Executive Compensation, Incentives, and the Role for Corporate Governance Regulation*

David L. Dicks
Kellogg School of Management

Job Market Paper

November 13, 2007

Abstract

This paper addresses the role for corporate governance regulation. I suggest that a pecuniary externality operating through executive compensation will motivate regulation. Governance mitigates agency costs, allowing firms to grant less incentive pay. Due to the competitive labor market, when a firm increases governance and lowers incentive pay to their manager, this allows other firms to lower their executive compensation as well. Because firms do not internalize this benefit, too little governance is implemented in the competitive outcome, and regulation can improve aggregate investor welfare. I also find that when regulation is implemented, large firms increase in value, while small firms decrease in value, and all firms will lower incentive pay. Finally, this paper suggests that corporate governance will be tight in recessions and lax in booms.

*I thank Michael Fishman for his patience and support. I also thank Andrew Hertzberg, Arvind Krishnamurthy, Michael Faulkender and Zhiguo He for their helpful comments. I can be reached at d-dicks@kellogg.northwestern.edu.

1 Introduction

This paper addresses the role for corporate governance regulation. I suggest that a pecuniary externality operating through executive compensation will motivate regulation. Specifically, governance mitigates agency costs, allowing firms to grant less incentive pay. Due to the competitive labor market, this change allows other firms to lower their executive compensation as well. Because firms do not internalize this benefit, too little governance is implemented in the competitive outcome, and regulation can improve investor welfare.

This paper has the following empirical implications. First, when forced to raise governance, firms will lower equity-based pay. This is because governance and incentive pay are substitutes in solving agency problems. Chhaochharia - Grinstein (2006) finds evidence of this after the passage of Sarbanes - Oxley. Second, when corporate governance regulation is tightened, the value of large firms will increase, while the value of small firms will decrease. Chhaochharia - Grinstein (2007) finds evidence of this after the passage of Sarbanes - Oxley as well.

This paper also shows that the dispersion of productivity of firms is linked to the aggregate use of governance. When the dispersion of firms increases, the amount of governance used in the economy increases. This result may have implications for levels of governance in different industries, different countries, and different points in the business cycle. For example, Eisfeldt - Rampini (2006) shows that the cross-sectional variance of productivity of firms is higher in recessions than in booms. Given this, my paper suggests that corporate governance should also be countercyclical. Many have suggested that during the 1990s, governance decreased because investors became lazy. Instead, my model suggests that this may be optimal. Though this paper uses a static model, it gives a framework to examine cycles in governance.

My model also shows that while compensation increases in the size of the firm, pay-performance sensitivity decreases. This is due to assortive matching in the labor market. In equilibrium, the more talented managers work at the larger firms, where their talent is used most productively. Because managers add value to the firm, corporate governance becomes more profitable where the better managers are. However, the competitive labor market forces them to be paid sufficiently not to leave the firm. Together, the competitive labor market and endogenous corporate governance implies the cross-sectional result that pay increases and pay-performance sensitivity decreases in the size of the firm. This is consistent with Murphy (1999). He (2007) also proves the result that pay-performance sensitivity decreases in the size of the firm.[1]

I also find the following policy implications to maximize investor welfare. First, the optimal regulation leaves the smallest firms alone, and is strictly increasing in the size of the firm. Second, the optimal regulation can be implemented with a subsidy of governance costs, and this subsidy can be funded by a corporate income tax. Further, implementing regulation that improves investor welfare increases investment ex ante and lowers marginal hurdle-rate premiums.

Many claim poor governance is the cause of large levels of executive compensation granted by companies. They argue the CEO is able to extract more from the shareholders due to their lack of control. I argue in this paper that there is a more subtle link between compensation and governance. Specifically, I show that in a superstar market for CEOs with agency problems, small firms will govern too little and pay too much, forcing larger firms to pay their CEO inefficiently large amounts of money. Thus, bad governance will appear to spread up from the smaller firms.

Why is corporate governance regulated? In the current literature, there are two competing reasons why regulation might be necessary: suboptimal contracting and information externalities. This paper suggests that pecuniary externalities should be considered when setting corporate governance regulation.

With rational parties and optimal contracting, regulation will only improve welfare if

[1] The result in He (2007) comes from the assumption that the agent is risk averse. Specifically, in his model, the agent has CARA utility.

there is an externality.[2] Much research has been done concerning information externalities, such as Admati - Pfleiderer (2000). The idea behind this literature is that information about the firm's cash flows has benefits for those outside the firm – either for portfolio choice or real investment decisions. Though this is an important concern, I instead develop a pecuniary externality that leads to a role for corporate governance regulation.

In this paper, each firm hires a CEO who adds value to the company, but has the opportunity to divert resources to personal uses. The company can deal with this moral hazard problem either by paying the CEO enough to discourage diversion, or can exercise governance to make diversion more costly to the CEO. Governance, however, is costly to the firm. Thus, firms face a trade-off between executive compensation and corporate governance.

I introduce this moral hazard problem into a superstar model similar to Rosen (1992), Tervio (2007), and Gabaix - Landier (2008). Each manager has a level of talent, and the value of a manager's talent is proportional to the size of the firm where the manager works. Further, the level of each CEO's talent is common knowledge, and talent is movable across firms, so each firm needs to pay their CEO enough to keep the manager from going to another firm. However, this leads to an accelerated growth in pay, because the better managers work at larger firms where their talents are used more efficiently.[3]

I find a role for corporate governance regulation due to a pecuniary externality. Small firms will handle the moral hazard problem by granting the CEO enough equity to report truthfully, while large firms would prefer to exercise governance. Because each firm must pay their CEO enough to stay, the small firms force the large firms to pay their CEOs more. The importance of this type of externality has likely increased, due to the widely perceived rise in importance of general purpose human capital for CEO, as shown in Frydman (2005).

[2] If firms do not use optimal contracts, then it may be possible to improve welfare by regulating firms. However, even if firms are not offering the best contracts to their executives, there is no reason to suggest that the government could do any better. Further, the threat of a takeover could lead to more efficient contracts, provided that anyone knows how to improve contracts.

[3] This model will not work in markets that tend to promote from within, like historically has been true in Japan. However, it seems that this is the kind of market that is dominant in American firms, since many firms hire their CEO from another firm, not from within.

I focus on governance that reduces the effectiveness of diverting company resources to personal use. Thus, this corresponds to monitoring types of governance, such as financial disclosure and accounting standards.[4] In this paper, I assume that directors maximize the value of their stock, so they are not biased in favor of management. This allows me to study a general equilibrium setting without managerial power, and I still find a role for governance regulation.

1.1 Related Literature

Generous executive compensation packages are granted to CEOs, leading some to suggest the presence of wrongdoing. Others suggest that the observed high pay is merely the market price for their labor.

On the wrongdoing side, there is the Managerial Power Perspective, championed in Bebchuk - Fried (2004) which suggests that the recent rise in executive compensation is due to lax governance. Specifically, since the board is often friendly with the CEO, they might be hesitant to disagree on a matter of pay. Supporting this view, Westphal - Stern (2007) finds that directors who are agreeable[5] are more likely to be placed on other boards, suggesting that directors may not have their incentives aligned with shareholders.

On the market price side of the issue, Gabaix - Landier (2008) find that the rise of CEO pay can be attributed to the increase in size of companies. They use a superstar model of the market for executives originally conceived in Rosen (1992), and calibrate the model so it fits the data. Thus, Gabaix - Landier (2008) attribute the recent rise in executive pay to the market working correctly.

This paper contributes to this literature, because I show that, even though the pay for CEOs is driven by a competitive market, rather than by managers imposing influence on

[4] Note that governance here does not refer to the GIM index from Gompers et al (2003). Because the GIM index measures the strengths of shareholder rights, focusing on anti-takeover provisions, it ignores independence of directors and committees, as well as auditing standards.

[5] The authors measure the agreeability of a director by their ingratiation scale, which includes such factor as directors avoiding disagreements with the CEO, doing a personal favor for the CEO or exercising flattery, etc. They used a survey to collect their data.

their own pay, there is still a role for corporate governance regulation due to the pecuniary externality that firms impose on each other in the way they handle agency problems.

This paper also relates to Hermalin - Weisbach (1998). In that paper, shareholders face a trade-off between paying the CEO and giving him job security. Corporate governance, though increasing the value of the company, hurts the CEO's job security. To motivate effort, the company can pay the CEO with cash or compensate the CEO with job security by lowering corporate governance. Thus, restricting firms from choosing these ex post sub-optimal corporate governance policies will hurt firm value ex ante. What drives this result, however, is that there is uncertainty over the ability of the CEO, and there is no asymmetric information. In my paper, talent is observable, but there is a moral hazard problem, resulting in the substitutability of governance and compensation.

In Section 2, I present a general equilibrium in the market for executives, with optimal executive compensation and levels of corporate governance. I examine governance regulation in Section 3. Section 4 contains extensions and applications, and I conclude in Section 5.

2 Model

2.1 Single-Firm Model

Consider the following principal-agent problem. A firm's cash flow equals STz, where S is the size of the firm, T is the talent of the manager, and z is random. Assume $z \geq 0$, $z \sim F$, $0 \in \text{support}(F)$, and $Ez = 1$. The firm can choose a level of governance $g \in [0, 1]$, at cost $\kappa g S$. Only the agent observes the cash flow, $y = STz$. The agent can report $\widehat{y} \leq y$, and enjoys private benefit of $\lambda(1 - g)(y - \widehat{y})$.[6] Governance, thus, is costly actions taken by the principal to decrease the attractiveness of diverting cash flows to personal uses. Both the principal and agent[7] are risk-neutral, the reported cash flows are contractible, and $0 < \lambda < 1$. The agent has an outside option with $U_0 > 0$. Let $C(\widehat{y})$ be the payment to agent when the

[6] The cash-diversion model used here is a linear case of the problem in Diamond(1984) and Lacker-Weinberg(1989), as used in DeMarzo-Fishman(2007). Solutions to this type of problem are isomorphic to solutions for effort problems with binomial effort and binomial output.

[7] Principal and firm are used equivalently throughout this paper, as are manager and agent.

agent reports and delivers cash flow $\widehat{y}$.

The manager chooses $\widehat{y}$ to maximize payoff, so

$$\widehat{y}(y;g) \in \arg\max_{y' \leq y} \{C(y') + \lambda(1-g)(y-y')\}$$

The principal maximizes the objective, which is:

$$E\{\widehat{y}(y;g) - C(\widehat{y}(STz))\} - \kappa g S$$

subject to the agent's incentive compatibility constraint, the agent's limited liability constraint ($C(\cdot) \geq 0$), and the agent's participation constraint, which is

$$E\{C(\widehat{y}) + \lambda(1-g)(y-\widehat{y})\} \geq U_0$$

Lemma 1 *There will be no diversion in equilibrium. Formally, $\widehat{y}(y;g) = y$ in equilibrium.*

Intuitively, $1 - \lambda(1-g) > 0$ for any level of g, so the principal is better off paying the manager his private benefit of diversion, and inducing the manager to report truthfully. Thus, the problem simplifies to the following:

$$\begin{aligned} \max_{C,g} \quad & E\{STz - C(STz)\} - \kappa g S \\ s.t. \quad & C(y) \geq C(\widehat{y}) + \lambda(1-g)(y-\widehat{y}) \qquad \forall \widehat{y} \leq y && \text{(IC)} \\ & E[C(y)] \geq U_0 && \text{(IR)} \\ & C(\cdot) \geq 0, \;\; g \in [0,1] && \text{(LL), (F)} \end{aligned}$$

Theorem 2 *For very large levels of reservation utility ($U_0 > ST$), the project is not carried out. For moderate levels of reservation utility ($\lambda ST \leq U_0 \leq ST$), the participation constraint is tight enough so that the IC constraint is lax, so*

$$C(y) = \frac{U_0}{ST} y, \;\; g = 0$$

For small levels of reservation utility, ($U_0 < \lambda ST$), the optimal contract depends on the level of talent. For $T \leq \frac{\kappa}{\lambda}$, the principal will pay agent enough to satisfy IC, setting

$$C(y) = \lambda y, \;\; g = 0$$

For $T > \frac{\kappa}{\lambda}$, principal governs closely, so participation constraint will bind, so

$$C(y) = \frac{U_0}{ST}y,\ g = 1 - \frac{U_0}{\lambda ST}$$

Note that the compensation to the agent is linear in the cash flow y. Thus, the solution to this problem is to give the agent a share of equity. For very large levels of reservation utility, it is inefficient to take the agent from other pursuits and have him run the project, because the project will produce less than the agent would have produced elsewhere. For moderate levels of reservation utility, there is no agency problem because the agent requires enough compensation to run the project that the principal can give him sufficient pay-performance sensitivity without paying more than is necessary.[8] Finally, for small levels of reservation utility, the principal gets to choose how to induce the agent to report truthfully. If the agent is less talented, the principal chooses to pay the agent enough to report truthfully. Thus, she grants the agent a share λ of the firm. However, if the agent is more talented, the principal finds it optimal to govern the agent as closely as possible and pay as little as possible. It is this case (small level of reservation utility) that will be relevant for the general equilibrium model presented in Section 2.2.

Consider the effect of governance regulation on this single firm. It is obvious that any regulation, either that requires an increase or decrease in governance, will hurt investor welfare, because the investors could have exercised the required level of governance without regulation. This logic extends to a case where we have a cross-section of firms and the participation constraint, $U(T)$, is an *exogenous* function of T. As I will show, however, this result is not robust to the setting where the principals compete for the agents, which is discussed next. The way that regulation can improve welfare is by relaxing the participation constraint at some firms.

[8] This works by the same intuition that if the principal sells the company to the agent there ceases to be an agency problem. The agent has a sufficient stake (λ) in the company that he will behave.

2.2 General Equilibrium Model

Now suppose that there is a continuum of firms, q to 1, where $q > 0$. Suppose $S'(x) > 0$ and $T'(x) > 0$, so that firm q is the smallest, and firm 1 is the largest, and manager q is least talented, and manager 1 is the most talented. In this section, I consider a market equilibrium for managerial talent. The key assumption is that talent is freely movable from firm to firm, and that it is equally useful at any firm. Also, talent is fully observable. Together, these assumptions allow firms to compete for CEOs. Random shocks, z, are independent across firms. Each firm must pick their level of governance prior to hiring their manager, but firms pick this level of governance optimally (in equilibrium).[9] Finally, I assume that $T(q) < \frac{\kappa}{\lambda} < T(1)$ so that, by Theorem 2, some firms pay to satisfy the moral hazard problems, while others exercise governance.

