Estimating Management Practice Complementarity between Decentralization and Performance Pay*

Bryan Hong†

Lorenz Kueng‡

Mu-Jeung Yang§

May 11, 2014 - Version 0.2

For the latest version, **click here**

Abstract

The existence of complementarity across management practices has been proposed as one potential explanation for the persistence of firm-level productivity differences. However, there are currently no conclusive population-level tests of the complementary joint adoption of management practices. Using unique detailed data on internal organization, occupational composition, and firm performance for a nationally representative sample of firms in the Canadian economy, we exploit regional variation in income tax progression as an instrument for the adoption of performance pay. We find systematic evidence for the complementarity of performance pay and decentralization of decision-making from principals to employees. Furthermore, in response to the adoption of performance pay, there is a concentration of decision-making at the level of managerial employees, as opposed to a general movement towards more decentralization in the organization. Finally, we find that adoption of performance pay is related to other types of organizational restructuring, such as greater use of outsourcing, Total Quality Management, re-engineering and a reduction in the number of layers in the hierarchy.

*We would like to thank Chad Syverson and the seminar participants at Emory University and the University of Washington for helpful comments. We are also grateful to Ergete Ferede for sharing his residual income progression data. The usual disclaimer applies.

†Ivey Business School, Western University

‡Kellogg School of Management, Northwestern University

§University of Washington, Seattle

Introduction

A central empirical fact regarding firm heterogeneity is that even within narrowly defined industries, productivity differences across firms are large and persistent, see Baily, Hulten and Campbell (1992), and Syverson (2011). This empirical pattern is also at the heart of the modern analysis of firm heterogeneity in diverse fields, such as industrial organization, macroeconomics, and international trade. Recent empirical evidence suggests that management practices are closely connected to firm productivity, see Bloom and Van Reenen (2007) and Bloom, Brynjolfsson, Jarmin, Saporta and Van Reenen (2014). However, little is known empirically about how different types of management practices interact to influence productivity. Specifically, complementarity among management practices is a potential cause of persistence in firm productivity. Under such complementarity, firms need to jointly adopt all complementary practices to reap the full productivity gains from these management practices. However, since the joint adoption of all these practices can be very difficult and costly to implement, low productivity firms might have insufficient incentive to do so. For this mechanism to potentially explain persistent firm productivity differences, management practice complementarity has to be shown to exist in the data. Therefore, we test for the existence complementary joint adoption of two key management practices at the heart of efficient decision-making in firms: performance pay and the degree of decentralization of decision-making. On the one hand, a vast literature starting out with Drucker (1949) and Penrose (1959) has argued that decentralized decision-making is important to promote firm growth, induce an optimal division of labor and enable the efficient use of decentralized employee information[1]. On the other hand, creating incentives for efficient decision making through performance pay has been at the core of principal-agent analysis, as exemplified by Jensen and Meckling (1976) and surveyed in Laffont and Martimort (2002).

Since the seminal work by Milgrom and Roberts (1990), there has been considerable interest in the complementarity of management or organizational practices, see Brynjolfsson and Milgrom (2013). Despite this interest, empirical evidence on the issue has been limited in scope. Two challenges posed significant hurdles to further empirical progress on this topic. First, large sample data on the internal organization of firms and use of management practices was limited. In fact, the most prominent empirical studies of management practice complementarity are either case studies such as Lincoln Electric or industry studies such as Ichniowski, Shaw and Prennushi (1997) or Kretschmer, Miravete and Pernias (2012). Second, even in datasets that provide larger samples, such as that used by Bloom and Van Reenen (2007), recovering causal estimates of complementarity proved to be difficult since joint adoption

[1] For examples see Aghion and Tirole (1997), Garicano (2000), Bloom, Sadun and Van Reenen (2012).

of management practices can be driven by correlation of adoption costs instead of complementarity, see Athey and Stern (1998).

In this paper, we address both these empirical challenges and provide the first large-sample evidence on management practice complementarity, representative for an entire economy. We are able to provide these estimates, as we have access to a unique establishment and firm-level panel dataset on management practices, organization, and firm performance that is representative of all business firms in the Canadian economy. To establish valid inferences regarding complementarity, we address the issues raised by Athey and Stern (1998) using an instrumental variables strategy. The basic idea of our instrumental variables strategy is that higher personal income tax progressivity makes the adoption of performance pay less likely and can therefore serve as an instrument for performance pay. To understand why, consider performance pay in a standard principal agent framework, such as Grossman and Hart (1983). In this model, the principal needs to pay the agent a higher wage in high output states to incentivize the agent to exert hidden effort. Since progressive taxation by definition exhibits higher marginal tax rates on higher income, more progressivity will tend to offset the incentive effects of performance pay. In other words, to induce a given level of effort, the firm has to pay the employee overproportionally more in high output states to counter the effects of tax progressivity.

We find that firms that adopt performance pay for exogenous reasons decentralize more decision-making from business owners and headquarters to employees. This suggests that performance pay and decentralization of control are complements. This result sheds light on a number of theoretical models that mostly came to ambiguous conclusions about the relation of performance pay and decentralization. Studies such as Prendergast (2002) and Van Den Steen (2010) described mechanisms through which performance pay and decentralization are complements, while Holmstrom and Milgrom (1991), Krishna and Morgan (2008) and Bester and Kremer (2008) provided theories showing that performance pay and decentralization are substitutes. It is worth emphasizing that our results are not a direct test of these models, since it is still possible that there are some industries or firms to which these theories apply. However, we note that our data allows us to characterize the average degree of complementarity between performance pay and decentralization in the entire population of firms, irrespective of industry, size class, exporter status etc. In this sense, one can interpret our findings as providing evidence on whether these theories apply on average in the population of business firms.

While our general result regarding performance pay and decentralization is of interest, it does not tell us much about the nature of decentralization of decision-making to employees. This decentralization could take very different forms that are indicative of the manner in which firms induce efficient decision-making. For example, if information of production line workers is crucial, one would expect decisions to be decentralized mostly to non managerial employees. If on the other hand,

managerial supervision and control is crucial, as suggested by Atalay, Hortascu and Syverson (2013), one would expect decision authority to be given to management. Motivated by these different views on the benefits of decentralization, we further analyze decentralization patterns by measuring decision control separately for managerial and non managerial employees. We find that firms that adopt performance pay for exogenous reasons concentrate control at the management level. That is, although these firms tend to decentralize decision control from HQ or business owners to employees, they also reallocate decision tasks from non managerial to managerial employees. Compatible with this concentration of control at the manager level, we find that these firms have systematically higher hiring rates for managers but not significantly higher overall hiring or firing rates. In other words, not only are decision responsibilities concentrated at the management level, but firms also hire more managers to deal with these responsibilities. Therefore, our findings are supportive of the hypothesis advanced by Atalay et al. (2013) that the nature of firms is intimately connected to their role in "mediating managerial supervision and control."

To explore whether unobserved regional factors drive our results, we compare establishments within the same location, which are part of multi-province firms. We show that establishments of firms with headquarters in high tax progression regions have a lower likelihood to adopt performance pay than establishments in the same location that have headquarters in low tax progression regions. Furthermore, we find that for these multi-province firms, the same patterns apply as in the general firm population: establishments which adopt performance pay for exogenous reasons decentralize decision making from principals to employees and reallocate decision tasks from non-managerial to managerial employees.

To obtain a more general understanding of how performance pay influences management practices, we also consider the impact of performance pay adoption on other organizational changes. We highlight three findings. First, firms systematically restructure business processes (re-engineering), without necessarily downsizing their workforce. Second, firms are more likely to adopt Total Quality Management (TQM). Third, firms are more likely to reduce the number of managerial layers (de-layering). We emphasize the importance of these results for theories of business processes and models of endogenous management layers.

The remainer of the paper is organized in the following manner. Section 1 develops a baseline theory that illustrates the incentive trade-offs that make decentralization and performance pay either complements or substitutes. Section 2 provides an overview of the establishment and firm data that is the basis for measurement. Section 3 describes our IV strategy and discusses our baseline empirical results. Section 4 provides a number of robustness checks.

Related Literature

This study contributes to the literature on the sources of persistent performance differences across firms, see Bloom and Van Reenen (2007), Syverson (2011) and Gibbons and Henderson (2013). In contrast to studies such as Bloom and Van Reenen (2007) and Bloom, Sadun and Van Reenen (2013b), which characterize the overall quality of management practices, we focus on the interaction between management practices.

In this context, we build on ideas that draw on literatures in several fields. After the initial seminal work by Milgrom and Roberts (1990), the importance of complementarity in a firm's management practices and organizational activities was introduced to business practitioners by Porter (1996) and prominently featured in the strategic management literature since Rivkin (2000). See also Porter and Siggelkow (2008) for a survey of this literature. Similar ideas regarding the importance of complementarity in production factors to explain persistent productivity differences have even been applied by Jones (2011) to total factor productivity (TFP) differences across countries. A major contribution to this literature is that our empirical results are representative for the entire population of business firms, irrespective of firm size and sector. Moreover, since our estimates are based on what Athey and Stern (1998) describe as "reduced form tests of complementarity based on exclusion restrictions," they provide a valid characterization of management practice complementarity despite potential threats to identification outlined by Athey and Stern (1998).

