Bureaucracy as a Mechanism to Generate Information

Walter Novaes

University of Washington

and

Luigi Zingales [*]

University of Chicago, NBER, & CEPR

http://gsblgz.uchicago.edu

April 6, 2000

Abstract

Firms that maintain no formal record of actions and events would hardly be considered well managed. Yet, corporate rules that require the recording of actions and the filling of reports are often labeled "bureaucratic" and inefficient. This paper argues that the thin line between efficiency and bureaucracy is crossed when firms introduce a managerial turnover policy to curb agency problems in a multi-layer hierarchy. Bureaucratic rules to record actions and events then arise to minimize the costs of managerial turnover. The model predicts that bureaucracy increases upon managerial turnover and it establishes a link between bureaucracy, incentive schemes, and the frequency of managerial turnover in a cross-section of firms.

[*]Novaes: Department of Finance, Box 353200, University of Washington, Seattle, WA 98195. phone: (206) 543-4435, fax (206) 685-9392, e-mail: novaes@u.washington.edu.

Weber (1968, p. 957) argues that \the management of the modern o±ce is based upon written documents (`the ¯les'), which are preserved in their original or draft form." In his view, a ¯rm that maintains no formal record of actions and events would hardly be considered well managed. Yet, recording actions and ¯lling reports cost time and e®ort and organizations that require their employees to do so are often labeled \bureaucratic" and ine±cient. This paper explores the balance between rules that allow for sound management practices and bureaucracy. We show that this thin line is crossed in order to curb managerial agency problems in a multi-layer hierarchy. To reduce the manager's rents, a ¯rm will ¯nd it optimal to adopt a turnover policy. Bureaucratic rules that require the recording of actions at a high level of detail are then introduced to minimize the costs of managerial turnover.

In our model, rules and procedures that force the recording of corporate events enhance value because they produce veri¯able information on employees' actions, allowing for more e®ective incentive schemes. As an example, consider a loan o±cer. He can be granted complete discretion in approving loans and be paid solely on the basis of his portfolio's returns. Nonetheless, since the ex-post performance of a portfolio of loans is only a noisy measure of a loan o±cer's actions, this compensation scheme imposes a great amount of risk on him. Alternatively, the bank can require the loan o±cer to ¯ll out reports on his clients' characteristics and to document the reasons why loans are granted, generating information on the \soundness" of the decision making process. As a result, the loan o±cer can be rewarded not only for the returns of his portfolio but also on the basis of input-based performance measures that are more tightly linked to his actions: Did he ¯ll the forms properly? Did he meet the guidelines? This broader compensation scheme imposes a lower risk on the loan o±cer, making it easier to elicit e®ort.

Filling out reports and documenting reasons for granting loans require time and e®ort, though. Hence, organizing a system of incentives involves a trade-o® between production e±ciency and the bene¯ts of improving incentives. Bureaucracy arises when this trade-o® is biased towards the information bene¯ts, implying corporate rules that require the recording of actions and events in too much detail. But why should this bias exist? We shall argue that the bias is an optimal response to managerial agency costs in a multi-layer organization.

Consider a three-layer hierarchy. The ¯rst layer (Board of Directors) chooses the incentive schemes for the second layer (manager) and the third layer (employee). The employee's contribution to the ¯rm's production consists of unobservable actions, which we henceforth call \e®ort". In turn, the manager organizes production, a task that includes corporate rules that ultimately determine the extent to which the employee will document his contribution to the ¯rm. In making this organizational choice, the manager is aware that rules mandating a record of actions improve the precision of an input based signal of the employee's performance at the cost of time and e®ort that could be used to enhance production.

Besides the standard problem of providing incentives for the employee to exert e®ort, we have an agency problem at the manager's level. As Shleifer and Vishny (1989) point out, managers may ine±ciently choose projects that rely more heavily on their skills to increase their bargaining power in a wage renegotiation.[1] Likewise, they may want to organize the ¯rm in a way that makes their continuation at the company's helm important to keep the employees properly motivated. To understand what type of distortion this agency problem may imply, suppose that the ¯rm is organized in a way that requires the employee to record his actions at a very high level of detail. The input based measure of performance will then be very precise, allowing the employee to feel con¯dent that his e®ort will be rewarded despite events, like managerial turnover, that are outside his control. As a ¯rst approximation, thus, the e®ectiveness of the system of incentives of a bureaucratized ¯rm is independent of the manager's identity.

In contrast, a less bureaucratized ¯rm has to rely more strongly on broader measures of performance (e.g. pro¯ts) to motivate the employee. This system of incentives is likely to be disrupted by managerial turnover. For instance, an employee under a pro¯t sharing compensation scheme may reduce e®ort if he realizes that pro¯ts will fall as a result of the hiring of a less e±cient new manager. The system of incentives of a less bureaucratized ¯rm, thus, is more severely disrupted by managerial turnover, making it more prone to be held up by the manager. In order to maximize her bargaining power vis-a-vis the Board of Directors, it is then in the manager's interest to bias the ¯rm's organization towards very little record keeping, that is, an excessive degree of informality.

Whether this bias translates into an excessively informal organization depends upon the

[1]See Edlin and Stiglitz (1995) and Prendergast (1995) for two other examples of how managers may reduce ¯rm's value to extract salary concessions.

manager's own incentive schemes. In fact, we shall show that linking the manager's compensation to the organizational choice assures efficiency. Such an incentive scheme allows for managerial rents, though. If the manager cannot pay for these rents up front (i.e. at the hiring), then shareholders have incentives to search for alternatives to reduce them. This is exactly what a performance based turnover policy does. The threat of being fired increases the manager's personal cost of distorting the organization in the direction of excess informality, decreasing her ability to hold up the Board. Thus, the combination of a compensation scheme with a managerial turnover policy increases shareholders' value in spite of the fact that, in the event that the turnover policy triggers the replacement of the incumbent manager, the firm's value will be reduced by the weakening of the employee's incentives.

Anticipating the weakening of incentives in a turnover event, the Board of Directors will try to mitigate its costs. As a result, the trade-off between production efficiency and generation of information shifts towards the latter to improve the precision of the input-based measure of performance. Bureaucratic rules that require the employee to record his actions in detail then arise to minimize the cost of the turnover policy. Ironically, the best way to address a manager's bias towards very little record keeping implies the distortion of the organization in the direction of too many bureaucratic rules.

One of the implications of our model is that the compensation schemes of more bureaucratized firms should rely more strongly on input based measures of performance. For instance, a more bureaucratic bank would make a loan officer's compensation scheme more sensitive to how well she followed the bank's credit procedures, while the compensation of a less bureaucratic bank would be more sensitive to the profits of the officer's portfolio. Accordingly, our paper yields a measure of bureaucracy that is potentially less difficult to estimate than the level of bureaucratic constraints, namely: the extent to which employees' pay depends on measures of input rather than output.

Our paper builds upon a growing literature that views bureaucracy as a restriction on employees' behavior meant to reduce the cost of providing them with incentives. The papers in this literature share a common view on the final benefits of bureaucracy, but they differ on how these benefits are achieved. In Holmstrom and Milgrom (1991), bureaucratic rules reduce the cost of providing incentives because they restrict other activities that the employees would like to do, and in so doing they reduce their opportunity cost of e®ort. In Tirole (1986 and 1992), rules reduce managers' discretion, decreasing the scope for collusion and, thus, reducing the cost of providing incentives.[2] In Prendergast and Topel (1996), bureaucracy is a rule that distorts the weight attributed to a manager's report on her subordinate's performance. This distortion reduces the scope for favoritism in the organization. Finally, in Milgrom (1988) and Milgrom and Roberts (1988), bureaucracy precludes employees from in°uencing their supervisors' decisions, reducing wasteful rent-seeking.

In all these papers, the role of bureaucracy is either to destroy or to commit to ignore relevant information. By contrast, in our model, bureaucratic procedures { like the management by written documents described by Weber { force employees to generate veri¯able information, which can be used to improve the e±ciency of incentive schemes. Thus, we focus on the role of bureaucracy as a mechanism to generate information. Perhaps more importantly, our model leads to di®erent empirical implications. For example, while Tirole (1986) suggests that bureaucracy should be negatively correlated with managerial turnover, our model can account for a positive correlation between the two, as found by Gouldner (1954).

The remainder of the paper is organized as follows. Section I describes the model. Section II characterizes the e±cient organizational choice and obtains a one to one mapping between a ¯rm's degree of bureaucratization and the extent to which employees' pay depends on measures of input rather than output. Section III explains why delegating the ¯rm's organization design to a self-interested manager generates an agency problem. Section IV shows how a managerial turnover policy helps solve the agency problem at the manager's level and how bureaucracy arises in equilibrium. Section V derives some comparative statics and discusses the empirical implications. Conclusions follow. Proofs of the propositions can be found in the appendix.

I Framework

We consider a ¯rm with a three-layer hierarchy: an employee whose contribution to production consists of his unobservable e®ort; a self interested manager who, by organizing production,

[2] The e®ect of rules on collusion is not unequivocal. If there is asymmetry of information between the manager and the employee, Felli (1996) shows that rules may actually increase collusion.

determines the extent to which the employee will document his contribution to the firm; and a Board of Directors that makes sure that the firm is under optimal incentive schemes.

A Cash-flow, preferences, and organization design

Our main interest in this paper is to investigate a trade-off between efficiency in production and generation of information. In order to characterize this trade-off, we have to relate the firm's cash-flow with the employee's incentives to exert effort, which are the ultimate reason for producing information in our model.

In one extreme, the manager can organize production in a way that all of the employee's effort is focused on enhancing production. If so, the firm's output is described by a stochastic production function, $x(a)$, which depends on the employee's effort, a. We assume that $x(a)$ is normally distributed with mean a and standard deviation $\sigma < \infty$:

$$x(a) \sim N(a; \sigma):$$

By organizing the firm in this way, the manager generates only one signal of the employee's effort: the output itself. This represents a problem to the Board because, as in the standard Principal-Agent models, effort is unobservable for outsiders and costly for the employee, whose preferences over effort, a, and income, w, are represented by a CARA utility function, with a coefficient of absolute risk aversion that is equal to r:

$$U(w; a) = -e^{-r(w - c(a))};$$

where $c : \Re_+ \to \Re_+$ is a convex and twice continuously differentiable cost function with $c(0) = 0$, $c'(a) > 0$ for $a > 0$, $c'(0) = 0$.

In order to elicit effort from the employee, her compensation would have to be linked to the sole signal of effort, that is, the noisy output x. Since the employee is risk averse, the incentive contract has to compensate not only for the disutility of effort but also for the risk that the contract imposes on the employee.

As Holmstrom (1979) demonstrates, the efficiency of the employee's contract can be enhanced by including measures of effort that are not perfectly correlated with output. Producing these other measures of performance is costly, though.

To illustrate the cost of producing alternative measures of the employee's performance, we revisit the loan officer example described in the introduction. In the case of a loan officer, an output based compensation contract pays according to the ex-post performance of his loan portfolio, which, of course, is only a noisy measure of his work. In addition to the return of his portfolio of loans, the officer's salary can be made contingent on the soundness of his decision making process. Let us then call y this input-based measure of the employee's performance, with

$$y \sim N(a; \lambda).$$

In order for y to be a valuable performance measure, an appropriate information system must support the loan officer's efforts to document his actions. For instance, credit memos will have to be properly stored to avoid ex-post manipulations and routine inspections have to be scheduled to appraise the loan officer's decision making process. In short, not only the loan officer will have to spend more time and effort recording his actions but the bank's job assignments will have to be adapted to support the effort of producing a useful input-based measure of performance.

We model the extent to which the firm adapts its line of production and the level of detail to which the employee will have to report his actions through the standard deviation λ of the signal y. An organizational structure that neither requires nor supports the employee's effort to document his actions will make y a useless signal of effort ($\lambda = \infty$). In contrast, an organizational choice that requires and supports the employee's documentation efforts makes y an informative signal (low λ).

Note, however, that filling in reports deviates the employee's effort from the truly productive tasks. In addition, organizing the firm to support the employee's documentation efforts is likely to move the firm away from the most efficient method of production. Generating informative performance measures is thus costly. We summarize these costs of decreasing λ (i.e. increasing the informational content of the input-based signal y) by assuming that, for given λ, total output, $x(a)$, is reduced by $\Gamma(\lambda)$, where $\Gamma : (0; \infty] \rightarrow \Re_+$ is a decreasing and convex function with $\Gamma(\infty) = \Gamma'(\infty) = 0$ and $\lim_{\lambda \rightarrow 0} \Gamma(\lambda) = \infty$. Also, we assume that $\Gamma(.)$ is twice continuously differentiable.

The firm's output is then a function of the employee's effort and the parameter λ that summarizes the organizational choice and the rules requiring the employee to document his actions:

$$x'(a;\theta) = x(a) - i(\theta):$$

For a given output level, x', the firm's revenue, π, is determined by the competitive price, p:

$$\pi(x';p) = p\,x';$$

where p is a random variable with a distribution function $F(p)$ and a density $f(p)$ that is strictly positive in the interval $[0;\bar{p}]$. Conditioned on the employee's effort, the price p, the output x', and the signal y are assumed to be mutually independent.