Define $W(T)$ to be the equilibrium expected wage for manager with talent T. I will endogenously derive this $W(T)$ later. By Lemma 1, the manager will report truthfully in equilibrium, so firms solve the following problem:

$$\begin{aligned}
\max_{C,g,T} \quad & E[y] - EC_T(y) - \kappa g S \\
s.t. \quad & C_T(y) \geq C_T(\widehat{y}) + \lambda(1-g(x))(y-\widehat{y}) \qquad \forall \widehat{y} < y && \text{(IC)} \\
& E[C_T(y)] \geq W(T) && \text{(IR)} \\
& y = STz && (1)
\end{aligned}$$

To solve the equilibrium, I need to find which managers works at which firm. Because any competitive equilibrium is Pareto optimal, the allocation of managers must be efficient, that is, it must maximize $\int_q^1 S(x)T(x)dx$. This will result in assortive matching of managers

[9] I need the assumption that firms commit to their level of governance before hiring the agent because if I do not make this assumption, the differential equation for equity share granted to the CEO explodes at x^*. When this assumption is not made for discrete distributions, the contagion of CEO pay is stronger, because off-equilibrium, firms are willing to pay a better CEO not only for their superior talent but also to compensate him for lowered cost of governance, since less governance is necessary due to the higher equity share. In the continuous version, however, this results in a singular point at x^* in the differential equation that defines β.

to firms. To see this, consider x and x', where $x > x'$.

$$\begin{aligned} S(x) &> S(x') \qquad T(x) > T(x') \\ &\Rightarrow [S(x) - S(x')]\,[T(x) - T(x')] > 0 \\ &\Rightarrow S(x)T(x) + S(x')T(x') > S(x')T(x) + S(x)T(x') \end{aligned}$$

Therefore, assortive matching will result from market equilibrium, so manager x will work for firm x.[10]

By Theorem 2, we can restrict attention to linear contracts. Thus, define $\beta(x)$ so that $C(STz) = \beta(x)S(x)T(x)z$. Given this change of variables, the problem simplifies to

$$\begin{aligned} \max_{\beta,g} \quad & S(x)T(x)(1-\beta(x)) - \kappa g(x)S(x) \\ s.t. \quad & \beta(x) \geq \lambda(1-g(x)) && \text{(IC)} \\ & \beta(x)S(x)T(x) \geq W(T(x)) && \text{(IR)} \end{aligned}$$

Define x^* such that $T(x^*) = \frac{\kappa}{\lambda}$. By Theorem 2, for $x > x^*$, $T(x) > \frac{\kappa}{\lambda}$, the firm would like to govern the manager and pay him nothing, so the IR binds. For $x < x^*$, however, $T(x) < \frac{\kappa}{\lambda}$, so the firm will set $\beta(x) = \lambda$ even when IR is nonbinding. The firm does not consider the effect of the pay granted on other firms, so this is the source of the pecuniary externality.

Suppose manager x was not getting paid enough by firm x. The manager's outside option would be to go to another firm. Suppose the manager wished to go to firm x', where $x' < x$. Define $\beta(x, x')$ to be the most that the manager could demand there, so that the firm would be willing to fire manager x'. For firm x' to be willing to replace manager x' with manager x, the profit of firm x' would need to increase by hiring manager x. Thus, it must be the case that

$$\begin{aligned} S(x')T(x)(1-\beta(x,x')) - \kappa g(x')S(x') &\geq S(x')T(x')(1-\beta(x')) - \kappa g(x')S(x') \\ \beta(x,x')S(x')T(x) &\leq S(x')[T(x) - T(x')] + \beta(x')S(x')T(x') \end{aligned}$$

[10] Specifically, I am assuming here that the set of CEOs hired by a set of firms must have equivalent measure. This is the continuous analogue of each firm hiring one manager.

Therefore, this will hold with equality (because $\beta(x, x')$ is the most that firm x' is willing to pay). Thus, endogenously,

$$W(T(x)) = \sup_{x'<x} [S(x')[T(x) - T(x')] + \beta(x')S(x')T(x')]$$

because a manager's outside option is to go to a smaller firm.[11]

Lemma 3 *Managers will prefer to move to the next largest firm, rather than jump down firms. That is, in threatening to leave their firm, managers find small moves preferable to large moves. Formally,*

$$W(T(x)) = \lim_{x' \to x^-} \{S(x')[T(x) - T(x')] + \beta(x')S(x')T(x')\}$$

So participation constraint becomes

$$\begin{aligned} \beta(x)S(x)T(x) - \beta(x-dx)S(x-dx)T(x-dx) &\geq S(x-dx)[T(x) - T(x-dx)] \\ \Leftrightarrow (\beta ST)' &\geq ST' \end{aligned}$$

This constraint shows that executive compensation must increase across firms at least as fast as the product of size and marginal talent. Thus, in general, each firm will set their pay in the following way. For $x \leq x^*$, $T(x) \leq \frac{\kappa}{\lambda}$, so the firm will always set $g = 0$, $\beta \geq \lambda$, and $(\beta ST) \geq ST'$, and by complementary slackness one of these will bind. For $x > x^*$, $T(x) > \frac{\kappa}{\lambda}$, so the participation constraint will bind, so $(\beta ST)' = ST'$, and $g = \max\{1 - \frac{\beta}{\lambda}, 0\}$. Thus, for $x > x^*$,

$$\beta(x)S(x)T(x) = \beta(x^*)S(x^*)T(x^*) + \int_{x^*}^{x} S(z)T'(z)dz$$

Note that pay is increasing across firms.

The previous lemma leads us to the role for corporate governance regulation. Firms do not internalize the benefit of corporate governance, because dealing with agency problems with governance allows not only that firm to lower its executive compensation, but also allows other firms to lower their executive compensation as well.

[11] Off-equilibrium a manager would not be able to induce a larger firm to hire him without taking less pay than from his firm in equilibrium, so a manager would never go to a larger firm.

Theorem 4 *Governance Regulation can improve welfare.*

The previous proposition holds for any twice differentiable, strictly increasing functions S and T. The intuition is that by forcing the marginal firm to exercise a little more governance, the regulation benefits all of the large firms a little, while only costing the marginal firm about the same as the benefit at each of the large firms. This leads to an increase in welfare. Also, it should be noted that if corporate governance is regulated, in general large firms increase in value and small firms lower in value relative to a no-regulation outcome. Chhaochharia - Grinstein (2007) showed that this happened with the passage of SOX.

2.3 Specification

Following Gabaix - Landier (2008), I will assume the following specification:

$$\begin{aligned} S(x) &= A\left(1+q-x\right)^{-a} \\ T(x) &= T_{Max} - \frac{B}{b}\left(1+q-x\right)^{b} \end{aligned}$$

where $b < a$. Gabaix - Landier (2008) estimate $b \approx \frac{2}{3}$ and $a \approx 1$ in their empirical section. This implies

$$\begin{aligned} S'(x) &= aA\left(1+q-x\right)^{-a-1} \\ \text{and} \quad T'(x) &= B\left(1+q-x\right)^{b-1} \end{aligned}$$

Consider the equilibrium in this context. When the endogenous participation constraint binds for all $x \in (x_1, x_2)$,

$$\beta(x_2)S(x_2)T(x_2) = \beta(x_1)S(x_1)T(x_1) + \int_{x_1}^{x_2} S(x)T'(x)dx$$

so under this specification,

$$\begin{aligned} \int_{x_1}^{x_2} S(z)T'(z)dz &= \int_{x_1}^{x_2} A\left(1+q-z\right)^{-a}\left(B\left(1+q-z\right)^{b-1}\right)dz \\ &= \frac{AB}{a-b}\left[\left(1+q-x_1\right)^{b-a} - \left(1+q-x_2\right)^{b-a}\right] \end{aligned}$$

Note that $(1+q-x)^{b-a}$ is increasing in x, because $b-a<0$. The first relevant cutoff is x^*: as demonstrated above, for all $x>x^*$, the participation constraint will bind. Therefore, for $x>x^*$,

$$\beta(x)S(x)T(x)=\beta(x^*)S(x^*)T(x^*)+\frac{AB}{a-b}\left((1+q-x)^{b-a}-(1+q-x^*)^{b-a}\right)$$

The key question is what happens for $x<x^*$, and thus also, what is $\beta(x^*)$? Specifically, at a point $x<x^*$, which constraint will bind: the incentive compatibility constraint, or the limited liability constraint. For $x=1$, there is no participation constraint by assumption, since the worst manager has nowhere else to go. However, since $x^*>q$ by assumption, the smallest firm, q, will still grant $\beta(q)=\lambda$ to manager q. For x near q, is it enough to grant manager x share $\beta(x)=\lambda$, or is more necessary? The following lemma answers this question.

Lemma 5 *In any equilibrium, there exists x^c such that IR is binding for $x<x^c$, IC is binding for $x\in[x^c,x^*]$, and IR is binding for $x>x^*$. Further, $x^c=q$ iff*

$$\lambda a T_{Max}\geq B\left(1+\lambda\left(\frac{a}{b}-1\right)\right)$$

Thus, for ease of exposition, I make the assumption stated below. Under this assumption on parameters, when companies choose to pay their managers enough to solve the moral hazard problem without governance, the endogenous participation constraint will be non-binding, because managers will be strictly better off staying at their current firms than going to smaller firms. This assumption thus guarantees monotonicity, and will hold throughout the remainder of the paper.

Assumption: $\lambda a T_{Max}\geq B\left(1+\lambda\left(\frac{a}{b}-1\right)\right)$

The equilibrium is summarized by the following proposition.

Theorem 6 *For $x\leq x^*$, $\beta(x)=\lambda$ and $g(x)=0$. For $x>x^*$,*

$$\begin{aligned}\beta(x)&=\frac{1}{S(x)T(x)}\left[\kappa S(x^*)+\frac{AB}{a-b}\left((1+q-x)^{b-a}-(1+q-x^*)^{b-a}\right)\right]\\ g(x)&=1-\frac{\beta(x)}{\lambda}\end{aligned}$$

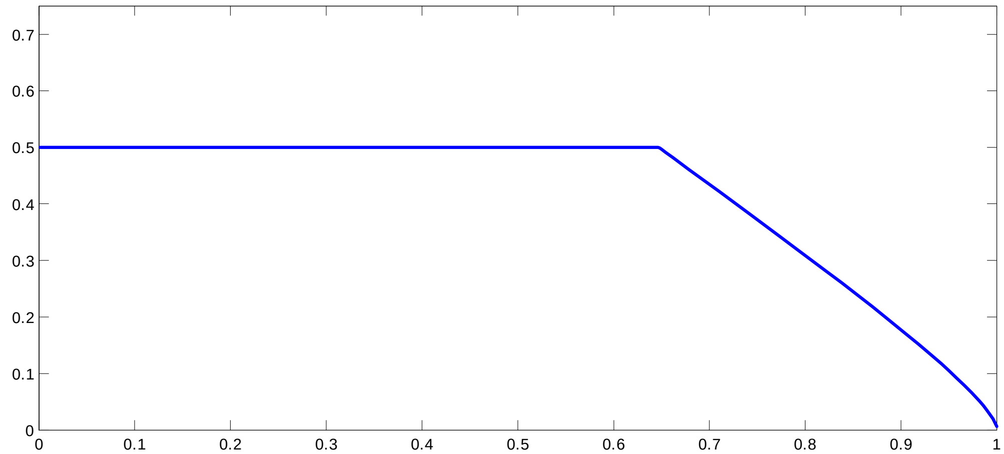

Figure 1: The share of inside equity across firms. Note that the size of firms is increasing in the index, so the largest firm is firm 1.

Pay-performance sensitivity across firms can be seen in Figure 1. Note that, in small companies ($x < x^*$) the manager's participation constraint is nonbinding, because the shareholders find it too expensive to govern the manager sufficiently. These small companies thus pay sufficiently to induce the CEO to report truthfully. However, these small companies do not take into account the effect that this has on larger companies: that the larger firms must pay their CEO more. This is a pecuniary externality, and thus there is a potential for regulation to improve investor welfare.

Observe the cross-sectional behavior of size, compensation, and governance: as firm size increases, compensation and governance increase. However, pay-performance sensitivity decreases. The properties are consistent with Murphy (1999).

Corollary 7 *Denote the objective function of each principal* $\Pi(x)$*. For* $x < x^*$*,*

$$\Pi(x) = (1 - \lambda)S(x)T(x)$$

For $x \geq x^*$*,*

$$\Pi(x) = S(x)\left(T(x) - \kappa\right) - \beta(x)S(x)T(x)\left(1 - \frac{\kappa}{\lambda T(x)}\right)$$

The corollary follows by substitution of $\beta(x)$ and $g(x)$ into $\Pi(x) = S(x)T(x) - \beta(x)S(x)T(x) - \kappa g(x)S(x)$. Note that large firms would like to set $g(x) = 1$, resulting in the $-\kappa$ term in the first coefficient. However, they must pay the manager some share of the company, which results in not only a cost to them, but also a savings of governance cost, which has a net cost of $\left(1 - \frac{\kappa}{\lambda T(x)}\right)$, which is positive on this region.

Now, consider time-series behavior of this equilibrium. Consider the share of firms in an industry that choose to exercise more than minimal governance. This is captured by $(1 - x^*)$. When x^* increases, less firms exercise voluntary governance, while when x^* decreases, more firms exercise voluntary governance, so $(1 - x^*)$ is the mass of firms exercising more than minimal governance.

Corollary 8 *The following comparative statics hold for the cutoff of firms that prefer to exercise governance rather than pay. The number of firms exercising voluntary governance is:*

$$\begin{aligned} \textit{increasing in the talent of CEOs} &: \frac{\partial (1 - x^*)}{\partial T_{Max}} > 0 \\ \textit{increasing in the severity of moral hazard problems} &: \frac{\partial (1 - x^*)}{\partial \lambda} > 0 \\ \textit{decreasing in cost of governance} &: \frac{\partial (1 - x^*)}{\partial \kappa} < 0 \end{aligned}$$

Finally, I examine the impact of a change in governance costs on the equilibrium levels of wages and governance.

Corollary 9 *When governance costs increase, the change in wages and governance levels depends on the level of the monitoring costs. For small levels of governance costs, ($\kappa\left(\frac{a}{b(1-\lambda)} + \frac{1}{\lambda}\right) < T_{Max}$), wages go down and governance levels go up when governance costs increase. For large levels of governance costs, ($\kappa\left(\frac{a}{b(1-\lambda)} + \frac{1}{\lambda}\right) > T_{Max}$), wages go up and governance levels go down when governance costs increase.*

3 Governance Regulation

In this section, I examine optimal governance regulation. Note that the compensation a CEO receives depends not only on his characteristics, but also on the amount paid to other CEOs. Because of this, the choice a company makes in their compensation and governance decision affects not only that company, but also larger companies. This creates a pecuniary externality. Thus, it is reasonable to consider how governance regulation can improve welfare.

3.1 Optimal Governance Regulation

In this section, I assume that the regulator can force each firm to carry out any level of governance, and seeks to maximize investor welfare. Also, I assume that the regulator can observe the size of firms, and knows the distribution of talent, but does not know the talent of an individual manager, though the companies do. I assume this so the equilibrium still has a competitive talent market, rather than the regulator assigning managers to firms.

Thus, the regulator solves the following to maximize investor welfare.[12]

$$\begin{aligned}
\max_{g_r(\cdot)} \quad & \int_q^1 \left\{ (1-\beta_r(x))\, T(x)S(x) - \kappa g_r(x)S(x) \right\} dx \\
s.t. \quad & \beta_r(x) \geq \lambda(1-g_r(x)) \\
& (\beta_r ST)' \geq ST' \\
& \beta_r \geq 0, g_r \in [0,1]
\end{aligned}$$

The first constraint is the reporting incentive compatibility constraint, the second is the endogenous participation constraint, and the third is limited liability and feasibility. Clearly, $g_r(x) = \max\{1 - \frac{\beta_r(x)}{\lambda}, 0\}$, because governance is costly.

Lemma 10 *Any regulation that lowers governance hurts investor welfare.*

The intuition behind this lemma is that lowering governance forces large firms to pay even more than they would under the no-regulation equilibrium. Thus, the no-regulation

[12] In this model, the regulator maximizes investor welfare because if investment is endogenized, improving the investors' payoff will increase investment and thus increase efficiency. See Section 4.2 for an illustration of this.

equilibrium dominates an equilibrium with regulation that lowers governance. Therefore, it is sufficient to consider regulation that increases governance.

