

Incentives, Supervision, and Limits to Firm Size

Ayush Gupta*

28 October 2026

[Please click here for the latest version](#)

Abstract

This paper studies whether incentive problems alone can limit the size of a firm. We develop a general model in which effort, supervision, and wages are endogenous. We establish a sufficient condition and a necessary condition for a firm to grow without bound. Our results show that incentive problems can limit firm size even without additional frictions such as costly communication or incomplete contracts. In particular, when the production function is not linear, the optimal size may be positive but finite. The proof of the necessary condition may be of independent interest for dynamic programming problems without discounting.

*University of Mannheim, Department of Economics; ayush.gupta@uni-mannheim.de

I am indebted to Dilip Mookherjee and Andy Newman for many helpful discussions and suggestions. I would also like to thank seminar participants at Boston University, Boston College, IOEA 2024, the Midwest Economic Theory Conference, and the Stony Brook Conference on Game Theory for helpful comments. Support by the German Research Foundation (DFG) through CRC TR 224 (Project B03) is gratefully acknowledged.

1 Introduction

It is often convenient to model a firm as a single agent that maximizes profits. However, firms are not monolithic entities; they are made up of individuals with diverse objectives. The workers' objectives often conflict with the owner's, and this creates an incentive problem. Established results tell us that incentive problems lead to workers choosing actions that are not optimal from the owner's perspective.¹ This is often called "loss of control" in the literature. Williamson (1967) showed that, if the loss of control is increasing in the number of workers, incentive problems alone will limit the number of workers that can be hired by a firm. This paper studies whether – and under what conditions – incentive problems result in increasing loss of control.

Our object of study is any contractual arrangement with one principal and many agents. While we present our results in the context of a firm, our model describes many other types of organizations as well.² We develop a general model in which worker effort, supervision, and wages are determined endogenously. The internal organization of the firm is also endogenous in our model. In particular, we allow firms to adopt a hierarchical structure: the owner supervises a subset of workers who, in turn, supervise other workers and so on. Our main results establish a sufficient condition and a necessary condition for a firm to grow infinitely large. We use these conditions to identify settings where loss of control is increasing and incentive problems alone limit firm size. As an illustration, we revisit the model of Calvo and Wellisz (1978) (hereafter "CW78") in which incentive problems do not limit the size of a firm. We show that their conclusion depends on the assumption of a linear production function. When the production function is concave, the optimal number of workers may be positive but finite. Proving our necessary condition requires a novel optimization technique, which may be of independent interest.

The existing literature has produced results on both sides of the question of whether incentive problems limit the size of a firm. CW78 showed that a firm can grow infinitely large as long as it is profitable to hire a single worker, and subsequent work supports this conclusion in more general settings. Papers which obtain a finite firm size do so under additional assumptions. For

¹See Mirrlees (1976); Shavell (1979); Holmström (1979); Grossman and Hart (1983).

²Conversely, not every organization that meets the legal definition of a firm will satisfy our assumptions.

example, a production function in which the effort of supervisors enters multiplicatively, or costly communication within the firm.³ It has therefore remained unclear whether incentive problems can limit firm size in general, or only in specific environments. Our results address this question directly. They show that incentive problems can limit firm size even in the absence of additional frictions such as incomplete contracts or communication costs. In settings where firm size is limited, we solve for an explicit bound on the optimal number of workers.

Our sufficient condition (Theorem 1) says that incentive problems do not limit firm size if a worker can produce more than his expected wage, while supplying as much supervision as he receives. That is because such a worker can supervise a subordinate who receives the same level of supervision, produces more than his expected wage, and supervises a subordinate of his own. The firm can thus add any number of workers in this manner and each additional worker increases the firm's total profit. Our necessary condition (Theorem 2) is the same requirement for an arbitrary group of workers. Intuitively, it is not sufficient for workers to produce more than their expected wage. Workers must also compensate the firm for the supervision they receive. Otherwise, incentive problems limit firm size because the owner has a finite allocation of time that can be spent on supervision and supervision is necessary for production.

Our results require two key assumptions. First, we assume that effort is costly so workers shirk on the job unless supervised by their manager. Second, we assume that supervision requires time, of which the owner has a finite endowment. Beyond these two assumptions, we impose minimal restrictions on the production function, the monitoring technology, or the type of employment contracts. However, we note that our analysis focuses on one side of the firm size trade-off. To that end, we take as given the benefits of organizing production within a firm. We also note that our model does not speak to the firm versus market distinction. Our results can help determine the optimal size of a firm, but not its scope.

The main technical contribution of this paper is the proof of our necessary condition. The firm's optimization problem can be set up as an infinite dynamic programming problem. There is no discounting, so the problem cannot be solved using the standard Bellman equation approach because

³Section 1.1 provides a more detailed overview of the existing literature.

the Bellman operator is not a contraction mapping. Instead, we apply the Kuhn-Tucker theorem in a novel way to bound the profit generated by each individual worker. The bound is uniform: it does not depend on the number of workers or on the structure of the firm. We then use this bound to prove our necessary condition. To our knowledge, this application of the Kuhn-Tucker theorem is novel in the literature. Moreover, it does not rely on the specifics of our model and can be used to study a general class of optimization problems.

Outside the firm setting, our assumptions require that enforcing a contract costs time and the owner has a finite endowment of it. She can contract with agents directly or through sub-contractors, sub-sub-contractors and so on. But, if our necessary condition is violated, incentive problems alone limit how many workers can be contracted with.

The remainder of this section reviews the related literature. Section 2 presents the CW78 framework and states our results in that setting. Section 3 uses a simple example to illustrate that, even in this simplified setting, incentive problems may limit firm size. Section 4 describes our general model and states our main results. Section 5 concludes.

1.1 Related Literature

This paper contributes to a very large literature initiated by Coase (1937). Coase proposed that firms exist to avoid transaction costs but cannot grow arbitrarily large because of decreasing returns to the entrepreneur function. Since then, both the costs and the benefits of a firm have been studied extensively. This paper focuses on the former.

The literature on incentive problems as a limit to firm size dates back to at least CW78, who endogenized loss of control using a model of supervision and wage incentives. They demonstrated that, in general, incentive problems do not limit the size of a firm: a firm can grow infinitely large, as long as it is profitable to hire a single worker. In the CW78 model, firm size is limited only if workers can identify the times at which their performance is being monitored. Subsequent papers have largely supported the CW78 result. For example, both Datta (1996) and Tsumagari (1999) showed that firm size remains unbounded even when supervisory effort is not directly observable.

Papers which obtain a finite firm size rely on additional assumptions. Qian (1994) assumed a production function in which the output produced by a worker is a multiplicative function of his own effort and the effort exerted by all of his (direct and indirect) supervisors. Unless all workers exert full effort, this assumption limits firm size by causing productivity to decline as the firm adds more hierarchical layers. Similarly, Keren and Levhari (1983) showed that the optimal size of a firm is finite but assumed that communication within a firm is costly. Relative to this literature, we provide a sufficient condition and a necessary condition which do not depend on the choice of production function or monitoring technology.