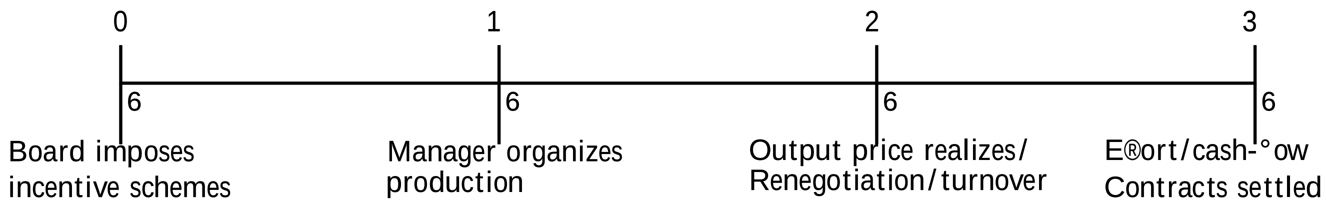
For simplicity, we assume that the manager and the shareholders are risky neutral. Hence, from the perspective of the shareholders, the optimal organizational choice maximizes the expected profits. The firm's organization is under the manager's control, though. Despite assuming that the act of organizing the firm does not directly affect the manager's utility function, we shall show that the manager has incentives to distort the organizational choice to extract salary concessions from the Board. Avoiding the distortion of the organization design, thus, requires an incentive scheme. Such incentive scheme can consist of a turnover policy that fires the incumbent manager in case the firm's expected profit is below some threshold and a salary, $s(\theta)$, that is contingent on the organizational choice, θ. We do not allow, however, the conflicts between the shareholders and the manager to be trivially solved by selling the firm to the latter. Accordingly, the manager is assumed to be credit constrained.

B Timing

The model has one production period and four dates, which are summarized in Figure 1.

The Board acts in the best interest of shareholders, choosing at date $t = 0$ the compensation schemes for the employee and the manager that maximize expected profits. If it is in the interest of shareholders, a managerial turnover policy can also be adopted at that time. Contracts can be written on three variables: the total output, x', the input-based signal y of the employee's effort, and the organizational choice θ, which adapts the firm's line of production to support rules that require the employee to document his contribution to the firm's output.

Figure 1: Timing of events



Changing a ¯rm's line of production is likely to require time. Since our framework has only one production period, we model the time lag for changing the line of production by assuming that the manager irreversibly chooses the organization design at date t = 1. For all practical purposes, thus, the extent to which the employee will record his actions is also sunk at this date. In contrast, the incentive contracts can be renegotiated at date t = 2 after the manager's fate at the company's helm has been determined by her organizational choice, the realization of the output price, and the ¯rm's turnover policy. Finally, the total output, x^0, and the input-based measure of performance, y, realize at date t = 3 after the employee's e®ort is exerted.

II The E±cient Organization Design and Bureaucracy

This section derives the e±cient organization design as the solution of a trade-o® between the costs and bene¯ts of improving the precision of the input-based measure of performance y. We solve this trade-o® backwards. For any given organizational choice ¸ and price level p, we ¯nd the optimal incentive scheme that can be imposed on the employee. This contract induces a value function, which is maximized by the e±cient organizational design. We will then say that a ¯rm is bureaucratized if the precision of the input-based measure of performance is higher than the level that is consistent with the e±cient organizational choice.

A The employee's contract

In order to simplify the analysis, we assume that wages are paid in units of the firm's output. As a result, the output price will not affect the employee's optimal contract, which, as characterized below, will be contingent on the firm's output and the input-based measure of performance,

$$\max_{\{w(.,.),a\}} pE[x - \Gamma(\lambda) - w(x - \Gamma(\lambda), y); a, \lambda, \sigma]$$

$$\text{s.t. (IR) } E[-e^{-r(w - c(a))}; a, \lambda, \sigma] \geq -1$$

$$\text{(IC) } a \in \operatorname{argmax}_{a \in \Re_+} E[-e^{-r(w - c(a))}; a, \lambda, \sigma].$$

The objective function is the firm's expected profit gross of the manager's wage. In solving for the employee's optimal contract, we can ignore the manager's salary because all of her actions will be sunk at the time that effort has to be elicited. Thus, any distortion that the manager's salary may impose on the employee's effort would be renegotiated away at the proper time. The participation constraint (IR) assures that the contract satisfies the employee's reservation value, which is assumed to be -1, while the incentive compatibility constraint (IC) provides the correct incentives for the employee to exert effort.

Given the employee's preferences and the joint distribution of (x, y), we can restrict attention to linear contracts.[3] The optimal incentive scheme, thus, is characterized by the coefficients $\{\alpha, \beta_1, \beta_2\}$ of the following compensation contract

$$w(x', y) = \alpha + \beta_1 x' + \beta_2 y.$$

Thanks to the linearity of the contract, we can replace the incentive compatibility constraint (IC) by the first order condition of the employee's effort decision. Furthermore, the participation constraint (IR) will be binding (because the objective function decreases with the manager's wage), allowing us to substitute it into the objective function. We can thus rewrite the program as

[3] See Holmstrom and Milgrom (1987).

$$\max_{\{\beta_1;\beta_2;a\}} a - c(a) - \frac{r}{2}[\beta_1^2\sigma^2 + \beta_2^2\theta^2] - \Gamma(\theta) \tag{1}$$
$$\text{s.t. } c'(a) = \beta_1 + \beta_2.$$

The first order conditions (which are also sufficient) of program (1) imply the standard under provision of effort: the optimal contract sacrifices some effort to reduce the dollar amount that should be paid to the employee as a compensation for the risk imposed by the incentive contract. The optimal weights of the linear contract satisfy:

$$\frac{\beta_2^\star(\theta;\sigma)}{\beta_1^\star(\theta;\sigma)} = \left(\frac{\sigma}{\theta}\right)^2. \tag{2}$$

The comparative statics of the optimal contract are quite intuitive. An increase in the volatility of output, σ, reduces the precision of the output-based measure of performance. Therefore, the optimal contract decreases the weight of output, β_1, in the employee's compensation. The effect on the weight of the input-based measure, β_2, is ambiguous, though. Increasing the volatility of output imposes more risk on the employee. To facilitate risk sharing, the optimal contract induces a lower level of effort which, in turn, reduces the urge for a high power incentive scheme. Thus, an increase in the volatility of output may result in a lower weight on the input-based performance measure y. Still, equation (2) implies that the weight of the input-based performance measure increases relatively to the noisier output-based measure of performance, that is, $\frac{\beta_2^\star(\theta;\sigma)}{\beta_1^\star(\theta;\sigma)}$ unambiguously increases. Analogous results follow for an increase in the volatility θ of the input-based measure of performance: the optimal effort and the weight on the input-based performance measure ($\beta_2^\star$) fall, $\frac{\beta_2^\star(\theta;\sigma)}{\beta_1^\star(\theta;\sigma)}$ decreases, while the effect on $\beta_1^\star$ is ambiguous.

B Organization design and bureaucracy

Given an organizational choice θ and the volatility σ of the firm's output, let the firm's expected profit (gross of the manager's salary) be $E[p]\{V(\theta;\sigma) - \Gamma(\theta)\}$, where

$$V(\theta;\sigma) \equiv E[x - w_\theta^\star; a_\theta^\star; \theta; \sigma] \equiv \max_{\{a;\beta_1;\beta_2\}} a - [c(a) + \frac{r}{2}(\beta_1^2\sigma^2 + \beta_2^2\theta^2)] \quad \text{s.t. } c'(a) = \beta_1 + \beta_2. \tag{3}$$

The subscript λ in $w^\star_\lambda$ and $a^\star_\lambda$ reminds us that the optimal employee's contract depends on the organizational design, whose efficient choice solves

$$\max_{\lambda \in (0,1]} E[p]\{V(\lambda, \sigma) - \Gamma(\lambda)\}$$

We assume a unique interior optimum, λ^{eff}, which must satisfy[4]

$$\frac{\partial V(\lambda^{eff}; \sigma)}{\partial \lambda} = \Gamma'(\lambda^{eff}):$$

We thus say that the firm is bureaucratized if the organizational design requires the employee to document his actions at a level of detail that allows the volatility of the input-based measure of performance to be lower than the efficient level, that is, $\lambda < \lambda^{eff}$. In contrast, the manager will run an excessively informal firm if the organizational choice implies that $\lambda > \lambda^{eff}$. In this case, the existing rules dictate the recording of actions and the filling of reports at a lower level of detail than efficiency would require.

From equation (2), the weights of the output and input-based measures of the employee's performance in an efficiently organized firm must satisfy

$$\frac{\beta^\star_2(\lambda^{eff})}{\beta^\star_1(\lambda^{eff})} = \left(\frac{\sigma}{\lambda^{eff}}\right)^2:$$

It then follows that, by choosing an excessively informal organization structure (i.e. $\lambda > \lambda^{eff}$), the employee's compensation will be biased towards the output-based measure of performance $\left(\frac{\beta^\star_2(\lambda)}{\beta^\star_1(\lambda)} = (\frac{\sigma}{\lambda})^2 < (\frac{\sigma}{\lambda^{eff}})^2\right)$. In contrast, a bureaucratized organization structure $(\lambda < \lambda^{eff})$ biases the employee's compensation towards the input-based measure of performance $\left(\frac{\beta^\star_2(\lambda)}{\beta^\star_1(\lambda)} = (\frac{\sigma}{\lambda})^2 > (\frac{\sigma}{\lambda^{eff}})^2\right)$. Conditioned on the volatility of output σ, there is a one-to-one correspondence between the firm's degree of bureaucratization and the relative slopes of the employee's incentive scheme. The more bureaucratized a firm is, the larger the importance of input-based measures of performance in the employee's compensation.

Is it in the manager's interest to deviate from the efficient organization choice? Incentives to distort would surely exist if the manager had any taste for a more or less bureaucratized

[4] This assumption holds, for example, if $c(a) = a^2$ and $\Gamma(\lambda) = \frac{1}{1000\lambda}$.

¯rm. It is hard to imagine, however, that, any such preference would survive the weakest of the incentive schemes. Nonetheless, we show in the next section that there may exist con°icts of interest in the organizational choice despite our assumption that the level of bureaucracy does not directly a®ect the manager's utility function.

III Con°icts of Interest in the Level of Bureaucracy

A The manager's quitting threat

Despite the existence of a competitive managerial labor market, once in control, a manager has many ways to acquire some power vis-à-vis the Board and use it to extract rents in a wage renegotiation.[5] In order to focus our attention on the costs and bene¯ts of producing information, we model the source of this power by assuming that the volatility of the ¯rm's output increases with the manager's departure. With a noisier measure of performance, the employee's incentive scheme becomes less e®ective, leading to a drop in pro¯tability. By threatening to impose this loss on the ¯rm, a manager can negotiate a higher wage.

Our loan o±cer example may help illustrate the interaction between managerial turnover and volatility of output. One of the roles of a loan o±cer is to detect early credit problems and take appropriate remedial actions (e.g. making sure that any collateral is properly maintained by the ¯rm and that it can be seized in case of default). Deciding whether to act on a signal of credit problems is often a subjective call, though. On the one hand, an early action may reduce the bank's loss in case the ¯rm actually becomes ¯nancially distressed. On the other hand, it may jeopardize a pro¯table business relationship. Accordingly, one would expect that loan o±cers share remedial management decisions with their supervisors, who, in turn, have to rely on the loan o±cer's information and beliefs to make the decision.

Conceivably, a conservative loan o±cer is likely to draw a less rosy picture of the client's ¯nancial situation than a more marketing oriented o±cer. Thus, supervisors should take into account their subordinates' characteristics when deciding whether to take early remedial actions. Clearly, a new supervisor who has not had enough time to know the loan o±cer

[5] Shleifer and Vishny (1989), for instance, suggest that an appropriate selection of which projects to undertake would make the manager more indispensable to the ¯rm's operations.

will be at disadvantage on this regard. Managerial turnover, thus, should imply a reduction in the expected return of the officer's portfolio and an increase in the volatility of returns. For simplicity, we ignore the mean effect and concentrate on the increased variance.[6] Output under a replacement manager is then equal to

$$x_R(a) - \Gamma(\lambda) \sim N(a - \Gamma(\lambda); \bar{\sigma});$$

with $\bar{\sigma} > \sigma$.

In the event of managerial turnover, the employee's optimal contract responds to the higher volatility of output by eliciting a lower level of effort. This reduction of effort moves the firm away from the first best, reducing its value by[7]

$$E[p]\{[V(\lambda;\sigma) - \Gamma(\lambda)] - [V(\lambda;\bar{\sigma}) - \Gamma(\lambda)]\} = E[p]\{V(\lambda;\sigma) - V(\lambda;\bar{\sigma})\} > 0.$$

The loss in value associated with managerial turnover gives some bargaining power to the incumbent manager via-à-vis the Board. As in Shleifer and Vishny (1989), the manager can threaten to quit in order to obtain a salary raise.

We assume that any salary renegotiation will be efficiently resolved, with the manager staying in the firm and capturing a fraction $\phi \in (0; 1]$ of the loss that her departure would have imposed on the firm's value. Accordingly, for any given organizational choice λ, the manager will renegotiate her contract whenever her expected salary is below

$$\phi E[p]\{V(\lambda;\sigma) - V(\lambda;\bar{\sigma})\}.$$

B Organization design and managerial rent

So far, we have just pointed out that, left unchecked, the manager's quitting threat assures her some rents. Do the rents affect the manager's choice of the organization design? We now show that, under a mild assumption, the manager can enhance the quitting threat by running

[6] This assumptions is consistent with Watts, Warner, and Wruck's (1988), who find an increase in the volatility of stock returns around a managerial turnover event but no abnormal stock returns.