Lemma 11 *In the optimal regulation equilibrium, the participation constraint will bind for all $x > x^*$. Thus, for $x > x^*$,*

$$\beta_r(x)T(x)S(x) = \beta_r(x^*)T(x^*)S(x^*) + \frac{AB}{a-b}\left[(1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right]$$

Not only is it privately optimal for firm $x > x^*$ to pay as little as possible, but also paying any more than required has a negative effect on larger firms. Thus, it is optimal to set $g_r(x)$ so that the participation constraint and incentive compatibility constraints bind for large firms ($x > x^*$). Thus, for $x > x^*$,

$$\begin{aligned} g_r(x) &= 1 - \frac{\beta_r(x)}{\lambda} \\ \Pi_r(x) &= (1-\beta_r(x))\,T(x)S(x) - \kappa g_r(x)S(x) \\ &= S(x)\left[T(x)-\kappa\right] \\ &\quad - \left\{\beta_r(x^*)T(x^*)S(x^*) + \frac{AB}{a-b}\left[(1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right]\right\}\left(1 - \frac{\kappa}{\lambda T(x)}\right) \end{aligned}$$

Note that $\Pi_r(x)$ is strictly decreasing in $\beta_r(x^*)$, and that $\Pi_r(x)$ is a function of only $\beta_r(x^*)$ and parameters. Thus, define $L(\cdot)$ so that

$$L(\beta_r(x^*)) = \int_{x^*}^{1} \Pi_r(x)dx$$

So, the objective becomes

$$\begin{aligned} \max_{g_r(\cdot)} \quad & L(\beta_r(x^*)) + \int_q^{x^*} \left\{(1-\beta_r(x))\,T(x)S(x) - \kappa g_r(x)S(x)\right\} dx \\ s.t \quad & \beta_r(x) \geq \lambda(1-g_r(x)) \\ & (\beta_r ST)' \geq ST' \\ & \beta_r \geq 0,\ g_r \in [0,1] \end{aligned}$$

In order to be a solution to this problem, it is necessary that $g_r(\cdot)$ solves

$$
\begin{aligned}
\max_{g_r(\cdot)} \quad & \int_q^{x^*} \left\{ (1-\beta_r(x))\, T(x)S(x) - \kappa g_r(x) S(x) \right\} dx \\
s.t \quad & \beta_r(x) \geq \lambda(1-g_r(x)) \\
& (\beta_r ST)' \geq ST' \\
& \beta_r(x^*) = \beta^* \\
& \beta_r \geq 0,\ g_r \in [0,1]
\end{aligned}
$$

Theorem 12 *Any solution to the previous problem will satisfy the following. For $x \leq x_1$, $\beta_r(x) = \lambda$, $g_r(x) = 0$. For $x > x_1$, the participation constraint binds everywhere, and $g_r(\cdot)$ is picked to support this, so*

$$
\begin{aligned}
\beta_r(x)S(x)T(x) &= \lambda S(x_1)T(x_1) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-x_1)^{b-a} \right) \\
g_r(x) &= 1 - \frac{\lambda S(x_1)T(x_1) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-x_1)^{b-a} \right)}{\lambda S(x)T(x)}
\end{aligned}
$$

The intuition behind this proposition is that the planner would like to have a discontinuity at x^*, so that $\beta_r(x) = \lambda$ for all $x < x^*$, and $\beta_r(x) = 0$ for all $x > x^*$. However, this is not feasible by the participation constraint, so that constraint binds. What remains is to solve the optimal choice of x_1, which I present in the following theorem.

Theorem 13 *x_1 is chosen so that*

$$
\int_{x_1}^{1} \left\{ 1 - \frac{\kappa}{\lambda T(x)} \right\} dx = 0
$$

Thus, for $x \leq x_1$, $\beta_r(x) = \lambda$, $g_r(x) = 0$, and for $x > x_1$, β_r follows the participation constraint, and $g_r(x) = 1 - \frac{\beta_r(x)}{\lambda}$. Further, if

$$
\int_{q}^{1} \left\{ 1 - \frac{\kappa}{\lambda T(x)} \right\} dx > 0
$$

the optimal governance policy is to set $g_r(q) = 1$, $\beta_r(q) = 0$, and allow β_r and g_r to follow participation constraint.

In summary, suppose $\int_q^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx < 0$. To find the optimal governance regulation, pick x_1 so that $\int_{x_1}^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx = 0$. Because for all $x < x_1$, $\beta_r(x) = \lambda$ and $g_r(x) = 0$, by continuity $\beta_r(x_1) = \lambda$ and $g_r(x_1) = 0$. For all $x > x_1$, the endogenous participation constraint binds, so

$$\beta_r(x)T(x)S(x) = \beta_r(x_1)T(x_1)S(x_1) + \frac{AB}{a-b}\left[(1+q-x)^{b-a} - (1+q-x_1)^{b-a}\right]$$

Also, the IC constraint will bind, so $\beta_r(x) = \lambda(1 - g_r(x))$, so by substitution,

$$\lambda(1 - g_r(x))T(x)S(x) = \lambda T(x_1)S(x_1) + \frac{AB}{a-b}\left[x^{b-a} - x_1^{b-a}\right]$$

To implement this, the regulator needs to require that, for firms $x \in [x_1, x^*]$, that each of those firms sets their governance at least as large as $g_r(x)$, where $g_r(\cdot)$ satisfies the above equation. This leads to the following Corollary.

Corollary 14 *When $\int_q^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx < 0$, the optimal governance regulation is increasing in firm size. Further, for the region of firms such that governance is required, the regulated governance minimum is strictly increasing in the size of the firm.*

Therefore, the optimal governance regulation has the following structure. For the smallest firms, the regulator should leave them alone. For the middle firms ($x \in [x_1, x^*]$), the regulator enforces strictly increasing governance requirements. For the largest firms, the regulator can still enforce the above governance schedule. However, she does not need to enforce regulation requirements at those firms, since those firms will govern as closely as possible. x_1 has the following comparative statics.

Corollary 15 *The following comparative statics hold for $(1 - x_1)$. The number of firms that will be forced to exercise governance under optimal regulation is:*

$$\begin{aligned} \textit{increasing in the talent of CEOs} &: \frac{\partial(1-x_1)}{\partial T_{Max}} > 0 \\ \textit{increasing in the severity of moral hazard problems} &: \frac{\partial(1-x_1)}{\partial \lambda} > 0 \\ \textit{decreasing in cost of governance} &: \frac{\partial(1-x_1)}{\partial \kappa} < 0 \end{aligned}$$

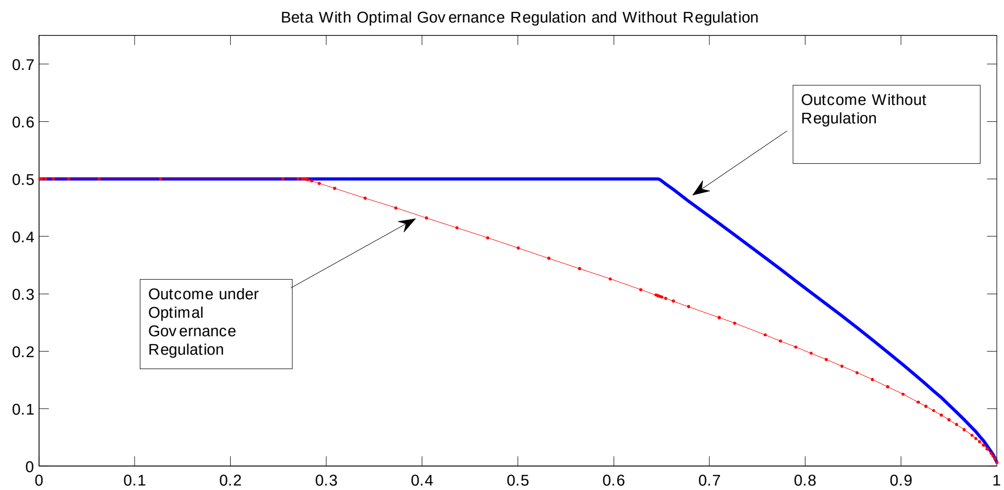

Figure 2: The share of inside equity across firms. The upper line is β in the absence of regulation, while the lower line is β_r in the presence of optimal regulation.

Because of the implementation of optimal governance, the share of inside equity is changed as shown in Figure 2.

3.2 Implementation of Optimal Governance Regulation

In the Section 3.1, I derived the optimal governance regulation. The optimal regulation was chosen based on the assumption that the regulator knows the size of each firm, knows the distribution of talent, and is able to regulate the level of governance at each firm. It may be difficult to enforce governance requirements at all, and very difficult to regulate at the firm level. Instead, in this section I will examine implementation through an alternate mechanism: a subsidy.

A subsidy has the advantage that, rather than forcing firms to do exercise governance, it changes their incentives so that they will exercise governance. The disadvantage of a subsidy is that it is costly for the regulator.

Suppose that the government subsides a share δ of governance costs. Thus, when a firm chooses g, it will cost them $\kappa(1-\delta)gS$. Because this lowers the cost of governance, it induces more firms to exercise governance. Define $x^*(\delta)$ so that

$$T\left(x^*\left(\delta\right)\right) = \frac{\kappa(1-\delta)}{\lambda}$$

Note that the arguments of Section 2 apply here, only substituting in $\kappa' = \kappa(1-\delta)$ for κ. The objective of the regulator is to maximize aggregate profit less aggregate cost of the subsidy:

$$\begin{aligned}
& \int_q^1 \Pi(x)dx - \int_q^1 \delta\kappa g(x)S(x)dx \\
= & \int_q^1 \left\{\left(1-\beta(x)\right)T(x)S(x) - \kappa(1-\delta)g(x)S(x)\right\}dx - \int_q^1 \delta\kappa g(x)S(x)dx \\
= & \int_q^1 \left\{\left(1-\beta(x)\right)T(x)S(x) - \kappa g(x)S(x)\right\}dx
\end{aligned}$$

Therefore, picking δ_1 so that $x^*(\delta_1) = x_1$, the optimal governance is implemented, because firms $[q, x_1]$ choose to exercise governance and firms $[x_1, 1]$ choose not to. This is summarized in the following corollary.

Corollary 16 *Suppose that the government can raise funds from investors efficiently. Then, the optimal governance can be implemented by a subsidy of governance costs. The percentage of governance costs paid by the regulator is* δ_1*, where* $T(x_1) = \frac{\kappa(1-\delta_1)}{\lambda}$ *and* x_1 *satisfies* $\int_{x_1}^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx = 0.$

This result is interesting because the optimal governance regulation can be implemented by a subsidy. This means that, even with all flexibility granted to the regulator, they cannot do any better than they could with this subsidy. It will take some sophistication in order to set the optimal level of the subsidy, but as the next corollary shows, a small subsidy will always improve welfare.

Corollary 17 *If the regulator knows nothing, welfare can still be improved by a small subsidy of governance costs.*

This corollary follows from the fact that, for $x^*(\delta) > x_1$, $\int_{x^*(\delta)}^{1} \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx > 0$, so the subsidy improves investor welfare. Again, this assumes that the investors efficiently pay for the subsidy. However, this may seem unreasonable, because tax collections usually result in dead-weight losses. Thus, it is reasonable to ask if there is a self-financing subsidy & tax scheme that implements the optimal governance.

Theorem 18 *A flat corporate income tax in this environment raises revenue efficiently, so the governance subsidy could be paid for with a flat corporate income tax to implement the optimal governance regulation.*

Thus, it is reasonable that a corporate income tax and corporate governance subsidy could implement the optimal corporate governance regulation. However, there are problems with a subsidy. First, if the subsidy is not self-financing, it may not be politically feasible, since some would consider it a gift to big business. Second, the subsidy may not actually lower the cost of governance. I discuss this in more detail in Section 4.5.

3.3 Optimal Governance Floor

In the Section 3.1, I derived the optimal governance regulation. However, due to practical or legal limitations, the regulator may not be able to regulate different levels of governance for different firms. This can be a real constraint, because in the United States, for example, Sarbanes - Oxley does not have different provisions for small firms. Suppose that the regulator is constrained, so that she must treat all firms the same, though the regulator can force all firms to maintain a minimum level of governance, γ. Further, suppose that the regulator is not allowed to use a subsidy as in Section 3.2. Thus, each firm solves the same problem as before, with the additional constraint that $g(x) \geq \gamma$. Thus, each firm solves

$$\begin{aligned} \max_{\beta_\gamma, g_\gamma} \quad & S(x)T(x)(1-\beta_\gamma(x)) - \kappa g_\gamma(x)S(x) \\ s.t. \quad & \beta_\gamma(x) \geq \lambda(1-g_\gamma(x)) \\ & \beta_\gamma(x)S(x)T(x) \geq W_\gamma(T(x)) \\ & g_\gamma(x) \geq \gamma \end{aligned}$$

which leads us to the next theorem, a generalization of Theorem 2.

Theorem 19 *For $U_0 > S(T-\kappa\gamma)$, project is not carried out. For $\lambda(1-\gamma)ST \leq U_0 \leq S(T-\kappa\gamma)$, the participation constraint is tight enough so that the IC constraint is lax, so*

$$\beta_\gamma(x) = \frac{U_0}{ST}, \;\; g_\gamma(x) = \gamma$$

For $U_0 < \lambda(1-\gamma)ST$, the optimal contract depends on the level of talent. For $T \leq \frac{\kappa}{\lambda}$, the principal will pay agent sufficiently to satisfy incentive problems, setting

$$\beta_\gamma(x) = \lambda(1-\gamma), \;\; g_\gamma(x) = \gamma$$

For $T > \frac{\kappa}{\lambda}$, principal governs closely, so participation constraint will bind, so

$$\beta_\gamma(x) = \frac{U_0}{ST}, \;\; g_\gamma(x) = 1 - \frac{U_0}{\lambda ST}$$

Note that the governance floor not only imposes a cost on firms, but also affect real investment decisions, because it makes an otherwise profitable project be rejected, because when $S(T-\kappa\gamma) \leq U_0 < ST$, the incentive constraint would have been lax, and the project would have been profitable for the firm, but the governance regulation is too costly.

Again, suppose the same specification from the main section, so $S(x) = A(1+q-x)^{-a}$ and $T(x) = T_{Max} - \frac{B}{b}(1+q-x)^b$. Further, I suppose that $T(q) \geq \kappa$, so that all firms will find it optimal to stay open, no matter how harsh the regulation. Note that Lemma 3 holds in this case, so a CEOs' preferred outside option will be to go to a slightly smaller firm rather than to a much smaller firm. This leads us to a generalization of Lemma 5, which leads us to a generalization of Theorem 6.

Lemma 20 *In any equilibrium, there exists x^c such that IR is binding for $x < x^c$, IC is binding for $x \in [x^c, x^*]$, and IR is binding for $x > x^*$. Further, $x^c = q$ iff*

$$\lambda (1 - \gamma) aT_{Max} \geq B \left(1 + \lambda (1 - \gamma) \left(\frac{a}{b} - 1\right)\right)$$

Finally, x^c is increasing in γ.

