aux particuliers). Others are considered as "skilled". Sourced from DADS.

Worker age: Average age of workers at the firm level (weighted by hours worked). Sourced from DADS.

[39] Note that the IT investment series (EAE/FUTE) are available in 1996 and 1997.

Technology: A pseudo-continuous variable of proportion of workers using micro-computers is constructed from information available in both Reponse 98 (relating to 1998) and COI (1997).

We use the FUTE dataset and the decomposition of the various activities of firms i in terms of amount of each kind of product l produced and sold (S_{il}) to construct the following indicators:

Firm-level Market Share:

$$MS_i = \sum_l \frac{S_{il}}{S_i} . \frac{S_{il}}{S_l}$$

Herfindahl Index:

$$HR_i = \sum_l \frac{S_{il}}{S_i} . HR_l, \quad HR_l = \sum_i \left(\frac{S_{il}}{\sum_{i'} S_{i'l}} \right)^2$$

This is constructed in the standard way at the industry level (H_l), but note that we weight this measure if a firm operates in more than one market (by a firm's market share in that sector ($\frac{S_{il}}{S_i}$), so it has a firm specific component).

Firm-level Diversification Indicator:

This simply indicates (the inverse of) the degree to which a firm operates across separate sectors

$$SPE_i = \sum_l \left(\frac{S_{il}}{S_i} \right)^2 .$$

B.2.7 Other Industry Level Information

All industry level variables are averaged over three years.

Sector Capital Intensity total capital stock in the four-digit industry divided by total number of workers in the industry. Sourced from FUTE.

Sector IT Investment: total IT investment in the four-digit industry divided by total number of workers in industry. Sourced from FUTE.

B.3 UK Data

B.3.1 Workplace Employee Relations Survey (WERS)

WERS 1998 is a survey of establishments in Britain conducted in 1998 (there were also surveys in 1980, 1984 and 1990).[40] It is described in detail in Cully *et al.* (1999). In one part of the questionnaire the establishment's senior manager is asked whether she "is able to make decisions without consulting" Head Quarters. Some of these decisions are relatively minor (such as staff appraisal). We focus on whether decisions over staff recruitment can be made by establishment's management without consulting someone higher in the corporate hierarchy as this is a key aspect of decentralized decision making (unfortunately the question on investment decisions used in France was not asked).

This question was only asked if the establishment was part of a larger multi-plant firm. Although the question was asked to all establishments we have to focus on manufacturing because the ABI data is only available for services from 1997 onwards so we would not be able to construct a robust measure of heterogeneity. The WERS data cannot be matched at the establishment-level to Census data so we are unable to condition on as rich a set of covariates as we can in France. In particular, we do

[40] The 1984 and 1990 panels were used by Caroli and Van Reenen (2001). There was also a WERS conducted in 2004, but the four-digit industry codes for this data set have not yet been released (there are only 12 industry divisions).

not have information on value added, profits or capital. Consequently we cannot include measures of the establishment's own productivity or Lerner Index in the regression. WERS does contain basic information on workers demographics (skill, age, female and part-timers) and we condition on these in the regressions (see Table notes). As a proxy for market power we used the question asked to managers whether the establishment faces no competitors, some competitors or many competitors (this is the same as Nickell, 1996).

B.3.2 ABI "Census" Data

We constructed heterogeneity and Frontier productivity terms for each UK four-digit industry. The UK and France share the European Unions' NACE classification system so this was straightforward. The only restriction was that industry averages with cell sizes below 25 are not allowed out of the Office of National Statistics (ONS).

Our base dataset is a panel of establishments covering almost all sectors of the UK private sector called the ABI (Annual Business Inquiry). This underlies many of the UK national statistics and is similar in structure to the US Longitudinal Research Database (LRD) being a population sample of large plants and a stratified random sample of smaller plants. The response rates to the ABI are high because it is illegal not to return the forms to the Office of National Statistics (ONS). The ABI contains all the basic information needed to estimate production functions (gross output, labour, materials, investment, etc.). For each firm we constructed value per worker and followed the same rules described for France to calculate heterogeneity (e.g. 90-10 of productivity growth rates) and the Frontier productivity (99th percentile). The UK data does not contain information on hours so the UK productivity measure is cruder than in France.