[7] To show that an increase in the volatility of output reduces value, use the envelope theorem to differentiate $V(\lambda;\sigma)$, equation (3), with respect to σ and obtain $\frac{\partial V(\lambda;\sigma)}{\partial \sigma} = -r(\beta_1^\star)^2\sigma < 0$.

an excessively informal ¯rm.

For any bargaining power Á > 0, the manager's gain in a salary renegotiation increases with the loss that managerial turnover causes to the ¯rm's value. Hence, the manager has incentives to run an informal ¯rm if bureaucracy reduces the e±ciency loss associated with her departure. Using the envelope theorem to di®erentiate the ¯rm's value (see equation (3)), one can check that this condition is satis¯ed if

$$E[p]\frac{@^2 V(¸;¾)}{@¸@¾}d¾d¸ = ¡2E[p]r¾¯_1(¸;¾)\frac{@¯_1(¸;¾)}{@¸}d¾d¸ < 0:$$

The above inequality holds if and only if $\frac{@¯_1(¸;¾)}{@¸} > 0$. Intuitively, a more informal organizational choice (higher ¸) increases the loss that the manager's departure imposes on shareholders if it makes the employee's contract rely more heavily on the output-based measure of performance, which will capture the increase in the volatily associated with managerial turnover. Since $\frac{@^2 V(¸;¾)}{@¸@¾}d¾d¸ = \frac{@^2 V(¸;¾)}{@¾@¸}d¸d¾$, it follows that

$$¡2r¾¯_1(¸;¾)\frac{@¯_1(¸;¾)}{@¸} = ¡2r¸¯_2(¸;¾)\frac{@¯_2(¸;¾)}{@¾}$$

and so $\frac{@¯_1(¸;¾)}{@¸} > 0$ if and only if $\frac{@¯_2(¸;¾)}{@¾} > 0$.

We have argued in section I that the signs of $\frac{@¯_1(¸;¾)}{@¸}$ and $\frac{@¯_2(¸;¾)}{@¾}$ are ambiguous. Nonetheless, Assumption 1 below provides a su±cient condition, which is satis¯ed by the quadratic cost function, for the weight of the input-based measure of performance, $¯_2(¸;¾)$, to increase with ¾. Roughly, $\frac{@¯_2(¸;¾)}{@¾}$ is ambiguous because an increase in volatility enhances the risk borne by the employee, which makes the optimal incentive scheme to reduce the amount of e®ort to be elicited. Less powerful incentive schemes then result, implying that $¯_2(¸;¾)$ may decrease despite the higher relative precision of the input-based measure of performance. Assumption 1 imposes a lower bound on the decline of the employee's e®ort when volatility increases. This lower bound assures that the decrease in $¯_2^?$ due to the decrease in e®ort does not fully o®set the increase that results from the relative improvement of y as a signal of the employee's e®ort.[8]

[8] The condition is su±cient but not necessary. Although an exponential cost function, $c(a) = e^a$, does not satisfy the su±cient condition, it is still true that $¯_2^?$ increases with ¾.

Assumption 1 For any effort a, the employee's cost of exerting effort satisfies $\frac{ac''(a)}{c'(a)} = k \in \Re$. Moreover, for any $(\lambda;\sigma)$, the elasticity of a^* with respect to σ satisfies $-\frac{\partial a^*}{\partial \sigma}\frac{\sigma}{a^*} < \frac{2}{k}$.

We thus have,

Proposition 1 Under Assumption 1, the manager has incentives to run an excessively informal organization. If $\tilde{\lambda}$ is the manager's unconstrained organizational choice then,

$$\lambda^{eff} < \tilde{\lambda} \in \operatorname{argmax}_{\lambda \in (0;1]} \phi E[p]\{V(\lambda;\sigma) - V(\lambda;\bar{\sigma})\}:$$

In order to prevent the manager from distorting the organizational design, the Board must give her an incentive contract. A salary $s(\lambda)$ that is contingent on the organizational choice appears as a natural candidate for an optimal incentive scheme. In particular, one might think that, under a competitive managerial labor market, a forced contract that pays the manager her reservation value if and only if she chooses the efficient organizational structure should obtain the first best.

Note, however, that, if we realistically assume that a contract that penalizes an employee for quitting is not enforceable, then the manager can choose the efficient organization structure and still threaten to quit. Therefore, a contract that elicits the efficient choice λ^{eff} must give the manager a rent that is at least equal to

$$\phi E[p]\{V(\lambda^{eff};\sigma) - V(\lambda^{eff};\bar{\sigma})\}:$$

Can the manager assure herself more than this lower bound? If managers benefit from limited liability in their employment relationships, the most that Boards can do if a contracted organization design is not implemented is to cut the managers' salaries. By threatening to quit, our manager is already accepting this cut, though. Therefore, a compensation contract that is contingent on the organizational choice will not necessarily stop the manager from distorting the firm's organization to maximize her expected gains in a salary renegotiation. In order to elicit the efficient organizational choice, the Board must offer the manager a compensation contract, $s(\lambda)$, such that

$$E[p]s(\lambda^{eff}) \geq \phi E[p]\{V(\tilde{\lambda};\sigma) - V(\tilde{\lambda};\bar{\sigma})\};$$

where $\tilde{\lambda}$ is the manager's unconstrained organizational choice as defined in Proposition 1.

It then follows that a contract that assures the efficient organizational choice leaves the manager's rents untouched. If the latter is unable to pay up front for the value of the future rents, then it is in the shareholders' interest to search for additional mechanisms to reduce them. The next section shows how a managerial turnover policy accomplishes this task and how it affects the Board's preferred organizational structure.

IV Bureaucracy and Turnover Policy

A Firing policy and managerial rents

Consider a turnover policy where the incumbent manager is fired whenever expected profits before the manager's compensation fall below some cut-off level T, that is, $p[V(\lambda;\sigma) - i(\lambda)] \leq T$ (we shall discuss later how this policy can be enforced). Given the organizational choice and the turnover policy, we can define a cut-off for the output price, $p(\lambda;T)$, such that the board fires the manager in all the states where the price falls in the interval $[0; p(\lambda;T))$:

$$p(\lambda;T) = \min\{\frac{T}{V(\lambda;\sigma) - i(\lambda)}; \bar{p}\}. \tag{4}$$

By replacing the manager when the output price is below the cut-off $p(\lambda;T)$, the Board increases the volatility of the firm's output, reducing the precision of the employee's output-based measure of performance. The optimal employee's contract, thus, should be contingent on the fate of the incumbent management, solving the following maximization program

$$\begin{aligned} &\max_{a_{i\lambda};w_{i\lambda}(\cdot)} pE[x - i(\lambda) - w_{i\lambda}(\cdot); a_{i\lambda}; \lambda; \sigma_i] \qquad (5)\\ &\text{s.t.} \quad E[-e^{-r(w_{i\lambda} - c(a_{i\lambda}))}; a_{i\lambda}; \lambda; \sigma_i] \geq -1 \\ &a_{i\lambda} \in \operatorname{argmax}_{a\in\Re_+} E[-e^{-r(w_{i\lambda} - c(a))}; a; \lambda; \sigma_i]; \end{aligned}$$

where $i \in \{C; R\}$, with $i = R$ in case the manager is replaced and $i = C$ otherwise; $\sigma_R = \bar{\sigma}$, and $\sigma_C = \sigma$.

The above program is identical to the one solved in section II, provided that one replaces the volatility of output σ by $\bar{\sigma}$ in case of turnover. As we have already argued, this increase in volatility reduces the e®ort that the optimal contract will want to elicit, moving the ¯rm's value away from the ¯rst best. Moreover, equation (2) implies that the higher volatility of output increases the relative importance of the input-based measure of performance in the optimal contract. We thus have

Proposition 2 Following managerial turnover, i) the employee's optimal contract increases the relative importance of the input-based measure of performance

$$\frac{\beta_2^C(\theta)}{\beta_1^C(\theta)} = \left(\frac{\sigma}{\theta}\right)^2 < \left(\frac{\hat{\sigma}}{\theta}\right)^2 = \frac{\beta_2^R(\theta)}{\beta_1^R(\theta)} \tag{6}$$

ii) employee's e®ort and ¯rm's value decrease:

$$a_{R_\theta} < a_{C_\theta} \quad \text{and} \quad V(\theta;\hat{\sigma}) < V(\theta;\sigma):$$

If managerial turnover decreases value, why should the Board of Directors adopt a turnover policy? A turnover policy that ¯res the incumbent when expected pro¯ts fall below a certain threshold makes the manager's decision to leave irrelevant in the low pro¯tability states. The amount of rents that the manager can extract from the quitting threat is then reduced to

$$\phi\{[\int_0^{p(\theta;T)} p dF(p) V(\theta;\hat{\sigma}) + \int_{p(\theta;T)}^{\bar{p}} p dF(p) V(\theta;\sigma)] - E[p]V(\theta;\hat{\sigma})\} < \phi E[p]\{V(\theta;\sigma) - V(\theta;\hat{\sigma})\}:$$

Hence, the optimal turnover policy trades o® a weakening of incentives when the manager is ¯red with the gains from reducing the manager's bargaining power. Of course, this trade-o® depends on the organizational choice, which the Board indirectly determines through the manager's compensation scheme. Our next task, thus, is obtain the turnover policy and the manager's compensation scheme that jointly maximize shareholders' value.

B Optimal turnover policy and bureaucracy

The Board's problem at date 0 is to choose an organization design $\theta^\star$, a compensation scheme that pays the manager $s(\theta^\star)$ if $\theta^\star$ is implemented and zero otherwise, and the maximum level of expect profits $T^\star$ that triggers managerial turnover that solve[9]

$$\max_{\{T;s(\cdot);\lambda\}} \int_0^{p(\lambda;T)} p\{V(\lambda;\bar\sigma) - \Gamma(\lambda) - s(\cdot)\}dF(p) + \int_{p(\lambda;T)}^{\bar p} p\{V(\lambda;\sigma) - \Gamma(\lambda) - s(\cdot)\}dF(p)$$

$$\text{subject to} \quad E[p]s(\cdot) \geq W(T) \tag{7}$$

$$W(T) = \max_{\lambda\in(0;1]} \phi\{[\int_0^{p(\lambda;T)} pdF(p)V(\lambda;\bar\sigma) + \int_{p(\lambda;T)}^{\bar p} pdF(p)V(\lambda;\sigma)] - E[p]V(\lambda;\bar\sigma)\} \tag{8}$$

$$p(\lambda;T) \equiv \min\{\frac{T}{V(\lambda;\sigma) - \Gamma(\lambda)};\bar p\}: \tag{9}$$

The objective function is the firm's expected profit at date $t = 0$. For any output price $p < p(\lambda;T)$, the manager is replaced and the standard deviation of the output increases to $\bar\sigma > \sigma$. As characterized by Program (5), the employee's contract will be $(w_{R_\lambda}(:); a_{R_\lambda})$, implying that the firm's expected profit is $p\{V(\lambda;\bar\sigma) - \Gamma(\lambda) - s(\cdot)\}$, where it is taken into account that the initial manager will be paid according to the terms of the compensation contract regardless of her departure. For $p \geq p(\lambda;T)$, the incumbent manager stays in power and the standard deviation of the output remains at σ. The employee's contract will be $(w_{C_\lambda}(:); a_{C_\lambda})$ and the expected profit is $p\{V(\lambda;\sigma) - \Gamma(\lambda) - s(\cdot)\}$. Constraint (7) is the manager's participation constraint, whose reservation value is characterized in equation (8) by her outside option of renegotiating the initial contract after distorting the organizational design. Finally, the last constraint defines the cut-off price $p(\lambda;T)$ from the choices of the organizational design, λ, and the maximum expected profit, T, that triggers managerial turnover.

Since the manager's compensation decreases the expected profits, the participation constraint (7) will be binding at the optimum and we can replace it in the objective function. We can then rewrite the program as

$$\max_{\{T;\lambda;s(\cdot)\}} \int_0^{p(\lambda;T)} p\{V(\lambda;\bar\sigma) - \Gamma(\lambda)\}dF(p) + \int_{p(\lambda;T)}^{\bar p} p\{V(\lambda;\sigma) - \Gamma(\lambda)\}dF(p) - W(T) \tag{10}$$

$$\text{subject to} \quad W(T) = \max_{\lambda\in(0;1]} \phi\{[\int_0^{p(\lambda;T)} pdF(p)V(\lambda;\bar\sigma) + \int_{p(\lambda;T)}^{\bar p} pdF(p)V(\lambda;\sigma)] - E[p]V(\lambda;\bar\sigma)\}$$

[9] In the objective function below, a new manager has no salary. This is consistent with the fact that the newcomer has no bargaining power because the threat of quitting is empty after the replacement of the initial manager has disrupted the firm's system of incentives. The new manager can then be put at her reservation value which, for simplicity, we assume to be zero.

$$p(\lambda;T) \equiv \min\{\frac{T}{V(\lambda;\sigma) - i(\lambda)}; 1\}.$$

It turns out that the solution of Program 10 implies bureaucracy, that is $\lambda^\star < \lambda^{eff}$, if managerial turnover happens with positive probability. The intuition is simple. Since managerial turnover weakens the employee's incentives, shareholders find it profitable to spend resources to reduce this cost. The possibility of managerial turnover, thus, distorts the organization of production to obtain more information on the employee's performance. Bureaucratic rules that require the employee to document his actions at a high level of detail arise and the input-based measure of performance will carry a stronger weight on the employee's compensation. Proposition 3 below formalizes this intuition.