Theorem 21 *If $x^* > x^c$, the three relevant regions are $[q, x^c)$, $[x^c, x^*)$, and $[x^*, 1]$. For $x \in [q, x^c)$,*

$$\begin{aligned} \beta_\gamma(x)S(x)T(x) &= \lambda(1-\gamma)A\left(T - \frac{B}{b}\right) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - 1\right) \\ g_\gamma(x) &= \gamma \end{aligned}$$

For $x \in [x^c, x^*)$,

$$\begin{aligned} \beta_\gamma(x) &= \lambda(1-\gamma) \\ g_\gamma(x) &= \gamma \end{aligned}$$

For $x \in [x^*, 1]$,

$$\begin{aligned} \beta_\gamma(x)S(x)T(x) &= \kappa(1-\gamma)S(x^*) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right) \\ g_\gamma(x) &= 1 - \frac{\beta_\gamma(x)}{\lambda} \end{aligned}$$

If $x^* \leq x^c$, for all x,

$$\begin{aligned} \beta_\gamma(x)S(x)T(x) &= \lambda(1-\gamma)A\left(T - \frac{B}{b}\right) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - 1\right) \\ g_\gamma(x) &= \max\left\{1 - \frac{\beta_\gamma(x)}{\lambda}, \gamma\right\} \end{aligned}$$

Therefore, I have found the general equilibrium behavior of all firms when in the presence of a governance floor. Now, to find the optimal governance floor, I merely have to optimize over γ. Note that, for $x \in [q, x^c]$, these firms are exercising inefficiently large amounts of governance, because $\beta_\gamma(x) > \lambda(1 - g_\gamma(x))$ and $g_\gamma(x) > 0$.

The profit of a firm x is given by Π_γ, which is given by

$$\Pi_\gamma(x) = (1 - \beta_\gamma(x))S(x)T(x) - \kappa g_\gamma(x)S(x)$$

where β_γ and g_γ are given by the previous proposition. Thus, the regulator's objective function is

$$I(\gamma) = \int_q^1 \Pi_\gamma(x)dx$$

Thus, the optimal regulation floor is solved by maximizing I over γ.

Theorem 22 *$I(\gamma)$ is continuous, differentiable, and weakly concave in γ. Further, $I(\gamma)$ is strictly concave for interior solutions. There exists a unique solution, γ^*, to the regulator's choice of floor. Also, if it is optimal to enforce a floor, there will be inefficiently large levels of governance exercised at some firms. Formally, $\gamma > 0$ implies that $x^c > q$, so that for $x \in (q, x^c)$, $g_\gamma(x) > 1 - \frac{\beta_\gamma(x)}{\lambda}$.*

Corollary 23 *It is optimal to impose a positive governance floor iff*

$$\kappa S(x^*) \int_{x^*}^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx > \kappa \int_q^{x^*} S(x)dx - \lambda \int_q^{x^*} S(x)T(x)dx$$

Corollary 22 shows that the optimal floor can be very unstable. For a very small change in parameters, from a set such that the previous inequality fails to a set of parameters such that it holds, the optimal γ^* will jump from 0 to above γ_0, where γ_0 satisfies

$$\lambda \left(1 - \gamma_0\right) aT_{Max} = B\left(1 + \lambda \left(1 - \gamma_0\right)\left(\frac{a}{b} - 1\right)\right)$$

This also implies that in some scenarios, the optimal floor is 0, which means that it may be optimal to not impose governance regulation if it must floor. This contrasts the previous section, where I found that a little governance will always improve welfare. The difference is that the floor is a very blunt instrument to regulate governance, while I supposed before that the regulation could be fine-tuned to target the middle firms which impose the pecuniary externality.

Also, note that the previous analysis holds because I assumed that $T(q) > \kappa$. This assumption guarantees that all firms will stay open regardless of the choice of regulation. However, if I allow $T(q) < \kappa$, then firms can endogenously choose to close if the governance regulation is too severe. For a governance floor γ, a firm will stay open only if

$$\begin{aligned} (1-\lambda(1-\gamma))S(x)T(x) - \kappa\gamma S(x) &\geq 0 \\ T(x) &\geq \frac{\kappa\gamma}{1-\lambda(1-\gamma)} \end{aligned}$$

Thus, define x_0 such that $T(x_0) = \frac{\kappa\gamma}{1-\lambda(1-\gamma)}$. Endogenizing x_0, however, makes the problem intractable. Though the FOCs can still be found, concavity is no longer satisfied, because we are no longer guaranteed that $\frac{dx^c}{d\gamma} > 0$. It is important to note, however, that the optimal choice of governance floor may force some firms to close.

Thus, it greatly improves welfare for the governance not to be required to take the form of a floor. A governance floor not only causes inefficiently high levels of governance, but it can also force some firms to close.

4 Discussion of Results and Policy Implications

4.1 Sarbanes - Oxley

Sarbanes - Oxley (SOX) was passed in 2002 following the fall of Enron as well as many other accounting scandals in 2001. Zhang (2005) documents a negative stock market response to the passage of SOX. Further, she finds that the market responded negatively to news that the law was more likely to pass or was going to be more severe. Thus, she concludes that the regulation was harmful.

In contrast, Hochberg, Sapienza, and Vissing-Jorgensen (2007) (HSVJ) finds that SOX had a positive effect on some firms' market value. They sort firms based on their lobbying decisions and find that the firms that lobbied against strict implementation of SOX had a positive return relative to other firms.

My model can explain why we see these two facts. Because CEOs at larger firms have bigger pay cuts when regulation is enforced, these CEOs would be the first to lobby against the strict implementation of SOX. This is supported by the fact that the lobbying firms tend to be much larger in HSVJ than those that do not lobby (because the externality is imposed by smaller firms on larger firms). These are the firms who have the best response to any regulation, so they have a positive return when matched with other firms. Taking these two papers together, it appears that SOX went too far, because regulation that was too aggressive would produce both of these effects. HSVJ shows that the regulation produced changes consistent with this paper, but Zhang shows that SOX was expected to do more harm than good.

Therefore, my model shows that in equilibrium, an competitive market for managerial talent will lead to a pecuniary externality. Thus, investor welfare can be improved by increasing governance that lowers private benefits of diversion. Supporting this view, HSVJ finds that the effect is concentrated in those who lobbied against enhanced financial disclosure, since these measures are the most likely to improve transparency and thus decrease private benefits of diversion.

One of the biggest problems with SOX is that it treats all firms the same, forcing all firms to meet stringent governance requirements (though the SEC has considered lowering requirements for small firms, to date they have decided against doing so). Their only outside option is to delist from their exchange. However, this is inefficient, since it hurts diversifiability, and thus investment. This model implies that the required level of governance should be increasing in firm size. Note that this is not because of fixed costs, but because the pecuniary externality that small firms impose on owners of large firms is small, while the externality imposed by medium firms is much larger. By enforcing an increasing governance requirement on firms, this will lessen the damage done to small firms, while still helping large firms.

Chhaochharia - Grinstein (2007) shows that large firms (impacted by SOX) improved in value, while small firms were hurt by SOX.[13] A reasonable question is if governance could really improve value, why were such measures not already carried out? It is difficult to believe that the large value gained by these firms would not have attracted corporate raiders, if a firm could unilaterally carried out this change. Instead, my paper suggests that the losses to small firms were crucial to the gains at large firms, because the governance dampens the pecuniary externality imposed on large firms. Indeed, Chhaochharia - Grinstein (2006) finds that when these changes were carried out, firms that were forced to increase their governance lowered the compensation paid to their CEOs, and specifically lowered the equity-based pay granted to their CEOs, which is also what would happen in this model.

In the model I assumed that there is a single market for CEO talent, and that this talent is transferable to other companies. However, it is unreasonable to suppose that a CEO of a manufacturing firm and a CEO of a financial firm could switch companies without any loss. What is more reasonable is that the market for CEOs is segmented – that there is a market for CEOs of financial firms, a market for CEOs of manufacturing firms, a market for CEOs of retail firms, and so on. Thus, one can interpret this model as expressing an equilibrium in a specific industry.[14]

Also, since each industry likely has a different market for CEOs, different governance laws for different industries will make sense. For example, in industries that are heavily regulated, such as utilities, it seems reasonable to believe that moral hazard problems are smaller, i.e. λ is smaller there.[15] Thus, it will be inefficient to force firms to carry out costly governance measures when there is smaller moral hazard problems, and thus less of a pecuniary externality. This may be difficult to implement, however. First, the traditional challenge that regulation leads to corruption is particularly relevant when the regulator has

[13] Due to the matching methodology used by Chhaochharia-Grinstein (2007), this is not inconsistent with the findings in Zhang.

[14] This interpretation is supported by the fact that the positive returns in HSVJ are dampened when controlling for the 1-digit industry category. My model suggests that most of the positive returns come from size and industry, because industry proxies for labor market conditions for CEOs.

[15] For example, because regulators are already paying very close attention to these firms, it would be more difficult to divert resources from these firms.

discretion. Second, there are legal boundaries to focusing laws on one specific industry, evidenced by the legal challenges against elevated minimum wage laws for big box retailers. However, regulators will likely get close to optimal regulation by conditioning governance regulation on size and 1-digit industry codes.

Finally, there is evidence that regulators underestimated the cost of these governance measures. There are two likely reasons for this. First, because of the fall of Arthur Andersen, the "Big Five" became the "Big Four," so past accounting fees were no longer a good prediction of future accounting fees. Further, regulation on auditor independence likely decreased firms' bargaining power. Thus, these together caused accounting costs to increase. Note that this is the same as the regulator underestimating κ in my model, which will lead to excessive levels of required governance. I discuss endogenous κ in Section 4.6.

The model suggests the following changes to SOX. First, lower governance requirements on the smallest firms. The positive externalities that these firms can spread to other firms is small, and the cost to small firms of constrained financing, because they delisted, is large. Second, implement governance requirements that are increasing in firm size, because this curtails the contagion of executive compensation in a more efficient manner. Finally, condition required governance on industry. If an industry had no moral hazard problem, then there would be no pecuniary externality, and thus the role for governance regulation would be seriously diminished.

4.2 Corporate Governance and Dispersion of Firm Productivity

In this subsection, I examine the impact of firm dispersion on the use of governance. An interesting question about corporate governance is why does its use change over time? The following comparative static gives an interesting explanation for why governance may change over time. In this section, I will assume that governance is not regulated. Recall that the size of firm x in my model is $S(x) = A\,(1+q-x)^{-a}$, so a captures dispersion in productivity.

Corollary 24 *For $x > x^*$, governance increases and pay-performance sensitivity decreases* *when dispersion of firms increases. Formally,* $\frac{\partial g}{\partial a} > 0$ *and* $\frac{\partial \beta}{\partial a} < 0$.

The intuition for this result is as follows. First, when firms are really close together, they force the market price for managerial talent high. The large firms, thus, are forced to solve their agency problems by granting the manager equity-based pay. However, when firms are more dispersed, firms do not need to grant as much equity to their managers, so they will exercise more governance.

This suggests several interesting implications. First, one could test this hypothesis across industries, because it is reasonable to suggest that the market for CEOs is segmented by industry. It is possible to use dispersion in q or factor productivity as a proxy for dispersion here. Second, one could test this hypothesis across countries. Finally, one could test this across time.

Eisfeldt - Rampini (2006) finds that dispersion of productivity is counter-cyclical. Thus, variance of productivity is higher in recessions than in booms. Given this, my model suggests that governance will also be counter-cyclical. That is, governance should be tighter in recessions and lax in booms. Many have suggested that governance became slack during the 90s, attributing this to investors becoming lazy because of large returns. The past corollary suggests, instead, that this may have been optimal.

4.3 Impact on Real Investment Decisions

In my model, I assumed that the objective of the regulator is to maximize investor welfare. A reasonable question is why is investor welfare important? In this section, I will show that expected return is critical in determining the level of investment. Specifically, I will endogenize A, where the size of the firm $S(x) = A\,(1 + q - x)^{-a}$.

First, for a level of A, define the profit of firm x without regulation as $\Pi_{nr}(x; A)$, so

$$\Pi_{nr}(x; A) = (1 - \beta(x))S(x)T(x) - \kappa g(x)S(x)$$

Thus, for $x \leq x^*$, $\Pi_{nr}(x;A) = (1-\lambda)S(x)T(x)$, and for $x > x^*$,

$$\begin{aligned}\Pi_{nr}(x;A) &= S(x)\left(T(x)-\kappa\right) \\ &- \left(\frac{AB}{a-b}\left[(1+q-x)^{b-a}-(1+q-x^*)^{b-a}\right]+\kappa S\left(x^*\right)\right)\left(1-\frac{\kappa}{\lambda T\left(x\right)}\right)\end{aligned}$$

Next, define the profit of firm x with optimal regulation as $\Pi_r(x;A)$, so

$$\Pi_r(x;A) = (1-\beta_r(x))S(x)T(x) - \kappa g_r(x)S(x)$$

Thus, for $x < x_1$, $\Pi_r(x;A) = (1-\lambda)S(x)T(x)$, and for $x > x_1$,

$$\begin{aligned}\Pi_r(x;A) &= S(x)\left(T(x)-\kappa\right) \\ &- \left(\frac{AB}{a-b}\left[(1+q-x)^{b-a}-(1+q-x_1)^{b-a}\right]+\lambda S\left(x_1\right)T\left(x_1\right)\right)\left(1-\frac{\kappa}{\lambda T\left(x\right)}\right)\end{aligned}$$

Note that $\Pi_{nr}(x;A) = A\Pi_{nr}(x;1)$, and $\Pi_r(x;A) = A\Pi_r(x;1)$, so thus profit scales in the parameter A.

Suppose that a continuum of identical investors have the choice between investing in an index fund[16] and consuming. So, by investing I_i, agent i will receive a share of $\frac{I_i}{\int_0^1 I_i di}$ of all companies. In the absence of regulation, the investor i will receive investment income

$$D_{nr}(I_i) = \frac{I_i}{\int_0^1 I_i di}\int_q^1 \Pi_{nr}(x;A)dx$$

Let the market for capital clear, so

$$\begin{aligned}\int_q^1 Ax^{-a}dx &= \int_0^1 I_i di \\ A &= \frac{\int_0^1 I_i di}{\int_q^1 x^{-a}dx}\end{aligned}$$

Thus, by scale invariance of profit, the investment income that the investor receives will be

$$\begin{aligned}D_{nr}(I_i) &= \frac{I_i}{\int_q^1 x^{-a}dx}\int_q^1 \Pi_{nr}(x;1)dx \\ D_{nr}(I_i) &= I_i R_{nr}\end{aligned}$$

[16] I need to assume that investors can only invest in a mutual fund for tractability.

where $R_{nr} = \frac{\int_q^1 \Pi_{nr}(x;1)dx}{\int_q^1 x^{-a}dx}$. Similarly, under optimal governance regulation, $D_r(I_i) = I_i R_r$, where $R_r = \frac{\int_q^1 \Pi_r(x;1)dx}{\int_q^1 x^{-a}dx}$. Note that $R_r > R_{nr}$. Further, investment income will be risk-free by the law of large numbers.

Thus, suppose that each investor maximizes

$$\begin{aligned} U &= u(c_0) + \beta u(c_1) \\ c_0 &= w - I_i \\ c_1 &= I_i R_g \end{aligned}$$

where R_g is the return under the prevailing governance regulation, $u' > 0$, $u'' < 0$, and u has relative risk aversion of less than 1. Substituting in, the investor's problem becomes

$$\max_{I_i} \{u(w - I_i) + \beta u(I_i R_g)\}$$

The first order condition for this problem is

$$u'(w - I_i) = \beta R_g u'(I_i R_g)$$

By differentiating both sides with respect to R_g, we find that I_i is increasing in R_g.

So, because $R_{nr} < R_r$, not only will returns under optimal investment be higher, but also that investment will be higher, because the agency costs are lower. Further, by the envelope theorem, when R_g increases by ε each investor is better off by $\beta u'(I_i R_g)\varepsilon$.