We also contribute to the recent literature on endogenous firm productivity and organization, such as Grossman and Helpman (2002), Chassang (2010), Gibbons, Holden and Powell (2012), and Powell (2013). This literature is important in that it promises insights into the sources of firm productivity differences that is complementary to R&D based models, such as Klette and Kortum (2004), Luttmer (2007) and Acemoglu, Akcigit, Bloom, Kerr and Van Reenen (2013). However, empirical evidence on how management practices interact to influence firm productivity has been scant. We hope to encourage more theory development in this important field and hope to guide some this new theory with representative stylized facts.

1 Theory

To clarify the mechanisms through which decentralization of control and implementation of performance pay can either be substitutes or complements, we develop a model of the trade-off between centralization and decentralization and how performance pay affects this trade-off. Like much of the recent empirical literature on management practices and decentralization, such as Bloom et al. (2012) and Bloom, Garicano, Sadun and Van Reenen (2013a), decentralization will be based on

the logic developed by Garicano (2000) and Garicano and Rossi-Hansberg (2012). Specifically, more decentralization in these models allows a firm to economize on the business owner's time, thereby allowing her to focus only on complex and rare problems. On the other hand, decentralization involves effort costs to train employees and communication costs if employees encounter an unsolvable problem and need to explain it to the business owner. We extend this basic framework by assuming that part of the effort costs spent by employees to solve problems is unobservable. This gives rise to a classical moral hazard problem, such as Grossman and Hart (1983), which can potentially be addressed using performance pay.

There are two players in our model, a principal P who can be considered as the business owner and an agent A, who is an employee. Firms face a unit measure of problems $z \in [0, 1]$ in their operations with cumulative distribution $F(z)$ and density $f(z)$. As is standard, we assume that these are ordered by frequency and complexity so that lower indices of z denote simple and very frequent problems and high values of z very complex and rare problems. This is formalized by assuming $f'(z) < 0$. Problems z need to be solved to produce output hence, every unsolved problem reduces output and therefore firm productivity.

To solve problems, both P and A have to exert effort to acquire knowledge to deal with a problem. This cost is higher the higher complexity of problems z, and we denote the overall cost by $a_i \cdot z$, where $i \in \{P, A\}$ and $a_i > 0$. As in Bloom et al. (2013a), we assume that P knows all production tasks that employees A know so that knowledge overlaps. If A cannot solve a particular problem, he will communicate the problem up the hierarchy to P. This entails a communication cost h. To introduce a basic moral hazard problem into this framework, we assume that the agent's effort costs have two parts. First, $a_{1,A} \cdot z$ are costs associated with observable effort in Bloom et al. (2012). Second, there is a part of the effort costs denoted by $a_{2,A} \cdot z$ that is associated with unobservable effort. That is, A has a hidden effort choice $e \in \{0, 1\}$ so that if he incurs effort $e = 1$, he pays an additional cost $a_{2,A} \cdot z$. As usual, while the effort choice e itself is hidden, the level of effort costs $a_{2,A} \cdot z$ is common knowledge.

Production. Output is generated as a result of solving problems, where z_A denotes the complexity of problems up to which employees are engaged in problem solving. If principal P is confronted with a problem, we assume she is able to solve it by incurring costs a_P. However, for employees, even for problems they in principle could solve, there is a random chance that the problem remains unsolved possibly due to unforeseen issues. Let x_A denote an indicator for whether the employee solved the problems he was confronted with. As is standard in principal agent problems, we assume that the employee can increase output by exerting the hidden effort choice e. In other words, if the agent exerts effort, the fraction of problems that he can

solve increases. Therefore, production is given by

$$\begin{aligned} \text{Principal:} \quad & 1 - F(z_A) \\ \text{Agent:} \quad & F(z_A) \cdot P\left(x_A = 1 \middle| e\right) \end{aligned}$$

Preferences. Business owner P is assumed to be risk neutral and the residual claimant of profit flows. Agent A is assumed to exhibit the following utility function

$$U_A(w_A, z_A, e) = \ln(w_A) - a_{1,A} z_A - e \cdot a_{2,A} z_A$$

where w_A is a (possibly state-contingent) wage. Since the effort associated with cost $a_{1,A} z_A$ is observable, we assume that this is directly paid for by the principal P.

Timing. The general timing of our model can be described as follows. At $t = 0$, P decides first on the degree of decentralization and then implements the optimal wage contract. That is, if performance pay is available, P will determine the optimal state-contingent wage payments, otherwise she will just set a constant wage to meet the agent's outside options. In $t = 1$, the agent faces the organizational choices made by P and decides on whether to exert effort. Production occurs and wage payments are made at the end.

Performance Pay. We begin with the solution of the moral hazard problem. If performance pay is available, P chooses to condition wage payments on whether the agent A solved a given problem. The optimal contract is designed to minimize expected wage payments subject to incentive-compatibility and participation constraints:

$$\begin{aligned} &\min_{w_A^1, w_A^0} P\left(x_A = 1 \middle| e = 1\right) \cdot w_A^H + P\left(x_A = 0 \middle| e = 1\right) \cdot w_A^L \\ &\qquad \text{subject to:} \\ &\text{(IC)} \quad E\left[U_A(w_A, z_A, e = 1) \middle| e = 1\right] \geq E\left[U_A(w_A, z_A, e = 0) \middle| e = 0\right] \\ &\text{(IR)} \quad E\left[U_A(w_A, z_A, e = 1) \middle| e = 1\right] \geq U_A^R \end{aligned} \tag{1}$$

To facilitate exposition, we assume that $P\left(x_A = 1 \middle| e = 0\right) = 1 - k$ and $P\left(x_A = 1 \middle| e = 1\right) = 1 - k + q$, where k is the probability that a random event renders the problem unsolvable and q is the incremental gain in successful "problem solving" if the agent exerts effort, with $q < k$. Wages w_A are contingent on problems being solved, so that $w_A = w_A^1$ if $x_A = 1$ and $w_A = w_A^0$ if $x_A = 0$.

By standard arguments, the solution to the contract design problem boils down to (IC) and (IR) holding exactly, pinning down contingent wages w_A^1, w_A^0. Given the functional form assumptions on utility and probabilities, performance pay is given

by

$$w_A^1(z_A) = \exp\left\{U_A^R + \frac{k}{q} \cdot a_{2,A} z_A\right\}$$
$$w_A^0(z_A) = \exp\left\{U_A^R - \frac{1-k}{q} \cdot a_{2,A} z_A\right\}$$

For notational convenience, let us define

$$\bar{w}(z_A) = E\left[w_A(z_A)\middle|e=1\right]$$

At this point, we note an important property of the optimal contract. In particular, the more complex the tasks of agents are, that is the higher z_A, the higher the costs of incentive pay. In the appendix, we show that the following inequality holds:

$$\bar{w}'(z_A) > 0 \tag{2}$$

The reason for this relationship is that for more complex problems the agent's effort costs are higher. Therefore to induce A to incur this effort, the principal has to strengthen the high powered incentives, which increases the costs of performance pay.

Decentralization without Performance Pay. The baseline case of our model without performance pay simplifies to the decentralization model of Bloom et al. (2012) and Bloom et al. (2013a), where one can replace the "trust" parameter in Bloom et al. (2012) with $P\left(x_A=1\middle|e=0\right)$. The degree of decentralization, z_A, is chosen to maximize profits:

$$\Pi(z_A,0) = \Big[1-F(z_A)\Big] + F(z_A)\cdot P\left(x_A=1\middle|e=0\right) - h\cdot a_P\cdot\Big[1-F(z_A)\Big] - a_{1,A}z_A - w_A^R$$

which implies the optimal degree of decentralization

$$f\left(z_A^{*,0}\right) = \frac{a_{1,A}}{h\cdot a_P - k} \tag{3}$$

where $z_A^{*,0}$ denotes the optimal degree of decentralization in a firm without performance pay. The comparative statics of this case follow from the fact that $f'(z) < 0$, as described in Bloom et al. (2012).