Proposition 3 If managerial turnover happens with positive probability, then the firm will be bureaucratized, that is, if $T > 0$, then $\lambda^\star < \lambda^{eff}$.

We now argue that managerial turnover will happen with positive probability if the manager's bargaining power is not too small. In other words, managerial turnover with bureaucracy will be an equilibrium outcome if the two-tier agency problem is severe enough.

Assume first that the manager has no bargaining power, say because the Board can commit not to renegotiate the manager's salary. In this case, there are no managerial rents to reduce and a standard compensation scheme suffices to elicit the efficient organization design. The shareholders' optimal choices when the manager has no bargaining power, that is $\phi = 0$, are then $T^\star = 0$ and $\lambda^\star = \lambda^{eff}$.

Suppose instead that the manager has all the bargaining power, that is, $\phi = 1$. The manager then captures all the efficiency gains of avoiding a turnover in a salary renegotiation. Aware of the outcome of a salary renegotiation, the Board will be very aggressive in the choice of the turnover policy. The proof of Proposition 4 shows that firing the manager with probability 1 is then optimal for shareholders, implying that the firm will be bureaucratized ($\lambda^\star < \lambda^{eff}$).

We have two polar cases. While the Board should not adopt a turnover policy when $\phi = 0$, firing the incumbent with probability 1 is optimal when $\phi = 1$. Standard continuity arguments thus imply that, for some level $\hat{\phi} \in (0;1)$ of the manager's bargaining power, a managerial turnover policy will be adopted if and only if $\phi \geq \hat{\phi}$. For levels of bargaining power above this cut-off, the double-tier agency problem is important enough to justify the introduction of a turnover policy. In order to minimize the costs of a turnover, the firm's organization will be distorted, giving rise to bureaucratic rules that force the employee to record her actions at a high level of detail.[10] In contrast, if the bargaining power is below this cut-off, then managerial rents do not justify imposing a turnover policy and the firm will be efficiently organized as if there were no agency problems at the manager's level. We thus have,

Proposition 4 There is a level $\hat{\phi} \in (0;1)$ of the manager's bargaining power such that the firm will be bureaucratized if and only if $\phi \geq \hat{\phi}$. For $\phi < \hat{\phi}$, the turnover policy will not be adopted and the firm will be efficiently organized, that is, $\lambda^{\star} = \lambda^{eff}$.

In case it is optimal to introduce a turnover policy at date $t = 0$, the Board faces a non-trivial implementation problem. Although it appears that managerial turnover after bad performance is an standard practice (see Kaplan (1994)), firing the management is not ex-post efficient in our model. After the manager chooses the organizational design, it is in the shareholders' interest to avoid managerial turnover to increase the efficiency of the employee's incentive scheme. The problem is that, anticipating the renegotiation, the manager would not be constrained by the turnover policy at the time she organizes the firm. As a result, the Board would not be able to use the threat of managerial turnover to curb rents. The next subsection shows, however, that the firm's capital structure provides a mechanism for the Board to commit to the turnover policy.

C Implementing the Optimal Mechanism

One can find many reasons why a performance based turnover policy could be enforced ex post. For one, it could be in the interest of the Board (although not necessarily of the shareholders) to fire the manager in case of bad performance in order to blame responsibility on the latter, saving the Board's reputation vis-à-vis the shareholders.[11] To formalize this idea, though, we

[10] Of course, our model is not the first one to point out that shareholders may distort their firm's organization to reduce employees' rents. Stole and Zwiebel (1996), for instance, show that hiring an excessive number of employees reduces the bargaining power of workers who try to capture a fraction of the firm's rents.

[11] Alternatively, shareholders might introduce some inefficiency in the Board's decisions by choosing board members with conflicting interests (see Hermalin and Weisbach, 1998).

would have to model a potential conflict of interests between board members and shareholders, a valuable endeavor, but one outside the scope of this paper.

The existence of risky debt gives an explanation of how the Board may be able to commit to the turnover policy that is more attuned with our model. As in Dewatripont and Tirole (1994) and Berkovitch and Israel (1996), we show that the firm's capital structure can be used to make the turnover policy ex-post optimal for the shareholders. In the presence of a fixed claim liability, inefficiently replacing the incumbent manager in low profitability states will be ex-post optimal for shareholders because the replacement cost will be shared with the debt holders, while the upside gains will be mostly captured by the shareholders. Anticipating this opportunistic behavior, debt holders will require a higher interest rate to compensate for the cost that they will bear in the low profitability states. Shareholders, however, will gladly pay the higher interest because the distortion that the interest rate will be pricing is more than offset by the reduced managerial rent. We thus have,[12]

Proposition 5 Let $T^\star$ be the ex-ante optimal cut-off for managerial turnover in terms of expected profits and $p(\theta;T^\star) \equiv \frac{T^\star}{V(\theta;\sigma) - i(\theta)}$ be the associated cut-off for managerial turnover in terms of the output price. Then, for any organization design θ that the manager may choose, there exists a debt level $D(\theta;T^\star)$ that makes it ex-post optimal for shareholders to fire the manager if and only if $p < p(\theta;T^\star)$.

To be sure, one could argue that, ex-post, the board and the debt holders have incentives to renegotiate the inefficient firing. The board, however, can make it more difficult for such renegotiation to succeed by having dispersed debt in their capital structure. As Bolton and Scharfstein (1996) point out, negotiating with dispersed debt holders is costly, allowing for the possibility of inefficient bargaining. As a result, the optimal mechanism can be implemented by issuing risky debt to dispersed debt holders and setting a policy of firing the incumbent whenever cash-flow is expected to fall below the optimal cut-off of expected profits.[13]

[12] Proposition 5 implies that the firm's debt level must be contingent on the manager's organizational choice (θ). Isagawa (1999) provides a mechanism that implements this mapping. Issue convertible debt with a strike price schedule that assures that debt holders will convert when the choice of θ requires a lower debt level than the initial one. Likewise, selling put options at a properly chosen strike price may assure an increase in leverage if some choice of θ so requires.

[13] Risky debt also allows shareholders to increase their bargaining power in a salary renegotiation with employees (e.g., Perotti and Spier (1993)). Introducing risky debt to the renegotiation game, however, would not

V Implications and Discussion

A An Empirical Measure of Bureaucracy

One of the contributions of our framework is to identify an empirical proxy for a company's degree of bureaucratization. While it is hard to measure a firm's level of bureaucratic constraints, it should be less difficult to measure the outcome of a bureaucratic system: the extent to which employees' pay depends on measures of input rather than output.

Conditioned on the volatility of output, the employee's optimal contract establishes a one-to-one correspondence between the organizational choice λ and the relative weights of the input and output measures of performance. If firm A is more bureaucratized than firm B, that is, $\lambda^A < \lambda^B < \lambda^{eff}$, then the optimal employee's contract of firm A will have a larger weight on the input-based measure of performance than the weight of firm B. Thus, we can measure a firm's level of bureaucracy by looking at its compensation schemes. For instance, a more bureaucratic bank would make a loan officer's compensation scheme more sensitive to how well she followed the bank's credit procedures, while the compensation of a less bureaucratic bank would be more sensitive to the profits of the officer's portfolio.

If we measure bureaucracy by the relative importance of the employee's input-based measure of performance, $\frac{\beta_2}{\beta_1}$, then our model predicts that bureaucracy increases in case the manager is replaced. Indeed, Proposition 2 implies that, for any λ,

$$\frac{\beta_2^C(\lambda)}{\beta_1^C(\lambda)} = \left(\frac{\sigma}{\lambda}\right)^2 < \left(\frac{\sigma_1}{\lambda}\right)^2 = \frac{\beta_2^R(\lambda)}{\beta_1^R(\lambda)}$$

This prediction is not shared by other theories of bureaucracy. For instance, Tirole (1986) points out that managerial turnover decreases the probability of collusion between the manager and the employees. Since the risk of collusion is the driving force in his model, we should expect a firm to become less bureaucratized after managerial turnover.

We could not find any study in economics investigating how bureaucracy changes when management is replaced. In the sociological literature, however, Gouldner (1954) documents an increased reliance on formal rules following managerial turnover. In the language of our

change the main results of our paper. In fact, as we show in the proof of Proposition 5, risky debt provides ex-post incentives to managerial turnover, implying a bias towards bureaucratization.

model, this corresponds to an increased importance of input-based performance measures in the ¯rm's compensation. Similarly, Grusky (1960) talks of an \organizational anomie" following managerial turnover. In our model, this corresponds to the weakening of the incentives leading to a decrease in the ¯rm's value.

B Comparative Statics

Our model links the degree of bureaucratization and the frequency of managerial turnover to the manager's bargaining power, Á, and the relative increase in the volatility of output upon managerial turnover, $\frac{¾}{¾}$. Note, however, that any regression that tries to test these comparative statics will have to control for the ¯rm's value (everything else being equal, managerial compensation in bigger ¯rms should have a lower sensitivity to output). If the ¯rm's value is inserted in the regression, though, the coe±cients of the manager's bargaining power and the volatility of output will re°ect only their \substitution e®ects" with the \income e®ects" being subsumed in the coe±cient of the ¯rm's value.

To isolate the substitution e®ects in the comparative statics, we add a \compensating income" that keeps the ¯rm's average value unchanged after a change in an exogenous variable. For instance, for an increase in ¾, the compensating income will assure that $\frac{@V(¸^?;¾)}{@¾} = 0$.

It then follows that the ¯ring threat becomes stronger (i.e. $T^?$ increases) if we increase Á or ¾ while keeping the ¯rm's value constant. In both cases, the cost to shareholders of a manager's out of equilibrium distortion of ¸ increases. A higher Á implies that the manager captures a larger share of the gains of not letting the manager to quit. Likewise, a higher ¾ implies that the cost of managerial turnover increases, allowing the manager to capture a bigger surplus when she renegotiates her salary. Not surprisingly, the higher cost of a distorted organization makes the optimal mechanism to use the ¯ring policy more aggressively. Therefore, $T^?$ increases whenever Á or ¾ increases.

The e®ects of a larger Á or a larger ¾ on the optimal organizational design are ambiguous. On the one hand, the larger threshold $T^?$ that triggers managerial turnover induces the optimal mechanism to increase bureaucracy in order to reduce the e±ciency loss in the turnover states. On the other hand, the larger $T^?$ makes it more important to enhance e±ciency if one wants to avoid the costly turnover. This latter e®ect pushes the mechanism towards less bureaucracy.

As Proposition 6 below shows, increasing bureaucracy to reduce the e±ciency loss in the turnover states is the dominant e®ect if the distortion that bureaucracy causes in the ¯rm's production is small to begin with. In contrast, reducing bureaucracy to increase e±ciency in the non-turnover states is the dominant e®ect if the distortion that bureaucracy causes is large. We thus have

Proposition 6 Keeping the ¯rm's average value constant, an increase in the volatility of output upon managerial turnover, ¾, or a higher bargaining power, Á, increase the cut-o® of expected pro¯ts, $T^{?}$, that triggers managerial turnover. The e®ect on the optimal organizational choice is ambiguous. If the output price is distributed uniformly in the interval [0; 1] and V (¸; ¾) ¡ ¡(¸) is quasi-concave with a unique maximizer, then bureaucracy increases (decreases) if the ine±ciency that bureaucracy causes was low (high) to begin with.

Proposition 6 tells us that if we were to run a cross-sectional regression of the frequency of managerial turnover on some proxies of the manager's bargaining power, we should obtain a positive coe±cient after controlling for ¯rm's value. Proposition 6 also tells us that a positive correlation between the frequency of managerial turnover and bureaucracy should obtain in a sample of ¯rms that are not too bureaucratized. We thus have a second test to di®erentiate our model from theories of bureaucracy based on collusion or favoritism. While these theories predict a negative cross-sectional correlation between bureaucracy and the frequency of managerial turnover, our model can also account for a positive correlation.

C Committing not to renegotiate salaries

In our model, bureaucracy arises to minimize the costs of a turnover policy that aims to reduce managerial rents. One could argue, though, that the Board could solve the rent extraction problem much more e±ciently by committing not to renegotiate the manager's incentive scheme. Free from the rent extraction problem, the Board could assure the e±cient organizational design by making the manager's salary contingent on the organizational choice. An important question, thus, is whether ¯rms can easily commit not to renegotiate the manager's salary.

As Stole and Zwiebel (1996) point out, labor contracts are nonbinding in nature and important employees have the ability to bargain directly with the Board. Moreover, a manager's ability to extract rents is not limited to her monetary compensation. By using the quitting threat, a manager may extract more perks or other indirect benefits from the corporation. Since it is difficult to imagine that a Board could commit along all these dimensions, a mechanism that commits the company to a specific wage schedule may not eliminate the manager's incentives to distort the organization. If so, a turnover policy might still be useful and bureaucracy would then arise.[14]

D The role of the turnover policy

To be sure, reducing a manager's bargaining power is not the only reason for introducing a turnover policy. An alternative explanation is that managers have no bargaining power and that managerial turnover is the outcome of a negative updating about the incumbent manager's skills. As it turns out, managerial agency costs imply bureaucracy in this case as well.

