

Occasional paper

36

Incomplete Contracts and the Internal Organisation of Firms

Phillipe Aghion
Nicholas Bloom
John Van Reenen

March 2013

Abstract

We survey the theoretical and empirical literature on decentralization within firms. We first discuss how the concept of incomplete contracts shapes our views about the organization of decision-making within firms. We then overview the empirical evidence on the determinants of decentralization and on the effects of decentralization on firm performance. A number of factors highlighted in the theory are shown to be important in accounting for delegation, such as heterogeneity and congruence of preferences as proxied by trust. Empirically, competition, human capital and IT also appear to foster decentralization. There are substantial gaps between theoretical and empirical work and we suggest avenues for future research in bridging this gap.

JEL classification: E3

Keywords: uncertainty

This paper was produced as part of the Centre's Productivity and Innovation Programme. The Centre for Performance is financed by the Economic and Social Research Council.

Acknowledgements

This paper was prepared for the "Grossman and Hart at 25" conference in Brussels. We are grateful for comments by our discussants, John Roberts and Bentley MacLeod, the editor and two anonymous referees, and research assistance from Megha Patnaik. We would like to thank the Economic and Social Research Council for their financial support through the Centre for Economic Performance. The empirical part builds heavily on joint work with Raffaella Sadun.

Phillipe Aghion is a Professor of Economics at Harvard University. Nicholas Bloom is an Associate of the Centre for Economic Performance, London School of Economics and Political Science and Professor of Economics at Stanford University. John Van Reenen is Director of the Centre for Economic Performance and Professor of Economics, London School of Economics and Political Science.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Occasional Paper should be sent to the editor at the above address.

1. Introduction

Grossman and Hart (1986) developed the incomplete contracts approach to analyze the costs and benefits of vertical integration, which could explain why firms have boundaries, and why not all transactions are taking place within a single firm. The basic idea is that contracts cannot specify all states of nature or all actions in advance, or there are states of nature or actions which cannot be verified *ex post* by third parties, and which therefore are not *ex ante* contractible. They used this approach to develop theories of ownership and vertical integration. When contracts cannot specify all possible uses of an asset, the contract *ex ante* must leave some discretion over the use of the assets: in other words it must allocate ownership of the asset to one or the other party. The benefit of integration is that the owner avoids hold-up by the other party, which in turn will enhance her incentives to invest in the relationship. The cost is that the other party will tend to under-invest in the relationship.

In this paper we show how the incomplete contract approach can be used to think about the *internal organization* of firms. The first half has a theoretical focus; we look at how formal and real authority are allocated between the firm's owner and its employees, between the top managers and sub-ordinates, and between firms and financiers. We also examine how the allocation of authority affects communication within the firm.

In the second half of the paper we analyze some of the empirical literature, examining first the determinants of organization (focusing on delegation/decentralization) from the perspective of the Grossman-Hart approach and its extensions, and second the effects of decentralization on firm performance. The emphasis is on looking at "stylized facts" from large-scale econometric studies of firms, rather than at case studies. Case studies are helpful in suggesting theoretical approaches and mechanisms, but are poor for hypotheses testing as they are small in number and highly selected. As with the theory, we also look at decentralization *within* firms, rather than the more commonly studied issue of boundaries of the firm or vertical integration which was the original motivation for Grossman and Hart (1986).¹ We find some support for aspects of the incomplete contracts approach in the importance of the congruence of preferences and firm heterogeneity for decentralization. However, we acknowledge that there are many other stylized facts from the empirical decentralization literature that may require alternative theoretical perspectives.

The structure of this paper is as follows. Section 2 examines theory. We look at delegation and authority (2.1), financial contracting (2.2), and delegation and the informational content of decision making (2.4).² Section 3 examines the empirical determinants of firm decentralization, focusing on some predictions of the theory (such as the importance of preference congruence as proxied by trust). Section 4 analyzes the effect of decentralization on firm performance. In Section 5 we conclude and suggest areas for future research.

¹See Lafontaine and Slade (2008) for a comprehensive review of the empirical literature on this.

²In Aghion et al. (2013) appendices A and C we also discuss work on academia and multi-divisional firms.

2. Theory

2.1. A model of delegation and authority in organizations

2.1.1. Basic model

Real authority, i.e. the ability to make decisions, requires information. But acquiring information in turn requires time and effort. Thus, for example, the CEO of a big holding company that consists of several horizontally integrated units, can only devote limited attention to each unit, which in turn implies that more real authority will lie with downstream agents in each unit. In fact, increasing the “span of control” is one way in which a top manager can commit to leave more real authority, and therefore more initiative, to her subordinates in various branches of activity. More generally, it is the design of the organization, together with the allocation of formal decision rights, that will determine how real authority is distributed within the firm.

The issue of real versus formal authority and of the implications of this distinction for the optimal design of firm organization, is addressed by Aghion and Tirole (1997) using an incomplete contracts/property rights approach.

Their basic framework involves two parties: P(principal) and A(agent). It is assumed that *formal authority* can be allocated contractually (e.g. shareholders allocate authority to the board of directors). In turn, boards allocate authority to management – and management to different layers of management, and so on. By contrast, *real authority* is exerted by the party which has information; this may be the party with formal authority, but not necessarily so. Contractual incompleteness is again key to the whole analysis: contracts signed ex ante between P and A cannot specify particular project choices, as these are not verifiable by third parties.

After the contract is signed, both P and A can invest in information acquisition: by investing effort $\frac{1}{2}E^2$, P acquires information with probability E . Similarly, by investing effort $\frac{1}{2}e^2$, A acquires information with probability e . An important parameter in the analysis of the costs and benefits of delegating formal (or real) authority to A is the degree of congruence between P’s and A’s preferences. Let α denote the probability that P’s preferred project is also A’s preferred project (call this congruence between the two parties’ preferences), and suppose that a party gets zero utility if the other party chooses her preferred project and preferences are not congruent. Finally, assume that an uninformed party will never pick a project at random as this might be too risky.

The timing of moves is as follows. First, the two parties sign a contract that allocates formal authority to one party, either P or A. Then, both parties invest in information acquisition, i.e. P and A choose E and e respectively. Then, if she acquires information, the party with no formal authority proposes a project to the party with formal authority. The party with formal authority then either picks her preferred project (if she herself has acquired information) or she picks the project proposed to her by the party without formal authority if she did not acquire information. It is in this latter case that *real authority* differs from *formal authority*, since the project is actually chosen by the party *without* formal authority (the party with formal authority is uninformed and therefore can only rubberstamp the other party’s project proposal).

For P to delegate formal authority to her agent A, involves a cost and a benefit. The cost is that the agent may choose a project which the principal does not prefer. This is the *loss*

of control effect. The benefit is that delegating formal authority to the agent encourages her to invest more effort (i.e. higher e) in information acquisition. This is the *initiative* effect. Which effect dominates, in turn depends upon the congruence parameter α : there exists a cut-off value α^* such that for $\alpha < \alpha^*$, the first effect dominates, and it is better for P to retain formal authority, whereas for $\alpha > \alpha^*$, the second effect dominates and it is better for P to delegate formal authority to A.

Since preference congruence turns out to be so critical to decentralization decisions, we examine this in the empirical section where there does appear to be some compelling evidence for the importance of empirical proxies for congruence, such as *trust*. Below we show that this is a robust prediction of generalizations to the basic theoretical approach.

2.1.2. Extensions

Subsequent to Aghion and Tirole (1997), AT, several papers have analyzed the allocation of formal authority internally in organizations. We shall describe some of these attempts in the next subsections. At this stage, let us mention a first attempt by Hart and Moore (1999), HM. HM analyze the optimal allocation of authority in multi-layer hierarchies. Their model is one where by assumption, upstream agents are less likely to get new ideas (getting an idea in HM is like obtaining information in AT) due to their higher span of control. However, when they do get an idea, this idea has higher potential because of their greater span. HM then show that it is optimal to have “chains of commands”, whereby whenever they have an idea, upstream agents (the “generalists”) have priority rights over implementing the idea; only if they don’t have an idea can downstream agents (the “specialists”) have their say on which action to implement. The intuition is that although upstream agents are less likely to have a new idea, having priority control rights makes sure that they are in control of all the assets downstream, which in turn allows them to fully realize the idea’s potential. But if they fail to have a new idea, then the next downstream agent on each branch of the hierarchy should have her say if she gets an idea, and so on, moving down in the hierarchy.

So far, we have concentrated on the allocation of formal authority within organizations. However, going back to AT, it could be that delegating formal authority to A is too costly to P, for example because with arbitrarily small probability, A might take some very costly action. In that case P will always want to retain formal authority, but yet she may want to commit herself not to invest too much in information acquisition, so as to preserve the A’s incentives to invest in e even though she keeps formal authority with herself. One way to achieve such commitment is through the choice of *span of control*. More specifically, by increasing the span of control, i.e. the number of agents and activities under her supervision, the principal will commit to limiting how much effort she devotes to acquiring information on each particular activity. This in turn will encourage initiative by agents on each activity, as they anticipate that the principal will ignore their proposals less often (as she will not have acquired the relevant information). The choice of the optimal span of control by the principal at date zero is in turn subject to the same trade-off between the principal’s loss of control and the agents’ initiatives as above. This trade-off also underlies other features of organizational design, such as the role of intermediaries, the costs and benefits of having multiple principals on some activities, or the optimal combination of tasks within teams.

The idea that the design of organizations can serve as a commitment device to delegate (real) authority, is further explored by Acemoglu et. al. (2007), who use it to test their theory

of the determinants of decentralization on French and British firm-level panel data (see Section 3 for more empirical detail). The model, closely related to AT, is one in which the owner of a firm in a given sector can learn about the outcome of an investment decision through observing other firms in the same sector, or by relying on the superior information of downstream agents (or on downstream agents' effort to acquire information) within the firm. The more precise the public information acquired through observing other firms in the same sector, the less a firm needs to delegate control to its better informed agent. This simple observation delivers a rich set of predictions on the determinants of decentralization. In particular it suggests that older firms should delegate less as these firms will have had time to learn from more predecessors in the same sector. It also suggests that the greater the firm heterogeneity in the same sector, the more a firm in that sector should delegate control as what it observe from other firms is less likely to be relevant for its own choices (if firms are very different it is harder for them to learn from each other). Finally, it suggests that a firm closer to the technological frontier should delegate control more, as it is more likely to face problems that have not been solved before by other firms in the same sector. In the empirical section below we discuss the empirical tests of these three predictions and show that they all receive support in the data.

Even if decentralization was the efficient choice due to characteristics of the firm's environment, Baker et. al. (1999) emphasize that delegation is often informal because the corporate head quarters (CHQ) must usually sign-off on decisions. The issue is whether the CHQ credibly commits to allowing the plant manager to effectively make important decisions and does not override the plant manager (in order to establish her reputation not to interfere). Thus, the extent of decentralization is the outcome of a repeated game between the CHQ and manager. Again, trust may facilitate a cooperative outcome to this repeated game, which suggests that regions or countries which have higher levels of trust should enjoy greater firm decentralization.

Another reason why delegating formal authority within firms may be difficult is explained by Bolton and Dewatripont (2011): nothing prevents a principal (the owner of a firm) who delegates formal authority to an agent to revert her decision at any time. Bolton and Dewatripont point to the so-called Business Judgment Rules, which prevent courts from enforcing contracts between several parties within the same firm. On the other hand, there are instances where transfer of formal authority can be enforced nevertheless. First, if this transfer is accompanied by a transfer of information from the principal to the agent, then this information transfer guarantees irreversibility of the transfer of control (see Aghion et al., (2004) for a more detailed discussion of this point). Second, in a more dynamic context, principals may want to establish a reputation for not reverting control allocation decisions over time, precisely to keep the option of credible control transfers in the future. Third, as stressed by Hart and Holmstrom (2010), taking control back from an agent causes the agent to become aggrieved which in turn may induce the agent to "shade", i.e. to take unobservable actions which are damaging to the firm. Finally, as pointed out by Bolton and Dewatripont (2011), there are at least two examples of contracting situations where Business Judgment Rules do not apply: financial contracting, i.e. contracts between a firm and its investors, which we discuss in the next sub-section, and universities, where the faculty are protected by contracts and tenure commitments which grants them academic freedom (see Aghion et al., 2008 and Appendix A of Aghion et al. 2013).

2.1.3. Summary and empirical implications

There are many other important theoretical aspects of firm delegation stemming from GH and AT. The next two sub-sections will focus on finance and information, as space constraints prevent us from elaborating on other important areas such as academia and multi-divisional firms (the Appendices of Aghion et al., 2012, summarize these contributions).

As a general point, although decentralization models in the GH-AT tradition generate a rich set of predictions, it is fair to say that relatively few of these have been subject to rigorous empirical examination. One main reason for this is that it is hard to develop empirical analogs of theoretical objects such as decentralization, information and communication in one firm, let alone in large-scale databases suitable for econometric analysis. In Section 3 we focus on some areas where empirical progress has been made.

