

Optimal contracts with reflection^{*}

Borys Grochulski[†]

Yuzhe Zhang[‡]

January 27, 2017

Abstract

In this paper, we show that whenever the agent's outside option is nonzero, the optimal contract in the continuous-time principal-agent model of Sannikov (2008) is reflective at the lower bound. This means the agent is never terminated or retired after poor performance. Instead, the agent is asked to suspend effort temporarily, as in Zhu (2013), which brings the agent's continuation value up. The agent is then asked to resume effort, and the contract continues. We show that a nonzero agent's outside option arises endogenously if the agent is allowed to quit and find a new firm. In addition, we find new dynamics of the reflection at the lower bound. In the baseline model, the reflection is slow, as in Zhu (2013), i.e., effort is suspended often. However, if the agent's disutility from the first unit of effort is zero, which is a standard Inada condition, or if his utility of consumption is unbounded below, the reflection becomes fast, i.e., effort is suspended seldom.

Keywords: dynamic moral hazard, quitting, random search, reflective dynamics, ODE splicing, sticky Brownian motion, fast reflection, instantaneous control

JEL codes: D82, D86, M52, C61

1 Introduction

In an important contribution to the literature on incentives, Sannikov (2008) develops and applies methods for computing optimal contracts in a class of continuous-time principal-agent problems. Sannikov (2008) makes this class of contracting problems tractable by providing a representation of the agent's continuation utility as a diffusion process and by identifying

^{*}The authors would like to thank Huberto Ennis, Boyan Jovanovic, Chris Phelan, Yuliy Sannikov, and Russell Wong for their helpful comments, and Jackson Evert for excellent research assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[†]Federal Reserve Bank of Richmond, borys.grochulski@rich.frb.org.

[‡]Texas A&M University, zhangeager@tamu.edu.

contracts with solutions to the Hamilton-Jacobi-Bellman equation, an ordinary differential equation (ODE). The solutions obtained in Sannikov (2008) provide several new insights into the structure of optimal dynamic incentive contracts. In particular, Sannikov (2008) shows that an optimal contract must terminate (retire the agent) after sufficiently poor performance, when the agent's continuation utility becomes so low that incentives can no longer be provided. This result is particularly interesting because it has no analog in the discrete-time literature.

In this paper, we qualify this result. We show that termination after poor performance is optimal only if the firm can drive the agent's continuation utility down all the way to the agent's minimax payoff of zero. If the firm faces a lower bound on the agent's continuation utility that is even slightly higher than the absolute lowest payoff the agent could get, then an optimal contract does not terminate after poor performance but rather reflects off the lower bound and continues to provide incentives to the agent. The agent's outside option is a natural lower bound on his continuation utility inside a contract. If the outside option is of any value to the agent, the minimax payoff cannot be implemented and, thus, the contract does not terminate after poor performance. Termination after poor performance occurs in the optimal contract only in the knife-edge case in which the outside option provides absolutely no value to the agent.

The benchmark model we study is exactly the same as in Sannikov (2008). In this principal-agent relationship, the agent's private actions are costly to the agent but enhance the output of the relationship. The output is publicly observable. The principal/firm is risk-neutral, while the agent is risk-averse. Both discount future payoffs at the same rate $r > 0$. The agent's continuation utility, W_t , is a sufficient state variable in this contracting problem. The agent has an outside option with value B , which imposes a lower-bound constraint on the contract: $W_t \geq B$ at all t . The agent's utility from consumption and disutility from actions/effort are such that his minimax payoff in the relationship is zero.

In this model, we show that with any $B > 0$, i.e., with an outside option even slightly higher than the agent's minimax payoff, the optimal contract does not terminate when W_t reaches B . Rather, W_t reflects off B and the contract continues. Reflection occurs because both the agent's compensation and his effort supply requirement are zero when W_t drops down to B . This means that both the sensitivity of W_t to the agent's performance and the agent's current utility flow are zero when $W_t = B$. With zero current utility flow, the contract delivers the agent's continuation utility $W_t = B > 0$ by promising him more in the future, i.e., by growing W_t at the agent's rate of time preference r . Thus, W_t moves up, away from B , and the contract continues without termination.

Figure 1 replicates the left panel of Figure 3 in Sannikov (2008). The faint vertical line indicates the exogenous lower bound B on the agent's continuation utility, which is taken to be 0.1 in

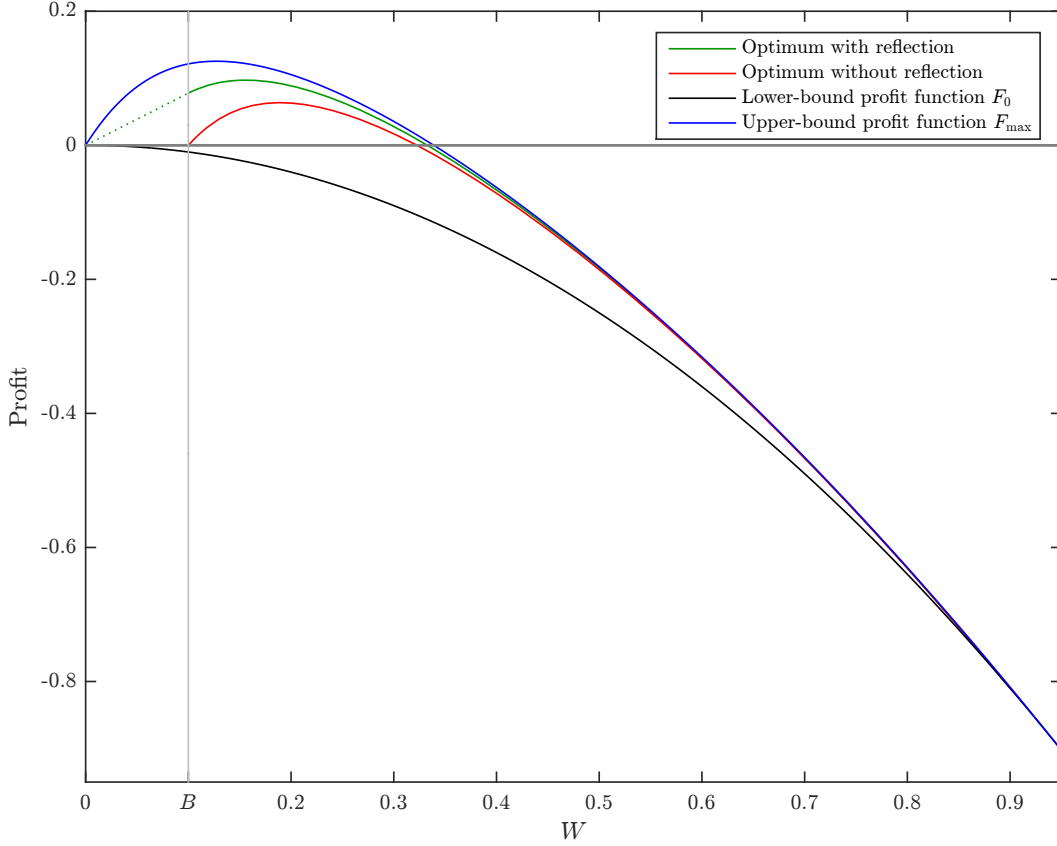


Figure 1: The firm's profit function with the agent's exogenous lower bound $B = 0.1$.

this example. The curve labeled “optimum without reflection” is the firm's profit function derived in Sannikov (2008) under the assumption that the contract terminates whenever the agent is asked to provide zero effort, so, in particular, the contract terminates at the lower bound B . The curve labeled “optimum with reflection” is the firm's profit function we obtain by relaxing this assumption. As we see, allowing for reflection leads to a better contract, which shows that the assumption of termination at $B > 0$ is restrictive. The upper-bound function $F_{\max}$ is the firm's optimal profit function when the agent's value of the outside option is his minimax payoff of zero. In this case, the assumption of termination at the lower bound B is not restrictive. The lower-bound function $F_0(W)$ is the (negative) profit the firm can make by retiring the agent, i.e., asking him to exert zero effort forever and paying him constant compensation at a level that delivers his promised utility W .

Our optimal contract with reflection is similar to the suspension contract studied in Zhu (2013). When W_t reaches B , contract calls for the agent to provide zero effort, but this suspension of effort is only temporary. Effort becomes positive again as soon as W_t leaves B . In the special case of $B = 0$, the contract never leaves B , and, thus, the suspension of effort becomes

permanent.¹ This explains why the assumption of termination at the lower bound is not restrictive in the special case of $B = 0$.

After characterizing the optimal contract subject to a given lower bound $B > 0$, we derive comparative statics results with respect to B . It is intuitive that the optimal contract is more efficient when the exogenous lower bound B is lower because in the contracting problem with a looser lower bound the firm can always use the same contract as in the problem with a tighter lower bound, but not vice versa. Indeed, the highest possible profit curve, shown in Figure 1 as $F_{\max}$, can be attained by the firm only if it faces the loosest possible lower-bound constraint with $B = 0$. The case of $B = 0$, therefore, is special in two ways: it allows the firm to attain higher profit than with any $B > 0$, and the optimal contract terminates at the lower bound $B = 0$ while it reflects at any lower bound $B > 0$.

To further examine reflection versus termination of the contract at the lower bound B , we consider the question of endogenous determination of B . In Section 6, we embed our contracting problem in a simple model of the labor market similar to Phelan (1995). In this model, there are a large number of firms having access to a common production technology. At any point in time, an agent matched with a firm is allowed to quit and rejoin the labor market, where he can rematch with a new firm after paying a search cost.² In this setting, the agent's outside option is determined endogenously. Due to competition between firms, unless the search cost is infinite, the agent's value of the outside option, B , is strictly larger than his minimax payoff. Therefore, the contract with reflection at the lower bound is used in equilibrium.

With reflection, the optimal contract has interesting dynamics around the lower bound. When we follow Sannikov (2008) in assuming that the agent's marginal disutility of effort is strictly positive even as effort goes to zero, the dynamics of the reflection are similar to that at the suspension contract studied in Zhu (2013). The contract calls for zero effort at the lower bound B but positive effort when W_t is just above B . Incentive compatibility, thus, implies that the volatility of W_t is zero at B and positive above B . Moreover, volatility is discontinuous at B : as soon as W_t leaves B , its volatility jumps, which makes W_t return to B frequently. Zhu (2013) dubs these dynamics "slow reflection." We show that the same dynamics arise in the model of Sannikov (2008) whenever the lower bound B is larger than the agent's minimax payoff.

However, if the marginal disutility of effort at zero is zero, which is a standard Inada condition,

¹The contract cannot move up from $B = 0$ because zero is the agent's minimax payoff in the Sannikov (2008) model. The only incentive compatible way to provide this value to the agent is to keep W_t at zero forever.

²We model the search cost as a time spell in which the agent is unmatched and receives zero flow utility. By allowing the expected duration of this spell to be anything between zero and infinity, we generalize Phelan (1995), where rematching is assumed to be immediate.

the dynamics of the reflection change. As W_t moves toward the lower bound, the volatility of W_t approaches zero smoothly, while the drift of W_t remains strictly positive. With volatility small near B , the contract does not return to B frequently. We dub these dynamics “fast reflection” and study them in detail in Section 7.

In Section 8, we study two extensions of the model. The first one follows Phelan (1995) in allowing the firm to break the contract (i.e., walk away without delivering to the agent his promised continuation utility W_t) upon incurring a deadweight expense of $K > 0$. We show that if K is not too large, a reflective upper bound emerges for the agent’s continuation value process in the optimal contract. When W_t becomes close to this bound, the firm’s continuation profit becomes close to $-K$. At the upper bound, the firm is indifferent between continuing and walking away. At this point, the optimal contract shows reflective dynamics that mirror those at the lower bound. The agent is asked to provide zero effort and W_t becomes insensitive to realized output. Compensation paid to the agent, however, is positive and high enough to make the drift of W_t negative. Thus, W_t moves down and away from the upper bound, allowing for resumption of effort and further continuation of the contract. Sannikov (2008) shows that if the firm has no outside option, or if K is sufficiently large, the contract has a high termination/retirement point, where the agent’s continuation utility is so high that incentives are too expensive to provide. We show that this retirement point vanishes from the optimal contract when K is low. Therefore, if both the agent and the firm have good outside options (positive B and low K), the state variable W_t has two reflective boundaries and the relationship between the firm and the agent never terminates.

The second extension relaxes the assumption that the agent’s utility from consumption is bounded from below. We show that the specification with utility unbounded below also gives rise to fast reflection at the lower bound. However, the dynamics of this reflection are different from those with the Inada condition on the marginal disutility of effort: the drift of the agent’s continuation value explodes to plus infinity as W_t approaches its lower bound B .

Relation to the literature We study optimal contracts in the model of Sannikov (2008), relaxing the assumption that effort must terminate when the state variable W_t hits its lower bound. We show that slow reflection discovered in Zhu (2013) is not specific to the risk-neutral principal-agent model with an impatient agent, where risk-sharing and intertemporal consumption smoothing motives are absent, but it also appears in principal-agent relationships with risk-aversion and identical discounting, where risk-sharing and consumption smoothing are valuable. In particular, we show that termination of the contract following the agent’s poor performance is never optimal, save the corner case in which the agent’s outside option is equal to his minimax payoff.

In a model with no option to replace the agent, the result of Sannikov (2008) showing the

optimality of retiring the agent at the lower bound B may be surprising, as it has no analog in discrete-time dynamic contracting models, e.g., Atkeson and Lucas (1995), Phelan (1995).³ We show that this result is specific to B being the agent’s minimax payoff. Our main result is that if B is even slightly larger than minimax, the optimal contract reflects off B rather than terminates.

We view this result as significant for two reasons. First, it brings the lessons from an important continuous-time contracting model closer to those obtained previously in discrete time, which lets us see the analogy between these two cases despite the differences in their solution methods. Second, our result helps us better understand the retirement result of Sannikov (2008) as a limiting case of slow reflection. In the contract with reflection, i.e., with $B > 0$, effort is suspended at B and drift of W_t equals $rB > 0$. When $B = 0$, effort is still suspended at B and drift of W_t still equals rB , which now is zero and, thus, slow reflection becomes permanent absorption.

Despite the qualitative similarity between the discrete- and continuous-time cases, continuous-time methods developed in Sannikov (2008) allow us to study the dynamics of the optimal contract in more detail than what has been done in discrete time. We show that, in addition to slow reflection discovered by Zhu (2013), the optimal contract can have fast-reflection dynamics of at least two kinds: one with volatility vanishing, and one with drift exploding at the lower bound.⁴ Further, we obtain comparative statics results with respect to the exogenous lower bound B : when B is larger, the value the agent can obtain from a take-it-or-leave-it offer made by the firm increases, but the firm’s profit and the total surplus of the relationship decrease. These comparative statics allow us to prove the existence and uniqueness of a market equilibrium with search, in which B is endogenized.

In addition, our paper develops a new method that allows for solving the model of Sannikov (2008) in cases in which the HJB equation fails to satisfy the standard Lipschitz condition. In particular, in Section 7 this failure occurs because the coefficient in front of the second derivative of the value function (i.e., volatility of the state variable) equals zero at the boundary. We develop a change-of-variable technique to restore Lipschitz continuity. This technique is applicable to other problems in which volatility of the state variable may become degenerate, e.g., Grochulski and Zhang (2016).

Organization Section 2 describes the contracting environment we study, which is exactly

³Termination of the relationship after poor performance is often optimal in dynamic contracting models with agent replacement, e.g., Spear and Wang (2005), Wang and Yang (2015), or the case considered in the right panel of Figure 3 in Sannikov (2008).

⁴Fast reflection occurs in many continuous-time models used in other branches of economics and finance, e.g., Cox et al. (1985), Brunnermeier and Sannikov (2014). We show that fast reflection can occur in an optimal dynamic contract under private information.

the same as that in Sannikov (2008). Section 3 recalls the optimal contracts without reflection derived in Sannikov (2008). Section 4 provides an informal description of how contracts with reflection can be constructed by combining low- and high-action ODEs by extending the method of Zhu (2013). Section 5 contains formal analysis proving these results. Section 6 endogenizes the lower bound on the agent's continuation value by embedding the contracting problem with one-sided commitment in a simple model of the labor market similar to Phelan (1995). Section 7 assumes an Inada condition on the marginal disutility of effort and studies fast dynamics of reflection. Section 8 studies two extensions that allow for, respectively, two-sided limited commitment and utility of consumption unbounded below. Section 9 concludes.

2 The principal-agent problem

The principal-agent contracting problem is the same as in Sannikov (2008). A principal/firm hires an agent, whose private effort influences the firm's output. Cumulative output produced up to date t , X_t , follows

$$dX_t = A_t dt + \sigma dZ_t,$$

where $A_t \in \mathcal{A}$ is the agent's action (effort), Z_t is a standard Brownian motion on $(\Omega, \mathcal{F}, P)$, and $\sigma > 0$ is a constant. We assume that the set of feasible actions $\mathcal{A}$ is a compact interval $[0, \bar{A}]$ for some $\bar{A} > 0$.⁵ The contract is a pair of progressively measurable processes $\{(C_t, A_t); 0 \leq t < \infty\}$, where A_t is the action recommended for the agent to take at t and C_t is his compensation. The agent and the firm evaluate the contract according to, respectively,

$$\mathbb{E} \left[r \int_0^\infty e^{-rt} (u(C_t) - h(A_t)) dt \right],$$

and

$$\mathbb{E} \left[r \int_0^\infty e^{-rt} (A_t - C_t) dt \right],$$

where $r > 0$. The agent's utility function $u : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is C^2 with $u' > 0$, $u'' < 0$, $\lim_{c \rightarrow 0} u'(c) = \infty$, and $u(0) = 0$. The function $h : \mathcal{A} \rightarrow \mathbb{R}_+$ represents the agent's disutility from effort. We assume that h is strictly increasing and strictly convex with $h(0) = 0$. In Section 7, we assume $\lim_{a \rightarrow 0} h'(a) = 0$, which is a standard Inada condition. Outside of Section 7, we follow Sannikov (2008) in assuming that $\lim_{a \rightarrow 0} h'(a) =: \gamma_0 > 0$.

Under a given contract (C, A) , the agent's continuation value process is

$$W_t := \mathbb{E}_t \left[r \int_t^\infty e^{-r(s-t)} (u(C_s) - h(A_s)) ds \right].$$

⁵That $\mathcal{A}$ includes a neighborhood of 0 is important in Section 7. Other parts of our analysis can be extended to the more general case studied in Sannikov (2008), where $\mathcal{A}$ is an arbitrary compact subset of $\mathbb{R}$ with the smallest element 0.

Sannikov (2008) shows that there exists a progressively measurable process $\{Y_t; 0 \leq t < \infty\}$ such that the agent's continuation value from the contract satisfies

$$dW_t = r(W_t - u(C_t) + h(A_t))dt + rY_t(dX_t - A_tdt), \quad (1)$$

where $dX_t - A_tdt$ is the agent's performance relative to the benchmark A_tdt . Here, Y_t represents the sensitivity of the agent's continuation value to his performance observed by the firm. The contract is incentive compatible (IC) at t if

$$A_t \in \operatorname{argmax}_{a \in \mathcal{A}} Y_t a - h(a). \quad (2)$$

The above IC condition implies that $A_t = 0$ is IC if and only if $Y_t \leq \gamma_0$, $A_t \in (0, \bar{A})$ is IC if and only if $Y_t = h'(A_t) > \gamma_0$, and $A_t = \bar{A}$ is IC if and only if $Y_t \geq h'(\bar{A})$.

In the recursive form, the firm's problem is to maximize the profit $F(W)$ that it can attain in the relationship with the agent when the agent is owed the continuation value W . The HJB equation for this problem is

$$F(W) = \max_{c, a, Y} a - c + F'(W)(W - u(c) + h(a)) + \frac{1}{2}F''(W)r\sigma^2Y^2, \quad (3)$$

where controls a and Y must jointly satisfy the IC constraint (2), and the state variable W_t must satisfy the lower-bound constraint

$$W_t \geq B \quad \text{at all } t \geq 0.$$

The lower bound $B \geq 0$, which represents the agent's outside option, is taken as exogenous by the firm. We follow Sannikov (2008) in assuming that the firm has the option of not employing the agent. Thus, we focus on contracts that give a nonnegative profit to the firm as of $t = 0$.

3 Optimal contracts without reflection

3.1 Optimal contract when $B = 0$

Theorem 1 in Sannikov (2008) solves the above contracting problem in the case of the lower bound $B = 0$. The optimal contract is obtained from the unique solution $F_{\max}$ to the HJB equation (3) that satisfies the following conditions: $F_{\max} \geq F_0$, $F_{\max}(0) = 0$, there exists W_{gp} such that $F_{\max}(W_{gp}) = F_0(W_{gp})$ and $F'_{\max}(W_{gp}) = F'_0(W_{gp})$, where F_0 is the retirement profit function $F_0(W) := -u^{-1}(W) \leq 0$. The retirement profit $F_0(W)$ is attained by a static contract, in which $A_t = 0$, $C_t = c$, and $W_t = W$ at all t , where c satisfies $u(c) = W$.

Until time $\tau := \inf\{t : W_t = 0 \text{ or } W_t = W_{gp}\}$, the optimal contract is constructed from the policy functions $c(\cdot)$, $a(\cdot)$, and $Y(\cdot)$ that attain the solution $F_{\max}$ in the HJB equation (3). In particular, for any $W_0 \in (0, W_{gp})$, $C_t = c(W_t)$ and $A_t = a(W_t)$, where W_t solves

$$dW_t = r(W_t - u(c(W_t)) + h(a(W_t)))dt + rY(W_t)(dX_t - a(W_t)dt) \quad \text{for } 0 \leq t < \tau. \quad (4)$$

Under this contract, the support of the state variable W_t is $[0, W_{gp}]$. The optimal effort $a(W)$ is strictly positive for all W in the interior of $[0, W_{gp}]$. Correspondingly, the volatility of the agent's continuation value is strictly positive in $(0, W_{gp})$. The end points of the support interval, 0 and W_{gp} , are referred to as retirement points. At either of these points, effort a is zero, and drift and volatility of W_t are zero as well. The dynamics of the state variable are stopped, i.e., 0 and W_{gp} are absorbing states for the process W_t .