Suppose that managers have no bargaining power but there is uncertainty about their skills. In this case, the Board of Directors will have to infer the manager's quality from her actions in the firm and/or the company's performance in order to decide whether to replace the incumbent. Regardless of managerial agency problems, thus, uncertainty about the manager's skills implies a turnover policy. To minimize the disruption of incentives in the event of a turnover, the firm will improve the precision of input-based measures of performance. As we have already argued, a set of rules will then be imposed on the employees, requiring the recording of actions and events. The extent of these rules, of course, depends on the importance of improving incentives, which, in turn, depends on the expected frequency of managerial turnover. As we argue below, bureaucracy arises because managerial agency costs imply a more aggressive turnover policy.

If Boards of Directors could trust managers to act on behalf of shareholders, one would expect a much smoother hiring process, where candidates for a manager's position would volunteer not only their strengths but also information that indicate their unfitness to the job. As Shleifer and Vishny (1989) point out, however, managers are likely to fight for their

[14] An anecdotal case may help illustrate the pressure to renegotiate. Stewart (1991) reports that, after the debacle of Michael Milken, Drexel guaranteed its main traders the same level of their previous year bonuses (regardless of their performance) in order to avoid their departure.

jobs. Candidates to a manager's position, thus, will not reveal all of their characteristics in their job interviews, making it more di±cult for the Board to select the best person. As a result, managerial agency costs reduce the Board's con¯dence in the ability of the manager that they selected, increasing their willingness to trigger managerial turnover in case, say, the ¯rm experiences low earnings. In response to the more aggressive turnover policy, bureaucratic rules will become more pervasive, requiring employees to record their actions at a higher level of detail.

VI Conclusions

In the economics literature, the term \bureaucracy" has been associated with ex-ante optimal constraints on the employee's actions: the requirement to work in the company's facilities, the denial of access to superiors, etc. In this literature, bureaucracy implies that information is either destroyed or ignored. Yet, the term `bureaucracy' is often associated with a system that generates rather than destroys information. Not only is this the sense of Weber's opening quote, but also the widespread perception that corporate rules that require the recording of actions and the ¯lling of forms make a company \too bureaucratic". Consistent with this idea, this paper explores the role of bureaucracy in generating information.

In our framework, the extent to which employees are required to document their actions emerges from a trade o® between the gains of improving incentives and production e±ciency. On the one hand, the recording of actions allows for input-based measures of performance that could increase the e®ectiveness of the ¯rm's system of incentives. On the other hand, recording actions and ¯lling reports cost time and e®ort that could be used to enhance production. Bureaucracy arises when this trade-o® is biased towards the bene¯ts of improving incentives, implying rules that require the employees to document their actions at a high level of detail.

Perhaps more importantly, the analysis shows that the system of incentives of more bureaucratized ¯rms rely more strongly on input-based performance measures. As a result, our paper yields a measure of bureaucracy - the extent to which employees' pay depends on measures of input rather than output - which we believe that can be used to test the model's implications that relate the level of bureaucracy to the frequency and costs of managerial turnover.

References

Bolton, P. and D. Scharfstein 1996, "Optimal Debt Structure and the Number of Creditors" Journal of Political Economy 104: 1 - 25.

Berkovitch, E. and R. Israel 1996 "The Design of Internal Control and Capital Structure," Review of Financial Studies 9: 209 - 240.

Dewatripont, M. and J. Tirole 1994, "A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence,". Quarterly Journal of Economics 109: 1027-1054.

Edlin, A. and J. Stiglitz 1995, "Discouraging Rivals: Managerial Rent Seeking and Economic Inefficiencies," American Economic Review 85: 1301-1312.

Felli, L. 1996, "Preventing Collusion Through Discretion," mimeo London School of Economics.

Gouldner, A. 1954, "Patterns of Industrial Bureaucracy", Glencoe, Illinois: The Free Press

Grusky, O. 1960, "Administrative Succession in Formal Organizations", Social Forces, XXXIX: 105-115.

Hermalin, B. and M. Weisbach 1998, "Endogenously Chosen Boards of Directors and their Monitoring of the CEO", American Economic Review forthcoming.

Holmstrom, B. 1979, "Moral Hazard and Observability", Bell Journal of Economics 10: 74-91.

Holmstrom, B. and P. Milgrom 1987, "Aggregation and Linearity in the Provision of Intertemporal Incentives", Econometrica 55: 303 - 328

Holmstrom, B. and P. Milgrom 1991, "Multitask principal-agent analysis: Incentive contracts, asset ownership and job design," Journal of Law, Economics, and Organizations 7: 24-52.

Isagawa, N. 1999, "Callable Convertible Debt under Managerial Entrenchment," mimeo Kobe University.

Jensen, M. and K. Murphy 1990, "Performance Pay and Top-Management Incentives", Journal of Political Economy, 98: 225-264.

Kaplan, S. 1994, "Top Executive Rewards and Firm Performance: A Comparison of Japan and the United States", Journal of Political Economy, 102: 510-546.

Milgrom, P., 1988, "Employment Contracts, Influence Activities, and Efficient Organization Design," Journal of Political Economy, 42: 42-61.

Milgrom, P. and J. Roberts, 1990, American Economic Review, Papers and Proceedings, 80: 154 - 159.

Prendergast, C. 1995, \A Theory of Responsibility in Organizations," Journal of Labor Economics 13: 387-400.

Prendergast, C. and R. Topel 1996, \Favoritism in Organizations", Journal of Political Economy, 104: 958 - 978.

Perotti, E. and K. Spier 1993, \Capital Structure as a Bargaining Tool: The Role of Leverage in Contract Renegotiation", American Economic Review 83: 1131-1141.

Shleifer, A. and Vishny, R. 1989, \Management Entrenchment: The Case of Manager-Speci¯c Assets," Journal of Financial Economics, vol.25: 123-140.

Stole, L. and J. Zwiebel 1996, \Organizational Design and Technology Choice under Intra-Firm Bargaining," American Economic Review XXCVI: 195-222.

Stewart, J.B. 1991 Den of Thieves, Simon & Schuster, New York.

Tirole, J. 1986, \Hierarchies and Bureaucracies: on the Role of Collusion in Organizations", Journal of Law, Economics, and Organization, 2: 181-214.

Tirole, J. 1992, \Collusion and the Theory of Organizations", in Advances in Economic Theory: Sixth World Congress, J.J. La®ont, ed.

Warner, J, R. Watts, and K. Wruck 1988,\Stock Prices and Top Management Changes," Journal of Financial Economics 20: 461-492.

Weber, M., 1968, Economy and Society University of California Press, Berkeley, CA.

Appendix

Proof of Proposition 1: It suffices to prove that, under Assumption 1, $\frac{\partial \beta_2(\lambda;\sigma)}{\partial \sigma} > 0$. The first order conditions for Program 3 with respect to effort, a, and the coefficients β_1 and β_2 are, respectively, $1 - c'(a)[1 + r\frac{(\sigma\lambda)^2}{\sigma^2+\lambda^2}c''(a)] = 0$, $\beta_1 = \frac{c'(a)}{1+(\frac{\sigma}{\lambda})^2}$, and $\beta_2 = \frac{c'(a)}{1+(\frac{\lambda}{\sigma})^2}$. Using the Implicit Function Theorem in the first order conditions yields

$$\frac{\partial \beta_2(\lambda;\sigma)}{\partial \sigma} = \frac{[1+(\frac{\lambda}{\sigma})^2]c''(a)\frac{da}{d\sigma} + c'(a)2(\frac{\lambda}{\sigma})^2\frac{1}{\sigma}}{[1+(\frac{\lambda}{\sigma})^2]^2} > 0 \quad ,$$

$$\frac{-c''(a)}{c'(a)}\frac{da}{d\sigma}\sigma < 2\frac{(\frac{\lambda}{\sigma})^2}{1+(\frac{\lambda}{\sigma})^2}$$

The sufficient condition is obtained by noting that $\frac{(\frac{\lambda}{\sigma})^2}{1+(\frac{\lambda}{\sigma})^2} < 1$ and replacing $\frac{c''(a)}{c'(a)}$ by $\frac{k}{a}$.

$\square$

Proof of Proposition 3: Shareholders' expected utility for any given $(\lambda;T)$ can be written as

$$\Pi(\lambda;T) = E[p]\{V(\lambda;\sigma) - \Gamma(\lambda)\} - \int_0^{p(\lambda;T)} p dF(p)[V(\lambda;\sigma) - V(\lambda;\bar{\sigma})] - W(T) \tag{11}$$

Assuming that $T > 0$ in the optimum, we show that $\Pi(\lambda^{eff};T) > \Pi(\lambda;T)$ for any $\lambda > \lambda^{eff}$. Hence, the turnover policy does not create incentives to increase λ beyond the efficient level λ^{eff}. To see this, suppose by contradiction that $\Pi(\lambda;T) \geq \Pi(\lambda^{eff};T)$ for some $\lambda > \lambda^{eff}$. Thus

$$\Pi(\lambda;T) - \Pi(\lambda^{eff};T) = E[p]\{[V(\lambda;\sigma) - \Gamma(\lambda)] - [V(\lambda^{eff};\sigma) - \Gamma(\lambda^{eff})]\} -$$

$$\{\int_0^{p(\lambda;T)} p dF(p)[V(\lambda;\sigma) - V(\lambda;\bar{\sigma})] - \int_0^{p(\lambda^{eff};T)} p dF(p)[V(\lambda^{eff};\sigma) - V(\lambda^{eff};\bar{\sigma})]\} \geq 0$$

Under Assumption 1, $V(\lambda^{eff};\sigma) - V(\lambda^{eff};\bar{\sigma}) < V(\lambda;\sigma) - V(\lambda;\bar{\sigma})$ for any $\lambda > \lambda^{eff}$.[15] Moreover, $p(\lambda;T) = \min\{\frac{T}{V(\lambda;\sigma)-\Gamma(\lambda)};\bar{p}\}$ implies that $p(\lambda;T) \leq p(\lambda';T)$ whenever $V(\lambda;\sigma) - \Gamma(\lambda) \geq V(\lambda';\sigma) - \Gamma(\lambda')$. Since λ^{eff} is a maximizer of $V(\lambda;\sigma) - \Gamma(\lambda)$, we have that $p(\lambda^{eff};T) \leq p(\lambda;T)$. It then follows that the second term in brackets is strictly positive. To obtain $\Pi(\lambda;T) - \Pi(\lambda^{eff};T) \geq 0$ it is then necessary that $V(\lambda;\sigma) - \Gamma(\lambda) > V(\lambda^{eff};\sigma) - \Gamma(\lambda^{eff})$, which cannot happen because λ^{eff} is a maximizer of $V(\lambda;\sigma) - \Gamma(\lambda)$.

Having shown that $\lambda^* \leq \lambda^{eff}$ when $T > 0$, we complete the proof by showing that there exists $\epsilon > 0$ such that λ^{eff} is dominated by any $\lambda \in (\lambda^{eff} - \epsilon; \lambda^{eff})$. Assume first that $p(\lambda^{eff};T) < \bar{p}$. Thus, $\Pi(\lambda;T)$ is differentiable with respect to λ at $\lambda = \lambda^{eff}$. So

$$\frac{\partial \Pi(\lambda;T)}{\partial \lambda} = E[p]\{\frac{\partial V(\lambda;\sigma)}{\partial \lambda} - \Gamma'(\lambda)\} - \int_0^{p(\lambda;T)} p dF(p)[\frac{\partial V(\lambda;\sigma)}{\partial \lambda} - \frac{\partial V(\lambda;\bar{\sigma})}{\partial \lambda}] -$$

$$p(\lambda;T)f(p(\lambda;T))\frac{\partial p(\lambda;T)}{\partial \lambda}[V(\lambda;\sigma) - V(\lambda;\bar{\sigma})]$$

Plugging $\lambda = \lambda^{eff}$ vanishes the first and the last term in the derivative because $\lambda = \lambda^{eff}$ maximizes $V(\lambda;\sigma) - \Gamma(\lambda)$, which implies that $\frac{\partial V(\lambda^{eff};T)}{\partial \lambda} - \Gamma'(\lambda^{eff}) = 0$, and so $\frac{\partial p(\lambda^{eff};T)}{\partial \lambda} = -\frac{T[\frac{\partial V(\lambda^{eff};\sigma)}{\partial \lambda} - \Gamma'(\lambda^{eff})]}{(V(\lambda;\sigma) - \Gamma(\lambda))^2} =$

[15] This follows because Assumption 1 assures $\beta_2^*(\lambda;\bar{\sigma}) > \beta_2^*(\lambda;\sigma)$ which implies that the derivative of $V(\lambda;\sigma) - V(\lambda;\bar{\sigma})$ with respect to λ is strictly positive.