A related literature focuses on incomplete contracts. Grossman and Hart (1986) and Hart and Moore (1990) showed that incomplete contracts create hold-up problems which limit firm size because ownership needs to be aligned with the incentive to invest. Our paper does not assume incomplete contracts. Instead, we assume that even complete contracts are costly to enforce because of incentive problems. In the limit case, the cost can be interpreted as the time needed to verify that the agent has fulfilled his obligations under the contract. This cost may be very small but we argue that it will never be zero. For example, if an employee is paid a piece rate, the owner will need to verify the quantity and quality of the goods produced. Doing so might be as simple as counting the number of units produced, but it will always require some positive amount of time.

Another related literature assumes that communication within a firm is costly. Radner (1993), Bolton and Dewatripont (1994), Garicano (2000), and many others have studied how costly communication affects the internal organization of a firm and shown how it may limit firm size. We consider these works to be complementary to our own. We assume that there are no communication costs only because we want to isolate the effects of incentive problems. Our results show that, even as communication costs approach zero, the optimal size of a firm may be constrained by incentive problems.

This paper also contributes to the broader literature on incentives within firms and organizations. Foundational work in agency theory shows how moral hazard and imperfect observability shape optimal contracts (Mirrlees, 1976; Holmström, 1979; Shavell, 1979; Grossman and Hart, 1983). Subsequent work studies how firms use wages, performance pay, monitoring, career concerns, and

job design to mitigate these incentive problems. Fama (1980) studies agency problems within firms, while Holmström and Milgrom (1991) and Holmström and Milgrom (1994) show how multitasking, asset ownership, and job design jointly determine the optimal strength of incentives. Prendergast (1999) provides an overview of this literature. Our paper builds on the same central premise: because workers' objectives differ from the firm's, firms must spend resources inducing workers to take actions which are valuable to the organization.

A large empirical literature shows that incentives have economically meaningful effects inside firms. Lazear (1986), Gibbons (1987), Lazear (2000), and Shearer (2004) study piece rates and performance pay. Nagin et al. (2002) studies the interaction between monitoring and opportunistic behavior in a field experiment. Bandiera et al. (2005), Bandiera et al. (2007), and Bandiera et al. (2009) show that social ties, managerial incentives, and workplace relationships affect the response to incentive schemes. Related papers study team production and group incentives, including Hamilton et al. (2003), Boning et al. (2007), and Friebe et al. (2017). These papers demonstrate that incentive problems are quantitatively important within firms. Our contribution is different but complementary. Instead of studying how a firm should design incentives, we ask how incentive problems constrain firms when incentives are optimally designed.

Our paper is also related to work on internal labor markets, hierarchy, and managerial incentives. Malcomson (1984) studies the role of work incentives in hierarchical labor markets. Baker et al. (1994a) and Baker et al. (1994b) use personnel data to study wage policy and promotion dynamics inside a firm, while Gibbons and Waldman (1999) and Gibbons and Waldman (2006) develop a theory of wage and promotion dynamics inside firms. Bull et al. (1987) provides experimental evidence on tournaments and piece rates. At the managerial and executive level, Jensen and Murphy (1990), Aggarwal and Samwick (2003), and Oyer and Schaefer (2005) study incentive compensation and performance incentives within firms. This literature treats the firm as an organization composed of many agents whose incentives must be managed. Our paper contributes by showing that these internal incentive problems can impose an endogenous upper bound on firm size.

A separate but related literature studies how organizational structure affects incentives, authority, and information flows. Aghion and Tirole (1997) distinguish between formal and real authority,

Dessein (2002) studies the trade-off between authority and communication, and Alonso et al. (2008) analyzes when coordination requires centralization. Empirically, Rajan and Wulf (2006) document the flattening of corporate hierarchies, while Bloom and Van Reenen (2007) and Bloom et al. (2013) show that management practices vary substantially across firms and have important effects on productivity. These papers emphasize that the internal organization of a firm matters. We share this perspective, but isolate a different mechanism: even when communication costs, coordination problems, and incomplete contracts are absent, costly supervision can make incentive problems increasingly severe as the organization grows.

2 Simplified Model

This section describes the CW78 model to provide intuition. Our general model relaxes many of the assumptions made here and can be applied to a much larger class of problems.

The object of study is a hierarchical contractual arrangement with one principal and many agents. To remain consistent with the existing literature, we call this arrangement a firm. However, as noted in the introduction, our assumptions are not equivalent to the legal definition of a firm: our model describes many non-firm organizations as well, but not every firm satisfies our assumptions.

We will assume that the solitary principal (hereafter “owner”) wants to maximize profit. She can operate alone or hire from an infinite pool of identical agents (hereafter “workers”). The value of the output produced by each worker is given by $f(p)$ where p is the worker’s level of productive effort. We assume that $f(p)$ is the value net of all costs except for the worker’s wage. We also assume:

1. $f(p)$ is continuous, increasing, and bounded.
2. $f(0) = 0$.
3. $\sup_{p \in (0,1]} \frac{f(p)}{p} < \infty$.

We are assuming that a worker’s production depends only on his own level of productive effort.

This means that there is no complementarity in worker effort. Furthermore, the firm's total output is the sum of the output produced by each worker and the owner. We make these assumptions deliberately because they ensure that, in the absence of incentive problems, the firm's production is linear in the number of workers. We thus rule out every other source of decreasing returns and ensure that any limit to firm size in our model is a consequence of incentive problems alone.

Our assumptions about the production process can be read as a tractable approximation of settings in which tasks are modular enough that one can meaningfully define a focal unit of work, even if completing that unit of work involves review, support, escalation, or handoffs by several team members. Natural examples include consulting and professional services, law firms, customer support and operations, claims processing or case management, many modular software or IT teams, software incident response, radiology or other diagnostic reading pools, and other environments in which cases or sub-tasks may be allocated across members.

Note that CW78 assumed $f(p) = \beta p$ where $\beta > 0$ is fixed. It is clear that our assumptions allow for a much larger class of production functions. Our first and second assumptions are ubiquitous in the literature but our third assumption merits special attention. It requires that production per unit of effort is bounded above. This assumption ensures that the firm's total production is finite whenever the total level of effort exerted by all workers is finite. Without this assumption, the firm could produce any level of output by dividing a fixed level of productive effort across an increasingly large number of workers.

The von Neumann-Morgenstern utility index of each worker is given by $u(w, e)$ where $w \geq 0$ is the wage received by the worker and $e \in [0, 1]$ is the level of total effort exerted by him. We interpret effort as the fraction of the day spent working rather than shirking. We assume:

1. $u(w, e)$ is continuous.
2. $u(w, e)$ is increasing in w .
3. $u(w, e)$ is decreasing in e .
4. The outside option gives zero utility.