3.2 Optimal contract without reflection when $B > 0$

The optimal contract with $B = 0$, described above, does not use the zero-effort action $a = 0$ unless $F(W) = F_0(W)$, which only happens at the (permanent) retirement points $W = 0$ and $W = W_{gp}$. Theorem 3 in Sannikov (2008) solves for an optimal contract with $B > 0$ under the assumption that, as in the case of $B = 0$, the zero-effort action $a = 0$ is not used prior to agent retirement or contract termination.⁶ In particular, this solution maintains that the agent is fired when $W_t = B$ (i.e., the agent takes his outside option), from which the firm makes a continuation profit of zero. Under this assumption, the optimal contract is obtained from the solution $\tilde{F}$ of the HJB equation that satisfies the following conditions: $\tilde{F} \geq F_0$, $\tilde{F}(B) = 0$, there exists W_{gp} such that $\tilde{F}(W_{gp}) = F_0(W_{gp})$ and $\tilde{F}'(W_{gp}) = F_0'(W_{gp})$. This solution curve is depicted in Figure 1 as the optimum without reflection.

4 Optimal contracts allowing for reflection: intuition

In this paper, we relax the assumption that the zero-effort action is not used prior to agent retirement or contract termination. In particular, we allow for the zero-effort action to be applied at the lower bound B without terminating the contract. We show that this possibility is in fact optimal when $B > 0$. The dynamics of the optimal contract are reflective at the lower bound: the process W_t moves up from B and the contract continues.

Our approach to solving the HJB equation (3) combines the approaches of Sannikov (2008)

⁶Wang and Yang (2015) make the same assumption.

and Zhu (2013). We can write the HJB equation (3) as

$$F(W) = \max \left\{ \max_{c,a} a - c + F'(W)(W - u(c) + h(a)) + \frac{1}{2}F''(W)r\sigma^2(h'(a))^2, \quad (5) \right.$$

$$\left. \max_c -c + F'(W)(W - u(c)) \right\}. \quad (6)$$

Following Zhu (2013), we will construct a solution to the HJB equation from solutions to the component ODEs (5) and (6).

The first component ODE, (5), is exactly the equation studied in Sannikov (2008). In this equation, volatility Y is set equal to $h'(a) > 0$ at all W . In addition to incentive compatibility of contracts constructed from solutions to (5), this ensures that the volatility of the implied process W_t is everywhere strictly positive.⁷ Although it could be natural to call (5) positive-volatility ODE, in order to emphasize the analogy with Zhu (2013), we will refer to (5) as the high-action ODE.

The second component ODE, (6), forces volatility Y to be zero. The only incentive compatible action under $Y = 0$ is the zero-effort action $a = 0$. We will call this ODE low-action ODE. The advantage of having $Y = 0$ is that along any solution to the low-action ODE, the dynamics of W_t are deterministic, i.e., W_t is not sensitive to output. This property will allow for reflection of W_t at the lower bound B .

We will first study (5) and (6) separately and then combine solutions to these two equations to construct an optimal contract.

4.1 Solutions to the low-action ODE

To distinguish the solutions to the two component ODEs, we will continue to use F to denote solutions to the high-action ODE (5) but will use L to denote solutions to the low-action ODE (6). Thus, we can write the low-action ODE simply as

$$L(W) = \max_{c \geq 0} -c + L'(W)(W - u(c)), \quad (7)$$

or, denoting $u(c)$ as W' and using the definition of F_0 , equivalently as

$$L(W) = \max_{W' \geq 0} F_0(W') + L'(W)(W - W'). \quad (8)$$

⁷In particular, action $a = 0$ is allowed in (5) only with positive volatility $h'(0) = \gamma_0 > 0$. The pair $(a, Y) = (0, \gamma_0)$ is never used in the optimal contract. In the optimal contract, which we present in Theorem 1, action $a = 0$ is used with volatility $Y = 0$ when the optimal contract is determined by a solution to the other component ODE, (6). See Lemma A.2 in the Appendix. The pair $(a, Y) = (0, \gamma_0)$ is allowed in (5) merely for technical reasons.

We will look for solutions to this ODE in the region $\mathcal{R} := \{(W, L) : L > F_0(W)\}$ as we know that $F_0(W)$ is a lower bound on the firm's profit function.

We can show that if $L(W)$ is a solution to (8) in $\mathcal{R}$, then $L(W)$ is a straight line. Indeed, if we denote the maximizer in (8) by W^* , we have $L(W) = F_0(W^*) + L'(W)(W - W^*)$. Clearly, if $W^* = W$, then $L(W) = F_0(W^*) = F_0(W)$, which is outside of $\mathcal{R}$. Thus, all solutions to (8) inside the region $\mathcal{R}$ must have $W^* \neq W$. Differentiating (8), we get $L''(W)(W - W^*) = 0$. This and $W^* \neq W$ imply that all solutions to (8) inside $\mathcal{R}$ must have $L''(W) = 0$, i.e., be straight lines.

With this result, we can classify solutions to (8) inside $\mathcal{R}$ by their constant slope $L'(W) =: \alpha$.

Case 1. Take any $\alpha \geq 0$. Then the optimal W^* in (8) is zero, and (8) reduces to $L(W) = 0 + \alpha(W - 0) = \alpha W$. Thus, the set of solutions with nonnegative slope consists of all straight lines out of the origin, restricted to $\mathcal{R}$.

It will be convenient for us to single out the solution with slope zero as a separate subcase. We will refer to straight lines out of the origin with $\alpha > 0$ as Case 1a solutions, and with $\alpha = 0$ as the Case 1b solution.

Case 2. Take any $\alpha < 0$. Then the optimal W^* in (8) satisfies the first-order condition $F'_0(W^*) = \alpha$, and (8) reduces to $L(W) = F_0(W^*) + \alpha(W - W^*) = F_0(W^*) + F'_0(W^*)(W - W^*)$. Thus, the set of solutions with negative slope consists of all lines tangent to F_0 , restricted to $\mathcal{R}$.

Along any solution to the low-action ODE, the dynamics of W_t are deterministic. In particular, along any Case-1 solution we have $0 = W' = u(c)$, which means $c = 0$. Using $c = 0$ together with $a = Y = 0$ in (1), we get

$$dW_t = rW_t dt, \tag{9}$$

i.e., along any Case-1 solution, W_t grows deterministically at the rate r .

In Case-2 solutions, $0 < W' = u(c)$, which means $c > 0$. Depending on the sign of $W - W^*$, W_t will grow along some solutions but decline along others. These cases are less important to us for now, as they will not be part of the optimal contract until we consider two-sided limited commitment in Section 8.1.

4.2 Combining high- and low-action solutions

Let us now discuss informally how solutions to the two ODEs can be combined to construct a contract with reflection that improves on the optimal contract without reflection. Figure 2 replicates again the optimum without reflection, $\tilde{F}$, for the example with $B = 0.1$ presented

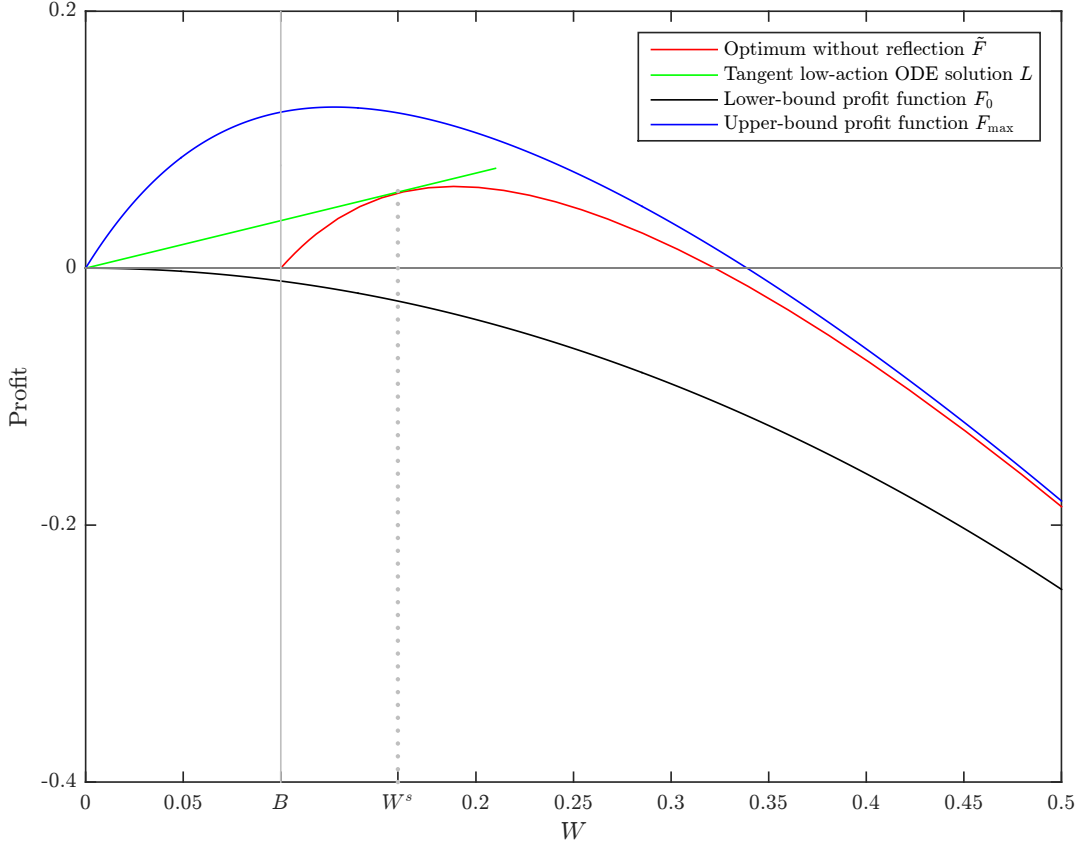


Figure 2: Combining the high- and low-action solutions. The splicing point is $W^s = 0.16 > 0.1 = B$. Positive effort is exerted to the right of W^s . Zero effort is exerted at and to the left of W^s .

in Figure 3 (left panel) of Sannikov (2008). As we see, there exists a unique Case-1 solution $L(W)$ to the low-action ODE that is tangent to $\tilde{F}$ at some $W^s > B$. Note that the slope of L is $\tilde{F}(W^s)/W^s$.

Consider a new contract (C, A) defined by using the optimal c from the solution L to the low-action ODE at all $W \in [B, W^s]$ and the optimal controls (c, a, Y) from the solution $\tilde{F}$ to the high-action ODE at all $W \in (W^s, W_{gp}]$. Because the two solutions satisfy at W^s the value matching and smooth pasting conditions, $L(W^s) = \tilde{F}(W^s)$ and $L'(W^s) = \tilde{F}'(W^s)$, this contract delivers to the firm profit $L(W)$ if $W \in [B, W^s]$ and $\tilde{F}(W)$ if $W \in (W^s, W_{gp}]$. Because $L(W) > \tilde{F}(W)$ for all $W \in [B, W^s)$, the new contract constitutes a Pareto improvement over the optimal contract without reflection.

By (9), the process W_t implied by this contract is deterministic in the interval $[B, W^s]$. The agent's continuation value W_t grows exponentially and moves out of $[B, W^s]$. Once W_t leaves $[B, W^s]$, it never drops below W^s , i.e., it reflects off W^s and stays in $[W^s, W_{gp}]$.

Note also in Figure 2 that the second derivatives of L and $\tilde{F}$ are not equal at W^s . Because

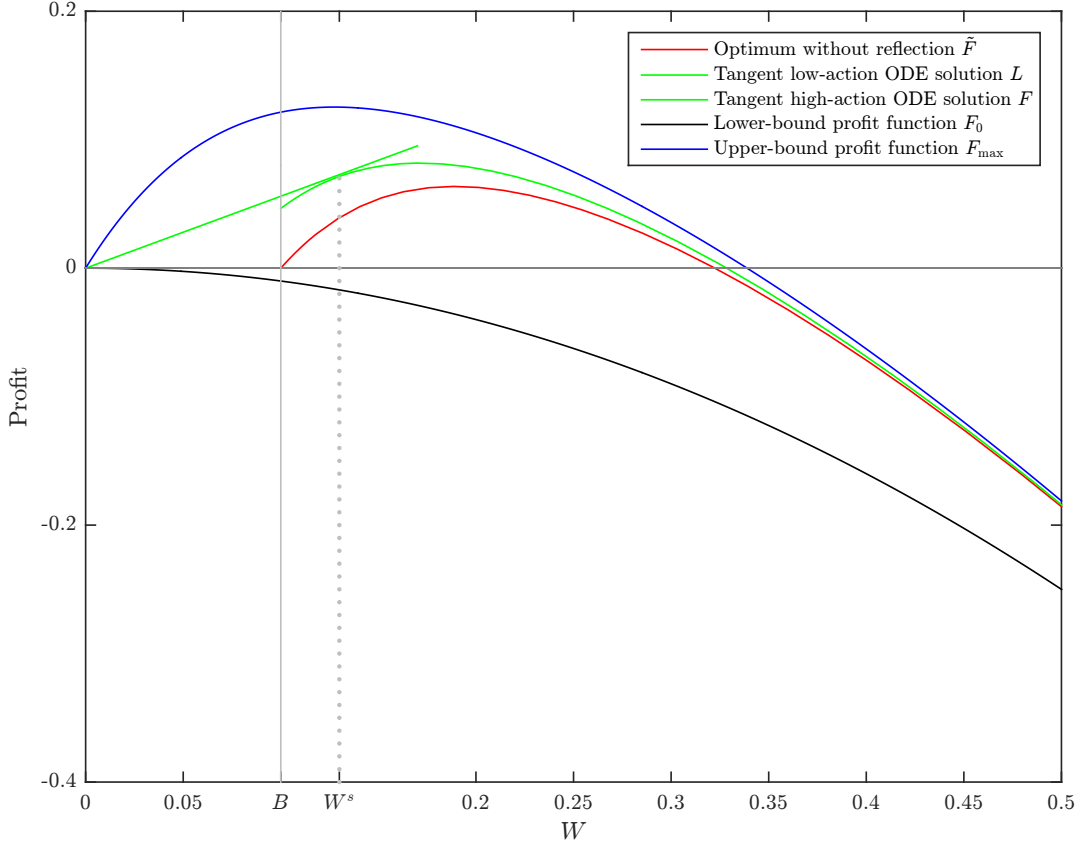


Figure 3: Combining the high- and low-action solutions at the splicing point $W^s = 0.13$ leads to a better contract than $\tilde{F}$ at all W .

$W^s > B$ is an interior point in the feasible support for the continuation value process, this means that the contract obtained by splicing L and $\tilde{F}$ at W^s is not an optimal contract. In fact, better combinations of low- and high-action ODE solutions exist. One such example is provided in Figure 3. In that example, the splicing point W^s is closer to the lower bound B , the low-action ODE solution has a higher slope, and the high-action ODE solution is everywhere above $\tilde{F}$. As before, the solutions to the two ODEs are spliced at a point where the smooth pasting conditions are satisfied, thus giving a consistent contract over the whole domain $[B, W_{gp}]$.⁸

The intuition for why a lower splicing point W^s allows the firm to attain a higher profit curve follows from the fact that the (endogenous) support $[W^s, W_{gp}]$ for the state variable W_t is larger when W^s is lower. Clearly, any feasible contract (C, A) remains feasible if the support for W_t is enlarged, so the firm cannot do worse with a lower W^s . In fact, the firm can do strictly better. At both W^s and W_{gp} , the contract must ask for zero effort (at W_{gp} permanently). With more

⁸ W_{gp} is not the same in these two examples. It is higher in the second case.

distance between W^s and W_{gp} , the contract can sustain positive effort for longer and/or ask for higher levels of effort while effort is positive, as the volatility of W_t that is necessary for higher agent effort does not bump W_t into W^s or W_{gp} as fast.

As we show formally next, the optimal contract obtains when the splicing point W^s is as low as possible, i.e., when it coincides with the exogenous lower bound B , as depicted in Figure 1. In this case, the splicing point cannot be moved further to the left, i.e., the endogenous support $[W^s, W_{gp}]$ cannot be made any larger.

5 Solution of the contracting problem

In this section, we provide a formal solution verifying the intuition given in the previous section.

As in Sannikov (2008), let us denote by W_{gp}^* the first-best effort shut-down threshold, i.e., $W_{gp}^* = u(c)$, where c solves $u'(c) = h'(0)$.

For $B \in [0, W_{gp}^*]$ and for two numbers y and y' , we will denote by $F_{(B,y,y')}$ the solution to the high-action ODE that starts at $W = B$ and satisfies the boundary condition $F(B) = y$, $F'(B) = y'$.⁹

We start out by examining solutions $F_{(B,0,0)}$, i.e., the high-action ODE solutions that start at the horizontal axis with the initial slope of zero, where the starting W is $B \in [0, W_{gp}^*]$. We are interested in these solutions because they can be pasted smoothly with the lowest of the low-action ODE solutions, i.e., the solution that follows the horizontal axis (solution Case 1b).

Lemma 1 (Largest lower bound) *There exists a unique $\bar{B} \in [0, W_{gp}^*]$ for which the solution $F_{(\bar{B},0,0)}$ satisfies a) $F_{(\bar{B},0,0)}(W) \geq F_0(W)$ for all $W \in [\bar{B}, W_{gp}^*]$, and b) $F_{(\bar{B},0,0)}(W_{gp}) = F_0(W_{gp})$ and $F'_{(\bar{B},0,0)}(W_{gp}) = F'_0(W_{gp})$ for some $W_{gp} \in [\bar{B}, W_{gp}^*]$.*

Conditions a) and b) in the above lemma are analogous to the conditions in Lemma 3 of Sannikov (2008). The above lemma identifies the largest lower bound, $\bar{B}$, at which the firm can find a contract that never violates this bound but also lets the firm break even in expectation as of $t = 0$. As we will see in Theorem 1 below, the optimal contract subject to the agent's quitting constraint at the lower bound $\bar{B}$ will be constructed by splicing the lowest of the positively sloped low-action solutions, i.e., the ray that follows the horizontal axis, with the high-action ODE solution $F_{(\bar{B},0,0)}$.

For $B > \bar{B}$, the solution curve $F_{(B,0,0)}$ stays strictly above F_0 for all $W \geq B$, i.e., it fails to satisfy condition b) of Lemma 1. This means that with $B > \bar{B}$ there is no contract such that

⁹Lemma 1 in Sannikov (2008) shows existence, uniqueness, concavity, and continuity in initial conditions of solutions to the high-action ODE.

$W_t \geq B$ at all t and $F(W_0) \geq 0$.

For $0 \leq B < \bar{B}$, the solution curve $F_{(B,0,0)}$ crosses the retirement profit curve F_0 at some $W > B$. By pasting $F_{(B,0,0)}$ with the horizontal low-action ODE solution at $W = B$, it is possible to obtain a feasible contract, i.e., a contract at which the agent's incentive and quitting constraints are satisfied and the firm breaks even. But because $F_{(B,0,0)}$ crosses F_0 , this contract would not be optimal. A better contract can be obtained if $F_{(B,0,0)}$ is replaced with $F_{(B,y,y')}$ such that $F_{(B,y,y')} > F_{(B,0,0)}$.

The next lemma describes how such a solution curve $F_{(B,y,y')}$ is obtained. Fix $B \in (0, \bar{B})$ and take some $y > 0$. The unique positively-sloped solution to the low-action ODE that goes through the point (B, y) has slope y/B . Using the value matching condition $F(B) = y$ and the smooth pasting condition $F'(B) = y/B$ as boundary conditions, we can solve the high-action ODE starting at the point (B, y) and obtain the solution curve $F_{(B,y,\frac{y}{B})}$. We look for the initial level y such that the curve $F_{(B,y,\frac{y}{B})}$ stays above the retirement profit curve F_0 and touches it.

Lemma 2 (Largest initial level and slope) *For each $B \in (0, \bar{B}]$, there exists a unique $0 \leq y < F_{\max}(B)$ such that the solution $F_{(B,y,\frac{y}{B})}$ satisfies a) $F_{(B,y,\frac{y}{B})}(W) \geq F_0(W)$ for all $W \in [B, W_{gp}^*]$, and b) $F_{(B,y,\frac{y}{B})}(W_{gp}) = F_0(W_{gp})$ and $F'_{(B,y,\frac{y}{B})}(W_{gp}) = F'_0(W_{gp})$ for some $W_{gp} \in [B, W_{gp}^*]$. If $B = \bar{B}$, then $y = 0$. For $B \in (0, \bar{B})$, $y > 0$.*

Let us denote the unique y pinned down in this lemma by $y^*(B)$. Also, we will denote the point W_{gp} pinned down by the smooth pasting condition between $F_{(B,y^*(B),\frac{y^*(B)}{B})}$ and F_0 by $W_{gp}(B)$.