0 as well. The second term is negative because $\frac{\partial V(\theta;\sigma)}{\partial\theta} - \frac{\partial V(\theta;\bar{\sigma})}{\partial\theta} = r\theta[(\beta_2^*(\theta;\bar{\sigma}))^2 - (\beta_2^*(\theta;\sigma))^2]$, and $\beta_2^*(\theta;\bar{\sigma}) > \beta_2^*(\theta;\sigma)$ under Assumption 1. It then follows that $\frac{\partial\Pi(\theta^{eff};T)}{\partial\theta} < 0$, implying that $\theta < \theta^{eff}$ in the optimum when $p(\theta^{eff};T) < \bar{p}$. If $p(\theta^{eff};T) \geq \bar{p}$, then $\frac{\partial\Pi(\theta^{eff};T)}{\partial\theta}$ reduces to $E[p]\{\frac{\partial V(\theta^{eff};\bar{\sigma})}{\partial\theta} - \Gamma'(\theta^{eff})\}$, which is negative because $\frac{\partial V(\theta^{eff};\sigma)}{\partial\theta} - \Gamma'(\theta^{eff}) = 0$ and $\frac{\partial V(\theta^{eff};\sigma)}{\partial\theta} > \frac{\partial V(\theta^{eff};\bar{\sigma})}{\partial\theta}$ under Assumption 1.

2

Proof of Proposition 4: Let us rewrite $\Pi(\theta;T)$ as follows

$$\Pi(\theta;T) = E[p]\{V(\theta;\bar{\sigma}) - \Gamma(\theta)\} + \{\int_{p(\theta;T)}^{\bar{p}} pdF(p)[V(\theta;\sigma) - V(\theta;\bar{\sigma})] - W(T)\}.$$

Recall that $W(T)$ is the maximum expected loss that the manager can impose on the shareholders by voluntarily quitting, that is,

$$W(T) = \max_{\theta\in(0;1]} \phi\{[\int_0^{p(\theta;T)} pdF(p)V(\theta;\bar{\sigma}) + \int_{p(\theta;T)}^{\bar{p}} pdF(p)V(\theta;\sigma)] - E[p]V(\theta;\bar{\sigma})\}$$

Letting $\phi = 1$, and collecting terms in the above equation we obtain

$$W(T) = \max_{\theta\in(0;1]} \int_{p(\theta;T)}^{\bar{p}} pdF(p)[V(\theta;\sigma) - V(\theta;\bar{\sigma})]$$

It then follows that the second term in brackets in $\Pi(\theta;T)$ is non-positive for any $(\theta;T)$. An upper bound of $\Pi(\theta;T)$ when $\phi = 1$, thus, maximizes $V(\theta;\bar{\sigma}) - \Gamma(\theta)$ and vanishes the term in brackets. Let $\hat{\theta} < \theta^{eff}$ be the maximizer of $V(\theta;\bar{\sigma}) - \Gamma(\theta)$.[16] We now show that $\hat{T} \equiv \bar{p}[V(\theta^{eff};\sigma) - \Gamma(\theta^{eff})]$ drops the term in brackets to 0 if $\theta = \hat{\theta}$. Indeed, setting $T = \hat{T}$ makes $p(\hat{\theta};\hat{T}) = \bar{p}$ implying that $\int_{p(\hat{\theta};\hat{T})}^{\bar{p}} pdF(p)[V(\theta;\sigma) - V(\theta;\bar{\sigma})] = 0$. Thus, $(\hat{\theta};\hat{T})$ obtains the upper bound. In contrast, a policy of avoiding managerial turnover with probability 1 is sub-optimal when $\phi = 1$. To see that $T = 0$ does not obtain the upper bound, note that, since $\hat{\theta} < \theta^{eff} < \tilde{\theta} = \arg\max_\theta V(\theta;\sigma) - V(\theta;\bar{\sigma})$, $\hat{\theta}$ does not maximize $V(\theta;\sigma) - V(\theta;\bar{\sigma})$. Therefore, the second term in brackets, $\{\int_{p(\theta;T)}^{\bar{p}} pdF(p)[V(\theta;\sigma) - V(\theta;\bar{\sigma})] - W(T)\}$, does not vanish when $T = 0$. We have then established that $T > 0$ when $\phi = 1$, with $T = \hat{T}$ optimal. Since $T^* = 0$ when $\phi = 0$, the existence of $\bar{\phi} \in (0;1)$ such that $T > 0$ is optimal follows from upper-hemicontinuity of the value function.[17]

Now we show that if $\bar{T} > 0$ is optimal for some $\bar{\phi} > 0$, then $T = 0$ is dominated for any $\phi > \bar{\phi}$. Suppose by contradiction that $(\theta(\bar{T});\bar{T})$ is optimal for $\bar{\phi}$ with $\bar{T} > 0$, and that there exists some $\phi > \bar{\phi}$,

---

[16] To see that $\hat{\theta} < \theta^{eff}$, differentiate the first order condition of $\max_\theta V(\theta;\sigma) - \Gamma(\theta)$ with respect to σ to obtain $\frac{d\theta^*}{d\sigma} = -\frac{-2r\theta\beta_2(\theta;\sigma)\frac{d\beta_2(\theta;\sigma)}{d\sigma}}{V''(\theta;\sigma) - \Gamma''(\sigma)}$. Since, the denominator is non-positive from the second order necessary condition, the sign of $\frac{d\theta^*}{d\sigma}$ is equal to the sign of $-\frac{d\beta_2(\theta;\sigma)}{d\sigma}$, which, under Assumption 1, is negative.

[17] Call $T(\phi)$ an optimizer for the maximization program $\max_T \Pi(\theta;T;\phi)$ where ϕ is a parameter of the objective function. We prove by contradiction that there exists $\bar{\phi} < 1$ where the optimal threshold $T(\bar{\phi}) \in (0;\hat{T})$. Suppose that this is not the case. Thus, either $T(\phi) = 0$ for any $\phi < 1$ (remember that $T(0) = 0$) or $T(\phi) = \hat{T}$ for any $\phi > 0$. For the former case, take any sequence $\phi_n \to 1$, with $\phi_n < 1$ for any n. Thus $\lim_{n\to 1} T(\phi_n) = 0 < \hat{T} = T(1)$, contradicting upper-hemicontinuity of $T(\phi)$. Otherwise, for any sequence $\phi_n \to 0$, with $\phi_n > 0$ for any n, $\lim_{n\to 1} T(\phi_n) = \hat{T} > 0 = T(0)$, also contradicting upper-hemicontinuity of $T(\phi)$.

with $(\lambda(T);T) = (\lambda^{eff};0)$ optimal. Thus, optimality of $(\lambda^{eff};0)$ with ϕ implies:

$$\int_0^{p(\lambda(\bar{T});\bar{T})} p\,dF(p)V(\lambda^{eff};\hat{\sigma}) + \int_{p(\lambda(\bar{T});\bar{T})}^{\bar{p}} p\,dF(p)V(\lambda^{eff};\sigma) - \phi\max_{\lambda}\int_{p(\lambda;\bar{T})}^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})] -$$

$$- i(\lambda(\bar{T}))E[p] \le E[p]V(\lambda^{eff};\sigma) - \phi\max_{\lambda}\int_0^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})] - i(\lambda^{eff})E[p] \;)$$

$$\int_0^{p(\lambda(\bar{T});\bar{T})} p\,dF(p)[V(\lambda^{eff};\sigma) - V(\lambda^{eff};\hat{\sigma})] - [i(\lambda^{eff}) - i(\lambda(\bar{T}))]E[p] \ge$$

$$\phi\{(\max_{\lambda}\int_0^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})]) - (\max_{\lambda}\int_{p(\lambda;\bar{T})}^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})])\}$$

Since $\tilde{\phi} < \phi$ and the difference of the two maxima is positive, the latter inequality implies

$$\int_0^{p(\lambda(\bar{T});\bar{T})} p\,dF(p)[V(\lambda^{eff};\sigma) - V(\lambda^{eff};\hat{\sigma})] - [i(\lambda^{eff}) - i(\lambda(\bar{T}))]E[p] >$$

$$\tilde{\phi}\{(\max_{\lambda}\int_0^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})]) - (\max_{\lambda}\int_{p(\lambda;\bar{T})}^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})])\}$$

But this inequality implies

$$E[p]\{V(\lambda^{eff};\sigma) - i(\lambda^{eff})\} - \tilde{\phi}\max_{\lambda}\int_0^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})] >$$

$$\int_0^{p(\lambda(\bar{T});\bar{T})} p\,dF(p)V(\lambda^{eff};\hat{\sigma}) + \int_{p(\lambda(\bar{T});\bar{T})}^{\bar{p}} p\,dF(p)V(\lambda^{eff};\sigma) - \tilde{\phi}\max_{\lambda}\int_{p(\lambda;\bar{T})}^{\bar{p}} p\,dF(p)[V(\lambda;\sigma) - V(\lambda;\hat{\sigma})] - i(\lambda(\bar{T}))E[p];$$

contradicting optimality of $\bar{T} > 0$ when $\phi = \tilde{\phi}$.

To finish the proof note that, conditioned on $T = 0$, the program collapses to $\max_{\lambda} E[p]\{V(\lambda;\sigma) - i(\lambda)\}$ which has $\lambda = \lambda^{eff}$ as the only maximizer.

$\square$

Proof of Proposition 5:

For any λ, let $h_{\lambda R} = x - i(\lambda) - w_{\lambda R}(x - i(\lambda); y) - s(\lambda)$ be the firm's (random) profit in units of output in case the incumbent manager is replaced.[18] The random variable $h_{\lambda R}$ induces a CDF $H_R(h|\lambda)$. Likewise, the firm's random profit in units of output in case the manager is not fired is $h_{\lambda C} = x - i(\lambda) - w_{\lambda C}(x - i(\lambda); y) - s(\lambda)$ with CDF $H_C(h|\lambda)$.

Lemma 1 The random variables $h_{\lambda R}$ and $h_{\lambda C}$ are normally distributed with means $\mu_{\lambda R} > 0$ and $\mu_{\lambda C} > 0$, respectively, and standard deviations $\sigma_{\lambda R}$ and $\sigma_{\lambda C}$. Moreover, $\mu_{\lambda C} > \mu_{\lambda R}$ and $\sigma_{\lambda R} > \sigma_{\lambda C}$.

Proof. Normality of $h_{\lambda R}$ and $h_{\lambda C}$ follows from w being linear on x and y, which, conditioned on the employee's effort, are independent and normally distributed random variables. $\mu_{\lambda R} < \mu_{\lambda C}$ follows from $\mu_{\lambda R} = V(\lambda;\hat{\sigma}) - i(\lambda) - s(\lambda) < V(\lambda;\sigma) - i(\lambda) - s(\lambda) = \mu_{\lambda C}$. Substituting the employee's optimal contract into $\sigma_{\lambda C} = Var(x - i(\lambda) - w(x - i(\lambda); y) - s(\lambda))$ yields $\sigma^2(1 - \beta_1)^2 + \beta_2^2(\lambda)^2$. Differentiating

[18] To simplify the analysis, we assume that the employee and the manager's salary are senior to debt and will be paid according to their binding contracts in any state of nature. This implies that shareholders put up an escrow account to insure the salary payments against low realizations of x.

$Var(x - \Gamma(\lambda) - w(x - \Gamma(\lambda); y) - s(\lambda))$ with respect to σ^2 we obtain

$$\frac{\partial Var(x - \Gamma(\lambda) - w(x - \Gamma(\lambda); y))}{\partial \sigma^2} = (1 - \beta_1)^2 - 2(1 - \beta_1)\sigma^2 \frac{d\beta_1}{d\sigma^2} + 2\lambda^2 \beta_2 \frac{d\beta_2}{d\sigma^2}$$

Note first that since the optimal contract elicits effort below the first best level, $c'(a) < 1$. Also, $\beta_i \geq 0$ for $i \in \{1, 2\}$ in the optimal contract. Thus, $c'(a) = \beta_1 + \beta_2$ implies that the first term in the derivative is strictly positive. Now, the second term is positive because $\beta_1 < 1$ and an increase in the standard deviation of output implies that the optimal contract places a lower weight on the output based measure of performance. To prove that the derivative of the variance is positive it then suffices to show that $\frac{d\beta_2}{d\sigma^2} \geq 0$ for any λ. But this is true under Assumption 1 as shown in the proof of Proposition 1. Thus, $\bar{\sigma} > \sigma$ implies $\sigma_{\lambda R} > \sigma_{\lambda C}$.

$\square$

Since the firm's cash-flow is distributed between minus infinity and plus infinity, it makes sense to describe the leverage policy as a debt level $D \in (-\infty, \infty)$. In this context, a negative debt level should be interpreted as hedging.[19] Given an output price $p \in (0, \bar{p}]$, λ, and a debt level D, the shareholders' expected payoff of firing the incumbent manager is

$$\Delta(D; \lambda; p) = \int_{\frac{D}{p}}^{\infty} [ph - D] dH_R(h|\lambda) - \int_{\frac{D}{p}}^{\infty} [ph - D] dH_C(h|\lambda). \quad (12)$$

In order to prove the Proposition, we will make use of the following result:

Lemma 2 Let $l(z) = \frac{n(z)}{1 - N(z)}$ be the likelihood ratio of a standard normal random variable, where $n(z)$ and $N(z)$ are, respectively, the density and the CDF of a normal distribution with mean zero and standard deviation 1. Then, for any $z \in \Re$, $l(z) > z$.