Therefore, when the regulator steps in and enforces optimal corporate governance regulation, investment increases. This is because investors make their investment decisions based upon how much they expect to get in return. Because the optimal governance regulation improves the payout to investors, more is invested and the economy is improved. This increase in efficiency is the result of reducing aggregate agency costs. Also, because R_f, the return from the choice of an optimal floor, satisfies $R_{nr} \leq R_f < R_r$, the optimal governance regulation will result in more investment than under the best governance floor, which results in more investment than the equilibrium without governance. Therefore, corporate governance regulation can have real effects, because if it is done efficiently it will increase investment by lowering agency costs. In the next section, I will discuss how this agency cost can also explain hurdle rates that seem too high.

4.4 Endogenous Hurdle Rates Premiums

A hurdle rate is often used in capital budgeting decisions. For example, Graham-Harvey (2001) finds that 56.94% of CFOs surveyed used a hurdle rate, making it the third most used technique behind net present value and internal rate of return. A hurdle rate criterion is similar to an internal rate of return, except that the project is accepted iff the internal rate of return is greater than the hurdle rate (the internal rate of return criterion uses the correct risk-adjusted rate).

The common criticism of the use of a hurdle rate is that the rates used are too high. Indeed, Meier - Tarhan (2007) documents that the average hurdle rate that firms reported in a survey was 14.1%, which is equivalent to a real rate of 11.6%. This is a premium over cost of capital of between 5% and 7.5%, depending on the level of the equity premium. Thus, firms appear to be using too large of a rate. Under the usual assumptions, this hurts firm value by biasing projects against long-term projects and by causing firms to reject profitable projects that have a smaller internal rate of return than the hurdle rate. Meier - Tarhan (2007) attribute the large hurdle rates to a number of frictions, including financial flexibility of the firm.

In this section, I will show how agency costs will lead to hurdle rates that exceed the cost of capital, and that the hurdle rate is correlated with the size of the firm, level of corporate governance, and the CEO's pay-performance sensitivity.

Suppose that after each firm enters into a contract with their CEO, which specifies the share of the firm granted to the CEO, $\beta(x)$, and the level of governance, $g(x)$, that the company unexpectedly finds the opportunity for a new project, independent of existing projects. Suppose that the project can either succeed or fail, and succeeds with probability

p. In case of success, the project pays R, but pays 0 if it fails. However, the CEO can report that the project failed and divert the cash flow R to personal consumption, which gives him private benefit $\lambda(1-g(x))$. In order to prevent him from doing this, the company must grant him a share of the new project at least as large as $\lambda(1-g(x))$. Thus, $\beta(x) = \lambda(1-g(x))$. Thus, supposing a unit discount factor, the project is taken iff it has positive payoff to investors after agency costs, which holds iff

$$\begin{aligned} p\left(R-\beta(x)R\right) &\geq I \\ \frac{pR}{(1+h(x))} &\geq I \end{aligned}$$

where

$$\frac{1}{1+h(x)} = 1-\beta(x)$$

Thus, $h(x)$ is the hurdle rate when the interest rate is 0.

First, note that $h(x) > 0$ for all x, so that all firms have a positive hurdle rate premium above the cost of capital. The reason for this is that I assume that the manager is unable to invest enough to buy his own share of the project, and thus the project must earn enough to not only pay back the investors' their money, but also enough to cover the agency costs. Thus, I have endogenously derived a role for the hurdle rate premium, because of agency costs. Further, this shows that h will be large when β is large and small when β is small. Therefore, we would expect that the hurdle rates will be larger at smaller companies, adjusted for risk, because smaller companies will need to devote a larger share of cash flows to agency problems. Finally, hurdle rates will be decreasing in the quality of governance and increasing in the manager's pay-performance sensitivity.

4.5 Mutual Funds and Governance

Governance is also important in mutual funds. It is commonly observed that mutual funds have poor governance, and that consumers are taken advantage of by them. This view is supported by Kuhnen (2007), which finds favoritism in the relationships of fund directors and advisory firms, indicating poor governance. Many attribute poor governance to entrenchment of fund managers. Because funds are easily transferable between mutual funds, it is difficult to imagine that investors do not have other outside options.

This paper sheds light on this issue. This model works for mutual funds and fund managers as well as it does for firms and executives. At small funds, it may not be worth it to exercise governance, so these funds will pay the manager generous fees. This affects large funds, however, because they must pay their managers large enough not to leave for a smaller fund. Because governance and compensation are substitutes, large funds will find it optimal to exercise less governance as well, resulting pervasive lax governance.

I should note that fund families do not alleviate this problem. Suppose that Fidelity decided to increase governance in one of their smaller funds, suboptimally, hoping to improve the profitability of the large funds. This would not work because the other small funds would still pay instead of exercise governance, still imposing the pecuniary externality on the larger funds. Thus, because there are many fund families, it is would be very difficult for these fund families to be able to coordinate among themselves sufficiently to solve the governance problem.

4.6 Problems with Subsidies and Endogenous Governance Costs

In Section 3.2, I show that the optimal governance regulation can be implemented by granting a subsidy for governance costs. However, this might not work in practice, because I assume that the cost of governance is $\kappa g S$, for a constant κ across firms and various regulatory environments. If we interpret the cost of corporate governance as accounting fees, the accounting firms may extract some of this benefit of this subsidy in higher rates.

Further, under governance schemes where corporate governance is required by the government, the parameter κ could change. Though the market for governance could be perfectly competitive, it very likely would have an upward sloping aggregate supply curve. Supposing that κ is a function of $\int_q^1 g(x)S(x)dx$, this would impact the optimal governance regulation.

Though the optimal regulation would still take the same form, the government would require less of it. Also, implementing a governance floor becomes even more difficult, because a floor induces inefficiently large levels of governance to be carried out. Further, regulation gets even worse if the regulation gives the accounting firms market power, because then they could extract rents from firms. There is some evidence that this happened after Sarbanes - Oxley passed, due to the large increase in compliance costs.

5 Conclusion

In this paper, I model a general equilibrium in the market for executives. At each firm, there is a moral hazard problem, which can be dealt with by paying the manager enough to induce truth telling or by governing the company very closely. Thus, pay and governance are substitutes. When a firm increases governance, that firm can lower its pay, and other firms can lower their pay as well. Thus, there is a pecuniary externality in this setting, so governance regulation is optimal for investors.

This paper predicts that firms will react to regulation in the following ways. First, the value of large firms will improve, but value of small firms will fall. Second, following governance regulation, equity-based compensation will fall, because governance and incentive pay are substitutes.

Further, this model suggests that the dispersion of firms may have a role in explaining differences in corporate governance across time, industries, and countries. Specifically, when productivities of firms are far apart, the outside option of the manager will not require that large of a share of the firm to be given to him, so the firm will exercise more governance. Due to the findings in Eisfeldt - Rampini (2006) that dispersion of productivity tends to be higher in recessions, this would suggest that governance would become more lax during booms and tighter during recessions.

This model predicts that the compensation of CEOs will be increasing in firm size, while the pay-performance sensitivity will decrease in firm size. These patterns in executive compensation are consistent with Murphy (1999). The empirical finding that pay-performance sensitivity is decreasing in firm size is usually attributed to risk aversion. All parties are risk-neutral in my paper, but I find this because the largest firms choose to exercise more governance, which substitutes for pay.

Also, I find that the larger firms should exercise more governance than smaller firms, and that the optimal governance regulation forces some of the medium-sized firms to govern more closely than they would like to in order to dampen the contagion of executive pay. Thus, the optimal governance policy should ignore the smallest firms. The optimal regulation will also be strictly increasing in the size of the firm.

If the regulator is restricted to a floor on governance, the optimal floor may be zero. Further, if the floor is positive, it will cause inefficient levels of governance to be used at some firms, because the endogenous participation constraint combined with the required level of governance will cause the incentive compatibility constraint to hold with strict inequality, implying that these firms could lower their levels of governance (if regulation was relaxed for that firm), improve their profit, and hurt no other firms. I also show that when regulation improves investor welfare, investment will increase, so long as investors are not too risk-averse.

Finally, I show that marginal hurdle rates will be higher at smaller, less governed firms than at larger, better governed firms, because marginal agency costs are higher at the smaller firms than at the larger firms. This addresses why there are financial constraints limiting capital to smaller firms.

Two important issues that this paper does not address are endogenous governance costs and delisting. In this paper, I assume that the cost of governance is independent of regulation. However, since the passage of SOX, accounting fees have grown. Because firms are required to carry out these measures, they lose bargaining power over price. Thus, it would be interesting to incorporate this into the model, perhaps by making κ, the cost of governance, increase when regulation increases (for example, by making κ an increasing function of aggregate regulated governance expenditures). Further, in this paper, I assume that if a firm does not want to follow regulation, their only alternative is to close. Under SOX, however, firms have the option to delist, yet still operate. Thus, another interesting extension would be to model this decision, and the impact that this option has on other firms.

Another important extension to this model would be endogenizing the size of the firms through a competitive capital market. For tractability, I assumed that the size of each firm was fixed, though in an extension I endogenized the size of the economy by scaling all firms together. It would be interesting to examine an equilibrium where each firm chose its size endogenously, then hired a manager in the market. It would also be interesting to examine the impact of long-term contracting in the dynamic extension to this environment.

References

[1] Admati, Anat and Paul Pfleiderer (2000) "Forcing Firms to Talk: Financial Disclosure Regulation and Externalities," *Review of Financial Studies*, **13**, 479-519.

[2] Bebchuk, Lucian and Jesse Fried (2004) *Pay Without Performance: The Unfulfilled Promise of Executive Compensation* (Harvard University Press).

[3] Chhaochharia, Yidhi and Yaniv Grinstein (2006) "CEO Compensation and Board Structure" Cornell University Working Paper

[4] Chhaochharia, Yidhi and Yaniv Grinstein (2007) "Corporate Governance and Firm Value: The Impact of the 2002 Governance Rules," *Journal of Finance*, **62**, 1789-1825.

[5] Connolly, Marie and Alan B. Krueger (2005) "Rockonomics: The Economics of Popular Music," Princeton Working Paper.

[6] DeMarzo, Peter M. and Michael J. Fishman (2007) "Agency and Optimal Investment Dynamics" *Review of Financial Studies*, **20**, 151-188.

[7] Diamond, Douglas W. (1984) "Financial Intermediation and Delegated Monitoring," *The Review of Economic Studies* **51**, 393-414.

[8] Dicks, David L. (2007) "A Theory of Excessive Perks," Kellogg School of Management Working Paper.

[9] Dye, Ronald (1990) "Mandatory versus Voluntary Disclosure: The Cases of Financial and Real Externalities," *The Accounting Review,* **65**, 1-24

[10] Edmans, Alex, Xavier Gabaix and Augustin Landier (2007) "A Calibratable Model of Optimal CEO Incentives in Market Equilibrium," Available at SSRN: http://ssrn.com/abstract=999096.

[11] Eisfeldt, Andrea L. and Adriano A. Rampini (2006) “Capital Reallocation and Liquidity,” *Journal of Monetary Economics*, **53**, 369-399.

[12] Fishman, Michael J. and Kathleen M. Hagerty (2003) “Mandatory Versus Voluntary Disclosure in Markets with Informed and Uninformed Customers,” *The Journal of Law, Economics, & Organization*, **19**, 45-63.

[13] Frydman, Carola (2005) "Rising Through the Ranks: The Evolution of the Market for Corporate Executives, 1936-2003," MIT Working Paper.

[14] Gabaix, Xavier and Augustin Landier (2008) “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics*, forthcoming.

[15] Gompers, Paul A., Joy Ishii, and Andrew Metrick (2003) “Corporate Governance and Equity Prices,” *Quarterly Journal of Economics*, **118**, 107-155.

[16] Graham, John and Campbell Harvey (2001) “The Theory and Practice of Corporate Finance: Evidence from the Field,” *Journal of Financial Economics*, **60**, 187-243.

[17] He, Zhiguo (2007) “Agency Problems, Firm’s Valuation, and Capital Structure,” Kellogg School of Management Working Paper.

[18] Hermalin, Benjamin E. and Michael S. Weisbach (1998) “Endogenously Chosen Boards of Directors and Their Monitoring of the CEO,” *American Economic Review*, **88**, 96-118.

[19] Hochberg, Yael, Paola Sapienza, and Annette Vissing-Jorgensen (2007) “A Lobbying Approach to Evaluating the Sarbanes - Oxley Act of 2002,” NBER Working Paper 12952.

[20] Krueger, Alan (2005) “The Economics of Real Superstars: The Market for Rock Concerts in the Material World,” *Journal of Labor Economics,* **23**, 1-30.

[21] Kuhnen, Camelia (2007) "Social Networks, Corporate Governance and Contracting in the Mutual Fund Industry," Kellogg School of Management Working Paper.

[22] Lacker, Jeffrey M. and John A. Weinberg (1989) "Optimal Contracts Under Costly State Falsification," *The Journal of Political Economy*, **97**, 1345-1363.

[23] Meier, Iwan and Vefa Tarhan (2007) "Corporate Investment Decision Practice and the Hurdle Rate Premium Puzzle," Loyola University Chicago Working Paper.

[24] Murphy, Kevin (1999) "Executive Compensation," *Handbook of Labor Economics Volume 3* (Elsevier).

[25] Rosen, Sherwin (1981) "The Economics of Superstars," *American Economic Review*, **71**, 845-858.

[26] Rosen, Sherwin (1982) "Authority, Control and the Distribution of Earnings," *Bell Journal of Economics*, **13**, 311-323.

[27] Rosen, Sherwin (1990) "Contracts and Market for Executives," University of Chicago - Economics Research Center 90-12.

[28] Tervio, Marko (2007) "The Difference That CEOs Make: An Assignment Model Approach," University of California-Berkeley Working Paper.

[29] Westphal, J.D. & Stern, I. (2007) "Flattery will get you everywhere (Especially if you are a male Caucasian): How ingratiation, boardroom behavior, and demographic minority status affect the likelihood of gaining board appointments at US Companies." *Academy of Management Journal*, Forthcoming.

[30] Zhang, Ivy (2005) "Economic Consequences of the Sarbanes - Oxley Act of 2002," University of Rochester Working Paper.

A Appendix A: Proofs

Proof of Lemma 1. Pick an optimal contract, and pick a set of positive measure $A \subset Y$ where the manager would report $\widehat{y}(y;\lambda) < y$. Set $C(y) = \lambda(1-g)(y-\widehat{y}) + C(\widehat{y})$ for all $y \in A$, which induces manager to tell the truth. This improves objective by at least $[1-\lambda(1-g)]\int_A (y-\widehat{y}(y,\lambda))\,dF > 0$. This is at least, because manager might report y instead of $\widehat{y}$ when $y' > y$ is observed by manager. Because this change improves objective, the contract cannot be optimal. Further, because this argument works for $\forall A \subseteq Y$, principal would always find it optimal to inducing truth-telling. ■

Proof of Theorem 2. When $U_0 > ST$ project ceases to be profitable, since efficiency is improved by letting agent accept his outside option. When $\lambda ST \leq U_0 \leq ST$, setting $C(y) = \frac{U_0}{ST}y$, clearly satisfies the IC (with $g = 0$) and LL. This is optimal because principal cannot lower agent's expected payoff, since the IR binds.