5.1 Optimal contract

Following Zhu (2013), let us define a function $V : [B, W_{gp}(B)] \rightarrow \mathbb{R}$ by splicing the high-action ODE solution $F_{(B,y^*(B),\frac{y^*(B)}{B})}(W)$ with, respectively, the low-action ODE solution $L(W) = \frac{y^*(B)}{B}W$ at B , and with F_0 at $W_{gp}(B)$. That is, let

$$V(W) := \begin{cases} L(W) & \text{for } W = B, \\ F_{(B,y^*(B),\frac{y^*(B)}{B})}(W) & \text{for } W \in (B, W_{gp}(B)), \\ F_0(W) & \text{for } W = W_{gp}(B). \end{cases}$$

Theorem 1 *For each $B \in (0, \bar{B}]$, the function V is the firm's value function in the contracting problem with the agent's lower bound B . The optimal controls c, A, Y attaining V define an optimal contract $C_t = c(W_t)$, $A_t = a(W_t)$, where $\{W_t; 0 \leq t < \infty\}$ is a solution to (4). In particular, $c(B) = a(B) = Y(B) = 0$, with $dW_t = rBdt > 0dt$ when $W_t = B$; $a(W) > 0$ and $Y(W) > 0$ for all $W \in (B, W_{gp}(B))$; and $c(W_{gp}(B)) > 0$, $a(W_{gp}(B)) = Y(W_{gp}(B)) = 0$, with $dW_t = 0$ when $W_t = W_{gp}(B)$.*

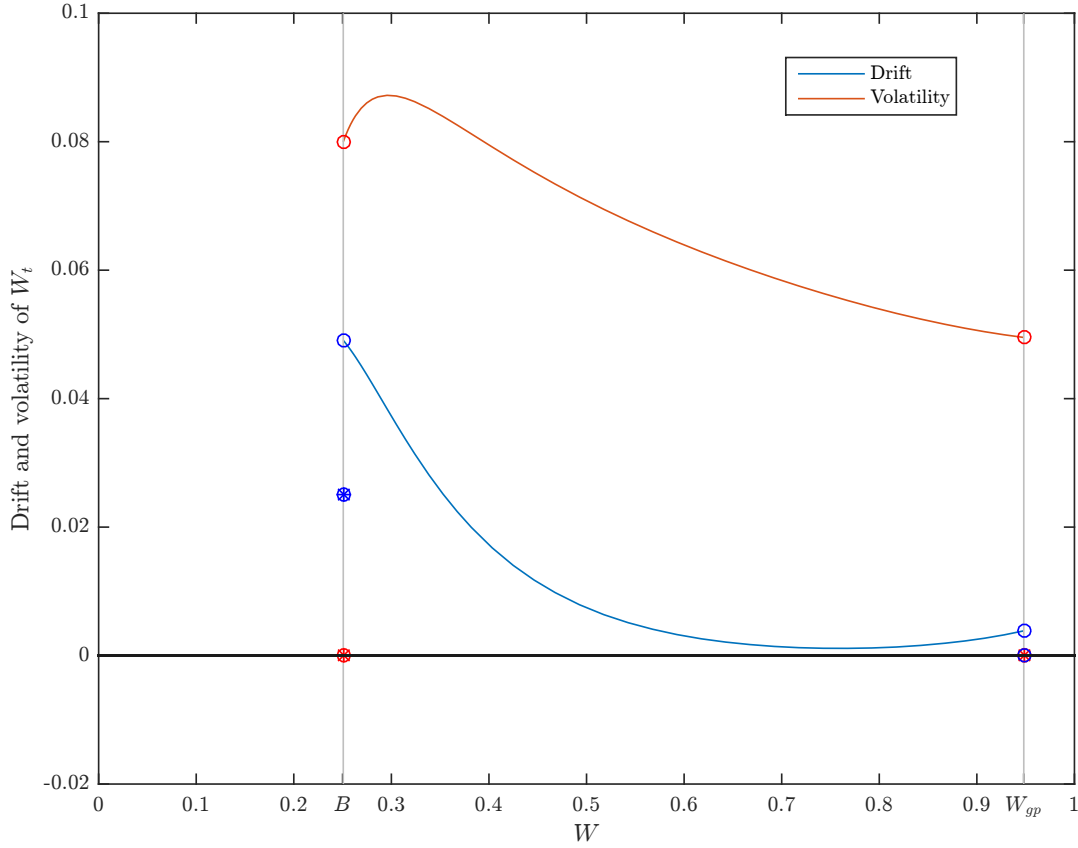


Figure 4: Drift and volatility of W_t . The volatility function is discontinuous, jumping down to zero at both boundaries. Drift is also discontinuous. It jumps down to zero at W_{gp} but is strictly positive at B .

The proof follows Sannikov (2008) very closely with two exceptions. The technical argument for the existence of a solution to (4) is modified to account for volatility of W_t vanishing at B , and the step verifying the optimality of the contract is modified to account for the reflection of the process W_t at B .

The drift and volatility functions of the process W_t , which determine the dynamics of the optimal contract, are shown in Figure 4. Because the super-contact condition is not satisfied at B or W_{gp} , the optimal controls Y and a jump at these points. In particular, volatility Y is extinguished at both boundaries. Since a jumps at the boundaries, drift of W_t is also discontinuous at these points. But despite this discontinuity, drift of W_t remains strictly positive at B , which generates reflection of the process W_t off B . The reflection is slow, as in Zhu (2013). That is, after hitting B , the process W_t returns to B frequently and spends a positive expected amount of time there. Section 7.5 provides additional discussion of the speed of reflection of W_t off B .

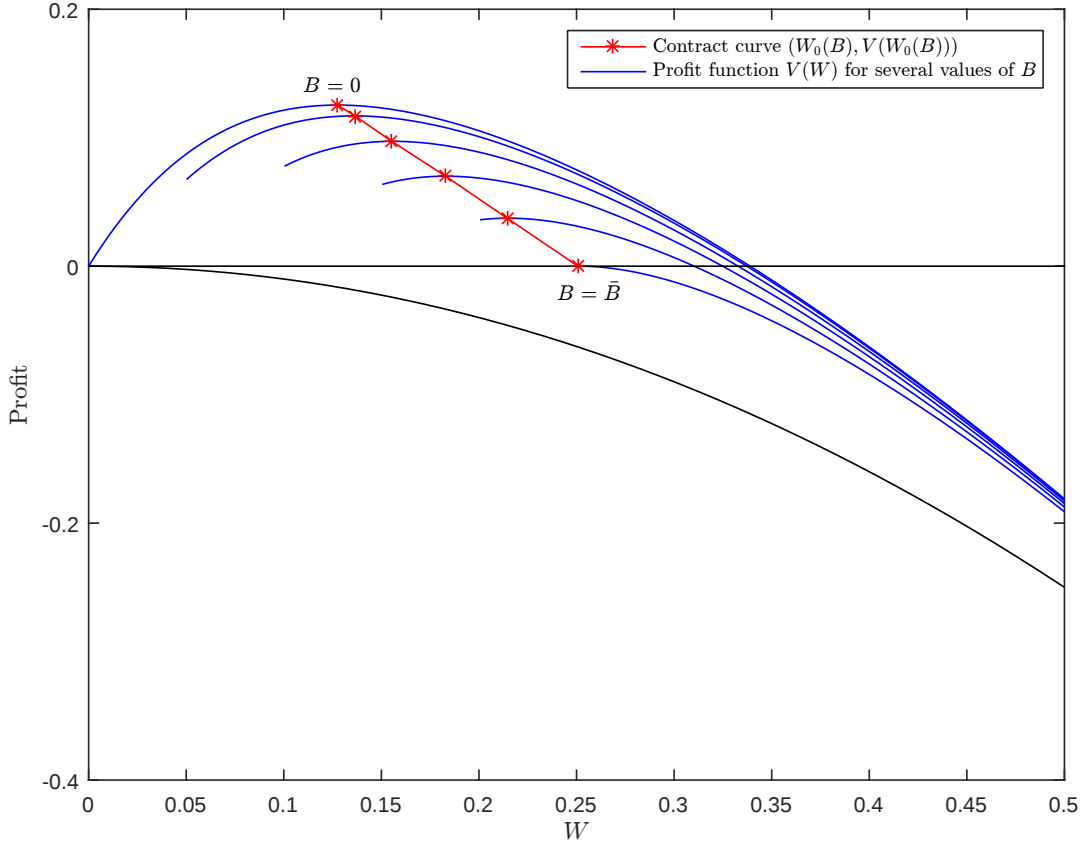


Figure 5: Comparative statics with respect to B .

5.2 Comparative statics

Next, we examine how the optimal contract depends on the value of the agent's outside option, B . In particular, we study how the agent's value from the contract, $W_0(B) := \arg\max_{B \leq W \leq W_{gp}(B)} V(W)$, and the firm's value, $V(W_0(B))$, depend on B .

Proposition 1 *For all $B \in [0, \bar{B}]$, $W_0(B)$ is strictly increasing in B but $W_0(B) - B$ is strictly decreasing in B . $V(W_0(B))$ is strictly decreasing in B , with $V(W_0(\bar{B})) = 0$.*

The agent's value net of his outside option, $W_0(B) - B$, represents his surplus from the contracting relationship, with the firm's surplus being $V(W_0(B))$. When the agent's outside option improves, the value he obtains from the contract, W_0 , increases. But the contracting relationship becomes less profitable as the support for the state variable, $[B, W_{gp}(B)]$, shrinks and the firm has less room to provide incentives. As a result, both the firm's surplus and the agent's surplus above his outside option are reduced. Figure 5 shows the inverse relationship between $W_0(B)$ and $V(W_0(B))$ graphically.

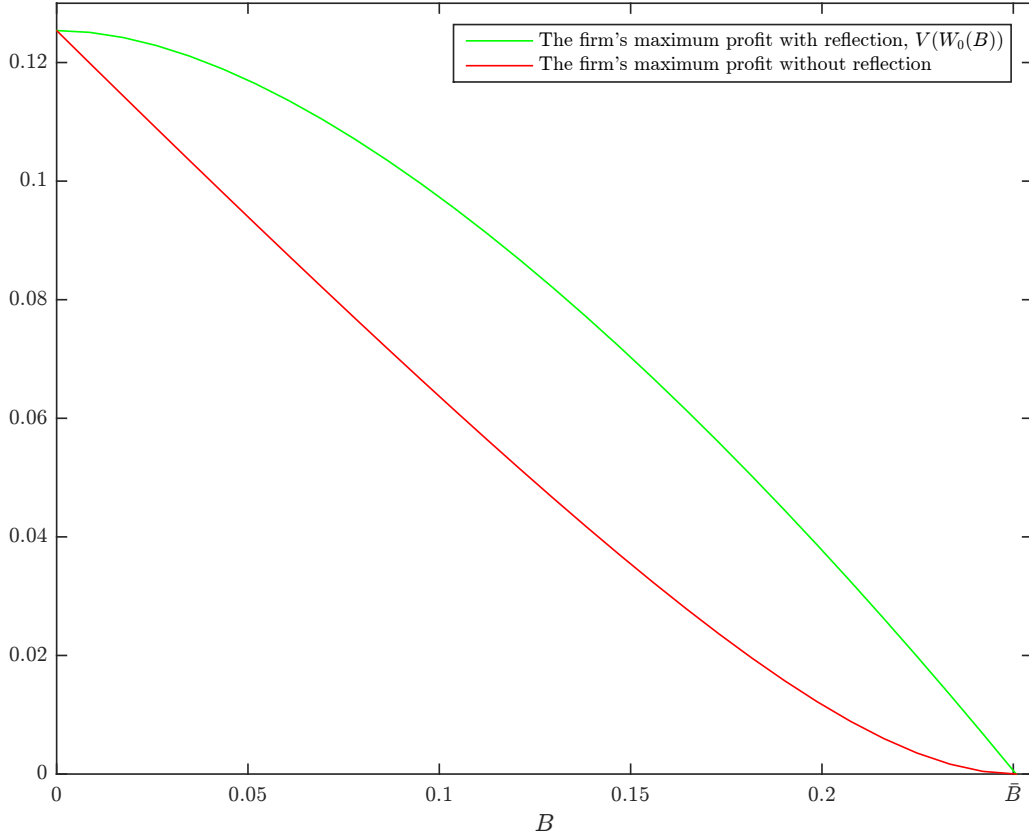


Figure 6: The peak of the firm's profit function with and without reflection, given any $B \in [0, \bar{B}]$.

5.3 The value of reflection

In this paper, we show that reflection at the lower bound $B > 0$ is more efficient than termination.¹⁰ With reflection, the added flexibility of being able to suspend the agent's effort temporarily allows the firm to move the contract out of the left corner of the state space, which lets it achieve higher profits.

However, the value of reflection is nonmonotonic in the agent's outside option B . Figure 6 plots the firm's maximum profit under the optimal contract with reflection against B , $V(W_0(B))$, alongside the maximum profit the firm can attain in the optimal contract without reflection. The firm's profit gain from using a contract with reflection is largest at intermediate B between 0 and $\bar{B}$.¹¹

Intuitively, reflection does not benefit the firm much if B is very small or very large. If B is

¹⁰If $B \geq \bar{B}$, it is optimal to not hire the agent at all, so the distinction between reflection and termination in a contract is not meaningful.

¹¹In relative terms, the gain is largest at B close to $\bar{B}$.

close to the minimax payoff of zero, the contract reflects off B very slowly, approximating the contract without reflection if B goes to zero, hence the firm's gain in profit due to reflection is small.¹² The firm's benefit of using the optimal contract with reflection is also small if B is close to $\bar{B}$. This is because the firm's surplus from the optimal contract is small when the agent's outside option is large, as we see in Proposition 1. If B is close to $\bar{B}$, therefore, the firm loses little by taking the continuation profit of zero and letting the agent take his outside option when W_t hits B , which is what the optimal contract without reflection calls for at that point.

At intermediate B , however, the gains from using the optimal contract with reflection are very significant. In the example presented in Figure 6, where $\bar{B} = 0.25$, the firm's profit nearly doubles at $B = 0.15$. Another example of significant gains is shown in Figure 1, where $B = 0.1$. Intuitively, at such intermediate values of B , the lower bound is large enough for reflection to be inexpensive to engineer, as W_t spends little time at B , and low enough for the firm to pass on the outside option as means of delivering utility to the agent.

6 Endogenous outside option

In this section, we endogenize the agent's outside option B by embedding the contracting problem in a simple model of the labor market similar to Phelan (1995). In a meeting between the firm and the agent, the firm designs the contract (C, A) and offers it to the agent as a take-it-or-leave-it proposal. If the agent rejects, he goes into the market, where he searches for a match with a new firm, while the old firm exits with the payoff of zero. While searching, the agent consumes zero and exerts zero effort.¹³ The new match arrives with Poisson intensity λ . The firm in the new match is identical to the firm in the previous match, i.e., it operates the same production technology and faces the same contract-design problem. If the agent accepts the contract, he is not committed to it, i.e., he can quit and go back to the market at any time.

Let W_0 denote the initial value that a new contract (C, A) delivers to the agent. Since the agent receives zero utility while searching, his post-separation utility comes exclusively from future matches, whose arrival time has density $\lambda e^{-\lambda t}$. The consistency condition between the agent's outside option B and the initial value W_0 therefore is

$$B = \int_0^\infty \lambda e^{-\lambda t} e^{-rt} W_0 dt = \frac{\lambda}{r + \lambda} W_0. \quad (10)$$

The firm takes B as given and solves the contracting problem so as to maximize its profit.

¹²Recall that with reflection the drift of W_t at B is rB .

¹³This normalizing assumption can easily be relaxed.

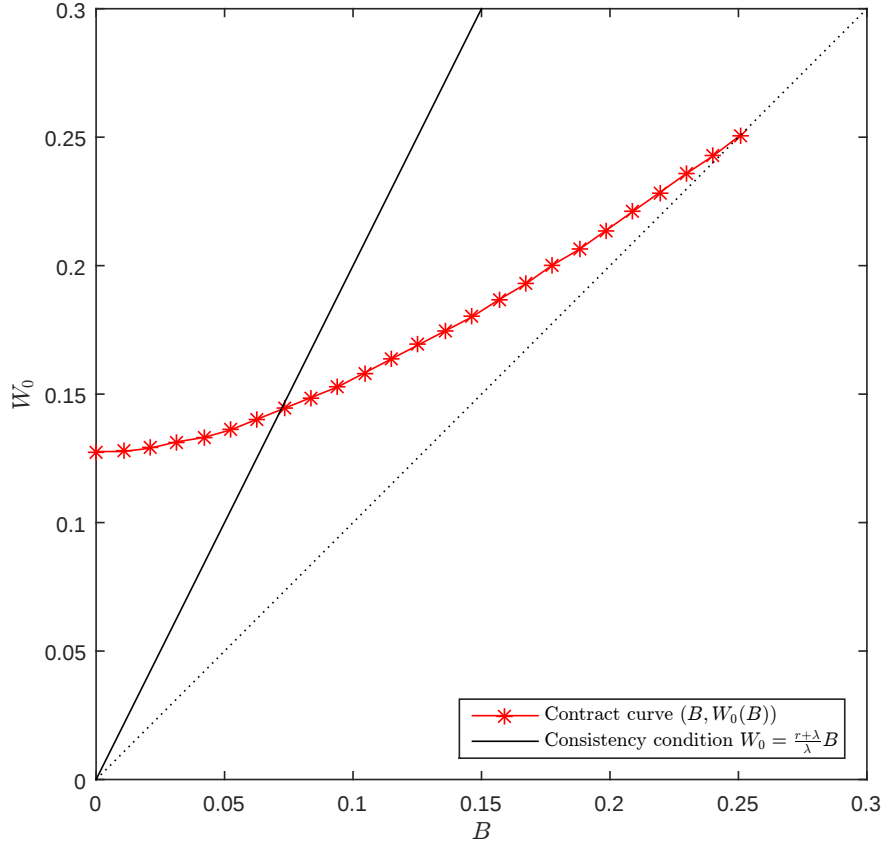


Figure 7: Equilibrium conditions.

Definition 1 Given $\lambda \geq 0$, competitive equilibrium consists of the agent's initial contract value $W_0 \geq 0$ and outside option $B \geq 0$ and of a function V such that 1) V is the firm's value function given the agent's outside option B , and 2) the consistency condition (10) holds with $W_0 \in \operatorname{argmax}_{B \leq W \leq W_{gp}(B)} V(W)$.

Proposition 2 For every $\lambda \in [0, \infty]$, there exists a unique competitive equilibrium (W_0, B, V) . The equilibrium values W_0 and B are strictly decreasing, but $V(W_0)$ is strictly increasing in the expected search time $1/\lambda$. If the expected search time is zero, then $W_0 = B = \bar{B}$ and $V(W_0) = 0$. If the expected search time is infinite, then $B = 0$, $W_0 = \operatorname{argmax}_{W \geq 0} F_{\max}(W)$ and $V(W_0) = \max_{W \geq 0} F_{\max}(W)$.

Proof of Proposition 2 follows from Figure 7. In that figure, the starred curve represents the relationship between the lower bound B and the agent's value W_0 delivered by the optimal contract, i.e., the function $W_0(B)$ studied in Proposition 1. We have $W_0(0) > 0$, $W_0(\bar{B}) = \bar{B}$, and the slope of $W_0(B)$ is everywhere smaller than one because the agent's surplus $W_0(B) - B$ is decreasing in B . The straight solid line represents the consistency condition $W_0 = \frac{r+\lambda}{\lambda} B$, the slope of which is larger than one. The unique intersection of these two lines determines the unique equilibrium.

Figure 7 also shows how the equilibrium values of B and W_0 depend on the (exogenous) new match arrival rate λ or, equivalently, on the expected search time $1/\lambda$. If the agent can get a new match immediately after quitting, i.e., if $\lambda = \infty$, then the consistency condition (10) coincides with the 45-degree line. The intersection with the contract curve determines the equilibrium value of B at its highest possible level, $\bar{B}$. With $W_0 = B = \bar{B}$, the firm makes zero profit in equilibrium, as $V(\bar{B}) = 0$. For all $0 < \lambda < \infty$, i.e., with the search cost $1/\lambda$ positive and finite, the slope of the equilibrium condition (10) is larger than one, which gives us equilibrium values of B and W_0 strictly between zero and $\bar{B}$.¹⁴ The lower bound B can in equilibrium be zero only if condition (10) is a vertical line, i.e., if the search cost $1/\lambda$ is infinite.

Since the equilibrium contract features termination of output and retirement of the agent at the lower bound only in the case of $B = 0$, our comparative statics analysis shows that such termination occurs only when the agent's search cost $1/\lambda$ is infinite. For all finite values of the search cost, the lower bound B determined in equilibrium is strictly positive, which implies that output is not terminated when $W_t = B$. Rather, the process W_t reflects off B and output is resumed.

7 Fast reflection

In this section, we assume that h has a bounded second derivative and that it satisfies the standard Inada condition $\lim_{a \rightarrow 0} h'(a) = 0$. We show that the dynamics of the reflection at the lower bound change. In particular, they become fast, meaning the process W_t spends zero time at B .

Assumption 1 $h(\cdot) \in C^3$, that is, its third derivative is continuous. $h(0) = h'(0) = 0$ and $0 < h''(0) < \infty$.

7.1 Boundary conditions

In this section, we derive boundary conditions for the high-action ODE (5) at an exogenous $B \geq 0$. As before, in order to obtain a contract that satisfies the domain restriction $W_t \geq B$ at all t , we will construct solutions to the HJB equation by splicing at B a high-action ODE solution F with a low-action ODE solution L . With $h'(0) = 0$, however, smooth pasting between L and F requires that the second derivative of F diverges to negative infinity as W approaches B .

¹⁴By Proposition 1, the firm's profit $V(W_0(B))$ is positive in this range.