2.2. Financial contracting and the role of contingent control allocations

Since the work of Modigliani and Miller (1958)—which suggested that the mix of debt and equity a firm has does not affect its value—economists have wondered why most firms have some combination of debt and equity financing. Debt has certain tax advantages (interest is typically tax deductible for the firm but dividend payments are not), yet corporate debt was prevalent even before corporate income tax existed. Why would firms have debt, then?

Aghion and Bolton (1992) develop an incomplete contracts model which provides a rationale for holding debt. They argue that the mix of debt and equity financing divide up the states of the world where debt and equity holders have control of the firm’s assets. When times are good, and cash flows are sufficient to meet interest payments, equity holders have control. When times are bad, debt holders get control.

Suppose that an entrepreneur needs to finance a project that costs $K = 10$. She does not have private wealth and thus needs funding from an outside investor. The investor cares only for monetary benefits, whereas the entrepreneur only draws private benefits from taking various actions. Actions are not verifiable by a third party and therefore cannot be contracted upon ex ante. Hence, all the initial contract can do is allocate control rights between the two parties. The timing of the relationship between the entrepreneur and the investor can be described as follows. At the contracting stage, the two parties write a financial contract which allocates control rights. The contract must be “feasible,” i.e. it must satisfy the investor’s ex ante participation constraint (she must get at least as much as her outside option in expectation). Then the state of nature is realized. Suppose there is a good state θ_g and a bad state θ_b , each of which can occur with probability $1/2$, and which state is realized is verifiable by a third party. Then an action must be chosen. Suppose that only two actions can be chosen: a_1 and a_2 . Action a_1 maximizes monetary revenue in all states, whereas action a_2 maximizes private benefits in all states. For example, let $\pi(a, \theta)$ denote the monetary profit from taking action a in state θ , and $B(a)$ denote the private benefit from taking action a in any state θ and suppose that

$$\begin{aligned}\pi(a_1, \theta_g) &= 11, \pi(a_1, \theta_b) = 13 \\ \pi(a_2, \theta_g) &= 10, \pi(a_2, \theta_b) = 6\end{aligned}$$

and

$$B(a_1) = 2 < B(a_2) = 4$$

The first best involves a_1 being chosen in state θ_b and a_2 being chosen in state θ_g , since

$$B(a_1) + \pi(a_1, \theta_b) = 15 > 10 = B(a_2) + \pi(a_2, \theta_b)$$

whereas

$$B(a_1) + \pi(a_1, \theta_g) = 13 < 14 = B(a_2) + \pi(a_2, \theta_g).$$

We then compare between three governance structures, which correspond to three types of financial contracts. Entrepreneur control (e.g. as implemented through issuing non-voting shares), would lead to action a_2 being chosen in all states, but this would violate the investor's participation constraint since

$$\frac{1}{2}(6) + \frac{1}{2}(10) = 8 < 10 = K.$$

The investor has cash which he might use to convince the entrepreneur to take action a_1 in state θ_b . However, if most of the bargaining power at the renegotiation stage lies with the entrepreneur, the prospect of ex post renegotiation will not help satisfy the investor's ex ante participation constraint. In this case, entrepreneur control is not ex ante feasible. How about investor control (e.g. as implemented through issuing voting equity)? In this case the investor will choose action a_1 in all states, even though action a_2 is the first best action in state θ_b , i.e. the action that maximizes the sum of monetary and private benefits in that state. Now, can the entrepreneur renegotiate the action from a_1 to a_2 in state θ_g ? The answer is no, simply because the entrepreneur has no cash she can use to bribe the investor into changing his choice of action. Investor control satisfies the investor's participation constraint, since

$$\frac{1}{2}(11) + \frac{1}{2}(13) = 12 > 10 = K.$$

However it is not first-best, as total surplus would be maximized by having action a_1 taken in state θ_b and action a_2 taken in state θ_g .

Now, consider a contract that specifies a contingent allocation of control – to the entrepreneur in state θ_g and to the investor in state θ_b . This contract will lead to action a_1 being taken in state θ_b and action a_2 being taken in state θ_g . Contingent control can in turn be implemented through a debt contract that transfers control from the entrepreneur to the investor in state θ_b .

The idea that contingent control can help align incentives goes beyond financial contracting. For example Bolton and Dewatripont (2011) restate the Aghion-Bolton model as one of a headquarter who commits to replace a divisional manager only in some states of nature (e.g., in state θ_b for the above notation). Under this reinterpretation $(B(a_2) - B(a_1))$ can be seen as the private cost for the division manager of closing down his unit and thus having his employees laid off, rather than maintaining the unit in operation. This contingent arrangement reduces the expected monetary losses of the headquarters and simultaneously internalizes the private costs of closing down units. In a context where the information about θ could be manipulated by the division managers, for example through account manipulation or through risky decision making (gambling for resurrection) so as to avoid being replaced, contingent control can help mitigate the problem by offering a guarantee to divisional managers. Another guarantee is to give managers a stake in the firm's profits even when they are being replaced (see Bolton and Dewatripont (2011) and Garicano (2000)).

2.3. Delegation as a way to improve the informational content of decision-making

Dessein (2002) analyzes how the allocation of control can help incorporate the agent's information into decision-making in a situation where the agent has private information. In contrast to Aghion and Tirole (1997), there is no information acquisition effort by the agent or the principal, therefore in Dessein's model the allocation of authority is not so much a tool to motivate the agent (as in Aghion and Tirole) or to give a supplier incentives to make relationship specific investments (as in Grossman and Hart). The main insight of Dessein (2002) is that in a world with asymmetric information and contractual incompleteness, the delegation of authority from a principal to an agent is often the best way to elicit the agent's private information.

In Dessein's setting the agent is assumed to be better informed, but with preferences over decisions that are not fully congruent with those of the principal. If the principal has authority (i.e. if she holds the decision rights), which Dessein refers to as "centralization", then the agent communicates his information "strategically" in order to tilt the principal's decision. Centralization thus results in information loss. However, while "delegating" control to the agent avoids this information loss, it also makes the agent's biased decision-making prevail. Delegation thus results in a loss of control. Dessein (2002) shows that for a broad range of parameters, the loss of control under delegation matters less than the loss of information under centralization. In particular, the smaller the agent's bias (i.e. the more congruent the principal's and agent's objectives are), or the larger the agent's informational advantage, or the more uncertainty there is, the more likely it is that delegation is optimal.

An important assumption in Dessein (2002) is that under "principal authority", the principal cannot commit to not taking the decision which she believes maximizes her expected utility. She can, however, commit to "delegate" control rights to the agent. This is consistent with the incomplete contracting assumption that actions or decisions are non-contractible, even when control allocation is contractible. Delegation can then be interpreted as a commitment device from the principal to use the agent's information in the way that best fits the agent's objectives. While delegation results in biased decision-making (this is the *loss of control* effect) it ensures that decision-making responds more to the agent's information. Compare this with centralized authority, which ensures an unbiased decision from the principal's point of view but encourages the agent to distort his information in order to influence the principal.

Information communication is modeled using Crawford and Sobel (1982)'s model of strategic communication. Since decisions are non-contractible and information is assumed to be soft, communication between the agent and principal takes the form of "cheap talk". However, strategic communication can be informative if there is sufficient preference alignment between the agent (the sender) and the principal (the receiver). Dessein's contribution can be seen as one of introducing control-rights considerations in strategic communication games, with the idea that delegation induces communication by the agent to become less "strategic", although at the cost of non-congruent decision-making.

The result that delegation dominates when the agent's bias is small is not fully obvious: communication between the agent and principal is also better under centralization in that case, as the agent gains less by distorting information when her preferences are more congruent with those of the principal. Yet Dessein (2002) shows that under general conditions, as long as the bias of the agent is sufficiently small, delegation is always strictly preferred over centralization. For some specific cases, such as the leading example in Crawford and Sobel

(1982), Dessein (2002) obtains the striking result that communication never takes place at the optimum: as long as preferences are sufficiently congruent, the principal is strictly better off avoiding communication (and the resulting information distortions) altogether, by committing to delegate control to the agent.

Overall, this paper provides an important rejoinder to the central insight of the property rights literature that the cost of integration (or centralization) are the reduced incentives to either make relationship-specific investments (Grossman and Hart) or to acquire information (Aghion and Tirole). In Dessein (2002), the cost of centralization (e.g. through integration), is the distortion and loss of information.

3. Empirical evidence on the determinants of decentralization

Recent empirical interest in decentralization has been stimulated by the growth of a substantial body of evidence that documents persistence performance differences among firms, even in narrowly defined industries (see Syverson, 2011; Gibbons and Roberts, 2012 or Aghion et al., 2013, Appendix C). Could these differences be due to different organizational structures within firms? This section is organized as follows. Sub-section 3.1 focuses on the measurement of firm decentralization, 3.2 on the impact of trust on decentralization, 3.3 on empirical implication of learning models for decentralization and 3.4 documents more general “stylized facts” of decentralization. We formally discuss identification issues in more detail in Appendix D of Aghion et al., 2013.

3.1. Measuring firm decentralization

A key factor in any organization is who makes decisions. A centralized firm is one where decisions are all taken at the top of the hierarchy, and a decentralized firm is where decision-making is more evenly dispersed throughout the corporate hierarchy.³

How can the concept of decentralization be implemented empirically? One way is to look at the organization charts of firms (“organogram”) as graphical representations of the formal authority structure. One of the best studies in this area is Rajan and Wulf (2006) who use the charts of over 300 large US corporations during 1987-1998 to examine the evolution of organizations (e.g. the number of people who directly report to the CEO as a measure of the span of control). Unfortunately, as Max Weber and (more recently as discussed above) Aghion and Tirole (1997) stressed, formal authority is not the same as real authority; the organogram may not reflect where power really lies.⁴

Observing whether a firm is decentralized into profit centers is useful, as this is a formal delegation of power - the head of such a business unit will be judged by the CEO on the basis of the unit’s profitability. Similarly if the firm is composed of cost (or revenue) centers this indicates less decentralization as only costs (or revenue) are likely to be under control of the

³An extreme case of decentralized “organization” is an idealized market economy where atomistic individuals make all the decisions and form spot contracts with each other. The origin of many of the debates on decentralization have their origins in the 1930’s, over the relative merits of a market economy relative to a centrally planned one.

⁴We focus on decentralization as distinct from managerial spans of control. These are distinct concepts – the span and depth (number of levels) of a hierarchy are compatible with different power relationships between the levels. Nevertheless, there is some evidence that the move towards delayering over the last twenty years has been associated with decentralization (see Rajan and Wulf, 2006).

manager. If the firm does not even delegate responsibility to a cost or revenue center this indicates a very centralized company. Acemoglu et al. (2007) use this distinction to classify firms in their empirical work on French and British firm panel data.

Whether a company is organized into profit centers is a rather crude indicator of decentralization. A better (but more costly) approach is to directly survey firms. Bloom et al. (2012b) measure decentralization between the plant manager and the corporate central headquarters (CHQ). They asked plant managers about their decisions over investment (maximum capital investment that could be made without explicit sign off from central headquarters), hiring, marketing and product introduction (the latter three on a scale of 1-5). This was conducted at the same time as the Bloom and Van Reenen (2007) management survey.

Bloom et al. (2012b) constructed an empirical summary of decentralization combining these four measures into a single index by z-scoring each individual indicator and z-scoring the average (so that the index has zero mean and unit standard deviation). We followed their approach but combine their data with additional waves of the World Management Survey conducted in 2007 and 2009 covering eight more countries.

Figure 1 shows that decentralization varies substantially across countries, with the US, the UK and Northern European countries being the most decentralized and Southern European and Asian countries the least. There is an even wider spread of decentralization across firms *within* every country, as shown in Figure 2.

Decentralization extends beyond just plant managers and the CHQ of course. For example, one can also consider the autonomy of workers from the plant manager following Bresnahan et al. (2002). Proxies for this dimension of decentralization include questions indicating greater worker control over the pace of work and over the allocation of tasks.

3.2. Trust (congruence of preferences)

The Aghion-Tirole approach offers a natural implementation of incomplete contracts to study decentralization within firms. One key parameter in fostering decentralization is the congruence of preferences between principal and agent. Other theoretical developments following Grossman-Hart such as Baker, Gibbons and Murphy (1999) also support this idea. Finding empirical proxies for the congruence parameter is challenging, but one possibility is to use measures of trust. In recent years, economists have started to take cultural factors more seriously in determining economic outcomes (Guiso et al., 2006; Greif, 1994). This stems in part from the influence of Putnam's (1993) work on the importance of social capital. Empirically, generalized social trust as a proxy for social capital has been found to be associated with many positive economic outcomes (e.g. see Knack and Keefer, 1997, on trust and growth or Guiso et al., 2009, on trust and foreign trade and investment).