Decentralization with Performance Pay. The equilibrium choice of decentralization under the existence of performance pay maximizes

$$\Pi(z_A,1) = \Big[1-F(z_A)\Big] + F(z_A)\cdot P\left(x_A=1\middle|e=1\right) - h\cdot a_P\cdot\Big[1-F(z_A)\Big] - a_{1,A}z_A - \bar{w}(z_A)$$

In particular, the optimal degree of decentralization under performance pay $z_A^{*,1}$ is given by

$$f\left(z_A^{*,1}\right) = \frac{a_{1,A} + \bar{w}'(z_A)}{h \cdot a_P - k + q} \tag{4}$$

Complementarity of Decentralization and Performance Pay. Whether decentralization and performance pay are complements or substitutes depends on the balance of two margins. First, performance pay can induce employees to make more efficient decisions by exerting effort. This margin enters through the term $q = P\left(x_A = 1 \middle| e = 1\right) - P\left(x_A = 1 \middle| e = 0\right)$ and is a force toward decentralization. This is comparable to more "trust" in the framework of Bloom et al. (2012). The reason is that performance pay incentivizes exertion of unobservable effort, which in turn makes higher output more likely. Second, optimal performance pay will depend on the degree of decentralization, as shown in equation (2). More decentralization – that is higher z_A – will increase effort costs of agents; it therefore requires more high-powered incentives to induce effort. This is a force toward centralization, since decentralized decision-making becomes more costly. If the incremental gain from hidden effort of employees q is relatively high, and the marginal performance pay costs of decentralization $\bar{w}'(z_A)$ are low, then $z_A^{*,1} > z_A^{*,0}$ and decentralization and performance pay are complements. However, if marginal performance pay costs of decentralization $\bar{w}'(z_A)$ are high and q is low, decentralization and performance pay will be substitutes, or $z_A^{*,1} < z_A^{*,0}$. We summarize the implications for our empirical section as follows

> **Proposition.** In our baseline model of decentralization and performance pay, suppose there is exogenous variation in performance pay. Specifically, let denote $z_A^{*,0}$ the optimal degree of decentralization without performance pay and $z_A^{*,1}$ the optimal degree of decentralization with performance pay. Then,
>
> $$z_A^{*,1} > z_A^{*,0} \text{ if and only if } \frac{q}{h \cdot a_P - k} > \frac{\bar{w}'\left(z_A^{*,1}\right)}{a_{1,A}}$$

In other words, in response to an exogenous adoption of performance pay, decentralization increases if the incremental efficiency gain from effort outweighs the increased cost of performance pay. The comparative static in the proposition clarifies the potential mechanisms underlying our empirical approach. First, as in the proposition, we will utilize exogenous variation in the adoption of performance pay to then trace out the optimal degree of decentralization. Second, note that in our model, decentralization and performance pay can be either complements or substitutes. That is our model nests both possibilities, so that depending on our empirical results, the model could subsequently be calibrated to investigate the quantitative impact of the mechanisms we described.

2 Data and Measurement

2.1 Data

To estimate the degree of complementarity between management practices, detailed data on the internal organization of firms and the implementation of management practices is necessary. The source of our data is the Workplace and Employee Survey (WES), which is a random stratified sample of establishments with the universe of Canadian firms as the target population and is conducted by Statistics Canada. The survey has a cross sectional dimension of approximately 6500 firms over the time period 1999 to 2006, with a low overall attrition rate of around 20%. Of these 6500 firms, we focus on the sample of around 5500 for-profit business firms. There are several advantages of this dataset compared to other existing micro-datasets on management practices and internal firm organization. First, the WES has a comprehensive sectoral coverage in contrast to either more narrow industry studies such as Ichniowski et al. (1997) or manufacturing based studies such as Bloom et al. (2014). Second, since the target population is the universe of business firms, the WES is not biased towards certain size classes as is the World Management Survey Data by Bloom and Van Reenen (2007). Third, a key advantage of the WES is that firm information is obtained from a culturally relatively homogeneous country such as Canada[2]. This is important as it allows us to exclude many unobserved cultural factors that might bias our results.

2.2 Measurement

Real Decision Authority. The measurement of decentralization requires a credible approach to quantify the degree of decision control by organizational layer. To understand the potential problems, take as example measures of formal decision authority such as can be found in occupational titles and organizational charts. As indicated by Aghion and Tirole (1997), managers higher up the hierarchy often only "rubberstamp" decisions actually taken by non managerial employees. Increased range of formal responsibilities and reporting relations can either imply increased control. Or it might imply even more limited attention to given decision tasks and de facto reduced control.

In contrast, the WES includes detailed information regarding real decisions on tasks across layers in the organizational hierarchy. These questions are similar to measures of worker autonomy in Bresnahan, Brynjolfsson and Hitt (2002) and Bloom et al. (2013a) in that they allow us to measure to which degree principals vs. agents are taking decisions across 12 potential tasks. Specifically, the survey question is: "Who normally makes decisions with respect to the following activities?". The re-

[2]The exception being Quebec, a fact for which we explicitly control later.

spondent is then given a choice of 12 possible activities from "Daily planning of individual work" over "Quality control" up to "Product and service development." There are six possible responses to the question of who makes decisions: non managerial employee, work group, work supervisor, senior manager, individual/group outside the workplace – typically headquarters for multi-establishment firms – and business owners. Table 1 provides some rough idea about the patterns of decision allocation across layers. For any of the 12 possible tasks, it shows how high up the hierarchy the tasks is typically decided and how many layers are involved. To calculate this, we assign non managerial employees the number 1, work groups the number 2, etc. As the first two columns show, decision-making on routine tasks like daily workplanning are typically decided decentralized in lower layers, while complex tasks such as product/service development are decided higher up the hierarchy. In particular, note that if we exclude firms that have any involvement of business owners in decisions this pattern becomes even stronger. The reason is that very small firms are naturally centralized as business owners are typically involved in all activities in the firm, while larger firms implement division of labor to economize on the business owner's time. The last column shows that, on average, not much more than one layer is involved in decision-making for most activities.

A particular advantage of the WES survey format is that it allows us to clearly separate principals from agents. Principals are defined as residual claimants of profit flows from the firm. In particular, since most firms are single-establishment entities, the separation between professional senior managers and business owners is important for identifying principals. Furthermore, for multi-unit firms decision-makers outside the establishment are typically HQ, so that we identify this HQ with the principal. On the other hand, agents are defined by any type of employee, including managers and non managerial employees. For the precise definition of our measures of the degree of decentralization, let us begin with defining the following sets: $D_{\texttt{Pcl}}$ is the set of activities/tasks that principals are involved in, $D_{\texttt{Mgr}}$ is set of activities that management is involved in, $D_{\texttt{NMgr}}$ is set of tasks that non managerial employees are involved in and $D_{\texttt{Agt}} = D_{\texttt{Mgr}} \cup D_{\texttt{NMgr}}$ is set of tasks that agents, that is managerial or non managerial employees are involved in. The allocation of decision control within the firm is therefore given by

$$\texttt{Pcl}_{\texttt{Crtl}} = \sum_d 1_{\left\{d \in D_{\texttt{Pcl}} \cap d \notin D_{\texttt{Agt}}\right\}}$$

$$\texttt{Mgr}_{\texttt{Crtl}} = \sum_d 1_{\left\{d \in D_{\texttt{Mgr}} \cap d \notin D_{\texttt{Pcl}} \cap d \notin D_{\texttt{NMgr}}\right\}}$$

$$\texttt{NMgr}_{\texttt{Crtl}} = \sum_d 1_{\left\{d \in D_{\texttt{NMgr}} \cap d \notin D_{\texttt{Pcl}} \cap d \notin D_{\texttt{Mgr}}\right\}}$$

The survey allows decision tasks to be decided by multiple layers. Since it is unclear how to allocate the actual control in such cases, our benchmark measures of decision

and control are mutually exclusive. This implies that a task is only counted to be decided by the principal, if no other decision layer is involved in the decision[3]

Performance Pay. The WES survey data offers a variety of information on performance-based compensation in firms. Specifically, it allows us to measure four different types of performance pay: individual incentive pay such as bonuses, commissions, piece-rates etc., group or team incentives, profit sharing agreements and stock-based compensation. Additionally, the survey also asks about the presence of merit pay, which is defined as compensation for advanced education such as a master's or a doctoral degree. In principle, the data allows us to measure these types of performance by occupation type such as managers, production workers, sales representatives, administrators, technical support etc. Since most firms adopt similar performance pay for all occupations, we start out using only the data on types of performance pay. To gauge how accurate the survey responses on performance pay are, we regress quantitative compensation measures against the indicators of the presence of different types of performance pay. These measures are average wages, defined as payroll divided by number of employees and average non wage benefits, defined as non payroll labor compensation divided by number of employees. Results are displayed in table 2. As expected, establishments with performance pay systematically pay higher average wages.

Standard principal agent analysis characterized very general forms of state contingent compensation contracts to solve the moral hazard problem. Consequently, we measure the presence of performance pay with an indicator that is one if any form of performance pay is present. We exclude stock compensation from this, since information on stock compensation is completely missing for one year and only a very small fraction of firms offers stock compensation to their employees.