Indeed, the value matching and smooth pasting conditions, $F(B) = L(B)$ and $F'(B) = L'(B)$, imply

$$\begin{aligned}
0 &= F(B) - L(B) \\
&= \left(\max_{c \geq 0, a \in \mathcal{A}} a - c + L'(B)(B - u(c) + h(a)) + \frac{1}{2}F''(B)r\sigma^2h'(a)^2 \right) \\
&\quad - \left(\max_{c \geq 0} -c + L'(B)(B - u(c)) \right) \\
&= \max_{a \in \mathcal{A}} a + L'(B)h(a) + \frac{1}{2}F''(B)r\sigma^2h'(a)^2.
\end{aligned}$$

Because $h'(0) = 0$, this condition cannot be met with a finite $F''(B)$. If $F''(B) > -\infty$, the objective under maximization attains the value of 0 at $a = 0$, but $a = 0$ is not a maximizer, which can be easily seen by differentiating the objective and evaluating the derivative at $a = 0$: $1 + L'(B)h'(a) + r\sigma^2F''(B)h'(a)h''(a)|_{a=0} = 1 > 0$. In order to satisfy the above smooth pasting condition, therefore, we must allow for $F''(B) = -\infty$.¹⁵

In sum, in order to satisfy the value matching and smooth pasting conditions at B , a solution F to the high-action ODE (5) must satisfy $F(B) = L(B)$, $F'(B) = L'(B)$, and $F''(B) = -\infty$. The singularity of F at B means that we can no longer invoke Lemma 1 of Sannikov (2008) for the existence and uniqueness of the solution curve F . In the Appendix, we present a change-of-variable technique that allows us to solve the high-action ODE forward from B despite this singularity.

7.2 The first-best benchmark

Because the firm and the agent discount at the same rate, the optimal first-best contract (i.e., with agent action a observable) is static: a and c are constant. The firm's first-best profit function is

$$F_{\text{fb}}(W) := \max_{c,a} \{ a - c : u(c) - h(a) = W \}.$$

Let us note that the first-best shut-down threshold W_{gp}^* does not exist with the Inada condition on h , because c that solves $u'(c) = h'(0)$ does not exist. That is, a is always strictly positive in the first best, even if W is large. Also note that $F_0 < F_{\text{fb}}$ because

$$F_0(W) = \max_{c,a} \{ a - c : u(c) - h(a) = W \text{ and } a = 0 \}$$

¹⁵Note that this problem does not arise in Section 5, where h' is bounded away from zero. Indeed, with $h'(a) \geq \gamma_0 > 0$ for all $a \in \mathcal{A}$ and with $a + L'(B)h(a)$ bounded on $\mathcal{A}$, a finite number $F''(B) < 0$ exists for which $\max_{a \in \mathcal{A}} a + L'(B)h(a) + \frac{1}{2}F''(B)r\sigma^2h'(a)^2 = 0$, i.e., the value matching and smooth pasting conditions can be met without an infinite $F''(B)$.

and the restriction $a = 0$ binds at all W . We will look for solutions F of the high-action ODE that satisfy $F_0(W) < F(W) < F_{fb}(W)$ for all $W \geq B$.

7.3 Classification of solution curves

For $B > 0$ and for two numbers y and y' , denote by $F_{(B,y,y',-\infty)}$ the solution to the high-action ODE (5) that satisfies boundary conditions $F(B) = y$, $F'(B) = y'$, and $F''(B) = -\infty$.

The next two lemmas adapt Lemma 1 and Lemma 2 to the case with $h'(0) = 0$. In particular, the optimal solution F to the high-action ODE is no longer pinned down by a tangency condition $F'(W_{gp}) = F'_0(W_{gp})$ at a finite $W_{gp} < W_{gp}^*$. Instead, the optimal F is pinned down by the asymptotic requirement $F_0(W) < F(W) < F_{fb}(W)$ for all $W \geq B$.

Lemma 3 (Largest lower bound) *There exists a unique $\bar{B} > 0$ for which the solution $F_{(\bar{B},0,0,-\infty)}$ satisfies*

$$F_0(W) < F_{(\bar{B},0,0,-\infty)}(W) < F_{fb}(W) \text{ for all } W \geq \bar{B}.$$

Lemma 4 (Largest initial level and slope) *For every $B \in (0, \bar{B}]$, there exists a unique y such that the solution $F_{(B,y,\frac{y}{B},-\infty)}$ satisfies*

$$F_0(W) < F_{(B,y,\frac{y}{B},-\infty)}(W) < F_{fb}(W) \text{ for all } W \geq B.$$

As before, we will denote the unique y pinned down in this lemma by $y^*(B)$.

7.4 Optimal contract

For each $B \in (0, \bar{B}]$, let us define a function $V : [B, \infty) \rightarrow \mathbb{R}$ by splicing at $W = B$ the high-action ODE solution $F_{(B,y^*(B),\frac{y^*(B)}{B},-\infty)}(W)$ with the low-action ODE solution $L(W) = \frac{y^*(B)}{B}W$. That is, let

$$V(W) := \begin{cases} L(W) & \text{for } W = B, \\ F_{(B,y^*(B),\frac{y^*(B)}{B},-\infty)}(W) & \text{for } W > B. \end{cases}$$

Theorem 2 *For each $B \in (0, \bar{B}]$, the function V is the firm's value function in the contracting problem with the agent's lower bound B . The optimal controls (c, a, Y) attaining V define an optimal contract $C_t = c(W_t)$, $A_t = a(W_t)$, where $\{W_t; 0 \leq t < \infty\}$ is a solution to (4). In particular, $c(B) = a(B) = Y(B) = 0$, with $dW_t = rBdt$ at $W_t = B$; and $a(W) > 0$ and $Y(W) > 0$ for all $W > B$.*

The above theorem verifies that the contract constructed from the low- and high-action ODE solutions spliced at B is indeed optimal. Note that the optimal contract never terminates, i.e., the agent is never retired. The low retirement point does not exist because the optimal contract reflects at B , as in Section 5. Moreover, the upper retirement point does not exist either because the optimal solution curve F stays above F_0 at all W .

7.5 The dynamics of reflection

When h satisfies the Inada condition $h'(0) = 0$, the process W_t is no longer Sticky Brownian Motion studied in Section 5 or in Zhu (2013). The key difference here is the volatility of W_t in the neighborhood of the lower bound B . In the no-Inada case in Section 5, Figure 4 shows that volatility of W_t is discontinuous at B , i.e., is zero at B but remains strictly positive in any neighborhood of B . Figure 8 shows drift and volatility of W_t in an example with h satisfying $h'(0) = 0$. As we see, volatility of W_t is continuous at B , i.e., becomes arbitrarily small in a sufficiently small neighborhood of B . This means that when W_t gets close to B , its drift remains positive while its volatility becomes extinguished. Intuitively, the closer W_t is to B , the more deterministic it becomes in its movement up and away from B . Thus, the reflection of W_t off B is faster in the Inada case than in the non-Inada case. This intuition is confirmed by the following result.

Proposition 3 *For any initial condition $W_0 \geq B$, the (Lebesgue) measure of time that the process W_t spends at B is zero almost surely.*

In Section 5, and in Zhu (2013), the continuation value process W_t spends a strictly positive amount of time at its lower bound, which defines slow reflection. Here, the reflection of W_t is faster, with W_t spending zero time at B . Small volatility of W_t in the neighborhood of B implies that W_t revisits B infrequently. With less frequent revisits to B , the total amount of time that W_t spends at B is smaller.

Fast reflection is optimal here because the firm is strongly averse to volatility in W_t near B . Recall that with the Inada condition on h we have $F''(B) = -\infty$. The firm's cost of volatility, $|F''(W_t)|$, therefore becomes large as W_t approaches B . It is thus optimal for the firm to keep volatility low near B , which implies infrequent returns of W_t to B and fast reflection.

Section 8.2 shows that fast reflection also occurs when the agent's consumption utility function, u , is unbounded below, even with marginal disutility of effort bounded away from zero, $h'(0) > 0$. The reason for fast reflection there is that low c has a strong impact on the drift of W_t , which makes reflection inexpensive to engineer.

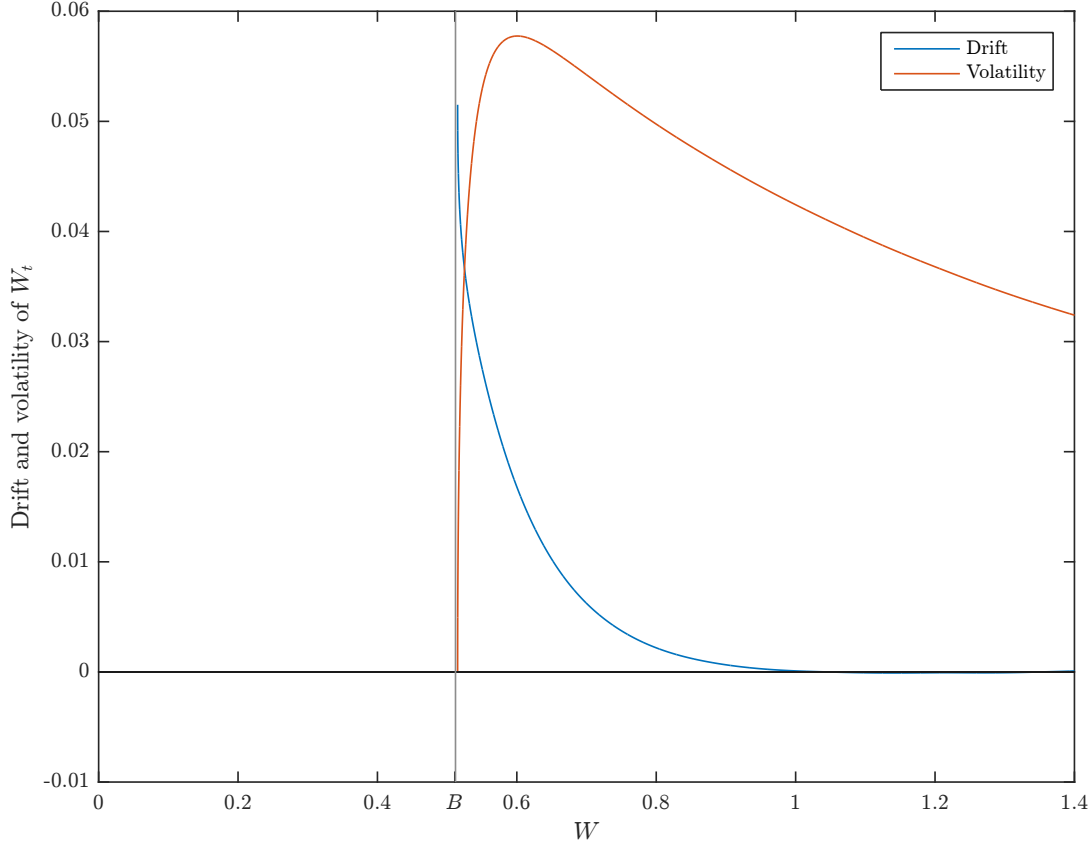


Figure 8: Drift and volatility of W_t with $h(a) = \frac{1}{2}a^2$. The lower bound is the largest equilibrium lower bound $\bar{B}$. An upper bound does not exist. Both drift and volatility are continuous on $[\bar{B}, \infty)$. Reflection is fast.

7.6 Remark on immiserization

In the dynamic contracting literature with moral hazard, starting with Thomas and Worrall (1990), typically $F'(W_t)$ is a martingale. This leads to the well-known result of immiserization, that is, the agent's continuation value converges to the lowest possible value almost surely. In our model, as in Phelan (1995), the immiserization result fails because the agent's continuation value is reflective at the lower bound. Yet the drift of $F'(W_t)$ is zero whenever $W_t > B$, i.e., $F'(W_t)$ continues to be a martingale at all but a single point in the support of W_t .¹⁶ The fact

¹⁶The drift and the volatility of W_t are, respectively, $r(W_t - u(C_t) + h(A_t))$ and $h'(A_t)r\sigma$. By Ito's lemma, the drift of $F'(W_t)$ can be calculated as

$$F''(W_t)r(W_t - u(C_t) + h(A_t)) + \frac{F'''(W_t)}{2}(h'(A_t)r\sigma)^2.$$

Differentiating the HJB equation with respect to W_t , we can show that $F''(W_t)r(W_t - u(C_t) + h(A_t)) + \frac{F'''(W_t)}{2}(h'(A_t)r\sigma)^2 = 0$. Cf. equation (12) in Sannikov (2008).

that $F'(W_t)$ is no longer a martingale when $W_t = B$ is key to eliminating the immiserization result. Because the state variable W_t has drift $rB > 0$ and zero volatility at B , by Ito's lemma, $F'(W_t)$ has drift $F''(B)rB < 0$ and zero volatility at B . One might be surprised that nonzero drift at just a single point, B , is strong enough to make the entire process $F'(W_t)$ fail to be a martingale.

The reason why this indeed is true is different in the case with and without the Inada condition on h . In Section 5, the subtlety lies in the slow reflection of Sticky Brownian Motion: even though B is a singleton, the process W_t spends a positive amount of time at B . In Section 7, where the Inada condition for h holds, W_t spends zero total amount of time at B , but the size of its drift at B , $|F''(B)rB|$, is infinity because $F''(B) = -\infty$. Infinite drift over a small time interval (of measure zero) has a nontrivial effect on the dynamics of the process.

8 Extensions

This section extends our analysis in two ways. First, we allow for limited commitment on the side of the firm. Second, we analyze the case with u unbounded below.

8.1 Two-sided limited commitment

In this section, we follow Phelan (1995) in introducing limited commitment also on the side of the firm. In particular, we allow the firm to get out of any contract upon incurring a deadweight cost of $K \geq 0$.¹⁷ This possibility adds another constraint to the contract-design problem:

$$F(W_t) \geq -K \text{ at all } t. \quad (11)$$

That is, a firm cannot credibly promise a contract that allows the firm's continuation value to drop below $-K$ in some state. In this section, we describe the solution to this contract-design problem. We treat B as exogenous, although it can be endogenized in the same way as in Section 6. As well, we return to the assumption $h'(0) = \gamma_0 > 0$, which gives us slow reflection, although it is also possible to study two-sided limited commitment with $h'(0) = 0$ and fast reflection.

The solution to this contracting problem is constructed as follows. At the lower bound $W = B$, we start solving the high-action ODE, in the direction of increasing W , with boundary conditions $F = y$ and $F' = y/B$ for some $y \geq 0$, as in Lemma 2. Recall that this choice of

¹⁷That is, if the firm pays K , it no longer has to meet its contractual obligation of delivering the continuation value W_t to the agent.

boundary conditions guarantees that the resulting solution, $F_{(B,y,\frac{y}{B})}$, pastes smoothly at B with a positively sloped solution to the low-action ODE (Case 1a in Section 4.1). We look for y such that $F_{(B,y,\frac{y}{B})}$ stays between F_0 and $F_{\max}$ until it drops down to the level $F_{(B,y,\frac{y}{B})}(D) = -K$ at some $D > B$.¹⁸ At $W = D$, we stop solving the high-action ODE and instead extend the solution as a straight line with slope $F'_{(B,y,\frac{y}{B})}(D) < 0$. This line can stay above F_0 , cut through F_0 , or become tangent to F_0 at some point $D' > D$. In this last case, the straight line is a Case-2 solution to the low-action ODE (see Section 4.1), and so D becomes a point of smooth pasting between the high-action ODE solution $F_{(B,y,\frac{y}{B})}$ and a low-action ODE solution tangent to F_0 at D' . Along this low-action solution, the agent's continuation value process has zero volatility and drift $W - u(c') + h(0)$, where $u(c') = D'$. For all $W < D'$, this drift is negative. In particular, it is negative at the splicing point D . With zero volatility and negative drift at D , the agent's continuation value process reflects downward off D , i.e., D becomes an upper bound for W_t .

Figure 9 provides one example of such a solution. Arguments similar to those in Theorem 1 can be used to verify that this solution indeed represents an optimal contract.

8.2 Unbounded utility

Thus far, our analysis of the optimal contract requires that u be bounded below. In particular, we have followed Sannikov (2008) in assuming that 0 is the lowest possible consumption level for the agent and $u(0) = 0$. This specification disallows many utility functions that are often used in applications, e.g., log or CRRA with the RRA coefficient larger than 1.

In this subsection, we allow for u unbounded below. In particular, we assume that $\lim_{c \downarrow 0} u(c) = -\infty$. We show that the optimal contract is reflective at an exogenous lower bound B .¹⁹ Further, reflection dynamics are fast, even with $h'(0) = \gamma_0 > 0$, and the firm's profit function is decreasing.

With $u(0) = -\infty$, by setting $c = a = Y = 0$, the firm can costlessly and instantaneously shift the agent's continuation utility upward. Indeed, with these controls, the expected flow of payoff to the firm is zero and the drift of the agent's utility is infinite. Since the firm can always apply this instantaneous control at the peak of its profit function, the process W_t will never go below the value at which the peak is attained, $W_0(B)$, which restricts the support

¹⁸If $F_{(B,y,\frac{y}{B})}$ becomes tangent to F_0 before dropping down to $-K$, then the point of tangency becomes $W_{gp}(B)$ from Lemma 2 and the firm's quitting constraint (11) does not bind. We will assume in this section that the firm's cost of breaking the contract, K , is low enough for this constraint to bind, i.e., $K < -F_0(W_{gp}(B))$. This assumption would be vacuous in the fast-reflection case with the Inada condition on h , where $W_{gp}(B)$ is infinite.

¹⁹The lower bound can be endogenized as in Section 6 but now under the assumption that the agent receives some positive flow of consumption while searching, to ensure $B > -\infty$.

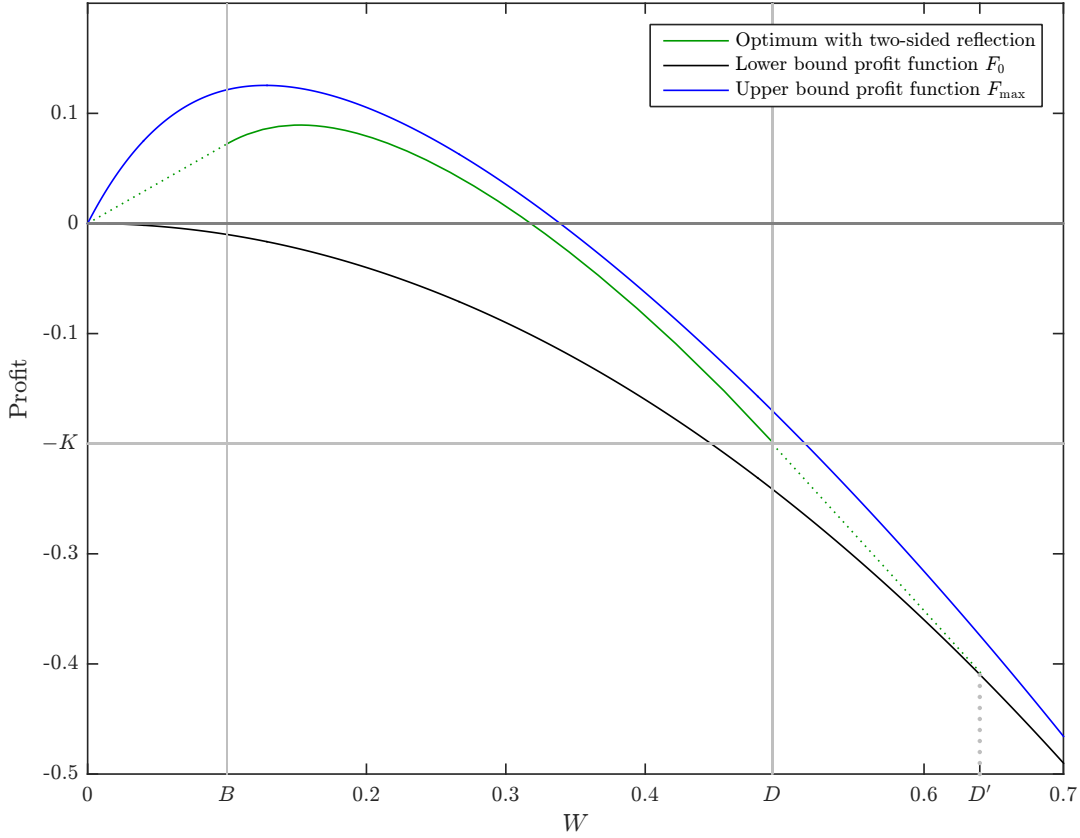


Figure 9: Optimal contract with two-sided limited commitment. The high-action ODE solution curve goes from the vertical line $W = B = 0.1$ to the horizontal line $F = -K = -0.2$. At both ends, it pastes smoothly with low-action ODE solutions. The higher splicing point $D = 0.49$ becomes a reflective upper bound for W_t .

of W_t to the values above $W_0(B)$. By the same logic as in Section 4.2, it is optimal to locate the peak of the profit function at the lower bound B , as this maximizes the support of W_t and allows the firm the most room to provide incentives and, hence, achieve highest profit. With the peak of F at B , $F'(B) = 0$ provides a boundary condition for the solution of the high-action ODE to the right of B .²⁰

²⁰Informally, we can think of the low-action ODE (7) as being solved to the left of B , with the solution having $c = 0$, infinite drift of W_t , and zero slope, so the low-action and high-action ODE solutions paste smoothly at B . Having infinite drift at B , the state variable W_t receives at that point a positive instantaneous shift familiar from the instantaneous control literature – see Stokey (2008). Since the firm’s cost of shifting W_t upward is zero, $F'(B) = 0$ satisfies the smooth pasting condition of the instantaneous-control problem, which requires that the marginal cost of shifting the state be equal to its marginal value. Instantaneous shifts of the state variable also occur in DeMarzo and Sannikov (2006) and Zhu (2013). There, the shifts are negative, carry a unit cost to the firm, and occur at the upper bound of the state variable, where the firm’s marginal profit equals negative one. Here, the shifts are positive, carry no cost to the firm, and occur at the lower bound of the state variable, where the firm’s marginal profit is zero.