C Robustness Checks

The robustness checks discussed in Section 5.4 are presented in Tables B1 and B2. Table B1 includes the results with alternative estimation strategies, additional controls and alternative samples. Table B2 includes the results of the instrumental variables strategy discussed in the text together with the corresponding first stages.

Table B1: Robustness checks on the probability of firm being decentralized
(Enquête COI)

Dependent variable (mean=0.304)	Firm decentralized into Profit Centers						
	Baseline (marginal effects) (1)	Logit ML (marginal effects) (2)	Linear probability model (3)	Frontier: TFP (4)	Firm rank in productivity distribution (5)	Frontier: 95^{th} percentile (6)	Frontier: 90^{th} percentile (7)
Heterogeneity	0.296	0.293	0.261	0.243	0.239	0.221	0.181
	(0.127)	(0.121)	(0.119)	(0.095)	(0.105)	(0.121)	(0.117)
Frontier term	-0.225	-0.228	-0.191	-0.066	-	-0.179	-0.104
(99^{th} percentile)	(0.045)	(0.044)	(0.040)	(0.023)		(0.077)	(0.091)
Labour productivity	0.141	0.144	0.128	0.202	-	0.137	0.132
(firm level)	(0.033)	(0.033)	(0.030)	(0.060)		(0.033)	(0.033)
Firm rank	-	-	-	-	-0.038	-	-
					(0.012)		
Firm age<5 years	0.172	0.172	0.151	0.157	0.164	0.170	0.170
	(0.041)	(0.044)	(0.034)	(0.042)	(0.041)	(0.041)	(0.041)
5≤ Firm age<10 years	0.066	0.061	0.057	0.075	0.02	0.065	0.065
	(0.022)	(0.023)	(0.017)	(0.023)	(0.022)	(0.022)	(0.022)
10≤Firm age<20 years	0.039	0.035	0.035	0.041	0.039	0.040	0.04
	(0.019)	(0.020)	(0.016)	(0.020)	(0.019)	(0.019)	(0.019)
Lerner index	-0.660	-0.672	-0.586	-0.543	-0.521	-0.636	-0.634
	(0.144)	(0.144)	(0.124)	(0.146)	(0.123)	(0.144)	(0.144)
Industry dummies	yes (73)	yes (73)	yes (73)	yes (73)	yes (73)	yes (73)	yes (73)
Observations	3,570	3,570	3,570	3,479	3,570	3,570	3,570

NOTES: All coefficients are marginal effects from probit maximum likelihood estimation except column (2) which is estimated by by Logit (marginal effects at the sample mean reported) and column (3) which is estimated by by OLS (linear probability model). Robust standard errors corrected for arbitrary variance-covariance matrix at the four-digit industry level in parentheses. Heterogeneity is defined as the dispersion of productivity growth rates within a four digit industry (the 90^{th} percentile less the 10^{th} percentile). Full set of firm and industry level controls included; see text for variable definitions. Exact definition of each experiment is described in text. In column (4), the proximity to frontier terms are expressed in TFP instead of labor productivity. In column (5), the indicator of proximity to frontier is the firm's rank (in terms of labor productivity) in its four-digit industry. In columns (6) and (7), the frontier term is defined as the 95^{th} or 90^{th} percentile of the four-digit industry labor productivity (instead of the 99^{th} in the main specification).

Table B1: (-cont.) Robustness checks on the probability of firm being decentralized
(Enquête COI)

Dependent variable (mean=0.304)	Firm decentralized into Profit Centers				
	Weighted: log(employment) (8)	Part of a larger group (9)	Joint stock firms (10)	Specialized firms (11)	French owned (12)
Heterogeneity	0.368	0.461	0.403	0.324	0.309
	(0.138)	(0.140)	(0.116)	(0.159)	(0.120)
Frontier term	-0.251	-0.303	-0.192	-0.179	-0.152
(99^{th} percentile)	(0.051)	(0.056)	(0.041)	(0.053)	(0.041)
Labour productivity	0.152	0.146	0.144	0.110	0.155
(firm level)	(0.040)	(0.052)	(0.042)	(0.034)	(0.032)
Firm age<5 years	0.151	0.125	0.224	0.184	0.158
	(0.045)	(0.053)	(0.042)	(0.050)	(0.048)
5≤ Firm age<10 years	0.060	0.059	0.084	0.073	0.067
	(0.025)	(0.035)	(0.025)	(0.024)	(0.024)
10≤Firm age<20 years	0.044	0.061	0.048	0.026	0.032
	(0.022)	(0.029)	(0.022)	(0.021)	(0.021)
Lerner index	-0.688	-0.827	-0.658	-0.533	-0.675
	(0.173)	(0.244)	(0.174)	(0.156)	(0.139)
Industry dummies	yes (73)	yes (67)	yes (67)	yes (73)	yes (73)
Observations	3,570	1,793	2,990	2,555	2,951