Lastly, consider $U_0 < \lambda ST$, and suppose g is optimal level of governance. Note WLOG $C(0) = 0$, since IC is slack at 0, so LL binds. Suppose, at the solution to the above problem, $\exists z \in \text{support}(F)$ such that $C(STz) = \lambda(1-g)STz + \delta$, for $\delta > 0$. The IC Constraint implies that $C(STz') \geq \lambda(1-g)STz' + \delta$ for all $z' \in \overline{Z}(z) = \{z'|z' \in \text{support}(F), z' \geq z\}$. Thus, by setting $C(STz) = \lambda(1-g)STz$, we can also decrease $C(STz')$ by δ as well, and thus improve objective by $\delta[1 - \lim_{\xi \to z^-} F(\xi)]$. Thus, $C(STz) = \lambda(1-g)STz$ almost surely in any optimal contract. The problem simplifies to

$$\begin{aligned} \max_{g} \quad & ST(1-\lambda(1-g)) - \kappa g S \\ s.t. \quad & \lambda(1-g)\,ST \geq U_0, \\ & g \in [0,1] \end{aligned}$$

Note that objective becomes $ST(1-\lambda) + S(\lambda T - \kappa)g$, so increasing g is beneficial to principal iff $T > \frac{\kappa}{\lambda}$. If $T \leq \frac{\kappa}{\lambda}$, $g = 0$, and thus $C(STz) = \lambda STz$. If $T > \frac{\kappa}{\lambda}$, IR binds, so $\lambda(1-g)ST = U_0$. Thus, $g = 1 - \frac{U_0}{\lambda ST}$, and $C(STz) = U_0 z$. ■

Proof of Lemma 3. Because $W(T(x)) \geq \sup_{x'<x}[S(x')[T(x) - T(x')] + \beta(x')S(x')T(x')]$,

$W(T(x)) \geq [S(x-dx)[T(x)-T(x-dx)]+\beta(x-dx)S(x-dx)T(x-dx)]$. Because $\beta(x)S(x)T(x) \geq W(T(x))$, $\beta(x)S(x)T(x) \geq [S(x-dx)[T(x)-T(x-dx)]+\beta(x-dx)S(x-dx)T(x-dx)]$, which implies $\beta(x)S(x)T(x)-\beta(x-dx)S(x-dx)T(x-dx) \geq S(x-dx)[T(x)-T(x-dx)]$, so $(\beta ST)' \geq ST'$. Thus, for $x'<x$, $\beta(x)S(x)T(x)-\beta(x')S(x')T(x') \geq \int_{x'}^{x} S(z)T'(z)dz$, so $\beta(x)S(x)T(x) \geq \beta(x')S(x')T(x')+\int_{x'}^{x} S(z)T'(z)dz > \beta(x')S(x')T(x')+S(x')(T(x)-T(x'))$, where strict inequality holds because S is strictly increasing. Thus, local moves down are better for the agent than big moves down. Firms with $x'>x$ will not consider manager x because in order to convince firm x' to accept manager x, manager x will have to accept a smaller share than $\beta(x')$ of the new firm, and because the firm cannot increase governance, manager x would not report truthfully. This assumes $\beta\left(x'\right) \leq \lambda$. If $\beta\left(x'\right)>\lambda$, then the differential equation holds with equality, so for $x''=x'-\delta$, $\beta(x')S\left(x'\right)T\left(x'\right)=\beta(x'')S(x'')T(x'')+\int_{x''}^{x'} S(z)T'(z)dz$, so $\beta(x')S\left(x'\right)T\left(x'\right)-\beta(x'')S(x'')T(x'')=\int_{x''}^{x'} S(z)T'(z)dz<S(x')[T(x')-T(x'')]$, so manager x'' could not induce firm x' to replace her manager with him without taking less pay than he originally had at firm x''. This logic extends to all $x<x'$. ■

Proof of Theorem 4. The proof will be structured in the following way. First, I will assume that $\beta(x)=\lambda$ for $x<x^*$ in a neighborhood of x^*. At the end of the proof, I will show that the argument extends to a general distribution of S and T.

Under the competitive equilibrium, large firms pay executive compensation equal to

$$\beta(x)S(x)T(x)=\beta(x^*)S(x^*)T(x^*)+\int_{x^*}^{x} S(u)T'(u)du$$

and must set governance equal to

$$g(x)=\max\left\{1-\frac{\beta(x)}{\lambda},0\right\}$$

Contingent on governance taking positive values, this governance will have cost

$$\begin{aligned}\kappa g(x)S(x) &= \kappa S(x)-\kappa\frac{\beta(x)S(x)T(x)}{\lambda T(x)} \\ &= \kappa S(x)-\frac{\kappa}{\lambda T(x)}\left[\beta\left(x^*\right)S\left(x^*\right)T\left(x^*\right)+\int_{x^*}^{x} S(u)T'(u)du\right]\end{aligned}$$

Thus, if the regulator enforces regulation that decreases $\beta(x^*)$ by a small amount, ε, this will lower $\beta(x)S(x)T(x)$ by $\varepsilon S(x^*)T(x^*)$ and raise governance costs by $\frac{\kappa}{\lambda T(x)}\varepsilon S(x^*)T(x^*)$ if $\beta(x) \leq \lambda$, and by 0 if $\beta(x) > \lambda$. Thus, the profit of large firms $(x > x^*)$ will increase by at least $\left(1 - \frac{\kappa}{\lambda T(x)}\right)\varepsilon S(x^*)T(x^*)$, so the aggregate profit of these large firms increase by at least $\varepsilon S(x^*)T(x^*)\int_{x^*}^{1}\left\{1 - \frac{\kappa}{\lambda T(x)}\right\}dx$. Thus, define $Benefit = \varepsilon S(x^*)T(x^*)\int_{x^*}^{1}\left\{1 - \frac{\kappa}{\lambda T(x)}\right\}dx$, though the benefit will be at least this big.

In order to regulate a decrease of ε in $\beta(x^*)$, the regulator can merely force the firms near x^* to raise governance. Setting the slope to minimize the number of affected firms, the regulator sets governance so that the compensation function, βST, is as steep as possible for regulated firms. This holds if, for regulated x,

$$\beta(x^*)S(x^*)T(x^*) = \beta(x)S\left(x\right)T(x) + \int_{x}^{x^*} S(z)T'(z)dz$$

Thus, define δ such that

$$\left(\lambda - \varepsilon\right)S(x^*)T(x^*) = \lambda S\left(x^* - \delta\right)T(x^* - \delta) + \int_{x^*-\delta}^{x^*} S(z)T'(z)dz$$

Because $\beta(x) > \lambda - \varepsilon$, $g(x) < \frac{\varepsilon}{\lambda}$ for $x < x^*$. The number of the firms that must be regulated will be $\delta = \varepsilon\frac{ST}{\lambda S'T - (1-\lambda)ST'}$, evaluated at x^*, by first order approximation (the denominator will be positive by assumption that endogenous IR is slack near x^*). Thus, the social cost of enforcing this regulation will be less than $\int_{x^*-\delta}^{x^*}\kappa g(x)S(x) < \kappa\left(\frac{\varepsilon}{\lambda}\right)S(x^*)\left(\varepsilon\frac{ST}{\lambda S'T-(1-\lambda)ST'}\right) = \varepsilon^2\frac{S^2T^2}{\lambda S'T-(1-\lambda)ST'}$. Thus define $Cost = \varepsilon^2\frac{S^2T^2}{\lambda S'T-(1-\lambda)ST'}$, though the cost will be less than this.

Note that

$$\frac{Cost}{Benefit} = \varepsilon\frac{\frac{S^2T^2}{\lambda S'T-(1-\lambda)ST'}}{S\left(x^*\right)T(x^*)\int_{x^*}^{1}\left\{1 - \frac{\kappa}{\lambda T(x)}\right\}dx}$$

which implies that

$$\lim_{\varepsilon\to 0}\left[\frac{Cost}{Benefit}\right] = 0$$

which proves that, for small ε, such a change improves welfare.

The intuition behind this proof is that such a change benefits a large number of firms for a proportional cost to a small number of firms. The argument gets even stronger if the condition that $\beta(x) = \lambda$ for x near x^* fails, because the set of firms receiving the benefit of lower wages gets even bigger, and these firms would not have to pay higher governance costs because $\beta(x) > \lambda$ at these firms, and it only takes regulating at an infinitesimally small set to improve the welfare of all of these firms. This will hold because, for any smooth strictly increasing S and T, either the measure of firms with $\beta(x) = \lambda$ will be zero or there will be a connected region such that $\beta(x) = \lambda$ on that region. Thus, the ε^2 type of argument works with any of these distributions as well, because either there will be a countable number of regions with measure of order ε that will be forced to exercise privately suboptimal governance. ■

Proof of Lemma 5. Consider the following function, ϕ:

$$\begin{aligned}\phi(x) &= \beta(q)S(q)T(q) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-q)^{b-a}\right) - \lambda S(x)T(x) \\ &= \lambda A(T_{Max} - \frac{B}{b}) + \frac{AB}{a-b}((1+q-x)^{b-a} - 1) - \lambda A\,(1+q-x)^{-a}\left(T_{Max} - \frac{B}{b}\,(1+q-x)^{b}\right)\end{aligned}$$

Note that $\phi(\cdot)$ is the difference between pay required to keep manager x from switching to firm $x - dx$, given that such contagion begins at $x = q$, and the amount firm x must pay manager x to solve the incentive problem. Note that

$$\begin{aligned}\frac{d\phi}{dx} &= AB\,(1+q-x)^{b-a-1} - \lambda a A\,(1+q-x)^{-a-1}\,(T_{Max} - \frac{B}{b}\,(1+q-x)^{b}) - \lambda AB\,(1+q-x)^{b-a-1} \\ &= A\,(1+q-x)^{-a-1}\left[B\left(1+\lambda\left(\frac{a}{b}-1\right)\right)(1+q-x)^{b} - \lambda a T_{Max}\right]\end{aligned}$$

Note that $\phi(q) = 0$, and that $\phi'(x) < 0$ iff $B(1+\lambda\left(\frac{a}{b}-1\right))\,(1+q-x)^{b} < \lambda a T_{Max}$, iff $x > 1 + q - \left[\frac{\lambda a T_{Max}}{B\left(1+\lambda\left(\frac{a}{b}-1\right)\right)}\right]^{1/b} \equiv x'$. Thus, for $x > x'$, $\phi'(x) < 0$, and $x < x'$, $\phi'(x) > 0$. If $\phi'(q) > 0$, then $x' > q$, so $\phi(x') > \phi(q) = 0$, so there exists a $x^c \in (x', 1+q)$ such that $\phi(x^c) = 0$, because $\lim\limits_{x\to 1+q} \phi(x) = -\infty$, so $\phi(x) > 0$ for all $x < x^c$ and $\phi(x) < 0$ for all $x > x^c$. When $\phi(x) > 0$, the IR from 1 is binding, so firm x must pay their manager x more than is necessary for IC constraint to keep her from going to firm $x - dx$.

Now suppose $x^* > x^c$, and consider $x \in (x^c, x^*)$. Consider IR constraint, that

$$\begin{aligned} (\beta ST)' &\geq ST' \\ \beta' ST + \beta S' T + \beta ST' &\geq ST' \end{aligned}$$

Now consider when this is satisfied by $\beta(x) = \lambda$, $\beta'(x) = 0$. The IR constraint is satisfied iff

$$\begin{aligned} \lambda S' T &\geq (1-\lambda) ST' \\ \lambda a T_{Max} &\geq \left(1 + \lambda \left(\frac{a}{b} - 1\right)\right) B \left(1 + q - x\right)^b \end{aligned}$$

which is identical to the condition on x such that $\phi'(x) < 0$. Thus, because $x > x^c > x'$, $\phi'(x) < 0$, so the IR constraint is slack for $x \in (x^c, x^*)$.

Finally, $x^c = q$ iff $\phi'(q) \leq 0$, which holds iff $B\left(1 + \lambda\left(\frac{a}{b} - 1\right)\right) \leq \lambda a T_{Max}$. ■

Proof of Theorem 6. By previous lemma, IR binds for $x \in [x^*, 1]$, and IC binds for $x \in (q, x^*]$, because $x^c = q$. Thus, for $x \leq x^*$, $\beta(x) = \lambda$ and $g(x) = 0$. Also, by definition, $T(x^*) = \frac{\kappa}{\lambda}$, so for all $x > x^*$, the endogenous IR constraint will bind, so

$$\begin{aligned} \beta(x)S(x)T(x) &= \beta(x^*)S(x^*)T(x^*) + \frac{AB}{a-b}((1+q-x)^{b-a} - (1+q-x^*)^{b-a}) \\ &= \lambda S(x^*)\frac{\kappa}{\lambda} + \frac{AB}{a-b}((1+q-x)^{b-a} - (1+q-x^*)^{b-a}) \\ &= \kappa S(x^*) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right) \end{aligned}$$

The results for $g(\cdot)$ holds because $g(x) = \max\{0, 1 - \frac{\beta(x)}{\lambda}\}$. ■

Proof of Corollary 8. Note that x^* is defined so that $T_{Max} - \frac{B}{b}\left(1 + q - x^*\right)^b - \frac{\kappa}{\lambda} = 0$. Thus, differentiating with respect to T_{Max}, I find that

$$\begin{aligned} 1 + B\left(1 + q - x^*\right)^{b-1} \frac{\partial x^*}{\partial T_{Max}} &= 0 \\ \frac{\partial x^*}{\partial T_{Max}} &= -\frac{1}{B\left(1 + q - x^*\right)^{b-1}} \end{aligned}$$

By similar logic,

$$\begin{aligned}
\frac{\partial x^*}{\partial \lambda} &= -\frac{\kappa}{\lambda^2}\frac{1}{B\left(1+q-x^*\right)^{b-1}} \\
\frac{\partial x^*}{\partial \kappa} &= \frac{1}{\lambda B\left(1+q-x^*\right)^{b-1}} \\
\frac{\partial x^*}{\partial B} &= \frac{1+q-x^*}{Bb} \\
\frac{\partial x^*}{\partial b} &= \left(1+q-x^*\right)\left[\frac{\log(1+q-x^*)}{b}-\frac{1}{b^2}\right]
\end{aligned}$$

Finally, note $\frac{\partial(1-x^*)}{\partial\theta} = -\frac{\partial x^*}{\partial\theta}$ for any parameter θ. ∎

Proof of Corollary 9. Recall that, when $g(x) > 0$, $g(x) = 1 - \frac{\beta(x)}{\lambda}$, and β satisfies

$$\begin{aligned}
\beta(x)S(x)T(x) &= \beta(x^*)S(x^*)T(x^*) + \frac{AB}{a-b}\left(\left(1+q-x\right)^{b-a} - \left(1+q-x^*\right)^{b-a}\right) \\
&= \kappa S(x^*) + \frac{AB}{a-b}\left(\left(1+q-x\right)^{b-a} - \left(1+q-x^*\right)^{b-a}\right)
\end{aligned}$$

because $\beta(x^*) = \lambda$ and $T(x^*) = \frac{\kappa}{\lambda}$. Thus,

$$\begin{aligned}
\frac{\partial\beta}{\partial\kappa} &= \frac{1}{S(x)T(x)}\left[S(x^*) + \kappa S'(x^*)\frac{\partial x^*}{\partial\kappa} - AB\left(1+q-x^*\right)^{b-a-1}\frac{\partial x^*}{\partial\kappa}\right] \\
&= \frac{S(x^*)}{\lambda S(x)T(x)}\left[-\left(1-\lambda\right) + \kappa\frac{a}{B\left(1+q-x^*\right)^{b}}\right]
\end{aligned}$$

because $\frac{\partial x^*}{\partial\kappa} = \frac{1}{\lambda B(1+q-x^*)^{b-1}}$. Thus, $\frac{\partial\beta}{\partial\kappa} > 0$ iff $\kappa a > \left(1-\lambda\right)B\left(1+q-x^*\right)^{b} = \left(1-\lambda\right)b\left(T_{Max} - \frac{\kappa}{\lambda}\right)$ iff

$$\kappa\left(\frac{a}{b\left(1-\lambda\right)} + \frac{1}{\lambda}\right) > T_{Max}$$