It is obvious that a worker who is offered a fixed wage will maximize his utility by exerting zero effort.⁴ The owner can resolve this incentive problem by offering employment contracts that are contingent on effort. But such a contract will not be effective if the owner never observes effort. Observing effort requires supervision which is both costly and imperfect.

CW78 assumed that, if $s \in [0, 1]$ is the level of supervision received by a worker, then s is both the fraction of the day that the owner spends supervising him and also the probability that his effort is observed perfectly. Intuitively, the owner knows more about a worker's level of effort when she spends more time supervising him. This section makes the same assumption but our general model allows for other forms of supervision.

The trade-off faced by the owner is that time spent on supervising a worker cannot be spent on other tasks. So one can think of supervision as being unproductive to the extent that it does not (directly) produce output. However, the owner can instruct a worker to supervise other workers, allowing for hierarchical firms with managers.

The CW78 model assumes that the owner offers a wage $\bar{w}$. The worker receives $\bar{w}$ if his effort level is not observed. If the worker's effort is observed, he receives $e\bar{w}$ where e is his level of effort. We maintain the same assumption throughout this section but Section 4 allows for other employment contracts.

The worker then exerts a level of effort that maximizes his expected utility:

$$U = \max_e su(e\bar{w}, e) + (1 - s)u(\bar{w}, e)$$

Let $\mathbf{E}$ be the set of maximizers. In general, $\mathbf{E}$ need not be a singleton but it will be a compact set. For the remainder of this paper, we will assume that the worker's chosen level of effort is $e = \max(\mathbf{E})$.⁵ This assumption allows us to define a function $e(\bar{w}, s)$ which gives the worker's choice of effort, conditional on $\bar{w}$ and s .

Now define $\bar{w}(e, s)$ to be the lowest wage that the owner must offer to induce effort e when the level

⁴A wage is said to be fixed if it does not depend on the worker's level of effort.

⁵We consider this to be a conservative assumption. If we allow the agent to pick a lower level of effort, it becomes *more* likely that incentive problems will limit firm size.

of supervision is s . When e cannot be induced, we define $\bar{w}(e, s) = \infty$. Then the worker's expected wage is:

$$w(e, s) = se\bar{w}(e, s) + (1 - s)\bar{w}(e, s)$$

We call $w(e, s)$ the expected wage function.

A worker can split his effort between production (productive effort) and supervising other workers (supervisory effort). We assume that effort is infinitely divisible so that $e = p + s'$ where p is the level of productive effort and s' is the level of supervisory effort. Note that s' can be split across multiple subordinates. We also assume that, conditional on the level of effort, workers are indifferent between productive work and supervisory work.

In this model all organizational decisions are made by the owner. The owner decides the size and structure of the firm i.e. how many workers to hire and who supervises whom. The owner also decides how much effort to induce from each worker and how to split that effort between productive work and supervisory work. The managers serve only to verify the level of effort exerted by their subordinates. In particular, we are assuming that managers cannot misreport what they observe so there is no possibility of collusion between workers and managers. The interaction is one-shot: the owner announces the structure of the firm, the level of supervision that each worker will receive, and the wage offered to each worker; workers accept or reject; those who accept choose their effort; supervision takes place; and wages are paid.

Before moving on to the results, it is helpful to introduce an example. Figure 1 shows the organizational structure of a hypothetical firm with twelve employees. In this firm, the owner directly supervises three employees (labelled "Managers" in Figure 1). In turn, each manager supervises three employees (labelled "Rank and File" in Figure 1). The rank and file workers do not supervise anyone so all of their effort is necessarily productive effort. The owner and the three managers may split their effort between productive and supervisory work in any ratio.

Note that we use the words worker and employee interchangeably to refer to any agent who is not the owner of the firm. The words manager and supervisor are used interchangeably to refer to any

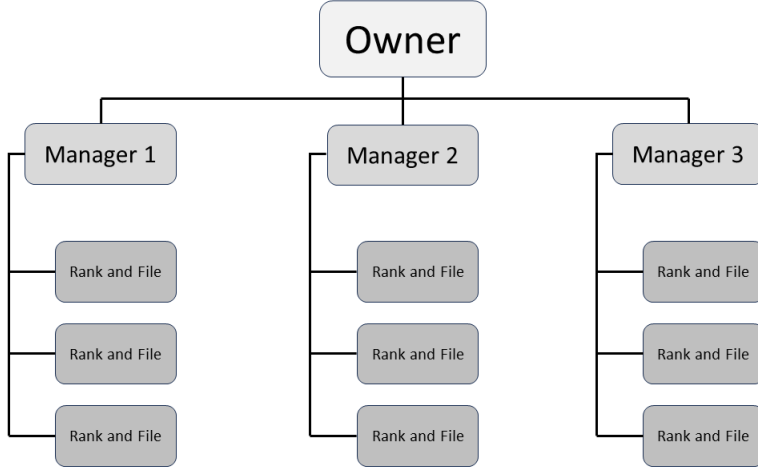


Figure 1: Example Firm Hierarchy

agent who supervises another agent.

For simplicity we will assume that the owner's production function is also $f(p)$ and that she always exerts $e = 1$ effort.⁶ She will produce $f(1)$ if she operates alone. If the owner hires exactly one worker, the firm will produce:

$$f(1 - s) + f(e) - w(e, s)$$

where the owner produces $f(1 - s)$ and the worker produces $f(e)$.⁷ The owner will hire the worker if there exists some $e^* \in [0, 1]$ and $s^* \in [0, 1]$ such that:

$$f(1) < f(1 - s^*) + f(e^*) - w(e^*, s^*)$$

$$f(1) - f(1 - s^*) < f(e^*) - w(e^*, s^*)$$

This observation provides some useful intuition. $f(e^*) - w(e^*, s^*)$ is the net profit generated by the worker. $f(1) - f(1 - s^*)$ is the reduction in the owner's production and can be interpreted as the opportunity cost of hiring and supervising a worker. This cost is a consequence of the incentive

⁶None of our results depend on the level of effort exerted by the owner. In particular, our results are valid even when there is more than one owner and they collectively exert more than one unit of effort.

⁷There is no one for the worker to supervise so all of his effort is productive effort.

problem which necessitates supervision.

Because of the incentive problem, it is not enough for the net profit generated by a worker to be positive. The owner will not hire the worker unless he generates enough profit to compensate for the opportunity cost of supervising him. As a result, there will be some settings in which the net profit generated by a worker is positive but he is not hired because the cost of supervising him is too high.

We now define $\delta = s' - s$ to be the *excess supervision* supplied by a worker. When δ is positive, the level of supervision received by a worker is less than the level of supervision he provides to his subordinates. The opposite is true when δ is negative.

Our results show that the effect of incentive problems depends on the relationship between the level of excess supervision supplied by a worker and the net profit generated by him.