3 Identification and Empirical Results

3.1 Instrumental Variables Strategy

Overview. To identify the degree of complementarity between performance pay and decentralization, we propose the use of regional variation of income taxes across Canadian provinces as an instrument for the adoption of performance pay. The basic idea is that income tax distortions increase the cost of incentive provision through performance pay. To implement a given level of effort, a firm has to pay a higher marginal cost, the higher income taxes are. This idea is succinctly summarized by Roberts (2004): "So, if income tax rates are lowered, it may be more attractive to

[3]In unreported regressions, we obtain similar results with non-exclusive categorizations of decision control.

increase the use of explicit performance pay, since the cost to the firm of providing intensity of incentives is reduced." This logic applies even more generally to income tax progression than the level of income taxes, see Alford (2003). The reason is that, fundamentally, performance pay contracts are nonlinear in that they reward higher marginal reward in high output states. A progressive income tax works exactly to offset this incentive, since high incomes result in higher marginal tax rates, see Gentry and Hubbard (2004). Thus, regional variation in income tax progression can be compared to a "factor price variation" where the factor of production is the adoption of performance pay. Therefore, systematic variation in other management practices, such as decentralization in response to exogenous variation in tax progression is comparable to a cross-price elasticity, in the spirit of Brynjolfsson and Milgrom (2013) and Aghion, Bloom and Van Reenen (2013). As emphasized by Aghion et al. (2013), for this identification strategy to be valid, our exclusion restriction is equivalent to the assumption that progressive income taxes do not directly impact decentralization decisions, except through the adoption of incentive pay. Under this exclusion restriction, we can utilize an IV strategy that provides a valid reduced form test of organizational complementarity, as argued by Arora (1996), Holmstrom and Milgrom (1994) and Athey and Stern (1998).

Income Tax Progression in Canada. There are several features that make Canada an ideal environment to apply the logic of our IV estimation. First, in general, the levels of income taxation in Canada are far higher than for example in the US, thereby causing stronger distortions in the economy. Although the top statutory income tax rates at the federal level in Canada are in fact lower than the top federal income tax in the US, Murphy, Clemens and Veldhuis (2013) indicate that tax brackets are uniformly lower, so that top marginal taxes tend to apply even at relatively modest income levels. Second, regional variation in income taxes is more important in Canada than in comparable federal countries. Murphy et al. (2013) state that "Compared to citizens in other peer countries, Canadians tend to pay a lower share of their total income taxes to the federal government, and a greater share to provincial governments." Since Canadian provinces have full autonomy on both income tax rates and brackets, these provincial income taxes also tend to be applicable at relatively modest income levels.

Construction of IV. Specifically, we use information on income tax progression in the panel of Canadian provinces from 1970 onwards. This data is from the annual publication "Finances of the Nation" and allows us to compute income taxes by province. Our measure of tax progression is residual income progression, a standard measure used in public finance to summarize tax progressivity, see Jakobsson (1976)

and Musgrave and Musgrave (1989). It is defined as

$$\rho = \frac{1 - \texttt{MTR}}{1 - \texttt{ATR}}$$

where MTR is the marginal tax rate and ATR the average tax rate. In this measure, $\rho = 1$ equals a flat tax system, while $\rho < 1$ implies that the tax system exhibits progressivity. Consequently, higher values of ρ imply less progressivity and will imply a higher incentive to adopt performance pay.

We adjust this basic measure of tax progression in two ways. First, the raw data on tax progression exhibits year-to-year fluctuations due to small random changes in statutory tax rates. Since we expect organizational changes, such as the implementation of performance pay, to exhibit long time lags, we smooth these annual fluctuations by taking 10-year averages of the tax progression measures. Second, for multi-province firms, we assign tax progression of the headquarter province as reference tax progression. We do this since the adoption of management practices in our data are mostly firm-wide. Firms that operate in multiple provinces have most of their establishments in the headquarter province, so it seems reasonable to assume that if firms adopt company-wide management practices, they use the same practices in other regions[4]. Note that the bulk of firms are mostly single-establishment firms. We will explicitly analyze multi-province firms in the robustness section.

3.2 Specification

Here, we describe the precise econometric specification as a backdrop for our discussion of empirical results. As discussed earlier, the endogenous variable of interest is the adoption of performance pay. We measure this management practice using an indicator denoted by $\texttt{PFP}_{i,t}$, for establishment i at time t, which is one if there is any form of "pay for performance." We are interested in the interaction of performance pay with several measures for decentralization, which we denoted as $\texttt{Pcl}_{\texttt{Crtl},i,t}, \texttt{Mgr}_{\texttt{Crtl},i,t}, \texttt{NMgr}_{\texttt{Crtl},i,t}$. As noted by Athey and Stern (1998), the possibility of correlated unobserved adoption costs across firms renders raw correlations invalid for tests of complementarity. Hence, based on the idea that income tax progression is a valid instrument for the endogenous variable $\texttt{PFP}_{i,t}$, the first stage of the IV regression is

$$\texttt{PFP}_{i,t} = \alpha_\rho \cdot \rho_{i,t} + \texttt{controls}_{i,t} + \varepsilon_{1,it}$$

where $\rho_{i,t}$ is the residual income progression measure described above. Since higher values of ρ capture less progressive income taxes, one would expect that $\alpha_\rho > 0$;

[4]In unreported regressions, we also assigned tax progressions based on the actual establishment locations. When tax progression variables based on actual location and on HQ location are included together to predict performance pay, only HQ-based tax progression measures are highly significant, while tax progression based on actual location is insignificant.

thus, less progression makes implementation of performance pay more likely. If the exclusion restriction holds, then one can use this first stage to obtain the following second stage IV estimator

$$\texttt{Pcl}_{\texttt{Crtl},i,t} = \beta_{\texttt{PFP}} \cdot \widehat{\texttt{PFP}}_{i,t} + \texttt{controls}_{i,t} + \varepsilon_{2,it}$$

where $\beta_{\texttt{PFP}}$ is the coefficient of interest. In particular, $\beta_{\texttt{PFP}} > 0$ implies that adoption of performance pay increases decision control by the principal. In this case, decentralization and performance pay would be substitutes. On the other hand, $\beta_{\texttt{PFP}} < 0$ implies that adoption of performance pay leads to less control by the principal and more delegation of decisions to employees. In this case, decentralization and performance pay are complements.

We use a number of standard variables to control for various obvious unobservable confounders. Among them are industry and year fixed effects to control for aggregate differences across industries or economy-wide shocks. Furthermore, we control for firm size as measured by the number of employees, multi establishment firms, establishment age, and exporter status. These are first-order controls as one might expect that firms have a natural tendency to decentralize activities the larger they grow; simultaneously, greater presence of more employees might make the implementation of performance pay more cost-effective, if the latter is associated with fixed costs of adoption.

We briefly note two properties of the estimator in use. First, to avoid running "forbidden regressions", we use a linear probability model for the first stage as recommended by Angrist (2001) and Angrist and Pischke (2009). Second, we note that since our estimator is just-identified, any weak instruments problem is likely to be diagnosed simply using correctly calculated second-stage standard errors, see Angrist and Pischke (2009). Table 3 provides first-stage estimates to ensure that the first stage is not insignificant, which is the only case for which the weak instruments problems could matter in the just-identified case. These estimates show that both in probit and linear probability models, the first stage is highly significant and exhibits the expected sign of α_ρ.

3.3 Results

As argued in section 1, predictions of a simple theory of decentralization with moral hazard are ambiguous on whether decentralization and performance pay are complements or substitutes. Our baseline results in table 4 suggest that they are complements. The first three columns of this table provide simple OLS estimates of our measures of decision control and the adoption of performance pay. The raw regression coefficients on performance pay are compatible with the presence of this complementarity, but as Athey and Stern (1998) indicate, they could also be generated by a positive correlation of adoption costs for both management practices. The

three right hand side columns report IV results, which confirm the presence of complementarity performance pay and the decentralization of decisions from principals to agents. In response to exogenous adoptions of performance pay, firms reallocate real decision authority systematically away from principals down the hierarchy. Note that the coefficient estimates for the IV estimates are mostly larger than the OLS estimates, thereby suggesting that unobservable costs of adoption are mostly not strongly positively correlated, as discussed by Athey and Stern (1998).

A notable feature of our empirical approach is our ability to separate agents into managerial employees, such as senior management and work supervisors, and non managerial employees, such as production workers and sales representatives. This allows us to answer the question of how far down the hierarchy firms decentralize in response to the adoption of performance pay. It is evident in table 4 that firms tend to concentrate decision making at the management level. That is, although they decentralize decision making down from principals to agents in general, they also centralize decision making from non managerial employees to managers. This is surprising for several reasons. First, one might have thought that since performance pay enables more efficient decision making at all levels of the organization, that adoption of performance pay induces a general movement toward decentralization. This prediction would have implied that the performance pay coefficients for both, $\texttt{Mgr}_{\texttt{Crtl},i,t}$ and $\texttt{NMgr}_{\texttt{Crtl},i,t}$ are positive, which is similar to the OLS estimates. Second, there exist popular views that middle management is a wasteful bureaucratic layer. This view emphasizes the importance of worker empowerment, so one would have expected that firms reallocate tasks from the management to non managerial employees. In that case, the coefficient on $\texttt{Mgr}_{\texttt{Crtl},i,t}$ should have been negative, while the coefficient on $\texttt{NMgr}_{\texttt{Crtl},i,t}$ would have been positive.