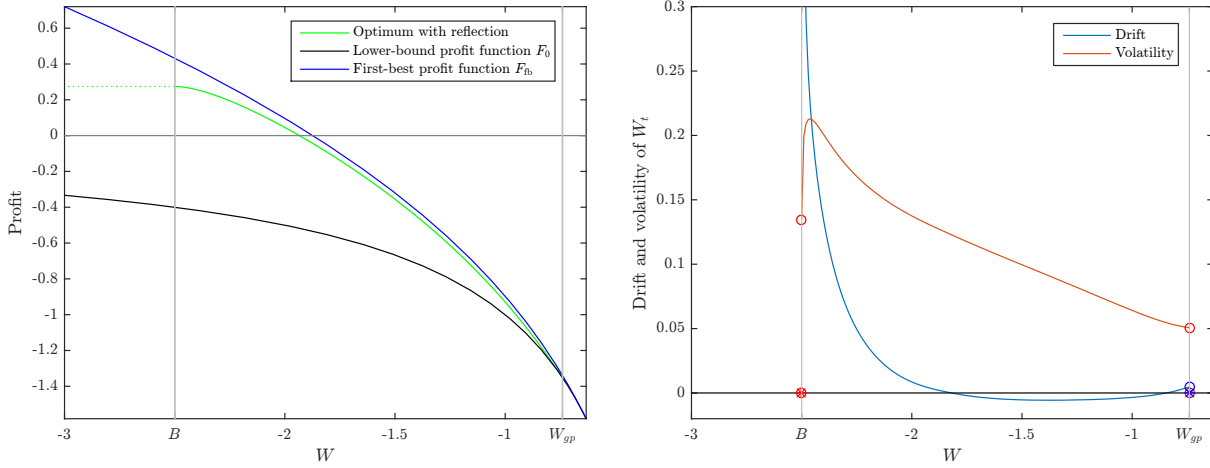


Figure 10: The case of unbounded u . Left panel: the firm's value function. Right panel: dynamics of the optimal contract.

The analog of Lemma 2 is as follows. At B , we start solving the high-action ODE with boundary conditions $F(B) = y \geq F_0(B)$ and $F'(B) = 0$.²¹ We search for an initial y such that the solution curve $F_{(B,y,0)}$ remains above F_0 and touches it at some $W_{gp} > B$. Since $F_{(B,y,0)}$ is strictly concave and $F'_{(B,y,0)}(B) = 0$, the firm's value function is strictly decreasing. This implies that the agent's compensation $c(W)$ is strictly positive everywhere outside of the lower bound $W = B$. As W_t approaches the lower bound, drift of W_t becomes unboundedly large, which implies that reflection of W_t off B is fast. Note the difference with Section 7, where fast reflection is due to volatility becoming small, rather than drift becoming large, as W_t approaches B . Figure 10 provides a computed example in which u is CRRA with relative risk aversion of 2.

9 Conclusion

We view this paper as making the following three contributions. First, we show that the optimal contract in the dynamic principal-agent model of Sannikov (2008) is reflective at the lower bound of the agent's continuation value process whenever this lower bound is greater than the agent's minimax payoff of zero. This means that the contract is never terminated (the agent is not fired or retired) after poor performance. Rather, the agent is temporarily asked to put in zero effort, which brings his continuation value up, and effort is resumed.

Second, we endogenize the lower bound of the agent's continuation value process by embedding

²¹Despite the singularity at $c = 0$, this solution can be advanced because $F'(W)u(c)$ remains bounded.

the contracting problem in a generalized version of the Phelan (1995) model of the labor market with one-sided commitment. In this model, the agent can quit at any time, so his outside option bounds his contract continuation value from below. We show that as long as the agent's search time for a new match is finite, this lower bound is strictly larger than the agent's minimax payoff, implying reflection at the lower bound in equilibrium.

Third, we find new dynamics of reflection of the agent's continuation value off its lower bound. We show that when the agent's disutility of effort satisfies the usual Inada condition (marginal disutility of the first unit of effort is zero), reflection off the lower bound is fast, i.e., the continuation value process spends zero total time at the lower bound. Another case of fast reflection obtains when the agent's utility function is unbounded below.

Appendix

Similar to Sannikov (2008), we can express the high-action ODE as

$$F''(W) = - \max_{a \in \mathcal{A}} \underbrace{\frac{a + F'h(a) - F + \max_c \{F'(W - u(c)) - c\}}{r\sigma^2(h'(a))^2/2}}_{H_a(W, F, F')}. \quad (12)$$

Lemma 1 in Sannikov (2008) shows that at each (W, F, F') a unique solution to (12) exists and is continuous in the initial conditions. Moreover, if the solution is strictly concave at one point, it is strictly concave everywhere.

The following lemma partially orders the solutions to (12).²²

Lemma A.1 *Consider two solutions F and $\tilde{F}$ to the high-action ODE that satisfy $F(W) \leq \tilde{F}(W)$ and $F'(W) \leq \tilde{F}'(W)$. If at least one of these inequalities is strict, then*

$$F'(W') < \tilde{F}'(W'), \quad \forall W' > W. \quad (13)$$

Proof This proof modifies the proof of Lemma 2 in Sannikov (2008). First, we show (13) in a small neighborhood of W . This holds trivially if $F'(W) < \tilde{F}'(W)$. If $F'(W) = \tilde{F}'(W)$ and $F(W) < \tilde{F}(W)$, then

$$F''(W) \leq -H_{\bar{a}}(W, F(W), F'(W)) < -H_{\bar{a}}(W, \tilde{F}(W), \tilde{F}'(W)) = \tilde{F}''(W).$$

²²This lemma generalizes Lemma 2 in Sannikov (2008) by including the comparison between F and $\tilde{F}$ such that $F'(W) = \tilde{F}'(W)$ and $F(W) < \tilde{F}(W)$. This additional flexibility will be useful for us, for example, in proving Lemma A.3.

where $\tilde{a}$ attains $\tilde{F}''(W)$ in (12). It follows from $F'(W) = \tilde{F}'(W)$ and $F''(W) < \tilde{F}''(W)$ that (13) holds in a small neighborhood of W .

Second, we show (13) for all $W' > W$ by contradiction. Suppose (13) does not hold, then there exists a smallest $\hat{W} > W$ at which $F'(\hat{W}) = \tilde{F}'(\hat{W})$. Since $F'(W') < \tilde{F}'(W')$ for all $W' \in (W, \hat{W})$, we have $F(\hat{W}) < \tilde{F}(\hat{W})$ and again

$$F''(\hat{W}) \leq -H_{\tilde{a}}(\hat{W}, F(\hat{W}), F'(\hat{W})) < -H_{\tilde{a}}(\hat{W}, \tilde{F}(\hat{W}), \tilde{F}'(\hat{W})) = \tilde{F}''(\hat{W}),$$

where $\tilde{a}$ attains $\tilde{F}''(\hat{W})$. It follows that $F'(\hat{W} - \epsilon) > \tilde{F}'(\hat{W} - \epsilon)$ for all sufficiently small $\epsilon > 0$, a contradiction. ■

Proof of Lemma 1

First, we show that, for each $B > 0$, the solution curve $F_{(B,0,0)}(W)$ is strictly concave. By Lemma 1 in Sannikov (2008), it is sufficient to show $F''_{(B,0,0)}(B) < 0$. Using $F_{(B,0,0)}(B) = F'_{(B,0,0)}(B) = 0$ in (12), we have

$$F''_{(B,0,0)}(B) = -\max_{a \in \mathcal{A}} \frac{a}{r\sigma^2(h'(a))^2/2} < 0.$$

Uniqueness: By contradiction, suppose that for some $B' < B$ both $F_{(B',0,0)}$ and $F_{(B,0,0)}$ satisfy conditions a) and b). Since $F_{(B',0,0)}$ is strictly concave, we have $F'_{(B',0,0)}(B) < 0 = F'_{(B,0,0)}(B)$ and $F_{(B',0,0)}(B) < 0 = F_{(B,0,0)}(B)$. Lemma A.1 thus implies that $F_{(B',0,0)}(W) < F_{(B,0,0)}(W)$ for all $W \geq B$. Since $F_{(B',0,0)}$ is weakly above F_0 , we have $F_0(W) \leq F_{(B',0,0)}(W) < F_{(B,0,0)}(W)$ for all $W \geq B$, i.e., the curve $F_{(B,0,0)}$ violates condition b), which is a contradiction.

Existence: Define

$$\bar{B} := \inf\{B \geq 0 : F_{(B,0,0)}(W) \geq F_0(W), \forall W \in [B, W_{gp}^*]\}.$$

The set $\{B \geq 0 : F_{(B,0,0)}(W) \geq F_0(W), \forall W \in [B, W_{gp}^*]\}$ is nonempty because $F_{(W_{gp}^*,0,0)}$ is above F_0 . Indeed, the domain of $F_{(W_{gp}^*,0,0)}$ is a singleton and $F_{(W_{gp}^*,0,0)}(W_{gp}^*) = 0 > F_0(W_{gp}^*)$. Because $F_{(\bar{B},0,0)}$ satisfies condition a) of the lemma directly from the definition of the above set, we only need to show that $F_{(\bar{B},0,0)}$ satisfies condition b) of the lemma.

If $\bar{B} = 0$, then we define $W_{gp} = 0$, as in Sannikov (2008). If $\bar{B} > 0$, then consider the sequence $\{\bar{B} - \frac{1}{n}\}_{n=1}^\infty$, which converges to $\bar{B}$ from the left. Since $\bar{B} - \frac{1}{n} \notin \{B \geq 0 : F_{(B,0,0)}(W) \geq F_0(W), \forall W \in [B, W_{gp}^*]\}$, there exists some $W_n \in [0, W_{gp}^*]$ such that $F_{(\bar{B}-\frac{1}{n},0,0)}(W_n) < F_0(W_n)$. Let W_{gp} be the limit of some subsequence $\{W_{n_k}\}_{k=1}^\infty$ of $\{W_n\}_{n=1}^\infty$, then $F_{(\bar{B},0,0)}(W_{gp}) = \lim_{k \rightarrow \infty} F_{(\bar{B}-\frac{1}{n_k},0,0)}(W_{n_k}) \leq \lim_{k \rightarrow \infty} F_0(W_{n_k}) = F_0(W_{gp})$. This and $F_{(\bar{B},0,0)}(W_{gp}) \geq F_0(W_{gp})$ from condition a) imply $F_{(\bar{B},0,0)}(W_{gp}) = F_0(W_{gp})$. We have $F'_{(\bar{B},0,0)}(W_{gp}) = F'_0(W_{gp})$ because

condition a) implies $(F_{(\bar{B},0,0)})'_+(W_{gp}) \geq (F_0)'_+(W_{gp})$ and $(F_{(\bar{B},0,0)})'_-(W_{gp}) \leq (F_0)'_-(W_{gp})$, where $(F)'_+$ and $(F)'_-$ denote the right and left derivatives of F . ■

Proof of Lemma 2

Uniqueness and existence arguments follow closely the steps in the proof of Lemma 1.

Uniqueness: By contradiction, suppose that for some $y_1 < y_2$ both $F_{(B,y_1,\frac{y_1}{B})}$ and $F_{(B,y_2,\frac{y_2}{B})}$ satisfy conditions a) and b). Lemma A.1 implies that $F_{(B,y_1,\frac{y_1}{B})}(W) < F_{(B,y_2,\frac{y_2}{B})}(W)$ for all $W \geq B$, because $F_{(B,y_1,\frac{y_1}{B})}(B) < F_{(B,y_2,\frac{y_2}{B})}(B)$ and $F'_{(B,y_1,\frac{y_1}{B})}(B) < F'_{(B,y_2,\frac{y_2}{B})}(B)$. Since $F_{(B,y_1,\frac{y_1}{B})}$ is weakly above F_0 , we have $F_0(W) \leq F_{(B,y_1,\frac{y_1}{B})}(W) < F_{(B,y_2,\frac{y_2}{B})}(W)$ for all $W \geq B$. This contradicts condition b) for $F_{(B,y_2,\frac{y_2}{B})}(W_{gp})$.

Existence: Define

$$y := \inf\{x \geq 0 : F_{(B,x,\frac{x}{B})}(W) \geq F_0(W), \forall W \in [B, W_{gp}^*]\}.$$

The above set is nonempty because for each $B \in (0, \bar{B}]$ we have $F_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(W) > F_0(W), \forall W \in [B, W_{gp}^*]$. Indeed, strict concavity of $F_{\max}$ implies $F'_{\max}(B) < \frac{F_{\max}(B) - F_{\max}(0)}{B - 0} = \frac{F_{\max}(B)}{B} = F'_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(B)$, and $F_{\max}(W) < F_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(W), \forall W > B$ follows from Lemma A.1. Because $F_{\max}(W) \geq F_0(W)$ for all $W \leq W_{gp}^*$, $F_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(W) > F_0(W)$ for all $W \in [B, W_{gp}^*]$, so the set is nonempty and $y < F_{\max}(B)$.

As $F_{(B,y,\frac{y}{B})}$ satisfies condition a) of the lemma, we only need to show that it satisfies condition b).

If $y = 0$, then the definition of $\bar{B}$ and the fact that $F_{(B,0,0)}$ satisfies condition a) imply that $B \geq \bar{B}$. The assumption $B \leq \bar{B}$ implies then that $B = \bar{B}$. The conclusion follows because Lemma 1 has shown that $F_{(\bar{B},0,0)}$ satisfies condition b). If $y > 0$, then consider the sequence $\{y - \frac{1}{n}\}_{n=1}^\infty$, which converges to y from below. Since $y - \frac{1}{n} \notin \{x \geq 0 : F_{(B,x,\frac{x}{B})}(W) \geq F_0(W), \forall W \in [B, W_{gp}^*]\}$, for each n there exists some $W_n \in [B, W_{gp}^*]$ such that $F_{(B,y-\frac{1}{n},\frac{y-\frac{1}{n}}{B})}(W_n) < F_0(W_n)$. Let W_{gp} be the limit of some subsequence $\{W_{n_k}\}_{k=1}^\infty$ of $\{W_n\}_{n=1}^\infty$, then $F_{(B,y,\frac{y}{B})}(W_{gp}) = \lim_{k \rightarrow \infty} F_{(B,y-\frac{1}{n_k},\frac{y-\frac{1}{n_k}}{B})}(W_{n_k}) \leq \lim_{k \rightarrow \infty} F_0(W_{n_k}) = F_0(W_{gp})$.

This and $F_{(B,y,\frac{y}{B})}(W_{gp}) \geq F_0(W_{gp})$ from condition a) imply $F_{(B,y,\frac{y}{B})}(W_{gp}) = F_0(W_{gp})$. We have $F'_{(B,y,\frac{y}{B})}(W_{gp}) = F'_0(W_{gp})$ because condition a) implies $(F_{(B,y,\frac{y}{B})})'_+(W_{gp}) \geq (F_0)'_+(W_{gp})$ and $(F_{(B,y,\frac{y}{B})})'_-(W_{gp}) \leq (F_0)'_-(W_{gp})$. ■

The next auxiliary lemma confirms that $a = 0$ is only used when the contract is determined by the solution to the low-action ODE. This lemma will also be useful in verification proofs.

Lemma A.2 Take $B \in [0, \bar{B}]$, and denote the solution $F_{(B,y^*(B),\frac{y^*(B)}{B})}$ simply by F . Then

1. F is strictly concave: $F''(W) < 0$ for all $W \geq B$.

2. F satisfies

$$\min_{c \geq 0} F(W) + c + F'(W)(u(c) - W) \geq 0 \quad \text{at all } W \geq B. \quad (14)$$

3. The optimal action a^* is nonzero everywhere along the high-action ODE solution F .

Proof

1. By Lemma 1 in Sannikov (2008), it is sufficient to show $F''(B) < 0$. By smooth pasting with a low-action ODE at B , we have $-F(B) + \max_c \{F'(B)(B - u(c)) - c\} = 0$ and $F'(B) \geq 0$, which imply

$$F''(B) = -\max_{a \in \mathcal{A}} \frac{a + F'(B)h(a)}{r\sigma^2(h'(a))^2/2} < 0.$$

2. We first show that all tangent lines to F are weakly above F_0 , i.e., for all $W \geq B$

$$F(W) + F'(W)(W' - W) \geq F_0(W') \quad \forall W' \geq 0. \quad (15)$$

If $W' \geq B$, then concavity of F implies $F(W) + F'(W)(W' - W) \geq F(W')$, which is above $F_0(W')$ because F satisfies condition a) of Lemma 2. If $W' < B$, then concavity of F implies $F(W) + F'(W)(W' - W) = F(W) + F'(W)(B - W) + F'(W)(W' - B) \geq F(B) + F'(W)(W' - B) \geq F(B) + F'(B)(W' - B) = F(B) + \frac{F(B)}{B}(W' - B) = \frac{W'}{B}F(B) \geq 0$, which is above $F_0(W')$ because $0 \geq F_0$. Inequality (14) follows now from (15) by changing the variable $W' \in [0, \infty)$ to $u(c) \in [0, \infty)$, where $-c = F_0(W')$.

3. It follows from

$$-\frac{a^* + F'h(a^*) - F + \max_c \{F'(W - u(c)) - c\}}{r\sigma^2(h'(a^*))^2/2} = F''(W) < 0$$

that $a^* + F'h(a^*) > \min_{c \geq 0} F + c + F'(u(c) - W) \geq 0$. This implies $a^* \neq 0$.

■

Proof of Theorem 1

First, we show that the profit achieved by any incentive compatible contract (C, A) is at most $F_{(B, y^*(B), \frac{y^*(B)}{B})}(W_0(C, A))$. To simplify the notation, we will often drop the subscript in $F_{(B, y^*(B), \frac{y^*(B)}{B})}$ and refer to this solution simply as F . Denote the agent's continuation value by $W_t = W_t(C, A)$, which follows (1). As in Sannikov (2008), it is without loss of generality to only consider contracts such that $u'(C_t) \geq \gamma_0$ at all t , with which we have that (C_t, A_t) belongs

to the compact set $[0, (u')^{-1}(\gamma_0)] \times \mathcal{A}$ at all t . By Lemma 4 in Sannikov (2008), the profit is at most $F_0(W_0) \leq F(W_0)$ if $W_0 \geq W_{gp}^*$. If $W_0 \in [B, W_{gp}^*]$, define

$$G_t := r \int_0^t e^{-rt} (A_s - C_s) ds + e^{-rt} F(W_t). \quad (16)$$

By Ito's lemma, the drift of G_t is

$$re^{-rt} \left(A_t - C_t - F(W_t) + F'(W_t)(W_t - u(C_t) + h(A_t)) + r\sigma^2 Y_t^2 \frac{F''(W_t)}{2} \right).$$

Let us show that the drift of G_t is always nonpositive. If $A_t > 0$, then incentive compatibility requires $Y_t = h'(A_t)$. Then the fact that F solves the high-action ODE implies that the drift of G is nonpositive. If $A_t = 0$, then (14) and $F'' < 0$ imply that the drift of G_t is nonpositive.

It follows that G_t is a bounded supermartingale until the stopping time τ (possibly ∞) defined as the time when W_t reaches W_{gp}^* . At time τ , the firm's future profit is less than or equal to $F_0(W_{gp}^*) \leq F(W_{gp}^*)$. Therefore, the firm's expected profit at time 0 is less than or equal to

$$\mathbb{E} \left[\int_0^\tau e^{-rt} (A_t - C_t) dt + e^{-r\tau} F(W_{gp}^*) \right] = \mathbb{E}[G_\tau] \leq G_0 = F(W_0).$$

Second, we show that the contract (C, A) described in the statement of the theorem achieves profit $F(W_0)$ if $W_0 \in [B, W_{gp}]$. Existence of a weak solution to (4) follows from Engelbert and Peskir (2014). Defining G_t as in (16), but now specifically for the stated contract, we have from Ito's lemma that the drift of G_t is

$$re^{-rt} \left(A_t - C_t - F(W_t) + F'(W_t)(W_t - u(C_t) + h(A_t)) + r\sigma^2 h'(A_t)^2 \frac{F''(W_t)}{2} \right) \text{ if } W_t > B,$$

and

$$re^{-rt} (-C_t - F(W_t) + F'(W_t)(W_t - u(C_t))) \text{ if } W_t = B.$$

Given the construction of $F_{(B, y^*(B), \frac{y^*(B)}{B})}$, the drift of G_t is zero in both cases. It follows that G_t is a bounded martingale until the stopping time τ' (possibly ∞) when W_t reaches W_{gp} . At time τ' , the firm's future profit is equal to $F_0(W_{gp}) = F(W_{gp})$. Therefore, the firm's expected profit at time 0 is equal to

$$\mathbb{E} \left[\int_0^{\tau'} e^{-rt} (A_t - C_t) dt + e^{-r\tau'} F(W_{gp}) \right] = \mathbb{E}[G_{\tau'}] = G_0 = F(W_0).$$

■

Proof of Proposition 1 (comparative statics)

Define $\mathcal{X}$ as the collection of points (W, V) starting from which the solution $F_{(W, V, 0)}$ stays above F_0 and touches it. That is, $\mathcal{X}$ is the set of possible peak points of the firm's value function in equilibrium. Precisely:

$$\begin{aligned} \mathcal{X} := \{ & (W, V) : V \geq 0, \text{ the solution } F_{(W, V, 0)} \text{ satisfies two conditions:} \\ & \text{a) } F_{(W, V, 0)}(W') \geq F_0(W') \text{ for all } W' \in [W, W_{gp}^*], \\ & \text{b) } F_{(W, V, 0)}(W_{gp}) = F_0(W_{gp}) \text{ and} \\ & \quad F'_{(W, V, 0)}(W_{gp}) = F'_0(W_{gp}) \text{ for some } W_{gp} \in [W, W_{gp}^*] \}. \end{aligned}$$

Note that solution curves $F_{(W, V, 0)}$ in the above definition are strictly concave. Indeed, if $F''_{(W, V, 0)}(W) \geq 0$, then by Lemma 1 in Sannikov (2008), $F_{(W, V, 0)}$ is a weakly convex function, which means its value must stay nonnegative and never reach F_0 . Thus, $F''_{(W, V, 0)}(W) < 0$. By Lemma 1 in Sannikov (2008) again, we have that $F_{(W, V, 0)}$ is a strictly concave function.