Definition 1.

$$\begin{aligned}
 V(\delta) &= \sup_{e, s, s'} f(e - s') - w(e, s) \\
 &\text{subject to } \delta = s' - s \\
 &e \in (0, 1] \\
 &s' \in [0, e] \\
 &s \in [0, 1]
 \end{aligned} \tag{1}$$

Notice that $V(\delta)$ is the supremum of the profit a worker can generate when the excess supervision supplied by him is exactly δ . The value of $V(0)$ is central to the analysis that follows. It is the highest profit that can be generated by a worker who supplies exactly as much supervision as he receives.

We say that incentive problems limit the size of a firm when the firm's total profit is bounded above. Our concern is therefore with limits to profitable expansion, rather than limits on the number of workers a firm can employ. The distinction is relevant because it may be possible for the firm to hire an infinite number of workers who exert zero effort, receive no supervision, and are paid nothing. This caveat applies to all previous statements about firm size.

Theorem 1. Incentive problems do *not* limit firm size if $V(0) > 0$.

First note that $V(0) > 0$ if and only if there exists some $e^* \in [0, 1]$ and $s^* \in [0, e^*]$ such that:

$$f(e^* - s^*) > w(e^*, s^*) \quad (2)$$

The proof of Theorem 1 utilizes a replication argument. Suppose there exists (e^*, s^*) such that $f(e^* - s^*) > w(e^*, s^*)$. Then a worker can generate positive profit while receiving s^* supervision and exerting s^* effort on supervising a subordinate. But then the subordinate can replicate his supervisor: he is receiving s^* supervision so he can also generate positive profit while exerting s^* effort on supervising his own subordinate. This replication process can be continued *ad infinitum* to create an infinitely long branch in which every worker is supervising exactly one subordinate and generating $f(e^* - s^*) - w(e^*, s^*) > 0$ profit.

The intuition is that, because each worker supplies as much supervision as he receives, the net supervisory cost of each worker is zero. As a result, hiring an additional worker does not require any existing member of the firm to supply more supervision. It is therefore possible to hire an infinite number of workers who are all acting identically. Importantly, it is *not* sufficient for a worker to produce more than his wage. He also needs to compensate the firm for the supervision he receives.

An alternate interpretation comes from rewriting Equation (2) as follows:

$$f(e^*) - w(e^*, s^*) > f(e^*) - f(e^* - s^*)$$

The left hand side of this inequality is the net profit generated by a worker when he exerts zero supervisory effort and receives s^* excess supervision. The right hand side is the reduction in that worker's production when he supplies s^* supervision and receives zero excess supervision. The right hand side can therefore be interpreted as the "value" of the excess supervision received by the worker. Theorem 1 then says that incentive problems do not limit firm size if workers can compensate the firm for the value of the excess supervision received by them.

Note that our sufficient condition reduces to the CW78 result when f is linear. However, in general, the two conditions are not equivalent. The distinction is illustrated by an example in the next section.

We now define $\bar{V}(\delta)$ to be the *concave envelope* of $V(\delta)$ i.e. the smallest concave function such that $\bar{V}(\delta) \geq V(\delta)$ for every δ . Concretely, we define $\bar{V}(\delta)$ to be the supremum of $\alpha V(\delta_1) + (1 - \alpha)V(\delta_2)$ over all $\alpha \in [0, 1]$ and all δ_1, δ_2 with $\alpha\delta_1 + (1 - \alpha)\delta_2 = \delta$.⁸

$\bar{V}(\delta)$ is the highest profit per worker that can be generated by a *group* of workers who supply exactly δ units of excess supervision on average. Compare this to $V(\delta)$ which represents the same quantity for a single worker.⁹ The remainder of our results focus on $\bar{V}(\delta)$ because, in general, a group of workers specializing in different tasks may be more productive than any individual worker. But recall that our model has two tasks: production and supervision. So it is sufficient to focus on groups with at most two types of workers: workers who exert more productive effort and workers who exert more supervisory effort.

Lemma 1. $\bar{V}(0) \geq \bar{V}(\delta)$ for any $\delta \geq 0$.

Theorem 2. Incentive problems do not limit the size of a firm only if $\bar{V}(0) > 0$.

First note that $\bar{V}(0)$ is the highest profit per worker that can be generated by a group that supplies exactly as much supervision as it receives. $\bar{V}(0) < 0$ means that such a group *cannot* generate positive profit. Lemma 1 implies that if $\bar{V}(0) < 0$, any group of workers must be a net recipient of supervision in order to be profitable.

Theorem 2 shows that incentive problems limit the size of a firm when $\bar{V}(0) < 0$. That is because the total excess supervision supplied by all members of the firm (including the owner) sums to zero. One might expect that by adding vertical layers, the owner can increase the supply of supervision but that is an illusion. Any worker who supplies supervision must himself be supervised. The owner is the only member of the firm who supplies supervision without receiving any herself. But

⁸This is the same construction as the concavification used in the literature on Bayesian persuasion (Kamenica and Gentzkow, 2011). There the concave envelope is taken over posterior beliefs; here it is taken over levels of excess supervision.

⁹ $\bar{V}(\delta) = V(\delta)$ whenever $V(\delta)$ is concave.

the owner has a finite endowment of time. As the firm grows, the excess supervision supplied by the owner is spread across more and more workers. In the limit, the average worker receives zero excess supervision and therefore generates no more than $\bar{V}(0)$. We show that, even for a small firm, the difference between the firm's profit and $n\bar{V}(0)$ is bounded.¹⁰ This bound is then used to prove Theorem 2.

Theorem 2 can also be interpreted similarly to Theorem 1. It is not sufficient for workers to produce more than what they are paid. They must be productive enough for a group of workers to compensate for the *value* of excess supervision received by them. Otherwise, incentive problems will limit firm size.

From a technical standpoint, the owner faces a dynamic programming problem. She chooses how many workers to supervise and the level of supervision for each subordinate. The same decision must then be made for each subordinate, for the workers they supervise, and so on. The challenge is that there is no discounting. As a result, this problem cannot be solved using standard techniques because the Bellman operator is not a contraction. Instead, we use the Kuhn-Tucker theorem to show that the profit per worker generated by any group is bounded above by a linear function of the level of excess supervision received by the group as a whole. Applying this bound to the entire firm shows that, as the firm hires more workers, the bound on profit per worker shrinks to $\bar{V}(0)$. This allows us to prove that the total profit generated by the firm is bounded when $\bar{V}(0) < 0$. To our knowledge, this application of the Kuhn-Tucker theorem is novel in the literature. Moreover, it does not rely on the specifics of our model and can be applied to a general class of optimization problems.

We stress that the condition in Theorem 2 is necessary for a firm to grow infinitely large. It may be profitable for a firm to hire a finite but positive number of workers even when the necessary condition is not satisfied. The next section provides an explicit example of such a firm. We also use the proof of Theorem 2 to obtain a bound on the optimal size of a firm that is constrained by incentive problems.

Corollary 1. If $\bar{V}(0) < 0$ and $V(\delta) > -\infty$ for some $\delta > 0$, the optimal size of a firm is at most

¹⁰ n is the total number of workers.