NOTES: (-cont.) In column (8), all firms are weighted (in the regression) by the log of their employment. In column (9) "Part of a larger group" means that the firm is part of a larger group (either French or foreign). In column (10) "Joint stock firm" means that the firm is a joint stock fim, as opposed to limited liability firms, etc. In column (11) "Specialized" indicates that the firm has at least 80% of its sales in one four-digit industry. In column (12) "French owned" means that the firm is part of a French group.

Table B2: Allowing for Potential Endogeneity of Heterogeneity and Proximity to Frontier

(Enquête COI)

Dependent variable	Profit centers	Heterogeneity (France)	Profit centers	Proximity (France)	Profit centers	Heterogeneity (France)	Proximity (France)	Profit centers
Assumptions over the exogeneity of heterogeneity and proximity	Both exogenous (1)	(First stage) (2)	Endogenous heterogeneity (3)	(First stage) (4)	Endogenous proximity (5)	(First stage) (6)	(First stage) (7)	Both endogenous (8)
UK Heterogeneity	-	0.156	-	-	-	0.157	-0.093	-
		(0.048)				(0.054)	(0.099)	
UK Frontier	-	-	-	-0.218	-	0.005	-0.217	-
(99^{th} percentile)				(0.053)		(0.027)	(0.057)	
Heterogeneity	0.230	-	1.185	-0.419	0.310	-	-	1.572
(France)	(0.120)		(0.681)	(0.230)	(0.130)			(0.699)
Proximity to frontier	0.167	-0.017	0.187	-	0.341	-	-	0.456
(France)	(0.029)	(0.008)	(0.185)		(0.123)			(0.194)
Firm age<5 years	0.166	0.007	0.156	-0.054	0.175	0.008	-0.057	0.165
	(0.041)	(0.003)	(0.042)	(0.027)	(0.042)	(0.003)	(0.027)	(0.042)
5≤ Firm age<10 years	0.071	0.003	0.066	-0.030	0.077	0.004	-0.031	0.073
	(0.022)	(0.003)	(0.022)	(0.015)	(0.023)	(0.003)	(0.015)	(0.022)
10≤Firm age<20 years	0.043	0.005	0.038	-0.006	0.044	0.005	-0.008	0.037
	(0.020)	(0.002)	(0.020)	(0.011)	(0.020)	(0.002)	(0.011)	(0.020)
Lerner index	-0.723	0.035	-0.767	2.896	-1.228	-0.015	2.903	-1.542
	(0.137)	(0.025)	(0.545)	(0.109)	(0.387)	(0.015)	(0.110)	(0.570)
Industry dummies	yes (71)	yes (71)	yes (71)	yes (71)	yes (71)	yes (71)	yes (71)	yes (71)
Observations	3,518	3,518	3,518	3,518	3,518	3,518	3,518	3,518

NOTES: All coefficients are marginal effects from probit (multivariate normal) maximum likelihood estimation. Robust standard errors corrected for arbitrary variance-covariance matrix at the four-digit industry level in parentheses. Heterogeneity is defined as the dispersion of productivity growth rates within a four digit industry (the 90^{th} percentile less the 10^{th} percentile). Full set of firm and industry level controls included; see text for variable definitions. Observations only kept when match between UK and French four-digit industry is possible. In column (1), heterogeneity and proximity to frontier are both assumed exogenous in the "profit centers" equation. In columns (2) and (3), heterogeneity is assumed endogenous, while proximity to frontier is assumed exogenous. In columns (4) and (5), proximity to frontier is assumed endogenous, while heterogeneity is assumed exogenous. In columns (6) to (8), heterogeneity and proximity to frontier are both assumed endogenous.