Next, we show that $\mathcal{X}$ is a strictly decreasing curve.

Lemma A.3 *Suppose $(W, V) \in \mathcal{X}$ and $(\tilde{W}, \tilde{V}) \in \mathcal{X}$. If $W = \tilde{W}$, then $V = \tilde{V}$. If $W < \tilde{W}$, then $V > \tilde{V}$.*

Proof We prove it by contradiction. If $W = \tilde{W}$, then suppose $V \neq \tilde{V}$. Assume without loss of generality that $V < \tilde{V}$. Lemma A.1 implies that $F_{(\tilde{W}, \tilde{V}, 0)}(W') > F_{(W, V, 0)}(W') \geq F_0(W')$, $\forall W' \geq W$, hence violating condition b) for $F_{(\tilde{W}, \tilde{V}, 0)}$.

If $W < \tilde{W}$, then suppose $V \leq \tilde{V}$. Then $F'_{(\tilde{W}, \tilde{V}, 0)}(\tilde{W}) = 0 = F'_{(W, V, 0)}(W) > F'_{(W, V, 0)}(\tilde{W})$ because $F_{(W, V, 0)}(\cdot)$ is strictly concave. Lemma A.1 again implies that $F_{(\tilde{W}, \tilde{V}, 0)}(W') > F_{(W, V, 0)}(W') \geq F_0(W')$, $\forall W' \geq \tilde{W}$, hence violating condition b) for $F_{(\tilde{W}, \tilde{V}, 0)}$. ■

Since for each W there is a unique V such that $(W, V) \in \mathcal{X}$, we will denote this unique V as $\mathcal{V}(W)$. Moreover, in this proof we will denote the solution curve $F_{(W, \mathcal{V}(W), 0)}$ as $\mathcal{F}_W$ to simplify the notation. To prove the proposition, we will use the following result.

Lemma A.4 *If $W > \tilde{W}$, then $\mathcal{F}'_W(W - x) > \mathcal{F}'_{\tilde{W}}(\tilde{W} - x)$ for any $x > 0$. $\mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x)$ is strictly decreasing in both W and x .*

Proof First, we show the inequality for small $x > 0$. It is sufficient to show that $\mathcal{F}''_W(W) < \mathcal{F}''_{\tilde{W}}(\tilde{W})$, which follows from (12) and $\mathcal{V}(W) < \mathcal{V}(\tilde{W})$.

Second, we show the inequality for any $x > 0$ by contradiction. Suppose the inequality does not hold: then there exists a smallest $x > 0$ at which $\mathcal{F}'_W(W - x) = \mathcal{F}'_{\tilde{W}}(\tilde{W} - x)$. Since $\mathcal{F}'_W(W - y) > \mathcal{F}'_{\tilde{W}}(\tilde{W} - y)$ for all $y \in (0, x)$ and $\mathcal{F}_W(W) < \mathcal{F}_{\tilde{W}}(\tilde{W})$, we have $\mathcal{F}_W(W - x) <$

$\mathcal{F}_{\tilde{W}}(\tilde{W} - x)$. This and $W > \tilde{W}$ imply

$$\mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x) < \mathcal{F}_{\tilde{W}}(\tilde{W} - x) - \mathcal{F}'_{\tilde{W}}(\tilde{W} - x)(\tilde{W} - x).$$

It follows from (12) again that $\mathcal{F}''_W(W - x) < \mathcal{F}''_{\tilde{W}}(\tilde{W} - x)$ and therefore $\mathcal{F}'_W(W - x + \epsilon) < \mathcal{F}'_{\tilde{W}}(\tilde{W} - x + \epsilon)$ for all sufficiently small $\epsilon > 0$, a contradiction.

Third, we show that $\mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x)$ is decreasing in W . Because

$$\begin{aligned} & \mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x) \\ &= \mathcal{F}_W(W) - \int_0^x \mathcal{F}'_W(W - y) dy - \mathcal{F}'_W(W - x)(W - x) \\ &= \mathcal{V}(W) - \int_0^W \min\{\mathcal{F}'_W(W - x), \mathcal{F}'_W(W - y)\} dy, \end{aligned}$$

the monotonicity of $\mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x)$ in W relies on the fact that $\mathcal{V}(W)$ is decreasing in W , and $\mathcal{F}'_W(W - x)$ and $\mathcal{F}'_W(W - y)$ are increasing in W . These properties have been shown above.

Finally, $\mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x)$ is decreasing in x because

$$\frac{\partial(\mathcal{F}_W(W - x) - \mathcal{F}'_W(W - x)(W - x))}{\partial x} = \mathcal{F}''_W(W - x)(W - x) < 0.$$

■

We can now prove Proposition 1.

Suppose $B > \tilde{B}$ and define $x := W_0(B) - B$, $\tilde{x} := W_0(\tilde{B}) - \tilde{B}$. Because both profit functions, $\mathcal{F}_{W_0(B)}$ and $\mathcal{F}_{W_0(\tilde{B})}$, satisfy the smooth pasting condition with a low-action ODE solution at their respective lower bounds, B and $\tilde{B}$, we have

$$\begin{aligned} 0 &= \mathcal{F}_{W_0(B)}(W_0(B) - x) - \mathcal{F}'_{W_0(B)}(W_0(B) - x)(W_0(B) - x) \\ &= \mathcal{F}_{W_0(\tilde{B})}(W_0(\tilde{B}) - \tilde{x}) - \mathcal{F}'_{W_0(\tilde{B})}(W_0(\tilde{B}) - \tilde{x})(W_0(\tilde{B}) - \tilde{x}). \end{aligned}$$

Lemma A.4 now implies that either $W_0(B) > W_0(\tilde{B})$ and $x < \tilde{x}$, or $W_0(B) \leq W_0(\tilde{B})$ and $x \geq \tilde{x}$. The latter case cannot occur because $W_0(B) - x = B > \tilde{B} = W_0(\tilde{B}) - \tilde{x}$. Thus, $W_0(B)$ is strictly increasing and $W_0(B) - B$ is strictly decreasing in B . That $V(W_0(B))$ is strictly decreasing in B follows from the fact that function $\mathcal{V}$ is strictly decreasing, which was shown in Lemma A.3. ■

Proofs for Section 7

Writing the high-action ODE (12) as a system of first-order equations, we have

$$\frac{dF}{dW} = F', \quad (17)$$

$$\frac{dF'}{dW} = -\max_{a \in \mathcal{A}} \underbrace{\frac{a + F'h(a) - F + \max_c \{F'(W - u(c)) - c\}}{r\sigma^2(h'(a))^2/2}}_{H_a(W, F, F')}. \quad (18)$$

As we discuss in Section 7.1, F'' must be $-\infty$ and effort a must be 0 at the lower bound $W = B$. The system (17)-(18) is therefore singular at $W = B$. We now use a change-of-variable technique to obtain an alternative system that is equivalent to (17)-(18) and well-behaved in the neighborhood of $W = B$.

Change of variable

Dependent variable S . Let us define

$$S := F(W) - \max_{c \geq 0} \{F'(W)(W - u(c)) - c\}.$$

By this definition, S is the difference between $F(W)$, which solves the HJB equation (3), and the solution $L(W)$ to the low-action ODE (7) that satisfies the smooth pasting condition $L'(W) = F'(W)$. As such, S represents the firm's value of the option to induce positive effort from the agent at W . Because the firm always has the option to ask for zero effort, S is always nonnegative.

Using S , the HJB equation (3) can be written as

$$S = \max_{a \in \mathcal{A}} \{a + F'(W)h(a) + \frac{1}{2}F''(W)r\sigma^2h'(a)^2\}, \quad (19)$$

which shows that the firm's value from the agent's effort comes from the expected output produced by the agent, a , less the cost of compensating the agent for his disutility of effort $h(a)$, less the firm's cost of having to induce volatility $Y = h'(a)$ in the state variable.

Further, we can express the HJB equation (19) as

$$F'' = -\max_{a \in \mathcal{A}} \frac{a + F'h(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}. \quad (20)$$

Independent variable X . Now, instead of treating W as the independent variable and F and F' as dependent variables, we change the independent variable to

$$X := -F',$$

and treat W and S as dependent variables.

The dynamics of W and S in terms of X are as follows. From $dX/dW = -F''$ we have

$$\frac{dW}{dX} = \frac{-1}{F''}, \quad (21)$$

and

$$\begin{aligned} \frac{dS}{dX} &= \frac{d}{dX} \left(F + XW - \max_{c \geq 0} \{ Xu(c) - c \} \right) \\ &= \frac{dF}{dW} \frac{dW}{dX} + W + X \frac{dW}{dX} - u(c(X)) \\ &= -X \frac{dW}{dX} + W + X \frac{dW}{dX} - u(c(X)) \\ &= W - u(c(X)), \end{aligned}$$

where $c(X) = \arg \max_{c \geq 0} \{ Xu(c) - c \}$. To simplify the notation, we will write $u(c(X))$ as $U(X)$.²³

Using the HJB equation (20) to eliminate F'' from (21), we get the following system:

$$\frac{dS}{dX} = W - U(X), \quad (22)$$

$$\frac{dW}{dX} = \frac{1}{\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}}. \quad (23)$$

In the new variables $(X, W(X), S(X))$, the initial conditions at the lower bound B are as follows. The lower end of the domain for X , which we will denote by $\underline{X}$, is unknown. Since $\underline{X} = -F'(B)$, this value is to be determined by forward shooting. We know that $\underline{X} \leq 0$, with equality only if $B = \bar{B}$. Despite not knowing $\underline{X}$, we know that $W(\underline{X}) = B > 0$, and $S(\underline{X}) = F(B) - F'(B)B = F - \frac{F}{B}B = 0$, where the first equality follows from $F'(B) \geq 0$ and the second from smooth pasting with a low-action ODE at B .

Regularity of the transformed system

The continuity of the right-hand side of (23) in (X, S) depends on whether $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}$ is equal to 0. Define $\psi(X) := \max_{a \in \mathcal{A}} a - Xh(a)$.²⁴ There are four possibilities for the value of $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}$:

1. If $S \leq 0$, then $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2} \geq \max_{a \in \mathcal{A}} \frac{a - Xh(a)}{\frac{1}{2}r\sigma^2(h'(a))^2} = \infty$.

²³Note that $U(X) = \arg \max_{U \geq 0} \{ XU + F_0(U) \}$. Clearly, $U(X) = 0$ for $X \leq 0$, and $U(X)$ solves $X + F'_0(U) = 0$ for $X > 0$.

²⁴Note that $\psi(X) = \bar{A} - Xh(\bar{A})$ if $X \leq 0$, and $\psi(X) = \bar{a} - Xh(\bar{a})$ if $X > 0$, where $\bar{a}$ satisfies $1 = Xh'(\bar{a})$.

2. If $S \in (0, \psi(X))$, then $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2} > 0$ is positive and finite.
3. If $S = \psi(X)$, then $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2} = 0$.
4. If $S > \psi(X)$, then $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}$ is negative and finite.

Although solutions to (23) may exist outside of $[0, \psi(X)]$, when we construct an optimal contract we only consider solutions that satisfy $S \in [0, \psi(X)]$. We do so because $S < 0$ implies that the requirement of a nonnegative option value of inducing effort is violated, while $S > \psi(X)$ implies $F'' > 0$, i.e., the solution curve is no longer concave.

Next we show that the transformed system (22)-(23) is Lipschitz-continuous at all $S \in [0, \psi(X))$ and the original system (17)-(18) is Lipschitz-continuous at $S = \psi(X)$.

Lemma A.5 *Fix a point $(\hat{X}, \hat{W}, \hat{S})$ in the domain of the ODE system (22)-(23).*

1. *If $\hat{S} \in [0, \psi(\hat{X}))$, then the ODE system in (22)-(23) satisfies the Lipschitz condition in a neighborhood of $(\hat{X}, \hat{W}, \hat{S})$. This system violates the Lipschitz condition if $\hat{S} = \psi(\hat{X})$.*
2. *If $\hat{S} = \psi(\hat{X})$, then the ODE system in (17)-(18) satisfies the Lipschitz condition in a neighborhood of $(\hat{W}, \hat{F}, \hat{F}')$, where $\hat{F} = \hat{S} + \max_{U \geq 0} \{F_0(U) - \hat{X}(\hat{W} - U)\}$ and $\hat{F}' = -\hat{X}$.*

Proof

1. In this proof, we denote $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{(h'(a))^2}$ by $M(X, S)$. Pick a small $\epsilon > 0$ such that $S < \psi(X)$ for all $(X, S) \in [\hat{X} - \epsilon, \hat{X} + \epsilon] \times [\hat{S} - \epsilon, \hat{S} + \epsilon]$. Lipschitz continuity of (23) in this neighborhood requires that $|\frac{1}{M(X, S_2)} - \frac{1}{M(X, S_1)}| \leq K(S_1 - S_2)$ for all $0 \leq S_1 < S_2$ in $[\hat{S} - \epsilon, \hat{S} + \epsilon]$ and for some $K > 0$. We do not consider negative values of S_1 or S_2 because $\frac{1}{M(X, S)} = 0$ for all $S \leq 0$. The rest of this proof relies on the following equation,

$$\frac{1}{M(X, S_2)} - \frac{1}{M(X, S_1)} = \int_{S_1}^{S_2} \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2} dS, \quad (24)$$

where $a^*(X, S)$ is the maximizer in $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{(h'(a))^2}$. We can understand (24) heuristically by the envelope theorem, because

$$\frac{\partial}{\partial S} \frac{1}{M(X, S)} = \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2}.$$

The technical issue here is that $M(X, S)$ is nondifferentiable when the maximizer $a^*(X, S)$ is not unique. To deal with this issue, we justify (24) at the end of this proof using a generalized envelope theorem from Milgrom and Segal (2002). Here we shall proceed assuming (24) is correct.

To show Lipschitz continuity, it is sufficient to show that the integrand in (24) is bounded when $(X, S) \in [\hat{X} - \epsilon, \hat{X} + \epsilon] \times [\hat{S} - \epsilon, \hat{S} + \epsilon]$. We show the boundedness as follows.

Because $\lim_{a \rightarrow 0} h'(a) = 0$, we can pick a small $\delta > 0$ such that for all $a \in (0, \delta)$ and $X \in [\hat{X} - \epsilon, \hat{X} + \epsilon]$,

$$1 - Xh'(a) > \frac{1}{2}. \quad (25)$$

First, if $a^*(X, S) \geq \delta$, then

$$\begin{aligned} \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2} &\leq \frac{1}{M(X, S)^2} \frac{1}{(h'(\delta))^2} \\ &\leq \frac{1}{(\max_{a \in \mathcal{A}} \frac{a - (\hat{X} + \epsilon)h(a) - (\hat{S} + \epsilon)}{(h'(a))^2})^2} \frac{1}{(h'(\delta))^2}, \end{aligned}$$

where the first inequality follows from $a^*(X, S) \geq \delta$, and the second inequality follows from the fact that $M(X, S)$ is decreasing in both S and X .

Second, suppose $a^*(X, S) \in (0, \delta)$. The first-order condition for an interior a^* is

$$\frac{(1 - Xh'(a^*))h'(a^*)^2 - (a^* - Xh(a^*) - S)2h'(a^*)h''(a^*)}{h'(a^*)^4} = 0,$$

which implies

$$a^* - Xh(a^*) - S = \frac{(1 - Xh'(a^*))h'(a^*)}{2h''(a^*)} > \frac{h'(a^*)}{4h''(a^*)},$$

where the inequality is from (25). Therefore, the integrand in (24) satisfies

$$\frac{1}{M(X, S)^2} \frac{1}{(h'(a^*))^2} = \frac{(h'(a^*))^2}{(a^* - Xh(a^*) - S)^2} \leq \frac{(h'(a^*))^2}{\left(\frac{h'(a^*)}{4h''(a^*)}\right)^2} = 16(h''(a^*))^2.$$

The last term, $16(h''(a^*))^2$, is bounded by $\max_{a \in [0, \delta]} 16(h''(a))^2$.

Finally, we prove (24) accounting for the fact that $a^*(X, S)$ may not be unique and $M(X, S)$ is not differentiable everywhere (with respect to S). In this case, $a^*(X, S)$ is any selection from $\{a : \frac{a - Xh(a) - S}{(h'(a))^2} = M(X, S)\}$. First, we consider the case of $S_1 > 0$. Because $\frac{a - Xh(a) - S}{(h'(a))^2}$ is supermodular in (a, S) , $a^*(X, S)$ is increasing in S , which implies

$$M(X, S) = \max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{(h'(a))^2} = \max_{a \geq a^*(X, S_1)} \frac{a - Xh(a) - S}{(h'(a))^2}, \quad \forall S \geq S_1.$$

Because $\frac{\partial}{\partial S} \frac{a - Xh(a) - S}{(h'(a))^2} = \frac{-1}{(h'(a))^2}$ is bounded and continuous on the compact interval $[a^*(X, S_1), A]$, Corollary 4 in Milgrom and Segal (2002) implies that $M(X, S)$ is absolutely continuous in S , and $\partial M(X, S)/\partial S$ exists and equals $\frac{-1}{(h'(a^*(X, S)))^2}$ almost everywhere. Because $M(X, S)$ is bounded away from zero when $S \in [S_1, S_2]$ (i.e., $M(X, S) \geq M(X, S_2) > 0$), $\frac{1}{M(X, S)}$ is also absolutely continuous in S with derivative $\frac{\partial}{\partial S} \frac{1}{M(X, S)} = \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2}$ almost everywhere. Absolute continuity of $\frac{1}{M(X, S)}$ implies that

$$\frac{1}{M(X, S_2)} - \frac{1}{M(X, S_1)} = \int_{S_1}^{S_2} \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2} dS. \quad (26)$$

Second, we consider the case of $S_1 = 0$. Taking limit $S_1 \downarrow 0$ in (26) yields

$$\frac{1}{M(X, S_2)} - \frac{1}{M(X, 0)} = \int_0^{S_2} \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2} dS.$$

2. To show that (18) satisfies the Lipschitz condition, we first show that the optimal $a^*(W, F, F')$ in $H_a(W, F, F')$ is uniformly bounded away from zero whenever (W, F, F') is in a neighborhood of $(\hat{W}, \hat{F}, \hat{F}')$. That is, $\max_{a \in \mathcal{A}} H_a(W, F, F') = \max_{a \geq \epsilon} H_a(W, F, F')$ for some $\epsilon > 0$. It follows from $\hat{S} = \psi(\hat{X}) > 0$ that $\hat{a}^* + \hat{F}'h(\hat{a}^*) - \hat{S} = 0$, where $\hat{a}^*$ stands for $a^*(\hat{W}, \hat{F}, \hat{F}') > 0$. Continuity and $\hat{a}^* + \hat{F}'h(\hat{a}^*) - \hat{S} = 0 > 0 + \hat{F}'h(0) - \hat{S}$ imply that there exists a small $\epsilon > 0$ such that if $(W, F, F') \in (\hat{W} - \epsilon, \hat{W} + \epsilon) \times (\hat{F} - \epsilon, \hat{F} + \epsilon) \times (\hat{F}' - \epsilon, \hat{F}' + \epsilon)$, then

$$\frac{\hat{a}^* + F'h(\hat{a}^*) - S}{r\sigma^2(h'(\hat{a}^*))^2/2} > \frac{a + F'h(a) - S}{r\sigma^2(h'(a))^2/2}, \quad \forall a \in [0, \epsilon).$$

This means $a \in [0, \epsilon)$ cannot be the optimal effort at (W, F, F') as it is dominated by $\hat{a}^*$. Therefore, $\max_{a \in \mathcal{A}} H_a(W, F, F') = \max_{a \geq \epsilon} H_a(W, F, F')$ whenever $(W, F, F') \in (\hat{W} - \epsilon, \hat{W} + \epsilon) \times (\hat{F} - \epsilon, \hat{F} + \epsilon) \times (\hat{F}' - \epsilon, \hat{F}' + \epsilon)$.

Second, function $H_a(W, F, F')$ is differentiable in (F, F') , with

$$\begin{aligned} \frac{\partial H_a(W, F, F')}{\partial F} &= \frac{-1}{r\sigma^2(h'(a))^2/2}, \\ \frac{\partial H_a(W, F, F')}{\partial F'} &= \frac{W - u(c) + h(a)}{r\sigma^2(h'(a))^2/2}. \end{aligned}$$

These derivatives are uniformly bounded over $(a, W, F') \in [\epsilon, A] \times (\hat{W} - \epsilon, \hat{W} + \epsilon) \times (\hat{F}' - \epsilon, \hat{F}' + \epsilon)$. Therefore, $\max_{a \geq \epsilon} H_a(W, F, F')$ is Lipschitz continuous in (F, F') .

■

The Lipschitz continuity shown in Lemma A.5 allows us to solve (22)-(23) in a neighborhood of $S = 0$, and solve (17)-(18) in a neighborhood of $S = \psi(X)$. If $S \in (0, \psi(X))$, we can solve either (17)-(18) or (22)-(23) because they are equivalent.

The next lemma lists basic properties of solutions to (22)-(23), which will be used in proving Lemmas 3 and 4 below.