N^* where N^* is the largest integer such that $\bar{V}(\frac{-s}{N^*}) \geq 0$ for some $s \in [0, 1]$.

It is worth emphasizing that this bound holds for every finite firm regardless of how that firm is organized. The workers can be arranged in any number of layers, with any span of control, and they can differ in their effort, their wages, and the supervision they receive.

3 Example

The significance of our results can be illustrated with a simple example. Consider:

$$\begin{aligned} f(p) &= 4p - 2p^2 \\ u(w, e) &= w - \frac{e^2}{2} \end{aligned}$$

Assume that the monitoring technology and the employment contracts are as described in the previous section. Then:

$$\begin{aligned} e(\bar{w}, s) &= \min \{ s\bar{w}, 1 \} \\ \bar{w}(e, s) &= \frac{e}{s} \\ w(e, s) &= e^2 - e + \frac{e}{s} \end{aligned}$$

We stress that the expected wage function depends on the choice of employment contract. This example restricts attention to a specific contract only to remain consistent with the CW78 model.

The owner operating alone will produce $f(1) = 2$ units of profit. She will hire at least one worker if there exist some $e^* \in [0, 1]$ and $s^* \in [0, 1]$ such that:

$$f(1 - s^*) + f(e^*) - w(e^*, s^*) > 2$$

It is easily verified that many such (e^*, s^*) exist. One example is $e^* = s^* = \frac{1}{2}$. So the owner will hire at least one worker. Next note that:

$$f(p) = 4p - 2p^2 \leq 4p$$

$$w(e, s) = e^2 - e + \frac{e}{s} \geq 4(e - s)$$

Now consider any group of workers such that the total excess supervision supplied by them is exactly zero. Use i to index the workers in this group. Then:

$$\begin{aligned} \sum_i p_i &= \sum_i [e_i - s'_i] \\ &= \sum_i e_i - \sum_i s_i \end{aligned}$$

where the second equality comes from the fact that $\sum_i s'_i = \sum_i s_i$ because the group as a whole supplies zero excess supervision. Therefore the total net profit generated by the group is:

$$\begin{aligned} \sum_i [f(p_i) - w(e_i, s_i)] &= \sum_i [4p_i - 2p_i^2] - \sum_i [e_i^2 - e_i + \frac{e_i}{s_i}] \\ &\leq 4 \sum_i p_i - 4 \sum_i [e_i - s_i] \\ &= 0 \end{aligned}$$

So no group of workers can generate positive net profit while supplying at least as much supervision as they receive.

We have thus shown that $\bar{V}(0) \leq 0$ and our necessary condition is not satisfied. Theorem 2 then says that incentive problems *will* limit the size of this firm and there is a limit to how much profit the firm can generate by hiring more workers.

Notice that we constructed our example to be consistent with the CW78 model *except* for the fact that we did not assume f to be linear. We then showed that it is optimal for the firm to hire at least one worker but its profit is bounded, which is never the case in the CW78 model. Intuitively, the CW78 result does not hold in this example because it is not profitable for the worker to replicate the owner. More precisely, it is profitable for the owner to exert s effort on supervising a subordinate who exerts total effort e . But it is not profitable for the worker to do the same. The reason is that f is concave. Because the owner exerts more total effort than the worker, the opportunity cost of

supervising a subordinate is lower for the owner than for the worker.

This example highlights the difference between our results and the CW78 result. CW78 find that the firm can grow infinitely large if a worker can compensate for the opportunity cost of supervising him. Our results show that the worker needs to compensate for the *value* of supervision received by him. The two conditions are equivalent only when f is linear. In general, the CW78 condition may be satisfied even though the necessary condition in Theorem 2 is not. That is precisely when a finitely large firm may be optimal.

This example also illustrates that our results are typically easier to use than their abstract formulation might suggest. In particular, we did not need to solve for V or construct its concave envelope. That is because expected wages always enter $\bar{V}(0)$ linearly – they are simply averaged for any group of workers. So we only needed to consider convex combinations of $f(p)$. This is true in general and simple bounds on production and expected wages are often sufficient.

We also emphasize that our choice of f does not automatically limit firm size. Even though the individual production function is assumed to be concave, the firm as a whole does not exhibit decreasing returns to scale. We illustrate this point by considering the extreme case where supervision is not necessary.

When supervision is unnecessary, the owner can pay workers just enough to compensate them for the disutility of effort. So:

$$w(e, s) = \frac{e^2}{2}$$

One can see that it is still profitable for the owner to hire a worker who exerts effort $e = \frac{1}{2}$. But now that supervision is unnecessary, there is nothing preventing the owner from hiring additional workers who also exert effort $e = \frac{1}{2}$. Importantly, the owner can hire an infinite number of such workers and each of them will generate positive net profit for the firm. The firm's profit is therefore linear in the number of workers, there is no limit to how many workers can be hired, and total profit is unbounded. It is therefore the necessity of supervision which imposes a binding constraint on the size of the firm, not our choice of f .

4 General Model

This section presents our general model which relaxes many of the assumptions made in Section 2.

We start by allowing for a more general class of monitoring technologies. To do so, we assume that managers observe one signal for each subordinate they manage. The monitoring technology is then defined as a mapping from the worker's level of effort and the level of supervision he receives to a distribution over the set of signals. Formally:

$$\Gamma: (e_i, s_i) \rightarrow \Delta Y$$

where Γ is the monitoring technology, e_i is the level of effort exerted by worker i , s_i is the level of supervision received by him, and Y is the set of possible signals. For example, under the monitoring technology in CW78, the set of signals is $Y = [0, 1] \cup \{\phi\}$ where ϕ represents no signal. The signal observed by the manager is either $e_i \in [0, 1]$ or ϕ .

Our key assumption is that some supervision is necessary to learn anything about a worker's level of effort. Formally:

$$\Gamma(e_i, s_i = 0) = \Gamma(e'_i, s_i = 0)$$

for any $e_i \neq e'_i$. Thus, the signal observed by the manager contains no information about a worker's effort level when the worker receives zero supervision. However, we do not impose any restrictions on what the manager observes when the level of supervision is positive. For example, our framework allows for a monitoring technology such that a worker's effort level is observed perfectly whenever $s_i > \epsilon$ where $\epsilon > 0$ is small. Such technologies are of particular interest because monitoring may be very cheap in some industries.

We now define an employment contract to be a mapping from signals to wages. Formally:

$$\Psi_i: y_i \rightarrow \mathbb{R}_+$$

where Ψ_i is the employment contract offered to worker i and y_i is the signal observed by his manager.

Notice that wages are assumed to be weakly positive. This is the *limited liability* constraint. Also implicit in this formulation is the assumption that employment contracts are static and that managers cannot misreport the signals they observe.

We do not impose any additional restrictions on the monitoring technology or the employment contracts. In particular, our results are not restricted to settings with effort contingent contracts. The signals may contain information about effort, output, or any other outcome that is a function of effort. In all cases, the combination of monitoring technology and employment contract will generate a distribution over wages for every possible level of effort and supervision. Our results depend only on this relationship.