Bloom et al. (2012b) examine the importance of trust, finding that higher levels of trust in the region where a plant is located is associated with a significantly greater degree of decentralization. As in many other papers, trust is measured using the standard indicators from the World Values Survey which asks random samples of individuals about generalized trust. These external measures are matched to plant locations to obtain a measure of trust in the region. The authors also exploit the fact that their data contains many subsidiaries of multinational firms to construct measures of trust in the country of origin (the multinational's headquarters) and the country of location (where the affiliate is located). Both of these seem to matter for decentralization, but the most powerful factor is the *bilateral* trust between

country pairs, i.e. the degree to which people from the subsidiary’s parent country trust people in the country where the plant is located. Multinationals locating in countries that are seen to be relatively highly trusted (after country location and origin dummies are removed), are significantly more likely to decentralize. In other words, even though the UK is overall a relatively high trust country, multinationals headquartered in the US (where Britain has a relatively good trust reputation) tend to decentralize more towards their British affiliates than equivalent multinationals headquartered in France (where Britain has a relatively bad trust reputation for historical reasons).

These results suggest that trust can affect the internal structures of global firms and that some aspects of organization are transplanted abroad, as suggested by recent theories of international trade (e.g. Helpman et al., 2004). It fits well the idea that the congruence of preferences is a major determinant of delegation.

Enforcement of contracts should further foster decentralization, and we do in fact observe more delegation where there is stronger rule of law⁵. However, contracts are never perfectly enforceable, which leaves a role for trust to help generate more delegation. Recently, Bloom et al. (2012) have been running field experiments on firms in India, and discovered that family size (in particular the number of adult male family members) was the key determinant of firm size, probably due to the importance of trust. Owners only trusted other male family members to make major managerial decisions as they worried that outsiders would steal from the firm. Hence, the supply of (trusted) male family member time was typically the binding factor for firm growth.

3.3. Learning

In the theory section we discussed an extension to the Aghion-Tirole approach when considering how firms could learn either from others or from themselves. Acemoglu et al. (2007) examine three predictions from their model (i) delegation should be greater when the industry is more heterogeneous (so it is harder to learn from others); (ii) the firm is close to the frontier (so that there are fewer other firms to learn from) and (iii) the firm is younger (so it has less experience to learn from its own mistakes). Acemoglu et al. (2007) measure decentralization using both formal measures of whether firms are organized into profit centers as discussed above (for French firms) and direct survey measures of the power managers have over hiring decisions (for British firms). In both samples they find decentralization is more likely in industries that are more heterogeneous and for firms that are younger or closer to the technological frontier.

These results are illustrated in Figures 3-5 where the y-axis has the average degree of decentralization in different bins of the relevant variables. In Figure 3 there is a reasonably clear positive relationship after the second decile between decentralization and heterogeneity (as measured by the dispersion of firm productivity growth). In Figure 4 decentralization appears to be higher among firms closer to the technological frontier (as measured by the distance of the firm’s productivity from the leading firm in the four digit industry). That more frontier firms should delegate more may also explain why subsidizing higher education, in particular graduate education, is more likely to be growth-enhancing if universities are more decentralized (see for example Aghion et al., 2010). Figure 5 shows older firms to look more

⁵See Bloom et al., (2009). More generally on the importance of law and reputation on contract enforceability see MacLeod (2007).

centralized than younger firms, possibly because they have learned better what to do, so there is less need of delegation to a local manager who is better informed but may not pursue the principal's interests.

3.4. Other empirical factors influencing decentralization

The development of the incomplete contracts approach as applied to firm decentralization appears to have some confirmation in the data. The congruence of preferences (as proxied by trust) and heterogeneity (making it harder to learn) both seem to foster decentralization. There are many other findings in literature on the empirical determinants of decentralization. These are not so obviously implications of the Grossman-Hart approach, but it is worth considering them.

3.4.1. Firm size and scope

Some basic factors determine decentralization. All else equal, a larger firm will require more decentralization than a small firm. A sole entrepreneur does not need to delegate because she is her own boss, but as more workers are added, doing everything herself is no longer feasible. Penrose (1959) and Chandler (1962) stressed that decentralization was a necessary feature of larger firms, because CEOs do not have the time to take every decision in large firms (see also Geanakoplos and Milgrom, 1991). Similarly as firms expand in their scope both geographically and in product space, local information will become more costly to transmit so this will also favor decentralization

Most empirical findings support this. Bloom et al. (2012b) find that firm size and plant size are both associated with a significant increase in their decentralization index. Furthermore, plant managers in subsidiaries of foreign multinationals have more autonomy than similar plants of domestic non-multinationals. They interpret this as an indicator that managing at a distance is harder, inducing headquarters to give more autonomy to local managers.

3.4.2. Human capital

Many models would predict that human capital should be associated with decentralization. For example, more skilled workers will have greater ability to take on more responsibility. When the environment changes due to new technologies and organizational change is required, skilled workers may be better at learning how to cope with the new organizational structure.

There is generally a robust and positive association of decentralization and skills. For example, Bloom et al. (2012b) measure skills by the proportion of people who hold a college degree and find this to be significantly correlated with decentralization. Caroli and Van Reenen (2001) examine the relationship between skills and organization in some detail, arguing in favor of "skill biased organizational change": i.e. increases in the supply of human capital will tend to increase delegation. To tackle the endogeneity problem, they use information on the differential price of skilled vs. unskilled labor in the local market (as indicated by the wage differential between college educated workers and other individuals). They argue that this skill premium is partially driven by exogenous shifts in the supply of unskilled workers. For their sample of UK and French firms they find that regions where skill premia are higher have a lower probability of decentralization.

3.4.3. Information and Communication Technologies (ICT)

Garicano (2000) formalizes the idea of the firm as a cognitive hierarchy. There are a number of problems to be solved and the task is how to solve them in the most efficient manner. The simplest tasks are performed by those at the lowest level of the hierarchy, and the “exceptional” problems are passed upwards to an expert. The cost of passing problems upwards is that communication is expensive. The benefit of passing the problem upwards is that it reduces the cognitive burden on lower level employees.

This framework was designed to address the impact of ICT on firm organization. Interestingly, information technologies have different implications for decentralization than communication technologies. Consider again the decentralization decision between the CHQ and plant manager. When communication costs fall through (for example the introduction of e-mail or company intranets), it is cheaper for the plant manager to refer more decisions to the CHQ. So communication technologies should cause *centralization*. By contrast, technologies that make it easier for the plant manager to acquire information (e.g. Enterprise Resource Planning software, ERP like SAP) means that *decentralization* should increase. An example in law firms would be Lexis Nexis that enables junior lawyers to quickly find relevant cases without consulting a more senior associate or partner.

Bloom et al. (2009) test this theory and find considerable empirical support. Computer networks (reducing communication costs) significantly decrease decentralization to plant managers whereas tools to help managers access more information (like ERP) significantly increase decentralization.⁶

3.4.4. Product market competition

Some authors such as Acemoglu et al. (2007) argue that a cause of the aggregate increase in more decentralized organizations is rapid technological change. An alternative explanation is that globalization and deregulation have increased the degree of product market competition which has in turn stimulated organizational change. The theory is ambiguous here. If competition has made swift decisions more important, then this will have increased the salience of local knowledge, leading to greater decentralization under the framework discussed above (e.g. Aghion and Tirole, 1997). Similarly if competition aligns the incentives of agents more with the principal then the costs of decentralization may also have fallen. There are countervailing forces, however. For example, a larger number of firms in an industry aids yardstick competition, but it may also help learning in the Acemoglu et al. (2007) framework, which will reduce the need to decentralize.

The empirical evidence, however, seems more clear cut. Bloom et al. (2010) find a robust positive association between competition and decentralization using industry import competition, the inverse of the industry Lerner index, or simply the number of perceived competitors. A similar positive correlation was reported in Acemoglu et al. (2007) and Marin and Verdier (2008, 2009). All of these are cross sectional studies, so the positive coefficient on competition could simply reflect unobserved variables correlated with competition. Guadalupe and Wulf (2009) try to tackle this endogeneity problem using the Rajan and Wulf (2006) panel dataset

⁶The magnitude of the effect is substantial. An increase in ERP usage by 60% (the average difference in ICT between Europe and the US) increases a plant manager’s autonomy by an amount equivalent to the increase in US college graduates between 1990 and 2000.

on the changing organizational structure of firms over time. They argue that the Canada-US Free Trade Agreement in 1989 constitutes an exogenous increase in competition for US firms in industries where tariffs were removed. Exploiting this policy experiment reveals that competition is associated with their proxy for decentralization.

4. Organizational practices and firm productivity

How can researchers identify the effects of organizational structure (e.g. decentralization) on firm performance?⁷

4.1. Correlations of performance and organizational practices: the basic identification problem

Consider the basic production function as:

$$q_{it} = \alpha_l l_{it} + \alpha_k k_{it} + a_{it} \quad (4.1)$$

Where q is $\ln(\text{output})$, l is $\ln(\text{labor})$ and k is $\ln(\text{capital})$ of firm i at time t . Assume that we can write the TFP term a_{it} as

$$a_{it} = \alpha_0 + \beta m_{it} + u_{it} \quad (4.2)$$

where m_{it} is an organizational feature of the firm (such as decentralization) and u_{it} is an unobserved error. Together these equations imply:

$$q_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \beta m_{it} + u_{it} \quad (4.3)$$

This of course contains several assumptions. It assumes that the relevant organizational factor enters linearly, whereas organization could instead be affecting the coefficients on the other factor inputs, and many theories (e.g. of complementarity) would generate the same predictions. We discuss these below. We will assume that we can deal with the econometric problems in estimating the coefficients on the production function so that we have a consistent measure of total factor productivity (see Akerberg, Benkard, Berry and Pakes, 2007, for a discussion of recent contributions). Note that OLS estimates of (4.1) will generally be biased, as $E(m_{it}u_{it}) \neq 0$.

The traditional strategy is to assume that m_{it} is a firm fixed effect. So one approach is simply to recover average firm TFP under this assumption and project it on some cross-sectional measure of management m_i . This will indicate whether there is an association between the two measures, but not whether the relationship is causal. For example, Bloom and Van Reenen (2007) show that there is a robust relationship between TFP and their measure of management quality, but they interpret this as an “external validity” test of the quality of the management data rather than as any causal relationship.

An analogous strategy if there are time varying measures of organization is to treat all the correlated unobservables as fixed, i.e. $u_{it} = \eta_i + \varepsilon_{it}$ with $E(m_{it}\eta_i) \neq 0$ but $E(m_{it}\varepsilon_{it}) = E(m_{it}\varepsilon_{it-1}) = 0$. Then the fixed effect model estimated in (say) first differences would be

$$\Delta a_{it} = \beta \Delta m_{it} + \Delta \varepsilon_{it} \quad (4.4)$$

⁷Aghion et al (2013, Appendix D) discuss general identification issues in this literature in more detail.

which can be consistently estimated by OLS.

There are a huge number of studies that have correlated various aspects of the firm’s performance on various aspects of its organizational form (e.g. the survey in Lazear and Oyer, 2012). The better studies use micro data and pay careful attention to measurement issues and need to control for many covariates. For example, Cappelli and Neumark (1999) and Black and Lynch (2001) examine various aspects of “high performance” workplaces, mostly relating to employee involvement, team work and meetings. Both papers look across many industries and find no direct effect of these measures on performance (in contrast to many case studies). As we discuss below, Ichniowski, Shaw and Prenushi (1997), however, examined management practices and performance in 37 US steel mills over time and found a link between upgrading to bundles of modern practices and improved performance, so the correlation evidence is mixed.

There remain several serious problems. First is the data constraint that measuring organization is hard and finding data with time series variation even harder. Second, the management proxies are measured with error, so this will cause attenuation towards zero if the measurement error is classical. This bias is exacerbated in first differences. Third, and most seriously, the factors that cause variation in the propensity to adopt organizational practices will also likely be correlated with those affecting TFP so the assumption is unlikely to hold in most cases. The bias could be upwards or downwards (e.g. if firms doing badly are more likely to innovate organizationally as argued by Nickell et al., 2001).

There is no simple solution to these endogeneity problems as we fundamentally need some exogenous identifying variation. Bloom et al. (2012) implemented a gold standard randomized control trial in Indian textile plants. The intervention was by high quality management consultants to improve a range of management practices (as in Bloom and Van Reenen, 2007) which appeared to dramatically improve productivity. Most of the quasi-experiments have been in labor economics. A good example is Lazear (2000) who looked at the introduction of a pay for performance system for windshield installers in the Safelite Glass Company. Lazear found that productivity increased by around 44%, with about half of this due to selection effects and half from the same individuals changing behavior. More recently, Bandiera et al. (2007, 2009) engineered a change in the incentive pay system for managers in farm. They have no contemporaneous control group, but can examine the behavior of individuals before and after the introduction of the incentive scheme. Productivity rose by 21% mainly with at least half due to improved selection (the managers allocated more fields to the ablest workers rather than to their colleagues).

4.2. Complementarities between organizational practices

One of the key reasons why firms may find it difficult to adjust their organizational form is that there are important complementarities between sets of organizational practices. Milgrom and Roberts (1990) build a theoretical structure where such complementarities (or more precisely, super-additivities) mean that firms optimally choose clusters of practices that “fit together”. When the environment change so that an entrant firm would use this group of optimal practices, incumbent firms will find it harder – they will either switch a large number of practices together, or none at all.