To support these basic findings on the concentration of decision control at the management level, we move beyond the analysis of decision control and analyze labor demand patterns in firms implementing performance pay. Table 5 shows the impact of an exogenous adoption of performance pay on the occupational composition of firms. In particular, our dependent variable is the proportion of the workforce at an establishment, that comprises of one of five occupational classes: managers (`PctMgr`), production workers (`PctPrdw`), professionals (`PctPrf`), sales representatives (`PctSal`), and administrative staff (`PctAdm`). The key result is that establishments implementing performance pay employ more managers as a proportion of their workforce. So not only do firms that adopt performance pay driven by exogenous reasons allocate more decision responsibility to management, they also maintain larger pools of managers to meet these responsibilities.

One question is whether firms that implement performance pay do not actually employ more managers, but might tend to lay-off non-managers, so that the fraction of non-managers in the workforce is higher. A simple mechanism through which this could occur is that managers begin cutting costs and fire non managerial employ-

ees. This might explain why non managerial employees are given less control over decisions – there might simply be less of them at the firm. However, table 6 shows that firms implementing incentive pay do not have systematically higher layoff or quit rates. Moreover, overall hiring rates at these establishments are also not significantly lower. Instead, the evidence suggests that firms that adopt performance pay shift their hiring towards managers, as is evident in table 7. This evidence is compatible with a view that the largest gains from implementation of performance pay occur through more efficient decision making of managers. Consequently, firms tend to not only concentrate control at the management level, but also tend to hire more managers to deal with the increased responsibilities. Overall, the data are compatible with the view that management plays a key role at the intersection of production and non-production decisions. Our findings are also supportive of the hypothesis advanced by Atalay et al. (2013) that the nature of firms is intimately connected to their role in "mediating managerial supervision and control."

3.4 Extension: Organizational Change

We documented how firms that adopt performance pay for exogenous reasons, concentrate control at the management level and hire more professional managers. However, what do these managers do with their wider range of responsibilities? This section attempts to provide some preliminary answers to this question. While we find this question intriguing, we have to caution readers at this point that our the wider application of our IV strategy to variables other than decision control potentially threatens identification. Depending on the dependent variable, tax progressivity might have a direct impact on organizational changes beyond its influence through performance pay. However, the plausibility of such a direct mechanism must be decided on a case-by-case basis.

Table 8 summarizes the IV results for seven management practices: decrease in the degree of centralization, re-engineering of business processes, downsizing, delayering of managerial hierarchies, TQM and outsourcing. In each of these cases, the dependent variable is a survey response to the question: "Has your workplace experienced any of the following forms or organizational change... (in the past year)?" An affirmative answer is coded as one, while a negative answer is coded as zero. The first notable result that provides another validity check on our results is that establishments that implement performance pay do not seem to systematically decrease the degree of centralization. This seems puzzling in light of the clear movement of the delegation of decision making from principals to agents in general. However, it is in fact completely consistent with the counter-movement of a centralization of decisions from non managerial employees to management.

One of the management practices more frequently used by managers when adequately incentivized with performance pay is business process re-engineering (or

re-engineering). We note that this question in the survey provides a brief description of this management practice as "redesigning processes to improve performance and cost." After being popularized in the US by Hammer (1990), business process re-engineering has often been criticized as a cover for downsizing instead of a valid approach to improve firm performance. While a comprehensive evaluation of re-engineering is beyond the scope of our study, our data does allow us to investigate whether managers implementing re-engineering downsize their workforce simultaneously. In particular, the survey defines downsizing as "reducing the number of employees on payroll to reduce expenses: it is part of a reorganization in the workplace and not simply a response to a drop in demand." The fourth column in table 8 shows that although firms implementing performance pay are more likely to adopt re-engineering, they are not more likely to downsize. The latter result is consistent with our earlier findings regarding the overall layoff and quit rate at firms. To summarize, performance pay seems to induce managers to seek more efficient business processes, but does not necessarily lead them to downsize the workforce.

It is also worth highlighting that exogenous implementation of performance pay leads to more likely adoption of TQM. This is surprising, since a widespread assumption regarding the effective implementation of this management practice is that it is complementary to worker empowerment or more decentralization towards non managerial employees. In contrast, our empirical results suggest that worker empowerment and TQM are not necessarily complements, but are potentially substitutes.

Up to this point, much of our discussion of the knock-on effects of implementation of performance pay on other management practices was focused on inferring information regarding which decisions managers might take. But are there possible implications of the implementation of performance pay for the number of managerial hierarchies themselves? A natural starting point for this question might be a version of Garicano (2000), in which the number of hierarchies is endogenously determined in contrast to our baseline model that keeps the number of hierarchical levels fixed. According to Garicano (2000), an organization will be able to deal with more problems and responsibilities if the number of layers is high. Therefore, one might conjecture that an increased concentration of decisions at the management level might also increase the number of managerial layers. In the last column of table 8 reduction in the number of managerial levels (delayering) is used as a dependent variable. It shows that the simple intuition from a model of endogenous hierarchies is not borne out in the data. Hence, although management exhibits greater control over decisions in the firm once performance pay is implemented, it does not deal with these responsibilities by adding more layers in the organization. On the contrary, there is a systematic tendency to delayer in these firms.

4 Robustness

In this robustness section we address two potential concerns. First, given that most of the exogenous variation in our instrument is driven by regional differences, is it possible that unobserved omitted variables at the local level render our identification invalid? Second, are there possible alternative explanations for the specific decentralization patterns rather than complementarity between decentralization and performance pay?

4.1 Within-Province Variation

To explore whether unobserved regional factors drive our results, we focus on the sample of firms that operate across multiple provinces. For these firms, we include a full set of province fixed effects that controls for unobserved factors in their location. At the same time, since their assigned tax progression measures are based on the HQ region, the instrument is still valid. In other words, we compare establishments which are part of multi-province firms within the same location and predict that establishments with headquarters in high tax progression regions have a lower likelihood to adopt performance pay than establishments in the same location that have headquarters in low tax progression regions. This alternative IV strategy takes full account of unobserved confounders in an establishment's location without compromising our identification strategy. The left columns in table 9 report the results. The estimation results are comparable to our benchmark estimates. Both the coefficient on decision tasks controlled by the principal and those tasks controlled by non-managerial employees stay significant. Note also that the coefficient on the manager tasks stay significant at the 10% level, although we cut our sample in this robustness check by 90%.

4.2 Industry-Province Interaction IV

Another way to address possible concerns about only using regional variation in our IV strategy is to exploit more industry level variation. To this end, we use 4-digit industry variation in the share of firms generating product innovations in 1999. As in Holmstrom (1989), the idea is that employee activities that are important for innovation are hard to monitor. As a result, firms that operate in industries where innovation is important are more likely to adopt performance pay, if tax progression is low. We will therefore use regional tax progression and the interaction of the share of innovating firms and tax progression as our instrument. Since the IV variation is now also based on 4 digit industry variation, we cluster our standard errors by year, region and industry, instead of the broad year-region clusters. The Craig-Donald F-statistic to test for weak IVs has a value of 32.75 and the first stage is as before,

highly significant. The results can be seen in the right panels of table 9. Note that our coefficient estimates are quantitatively similar to our benchmark results in table 4. On the other hand, standard errors do increase somewhat, indicating that the industry variation does add some noise.

4.3 Alternative Explanations

We argued that using provincial differences in income tax progression can be used to instrument for the adoption of performance pay, even in the presence of unobserved heterogeneity in adoption cost, as emphasized by Athey and Stern (1998). If our identification strategy is valid, the IV estimator recovers a valid estimate of the degree of complementarity between performance pay and decentralization. Here, we explore a number of alternative explanations for our results and argue that our basic findings are robust to extensions seeking to control for these explanations.

- *Industry trends in technology or competition.* One unobservable factor influencing performance pay and decentralization are unobserved industry trends in technology and competition. For example, suppose firms adopt better computer technology that allows firms to decentralize task more easily. At the same time, this technology might provide an independent noisy signal on the effort of employees, effectively lowering the cost of performance pay. Consequently, firms might implement decentralization and performance pay, driven by industry trends in technology. An alternative plausible industry trend is the intensification of competition. On one hand, this could lead to decentralization of decision making, as argued by Bloom, Sadun and Van Reenen (2010). On the other hand, competition might induce principals to adopt performance pay, as in the model of Raith (2003). To control for these types of mechanisms, we include a full set of industry-time interactions to capture unobserved industry trends in competition and technology. In case competition is not only national, but has an important regional component, we also use subjective data on competition. The WES asks respondents to indicate whether they think that local competition is important. We therefore include a dummy `local cmp`, which is one if the firm perceived local competition to be of relevance.

- *Overall number of tasks and organizational layers.* Firms might also differ in terms of the degree to which they face different complexity of problems overall. For example, some firms with simple business processes might not have to formally allocate certain decisions, such as customer service, to the business owner or employees. Other firms with complicated Just-In-Time production will need to assign a large number of tasks to decision-makers; thus, it is natural for them to decentralize. Further, such firms will also tend to implement

performance pay, because of moral hazard problems in these complex business processes. We use two proxies to control for this issue. The first is a direct measure of the overall number of tasks the firm is involved in. The second is a proxy for the number of decision layers the firm has. We base this measure on ideas of Garicano (2000) that organizations with more layers economize on managers time and therefore are able to solve more complex problems. We construct this measure in the following manner: for each possible decision layer, we check whether the firm has assigned this layer a decision task across out 12 categories. If it has done so, we infer that this layer exists. If not, we presume that the firm does use this layer in decision making problems and hence we ignore it.