Our restrictions on the production function and the workers' utility index are unchanged from Section 2. We restate them here for the reader's convenience:

1. $f(p)$ is continuous, increasing, and bounded.
2. $f(0) = 0$.
3. $\sup_{p \in (0,1]} \frac{f(p)}{p} < \infty$.
4. $u(w, e)$ is continuous.
5. $u(w, e)$ is increasing in w .
6. $u(w, e)$ is decreasing in e .
7. The outside option gives zero utility.

Given a monitoring technology, employment contract, and level of supervision, workers exert a level of effort that maximizes their expected utility. We say that a contract Ψ_i *induces* effort e_i at supervision level s_i if e_i maximizes worker i 's utility:

$$e_i \in \arg \max_{\tilde{e}} \mathbb{E}_{y \sim \Gamma(\tilde{e}, s_i)} \left[u(\Psi_i(y), \tilde{e}) \right]$$

As before, if the set of maximizers is not a singleton, we assume that the worker exerts the highest level of effort that maximizes his expected utility.

We define $w(e, s)$ to be the lowest expected wage that the owner has to pay a worker to induce effort e at supervision level s . If no available contract can induce effort e under supervision s , we define $w(e, s) = \infty$.

Notice that $w(e, s)$ captures all relevant details about the monitoring technology, employment contracts, and the utility index. The owner's problem is identical across all environments that generate the same expected wage function. While the monitoring technology and the utility index are assumed to be exogenous, the employment contract is chosen by the owner. The expected wage function is therefore an endogenous object. Nevertheless, we state our results in terms of $w(e, s)$ because doing so allows us to provide more general results. To do otherwise would require a different result for every combination of monitoring technology and utility index. That is both infeasible and uninformative.

We now show that our results from Section 2 extend to the more general setting considered here. All proofs are contained in the appendix but we restate our results for clarity.

Theorem 1. Incentive problems do *not* limit firm size if $V(0) > 0$.

Lemma 1. $\bar{V}(0) \geq \bar{V}(\delta)$ for any $\delta \geq 0$.

Theorem 2. Incentive problems do not limit the size of a firm only if $\bar{V}(0) > 0$.

Corollary 1. If $\bar{V}(0) < 0$ and $V(\delta) > -\infty$ for some $\delta > 0$, the optimal size of a firm is at most N^* where N^* is the largest integer such that $\bar{V}(\frac{-s}{N^*}) \geq 0$ for some $s \in [0, 1]$.

We emphasize again that our results are independent of the firm's structure. If the necessary condition is not satisfied, then incentive problems alone will limit firm size, regardless of how the firm is organized. Similarly, the bound in Corollary 1 is a universal bound. Reorganizing the firm may change its profits, and may determine how close it can come to this bound, but it cannot eliminate the bound itself. Thus, our results should not be interpreted as identifying an optimal hierarchy or span of control. Rather, they place a restriction on what any organizational structure can achieve. Even if the firm is free to choose its hierarchy, allocate supervisory responsibilities optimally, and redesign its organization as it grows, it cannot become arbitrarily large when the

necessary condition fails. Conversely, if the sufficient condition is satisfied, the firm can grow without bound even with a very simple organizational structure.

Our results also yield transparent comparative statics. For example, an increase in worker productivity raises $\bar{V}(\delta)$ and relaxes the constraint on firm size. As a result, incentive problems alone limit less productive firms to a smaller size. Similarly, any change in the utility index, the monitoring technology, or the set of available contracts that lowers the expected wage has the same effect as an increase in productivity. Examples include monitoring technologies that make effort or output easier to verify, workers who are less tempted to shirk (for instance because they are intrinsically motivated), and a richer set of contracts (for instance because firms are able to offer nonlinear performance bonuses rather than just linear piece rates). Changes in the opposite direction tighten the constraint. What matters is therefore not productivity itself but productivity relative to the cost of providing incentives. Two industries with the same output per worker can face very different limits on firm size if effort is easier to verify in one than in the other.

5 Conclusion

This paper studies whether incentive problems alone can limit the size of a firm. We show that they can. When workers cannot compensate the firm for the supervision they receive, the owner's finite capacity to supervise places an upper bound on firm size. This bound depends on worker productivity, the monitoring technology, the severity of the incentive problem, and the contracts the firm can offer. A change in any of these can relax or tighten the constraint on firm size. Reorganizing the firm cannot.

Our results qualify the conclusion of Calvo and Wellisz (1978) that incentive problems do not limit firm size in general. We show that their conclusion depends on the assumption of a linear production function. When production is concave, hiring a worker can be profitable even though incentive problems limit the size of the firm. No additional friction, such as communication costs or complementarities in production, is needed. Our results can therefore be read as identifying one source of the decreasing returns to the entrepreneur function that Coase (1937) proposed as the

limit to firm size.

Our model rules out collusion between managers and their subordinates, and it assumes that contracts are static. Collusion would make supervision by workers less reliable than supervision by the owner, and we expect it to tighten the constraint on firm size. Dynamic contracts may work in the opposite direction. The threat of dismissal or the prospect of promotion can lower the wage needed to induce effort, although neither can remove the need for supervision altogether. We leave both extensions to future work.

References

- Aggarwal, Rajesh K. and Andrew A. Samwick**, “Performance Incentives within Firms: The Effect of Managerial Responsibility,” *The Journal of Finance*, 2003, 58 (4), 1613–1650.
- Aghion, Philippe and Jean Tirole**, “Formal and Real Authority in Organizations,” *Journal of Political Economy*, 1997, 105 (1), 1–29.
- Alonso, Ricardo, Wouter Dessein, and Niko Matouschek**, “When Does Coordination Require Centralization?,” *American Economic Review*, 2008, 98 (1), 145–179.
- Baker, George, Michael Gibbs, and Bengt Holmström**, “The Internal Economics of the Firm: Evidence from Personnel Data,” *The Quarterly Journal of Economics*, 1994, 109 (4), 881–919.
- , —, and —, “The Wage Policy of a Firm,” *The Quarterly Journal of Economics*, 1994, 109 (4), 921–955.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul**, “Social Preferences and the Response to Incentives: Evidence from Personnel Data,” *The Quarterly Journal of Economics*, 2005, 120 (3), 917–962.
- , —, and —, “Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment,” *The Quarterly Journal of Economics*, 2007, 122 (2), 729–773.