Lemma A.6 *Suppose the initial conditions are $S = 0$ and $\underline{X} \leq 0$, and suppose $F(B) < F_{fb}(B)$.*

1. *There exists $\epsilon > 0$ such that the solution to (22)-(23) satisfies $S > 0$ and $\frac{dW}{dX} > 0$ for $X \in (\underline{X}, \underline{X} + \epsilon)$.*

2. For $X \geq \underline{X} + \epsilon$, the solution to (22)-(23) belongs to one of three cases: (1) $S \in (0, \psi(X))$ for all $X \geq \underline{X} + \epsilon$, (2) $S = \psi(X)$ at some finite $X = \bar{X}$, and (3) $S = 0$ at some finite $X = \bar{X}$. In case (1), $\lim_{X \rightarrow \infty} W(X) = \infty$. In case (2), $\lim_{X \rightarrow \bar{X}} W(X) = \infty$. In case (3), $W(\bar{X}) < U(\bar{X}) < \infty$. Thus, the function $F(\cdot)$ implied by the solution to (22)-(23) is a global solution in the first two cases, but not in case (3). In case (3), the solution curve cannot be extended to $X > \bar{X}$ because $\frac{dS}{dX}|_{X=\bar{X}} < 0$, which would violate $S \geq 0$.
3. If $S > 0$ for all $X > \underline{X}$, then F is a global solution and stays above F_0 .
4. If there exists a smallest $\hat{W} \geq B$ such that $F(\hat{W}) = F_{\text{fb}}(\hat{W})$, then $F'(W) > (F_{\text{fb}})'(\hat{W})$ for all $W \geq \hat{W}$. Moreover, case (2) in part 2 applies, which implies that $S > 0$ for all $W > B$.

Proof

1. It follows from $\frac{dS}{dX}|_{X=\underline{X}} = W - U(\underline{X}) = B > 0$ that $S > 0$ for $X \in (\underline{X}, \underline{X} + \epsilon)$ and some small $\epsilon > 0$. If $S \in (0, \psi(X))$, then $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{r}{2}\sigma^2(h'(a))^2}$ is positive and finite. Therefore,

$$\frac{dW}{dX} = \frac{1}{\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S}{\frac{r}{2}\sigma^2(h'(a))^2}} > 0. \quad (27)$$

2. In case (1), suppose by contradiction that $\lim_{X \rightarrow \infty} W(X) < I$ for some finite $I > 0$. Then pick a sufficiently large $\hat{X}$ such that $U(X) > I + 1$ for all $X \geq \hat{X}$. Then $\frac{dS}{dX} = W(X) - U(X) < I - (I + 1) = -1, \forall X \geq \hat{X}$. This contradicts the assumption that S is always positive.

In case (2), by contradiction, suppose $\bar{W} := W(\bar{X}) < \infty$. Then (18) implies $F''(\bar{W}) = 0$. We can verify that the straight line $\tilde{F}(W) := F(\bar{W}) + F'(\bar{W})(W - \bar{W})$ is a solution to (17)-(18). On the other hand, the solution to (22)-(23) satisfies $S < \psi(X)$ for $X < \bar{X}$, which implies that $F''(W) < 0$ for $W < \bar{W}$. Since $F \neq \tilde{F}$, we have two solutions to (17)-(18) at $(\bar{W}, F(\bar{W}), F'(\bar{W}))$, contradicting the result that (17)-(18) satisfies the Lipschitz condition (part 2 of Lemma A.5).

In case (3), we show that $\frac{dS}{dX}|_{X=\bar{X}} = W - U(\bar{X}) < 0$. First, we show $\bar{X} > 0$. If $\bar{X} \leq 0$, then $U(X) = 0, \forall X \in [\underline{X}, \bar{X}]$ and

$$\begin{aligned} S(\bar{X}) &= S(\underline{X}) + \int_{\underline{X}}^{\bar{X}} S'(X) dX = S(\underline{X}) + \int_{\underline{X}}^{\bar{X}} (W(X) - U(X)) dX \\ &= 0 + \int_{\underline{X}}^{\bar{X}} W(X) dX > 0, \end{aligned}$$

which contradicts the assumption that $S = 0$ at $X = \bar{X}$. Second, we show that $W - U(\bar{X}) < 0$. By contradiction, suppose $W - U(\bar{X}) \geq 0$. If $W - U(\bar{X}) > 0$, then $\frac{dS}{dX}|_{X=\bar{X}} >$

0, contradicting the fact that $S > 0$ for X slightly below $\bar{X}$. If $W - U(\bar{X}) = 0$, then

$$\frac{d^2 S}{dX^2}|_{X=\bar{X}} = \frac{dW}{dX}|_{X=\bar{X}} - \frac{dU}{dX}|_{X=\bar{X}} = 0 - \frac{dU}{dX}|_{X=\bar{X}} < 0,$$

where the inequality follows from $\bar{X} > 0$ shown in the first step. This again contradicts the fact that $S > 0$ for X slightly below $\bar{X}$.

3. F is a global solution because case (3) in part 2 does not apply. We have $F_0(W) < F(W)$ because

$$\begin{aligned} 0 < S(X) &= F(W) + XW - \max_U (F_0(U) + XU) \\ &\leq F(W) + XW - (F_0(W) + XW) = F(W) - F_0(W). \end{aligned}$$

4. It follows from $F(W) \leq F_{\text{fb}}(W), \forall W \leq \hat{W}$ that $F'(\hat{W}) \geq F'_{\text{fb}}(\hat{W}) =: -X$. To show $F'(\tilde{W}) > (F_{\text{fb}})'(\hat{W}), \forall \tilde{W} \geq \hat{W}$, suppose by contradiction $W \in [\hat{W}, \infty)$ is the smallest W such that $F'(W) = (F_{\text{fb}})'(\hat{W}) = -X$. Since F is concave, $F(W) + F'(W)(\hat{W} - W) \geq F(\hat{W}) = F_{\text{fb}}(\hat{W})$, which implies

$$F(W) + XW \geq F_{\text{fb}}(\hat{W}) + X\hat{W} = F_0(U(X)) + XU(X) + \psi(X),$$

or $S \geq \psi(X)$. The fact that W is finite contradicts case (2) in part 2. Therefore, $F'(\tilde{W}) > (F_{\text{fb}})'(\hat{W}), \forall \tilde{W} \geq \hat{W}$ and the solution curve with bounded $-F'$ does not belong to case (1) in part 2.

To rule out case (3) in part 2, suppose by contradiction that $S = 0$ at some $\bar{X} \geq -F'(\hat{W})$. The properties $U(\bar{X}) > W(\bar{X}) \geq \hat{W}$ (shown in the proof of case (3) in part 2) and $F'(W) > (F_{\text{fb}})'(\hat{W}), \forall W \geq \hat{W}$ imply

$$\begin{aligned} &F(W(\bar{X})) + F'(W(\bar{X}))(U(\bar{X}) - W(\bar{X})) \\ &\geq F_{\text{fb}}(\hat{W}) + (F_{\text{fb}})'(\hat{W})(W(\bar{X}) - \hat{W}) + F'(W(\bar{X}))(U(\bar{X}) - W(\bar{X})) \\ &\geq F_{\text{fb}}(\hat{W}) + (F_{\text{fb}})'(\hat{W})(U(\bar{X}) - \hat{W}) \\ &\geq F_{\text{fb}}(U(\bar{X})) > F_0(U(\bar{X})), \end{aligned}$$

which contradicts $F(W(\bar{X})) + \bar{X}(W(\bar{X}) - U(\bar{X})) - F_0(U(\bar{X})) = S = 0$.

We now conclude that case (2) of part 2 holds, because cases (1) and (3) have been ruled out.

■

Proof of Lemma 3

Let $M > 0$ be the intersection between F_{fb} and the x-axis, i.e., $F_{\text{fb}}(M) = 0$. Define $\mathcal{U}$ as the set of initial conditions under which the solution crosses the upper bound F_{fb} ,

$$\mathcal{U} := \{B \in (0, M] : \text{there exists } \hat{W} \geq B \text{ such that } F_{(B,0,0,-\infty)}(\hat{W}) = F_{\text{fb}}(\hat{W})\}.$$

Define $\mathcal{L}$ as the set of initial conditions under which the solution returns to $S = 0$,

$$\mathcal{L} := \{B \in (0, M] : \text{there exists } W > B \text{ such that } F''_{(B,0,0,-\infty)}(W) = -\infty, \text{ i.e., } S = 0\}.$$

It follows from part 4 in Lemma A.6 that $\mathcal{U} \cap \mathcal{L} = \emptyset$. The rest of this proof consists of the following six steps.

1. Both $\mathcal{U}$ and $\mathcal{L}$ are nonempty. If $B = M$, then $F_{(B,0,0,-\infty)}(B) = 0 = F_{\text{fb}}(B)$. Therefore, $M \in \mathcal{U}$ and $\mathcal{U} \neq \emptyset$.

To show $\mathcal{L} \neq \emptyset$, we show that $B \in \mathcal{L}$ when B is sufficiently small. If $B = 0$, then $\frac{dS}{dX}|_{X=0} = B - U(X) = B = 0$, and $\frac{d^2S}{dX^2}|_{X=0} = \frac{dW}{dX} - \frac{dU(X)}{dX} < 0$. This implies that $S(\hat{X}) < 0$ for some small $\hat{X} > 0$. Because the ODE system (22)-(23) satisfies the Lipschitz condition, its solution is continuous in the initial condition B . That is, $S(\hat{X}) < 0$ whenever $B > 0$ is sufficiently small. The intermediate value theorem implies $S(X) = 0$ at some $X > 0$ whenever $B > 0$ is sufficiently small. Therefore, $B \in \mathcal{L}$.

2. Both $\mathcal{U}$ and $\mathcal{L}$ are open subsets of $(0, M]$.

If $B \in \mathcal{U}$, then $F_{(B,0,0,-\infty)}(\bar{W}) > F_{\text{fb}}(\bar{W})$ at some $\bar{W} > B$. Denote the solution to (22)-(23) as $(W(X, B), S(X, B))$ and suppose $\bar{W} = W(\bar{X}, B)$ for some $\bar{X}$. Since the solution to (23) is continuous in its initial conditions, $(W(\bar{X}, B), S(\bar{X}, B))$ are continuous in B . Moreover, $F_{\text{fb}}(W)$ is continuous in W and $F_{(B,0,0,-\infty)}(W(\bar{X}, B)) = S(\bar{X}, B) + F_0(U(\bar{X})) + \bar{X}(U(\bar{X}) - W(\bar{X}, B))$ is continuous in B . Therefore, $F_{(B,0,0,-\infty)}(W(\bar{X}, B)) - F_{\text{fb}}(W(\bar{X}, B))$ is continuous in B . There exists $\epsilon > 0$ such that $F_{(\tilde{B},0,0,-\infty)}(W(\bar{X}, \tilde{B})) - F_{\text{fb}}(W(\bar{X}, \tilde{B})) > 0$ for all $\tilde{B} \in (B - \epsilon, B + \epsilon)$. Therefore, $(B - \epsilon, B + \epsilon) \subset \mathcal{U}$.

$\mathcal{L}$ is an open subset of $(0, M]$. If $B \in \mathcal{L}$, then $S(X) = 0$ for some $X > 0$. Part 3 of Lemma A.6 shows that $S(\bar{X}) < 0$ if $\bar{X}$ is slightly above X . Since the solution to (22)-(23) is continuous in its initial conditions, $S(\bar{X}, B)$ is continuous in B . There exists $\epsilon > 0$ such that $S(\bar{X}, \tilde{B}) < 0$ for all $\tilde{B} \in (B - \epsilon, B + \epsilon)$. For each $\tilde{B} \in (B - \epsilon, B + \epsilon)$, it follows from $S(\bar{X}, \tilde{B}) < 0$ and $S(\tilde{X}, \tilde{B}) > 0$ for sufficiently small $\tilde{X} > 0$ that $S(\cdot, \tilde{B})$ reaches 0. Therefore, $(B - \epsilon, B + \epsilon) \subset \mathcal{L}$.

3. Because of the above properties, $\mathcal{U} \cup \mathcal{L}$ is not equal to the connected set $(0, M]$. Therefore, there exists $\bar{B} \in (0, M] \setminus (\mathcal{U} \cup \mathcal{L})$. Obviously, $F_{(\bar{B},0,0,-\infty)}$ is below F_{fb} . Part 3 in Lemma A.6 shows that $F_{(\bar{B},0,0,-\infty)}$ is above F_0 , because S is always positive.

4. $\bar{B}$ is unique. By contradiction, suppose $\bar{B}^1 < \bar{B}^2$ both belong to $(0, M] \setminus (\mathcal{U} \cup \mathcal{L})$. Lemma A.1 implies that $F_{(\bar{B}^2, 0, 0, -\infty)}(W) - F_{(\bar{B}^1, 0, 0, -\infty)}(W)$ is increasing in W . For all $W > \bar{B}^2$,

$$\begin{aligned} 0 < F_{(\bar{B}^2, 0, 0, -\infty)}(\bar{B}^2) - F_{(\bar{B}^1, 0, 0, -\infty)}(\bar{B}^2) &< F_{(\bar{B}^2, 0, 0, -\infty)}(W) - F_{(\bar{B}^1, 0, 0, -\infty)}(W) \\ &< F_{\text{fb}}(W) - F_0(W) \leq a^*(W), \end{aligned}$$

where $a^*(W)$ is the optimal effort in $F_{\text{fb}}(W)$. Taking limit $W \rightarrow \infty$, we have

$$F_{(\bar{B}^2, 0, 0, -\infty)}(\bar{B}^2) - F_{(\bar{B}^1, 0, 0, -\infty)}(\bar{B}^2) \leq \lim_{W \rightarrow \infty} a^*(W) = 0,$$

which is a contradiction.

5. $\mathcal{U} = (\bar{B}, M]$. By contradiction, suppose $B \in (\bar{B}, M]$ is in $\mathcal{L}$, then $F_{(B, 0, 0, -\infty)}$ reaches $S = 0$ at some W . Lemma A.1 implies that $F_{(\bar{B}, 0, 0, -\infty)}(W) < F_{(B, 0, 0, -\infty)}(W)$ and $F'_{(\bar{B}, 0, 0, -\infty)}(W) < F'_{(B, 0, 0, -\infty)}(W) =: -X$. Therefore,

$$\begin{aligned} &F_{(\bar{B}, 0, 0, -\infty)}(W) + F'_{(\bar{B}, 0, 0, -\infty)}(W)(U(X) - W) \\ &< F_{(B, 0, 0, -\infty)}(W) + F'_{(B, 0, 0, -\infty)}(W)(U(X) - W) = F_0(U(X)), \end{aligned}$$

where the inequality uses $U(X) > W$. This contradicts the fact that $F_{(\bar{B}, 0, 0, -\infty)}$ satisfies $S > 0$ at all W .

6. $\mathcal{L} = (0, \bar{B})$. By contradiction, suppose $B \in (0, \bar{B})$ is in $\mathcal{U}$, then $F_{(B, 0, 0, -\infty)}(\hat{W}) = F_{\text{fb}}(\hat{W})$ for some $\hat{W}$. Lemma A.1 implies that $F_{(\bar{B}, 0, 0, -\infty)}(\hat{W}) > F_{(B, 0, 0, -\infty)}(\hat{W}) = F_{\text{fb}}(\hat{W})$, which means that $\bar{B} \in \mathcal{U}$, a contradiction.

■

Proof of Lemma 4

Similar to the proof of Lemma 3, this proof proceeds in six steps. Define $\mathcal{U}$ as the set of initial conditions under which the solution crosses the upper bound F_{fb} ,

$$\mathcal{U} := \{y \in [0, F_{\text{fb}}(B)] : \text{there exists } \hat{W} \geq B \text{ such that } F_{(B, y, \frac{y}{B}, -\infty)}(\hat{W}) = F_{\text{fb}}(\hat{W})\}.$$

Define $\mathcal{L}$ as the set of initial conditions under which the solution reaches $S = 0$,

$$\mathcal{L} := \{y \in [0, F_{\text{fb}}(B)] : \text{there exists } W > B \text{ such that } F''_{(B, y, \frac{y}{B}, -\infty)}(W) = -\infty, \text{ i.e., } S = 0\}.$$

It follows from part 4 in Lemma A.6 that $\mathcal{U} \cap \mathcal{L} = \emptyset$.

1. Both $\mathcal{U}$ and $\mathcal{L}$ are nonempty. If $y = F_{\text{fb}}(B)$, then $F_{(B, y, \frac{y}{B}, -\infty)}(B) = F_{\text{fb}}(B)$. Therefore, $F_{\text{fb}}(B) \in \mathcal{U}$. $\mathcal{L}$ is nonempty because $y = 0 \in \mathcal{L}$. Because $B < \bar{B}$, the proof of Lemma 3 implies $F_{(B, 0, 0, -\infty)}$ reaches $S = 0$.

2. Both $\mathcal{U}$ and $\mathcal{L}$ are open subsets of $[0, F_{\text{fb}}(B)]$.

If $y \in \mathcal{U}$, then $F_{(B, y, \frac{y}{B}, -\infty)}(\bar{W}) > F_{\text{fb}}(\bar{W})$ at some $\bar{W} > B$. Denote the solution to (22)-(23) as $(W(X, y), S(X, y))$ and suppose $\bar{W} = W(\bar{X}, y)$ for some $\bar{X}$. Since the solution to (22)-(23) is continuous in its initial conditions, $(W(X, y), S(X, y))$ are continuous in y . Moreover, $F_{\text{fb}}(W)$ is continuous in W and $F_{(B, y, \frac{y}{B}, -\infty)}(W(\bar{X}, y)) = S(\bar{X}, y) + F_0(U(\bar{X})) + \bar{X}(U(\bar{X}) - W(\bar{X}, y))$ is continuous in y . Therefore, $F_{(B, y, \frac{y}{B}, -\infty)}(W(\bar{X}, y)) - F_{\text{fb}}(W(\bar{X}, y))$ is continuous in y . There exists $\epsilon > 0$ such that $F_{(B, \tilde{y}, \frac{\tilde{y}}{B}, -\infty)}(W(\bar{X}, \tilde{y})) - F_{\text{fb}}(W(\bar{X}, \tilde{y})) > 0$ for all $\tilde{y} \in (y - \epsilon, y + \epsilon)$. Therefore, $(y - \epsilon, y + \epsilon) \subset \mathcal{U}$.

$\mathcal{L}$ is an open subset of $[0, F_{\text{fb}}(B)]$. If $y \in \mathcal{L}$, then $S(X) = 0$ for some $X > -\frac{y}{B}$. Step 3 in Lemma A.6 shows that $S(X') < 0$ if X' is slightly above X . Since the solution to (18) is continuous in its initial conditions, $S(X', y)$ is continuous in y . There exists an $\epsilon > 0$ such that for all $\tilde{y} \in (y - \epsilon, y + \epsilon)$, $S(X', \tilde{y}) < 0$. For each $\tilde{y} \in (y - \epsilon, y + \epsilon)$, $S(X', \tilde{y}) < 0$ and $S(\tilde{X}, \tilde{y}) > 0$ when $\tilde{X}$ is slightly above $-\frac{\tilde{y}}{B}$, and therefore $S(\cdot, \tilde{y})$ reaches 0. Thus, $(y - \epsilon, y + \epsilon) \subset \mathcal{L}$.

3. Because of the above properties, $\mathcal{U} \cup \mathcal{L}$ is not equal to the connected set $[0, F_{\text{fb}}(B)]$. Therefore, there exists $y^* \in [0, F_{\text{fb}}(B)] \setminus (\mathcal{U} \cup \mathcal{L})$. Obviously, $F_{(B, y^*, \frac{y^*}{B}, -\infty)}$ is below F_{fb} . Part 3 in Lemma A.6 shows that $F_{(B, y^*, \frac{y^*}{B}, -\infty)}$ is above F_0 .
4. y^* is unique. By contradiction, suppose $y^* < \tilde{y}^*$ both belong to $[0, F_{\text{fb}}(B)] \setminus (\mathcal{U} \cup \mathcal{L})$. Lemma A.1 implies that $F_{(B, \tilde{y}^*, \frac{\tilde{y}^*}{B}, -\infty)}(W) - F_{(B, y^*, \frac{y^*}{B}, -\infty)}(W)$ is increasing in W . In particular,

$$\tilde{y}^* - y^* < F_{(B, \tilde{y}^*, \frac{\tilde{y}^*}{B}, -\infty)}(W) - F_{(B, y^*, \frac{y^*}{B}, -\infty)}(W) < F_{\text{fb}}(W) - F_0(W) \leq a^*(W),$$

where $a^*(W)$ is the optimal effort in $F_{\text{fb}}(W)$. Taking limit $W \rightarrow \infty$, we have

$$\tilde{y}^* - y^* \leq \lim_{W \rightarrow \infty} a^*(W) = 0,$$

which is a contradiction.

5. $\mathcal{U} = (y^*, F_{\text{fb}}(B)]$. By contradiction, suppose $y \in (y^*, F_{\text{fb}}(B)]$ is in $\mathcal{L}$, then $F_{(B, y, \frac{y}{B}, -\infty)}$ reaches $S = 0$ at some W . Lemma A.1 implies that $F_{(B, y^*, \frac{y^*}{B}, -\infty)}(W) < F_{(B, y, \frac{y}{B}, -\infty)}(W)$ and $F'_{(B, y^*, \frac{y^*}{B}, -\infty)}(W) < F'_{(B, y, \frac{y}{B}, -\infty)}(W) =: -X$. Therefore,

$$\begin{aligned} & F_{(B, y^*, \frac{y^*}{B}, -\infty)}(W) + F'_{(B, y^*, \frac{y^*}{B}, -\infty)}(W)(U(X) - W) \\ & < F_{(B, y, \frac{y}{B}, -\infty)}(W) + F'_{(B, y, \frac{y}{B}, -\infty)}(W)(U(X) - W) = F_0(U(X)), \end{aligned}$$

where the inequality uses $U(X) > W$. This contradicts the fact that $F_{(B, y^*, \frac{y^*}{B}, -\infty)}$ satisfies $S > 0$ at all W .