This has important implications for productivity analysis. The effects of introducing a single practice will be heterogeneous between firms and depend on what practices they already use. This implies that linear regressions of the form of equation (4.3) may be misleading. To

see this, consider two practices, m^1 and m^2 , whose relationship with productivity is such that TFP increases only when both are used together.

$$a_{it} = a_0 + \beta_1 m_{it}^1 + \beta_2 m_{it}^2 + \beta_{12} (m_{it}^1 \times m_{it}^2) + u_{it} \quad (4.5)$$

One version of the complementary hypothesis is $\beta_1 > 0$, $\beta_2 < 0$, and $\beta_{12} > 0$, i.e. the disruption caused by just using one practice (m_{it}^2) could actually reduce productivity. A regression which omits the interaction term may $m_{it}^1 \times m_{it}^2$ find only a zero coefficient on the linear terms.

The case study literature emphasizes the importance of complementarities. Testing for their existence poses some challenges, however, as pointed out most clearly by Athey and Stern (1998). A common approach is a regression of practice 1 (m_{it}^1) on practice 2 (m_{it}^2) with a positive covariance (conditional on other factors) indicating complementarity. It is true that complements will tend to covary positively, but this is a very weak test. There could be many other unobservables causing the two practices to move together. We need an instrumental variable for one of the practices (e.g. Van Biesebroeck, 2007), but this is hard to obtain as it is unclear what such an instrument would be, i.e. could it be legitimately excluded from the second stage equation? In classical factor demand analysis we would examine the cross price effects to gauge the existence of complements versus substitutes, i.e. does demand for practice 1 fall when the price of practice 2 rises (all else equal). There still remains the concern that the price shocks could be correlated with the productivity shocks, but such an assumption is weaker than assuming unobserved shocks to the firm's choice of practices are uncorrelated. Unfortunately, such tests are particularly hard to implement because there are generally no market prices for the organizational factors typically considered.

An alternative strategy is to work straight from the production function (or performance equation more generally). Consider the productivity equation after substituting in multiple practices:

$$q_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \beta_1 m_{it}^1 + \beta_2 m_{it}^2 + \beta_3 (m_{it}^1 \times m_{it}^2) + u_{it} \quad (4.6)$$

Ichniowski et al. (1997) estimate a version of equation (4.6) using very disaggregate panel data on finishing lines in US steel mills, using eleven human resource practices (including incentive pay, recruitment, teamwork, job flexibility and rotation). Their measure of productivity is based on downtime - the less productive lines were idle for longer. They find that introducing one or two practices has no effect on productivity, but introducing a large number together significantly raises productivity. Although the endogeneity problem is not eliminated, the controls for fixed effects, looking within one firm and using performance data, helps reduce some of the more obvious sources of bias.

4.3. The role of ICT again

One of the key productivity puzzles of recent years has been why the returns to the use of information and communication technologies appear to be so high and so heterogeneous between firms and between countries. For example, Brynjolfsson and Hitt (2003) find that the elasticity of output with respect to ICT capital is far higher than its share in gross output (see also Stiroh, 2004). One explanation for this is that effective use of ICT also requires significant changes in firm organization. Changing the notation of (4.6) slightly, we could write

$$q_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \beta_c c_{it} + \beta_m m_{it}^2 + \beta_{cm} (c_{it} \times m_{it}) + u_{it} \quad (4.7)$$

with the hypothesis that $\beta_{cm} > 0$. This is broadly the position of papers in the macro literature explaining the faster productivity growth of the US than Europe after 1995 (e.g. Jorgenson et al., 2008).

Bresnahan et al. (2002) try to test this directly by surveying large US firms on decentralization and team work (for a cross section) and combining this with data on ICT (from a private company Harte-Hanks), and productivity from Compustat. They find evidence that $\beta_{cm} > 0$, i.e. that computer capital is more productive when firms have greater decentralization and team work. Bloom et al. (2012a) broaden the sample to cover both the US and firms in seven European countries, and find evidence of complementarity of ICT with people management. They also show that their results are robust to controlling for firm fixed effects. Careful econometric case studies (e.g. Baker and Hubbard, 2004; Bartel et al., 2007) also identify differential productivity effects of ICT depending on organization form. Lemieux et al. (2009) show that one particular people management practice, performance pay, is becoming increasingly important in the US and has a significant impact on widening inequality. They suggest that the spread of ICT innovations has facilitated the adoption of performance pay techniques.

4.4. The role of human capital

One of the reasons for the renewed interest in organizational change by labor economists was the attempt to understand why technology seemed to increase the demand for human capital, thus contributing to the rise in wage inequality experienced by the US, UK and other countries since the late 1970s. Many theories have been proposed (see Autor et al., 2003, for a review), but one hypothesis is that lower IT prices increased decentralization for the reasons outlined in Garicano (2000), and decentralization leads to an increase in inequality (Garicano and Rossi-Hansberg, 2006). Further, decentralization is complementary with skills for at least three reasons. First, skilled workers are more able to analyze and synthesize new pieces of knowledge, so the benefits of local processing of information are enhanced. Additionally, skilled workers are better at communicating, which reduces the risk of duplication of information. Second, the cost of training them for multi-tasking is lower, and they are more autonomous and less likely to make mistakes. Finally, workers who are better educated may be more likely to enjoy job enrichment, partly because they expect more from their job in terms of satisfaction.

This has three main implications:

1. Decentralization leads to skill upgrading within firms. This is due to the fact that the return to new work practices is greater when the skill level of the workforce is higher.
2. A lower price of skilled labor will accelerate the introduction of organizational changes.
3. Skill intensive firms will experience greater productivity growth when decentralizing.

Caroli and Van Reenen (2001) find support for all three predictions. They estimate production functions (with the relevant interactions), skill share equations and organizational design equations. A novel feature of this approach is that because labor is traded in a market, it is possible to use local skill price variation to examine complementarity issues. They find that higher skill prices make decentralization less likely, consistent with “skill-biased organizational change”.

5. Conclusions

We began by surveying the theoretical literature on the organization of firms and the optimal decentralization of decision rights within firms. We discussed how the concept of incomplete contracts shapes this organization of decision making within firms. In particular, the inability to contract over all possible states of the world leads principals to delegate control to agents as a way to ex ante commit to letting agents expropriate some of the returns from costly activities, like collecting information on the best actions to take.

We then overviewed some of the empirical evidence on the organization and management of firms, focusing on decentralization. We looked at within firm organization, especially decentralization following the theoretical survey, and we also focus on econometric studies of large scale firm databases rather than case studies. There has recently been a number of papers measuring management and organizational practices across firms and countries. Like productivity, decentralization varies a lot across firms and countries (e.g. Scandinavian and Anglo-Saxon firms are more decentralized than those from Asia and Southern Europe). A number of factors highlighted in the incomplete contract theory are shown to be important for accounting for differences in firm organization - in particular heterogeneity and the congruence of preferences (as proxied by trust). Several other factors appear robustly positively correlated with decentralization, such as product market competition, human capital and firm size.

In terms of future work, we see two areas of opportunity. First, there is a need to match up empirical work more closely with the theory. Until recently, comprehensive datasets on measures of organization across many firms and countries were unavailable. Now that this gap in the core data infrastructure is being covered, there is a great opportunity for testing some of the theories of organizational economics. This is challenging, first because many of the important aspects in the environment emphasized in the theoretical literature are hard to match into empirical counterparts, and second because identifying the causal relationship between organizational changes and firm outcomes like productivity and growth is difficult. Even with strong measurement and tight links to theory, it is essential to identify the direction of causality in some of stylized empirical results we have identified above - for example, do more skilled managers enable decentralized decision making, or are skilled managers attracted to decentralized firms? We see the theory and empirics of the organization of the firm as one of the key growth areas in economics over the next twenty-five years.

References

- Aghion, P. and J. Tirole**, “Formal and real authority in organizations,” *Journal of Political Economy*, 1997, pp. 1–29.
- Aghion, P., Bloom, N. and Van Reenen, J.** (2013), “Incomplete contracts and the Internal Organization of Firms”, NBER working paper
- and **P. Bolton**, “An incomplete contracts approach to financial contracting,” *The Review of Economic Studies*, 1992, 59 (3), 473.
- , **M. Dewatripont, and P. Rey**, “Transferable control,” *Journal of the European Economic Association*, 2004, 2 (1), 115–138.
- Alonso, R., W. Dessein, and N. Matouschek**, “When does coordination require centralization?,” *The American Economic Review*, 2008, 98 (1), 145–179.
- Athey, S. and S. Stern**, “An empirical framework for testing theories about complementarity in organizational design,” 1998.
- Atkeson, A. and P.J. Kehoe**, “Modeling and Measuring Organization Capital,” *Journal of Political Economy*, 2005, 113 (5).
- Autor, D.H., F. Levy, and R.J. Murnane**, “The Skill Content of Recent Technological Change: An empirical exploration,” *Quarterly journal of economics*, 2003, 118 (4), 1279–1333.
- Baily, M.N., C. Hulten, and D. Campbell**, “Productivity dynamics in manufacturing establishments,” *Brookings Papers on Economic Activity: Microeconomics*, 1992, 4, 187–249.
- Baker, G., R. Gibbons, and K.J. Murphy**, “Informal authority in organizations,” *Journal of Law, Economics, and Organization*, 1999, 15 (1), 56.
- Baker, G.P. and T.N. Hubbard**, “Contractibility And Asset Ownership: On-Board Computers and Governance In US Trucking,” *Quarterly Journal of Economics*, 2004, 119 (4), 1443–1479.
- Bandiera, O., I. Barankay, and I. Rasul**, “Incentives for Managers and Inequality Among Workers: Evidence from a Firm Level Experiment,” *Quarterly Journal of Economics*, 2007, 122 (2), 729–773.
- , ———, and ———, “Social connections and incentives in the workplace: Evidence from personnel data,” *Econometrica*, 2009, 77 (4), 1047–1094.
- Bartel, AC, Ichiniowski C., and K. Shaw**, “How Does Information Technology affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills,” *Quarterly Journal of Economics*, 2007, 122 (4), 1721–1758.
- Bartelsman, E.J. and M. Doms**, “Understanding productivity: Lessons from Longitudinal Microdata,” *Journal of Economic literature*, 2000, 38 (3), 569–594.
- and **P.J. Dhrymes**, “Productivity dynamics: US Manufacturing Plants, 1972–1986,” *Journal of Productivity Analysis*, 1998, 9 (1), 5–34.
- , **J.C. Haltiwanger, and S. Scarpetta**, “Cross-country differences in productivity: the role of allocation and selection,” 2009.
- Biesebroeck, J. Van**, “Complementarities in automobile production,” *Journal of Applied Econometrics*, 2007, 22 (7), 1315–1345.
- Black, S.E. and L.M. Lynch**, “How to compete: the impact of workplace practices and information technology on productivity,” *Review of Economics and statistics*, 2001, 83 (3), 434–445.
- Bloom, N. and Sadun R. and Van Reenen, J.**, “Americans do I.T. better: US Multinationals and the Productivity Miracle,” *American Economic Review*, 2012a, 102 (1), 167–201

- Bloom, N. and Sadun R. and Van Reenen, J.**, “The organization of firms across countries” *Quarterly Journal of Economics*, 2012b, 127(4): 1663-1705
- , **Genakos, C. Sadun, R. and Van Reenen, J.** “Management Practices across firms and countries” *Academy of Management Perspectives* 26 (1) 12-33
- and **J. Van Reenen**, “Measuring and Explaining Management Practices across Firms and Countries,” *Quarterly Journal of Economics*, 2007, 122 (4), 1351–1408.
- , **B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts**, “Does management matter? Evidence from India,” forthcoming *Quarterly Journal of Economics*, 2012.
- , **C. Genakos, R. Martin, and R. Sadun**, “Modern Management: Good for the Environment or Just Hot Air?,” *The Economic Journal*, 2010, 120 (544), 551–572.
- , **C. Propper, S. Seiler, and J. Van Reenen**, “The impact of competition on management quality: evidence from public hospitals,” Technical Report, National Bureau of Economic Research 2010.
- , **L. Garicano, R. Sadun, and J. Van Reenen**, “The distinct effects of information technology and communication technology on firm organization,” 2009.
- , **R. Sadun, and J. Van Reenen**, “The organization of firms across countries,” 2009.
- , ———, and ———, “Does product market competition lead firms to decentralize?,” *American economic review*, 2010, 100 (2), 434–438.
- , **T. Kretschmer, and J. Van Reenen**, “Work-life balance, management practices and productivity,” 2009.
- Bolton, P. and M. Dewatripont**, “Authority in Organizations: A Survey!,” 2011.
- Bresnahan, T.F., E. Brynjolfsson, and L.M. Hitt**, “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence,” *Quarterly Journal of Economics*, 2002, 117 (1), 339–376.
- Cappelli, P. and D. Neumark**, “Do high-performance work practices improve establishment-level outcomes,” *Industrial and Labor Relations Review*, 2001, 54, 737–775.
- Caroli, E. and J. Van Reenen**, “Skill-Biased Organizational Change? Evidence from a Panel of British and French Establishments,” *Quarterly Journal of Economics*, 2001, 116 (4), 1449–1492.
- Chandler, A.D.**, *Strategy and structure: Chapters in the history of American enterprise*, Cambridge: M.I.T. Press, 1962.
- Crawford, V.P. and J. Sobel**, “Strategic information transmission,” *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.
- Criscuolo, C., J. Haskel, and R. Martin**, “Building the evidence base for productivity policy using business data linking,” *Economic Trends*, 2003, 600, 39–51.
- Daron, A., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti**, “Technology, Information and the Decentralization of the Firm,” *Quarterly Journal of Economics*, 2006, 122 (4), 1759–1799.
- Dessein, W.**, “Authority and communication in organizations,” *The Review of Economic Studies*, 2002, 69 (4), 811.
- DiNardo, J. and D.S. Lee**, “Economic Impacts of New Unionization on Private Sector Employers: 1984-2001,” *Quarterly Journal of Economics*, 2004, pp. 1383–1441.
- Foster, L., J. Haltiwanger, and C. Syverson**, “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?,” *The American Economic Review*, 2008, pp. 394–425.