Further, $\frac{\partial g}{\partial\kappa} = -\frac{1}{\lambda}\frac{\partial\beta}{\partial\kappa}$, thus concluding the proof. ∎

Proof of Lemma 10. Note that in the no regulation equilibrium, firms $x < x^*$ set $g(x) = 0$, so the regulator cannot force such a firm to decrease governance. Suppose the regulator forces firm $\widetilde{x}$, where $\widetilde{x} > x^*$, to lower optimal governance from $g(\widetilde{x})$ to $g(\widetilde{x}) - \varepsilon$. The IC constraint forces $\beta(\widetilde{x})$ to increase to $\beta(\widetilde{x}) + \varepsilon\lambda$. Because the participation constraint is binding for all $x > \widetilde{x}$, so $\beta(x)T(x)S(x) = \frac{AB}{a-b}\left(\left(1+q-x\right)^{b-a} - \left(1+q-\widetilde{x}\right)^{b-a}\right) + \beta(\widetilde{x})T(\widetilde{x})S(\widetilde{x})$. Thus, the change forces firms to switch from paying $\beta(x)$ to paying $\beta(x) + \varepsilon\lambda\frac{T(\widetilde{x})S(\widetilde{x})}{T(x)S(x)}$. Thus, the governance used by firm x decreases from $g(x)$ to $g(x) - \varepsilon\frac{T(\widetilde{x})S(\widetilde{x})}{T(x)S(x)}$. Thus, the regulation changes the profit of firm x to

$$\begin{aligned}
\Pi_r(x) &= (1-\beta_r(x))\,T(x)S(x) - \kappa g_r(x)S(x) \\
&= \Pi_{nr}(x) - \varepsilon\lambda T(\widetilde{x})S(\widetilde{x})\left[1 - \frac{\kappa}{\lambda T(x)}\right] \\
&< \Pi_{nr}(x)
\end{aligned}$$

The final inequality holds because $T(x) > \frac{\kappa}{\lambda}$ for $x > x^*$. ■

Proof of Lemma 11. By similar logic to the proof of the Lemma 10, note that a decrease in governance and increase in pay in these large firms hurts their profit (by Theorem 2) and the profits of all larger firms (by tightening their participation constraint), and leaves small firms unaffected. Thus, it is optimal to allow the firms $x > x^*$ to pay as little as possible, so the endogenous participation constraint will bind for $x > x^*$ in equilibrium under optimal regulation. ■

Proof of Theorem 12. Consider the profit of firm $x < x^*$, when forced to carry out governance $g_0(x)$. Assume that the participation constraint is nonbinding, so the firm will set $\beta(x) = \lambda(1 - g_0(x))$, because they prefer to pay rather than govern by Theorem 2. Thus,

$$\begin{aligned}
\Pi(x) &= S(x)T(x) - \beta(x)S(x)T(x) - \kappa g_0(x)S(x) \\
&= (1-\lambda)S(x)T(x) - g_0(x)S(x)\left[\kappa - \lambda T(x)\right] \\
&= (1-\lambda)S(x)T(x) - g_0(x)S(x)\left[\kappa - \lambda T(x)\right]
\end{aligned}$$

Note that the profit in a small firm is strictly decreasing in required governance, $g_0(x)$.

Next, consider the endogenous participation constraint, $(\beta TS) \geq ST'$. Thus, for $x \in [q, x^*]$

$$\begin{aligned}
\int_x^{x^*} (\beta TS)'\,du &\geq \int_x^{x^*} AB\,(1+q-u)^{b-a-1}\,du \\
\beta(x^*)T(x^*)S(x^*) - \beta(x)T(x)S(x) &\geq \frac{AB}{a-b}\left((1+q-x^*)^{b-a} - (1+q-x)^{b-a}\right)
\end{aligned}$$

Further, because the firm will set $\beta(x) = \lambda(1 - g_0(x))$,

$$\lambda\left[1 - g_0(x)\right]T(x)S(x) \leq \beta(x^*)T(x^*)S(x^*) - \frac{AB}{a-b}\left[(1+q-x^*)^{b-a} - (1+q-x)^{b-a}\right]$$

Note that if this constraint is lax, the social planner could demand less governance from firm x, as well as from firms near x, by switching to the governance plan that makes the participation constraint bind, and thus improve welfare. Thus, holding $\beta(x^*)$ constant at β^*, a deviation from the governance policy that makes participation constraint bind for all x such that $g(x) > 0$ results in making firms exercise inefficiently large amounts of governance without benefit. ■

Proof of Theorem 13. Note that, by the previous theorem, $\exists x_1$ so that $\forall x > x_1$, the participation constraint binds, and for $x < x_1$, $g_r(x) = 0$. First, suppose that $x_1 > q$, so that $\beta_r(x_1) = \lambda$, and $g_r(x_1) = 0$. Thus, for $x > x_1$,

$$\begin{aligned} \beta_r(x)T(x)S(x) &= \lambda S(x_1)T(x_1) + \frac{AB}{a-b}\left[(1+q-x)^{b-a} - (1+q-x_1)^{b-a}\right] \\ g_r(x) &= 1 - \frac{\beta_r(x)}{\lambda} \end{aligned}$$

The profit of firm $x < x_1$ is given by $\Pi_r(x) = (1-\lambda)\, S(x)T(x)$, and the profit of firm $x > x_1$ is given by

$$\begin{aligned} \Pi_r(x) &= (1-\beta_r(x))S(x)T(x) - \kappa g_r(x)S(x) \\ &= S(x)\left(T(x)-\kappa\right) - \left[\lambda S(x_1)T(x_1) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-x_1)^{b-a}\right)\right]\left(1 - \frac{\kappa}{\lambda T(x)}\right) \end{aligned}$$

Given a choice of firms, x_1, let the profit of firm x be given by $\Pi(x, x_1)$. Note that, because $\Pi(x, x_1)$ is continuous in x,

$$\frac{d}{dx_1}\int_q^1 \Pi(x,x_1)dx = \int_q^1 \frac{d}{dx_1}\Pi(x,x_1)dx$$

Also, note that $\frac{d}{dx}\Pi(x, x_1) = 0$ for $x < x_1$, because they are unaffected by this change in governance. For $x > x_1$,

$$\frac{d}{dx_1}\Pi(x,x_1) = -A\,(1+q-x_1)^{-a-1}\left\{\lambda a T_{Max} - B\,(1+q-x_1)^b\left(1+\lambda\left(\frac{a}{b}-1\right)\right)\right\}\left(1-\frac{\kappa}{\lambda T(x)}\right)$$

Thus, the derivative of the objective function, is given by

$$\frac{d}{dx_1}\int_q^1 \Pi(x,x_1)dx = -A\,(1+q-x_1)^{-a-1}\left[\lambda a T_{Max} - B\,(1+q-x_1)^b\left(1+\lambda\left(\frac{a}{b}-1\right)\right)\right]\int_{x_1}^1\left(1-\frac{\kappa}{\lambda T(x)}\right)$$

Note that, by assumption, $\lambda a T_{Max} \geq B\left(1+\lambda\left(\frac{a}{b}-1\right)\right)$, and because $1+q-x_1<1$, $\lambda a T_{Max} > B\left(1+q-x_1\right)^b\left(1+\lambda\left(\frac{a}{b}-1\right)\right)$, so the first order conditions are satisfied iff

$$\int_{x_1}^{1}\left(1-\frac{\kappa}{\lambda T(x)}\right)dx=0$$

Note that, for all $x>x^*$, $T(x)>\frac{\kappa}{\lambda}$, so $1-\frac{\kappa}{\lambda T(x)}>0$, and thus $\int_{x^*}^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx>0$. Further, for all $x<x^*$, $T(x)<\frac{\kappa}{\lambda}$, so $1-\frac{\kappa}{\lambda T(x)}<0$, thus for $x<x^*$, $\int_x^1\left(1-\frac{\kappa}{\lambda T(u)}\right)du$ is strictly increasing in z. Thus, suppose that there exists x_1 such that $\int_{x_1}^1\left(1-\frac{\kappa}{\lambda T(u)}\right)du=0$. Note that, for all $x\in(x_1,1]$, $\int_x^1\left(1-\frac{\kappa}{\lambda T(u)}\right)du>0$, so $\frac{d}{dx_1}\int_q^1\Pi(x,x_1)dx<0$, and for all $x\in[q,x_1)$, $\int_x^1\left(1-\frac{\kappa}{\lambda T(u)}\right)du<0$, so $\frac{d}{dx_1}\int_q^1\Pi(x,x_1)dx>0$. Therefore, x_1 is the unique solution to the maximization problem.

To this point, I have show that such an x_1 is unique and optimal, if it exists. If $\int_q^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx<0$, such an x_1 must exist, since $\int_{x^*}^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx>0$ and $\int_z^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx$ is strictly increasing in z for $z\in[q,x^*]$. Finally, allow me to consider what happens when $\int_q^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx>0$.

When $\int_q^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx>0$, it is strictly optimal to make participation constraint bind for all $x<1$. The only variable left is $g(q)$, the required governance of firm q. For all $x>q$,

$$\begin{aligned}
\beta(x)T(x)S(x) &= \beta(q)S(q)T(q)+\frac{AB}{a-b}\left((1+q-x)^{b-a}-1\right)\\
g(x) &= 1-\frac{\beta(x)}{\lambda}
\end{aligned}$$

Thus, for all x,

$$\begin{aligned}
\Pi(x) &= S(x)\left(T(x)-\kappa\right)-\left(1-\frac{\kappa}{\lambda T(x)}\right)\left[\beta(q)S(q)T(q)+\frac{AB}{a-b}\left((1+q-x)^{b-a}-1\right)\right]\\
&= S(x)\left(T(x)-\kappa\right)-\left(1-\frac{\kappa}{\lambda T(x)}\right)\left[\lambda\left(1-g(q)\right)S(q)T(q)+\frac{AB}{a-b}\left((1+q-x)^{b-a}-1\right)\right]
\end{aligned}$$

Thus,

$$\frac{d\Pi}{dg(q)}=\left(1-\frac{\kappa}{\lambda T(x)}\right)\lambda S(q)T(q)$$

Therefore, the FOC for $g(q)$ is

$$\lambda S(q)T(q)\int_q^1\left(1-\frac{\kappa}{\lambda T(x)}\right)dx>0$$

Thus, when $\int_q^1 \left(1 - \frac{\kappa}{\lambda T(x)}\right) dx > 0$, it is optimal to set $g(q) = 1$. ■

Proof of Corollary 15. x_1 satisfies $\psi(x_1) = 0$, where $\psi(z) = \int_z^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx = 0$. Thus, by total differentiation with respect to κ at x_1,

$$\begin{aligned} -\int_{x_1}^1 \frac{1}{\lambda T(x)} dx - \left[1 - \frac{\kappa}{\lambda T(x_1)}\right] \frac{\partial x_1}{\partial \kappa} &= 0 \\ \frac{\partial x_1}{\partial \kappa} &= \frac{\int_{x_1}^1 \frac{1}{\lambda T(x)} dx}{\frac{\kappa}{\lambda T(x_1)} - 1} \end{aligned}$$

Thus, $\frac{\partial x_1}{\partial \kappa} > 0$ because $T(x_1) < \frac{\kappa}{\lambda}$. By similar logic,

$$\begin{aligned} \frac{\partial x_1}{\partial \lambda} &= -\frac{\int_{x_1}^1 \frac{\kappa}{\lambda^2 T(x)} dx}{\frac{\kappa}{\lambda T(x_1)} - 1} < 0 \\ \frac{\partial x_1}{\partial T_{Max}} &= -\frac{\int_{x_1}^1 \frac{\kappa}{\lambda [T(x)]^2} dx}{\frac{\kappa}{\lambda T(x_1)} - 1} < 0 \end{aligned}$$

Finally, note $\frac{\partial(1-x_1)}{\partial \theta} = -\frac{\partial x_1}{\partial \theta}$ for any parameter θ. ■

Proof of Theorem 18. The tax in this proof will be a flat corporate income tax, which will fund a subsidy of governance. First, notice that the flat corporate income tax will not distort decisions, because

$$\begin{aligned} &\max(1-\tau)\left[S(x)T(x)\left(1-\beta(x)\right) - \kappa(1-\delta)g(x)S(x)\right] \\ = \;&(1-\tau)\max\left[S(x)T(x)\left(1-\beta(x)\right) - \kappa(1-\delta)g(x)S(x)\right] \end{aligned}$$

as long as $\tau < 1$. Thus, for any level of income tax, the corporate income tax will collect revenue without distorting firms' behavior, so by Corollary 16, the optimal governance regulation can be implemented with a flat corporate income tax and subsidy. ■

Proof of Theorem 19. This proof is very similar to Theorem 2, and will borrow heavily from it. Because the regulator requires governance level $g(x) \geq \gamma$, if $U_0 + \kappa\gamma S(x) > S(x)T(x)$, the project is too expensive to carry out. If $\lambda(1-\gamma)S(x)T(x) \leq U_0 \leq S(x)\left(T(x) - \kappa\gamma\right)$, setting $C(y) = \frac{U_0}{ST} y$ and $g(x) = \gamma$ will satisfy the IC constraint. When $U_0 < \lambda(1-\gamma)S(x)T(x)$, by identical logic to proof of Theorem 2, the principal prefers to pay if $T(x) < \frac{\kappa}{\lambda}$ and to govern closely if $T(x) \geq \frac{\kappa}{\lambda}$. Thus, for low talented CEOs, the IC constraint will bind and the

IR will be slack, and for high talented CEOs, the firm will govern close enough until both the IC and IR constraints will bind. ■

Proof of Lemma 20. This proof is similar to the proof of Lemma 4, except that we substitute $\widetilde{\lambda} = \lambda(1-\gamma)$ into the function ϕ. Thus, consider $\phi(\cdot)$

$$\begin{aligned}
\phi(x) &= \beta(q)S(q)T(q) + \frac{AB}{a-b}\left((1+q-x)^{b-a}-1\right) - \lambda(1-\gamma)S(x)T(x) \\
&= \lambda(1-\gamma)A(T_{Max} - \frac{B}{b}) + \frac{AB}{a-b}((1+q-x)^{b-a}-1) \\
&\quad -\lambda(1-\gamma)A\left(1+q-x\right)^{-a}(T_{Max} - \frac{B}{b}\left(1+q-x\right)^{b})
\end{aligned}$$

which is the difference between the amount needed to pay CEO x to keep him from switching to firm $x-dx$, when such contagion begins at q, and the amount needed to pay the CEO to induce him to report truthfully given firm is exercising minimal governance, γ. By identical logic to Lemma 4, the results hold (substituting $\beta(x) = \lambda(1-\gamma)$ for $\beta(x) = \lambda$ when proving that line is flat between x^* and x^c). Finally, I need to prove that x^c is increasing in γ. First note that $x^c = q$ iff

$$\lambda\left(1-\gamma\right)aT_{Max} \geq B\left(1+\lambda\left(1-\gamma\right)\left(\frac{a}{b}-1\right)\right)$$

Note that as we increase γ, the left hand side goes to zero, while the right goes to B, so as we increase γ, there exists a γ_0 so that for $\gamma > \gamma_0$, the above inequality fails, and $x^c > q$. Next consider interior x^c. x^c is defined as the solution to $\phi(x^c) = 0$. Thus, by implicit function theorem,

$$\frac{dx^c}{d\gamma} = -\frac{\frac{\partial\phi}{\partial\gamma}|_{x=x^c}}{\frac{\partial\phi}{\partial x}|_{x=x^c}}$$