- , —, and —, “Social Connections and Incentives in the Workplace: Evidence from Personnel Data,” *Econometrica*, 2009, 77 (4), 1047–1094.
- Bloom, Nicholas and John Van Reenen**, “Measuring and Explaining Management Practices across Firms and Countries,” *The Quarterly Journal of Economics*, 2007, 122 (4), 1351–1408.
- , **Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts**, “Does Management Matter? Evidence from India,” *The Quarterly Journal of Economics*, 2013, 128 (1), 1–51.
- Bolton, Patrick and Mathias Dewatripont**, “The Firm as a Communication Network,” *The Quarterly Journal of Economics*, 1994, 109 (4), 809–839.
- Boning, Brent, Casey Ichniowski, and Kathryn Shaw**, “Opportunity Counts: Teams and the Effectiveness of Production Incentives,” *Journal of Labor Economics*, 2007, 25 (4), 613–650.
- Bull, Clive, Andrew Schotter, and Keith Weigelt**, “Tournaments and Piece Rates: An Experimental Study,” *Journal of Political Economy*, 1987, 95 (1), 1–33.
- Calvo, Guillermo A. and Stanislaw Wellisz**, “Supervision, Loss of Control, and the Optimum Size of the Firm,” *Journal of Political Economy*, 1978, 86 (5), 943–952.
- Coase, Ronald H.**, “The Nature of the Firm,” *Economica*, 1937, 4 (16), 386–405.
- Datta, Saikat**, “On Control Losses in Hierarchies: A Synthesis,” *Rationality and Society*, 1996, 8 (4), 387–412.
- Dessein, Wouter**, “Authority and Communication in Organizations,” *The Review of Economic Studies*, 2002, 69 (4), 811–838.
- Fama, Eugene F.**, “Agency Problems and the Theory of the Firm,” *Journal of Political Economy*, 1980, 88 (2), 288–307.
- Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov**, “Team Incentives and Performance: Evidence From a Retail Chain,” *American Economic Review*, 2017, 107 (8), 2168–2203.

- Garicano, Luis**, “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 2000, 108 (5), 874–904.
- Gibbons, Robert**, “Piece-Rate Incentive Schemes,” *Journal of Labor Economics*, 1987, 5 (4), 413–429.
- **and Michael Waldman**, “A Theory of Wage and Promotion Dynamics Inside Firms,” *The Quarterly Journal of Economics*, 1999, 114 (4), 1321–1358.
- **and —**, “Enriching a Theory of Wage and Promotion Dynamics Inside Firms,” *Journal of Labor Economics*, 2006, 24 (1), 59–108.
- Grossman, Sanford J. and Oliver D. Hart**, “Implicit Contracts Under Asymmetric Information,” *The Quarterly Journal of Economics*, 1983, 98 (Supplement), 123–156.
- **and —**, “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,” *Journal of Political Economy*, 1986, 94 (4), 691–719.
- Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan**, “Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation,” *Journal of Political Economy*, 2003, 111 (3), 465–497.
- Hart, Oliver and John Moore**, “Property Rights and the Nature of the Firm,” *Journal of Political Economy*, 1990, 98 (6), 1119–1158.
- Holmström, Bengt**, “Moral Hazard and Observability,” *The Bell Journal of Economics*, 1979, 10 (1), 74–91.
- **and Paul Milgrom**, “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *The Journal of Law, Economics, and Organization*, 1991, 7 (Special Issue), 24–52.
- **and —**, “The Firm as an Incentive System,” *American Economic Review*, 1994, 84 (4), 972–991.
- Jensen, Michael C. and Kevin J. Murphy**, “Performance Pay and Top-Management Incentives,” *Journal of Political Economy*, 1990, 98 (2), 225–264.

- Kamenica, Emir and Matthew Gentzkow**, “Bayesian Persuasion,” *American Economic Review*, 2011, 101 (6), 2590–2615.
- Keren, Michael and David Levhari**, “The Internal Organization of the Firm and the Shape of Average Costs,” *The Bell Journal of Economics*, 1983, 14 (2), 474–486.
- Lazear, Edward P.**, “Salaries and Piece Rates,” *The Journal of Business*, 1986, 59 (3), 405–431.
- , “Performance Pay and Productivity,” *American Economic Review*, 2000, 90 (5), 1346–1361.
- Malcomson, James M.**, “Work Incentives, Hierarchy, and Internal Labor Markets,” *Journal of Political Economy*, 1984, 92 (3), 486–507.
- Mirrlees, James A.**, “The Optimal Structure of Incentives and Authority Within an Organization,” *The Bell Journal of Economics*, 1976, 7 (1), 105–131.
- Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor**, “Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment,” *American Economic Review*, 2002, 92 (4), 850–873.
- Oyer, Paul and Scott Schaefer**, “Why Do Some Firms Give Stock Options to All Employees? An Empirical Examination of Alternative Theories,” *Journal of Financial Economics*, 2005, 76 (1), 99–133.
- Prendergast, Canice**, “The Provision of Incentives in Firms,” *Journal of Economic Literature*, 1999, 37 (1), 7–63.
- Qian, Yingyi**, “Incentives and Loss of Control in an Optimal Hierarchy,” *The Review of Economic Studies*, 1994, 61 (3), 527–544.
- Radner, Roy**, “The Organization of Decentralized Information Processing,” *Econometrica*, 1993, 61 (5), 1109–1146.
- Rajan, Raghuram G. and Julie Wulf**, “The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies,” *The Review of Economics and Statistics*, 2006, 88 (4), 759–773.

Shavell, Steven, “Risk Sharing and Incentives in the Principal and Agent Relationship,” *The Bell Journal of Economics*, 1979, 10 (1), 55–73.

Shearer, Bruce, “Piece Rates, Fixed Wages, and Incentives: Evidence From a Field Experiment,” *The Review of Economic Studies*, 2004, 71 (2), 513–534.

Tsumagari, Masatoshi, “Incentives, Monitoring, and Firm Structure.” PhD dissertation, Boston University 1999.

Williamson, Oliver E., “Hierarchical Control and Optimum Firm Size,” *Journal of Political Economy*, 1967, 75 (2), 123–138.

Appendix

We prove our results in the general setting described in Section 4. The setting described in Section 2 is a special case.

Theorem 1. Incentive problems do *not* limit firm size if $V(0) > 0$.

Proof. Assume $V(0) > 0$.

Then there exists e^* and s^* such that:

$$f(e^* - s^*) - w(e^*, s^*) > 0$$

Define $M = f(e^* - s^*) - w(e^*, s^*)$ and define $M_0 = f(1 - s^*) + f(e^*) - w(e^*, s^*)$.

Notice that M_0 is the total profit generated by a firm in which the owner employs exactly one worker who exerts e^* effort and receives s^* supervision.

Consider a firm with two workers. Suppose both workers exert e^* effort; Worker 2 is supervised by Worker 1 and receives s^* supervision; Worker 1 is supervised by the owner. Then the total profit of this firm is given by:

$$f(1 - s^*) + f(e^* - s^*) + f(e^*) - w(e^*, s^*) - w(e^*, s^*) = M_0 + M$$

The owner can continue adding employees in this manner. The total profit of a firm with n employees will be given by $M_0 + (n - 1)M$.

Since $M > 0$, total profit is strictly increasing in n . Thus, the firm can grow infinitely large and generate infinite profit.

□

Lemma 1. $\bar{V}(0) \geq \bar{V}(\delta)$ for any $\delta \geq 0$.