Garicano, L., “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 2000, 108 (5), 874–904.

——— and **E. Rossi-Hansberg**, “Organizing growth,” *Journal of economic theory*, 2010.

——— and **T. Hubbard**, “Earnings inequality and coordination costs: Evidence from US law firms,” 2009.

Geanakoplos, J. and P. Milgrom, “A theory of hierarchies based on limited managerial attention,” *Journal of the Japanese and International Economies*, 1991, 5 (3), 205–225.

Gibrat, R., “Les Inégalités Economiques (Sirey, Paris),” 1931.

Gibbons, R. and J. Roberts (2012) *Handbook of Organizational Economics*, Princeton: Princeton University Press

Greif, A., “Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies,” *Journal of political economy*, 1994, pp. 912–950.

Grossman, S.J. and O.D. Hart, “The costs and benefits of ownership: A theory of vertical and lateral integration,” *The Journal of Political Economy*, 1986, pp. 691–719.

Guadalupe, M. and J. Wulf, “The flattening firm and product market competition: The Effect of Trade Liberalization on Corporate Hierarchies,” *American Economic Journal: Applied Economics*, 2010, 2 (4), 105–127.

Guiso, L., P. Sapienza, and L. Zingales, “People’s Opium? Religion and Economic Attitudes,” *Journal of Monetary Economics*, 2003, 50 (1), 225–282.

———, ———, and ———, “The role of capital social in financial development,” *American Economic Review*, 2004, 94 (3), 526–556.

———, ———, and ———, “Does culture affect economic outcomes?,” *The Journal of Economic Perspectives*, 2006, 20 (2), 23–48.

———, ———, and ———, “Cultural Biases in Economic Exchange?,” *Quarterly Journal of Economics*, 2009, 124 (3), 1095.

Guiso, S., P. Sapienza, and L. Zingales, “Cultural Bias in Economic Exchange,” *Quarterly Journal of Economics*, 2009.

Hart, O. and B. Holmstrom, “A theory of firm scope,” *Quarterly Journal of Economics*, 2010, 125 (2), 483–513.

Hart, O.D. and J. Moore, “On the design of hierarchies: coordination versus specialization,” 1999.

Helpman, E., M.J. Melitz, and S.R. Yeaple, “Export versus FDI with heterogeneous firms,” *The American Economic Review*, 2004, 94 (1), 300–316.

Hsieh, C.T. and P.J. Klenow, “Misallocation and manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 2009, 124 (4), 1403.

Ichniowski, C., K. Shaw, and G. Prennushi, “The effects of human resource management practices on productivity: A study of steel finishing lines,” *American Economic Review*, 1997, 87 (3), 291–313.

Jorgenson, D.W., M.S. Ho, and K.J. Stiroh, “A retrospective look at the US productivity growth resurgence,” *The Journal of Economic Perspectives*, 2008, 22 (1), 3–24.

Klette, T.J. and Z. Griliches, “The inconsistency of common scale estimators when output prices are unobserved and endogenous,” *Journal of Applied Econometrics*, 1996, 11, 343–361.

Knack, S. and P. Keefer, “Does Social Capital Have An Economic Payoff? A Cross-Country Investigation,” *Quarterly journal of economics*, 1997, 112 (4), 1251–1288.

Lafontaine, F. and M.E. Slade, “Inter-Firm Contracts: Evidence1,” *Handbook of Organizational Economics* (Gibbons, E. and Robert, D. J. eds), forthcoming, 2010.

Lazear, E., “Performance pay and productivity,” *American Economic Review*, 2000, 90 (5), 1346–1361.

Lazear, E. and P. Oyer, “Personnel Economics,” *Handbook of Organizational Economics* (Gibbons, R. and Robert, D. J. eds), 2012.

Lemieux, T., W.B. MacLeod, and D. Parent, “Performance Pay and Wage Inequality,” *Quarterly Journal of Economics*, 2009, 124 (1), 1–49.

MacLeod, W.B. “Reputations, Relationships and Contract Enforcement” *Journal of Economic Literature*, 2007, 45 (3), 597–630.

Marin, D. and T. Verdier, “Power inside the firm and the market: A general equilibrium approach,” *Journal of the European Economic Association*, 2008, 6 (4), 752–788.

—— and ——, “Power in the multinational corporation in industry equilibrium,” *Economic Theory*, 2009, 38 (3), 437–464.

Milgrom, P. and J. Roberts, “The economics of modern manufacturing: Technology, strategy, and organization,” *The American Economic Review*, 1990, pp. 511–528.

Modigliani, F. and M.H. Miller, “The cost of capital, corporation finance and the theory of investment,” *The American Economic Review*, 1958, 48 (3), 261–297.

Nickell, S., D. Nicolitsas, and M. Patterson, “Does doing badly encourage management innovation?,” *Oxford Bulletin of Economics and Statistics*, 2001, 63 (1), 5–28.

Penrose, ET, *The theory of the growth of the firm*, New York: Wiley, 1959.

Prescott, E.C. and M. Visscher, “Organization capital,” *The Journal of Political Economy*, 1980, pp. 446–461.

Rajan, R.G. and J. Wulf, “The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies,” *The Review of Economics and Statistics*, 2006, 88 (4), 759–773.

Solow, R.M., “Technical change and the aggregate production function,” *The Review of Economics and Statistics*, 1957, 39 (3), 312–320.

Stein, J. “Internal Capital Markets and the Competition for Corporate Resources”, *Journal of Finance*, 1997, 52, 111-133.

Stiroh, K., “Reassessing the Role of IT in the Production Function: A Meta-Analysis,” *Federal Reserve Bank of New York Working Paper*, 2002.

Syverson, C., “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy*, 2004, 112 (6).

——, “Product substitutability and productivity dispersion,” *Review of Economics and Statistics*, 2004, 86 (2), 534–550.

——, “What determines productivity?,” *Journal of Economic Literature*, forthcoming, 2010.

Walker, F.A., “On the Source of Business Profits: A Reply to Mr. Macvane,” *Quarterly Journal of Economics*, 1888, 2 (3), 263–296.

A. Academia

In this subsection we explain why academic freedom, defined as the fact that researchers have formal and real control rights on their research agendas, can be socially desirable in early stages of a new research line, i.e. for more basic research, whereas allocating those control rights to the employing firm, may dominate further down the line, i.e. for more applied research. This section draws on Aghion, Dewatripont and Stein (2008), henceforth ADS.

For simplicity, consider a two-stage research line with each line starting with an initial idea I_0 , and eventually generating a marketable product with value V after 2 successful stages. Each stage requires one researcher. We assume that it is sufficient to hire one researcher per stage, and that this researcher obtains a probability of success equal to $p < 1$ at any stage if he follows the success-maximizing (“practical”) research strategy at that stage. But instead of the practical strategy, the researcher is free to follow an “alternative” strategy. If we assume that the scientist has a zero individual probability of success following this approach, then this alternative strategy amounts to the scientist working on an activity that he enjoys more but that does not pay off in monetary terms. However, we can reinterpret this alternative strategy as the case in which the scientist works on an activity that may help initiate new lines but does not generate progress on that particular line.

There is an infinite supply of researchers at each stage, each of whom has an outside option w . After being hired at stage j , the scientist is exposed to idea I_{j-1} , and then learns whether he would prefer following the practical strategy or the alternative strategy. If he is able to undertake his favored strategy, he suffers no disutility from working. If, however, the scientist has to undertake the strategy that he likes less, he suffers disutility of z . The ex ante probability that a scientist prefers to follow the practical strategy is given by α . Assume further that the choice of the practical vs. alternative strategy is ex ante non-contractible.⁸

Academic research (or freedom) differs from private-sector research in that it leaves control rights over the research strategy in the hands of the researcher. Thus if a research line is pursued in academia, the researcher is paid wage w and always works on his preferred strategy. This implies that with probability α , the scientist works on the practical strategy, and with probability $(1 - \alpha)$, he works on the alternative strategy. Thus the ex ante probability of advancing to the next stage is given by αp . Now consider a researcher employed by the private sector. Whether the researcher prefers the practical or the alternative strategy, becomes evident once the researcher has been hired by the firm and has been given access to the idea by the firm owner. Yet ex post, the firm owner has the authority to force the scientist to work on the practical strategy. Anticipating this, the researcher will demand a wage of $w_p = w + (1 - \alpha)z$ in order to work in the private sector. The $(1 - \alpha)z$ markup over the academic wage represents compensation for loss of creative freedom—the fact that scientists now always have to adopt the practical strategy, whether this turns out to coincide with their preferences or not.

A.0.1. When is freedom optimal?

A main finding in ADS is that academic freedom tends to dominate private sector focus at earlier stages on a research line. To see this, take a research line involving 2 stages, and suppose that the first stage has been successful, so that we are now at stage 2, with one more stage to be completed in order to generate a payoff of V . If this last stage of research is done in the private sector, the expected payoff is equal to $E(\pi_2^p) = pV - w_p$. If instead the last stage is done in academia, the expected payoff is equal to $E(\pi_2^a) = \alpha pV - w$. This means that private sector research will yield a higher payoff than free (academic) research and only if $(1 - \alpha)pV > (w_p - w)$, or equivalently $pV > z$.

Now, let Π_2 denote the maximum of $E(\pi_2^p)$ and $E(\pi_2^a)$. Moving back to stage 1, we now compare between $E(\pi_1^p) = p\Pi_2 - w_p$ and $E(\pi_1^a) = \alpha p\Pi_2 - w$. Private sector research will yield

⁸In other words, one cannot write a contract that promises a scientist a bonus for following the practical strategy, because the nature of what kind of work that strategy entails cannot be adequately described ahead of time.

a higher payoff than free (academic) research at stage 1 if and only if $p\Pi_2 > z$.

Since $\Pi_2 < V$, it follows that private sector research is value-maximizing at stage 1, it is also value-maximizing at stage 2. In particular it cannot be value maximizing to have academic freedom operate at later stages than private sector research. The key result is therefore that academic freedom will be the optimal governance structure at earlier stages and private sector research will be optimal at later stages. The intuition is that while academia's wage cost advantage stays constant over research stages, its lower probability of success becomes more problematic as one approaches the final value V .

This result can be generalized to lines of any length k : if Π_i denotes the NPVs of the line of length k as of stage i , we have:

$$\Pi_i = \max\{E(\pi_i^p) = p\Pi_{i+1} - w_p, E(\pi_i^a) = \alpha p\Pi_{i+1} - w\} < \Pi_{i+1}.$$

This monotonicity property, together with the fact that research should be pursued under academic freedom if and only if $p\Pi_{i+1} > z$, yields the desired result.

A.0.2. Discussion

Note that the model so far provides a rationale for free (academic) research even in the extreme case where the alternative strategy has no value beyond saving the researcher the disutility of pursuing the practical strategy. In reality however there is value in experimenting with ideas that may lead to an entirely new research lines, consistently with the idea that scientific discoveries do not follow a purely "linear" model. This does not alter the relative optimality of academia (vs. private research) in earlier (vs. later) stages of research. It does, however, raise the desirability of freedom in general (and academia as the institutional regime that supports such freedom), if we make the realistic assumption that pursuing the alternative strategy confers a higher probability of generating entirely new research lines than pursuing the practical strategy (note that, realistically, the probability of such an event, possibly the result of an "accidental" discovery, is nonzero for both strategies).

Finally, how do universities commit to leaving control rights with their researchers? Our answer is that the duration of employment contracts in academia are verifiable and enforceable by third parties, and prevent the university from firing the faculty without very good reasons. In particular the tenure mechanism ensures that a faculty cannot be fired at all unless found out in fault. This we see as a very effective way to credibly allocate formal authority to the academic researcher over his research agenda.