6. $\mathcal{L} = [0, y^*)$. By contradiction, suppose $y \in [0, y^*)$ is in $\mathcal{U}$, then $F_{(B, y, \frac{y}{B}, -\infty)}(\hat{W}) = F_{\text{fb}}(\hat{W})$ for some $\hat{W}$. Lemma A.1 implies that $F_{(B, y^*, \frac{y^*}{B}, -\infty)}(\hat{W}) > F_{(B, y, \frac{y}{B}, -\infty)}(\hat{W}) = F_{\text{fb}}(\hat{W})$, which means that $y^* \in \mathcal{U}$, a contradiction.

■

Proof of Theorem 2 (verification in the Inada case)

First, we show that the profit achieved by any incentive compatible contract (C, A) is at most $F_{(B, y^*(B), \frac{y^*(B)}{B}, -\infty)}(W_0(C, A))$. We will refer to $F_{(B, y^*(B), \frac{y^*(B)}{B}, -\infty)}$ simply as F . Let $\epsilon > 0$ be a small number. Since F asymptotically approaches F_{fb} as $W \rightarrow \infty$, there is a large $\bar{W} > W_0(C, A)$ such that $F(\bar{W}) + \epsilon \geq F_{\text{fb}}(\bar{W})$. Define

$$G_t := r \int_0^t e^{-rt} (A_s - C_s) ds + e^{-rt} F(W_t). \quad (28)$$

By Ito's lemma, the drift of G_t is

$$re^{-rt} \left(A_t - C_t - F(W_t) + F'(W_t)(W_t - u(C_t) + h(A_t)) + r\sigma^2 Y_t^2 \frac{F''(W_t)}{2} \right).$$

Let us show the drift of G_t is always nonpositive. If $A_t > 0$, then incentive compatibility requires $Y_t = h'(A_t)$. Then the fact that F solves the high-action ODE implies that the drift of G is nonpositive. If $A_t = 0$, the same argument as in Lemma A.2 implies that (14) holds also for $F = F_{(B, y^*(B), \frac{y^*(B)}{B}, -\infty)}$. Inequality (14) and $F'' < 0$ imply that the drift of G_t is nonpositive also when $A_t = 0$.

It follows that G_t is a supermartingale until the stopping time τ (possibly ∞), defined as the time when W_t reaches $\bar{W}$. For all $t < \infty$,

$$\begin{aligned} & \mathbb{E} \left[r \int_0^\tau e^{-rt} (A_t - C_t) dt + e^{-r\tau} F(\bar{W}) \right] \\ &= \mathbb{E} \left[G_{t \wedge \tau} + 1_{t \leq \tau} \left(r \int_t^\tau e^{-rs} (A_s - C_s) ds + e^{-r\tau} F(\bar{W}) - e^{-rt} F(W_t) \right) \right] \\ &= \mathbb{E}[G_{t \wedge \tau}] + e^{-rt} \mathbb{E} \left[1_{t \leq \tau} \left(r \int_t^\tau e^{-r(s-t)} (A_s - C_s) ds + e^{-r(\tau-t)} F(\bar{W}) - F(W_t) \right) \right] \\ &\leq G_0 + e^{-rt} \left[\bar{A} + 2 \max_{W \in [B, \bar{W}]} F(W) \right]. \end{aligned}$$

Taking $t \rightarrow \infty$ yields $\mathbb{E} \left[r \int_0^\tau e^{-rt} (A_t - C_t) dt + e^{-r\tau} F(\bar{W}) \right] \leq G_0 = F(W_0)$. At time τ , the firm's future profit is less than or equal to $F_{\text{fb}}(\bar{W}) \leq F(\bar{W}) + \epsilon$. Therefore, the firm's expected profit at time 0 is less than or equal to

$$\mathbb{E} \left[r \int_0^\tau e^{-rt} (A_t - C_t) dt + e^{-r\tau} (F(\bar{W}) + \epsilon) \right] \leq F(W_0) + \epsilon.$$

Since ϵ is arbitrary, we conclude that the firm's expected profit at time 0 is less than or equal to $F(W_0)$.

Second, we show that the contract (C, A) described in the statement of the theorem achieves profit $F(W_0)$ if $W_0 \in [B, \infty)$. Defining G_t as in (28), but now specifically for the stated contract, we have from Ito's lemma that the drift of G_t is

$$re^{-rt} \left(A_t - C_t - F(W_t) + F'(W_t)(W_t - u(C_t) + h(A_t)) + r\sigma^2 h'(A_t)^2 \frac{F''(W_t)}{2} \right) \text{ if } W_t > B,$$

and

$$re^{-rt} (-C_t - F(W_t) + F'(W_t)(W_t - u(C_t))) \text{ if } W_t = B.$$

Given the construction of $F = F_{(B, y^*(B), \frac{y^*(B)}{B}, -\infty)}$, the drift of G_t is zero in both cases. It follows that G_t is a martingale. For all $t < \infty$,

$$\mathbb{E} \left[r \int_0^t e^{-rs} (A_s - C_s) ds + e^{-rt} F(W_t) \right] = G_0 = F(W_0),$$

which implies

$$\mathbb{E} \left[r \int_0^t e^{-rs} (A_s - C_s) ds \right] \geq F(W_0) - e^{-rt} \max_W F(W). \quad (29)$$

Since $r \int_0^t e^{-rs} (A_s - C_s) ds$ converges to $r \int_0^\infty e^{-rt} (A_t - C_t) dt$ almost surely and is bounded above by $\bar{A}$, Fatou's lemma and (29) imply

$$\begin{aligned} \mathbb{E} \left[r \int_0^\infty e^{-rt} (A_t - C_t) dt \right] &\geq \lim_{t \rightarrow \infty} \mathbb{E} \left[r \int_0^t e^{-rs} (A_s - C_s) ds \right] \\ &\geq \lim_{t \rightarrow \infty} \left(F(W_0) - e^{-rt} \max_W F(W) \right) = F(W_0). \end{aligned}$$

Therefore, the firm's expected profit at time 0, $\mathbb{E} [r \int_0^\infty e^{-rt} (A_t - C_t) dt]$, is at least $F(W_0)$. ■

Proof of Proposition 3 (fast reflection)

This proof relies heavily on the discussion of exit and entrance points in Borodin and Salminen (2002). The continuation utility W_t is a diffusion process:

$$dW_t = r(W - u(c(W_t)) + h(a(W_t)))dt + r\sigma h'(a(W_t))dZ_t,$$

where Z_t is a standard Brownian motion. Let m and s denote, respectively, the speed measure and the scale function of the diffusion process W_t .²⁵ Using the formula on page 17 of Borodin

²⁵Since there is no killing of the process W_t in our model, the killing measure k is zero.

and Salminen (2002), we calculate the density function of the speed measure m as follows,

$$\begin{aligned} m(x) &= \frac{2e^{\mathcal{B}(x)}}{(r\sigma h'(a(x)))^2}, \quad \forall x > B, \\ m((x, z)) &= \int_x^z m(y)dy, \quad \forall z > x > B, \end{aligned}$$

where $\mathcal{B}(x) := \int_D^x \frac{2r(W-u(c(W))+h(a(W)))}{(r\sigma h'(a(W)))^2} dW$ and $D \in (B, \infty)$ is a constant. The scale function s satisfies

$$s(x) = \int_D^x s'(y)dy = \int_D^x e^{-\mathcal{B}(y)} dy.$$

First, we show that W_t reaches B with positive probability by checking that B is nonsingular, that is, B is both an exit and an entrance for the diffusion W_t . Recall that B is an exit if $\int_B^z m((x, z))s'(x)dx < \infty$, and is an entrance if $\int_B^z (s(z) - s(x))m(dx) < \infty$. To verify these inequalities, we must estimate upper and lower bounds of the aforementioned functions. By Lemma A.7, $\lim_{W \downarrow B} \frac{h'(a(W))^2}{W-B} = \frac{4B}{r\sigma^2} > 0$. Therefore,

$$\lim_{W \downarrow B} \frac{\frac{2r(W-u(c)+h(a))}{(r\sigma h'(a))^2}}{\frac{1}{W-B}} = \frac{1}{2}.$$

Consequently, for any $\epsilon > 0$, if W is sufficiently close to B ,

$$\begin{aligned} \left(\frac{1}{2} - \epsilon\right) \frac{1}{W-B} &\leq \frac{2r(W-u(c)+h(a))}{(r\sigma h'(a))^2} \leq \left(\frac{1}{2} + \epsilon\right) \frac{1}{W-B} \\ \left(\frac{4B}{r\sigma^2} - \epsilon\right)(W-B) &\leq h'(a(W))^2 \leq \left(\frac{4B}{r\sigma^2} + \epsilon\right)(W-B). \end{aligned}$$

Then, if D is sufficiently close to B and $x < D$,

$$\begin{aligned} \left(\frac{1}{2} - \epsilon\right) (\ln(D-B) - \ln(x-B)) &= \left(\frac{1}{2} - \epsilon\right) \int_x^D \frac{1}{W-B} dW \\ &\leq -\mathcal{B}(x) := \int_x^D \frac{2r(W-u(c)+h(a))}{(r\sigma h'(a))^2} dW \\ &\leq \left(\frac{1}{2} + \epsilon\right) \int_x^D \frac{1}{W-B} dW \\ &= \left(\frac{1}{2} + \epsilon\right) (\ln(D-B) - \ln(x-B)). \end{aligned}$$

Then, $\lim_{x \downarrow B} s(x) > -\infty$ because

$$\begin{aligned} \lim_{x \downarrow B} \int_D^x e^{-\mathcal{B}(y)} dy &= -\lim_{x \downarrow B} \int_x^D e^{-\mathcal{B}(y)} dy \geq -\lim_{x \downarrow B} \int_x^D e^{(\frac{1}{2}+\epsilon)(\ln(D-B)-\ln(y-B))} dy \\ &= -\lim_{x \downarrow B} \int_x^D \left(\frac{y-B}{D-B}\right)^{-(\frac{1}{2}+\epsilon)} dy > -\infty. \end{aligned}$$

Moreover, if $x < D$,

$$m(x) := \frac{2e^{\mathcal{B}(x)}}{(r\sigma h'(a(x)))^2} \leq \frac{2\left(\frac{x-B}{D-B}\right)^{\left(\frac{1}{2}-\epsilon\right)}}{r^2\sigma^2\left(\frac{4B}{r\sigma^2}-\epsilon\right)(x-B)}.$$

Then, $\lim_{x \downarrow B} m((x, z)) < \infty$ because

$$\lim_{x \downarrow B} \int_x^z m(y) dy \leq \lim_{x \downarrow B} \int_x^z \frac{2\left(\frac{y-B}{D-B}\right)^{\left(\frac{1}{2}-\epsilon\right)}}{r^2\sigma^2\left(\frac{4B}{r\sigma^2}-\epsilon\right)(y-B)} dy < \infty.$$

B is an exit because

$$\int_B^z m((x, z)) s'(x) dx \leq m((B, z)) \int_B^z s'(x) dx = m((B, z))(s(z) - s(B)) < \infty.$$

B is an entrance because

$$\int_B^z (s(z) - s(x)) m(dx) \leq (s(z) - s(B)) \int_B^z m(dx) = (s(z) - s(B)) m((B, z)) < \infty.$$

Second, we show that the diffusion process W_t is *reflecting* by checking $m(\{B\}) = 0$. A reflecting process has the property that $P_x(\text{Leb}(\{t \geq 0 : W_t = B\}) = 0) = 1$ for all $x \geq B$. Pick a bounded and smooth function f , and define $f^+(x) := \frac{f'(x)}{s'(x)}$ and

$$g := \mathcal{G}f = \frac{1}{2}(r\sigma h'(a(x)))^2 f''(x) + r(x - u(c(x)) + h(a(x))) f'(x).$$

Equation (c) on page 16 of Borodin and Salminen (2002) implies that

$$m(\{B\}) = \frac{\lim_{x \downarrow B} f^+(x)}{g(B)} = \frac{\frac{f'(B)}{\lim_{x \downarrow B} s'(x)}}{rB f'(B)} = 0,$$

where the last equality follows from $\lim_{x \downarrow B} s'(x) = \lim_{x \downarrow B} e^{-\mathcal{B}(x)} \geq \lim_{x \downarrow B} \left(\frac{D-B}{x-B}\right)^{\frac{1}{2}-\epsilon} = \infty$. ■

The next lemma is used in the proof of Proposition 3.

Lemma A.7 *Under Assumption 1, $\lim_{X \downarrow -F'(B)} \frac{h'(a(X))^2}{W(X)-B} = \frac{4B}{r\sigma^2} > 0$.*

Proof First, we show $\lim_{X \downarrow -F'(B)} \frac{da}{dX} = 2B > 0$. The first-order condition with respect to a taken in equation (23) is

$$h'(a) - X(h'(a))^2 + 2(S - a + Xh(a))h''(a) = 0. \quad (30)$$

The sufficiency of this first-order condition is shown in equation (32) in Lemma A.8. Using (22), we now totally differentiate (30) with respect to X

$$\begin{aligned} 0 &= h''(a) \frac{da}{dX} - (h'(a))^2 - X2h'(a)h''(a) \frac{da}{dX} + 2(S - a + Xh(a))h'''(a) \frac{da}{dX} \\ &\quad + 2\left(W - u(c(X)) - \frac{da}{dX} + h(a) + Xh'(a) \frac{da}{dX}\right) h''(a) \\ &= -h''(a) \frac{da}{dX} - (h'(a))^2 + 2(S - a + Xh(a))h'''(a) \frac{da}{dX} + 2(W - u(c(X)) + h(a)) h''(a). \end{aligned}$$

Solving for $\frac{da}{dX}$, we get

$$\frac{da}{dX} = \frac{2(W - u(c(X)) + h(a))h''(a) - (h'(a))^2}{h''(a) - 2(S - a + Xh(a))h'''(a)},$$

which converges to $\frac{2(B - u(0) + h(0))h''(0) - 0}{h''(0) - 0} = 2B$ as X converges to $-F'(B)$.

Second, we show $\lim_{X \downarrow -F'(B)} \frac{h'(a)^2}{W - B} = \frac{4B}{r\sigma^2} > 0$. Using L'Hôpital's rule and $\lim_{X \downarrow -F'(B)} \frac{da}{dX} = 2B$, we get

$$\begin{aligned} \lim_{X \downarrow -F'(B)} \frac{h'(a)^2}{W - B} &= \lim_{X \downarrow -F'(B)} \frac{2h'(a)h''(a)\frac{da}{dX}}{\frac{dW}{dX}} \\ &= 4Bh''(0) \lim_{X \downarrow -F'(B)} \frac{h'(a)}{\frac{dW}{dX}} \\ &= \frac{8Bh''(0)}{r\sigma^2} \lim_{X \downarrow -F'(B)} \frac{a - Xh(a) - S}{h'(a)}, \end{aligned}$$

where the last equality uses (23). Using L'Hôpital's rule again,

$$\begin{aligned} \lim_{X \downarrow -F'(B)} \frac{h'(a)^2}{W - B} &= \frac{8Bh''(0)}{r\sigma^2} \lim_{X \rightarrow -F'(B)} \frac{\frac{da}{dX} - h(a) - Xh'(a)\frac{da}{dX} - (W - u(c(X)))}{h''(a)\frac{da}{dX}} \\ &= \frac{8Bh''(0)}{r\sigma^2} \frac{2B - 0 + F'(B)0(2B) - (B - u(0))}{h''(0)2B} = \frac{4B}{r\sigma^2}. \end{aligned}$$

■

The last lemma justifies the first-order condition (30) used in Lemma A.7.

Lemma A.8 *Pick $\delta > 0$ such that $-2h(a)h'''(a) < \frac{1}{\max_{X \in [\underline{X}, \underline{X}+1]} |X|} \frac{h''(0)}{4}$, $2S(X)h'''(a) < \frac{h''(0)}{4}$, and $-h''(a) - 2ah'''(a) < -\frac{3}{4}h''(0)$ for all $(a, X) \in [0, \delta] \times (\underline{X}, \underline{X} + \delta]$, where $\underline{X} = -F'(B)$. There exists $\epsilon \in (0, \delta)$ such that for all $X \in (\underline{X}, \underline{X} + \epsilon]$, $\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S(X)}{\frac{r}{2}\sigma^2(h'(a))^2}$: 1) achieves its maximum in $(0, \delta)$, and 2) the maximizer is unique, interior, and given by the first-order condition.*

Proof It follows from $\lim_{X \downarrow \underline{X}} \max_{a \in \mathcal{A}} \frac{a - Xh(a) - S(X)}{\frac{r}{2}\sigma^2(h'(a))^2} = \lim_{X \downarrow \underline{X}} (-F''(W)) = \infty$ and the finiteness of $\max_{(a, X) \in [\delta, \bar{A}] \times [\underline{X}, \underline{X}+1]} \frac{a - Xh(a) - S(X)}{\frac{r}{2}\sigma^2(h'(a))^2}$ that there exists $\epsilon \in (0, \delta)$ such that for all $X \in (\underline{X}, \underline{X} + \epsilon]$,

$$\max_{a \in \mathcal{A}} \frac{a - Xh(a) - S(X)}{\frac{r}{2}\sigma^2(h'(a))^2} > \max_{(a, X) \in [\delta, \bar{A}] \times [\underline{X}, \underline{X}+1]} \frac{a - Xh(a) - S(X)}{\frac{r}{2}\sigma^2(h'(a))^2}. \quad (31)$$

The first conclusion of the lemma follows because (31) implies that if $X \in (\underline{X}, \underline{X} + \epsilon]$, then $\frac{a - Xh(a) - S(X)}{\frac{r}{2}\sigma^2(h'(a))^2}$ cannot achieve a maximum in $a \in [\delta, \bar{A}]$.

For the second conclusion, we show that the maximizer is interior and unique. The derivative of $\frac{a - Xh(a) - S(X)}{(h'(a))^2}$ with respect to a is

$$\begin{aligned} & \frac{(1 - Xh'(a))h'(a)^2 - (a - Xh(a) - S(X))2h'(a)h''(a)}{(h'(a))^4} \\ = & \frac{(1 - Xh'(a))h'(a) - (a - Xh(a) - S(X))2h''(a)}{(h'(a))^3}. \end{aligned}$$

Because $((1 - Xh'(a))h'(a) - (a - Xh(a) - S(X))2h''(a))|_{a=0} = S(X)2h''(0) > 0$ for all $X \in (\underline{X}, \underline{X} + \epsilon]$, the above derivative is positive at $a = 0$ and hence the optimal a is in $(0, \delta)$. To show uniqueness of the maximizer, it is sufficient to show that the first-order condition for a ,

$$(1 - Xh'(a))h'(a) - (a - Xh(a) - S(X))2h''(a) = 0, \quad (32)$$

has a unique solution. We verify that $(1 - Xh'(a))h'(a) - (a - Xh(a) - S(X))2h''(a)$ is strictly decreasing in a by showing its derivative is negative.

$$\begin{aligned} & \frac{\partial}{\partial a} ((1 - Xh'(a))h'(a) - (a - Xh(a) - S(X))2h''(a)) \\ = & (1 - Xh'(a))h''(a) - Xh''(a)h'(a) - (a - Xh(a) - S(X))2h'''(a) - (1 - Xh'(a))2h''(a) \\ = & -h''(a) - 2ah'''(a) - 2Xh(a)h'''(a) + 2S(X)h'''(a) \\ < & -\frac{3}{4}h''(0) + \frac{1}{4}h''(0) + \frac{1}{4}h''(0) = -\frac{1}{4}h''(0) < 0, \end{aligned}$$

where the first inequality follows from the definition of δ . ■

References

- Atkeson, A. and R. E. Lucas, Jr. (1995). Efficiency and equality in a simple model of efficient unemployment insurance. *Journal of Economic Theory* 66(1), 64–88.
- Borodin, A. and P. Salminen (2002). *Handbook of Brownian Motion - Facts and Formulae*, 2nd ed. Birkhauser.
- Brunnermeier, M. K. and Y. Sannikov (2014). A macroeconomic model with a financial sector. *The American Economic Review* 104(2), 379–421.
- Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica* 53(2), 385–407.
- DeMarzo, P. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *The Journal of Finance* 66(6), 2681–2724.

- Engelbert, H. and G. Peskir (2014). Stochastic differential equations for sticky brownian motion. *Stochastics* 86(6), 993–1021.
- Grochulski, B. and Y. Zhang (2016). Market-based incentives. *FRBR Working Paper 13-05R*.
- Milgrom, P. and I. Segal (2002, March). Envelope theorems for arbitrary choice sets. *Econometrica* 70(2), 583–601.
- Phelan, C. (1995). Repeated moral hazard and one-sided commitment. *Journal of Economic Theory* 66(2), 488–506.
- Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies* 75(3), 957–984.
- Spear, S. E. and C. Wang (2005). When to fire a CEO: Optimal termination in dynamic contracts. *Journal of Economic Theory* 120(2), 239–256.
- Stokey, N. L. (2008). *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. Princeton University Press.
- Thomas, J. and T. Worrall (1990). Income fluctuations and asymmetric information: An example of a repeated principal-agent problem. *Journal of Economic Theory* 51(2), 367–390.
- Wang, C. and Y. Yang (2015). Equilibrium matching and termination. *Journal of Monetary Economics* 76, 208–229.
- Zhu, J. Y. (2013). Optimal contracts with shirking. *Review of Economic Studies* 80(2), 812–839.