- *Unobserved employee skill.* Employee skill can lower costs of both decentralization and performance pay. Highly skilled employees might be more likely to require little training to make decisions on relatively complicated tasks. Further, performance pay might then just be present to provide high wage payments to very skilled employees to induce a sorting toward high-skill employees, as shown by Lazear (2000). To control for this explanation, we include the ratio of payroll to number of total workers and the number of workers in the to earnings bracket to proxy for the presence of highly skilled workers.

- *Novelty business strategy.* The joint adoption of performance pay and decentralization might also result from a firm's business strategy that is based on innovating and introducing new products. For example, a firm trying to innovate might decentralize decision making to benefit from information acquisition of employees, as in Aghion and Tirole (1997). At the same time, the firm might want to implement performance pay to give employees incentives to participate in innovation. To control for this mechanism, we include a dummy that takes the value of one if a firm's top priority in its business strategy is innovation. This measure is more fully developed and explained in Yang, Kueng and Hong (2014).

- *Risk.* In response to empirical studies suggesting that there is a positive correlation between risk and performance pay, theoretical models such as Prendergast (2002) have argued that risk can drive both, decentralization and performance pay. Therefore, a potential confounder potentially biasing results towards a finding of complementarity between performance pay and decentralization could be risk. We include the standard deviation of operating margin at the establishment level as an explicit measure of risk.

- *French culture.* For most of our sample, it is safe to assume that the cultural differences across Canadian provinces are minor. The obvious exception

is Quebec. There are several ways in which the presence of French culture in Canada could have an influence on our results. First, regarding decentralization, Bloom et al. (2012) note that there is a tendency to centralize decision-making in French culture. Second, French culture as part of continental European culture has been argued to exhibit a larger degree of inequity aversion based on different equilibrium beliefs regarding luck vs. meritocracy. See Alesina and Angeletos (2005). This inequity aversion might prevent firms under French influence to implement performance pay. We directly control for the effects of French culture by including a dummy for the location of an establishment being in Quebec.

Tables 10-12 display the results from the robustness checks of our results. The specifications include all the standard controls we used in previous specifications. These controls are omitted here due to space constraints. As these tables show, all baseline results are robust, whether the control variables are included sequentially or all together.

Conclusion

We provided first population level evidence on management practice complementarity, combining regional variation on income tax progression with firm level data on internal organization from Canada. We see promise in both the novel data we exploited as well as in the theoretical possibilities opened by our estimates.

Our estimates of the average degree of complementarity between performance pay and decentralization are potentially useful to calibrate quantitative models of endogenous productivity and management practices. One example of such models are the agent-based models, such as Rivkin (2000). An alternative avenue that we currently explore is to embed management practice complementarity in a model of firm dynamics with financial frictions. In such a setup, management practice complementarity is likely to strongly interact with credit frictions, so that credit constraints can have a large impact on firm productivity differences and aggregate productivity.

We are also currently working on providing direct estimates of the importance of complementarity for productivity. As Athey and Stern (1998) show, if management practice complementarity is important, then this will have important consequences for productivity estimation. Complementary management practices have to be evaluated together to recover valid estimates of the productivity gains from adopting management practices. We plan to address this issue by combining our IV strategy with the structural estimator proposed by Athey and Stern (1998) and Kretschmer et al. (2012) to fully characterize the impact of performance pay and decentralization on productivity.

One of the most notable findings of this study where the patterns of decentralization implied by an exogenous adoption of performance pay. These results suggests the deployment of professional management is a key function of firms in the spirit of Atalay et al. (2013). In Hong, Kueng and Yang (2014) we explicitly test their hypothesis by analyzing whether multi-establishment, multi-sector firms deploy managerial resources differently from multi-establishment, single-sector or single-establishment firms. This research not only complements the decentralization patterns documented here, but also promises additional insights into the nature of firms.

Table 1: Summary statistics on the average and total number of hierarchical layers

	Dependent Variable		
Tasks/Activities	Avg. layer	Avg. layer (ex. bus. owner)	No. of layers
Daily workplanning	3.24	2.85	1.16
Weekly workplanning	3.32	2.97	1.17
Follow-up of results	3.56	3.25	1.17
Customer relations	3.47	3.2	1.25
Quality control	3.55	3.37	1.25
Purchase of supplies	3.34	3.09	1.17
Equipment maintenance	3.4	3.13	1.17
Set staffing levels	3.87	3.71	1.07
Filling vacancies	3.84	3.65	1.08
Training	3.63	3.5	1.17
Production technology choice	3.86	3.84	1.09
Product/service development	3.85	3.8	1.12

Notes: Organizational layers are coded as 1 (non managerial employees), 2 (work group), 3 (work supervisor), 4 (senior manager), 5 (HQ), 6 (business owner). All data is for the 2003 cross section. The first column displays average layer involved in task across 5,300 firms. The second column displays the average number of layers involved in a task across 2,200 firms where the business owner is not involved at all in decision making. All summary statistics are weighted using survey weights to make results representative for all business firms in Canada. The third column displays the number of layers involved in a given task, averaged across all firms.

Table 2: Performance pay and average establishment level wage and non-wage benefits

	Dependent Variable	
Independent Variable	`avg wage`	`avg non-wage benefits`
`Ind Incen`	4.479***	0.131
	(0.887)	(0.139)
`Team Incen`	2.205*	0.262
	(1.163)	(0.179)
`Profit Sharing`	3.364**	0.171
	(1.320)	(0.162)
`Merit`	0.493	0.138
	(1.267)	(0.252)
`Stock`	4.182*	0.852*
	(2.385)	(0.454)
$\log(size)$	0.934**	-0.121
	(0.409)	(0.103)
$\log(age)$	1.382***	0.469***
	(0.375)	(0.084)
`multi-unit`	1.695	1.079**
	(1.702)	(0.425)
`exporter`	2.374**	0.093
	(0.981)	(0.151)
Industry FE	Yes	Yes
Time FE	Yes	Yes
N	17,104	12,568

Notes: Dependent variables are: payroll/(total employees) in column 1 and (non-payroll labor compensation)/(total employees) in column 2. Independent variables are dummies for different types of performance pay and control variables. Standard errors clustered by sampling strata, which are broad industry-size-region categories. Regressions use sampling weights. Additional controls: establishment size, age, multi-unit flag and exporter status. Data is for years 2001, 2003 and 2005, since 1999 does not include data in stock compensation.

Table 3: IV First Stage: Tax progression and adoption of performance pay

	Dependent Variable	
Independent Variable	`Probit PFP`	`Linear Prob PFP`
ρ (`Tax Progression`)	5.435***	1.433***
	(0.658)	(0.130)
$\log(size)$	0.488***	0.140***
	(0.019)	(0.005)
$\log(age)$	-0.042**	-0.011**
	(0.022)	(0.005)
`multi-unit`	0.460***	0.155***
	(0.082)	(0.024)
`exporter`	0.087	0.028
	(0.065)	(0.020)
Industry FE	Yes	Yes
Time FE	Yes	Yes
N	22,285	22,285

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variable is dummy for the adoption of performance pay. Measure of tax progression is residual income progression as defined in the text, where tax region is defined to be location of HQ. Years include 1999, 2001, 2003 and 2005.

Table 4: Decentralization and Performance Pay: OLS and IV Estimates

	OLS			IV		
Independent Variable	$\texttt{Pcl}_{\texttt{Crtl}}$	$\texttt{Mgr}_{\texttt{Crtl}}$	$\texttt{Nmgr}_{\texttt{Crtl}}$	$\texttt{Pcl}_{\texttt{Crtl}}$	$\texttt{Mgr}_{\texttt{Crtl}}$	$\texttt{Nmgr}_{\texttt{Crtl}}$
`Performance Pay`	-1.943***	0.892***	0.049	-5.838***	5.290***	-1.966***
	(0.359)	(0.191)	(0.089)	(1.128)	(1.187)	(0.733)
$\log(size)$	-1.749***	0.622***	-0.134**	-1.192***	-0.000	0.147
	(0.119)	(0.095)	(0.052)	(0.210)	(0.124)	(0.121)
$\log(age)$	0.042	0.020	0.065*	-0.096*	0.141***	0.007
	(0.090)	(0.097)	(0.033)	(0.058)	(0.043)	(0.037)
`multi-unit`	-0.343	-0.646***	-0.329***	0.353	-1.427***	0.008
	(0.236)	(0.248)	(0.114)	(0.252)	(0.292)	(0.162)
`exporter`	0.383	-0.373**	0.064	0.348	-0.402**	0.036
	(0.323)	(0.187)	(0.084)	(0.269)	(0.197)	(0.143)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,684	11,684	11,684	11,684	11,684	11,684

Notes: Standard errors clustered by sampling strata for OLS, which are broad industry-size-region categories. Clustering for IV is by region and year. Regressions use sampling weights. Dependent variables are number of decisions made by principals (`Pcl`), management (`Mgr`) and non managerial employees (`Nmgr`). OLS-columns use dummy for performance pay as independent variable. IV-columns use a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 2003 and 2005 as information distinguishing business owners from senior management is only available in these years.