A complementary answer is that research universities are basically non-profit organizations, which implies in particular that they do not have the resources to micro-manage the research of their faculty members, i.e. to monitor what the faculty do. In terms of the Aghion-Tirole model, university administrations and presidents do not have the means nor the incentives to invest in information acquisition about the kind of research their faculty engage in (this latter point is formalized in ADS).

A.1. Information, communication and the allocation of authority in multi-divisional firms

Alonso, Dessein and Matouschek (2008), henceforth ADM, apply the methodology and approach of Dessein (2002) to study the allocation of decision-making authority in a multi-divisional organization. The main question addressed in their paper is how the need for coordination affects the allocation of control in large multi-national or multi-unit organizations. ADM consider two business units, each run by a different division manager, which share a common central office. Business-units are profit centers, and division managers are (exogenously) assumed to care more about "local" profits in their own unit, than about the firm's overall profits. For example, as in Stein (1997), the manager may be able to siphon off some of the profits of the unit he runs. Each manager also observes some independent private information which affects the optimal way to run his own unit. The central office, or "headquarters", in contrast, does not generate any profits of its own, yet the headquarter

manager may choose to keep control rights over decisions relevant to the two business units. Alternatively, headquarters may choose to leave these control rights with the division managers. The headquarter is assumed to maximize total firm profits (think of the headquarter manager as being the owner of the overall firm).

Each business unit must take a (non-contractible) action which should ideally be both, responsive to the local information observed by the unit manager, and coordinated with the decision taken by the other unit. Overall profit maximization thus entails both, adaptation to local information, and coordination with other unit. Should headquarters keep control over these decisions which involve coordination externalities, or should they decentralize decision-making to the business units managers so as to make it more responsive to local information?

As in Dessein (2002), the cost of centralization is that each unit manager will typically distort the local information he receives in order to influence the headquarter's decision-making process and tilt it towards maximizing profits in his own unit. On the other hand, centralization facilitates coordination between the decisions in the two units. A main difference with Dessein (2002) is that here also under decentralization unit managers have an incentive to distort their local information, because of the strategic interaction between the two units. Under decentralization, the horizontal communication between the two unit managers will also be strategic and distorted, as division managers have different objectives and aim to influence each other's decisions. Hence, it all boils down to whether decentralized coordination (whereby division manager communicate horizontally with each other) is preferred over centralized coordination (whereby division managers communicate vertically with the headquarters which then makes the final decisions).

A first result is that vertical communication is always more informative than horizontal communication as the conflict of interest is larger between the two divisions than between either division and the headquarters. As long as the need for coordination is not too large (in other words when the externalities between the two divisions are sufficiently small), decentralization is strictly optimal, even when division managers only care about the profits of their own business unit. As coordination considerations become more important, i.e. as externalities between the two units increase, decentralized decision-making remains optimal provided the two division managers put enough weight on joint profits. It is only when both coordination considerations are important *and* division managers are sufficiently biased towards maximizing profits in their own unit centralized control by headquarters becomes optimal.

The idea that centralization helps coordination, is also developed by Hart and Holmstrom (2010), henceforth HH. Like ADM, HH consider a two-unit firm where control rights on how to run each unit can either be allocated to the unit manager or to an outside "headquarter" manager. Unlike in ADM where informational considerations feature prominently, in HH the trade-off between centralization and delegation can be summarized as follows: under centralization, the headquarter maximizes the sum of monetary benefits of the two units, in particular she will always make sure that the two units coordinate whenever coordination maximizes the sum of monetary benefits; however the cost of centralization is that private benefits are not internalized by the decision maker. By contrast, under decentralization, each unit manager seeks to maximize his own monetary plus private benefits, and in particular he will oppose coordination whenever coordinating does not maximize this sum, even if coordination happens to maximize the sum of monetary benefits. In case the principal has the possibility to ensure coordination after all by taking control back from unit managers, delegation has the further cost that its reversal generates grievance and therefore shading by the unit manager (s).

B. The empirics of performance heterogeneity across countries and firms

After our theoretical overview of incomplete contracting and the organization of firms we present an overview of the empirical literature on the topic. Before doing this we first examine some of the literature motivating the study of the organization of firms, which examines the incredible differences in productivity between firms and countries.

B.1. Productivity patterns at the country level

In terms of aggregate growth, Solow (1957)’s initially surprising finding that a large fraction of the growth of the US economy was due to growth in residual total factor productivity (TFP) rather than the accumulation of factor inputs such as labor or capital, has now been replicated in numerous countries. In growth theory TFP is assigned to the generation and diffusion of “hard” technological innovations such as hybrid corn, beta blockers and information and communication technologies (ICT). With the advent of better micro-economic data on plant productivity, researchers such as Bailey, Hulten and Campbell (1992) were able to show that about half of aggregate US productivity growth was due to the reallocation of output among plants rather than simply increases in productivity of ongoing incumbent plants. Thus, the idea of a representative plant or firm increasing its productivity can be very misleading as a description of the creative destruction at play in the wider economy with much aggregate TFP growth coming from inefficient low productivity establishments shrinking and exiting the economy in conjunction with more productive plants growing .

More recently, what is true in the time series also appears to be true in the cross section. Figure A1 shows the correlation between GDP per capita and TFP for a large number of countries. It is clear that those countries with high TFP are also the countries with high GDP per capita, suggesting that TFP is critical for understanding cross-country success. But this aggregate TFP difference is also strongly influenced by how different economies allocate output to plants of heterogeneous productivity levels. For example, Figure A2 shows the estimated productivity distribution of the manufacturing sectors in the US and India from Hsieh and Klenow (2009). Compared to the US, India appears to have a much longer “left tail” of low productivity plants. This suggests that there is something about the structure of the Indian economy that is more forgiving of low productivity firms than in the US. A large number of possible explanations present themselves that we will later examine such as competitive intensity in the product market, labor regulations, distortions due to corruption and tax, etc.

What these lines of research show is that productivity dispersion at the micro-economic level is fundamental to understanding the macro-economic patterns. But this only pushes the question one level deeper: what causes firm heterogeneity?

B.2. Productivity heterogeneity

Research on firm heterogeneity has a long history in social science (Syverson (2010) offers a longer survey). Systematic empirical analysis first focused on the firm size distribution measured by employment, sales or assets. Most famously, Gibrat (1931), characterized the size distribution as approximately log normal and sought to explain this with reference to simple statistical models of growth (i.e. Gibrat’s Law that firm growth is independent of size). In recent decades access to large scale (frequently near population) databases in the government (e.g. the Longitudinal Business Database of US establishments) and private sector (e.g. Bureau Van Dijk’s Orbis database of firm accounts) has enabled researchers to look more directly at firm productivity. The growing availability of plant-level data from the US Census Bureau and other nations combined with rapid increases in computer power (allowing much cheaper storage and ability to analyze such databases) has facilitated this development. Bartelsman, Haltiwanger and Scarpetta (2008) offer many examples of the cross country micro-datasets now being used for productivity analysis.

One of the robust facts emerging from these analyses is the very high degree of heterogeneity between business units (see Bartelsman and Doms, 2000). For example, Syverson (2004a) analyses labor productivity (output per worker) in US manufacturing establishments in the 1997 Economic Census and shows that on average, a plant at the 90th percentile of the productivity distribution is over four times as productive as a plant at the 10th percentile in the same four digit sector. Similarly, Criscuolo, Haskel and Martin (2003) show that in the UK in 2000 there is a fivefold difference in productivity between these deciles.

What could explain these differences in productivity, and how can they persist in a competitive industry? One explanation is that if we accounted properly for the different inputs in

the production function there would be little residual productivity differences. It is certainly true that moving from labor productivity to total factor productivity (TFP) reduces the scale of the difference. For example, in Syverson (2004b) the 90-10 productivity difference falls from a factor of 4 to a factor of 1.9, but it does not disappear.

These differences show up clearly even for quite homogeneous goods. An early example is Salter (1960) who studied the British pig iron industry between 1911 and 1926. He showed that the best practice factory produced nearly twice as many tons per hour as the average factory.

A major problem in measuring productivity is the fact that researchers rarely observe plant level prices so an industry price deflator is usually used. Consequently, measured TFP typically includes an element of the firm-specific price-cost margin (e.g. Klette and Griliches, 1996). Foster, Haltiwanger and Syverson (2009) study 11 seven-digit homogeneous goods (including block ice, white pan bread, cardboard boxes and carbon black) where they have access to plant specific output (and input) prices. They find that conventionally measured revenue based TFP (“TFPR”) numbers actually understate the degree of true productivity dispersion (“TFPQ”) especially for newer firms as the more productive firms typically have lower prices and are relatively larger.

Higher TFP is positively related to firm size, growth and survival probabilities. Bartelsman and Dhrymes (1998, Table A.7) show that over a five year period around one third of plants stay in their productivity quintile. This suggests that productivity differences are not purely transitory shocks or measurement errors, but partially persist.

In summary, there is a substantial body of evidence of persistent firm-level heterogeneity in firm productivity (and other dimensions of performance) in narrow industries in many countries and time periods. Differential observable inputs, heterogeneous prices and idiosyncratic stochastic shocks are not able to adequately account for the remarkable dispersion of productivity. So what could account for this? One long suggested factor is management practices, with authors going back at least to Walker (1887) suggesting that management practices play an essential role in explaining differences in performance across firms.

B.3. Management measurement and heterogeneity

Progress in understanding the role of management has been severely limited by the absence of high firm-level quality data. Recently, Bloom and Van Reenen (2007) developed a survey tool that can in principle be used to directly quantify management practices across firms, sectors and countries. Fundamentally, the aim is to measure the overall managerial quality of the firm by benchmarking it against a series of global best practices (see next section for a discussion of what these mean).

Bloom and Van Reenen (2007) use an interview-based evaluation tool that defines and scores, from one (“worst practice”) to five (“best practice”), eighteen basic management practices. This evaluation tool was developed by an international consulting firm to target practices they believed were associated with better performance, covering three broad areas:

1. **Monitoring:** how well do companies track what goes on inside their firms, and use this for continuous improvement? For example, is product quality regularly monitored so that any production defects are quickly addressed rather than left to damage large volumes of output.
2. **Target setting:** do companies set the right targets, track the right outcomes and take appropriate action if the two are incongruent? For example, are individual production targets calibrated to be stretching but achievable, rather than incredibly easy or impossibly hard.
3. **People:** are companies promoting and rewarding employees based on ability and effort, and systematically trying to hire and keep their best employees? For example, are employees that perform well, work hard and display high ability promoted faster than employees who under-perform, are lazy and appear incompetent.

The management survey tool excludes practices whose performance impact clearly depends on individual firms circumstance – for example, setting lower prices or M&A activity.

To obtain accurate responses from firms they interview production plant managers using a ‘double-blind’ technique. One part of this double-blind technique is that managers are not told they are being scored or shown the scoring grid. They are only told they are being “interviewed about management practices for a research project”. To run this blind scoring, open questions were used since these do not tend to lead respondents to a particular answer. For example, the first monitoring question starts by asking “tell me how you monitor your production process” rather than a closed question such as “do you monitor your production daily (yes/no)”. Interviewers also probed for examples to support assertions. The other side of the double-blind technique is interviewers are not told in advance anything about the firm’s performance to avoid prejudice. They are only provided with the company name, telephone number and industry. Since the survey covers medium-sized firms (defined as those employing between 100 and 10,000 workers) these would not be usually known *ex ante* by the interviewers. The survey was targeted at plant managers, who are senior enough to have an overview of management practices but not so senior as to be detached from day-to-day operations. The sample response rate was 45% and this was uncorrelated with measures of firm performance.

One way to summarize firm-specific quality is to z-score each individual question and take an average across all eighteen questions. This management practice score is strongly correlated with firm performance (total factor productivity, profitability, growth rates, and Tobin’s Q and survival rates) as well as firm size. These performance data were taken from independently collected company accounts and imply that the managers’ responses contained real information. Figure A3 shows the correlation between the management score and labor productivity, for example. Firms with higher management scores tend to have higher sales per worker relative to the industry and country average. These correlations should by no means be taken as causal but do suggest that the management data contains useful information. We examine the identification issues in detail in Section 5 below. Other research shows that better management is also associated with more energy efficient production (Bloom, Genakos, Martin and Sadun, 2010), better patient outcomes in hospitals (Bloom, Propper, Seiler and Van Reenen, 2010) and improved work-life balance indicators (Bloom, Kretschmer and Van Reenen, 2009).

The bar chart in Figure A4 plots the average management practice scores across countries from the interviews. The US has the highest average management practice scores, with Germany, Japan and Sweden below, followed by a block of mid-European countries (e.g. France, Italy, the UK and Poland) and Australia, with Southern Europe and developing countries Brazil, China, Greece and India at the bottom. In one sense this cross-country ranking is not surprising since it approximates the cross-country distribution of productivity (recall Figure A1).