Proof. Suppose $V(b) > V(0)$ for some $b > 0$.

Then there exists some $\tilde{e}$, $\tilde{s}$, and $\tilde{s}'$ such that $\tilde{s}' - \tilde{s} = b$ and:

$$f(\tilde{e} - \tilde{s}') - w(\tilde{e}, \tilde{s}) > V(0)$$

But recall that f is increasing. So:

$$\begin{aligned} f(\tilde{e} - \tilde{s}') - w(\tilde{e}, \tilde{s}) &\leq f(\tilde{e} - \tilde{s}) - w(\tilde{e}, \tilde{s}) \\ &\leq V(0) \end{aligned}$$

This is a contradiction. So $V(0) \geq V(b)$ for any $b \geq 0$.

Now suppose there exists some $\delta > 0$ such that $\bar{V}(0) < \bar{V}(\delta)$.

Fix any $a < 0$ and pick $\alpha \in (0, 1)$ such that $\alpha a + (1 - \alpha)\delta = 0$. Then by concavity:

$$\bar{V}(0) \geq \alpha \bar{V}(a) + (1 - \alpha) \bar{V}(\delta)$$

Therefore:

$$\bar{V}(0) > \bar{V}(a) \geq V(a) \tag{3}$$

We already know that for every $b \geq 0$:

$$\bar{V}(0) \geq V(0) \geq V(b) \quad (4)$$

Recall that $a < 0$ can be any. So combining Equation (3) and Equation (4) gives that for every $\delta \in [-1, 1]$:

$$\bar{V}(0) \geq V(\delta)$$

But then for all δ :

$$\bar{V}(0) \geq \bar{V}(\delta)$$

This is the desired contradiction and proves that $\bar{V}(0) \geq \bar{V}(\delta)$ for all $\delta \geq 0$.

□

Theorem 2. Incentive problems do not limit the size of a firm only if $\bar{V}(0) > 0$.

Proof. Consider the following maximization problem:

$$\begin{aligned} & \underset{\delta}{\text{maximize}} && \bar{V}(\delta) \\ & \text{subject to} && \delta \geq 0 \end{aligned}$$

Lemma 1 tells us that the maximum is $\bar{V}(0)$.

Now there are two possibilities:

First consider the case when $\bar{V}(\delta) > -\infty$ for some $\delta > 0$.

Then $\delta = 0$ is in the relative interior of the domain of $\bar{V}(\delta)$ and the Kuhn-Tucker theorem tells us that there exists a finite $\lambda \geq 0$ such that:

$$\bar{V}(0) \geq \bar{V}(\delta) + \lambda\delta$$

Now consider any firm with n workers. Let e_i be the total level of effort exerted by worker i . Let s_i be the level of supervision received by him. Let s'_i be the total level of supervision he provides to his subordinates (if any). Let σ be the total level of supervisory effort exerted by the owner.

The total profit generated by this firm is:

$$\begin{aligned}
f(1 - \sigma) + \sum_{i=1}^n \left[f(e_i - s'_i) - w(e_i, s_i) \right] &\leq f(1 - \sigma) + \sum_{i=1}^n \left[\bar{V}(s'_i - s_i) \right] \\
&\leq f(1 - \sigma) + \sum_{i=1}^n \left[\bar{V}(0) - \lambda(s'_i - s_i) \right] \\
&= f(1 - \sigma) + n\bar{V}(0) - \lambda \sum_{i=1}^n (s'_i - s_i) \\
&= f(1 - \sigma) + n\bar{V}(0) + \lambda\sigma \\
&\leq f(1 - \sigma) + n\bar{V}(0) + \lambda
\end{aligned}$$

Both $f(1 - \sigma)$ and λ are finite. So a firm can generate infinite profit only if $\bar{V}(0) > 0$.

Now consider the case when $\bar{V}(\delta) = -\infty$ for all $\delta > 0$. Then λ need not be finite.

But that also means $e_i > s_i$ cannot be induced. Then:

$$\begin{aligned}
\sum_i p_i &= \sum_i (e_i - s'_i) \\
&\leq \sum_i (s_i - s'_i) \\
&= \sigma
\end{aligned}$$

Now define $M = \sup_{p \in (0,1]} \frac{f(p)}{p}$. Recall that $M < \infty$ by assumption.

Then $f(p) \leq Mp$ for all $p > 0$.

The firm's total profit is therefore:

$$\begin{aligned}
f(1 - \sigma) + \sum_i \left[f(e_i - s'_i) - w(e_i, s_i) \right] &\leq f(1 - \sigma) + \sum_i f(e_i - s'_i) \\
&\leq f(1 - \sigma) + \sum_i M(e_i - s'_i) \\
&\leq f(1 - \sigma) + M\sigma \\
&\leq f(1) + M
\end{aligned}$$

Since $f(1)$ and M are both finite, the firm's total profit is again bounded above.

□

Corollary 1. If $\bar{V}(0) < 0$ and $V(\delta) > -\infty$ for some $\delta > 0$, the optimal size of a firm is at most N^* where N^* is the largest integer such that $\bar{V}(\frac{-s}{N^*}) \geq 0$ for some $s \in [0, 1]$.

Proof. Fix any firm with n workers exerting positive effort. Exclude workers exerting zero effort.

Let δ_i be the excess supervision supplied by worker i . Let σ be total level of supervisory effort exerted by the owner.

Recall (from the previous proof) that the total profit generated by this firm is at most:

$$f(1 - \sigma) + \sum_{i=1}^n \bar{V}(\delta_i)$$

We know that $\sum_i \delta_i = -\sigma$ and $\bar{V}$ is concave. So Jensen's inequality gives:

$$f(1 - \sigma) + \sum_{i=1}^n \bar{V}(\delta_i) \leq f(1 - \sigma) + n \bar{V}\left(\frac{-\sigma}{n}\right)$$

The owner generates $f(1)$ units of profit when she operates alone. So hiring n workers is optimal only if:

$$\begin{aligned}
& f(1 - \sigma) + n\overline{V}\left(\frac{-\sigma}{n}\right) \geq f(1) \\
\implies & n\overline{V}\left(\frac{-\sigma}{n}\right) \geq f(1) - f(1 - \sigma) > 0 \\
\implies & \overline{V}\left(\frac{-\sigma}{n}\right) \geq 0
\end{aligned}$$

Recall that $\overline{V}$ is concave by definition. If $V(\delta) > -\infty$ for some $\delta > 0$ then $\overline{V}(\delta) > -\infty$ for the same $\delta > 0$ and $\delta = 0$ is in the relative interior of the domain of $\overline{V}$.

We know that a finite concave function is continuous over its relative interior so $\overline{V}$ will be continuous at $\delta = 0$.

Therefore there exists some $\overline{N}$ such that $\overline{V}\left(\frac{-\sigma}{n}\right) < 0$ for all $n > \overline{N}$.

Defining N^* as the largest integer that is weakly smaller than $\overline{N}$ completes the proof.

□