Note that $\frac{\partial\phi}{\partial x}|_{x=x^c} = \phi'(x^c) < 0$ because $x^c > x'$. (Recall that, when $x^c > q$, x' is the unique point such that $\phi'(x') = 0$, and, by Lemma 4, for all $x < x'$, $\phi'(x) > 0$, and for $x > x'$, $\phi'(x) < 0$) Further,

$$\begin{aligned}
\frac{\partial\phi}{\partial\gamma} &= -\lambda A\left(T_{Max} - \frac{B}{b}\right) + \lambda A\left(1+q-x\right)^{-a}\left(T_{Max} - \frac{B}{b}\left(1+q-x\right)^{b}\right) \\
&= \lambda\left[S(x)T(x) - S(q)T(q)\right]
\end{aligned}$$

Because S and T are strictly increasing functions, ST is also strictly increasing, so $S(x)T(x) > S(q)T(q)$, and thus $\frac{\partial \phi}{\partial \gamma}|_{x=x^c} > 0$, and therefore $\frac{dx^c}{d\gamma} > 0$. ■

Proof of Theorem 21. The result follows from Lemma 20 and from noting that $T(x^*) = \frac{\kappa}{\lambda}$, by definition. Everything else follows from substitution. ■

Proof of Theorem 22. β, S, T, and g are continuous, finite-valued functions on $[q, 1]$. Further, S and T are continuously differentiable, while β and g and differentiable almost everywhere (the exceptions being at x^* and x^c). Thus, by the Fundamental Theorem of Calculus, I will be continuous and differentiable for almost all x, and twice-differentiable for almost all x. Solving for the optimal floor γ^*, the FOCs are given by $I'(\gamma^*) = 0$. By the Fundamental Theorem of Calculus,

$$I'(\gamma) = \int_q^1 \frac{d}{d\gamma}\Pi_\gamma(x)dx$$

because Π_γ is continuous in x. Thus, I need to split up the cases of when $x^* \geq x^c$ and when $x^* < x^c$. First, suppose that $x^* \geq x^c$. For $x < x^c$,

$$\begin{aligned}
\Pi_\gamma(x) &= S(x)\left(T(x) - \kappa\gamma\right) - \lambda\left(1-\gamma\right)S(q)T(q) - \frac{AB}{a-b}\left((1+q-x)^{b-a} - 1\right) \\
\frac{d}{d\gamma}\Pi_\gamma(x) &= -\kappa S(x) + \lambda S(q)T(q)
\end{aligned}$$

For $x \in [x^c, x^*]$,

$$\begin{aligned}
\Pi_\gamma(x) &= (1-\lambda(1-\gamma))S(x)T(x) - \kappa\gamma S(x) \\
\frac{d}{d\gamma}\Pi_\gamma(x) &= \left[\lambda T(x) - \kappa\right]S(x)
\end{aligned}$$

For $x > x^*$,

$$\begin{aligned}
\Pi_\gamma(x) &= S(x)\left(T(x) - \kappa\right) \\
&\quad - \left[\kappa(1-\gamma)S(x^*) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right)\right]\left[1 - \frac{\kappa}{\lambda T(x)}\right] \\
\frac{d}{d\gamma}\Pi_\gamma(x) &= \kappa S(x^*)\left[1 - \frac{\kappa}{\lambda T(x)}\right]
\end{aligned}$$

Thus, for a solution γ^* such that $x^* \geq x^c$, the FOC is that $I'(\gamma) = \int_q^1 \frac{d}{d\gamma}\Pi_\gamma(x)dx = 0$. Further, $I'(\gamma) = \int_q^1 \frac{d}{d\gamma}\Pi_\gamma(x)dx = \int_q^{x^c} \frac{d}{d\gamma}\Pi_\gamma(x)dx + \int_{x^c}^{x^*} \frac{d}{d\gamma}\Pi_\gamma(x)dx + \int_{x^*}^1 \frac{d}{d\gamma}\Pi_\gamma(x)dx$. Thus,

$$\begin{aligned} I'(\gamma) &= \lambda S(q)T(q)(x^c - q) + \lambda T_{Max} \int_{x^c}^{x^*} S(x)dx \\ &\quad -\lambda \frac{AB}{b} \frac{1}{1+b-a} \left[(1+q-x^c)^{b-a+1} - (1+q-x^*)^{b-a+1} \right] \\ &\quad -\kappa \int_q^{x^*} S(x)dx + \kappa S(x^*) \int_{x^*}^1 \left\{ 1 - \frac{\kappa}{\lambda T(x)} \right\} dx \end{aligned}$$

Intuitively, the first element is the marginal benefit to the small firms of lowering the executive compensation. The next two terms are the marginal benefit to middle firms of lowering executive compensation. The fourth piece is the marginal cost of governance to the small and middle firms. Because these are not large firms, these four elements will sum to a negative number, because this change hurts these firms. The final term is the net benefit to large firms of this change. This piece is positive, because these firms benefit from this regulation. Notice also that the $I'(\cdot)$ is a function of parameters and x^c, which I proved in Lemma 19 is increasing in γ. Thus,

$$\begin{aligned} I''(\gamma) &= \lambda S(q)T(q)\frac{dx^c}{d\gamma} - \lambda T_{Max} S(x^c)\frac{dx^c}{d\gamma} + \lambda \frac{AB}{b} (1+q-x^c)^{b-a} \frac{dx^c}{d\gamma} \\ &= -\lambda \left[S(x^c)T(x^c) - S(q)T(q) \right] \frac{dx^c}{d\gamma} \end{aligned}$$

Because $S(x^c)T(x^c) > S(q)T(q)$, $\lambda > 0$, and $\frac{dx^c}{d\gamma} > 0$, $I'' < 0$, so this is locally concave on the region that $x^* < x^c < 1$. Note that when $x^c = q$, $\frac{dx^c}{d\gamma} = 0$, so $I(\cdot)$ is locally linear.

Consider when $x^c > x^*$. Then, the relevant regions are $[q, x^c]$, and $[x^c, 1]$. For $x < x^c$,

$$\begin{aligned} \Pi_\gamma(x) &= S(x)\left(T(x) - \kappa\gamma\right) - \lambda(1-\gamma)S(q)T(q) - \frac{AB}{a-b}\left((1+q-x)^{b-a} - 1 \right) \\ \frac{d}{d\gamma}\Pi_\gamma(x) &= \lambda S(q)T(q) - \kappa S(x) \end{aligned}$$

For $x > x^c$,

$$\begin{aligned} \Pi_\gamma(x) &= S(x)\left(T(x) - \kappa\right) - \left[\lambda(1-\gamma)S(q)T(q) + \frac{AB}{a-b}\left((1+q-x)^{b-a} - 1 \right) \right] \left[1 - \frac{\kappa}{\lambda T(x)} \right] \\ \frac{d}{d\gamma}\Pi_\gamma(x) &= \lambda S(q)T(q) \left[1 - \frac{\kappa}{\lambda T(x)} \right] \end{aligned}$$

Again, because $I'(\gamma) = \int_q^1 \frac{d}{d\gamma}\Pi_\gamma(x)dx = \int_q^{x^c} \frac{d}{d\gamma}\Pi_\gamma(x)dx + \int_{x^c}^1 \frac{d}{d\gamma}\Pi_\gamma(x)dx$,

$$I'(\gamma) = \int_q^{x^c} \{\lambda S(q)T(q) - \kappa S(x)\}\, dx + \lambda S(q)T(q) \int_{x^c}^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx$$

Again, the first term is the net marginal benefit to smaller firms of increasing the floor, which is negative. The second term is the net marginal benefit to large firms of increasing the floor, which is positive. For the second order conditions,

$$\begin{aligned} I''(\gamma) &= [\lambda S(q)T(q) - \kappa S(x^c)] \frac{dx^c}{d\gamma} - \lambda S(q)T(q) \left[1 - \frac{\kappa}{\lambda T(x^c)}\right] \frac{dx^c}{d\gamma} \\ &= -\frac{\kappa}{T(x^c)} \left[S(x^c)\, T(x^c) - S(q)T(q)\right] \frac{dx^c}{d\gamma} \end{aligned}$$

Because ST is increasing, $I''(\gamma) < 0$, and thus I is locally concave on $x^c > x^*$. Finally, note that for $x^c = 1$, I is locally linear, because $\frac{dx^c}{d\gamma} = 0$.

To complete the proof, allow me to define γ_0, γ_1, and γ_2. Define γ_0 as the floor of governance such that for all $\gamma < \gamma_0$, $x^c(\gamma) = q$, and for $\gamma > \gamma_0$, $x^c(\gamma) > q$. Define γ_1 such that for $x^c(\gamma_1) = x^*$. (Recall that x^* does not depend on γ, because $T(x^*) = \frac{\kappa}{\lambda}$) Finally, define γ_2 such that for all $\gamma < \gamma_2$, $x^c(\gamma) < 1$, and for all $\gamma > \gamma_2$, $x^c(\gamma) = 1$. These governance levels, γ_0, γ_1, and γ_2 are unique by Lemma 19. Further, by above section of this proof, I have shown that I is linear on $[0, \gamma_0)$, strictly concave on (γ_0, γ_1), strictly concave on (γ_1, γ_2), and linear on $(\gamma_2, 1]$. Thus, to prove global concavity, all I need to show is that

$$I'_+(\gamma_1) \leq I'_-(\gamma_1)$$

because then $I'(\cdot)$ will be a weakly decreasing function. Recall that $I'_-(\gamma_1) = \lim_{\gamma \to \gamma_1^-} I'(\gamma)$, and $I'_+(\gamma_1) = \lim_{\gamma \to \gamma_1^+} I'(\gamma)$. Note that $\lim_{\gamma \to \gamma_1} x^c(\gamma) = x^*$. Thus,

$$\begin{aligned} I'_-(\gamma_1) &= \lambda S(q)T(q)(x^* - q) - \kappa \int_q^{x^*} S(x)dx + \kappa S(x^*) \int_{x^*}^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx \\ I'_+(\gamma_1) &= \int_q^{x^*} \{\lambda S(q)T(q) - \kappa S(x)\}\, dx + \lambda S(q)T(q) \int_{x^*}^1 \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx \end{aligned}$$

Therefore, I will be globally concave iff $I'_-(\gamma_1) - I'_+(\gamma_1) \geq 0$.

$$\begin{aligned} I'_-(\gamma_1) - I'_+(\gamma_1) &= [\kappa S(x^*) - \lambda S(q)T(q)] \int_{x^*}^{1} \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx \\ &= \lambda [S(x^*)T(x^*) - S(q)T(q)] \int_{x^*}^{1} \left\{1 - \frac{\kappa}{\lambda T(x)}\right\} dx \end{aligned}$$

The second line holds because $T(x^*) = \frac{\kappa}{\lambda}$. Thus, because $\lambda > 0$, $S(x^*)T(x^*) > S(q)T(q)$, and for all $x > x^*$, $T(x) > \frac{\kappa}{\lambda}$, so $\left\{1 - \frac{\kappa}{\lambda T(x)}\right\} > 0$, $I'_-(\gamma_1) - I'_+(\gamma_1) > 0$, and thus I will be globally concave.

Therefore, without loss of generality, $\gamma^* \in \{0\} \cup (\gamma_0, \gamma_2] \cup \{1\}$. Thus, whenever the regulator imposes a governance floor $\gamma^* > 0$, $x^c > q$ and thus inefficient governance is being carried out at the smallest firms. ■

Proof of Corollary 24. Recall that, for $x > x^*$, in the competitive outcome,

$$\begin{aligned} \beta(x)S(x)T(x) &= \beta(x^*) S(x^*) T(x^*) \\ &\quad + \frac{AB}{a-b}\left[(1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right] \\ &= \kappa A(1+q-x^*)^{-a} \\ &\quad + \frac{AB}{a-b}\left[(1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right] \end{aligned}$$

Thus, dividing both sides by $S(x)$,

$$\begin{aligned} \beta(x)T(x) &= \kappa \left(\frac{1+q-x^*}{1+q-x}\right)^{-a} \\ &\quad + \frac{B}{a-b}\left[(1+q-x)^{b} - (1+q-x^*)^{b}\left(\frac{1+q-x^*}{1+q-x}\right)^{-a}\right] \end{aligned}$$

Thus, differentiating both sides with respect to a,

$$\begin{aligned}
\frac{\partial \beta(x) T(x)}{\partial a} &= \kappa \left(\frac{1+q-x^*}{1+q-x}\right)^{-a} \left[-\log\left(\frac{1+q-x^*}{1+q-x}\right)\right] \\
&\quad -\frac{B}{(a-b)^2}\left[(1+q-x)^b - (1+q-x^*)^b \left(\frac{1+q-x^*}{1+q-x}\right)^{-a}\right] \\
&\quad +\frac{B}{a-b}\left(-(1+q-x^*)^b\right)\left(\frac{1+q-x^*}{1+q-x}\right)^{-a}\left[-\log\left(\frac{1+q-x^*}{1+q-x}\right)\right] \\
&= \kappa \left(\frac{1+q-x^*}{1+q-x}\right)^{-a} \left[-\log\left(\frac{1+q-x^*}{1+q-x}\right)\right] \\
&\quad -\frac{B(1+q-x)^a}{(a-b)^2}\left[(1+q-x)^{b-a} - (1+q-x^*)^{b-a}\left(1+(a-b)\log\left(\frac{1+q-x^*}{1+q-x}\right)\right)\right]
\end{aligned}$$

Note that because $x > x^*$, $1+q-x < 1+q-x^*$, so $\log\left(\frac{1+q-x^*}{1+q-x}\right) > 0$. Also, note that $T(x) = T_{Max} - \frac{B}{b}(1+q-x)^b$, so $\frac{\partial T}{\partial a} = 0$. Define $l(x)$ for $x > x^*$ so that

$$l(x) = (1+q-x)^{b-a} - (1+q-x^*)^{b-a}\left(1+(a-b)\log\left(\frac{1+q-x^*}{1+q-x}\right)\right)$$

Thus, if $l(x) > 0$, then $\frac{\partial \beta}{\partial a} < 0$. Note that $l(x^*) = 0$, so if $l' > 0$, the proof is complete.

$$\begin{aligned}
l'(x) &= (a-b)(1+q-x)^{b-a-1} - (1+q-x^*)^{b-a}(a-b)\left(\frac{1+q-x}{1+q-x^*}\right)\left(-\frac{1+q-x^*}{(1+q-x)^2}\right)(-1) \\
&= (a-b)\frac{1}{1+q-x}\left[(1+q-x)^{b-a} - (1+q-x^*)^{b-a}\right]
\end{aligned}$$

Recall that $f(x) = (1+q-x)^{b-a}$ is strictly increasing in x for $b < a$, so $l'(x) > 0$ for all $x > x^*$, so $l(x) > 0$ for all $x > x^*$. Therefore,

$$T(x)\frac{\partial \beta}{\partial a} = -\kappa\left(\frac{1+q-x^*}{1+q-x}\right)^{-a}\log\left(\frac{1+q-x^*}{1+q-x}\right) - \frac{B}{(a-b)^2}\frac{1}{(1+q-x)^{-a}}l(x)$$

for all $x > x^*$. Therefore, $\frac{\partial \beta}{\partial a} < 0$ for all $x > x^*$.

Finally, noting that, for $x > x^*$, $g(x) = 1 - \frac{\beta(x)}{\lambda}$, the proof is concluded. ■