The between country variation in average management scores are swamped by the within country variation. Figure A5 plots out the histogram of management scores across firms for each country in our sample. As with productivity there is huge variation in management practices with many badly managed firms in all countries, even those with high average scores. An interesting feature of Figure A5 is that the US appears to have little or no left tail of very badly managed firms (say scores less than 2). In fact, a large fraction of the cross country differences in Figure A4 can be accounted for by whether a country has a thick or thin left tail of badly managed firms. This immediately suggests that one factor causing cross country differences in productivity is the strength of the forces of selection which aggressively weed out badly managed firms in countries like the US to a much greater extent than in India. This would be consistent with the productivity data shown in Figure 2.

There are alternative approaches to measuring management or, more generally, attempts to measure “intangible capital” or “organizational capital” (Prescott and Visscher, 1980). First, one could try and infer these as residuals using relatively weak conditions (variants of TFP) or more tightly specified structures (e.g. Atkeson and Kehoe, 2005). Secondly, one can use past expenditures to build up intangible stocks exactly as would be done for tangible capital (e.g. through the perpetual inventory method). This is frequently done for R&D and

advertising, but it is far harder to accomplish for management as there is no clear data on such expenditures. Third, more conventional paper and pen surveys can be used with closed (limited number of responses like "yes" or "no") type questions (e.g. Black and Lynch, 2001). There are a large number of these for individual countries, but very few that are comparable across international boundaries.

C. Identification

We end this section with a discussion of identification issues. The empirical discussion above was mainly in terms of partial correlations: what factors are associated with the variation in firm organizational form. Incomplete contract (and other) theories suggest what we should expect to observe, in equilibrium, covarying with decentralization. Such type of analysis is useful and potentially consistent (or inconsistent) with theories. However, there is always an "endogeneity" concern with examining such correlations. Generally in social science we are concerned whether a change in one variable causally changes the outcome in another. For example, in Bloom, Sadun and Van Reenen (2012b) we are concerned whether an exogenous change in trust causes an increase in delegation. We observe a positive correlation even after controlling for a host of other confounding influences. Nevertheless, it could still be the case that there is an omitted variable. For example, different strong local enforcement of the rule of law could be correlated with both more trust and more decentralization. Alternatively, there could be direct reverse causality: an increase in delegation could foster greater trust, rather than the other way around.

There is a vast literature in statistics and econometrics on the identification of causal relationships and this has become a strong focus of interest in economics over the last decade. Consider a binary variable of interest $D_i = \{0, 1\}$, e.g. whether trust is high in firm i or not and an outcome of interest Y_i , e.g. the degree of decentralization. Define the potential outcomes for firm i as Y_{1i} if $D_i = 1$ and Y_{0i} if $D_i = 0$. We may be interested in $Y_{1i} - Y_{0i}$, i.e. the degree of decentralization that occurs from shifting a firm from a low to high trust environment. Essentially the problem is constructing a valid counterfactual: i.e. what would have happened to the same treated firm in the absence of the treatment (and vice versa)? Since this is not directly observed the design of the empirical study has to reconstruct this counterfactual.

The naive approach would be to compare the outcomes of those who were treated with those who were not, i.e. $E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$. Unfortunately this is unlikely to get us what we want because

$$E(Y_i|D_i = 1) - E(Y_i|D_i = 0) = \{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)\} + \{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)\} \quad (\text{C.1})$$

The first term is on the right hand side, $\{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)\}$, is the "average treatment effect on the treated" and an object of interest. But the second term $\{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)\}$ is a "selection bias" which will in general not be zero. The firms who have high trust are not randomly allocated.

There are several ways to construct the counterfactual. First, we could run a field experiment. This has been surprisingly under-utilized in organizational economics (except in the lab with students) even though randomized control trials are becoming common in some areas of applied economics such as development and labor. The benefit of a field experiment is that if properly implemented (and this is a big "if") the control group is the valid counterfactual. For the most part we have to rely on non-experimental approaches to construct the counterfactual. Classically, researchers will condition on many other variables to deal with the endogeneity concern, so we look at the object

$$E(Y_i|D_i = 1, X_i) - E(Y_i|D_i = 0, X_i)$$

This may help so long as the X 's are exogenous (if not then we may make the biases worse - the "bad control" problem). But unless we have a very rich set of controls in the X -vector

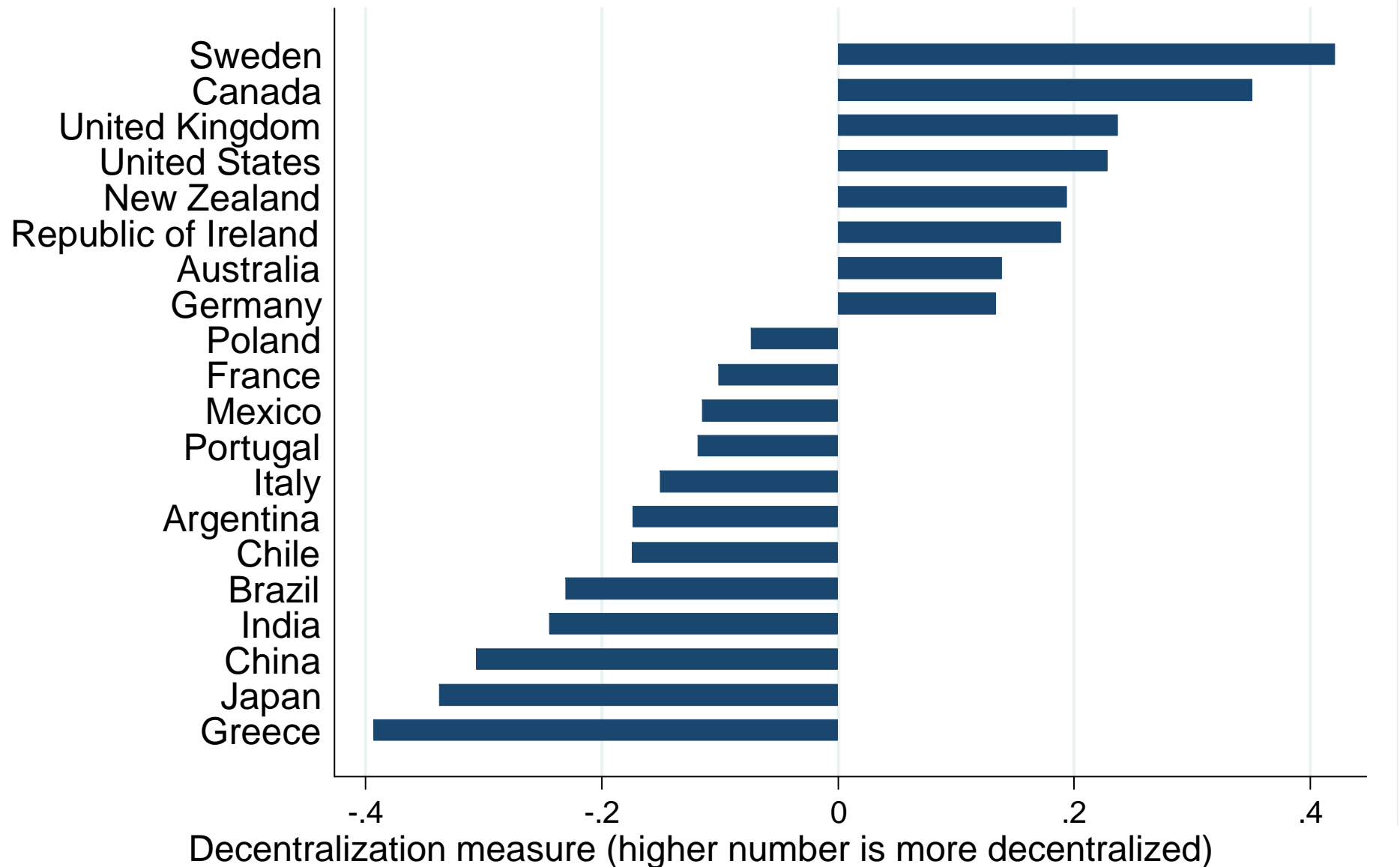
there could still be “selection on the unobservables” implying that there will still be a bias remaining. Matching estimators (e.g. on the propensity score) to obtain the counterfactual helps ensure a common support between treatment and control groups. But if there are still correlated unobservables then matching will not eliminate the selection bias in equation.

There are various identification strategies to eliminate the bias. First, controlling for firm fixed effects and time dummies (“Difference in Differences” or DD) is one approach if there is panel data. However, if there is a correlated unobservable shock there will still be bias. Second, instrumental variables can be used (and combined with DD). Essentially this requires a variable correlated with D_i but uncorrelated with the error term. A randomized control trial generates the instrument which is whether a firm has been randomly assigned to treatment or control (see Bloom et al, 2010, below). In the context of the decentralization-trust relationship, Bloom, Sadun and Van Reenen (2012b) use religious differences between countries to instrument bilateral trust following Guiso, Sapienza and Zingales (2010). As noted above, Guadalupe and Wulf (2010) exploit a policy experiment of trade liberalization to obtain exogenous variation in competition as US firms who operated more intensively in product markets where tariffs fell the most would be (exogenously) more strongly affected by the policy change. Finally, regression discontinuity designs have become a popular way of constructing counterfactuals in designs where there is a discontinuous break in a variable of interest. For example, in examining the effects of US labor unions on firm performance DiNardo and Lee (2004) examine the outcomes of union elections where workers narrowly voted in favor of the union compared to elections where they voted narrowly against. Because the unobservables should be balanced on either side of the discontinuity this can in principle deal with the selection problem.

The current literature on empirical organizational economics has not emphasized the identification issues discussed here, partially because of the difficulty of addressing causality. Sometimes a structural model is proposed and the parameters are estimated from the data (e.g. Garicano and Hubbard, 2009). There is nothing in principle wrong with this and any model makes some assumptions. A concern with some structural work, of course, is that the model may be rather inflexible and imposed when it is at stark variance with some main features of the data. The identification procedures discussed here can be combined with structural estimation, although this has rarely happened in practice in empirical economics.

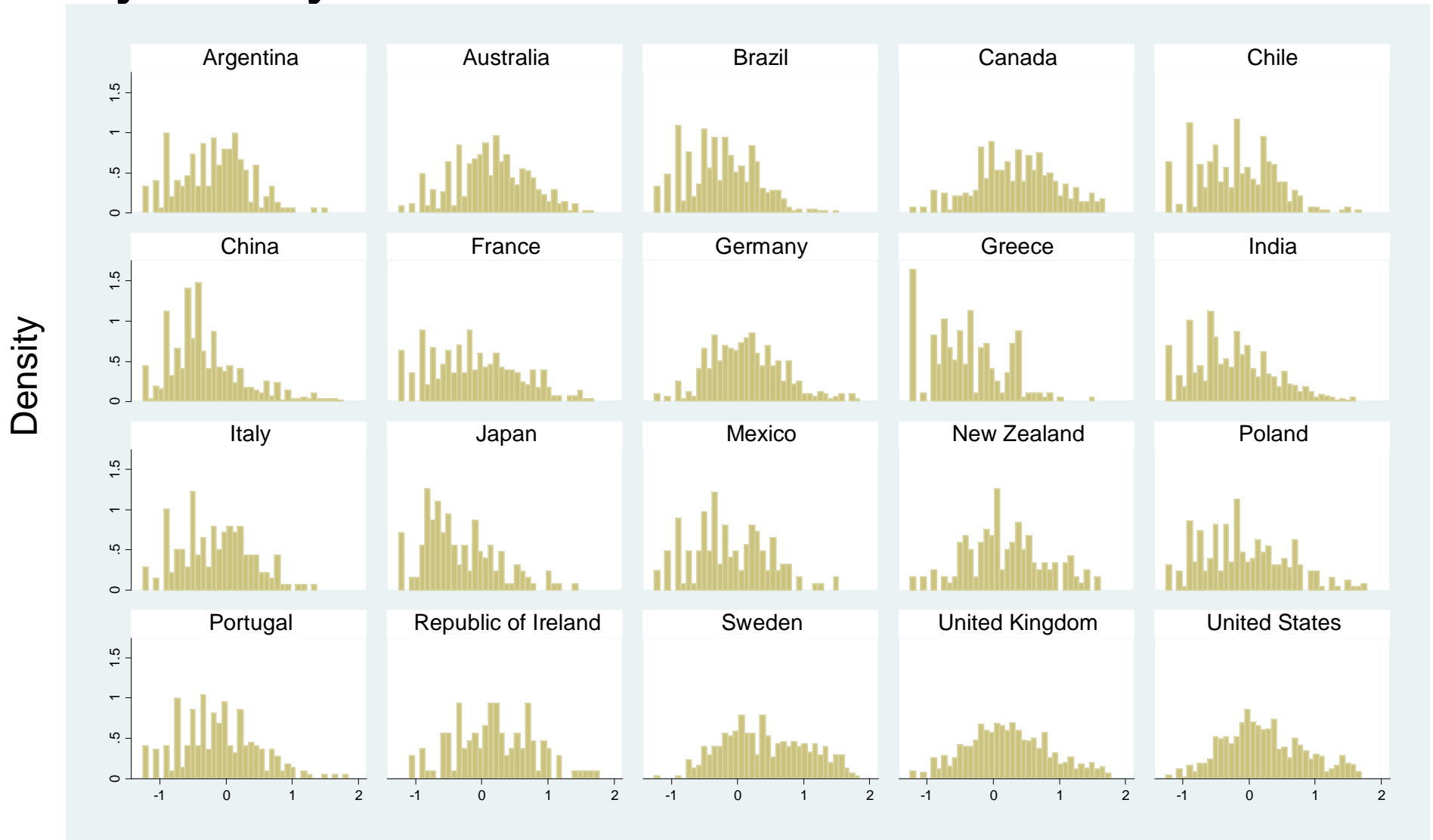
It is likely that more papers on credible identification in organizational economics will be a growth area in the future.