Table 5: IV Regressions: Occupational Composition

	Dependent Variable				
Independent Variable	PctMgr	PctPrdw	PctPrf	PctSal	PctAdm
Performance Pay	0.300***	-0.233***	0.046*	0.014	-0.007
	(0.060)	(0.090)	(0.027)	(0.038)	(0.027)
log(*size*)	-0.088***	0.076***	-0.003	0.010	-0.018**
	(0.010)	(0.013)	(0.005)	(0.007)	(0.007)
log(*age*)	0.002	0.005	-0.003	-0.010***	0.009**
	(0.005)	(0.006)	(0.002)	(0.003)	(0.004)
multi-unit	-0.065***	0.040	-0.005	0.041***	-0.045***
	(0.018)	(0.024)	(0.009)	(0.015)	(0.013)
exporter	-0.006	0.005	0.025***	-0.012	-0.006
	(0.010)	(0.015)	(0.007)	(0.010)	(0.007)
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
N	22,285	22,285	22,285	22,285	22,285

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variables are percentages of different occupations: managers (PctMgr), production workers (PctPrdw), professionals PctPrf, salesworkers PctSal, and administrative staff (PctAdm). Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 1999, 2001, 2003 and 2005.

Table 6: IV Regressions: Hiring and firing

	Dependent Variable			
Independent Variable	`Layoff rate`	`Quit rate`	`Dismissal rate`	`Hiring rate`
`Performance Pay`	-0.139	0.029	-0.002	-0.123
	(0.100)	(0.088)	(0.040)	(0.162)
log($size$)	-0.009	-0.006	0.001	0.015
	(0.016)	(0.020)	(0.005)	(0.030)
log(age)	-0.008	-0.008	-0.004*	-0.044***
	(0.007)	(0.010)	(0.002)	(0.012)
`multi-unit`	0.022	0.021	-0.008	0.014
	(0.015)	(0.025)	(0.010)	(0.036)
`exporter`	0.019	-0.008	-0.006	-0.022
	(0.016)	(0.017)	(0.005)	(0.023)
Industry FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
N	22,285	22,285	22,285	22,285

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variables are gross labor turnover rates: layoffs relative to labor force, quits relative to labor force, dismissals relative to labor force and hires relative to labor force. Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 1999, 2001, 2003 and 2005.

Table 7: IV Regressions: Hiring rates by occupation

	Dependent Variable				
Independent Variable	HrateMgr	HratePrdw	HratePrf	HrateSal	HrateAdm
Performance Pay	0.039**	-0.230**	0.009	0.124***	-0.001
	(0.017)	(0.091)	(0.013)	(0.047)	(0.024)
log(*size*)	-0.004*	0.032**	0.002	-0.005	-0.003
	(0.002)	(0.015)	(0.003)	(0.006)	(0.003)
log(*age*)	-0.006***	0.005	-0.004***	-0.018***	-0.005
	(0.001)	(0.009)	(0.001)	(0.004)	(0.004)
multi-unit	-0.014**	0.033	-0.004	-0.011	-0.009
	(0.006)	(0.025)	(0.004)	(0.018)	(0.009)
exporter	0.003	-0.014	0.007***	-0.016*	0.003
	(0.005)	(0.014)	(0.003)	(0.010)	(0.006)
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
N	22,285	22,285	22,285	22,285	22,285

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variables are hiring rates for different occupations, defined as hires relative to all employees: managers (HrateMgr), production workers (HratePrdw), professionals (HratePrf), sales workers (HrateSal) and administrative staff (HrateAdm). Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 1999, 2001, 2003 and 2005.

Table 8: IV Regressions: Organizational Change

	Dependent Variable					
Independent Variable	Decentralization	Reengineering	Downsizing	TQM	Outsourcing	Delayering
Performance Pay	-0.017	0.334***	0.026	0.183***	0.184*	0.063***
	(0.031)	(0.066)	(0.054)	(0.062)	(0.100)	(0.023)
log(*size*)	0.016***	0.027***	0.014*	0.020**	-0.002	0.006*
	(0.004)	(0.009)	(0.009)	(0.008)	(0.012)	(0.003)
log(*age*)	-0.007**	-0.008	-0.003	-0.019***	-0.005	0.003
	(0.003)	(0.007)	(0.005)	(0.006)	(0.003)	(0.002)
multi-unit	0.038**	-0.004	0.024	-0.011	-0.028	0.022
	(0.018)	(0.032)	(0.019)	(0.026)	(0.021)	(0.016)
exporter	0.004	0.047***	0.024**	0.014	0.048***	0.000
	(0.006)	(0.014)	(0.012)	(0.012)	(0.012)	(0.008)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
N	22,285	22,285	22,285	22,285	22,285	22,285

Notes: Standard errors clustered by sampling strata, which are broad industry-size-region categories. Regressions use sampling weights. Dependent variables are dummies for organizational changes, described in the main text. Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 1999, 2001, 2003 and 2005

Table 9: Decentralization and Performance Pay: Within-province and Interaction-IV estimates

	within-province			interaction IV		
Independent Variable	$\texttt{Pcl}_{\texttt{Crtl}}$	$\texttt{Mgr}_{\texttt{Crtl}}$	$\texttt{Nmgr}_{\texttt{Crtl}}$	$\texttt{Pcl}_{\texttt{Crtl}}$	$\texttt{Mgr}_{\texttt{Crtl}}$	$\texttt{Nmgr}_{\texttt{Crtl}}$
`Performance Pay`	-8.495***	3.225*	-1.975**	-5.573***	4.747***	-1.900**
	(2.089)	(1.918)	(0.925)	(1.378)	(1.459)	(0.743)
log($size$)	-0.105	0.315	0.023	-1.229***	0.075	0.137
	(0.247)	(0.320)	(0.102)	(0.224)	(0.188)	(0.110)
log(age)	0.419	-0.221	-0.007	-0.088	0.124	0.009
	(0.327)	(0.306)	(0.083)	(0.086)	(0.077)	(0.045)
`multi-unit`	-1.448***	-0.126	-0.299	0.306	-1.330***	-0.004
	(0.364)	(0.646)	(0.279)	(0.412)	(0.305)	(0.165)
`exporter`	0.772	-0.464	0.048	0.349	-0.403**	0.036
	(0.569)	(0.587)	(0.173)	(0.296)	(0.201)	(0.126)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	No	No	No
Observations	1,540	1,540	1,540	11,684	11,684	11,684

Notes: Standard errors for the within-province regressions are clustered by region and year. Standard errors for the interaction-IV regressions are clustered by region-year-industry. Regressions use sampling weights. Dependent variables are number of decisions made by principals (`Pcl`), management (`Mgr`) and non managerial employees (`Nmgr`). IV-columns use a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 2003 and 2005 as information distinguishing business owners from senior management is only available in these years.

Table 10: IV Regressions: Robustness (1)

	Dependent Variable: $\texttt{Pcl}_{\texttt{Crtl}}$							
Performance Pay	-5.630***	-5.661***	-6.227***	-5.212***	-5.569***	-5.369***	-8.630***	-7.640***
	(1.125)	(1.125)	(0.952)	(1.360)	(1.093)	(0.855)	(1.828)	(1.573)
local cmp		0.047						0.168
		(0.133)						(0.196)
taskscope			0.788***					0.732***
			(0.035)					(0.037)
decision layers			-1.294***					-1.139***
			(0.152)					(0.167)
avg wage				-0.012**				-0.002
				(0.005)				(0.005)
Pct80K+				1.978**				2.498**
				(0.886)				(1.016)
$\texttt{novelty}_{top}$					0.228			0.540*
					(0.260)			(0.323)
Risk						42.844		23.482
						(37.670)		(30.031)
$I_{\{\texttt{Quebec}\}}$							-0.826***	-0.490*
							(0.213)	(0.276)
Ind× Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,684	11,684	11,602	11,684	11,442	11,684	11,684	11,301

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variable is number of decisions made by principals. Main controls capture various alternative mechanisms and are discussed in the main text. Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 2003 and 2005 as information distinguishing business owners from senior management is only available in these years. Additional controls not displayed in the table but included in the regression: log size, log age, multi-unit flag, exporter dummy.