Proof. The normal distribution satisfies the monotone likelihood ratio property with $l(z) > 0$ and $l'(z) > 0$ for any $z \in (-\infty, \infty)$. We now show that $l'(z) > 0 \Leftrightarrow l(z) > z$. Suppose not. Then, for some z, $l(z) \leq z \Rightarrow$

$$\frac{e^{\frac{-z^2}{2}}}{\int_z^{\infty} e^{\frac{-t^2}{2}} dt} \leq z \Rightarrow e^{\frac{-z^2}{2}} \leq \int_z^{\infty} z e^{\frac{-t^2}{2}} dt$$

Note that $z e^{\frac{-t^2}{2}} < t e^{\frac{-t^2}{2}}$ for any $t > z$. Therefore,

$$e^{\frac{-z^2}{2}} \leq \int_z^{\infty} z e^{\frac{-t^2}{2}} dt < \int_z^{\infty} t e^{\frac{-t^2}{2}} dt = e^{\frac{-z^2}{2}}.$$

Contradiction.

$\square$

Lemma 3 For any $p \in (0, \bar{p}]$ and λ, there exists a unique $D(\lambda; p)$ such that shareholders have incentives to fire the incumbent manager if and only if $D < D(\lambda; p)$, with indifference at $D = D(\lambda; p)$.

Proof. We start showing that the shareholders' expected payoff of firing the incumbent, $\Delta(D; \lambda; p)$, has a unique interior maximum, $\bar{\Delta} > 0$, at a positive debt level. Taking the derivative of $\Delta(D; \lambda; p)$

[19] A negative debt level is then equivalent to the premium the company pays to eliminate part of the liability associated with a negative realization of x.

(equation (12)) with respect to D yields[20]

$$\frac{\partial\Delta(D;.)}{\partial D} = -\int_{\frac{D}{p}}^{\infty} dH_R(h) + \int_{\frac{D}{p}}^{\infty} dH_C(h) = H_R(\frac{D}{p}) - H_C(\frac{D}{p});$$

which can be rewritten as

$$\frac{\partial\Delta(D;.)}{\partial D} = N(\frac{\frac{D}{p} - \mu_R}{\sigma_T}) - N(\frac{\frac{D}{p} - \mu_C}{\sigma_C});$$

where $N(.)$ is the CDF of a standard normal distribution. Let $\hat{D}$ solve

$$\frac{\frac{\hat{D}}{p} - \mu_R}{\sigma_R} = \frac{\frac{\hat{D}}{p} - \mu_C}{\sigma_C} \Rightarrow \hat{D} = p\{\frac{\frac{\mu_C}{\sigma_C} - \frac{\mu_R}{\sigma_R}}{\frac{1}{\sigma_C} - \frac{1}{\sigma_T}}\} > 0.$$

One can easily check that for $D < \hat{D}$, $\frac{\frac{D}{p} - \mu_R}{\sigma_R} > \frac{\frac{D}{p} - \mu_C}{\sigma_C}$, with the reverse inequality for $D > \hat{D}$. Hence, $N(\frac{\frac{D}{p} - \mu_R}{\sigma_R}) > N(\frac{\frac{D}{p} - \mu_C}{\sigma_C}) \Rightarrow \frac{\partial\Delta(D;.)}{\partial D} > 0$ for any $D < \hat{D}$, with the reverse inequality for $D > \hat{D}$, and equality for $D = \hat{D}$. This proves that $\Delta(D;\lambda;p)$ has a unique maximum at $D = \hat{D} > 0$. To investigate the sign of $\Delta(\hat{D};\lambda;p)$, we rewrite $\Delta(D;\lambda;p)$ in terms of conditional expectations

$$p\{(1 - H_R(\frac{D}{p}))E_R[h - \frac{D}{p} | h \geq \frac{D}{p}] - (1 - H_C(\frac{D}{p}))E_C[h - \frac{D}{p} | h \geq \frac{D}{p}]\}.$$

Since H_R and H_C are CDFs of normal distributions, we can use the expectation of a truncated normal to compute the difference of the conditional expectations

$$\Delta(D;\lambda;p) = p\{[1 - N(\frac{\frac{D}{p} - \mu_R}{\sigma_R})][\mu_R + \sigma_R \frac{n(\frac{\frac{D}{p} - \mu_R}{\sigma_R})}{1 - N(\frac{\frac{D}{p} - \mu_R}{\sigma_R})} - \frac{D}{p}] - [1 - N(\frac{\frac{D}{p} - \mu_C}{\sigma_C})][\mu_C + \sigma_C \frac{n(\frac{\frac{D}{p} - \mu_C}{\sigma_C})}{1 - N(\frac{\frac{D}{p} - \mu_C}{\sigma_C})} - \frac{D}{p}]\};$$

where $n(h)$ is the density of a standard normal distribution. Therefore, the fractions above are likelihood ratios evaluated at $\frac{D - \mu_R}{\sigma_R}$ and $\frac{D - \mu_C}{\sigma_C}$. Evaluating $\Delta(D;\lambda;p)$ at $\hat{D}$:

$$\Delta(\hat{D};\lambda;p) = p\{[1 - N(\frac{\frac{\hat{D}}{p} - \mu_R}{\sigma_R})][\mu_R + \sigma_R l(\frac{\frac{\hat{D}}{p} - \mu_R}{\sigma_R}) - \frac{\hat{D}}{p}] - [1 - N(\frac{\frac{\hat{D}}{p} - \mu_C}{\sigma_C})][\mu_C + \sigma_C l(\frac{\frac{\hat{D}}{p} - \mu_C}{\sigma_C}) - \frac{\hat{D}}{p}]\}.$$

By construction of $\hat{D}$, we have that $\frac{\frac{\hat{D}}{p} - \mu_R}{\sigma_R} = \frac{\frac{\hat{D}}{p} - \mu_C}{\sigma_C} \equiv D^R$. Therefore,

$$\Delta(\hat{D};\lambda;p) = p[1 - N(D^R)]\{[\mu_R + \sigma_R l(D^R)] - [\mu_C + \sigma_C l(D^R)]\} > 0 \Leftrightarrow$$

$$\mu_R + \sigma_R l(D^R) > \mu_C + \sigma_C l(D^T) \Leftrightarrow \mu_C - \mu_R < (\sigma_R - \sigma_C) l(D^R) \Leftrightarrow$$

$$l(D^R) > \frac{\mu_C - \mu_R}{\sigma_R - \sigma_C}.$$

Since $l(D^R) > D^R$, the above inequality is satisfied if $D^R \geq \frac{\mu_C - \mu_R}{\sigma_R - \sigma_C}$. In fact, this latter inequality holds

[20] From now on we will drop λ from the notation except when necessary to avoid ambiguity.

with equality:

$$D^R = \frac{\frac{\hat{D}}{p} - \mu_C}{\sigma_C} = \frac{\frac{p\{\frac{\frac{\mu_C}{\sigma_C} - \frac{\mu_R}{\sigma_R}}{\frac{1}{\sigma_C} - \frac{1}{\sigma_R}}\}}{p} - \mu_C}{\sigma_C} = \frac{\frac{\sigma_R \mu_C - \sigma_C \mu_R}{\sigma_R - \sigma_C} - \mu_C}{\sigma_C} = \frac{\mu_C - \mu_R}{\sigma_R - \sigma_C}.$$

Now we show that there is $\underline{D} < \hat{D}$, such that $\Delta(\underline{D}; \lambda; p) < 0$. Since $\Delta(D; \lambda; p)$ is continuous on D and $\Delta(\hat{D}; \lambda; p) > 0$, existence of $D(\lambda; p) \in (\underline{D}; \hat{D})$ with $\Delta(D(\lambda; p); \lambda; p) = 0$ follows from the Intermediate Value Theorem. Moreover, because $\frac{\partial \Delta(D;.;.)}{\partial D} > 0$ for any $D < \hat{D}$, we have that $\Delta(D; \lambda; p) < 0$ for any $D < D(\lambda; p)$, with the reverse inequality for any $D \in (D(\lambda; p); \hat{D}]$.

To prove existence of $\underline{D}$ with $\Delta(\underline{D}; \lambda; p) < 0$, note that

$$\Delta(D; \lambda; p) = \int_{\frac{D}{p}}^{\infty} [ph - D]dH_R(h) - \int_{\frac{D}{p}}^{\infty} [ph - D]dH_C(h) \;,$$

$$\Delta(D; \lambda; p) = (\int_{\frac{D}{p}}^{\infty} ph\, dH_R(h) - \int_{\frac{D}{p}}^{\infty} ph\, dH_C(h)) + D(H_R(\frac{D}{p}) - H_C(\frac{D}{p})). \qquad (13)$$

For any $D < 0$ and $p > 0$, $H_R(\frac{D}{p}) - H_C(\frac{D}{p}) > 0 \Rightarrow D(H_R(\frac{D}{p}) - H_C(\frac{D}{p})) < 0$. Now, $\lim_{D \to -\infty} \{\int_{\frac{D}{p}}^{\infty} ph\, dH_R(h) - \int_{\frac{D}{p}}^{\infty} ph\, dH_C(h)\} = p\{E[h_{\lambda R}] - E[h_{\lambda C}]\} < 0$. Since the first term in brackets in $\Delta(D; \lambda; p)$ converges to a negative number when $D \to -\infty$ and the second term is negative for any $D < 0$, existence of $\underline{D}$ with $\Delta(\underline{D}; \lambda; p) < 0$ follows for $\underline{D}$ sufficiently small. Therefore, shareholders are strictly better of retaining the incumbent if $D \leq D(\lambda; p)$, where $D(\lambda; p)$ is the unique $D \in (-\infty; \hat{D}]$ with which $\Delta(D; \lambda; p) = 0$.

To finish the proof of the lemma, we show that $\Delta(D; \lambda; p) > 0$ for any $D > \hat{D}$. Hence, $D(\lambda; p)$ is the only debt level that makes shareholders indifferent on whether to fire the incumbent, with an strict preference for retaining her when $D < D(p; \lambda)$ and an strict preference for firing her when $D > D(\lambda; p)$. To prove that $\Delta(D; \lambda; p) > 0$ for any $D > \hat{D}$, we show that $\lim_{D \to \infty} \Delta(D; \lambda; p) = 0$. Since, $\frac{\partial \Delta(D; \lambda; p)}{\partial D} < 0$ for any $D > \hat{D}$, the limit being zero is only possible if $\Delta(D; \lambda; p) > 0$ for any $D > \hat{D}$.

Consider $\Delta(D; \lambda; p)$ as written in equation (13). The first term in parentheses clearly vanishes when $D \to \infty$. To obtain the limit of the second term we re-write it as

$$\lim_{D \to \infty} \frac{N(\frac{\frac{D}{p} - \mu_R}{\sigma_R}) - N(\frac{\frac{D}{p} - \mu_C}{\sigma_C})}{\frac{1}{D}}$$

Using L'Hospital rule to evaluate the above limit

$$\lim_{D \to \infty} -D^2 \frac{1}{p} \frac{1}{\sqrt{2\pi}} \{\frac{1}{\sigma_R} e^{\frac{-(\frac{D}{p} - \mu_R)^2}{2\sigma_T^2}} - \frac{1}{\sigma_C} e^{\frac{-(\frac{D}{p} - \mu_C)^2}{2\sigma_C^2}}\}$$

The above limit is zero because

$$\lim_{D \to \infty} -D^2 e^{\frac{-(\frac{D}{p} - \mu_i)^2}{2\sigma_i^2}} = \lim_{D \to \infty} -\frac{D^2}{e^{\frac{(\frac{D}{p} - \mu_i)^2}{2\sigma_i^2}}} = \lim_{D \to \infty} -\frac{2D}{(\frac{\frac{D}{p} - \mu_i}{\sigma_i})\frac{1}{p\sigma_i} e^{\frac{(\frac{D}{p} - \mu_i)^2}{\sigma_i^2}}} = 0$$

where we made another use of the L'Hospital rule in the second to last equality.

$\square$

Lemma 3 showed that, for any cut-off T and $p \in (0; \bar{p}]$, there is a unique debt level, $D(\lambda; p)$, that makes shareholders indifferent on whether to fire the manager. In particular, this is true for $p = p(\lambda; T)$. The next Lemma proves that, with $D(\lambda; p(\lambda;T))$, shareholders strictly prefer to fire the manager if and only if $p < p(\lambda;T)$. This result establishes the shareholder's ability to commit to an optimal strategy of firing the manager for any λ that the latter may choose in or out of equilibrium.

Lemma 4 With debt level $D(\lambda; p(\lambda;T))$, shareholders are strictly better off firing the manager if and only if $p < p(\lambda;T)$. Moreover, shareholders are strictly better off retaining the manager if and only if $p > p(\lambda;T)$.