Figure 1: Decentralization is higher in Scandinavia and Anglo Saxon countries, and lower in Asia and Southern Europe



Source: Histogram of decentralization across firms within each country. The definition of decentralization follows Bloom, Sadun and Van Reenen (2012b), with data updated to 2010 from www.worldmanagementsurvey.com

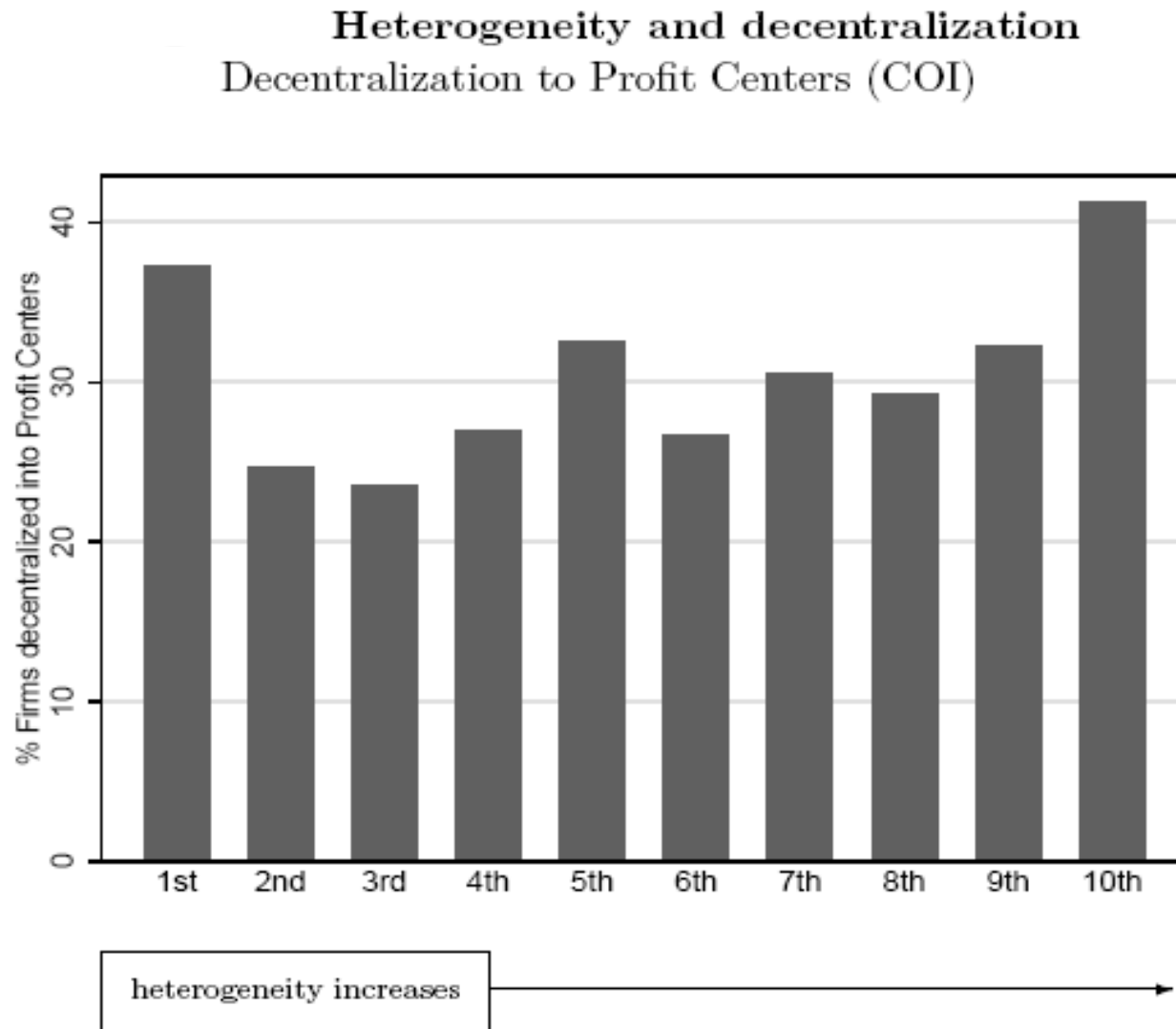
Figure 2: Decentralization varies heavily across firms within every country



Decentralization measure (higher number is more decentralized)

Source: Histogram of decentralization across firms within each country. The definition of decentralization follows Bloom, Sadun and Van Reenen (2012b), with data updated to 2010 from www.worldmanagementsurvey.com

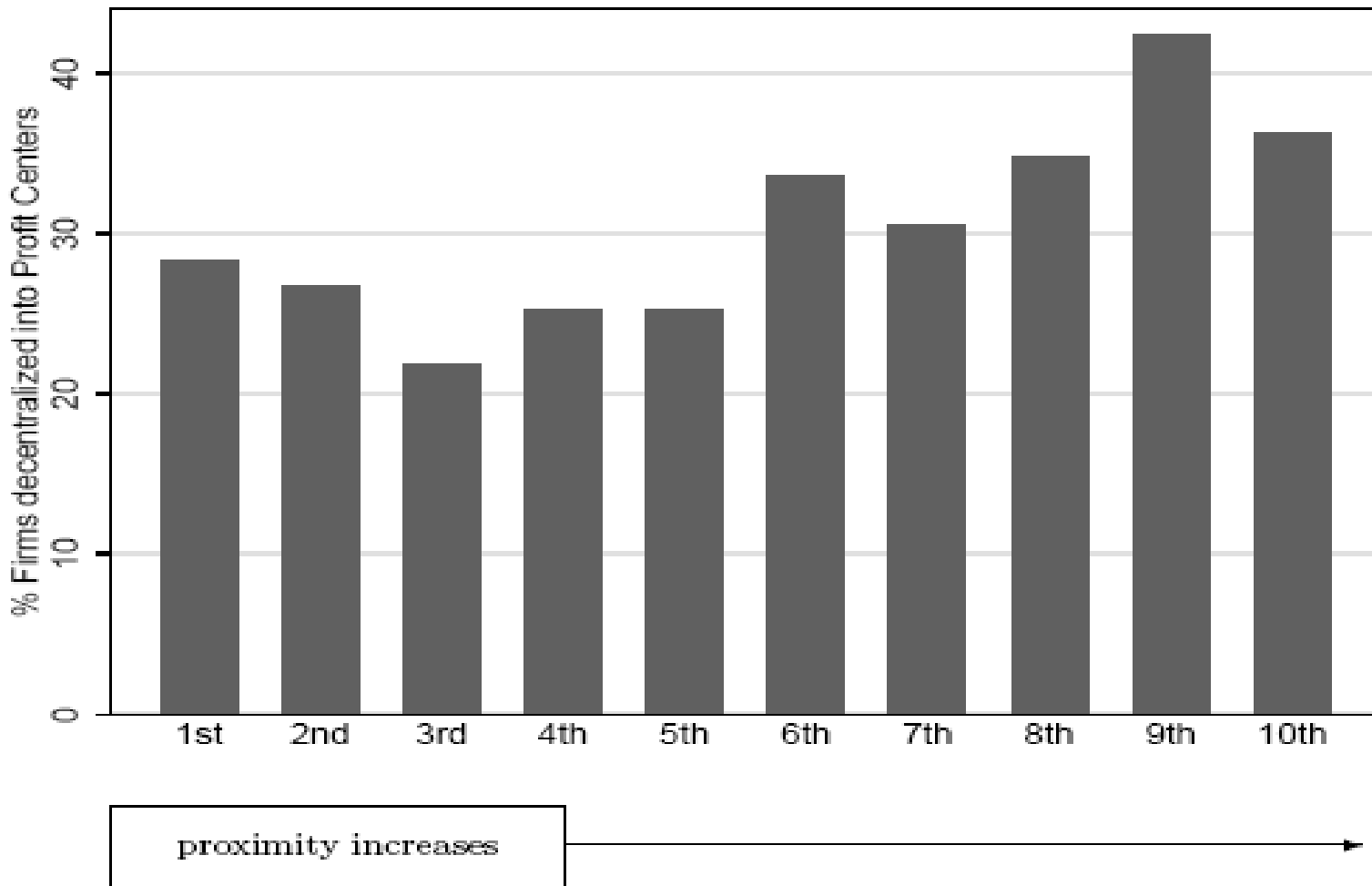
Figure 3, Decentralization is higher when heterogeneity is greater



Source: Acemoglu, Aghion, Lelarge, Van Reenen and Zilibotti (2007)

Figure 4, Decentralization is higher when plants are closer to the TFP frontier

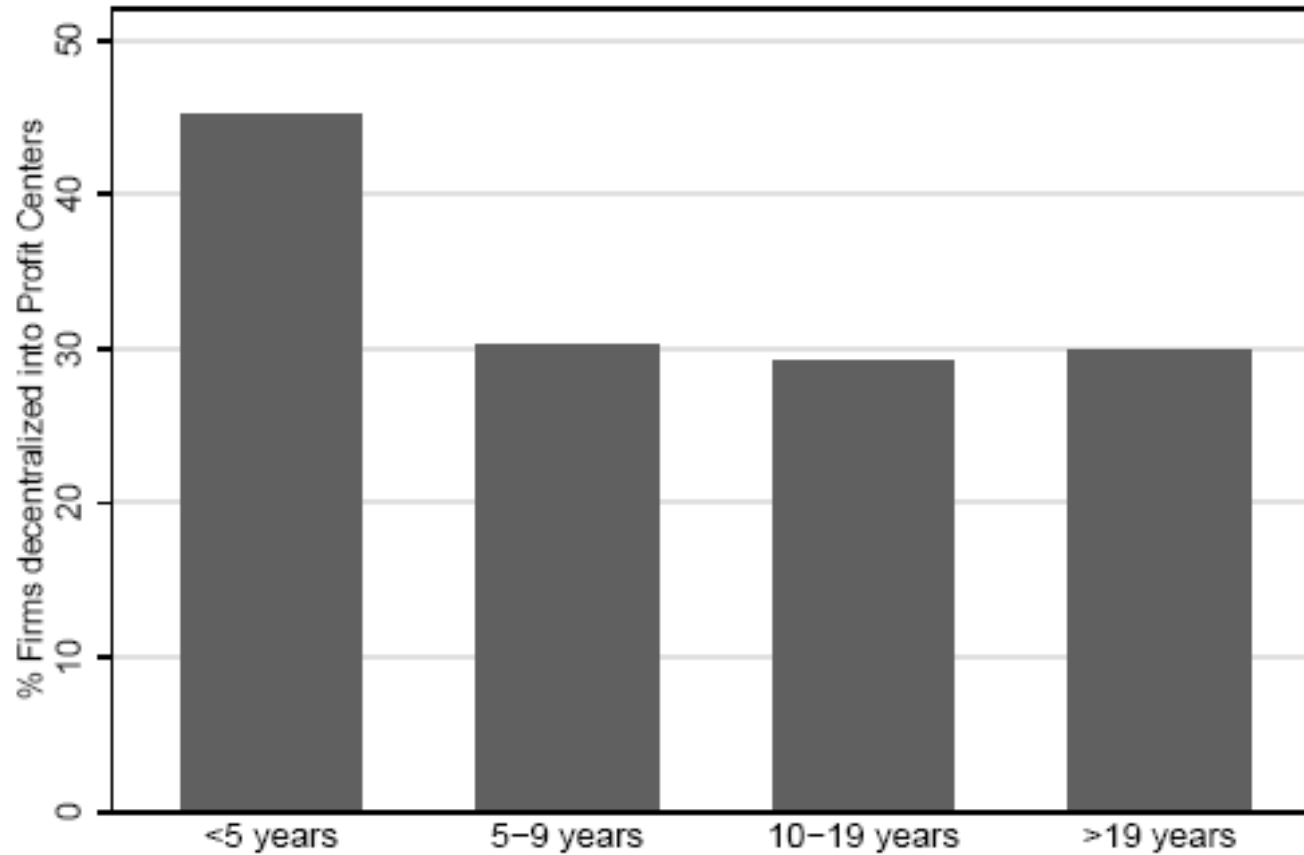
Proximity to frontier and decentralization
Decentralization to Profit Centers (COI)



Source: Acemoglu, Aghion, Lelarge, Van Reenen and Zilibotti (2007)

Figure 5, Decentralization is higher in younger plants

Age and decentralization
Decentralization to Profit Centres (COI)



Source: Acemoglu, Aghion, Lelarge, Van Reenen and Zilibotti (2007)