Table 11: IV Regressions: Robustness (2)

	Dependent Variable: $\texttt{Mgr}_{\texttt{Crtl}}$							
Performance Pay	5.140***	5.141***	4.878***	5.288***	5.047***	5.006***	6.276***	5.697***
	(1.012)	(1.002)	(0.833)	(1.128)	(0.949)	(1.037)	(1.990)	(1.675)
local cmp		-0.295**						-0.324**
		(0.121)						(0.150)
taskscope			0.257***					0.290***
			(0.040)					(0.046)
decision layers			-0.741***					-0.833***
			(0.171)					(0.213)
avg wage				-0.003				0.000
				(0.003)				(0.004)
Pct80K+				-0.887				-1.273*
				(0.686)				(0.731)
$\texttt{novelty}_{top}$					-0.345			-0.271
					(0.314)			(0.352)
Risk						78.294***		76.689***
						(24.843)		(24.466)
$I_{\{\texttt{Quebec}\}}$							0.313	0.300
							(0.265)	(0.235)
Ind× Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,684	11,684	11,602	11,684	11,442	11,684	11,684	11,301

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variable is number of decisions made by managerial employees. Main controls capture various alternative mechanisms and are discussed in the main text. Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 2003 and 2005 as information distinguishing business owners from senior management is only available in these years. Additional controls not displayed in the table but included in the regression: log size, log age, multi-unit flag.

Table 12: IV Regressions: Robustness (3)

	Dependent Variable: $\texttt{Nmgr}_{\texttt{Crtl}}$							
Performance Pay	-2.011***	-2.023***	-1.912***	-2.184***	-1.992***	-1.984***	-3.669***	-3.909***
	(0.668)	(0.679)	(0.606)	(0.728)	(0.637)	(0.649)	(1.186)	(1.220)
local cmp		0.017						0.111
		(0.048)						(0.077)
taskscope			-0.011					-0.036
			(0.018)					(0.028)
decision layers			0.445***					0.598***
			(0.070)					(0.116)
avg wage				0.006***				0.006**
				(0.002)				(0.003)
Pct80K+				0.158				0.769*
				(0.316)				(0.426)
$\texttt{novelty}_{top}$					0.070			0.213
					(0.142)			(0.204)
Risk						-38.714**		-36.589
						(17.316)		(26.874)
$I_{\{\texttt{Quebec}\}}$							-0.457***	-0.547***
							(0.152)	(0.178)
Ind× Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,684	11,684	11,602	11,684	11,442	11,684	11,684	11,301

Notes: Standard errors clustered by region and year. Regressions use sampling weights. Dependent variable is number of decisions made by non managerial employees. Main controls capture various alternative mechanisms and are discussed in the main text. Specification is a LIML IV estimator with tax progression as instrument and the linear probability first stage displayed in table 3. Years are 2003 and 2005 as information distinguishing business owners from senior management is only available in these years. Additional controls not displayed in the table but included in the regression: log size, log age, multi-unit flag.

References

Acemoglu, D., U. Akcigit, N. Bloom, W. Kerr, and J. Van Reenen, "Innovation, reallocation and growth," *mimeo, Stanford University*, 2013.

Aghion, P. and J. Tirole, "Formal and real authority in organizations," *Journal of Political Economy*, 1997.

_ , N. Bloom, and J. Van Reenen, "Incomplete contracts and the internal organization of firms," *Journal of Law, Economics and Organization*, 2013.

Alesina, A. and M. Angeletos, "Fairness and Redistribution: US vs. Europe," *American Economic Review*, 2005.

Alford, S., "The effect of personal income tax on corporate agency costs," *mimeo, Asper School of Business, University of Manitoba*, 2003.

Angrist, J., "Models with dummy endogenous regressors: simple strategies for empirical practice," *Journal of Business and Economic Statistics*, 2001.

_ and S. Pischke, *Mostly harmless econometrics*, Prince, 2009.

Arora, A., "Testing for complementarities in reduced form regressions: a note," *Economics*, 1996.

Atalay, E., A. Hortascu, and C. Syverson, "Vertical integration and input flows," *American Economic Review*, 2013.

Athey, S. and S. Stern, "An empirical framework for testing ttheorie about complementarity in ororganization design," *NBER Working Paper*, 1998.

Baily, M., C. Hulten, and D. Campbell, "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity: Microeconomics*, 1992.

Bester, H. and D. Kremer, "Delegation and Incentives," *Rand Journal of Economics*, 2008.

Bloom, N. and J. Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Jounral of Economics*, 2007.

_ , L. Brynjolfsson E.and Foster, R. Jarmin, I. Saporta, and J. Van Reenen, "Management and IT in America," *Working Paper, Stanford University*, 2014.

_ , L. Garicano, R. Sadun, and J. Van Reenen, "The distinct effects of information technology and communication technology on firm organization," *Management Science*, 2013.

_ , R. Sadun, and J. Van Reenen, "Does product market competition lead firms to decentralize?," *American Economic Review Papers and Proceedings*, 2010.

_ , _ , and _ , "The organization of firms across countries," *Quarterly Jounral of Economics*, 2012.

_ , _ , and _ , "Management as Technology?," *mimeo, Stanford University*, 2013.

Bresnahan, T., E. Brynjolfsson, and L. Hitt, "Information technology, workplace organization and the demand for skilled labor: firm-level evidence," *Quarterly Jounral of Economics*, 2002.

Brynjolfsson, E. and P. R. Milgrom, "Complementarity in Organizations," *Handbook of Organizational Economics*, 2013.

Chassang, S., "Building Routines: Learning, Cooperation and the Dynamics of Incomplete Relational Contracts," *American Economic Review*, 2010.

Drucker, P., *The Concept of the Corporation*, Transaction Publishers, 1949.

Garicano, L., "Hierachies and the organization of knowledge in production," *Journal of Political Economy*, 2000.

_ and E. Rossi-Hansberg, "Organizing Growth," *Journal of Economic Theory*, 2012.

Gentry, W. and G. Hubbard, "The Effects of Progressive Income Taxation on Job Turnover," *Journal of Public Economics*, 2004.

Gibbons, R. and R. Henderson, "What do managers do?," *Handbook of Organizational Economics*, 2013.

_ , R. Holden, and M. Powell, "Organization and information: firms' governance choices in rational-expectations equilibrium," *Quarterly Journal of Economics*, 2012.

Grossman, G. and E. Helpman, "Integration vs. outsourcing in industry equilibrium," *Quarterly Journal of Economics*, 2002.

_ and O. Hart, "An Analysis of the principal-agent problem," *Econometrica*, 1983.

Hammer, M., "Reengineering work: don't automate, obliterate!," *Harvard Business Review*, 1990.

Holmstrom, B. and P. Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," *Jounral of Law, Economics and Organization*, 1991.

_ and _ , "The firm as an incentive system," *American Economic Review*, 1994.

Holmstrom, Bengt, "Agency Costs and Innovation," *Journal of Economic Behavior and Organization*, 1989.

Hong, Bryan, Lorenz Kueng, and Mu-Jeung Yang, "Firms as Allocators of Managerial Supervision and Control," *mimeo, University of Washington*, 2014.

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.

Jakobsson, U., "On the measurement of the degree of progression," *Journal of Public Economics*, 1976.

Jensen, M.C. and W.H. Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structures," *Journal of Financial Economics*, 1976, *3* (4), 305–360.

Jones, Charles I., "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macro*, 2011.

Klette, T. and S. Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004.

Kretschmer, T., E. Miravete, and J. Pernias, "Competitive pressure and the adoption of complementary innovations," *American Economic Review*, 2012.

Krishna, V. and J. Morgan, "Contracting for information under imperfect committment," *Rand Journal of Economics*, 2008.

Laffont, J. and D. Martimort, *The theory of incentives: the principal-agent model*, Princeton University Press, 2002.

Lazear, E., "Performance Pay and Productivity," *American Economic Review*, 2000.

Luttmer, E., "Selection, growth and the size distribution of firms," *Quarterly Journal of Economics*, 2007.

Milgrom, P. and J. Roberts, "The economics of modern manufacturing: technology, strategy and organization," *American Economic Review*, 1990.

Murphy, R., J. Clemens, and N. Veldhuis, "The economic costs of increased marginal tax rates in Canada," *Studies in Budget and Tax Policy, Frazier Institute*, 2013.

Musgrave, R. and P. Musgrave, *Public finance in theory and practice*, McGraw-Hill, 1989.

Penrose, E., *The theory of the growth of the firm*, Oxford University Press, 1959.

Porter, M., "What is strategy?," *Harvard Business Review*, 1996.

_ and N. Siggelkow, "Contextuality within activity systems and substainability of competitive advantage," *Academy of Management Perspectives*, 2008.

Powell, M., "Productivity and credibility in industry equilibrium," *mimeo, Kellogg School of Management, Northwestern University*, 2013.

Prendergast, C., "The tenuous trade-off between risk and incentives," *Journal of Political Economy*, 2002.

Raith, M., "Competition, risk and managerial incentives," *American Economic Review*, 2003.

Rivkin, J., "Imitation of complex strategies," *Management Science*, 2000.

Roberts, J., *The Modern Firm*, Oxford, 2004.

Steen, E. Van Den, "Disagreement and the Allocation of Control," *Journal of Law, Economics and Organization*, 2010.

Syverson, C., "What determines productivity?," *Journal of Economic Literature*, 2011.

Yang, Mu-Jeung, Lorenz Kueng, and Bryan Hong, "Business Strategy and the Management of Firms," *mimeo, University of Washington*, 2014.