Proof. Let $D(\lambda; p(\lambda;T)) \equiv D_T(\lambda)$. To analyze the shareholders' incentives under $D_T(\lambda)$ for $p \neq p(\lambda;T)$, we differentiate $\Delta(D_T(\lambda);\lambda;p)$ with respect to p:

$$\frac{\partial \Delta(D_T(\lambda);\lambda;p)}{\partial p} = \int_{\frac{D_T(\lambda)}{p}}^{\infty} h\,dH_R(h) - \int_{\frac{D_T(\lambda)}{p}}^{\infty} h\,dH_C(h).$$

By construction of $D_T(\lambda)$

$$\Delta(D_T(\lambda);\lambda;p(\lambda;T)) = p(\lambda;T)\left\{\int_{\frac{D_T(\lambda)}{p(\lambda;T)}}^{\infty}\left[h - \frac{D_T(\lambda)}{p(\lambda;T)}\right]dH_R(h) - \int_{\frac{D_T(\lambda)}{p(\lambda;T)}}^{\infty}\left[h - \frac{D_T(\lambda)}{p(\lambda;T)}\right]dH_C(h)\right\} = 0 \iff$$

$$p(\lambda;T)\left[\int_{\frac{D_T(\lambda)}{p(\lambda;T)}}^{\infty} h\,dH_R(h) - \int_{\frac{D_T(\lambda)}{p(\lambda;T)}}^{\infty} h\,dH_C(h)\right] + \frac{D_T(\lambda)}{p(\lambda;T)}\left[H_R\left(\frac{D_T(\lambda)}{p(\lambda;T)}\right) - H_C\left(\frac{D_T(\lambda)}{p(\lambda;T)}\right)\right] = 0.$$

Since $\frac{D_T(\lambda)}{p(\lambda;T)} < \frac{\hat{D}}{p(\lambda;T)}$, then $H_R\left(\frac{D_T(\lambda)}{p(\lambda;T)}\right) > H_C\left(\frac{D_T(\lambda)}{p(\lambda;T)}\right)$, implying that the second term in brackets is positive. In order to obtain $\Delta(D_T(\lambda);\lambda;p(\lambda;T)) = 0$, the first term in brackets, which is $\frac{\partial \Delta(D_T(\lambda);\lambda;p)}{\partial p}$ evaluated at $p(\lambda;T)$, must be negative. Therefore, for some $\epsilon > 0$, shareholders have incentives to fire the incumbent for any $p \in (p(\lambda;T) - \epsilon; p(\lambda;T))$, with incentives to retain her for any $p \in (p(\lambda;T); p(\lambda;T) + \epsilon)$.

Now, suppose by contradiction that, for some $p > p(\lambda;T)$, shareholders have incentives to fire the incumbent, that is $\Delta(D_T(\lambda);\lambda;p) \geq 0$. From the previous paragraph, we know that for $\hat{p} \in (p(\lambda;T); p(\lambda;T) + \epsilon)$, $\Delta(D_T(\lambda);\lambda;\hat{p}) < 0$. By the Intermediate Value Theorem, there exists $\bar{p} > p(\lambda;T)$ (possibly p itself if $\Delta(D_T(\lambda);\lambda;p) = 0$) such that

$$\Delta(D_T(\lambda);\lambda;\bar{p}) = \bar{p}\left\{\int_{\frac{D_T(\lambda)}{\bar{p}}}^{\infty}\left[h - \frac{D_T(\lambda)}{\bar{p}}\right]dH_T(h) - \int_{\frac{D_T(\lambda)}{\bar{p}}}^{\infty}\left[h - \frac{D_T(\lambda)}{\bar{p}}\right]dH_C(h)\right\} = 0.$$

But then $\bar{D} = \left(\frac{p(\lambda;T)}{\bar{p}}\right)D_T(\lambda)$ would obtain $\Delta(\bar{D};\lambda;p(\lambda;T)) = 0$, contradicting $D_T(\lambda)$ as the unique debt level that obtains indifference at $p = p(\lambda;T)$. We conclude that $\Delta(D_T(\lambda);\lambda;p) < 0$ for any $p > p(\lambda;T)$. The same type of argument can be used to prove that $\Delta(D_T(\lambda);\lambda;p) > 0$ for any $p < p(\lambda;T)$.

□

Proof of Proposition 6: We prove the comparative statics for σ. The comparative statics for ϕ can be proved in a similar way. Assuming an interior solution, the first order conditions for λ and T (Program 10) are, respectively

$$\frac{\partial p(\lambda;T)}{\partial \lambda}p(\lambda;T)f(p(\lambda;T))[V(\lambda;\bar{\sigma}) - V(\lambda;\sigma)] + \int_0^{p(\lambda;T)} p\frac{\partial V(\lambda;\bar{\sigma})}{\partial \lambda}dF(p) + \int_{p(\lambda)}^{\bar{p}} p\frac{\partial V(\lambda;\sigma)}{\partial \lambda}dF(p) - \frac{\partial \Gamma(\lambda)}{\partial \lambda}E[p] = 0$$

$$\frac{\partial p(\lambda;T)}{\partial T}p(\lambda;T)f(p(\lambda;T))[V(\lambda;\bar{\sigma}) - V(\lambda;\sigma)] - \frac{\partial W(T;\bar{\sigma};\phi)}{\partial T} = 0$$

For convenience, we re-write the first order conditions as

$$G(\lambda;T;\bar{\sigma};\phi) = 0$$

$$L(\theta;T;\bar{\sigma};\phi)=0$$

By the implicit function theorem, we have that

$$\frac{\partial\theta}{\partial\bar{\sigma}}=\frac{-\frac{\partial G}{\partial\bar{\sigma}}\frac{\partial L}{\partial T}+\frac{\partial G}{\partial T}\frac{\partial L}{\partial\bar{\sigma}}}{\frac{\partial G}{\partial\theta}\frac{\partial L}{\partial T}-\frac{\partial G}{\partial T}\frac{\partial L}{\partial\theta}}.$$

$$\frac{\partial T}{\partial\bar{\sigma}}=\frac{-\frac{\partial G}{\partial\theta}\frac{\partial L}{\partial\bar{\sigma}}+\frac{\partial G}{\partial\bar{\sigma}}\frac{\partial L}{\partial\theta}}{\frac{\partial G}{\partial\theta}\frac{\partial L}{\partial T}-\frac{\partial G}{\partial T}\frac{\partial L}{\partial\theta}}.$$

Note that the denominators in $\frac{\partial\theta}{\partial\bar{\sigma}}$ and $\frac{\partial T}{\partial\bar{\sigma}}$ are equal and non-negative (they are determinants of the principal minor of order two of a negative semidefinite matrix). We, then have that

$$\mathrm{sign}[\frac{\partial\theta}{\partial\bar{\sigma}}]=\mathrm{sign}[-\frac{\partial G}{\partial\bar{\sigma}}\frac{\partial L}{\partial T}+\frac{\partial G}{\partial T}\frac{\partial L}{\partial\bar{\sigma}}]$$

$$\mathrm{sign}[\frac{\partial T}{\partial\bar{\sigma}}]=\mathrm{sign}[-\frac{\partial G}{\partial\theta}\frac{\partial L}{\partial\bar{\sigma}}+\frac{\partial G}{\partial\bar{\sigma}}\frac{\partial L}{\partial\theta}]$$

Let us first look at the sign of $\frac{\partial T}{\partial\bar{\sigma}}$. The second term vanishes because $\frac{\partial V(\theta;\bar{\sigma})}{\partial\bar{\sigma}}=0$ implies $\frac{\partial\{\frac{\partial V(\theta;\bar{\sigma})}{\partial\bar{\sigma}}\}}{\partial\theta}=0$, which in turn leads to $\frac{\partial G}{\partial\bar{\sigma}}=0$. Now, using $\frac{\partial V(\theta;\bar{\sigma})}{\partial\bar{\sigma}}=0$ once more, we have that

$$\frac{\partial L}{\partial\bar{\sigma}}=-\frac{\partial\{\frac{\partial W(T;\bar{\sigma};\phi)}{\partial T}\}}{\partial\bar{\sigma}}=-\frac{\partial\{\frac{\partial W(T;\bar{\sigma};\phi)}{\partial\bar{\sigma}}\}}{\partial T}$$

Where $W(T;\bar{\sigma};\phi)=\max_\theta\phi\int_{p(\theta;T)}^{1}p\,dF(p)[V(\theta;\sigma)-V(\theta;\bar{\sigma})]$ and $\tilde{\theta}$ being the solution of this maximization program. Thus

$$\frac{\partial L}{\partial\bar{\sigma}}=-\frac{\partial\{-\phi\int_{p(\tilde{\theta};T)}^{1}p\,dF(p)\frac{\partial V(\tilde{\theta};\bar{\sigma})}{\partial\bar{\sigma}}\}}{\partial T}=-\phi\frac{\partial V(\tilde{\theta};\bar{\sigma})}{\partial\bar{\sigma}}\frac{\partial p(\tilde{\theta};T)}{\partial T}p(\tilde{\theta};T)f(p(\tilde{\theta};T))>0;$$

where the positive sign follows from $\frac{\partial p(\theta;T)}{\partial T}=\frac{1}{V(\theta;\sigma)-i(\theta)}>0$ for any θ and $\frac{\partial V(\tilde{\theta};\bar{\sigma})}{\partial\bar{\sigma}}<0$ because $\frac{\partial V(\theta;\bar{\sigma})}{\partial\bar{\sigma}}=0$ after the compensating income is taken into account, $\tilde{\theta}>\theta$ and $\frac{\partial^2V(\theta;\sigma)}{\partial\sigma\partial\theta}<0$ under Assumption 1. It then follows that the sign of $\frac{\partial T}{\partial\bar{\sigma}}$ is equal to the sign of $-\frac{\partial G}{\partial\theta}$, which is non-negative because $\frac{\partial G}{\partial\theta}\leq 0$ (it is the determinant of the principal minor of order one of a negative semidefinite matrix), proving that $\frac{\partial T}{\partial\bar{\sigma}}\geq 0$. Consider now the sign of $\frac{\partial\theta}{\partial\bar{\sigma}}$ which is equal to the sign of $-\frac{\partial G}{\partial\bar{\sigma}}\frac{\partial L}{\partial T}+\frac{\partial G}{\partial T}\frac{\partial L}{\partial\bar{\sigma}}$. Since $\frac{\partial G}{\partial\bar{\sigma}}=0$, the first term vanishes. From $\frac{\partial L}{\partial\bar{\sigma}}>0$, it follows that $\mathrm{sign}[\frac{\partial\theta}{\partial\bar{\sigma}}]=\mathrm{sign}[\frac{\partial G}{\partial T}]$. To sign $\frac{\partial G}{\partial T}$ we assume that p is uniformly distributed over $[0;1]$ and that $V(\theta;\sigma)-i(\theta)$ is quasi-concave with respect to θ with a unique maximizer $\theta=\theta^{eff}$. If so,

$$\frac{\partial G}{\partial T}=[V(\theta;\bar{\sigma})-V(\theta;\sigma)]\{\frac{\partial^2p(\theta;T)}{\partial\theta\partial T}p(\theta;T)+\frac{\partial p(\theta;T)}{\partial\theta}\frac{\partial p(\theta;T)}{\partial T}\}+\frac{\partial p(\theta;T)}{\partial T}p(\theta;T)[\frac{\partial V(\theta;\bar{\sigma})}{\partial\theta}-\frac{\partial V(\theta;\sigma)}{\partial\theta}].$$

Now, the first term is positive because $V(\theta;\bar{\sigma})-V(\theta;\sigma)<0$ and $\frac{\partial p(\theta;T)}{\partial\theta}=-T\frac{\frac{\partial V(\theta;\sigma)}{\partial\theta}-i'(\theta)}{(V(\theta;\sigma)-i(\theta))^2}<0$, because $\theta>\theta^{eff}$, and $\frac{\partial V(\theta^{eff};\sigma)}{\partial\theta}-i'(\theta^{eff})=0$ with $V(\theta;\sigma)-i(\theta)$ quasi-concave with respect to θ. Likewise, $\frac{\partial^2p(\theta;T)}{\partial\theta\partial T}=-\frac{\frac{\partial V(\theta;\sigma)}{\partial\theta}-i'(\theta)}{(V(\theta;\sigma)-i(\theta))^2}<0$. This first positive term captures the incentives to increase efficiency that arise when an increase in $\bar{\sigma}$ induces the optimal mechanism to increase T. As a response to the higher threshold in expected profits, the mechanism reduces bureaucracy to increases efficiency and so decreasing the likelihood of a costly managerial turnover. Note that if the organizational inefficiency is small to begin with, that is $\theta\approx\theta^{eff}$, then $\frac{\partial p(\theta;T)}{\partial\theta}\approx\frac{\partial^2p(\theta;T)}{\partial\theta\partial T}\approx 0$, and this efficient effect is of second order. Now, since $\frac{\partial p(\lambda;T)}{\partial T} = \frac{1}{V(\lambda;\sigma) - i(\lambda)} > 0$, the sign of second term is determined by $\frac{\partial V(\lambda;\sigma')}{\partial \lambda} - \frac{\partial V(\lambda;\sigma)}{\partial \lambda} = r_{\lambda}[\beta_2^2(\lambda;\sigma) - \beta_2^2(\lambda;\sigma')]$ which is negative under Assumption 1. This second term captures the incentives to increase bureaucracy in order to reduce the higher expected costs of managerial turnover due to the higher T induced by the increase in σ'. Note that if the organizational inefficiency is high to begin with (i.e. $a \approx 0 \Rightarrow \beta_2(\lambda;\sigma) \approx \beta_2(\lambda;\sigma') \approx 0$), then the incentives to increase bureaucracy are of second order and the incentives to enhance efficiency dominate.

$\square$