Working Paper Series

Optimal Contracts with Reflection

WP 16-14

Borys Grochulski
Federal Reserve Bank of Richmond

Yuzhe Zhang
Texas A&M University

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Optimal Contracts with Reflection*

Borys Grochulski[†] Yuzhe Zhang[‡]

September 29, 2016
Working Paper No. 16-14

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 put zero effort temporarily, which brings his 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 (after a random search time of finite expected duration). In addition, we find new dynamics of the reflection at the lower bound. In the baseline model, the dynamics of the reflection are slow, as in Zhu (2013), i.e., the zero-action is used 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., the zero-effort action is used seldom.

1 Introduction

In an important contribution to the literature on incentives, Sannikov (2008) develops and applies methods for computing optimal contracts in 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 value as a diffusion process and by identifying contracts with solutions to the Hamilton-Jacobi-Bellman equation, an ODE. The solutions obtained in Sannikov (2008) provide several new insights into the structure of optimal dynamic

*The authors would like to thank Huberto Ennis, Boyan Jovanovic, Chris Phelan, Yuliy Sannikov, and Russell Wong for their helpful comments. 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.

incentive contracts. In particular, Sannikov (2008) shows that the optimal contract terminates after sufficiently poor performance.

In this paper, we qualify this result. We show that termination after poor performance is only optimal when the agent's outside option is his minimax payoff of zero. If the agent's outside option is even sightly better than the absolute lowest payoff the agent could get, the optimal contract reflects off of the lower bound and does not terminate. Termination after poor performance, thus, occurs in the optimal contract only under the assumption that the agent's outside option value is in a corner.

The agent's continuation utility W_t is the state variable in this contracting problem. When W_t reaches its lower bound $B > \text{minmax}\{W_t\} = 0$, the agent's compensation and effort supply requirement are zero. 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 increasing W_t at the agent's rate of time preference $r > 0$. Thus, W_t moves up away from $W_t = B$, which allows the contract to continue without termination.

Figure 1 replicates Figure 3 from Sannikov (2008). The faint vertical line indicates the exogenous lower bound on the agent's continuation utility, which is taken to be 0.1 in this example. The curve labeled "optimum without reflection" is 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. The curve labeled "optimum with reflection" is the optimal profit for the firm when we relax this assumption. As we see, allowing for reflection leads to a better contract. The upper bound $F_{\max}$ 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 optimal contract does not reflect even if allowed to. The lower bound F_0 is the (negative) profit the firm can make by never asking the agent to exert effort and just paying him a constant compensation that delivers his promised utility W.

In Section 6, we embed this 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. 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 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

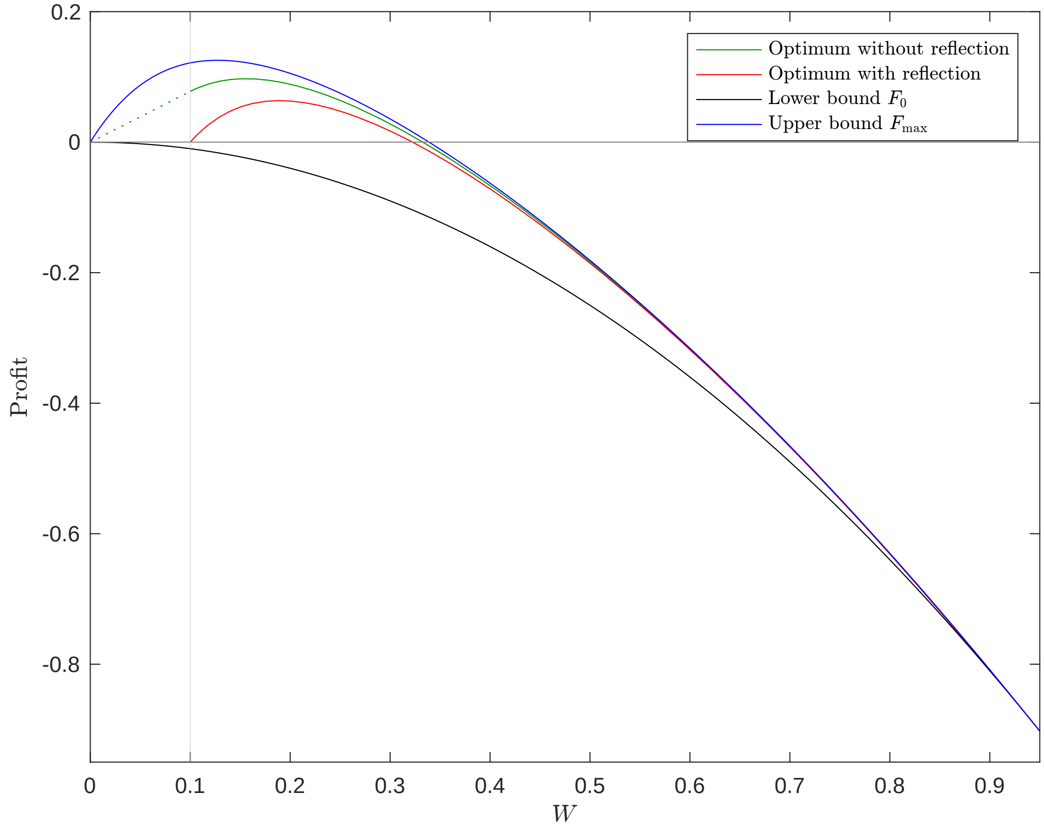

Figure 1: The firm's profit function with the agent's exogenous lower bound $B = 0.1$.

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 but positive effort when W_t is just above the lower bound B. Incentive compatibility, thus, implies that the volatility of W_t is zero at B and positive above B. Moreover, optimal effort is discontinuous at B, which means the volatility of W_t jumps at B as well. The contract, thus, returns to B frequently, which Zhu (2013) dubs slow reflection. We show that the same dynamics arise in the Sannikov (2008) model with nonzero lower bound.

However, if the marginal disutility of effort at zero is zero, which is a standard Inada condition, the dynamics of the reflection change. As W_t moves toward the lower bound, effort decreases toward zero continuously and thus the volatility of W_t approaches zero smoothly, while the drift of W_t remains strictly positive. With volatility extinguished 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, which is incentive compatible because W_t becomes insensitive to realized output. Compensation paid to the agent, however, is positive and set at a level high enough to guarantee a negative drift of W_t, which makes W_t move down and away from the upper bound, allowing for resumption of effort and further continuation of the contract. In this extension, thus, the optimal contract does not terminate. Under the optimal contract in the model with positive agent outside option and a finite-cost firm outside option, neither the low nor high agent retirement points characterized by Sannikov (2008) exist.

The second extension relaxes the assumption that the agent's utility from consumption is bounded from below. We show that this specification 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.

Contribution relative to the existing literature Our paper extends the analysis of Sannikov (2008) building on Zhu (2013). We show that slow reflection discovered by Zhu (2013) is not specific to the risk-neutral principal-agent model without risk-sharing or intertemporal consumption smoothing, which Zhu (2013) considers, but 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, outside of the corner case in which the agent's outside option is equal to his minimax payoff. Given that termination of output of the relationship after poor performance does not appear in dynamic contracting models in discrete time, e.g., Atkeson and Lucas (1995) or Phelan (1995), the result of Sannikov (2008) showing that productivity of the relationship seizes for incentives reasons when the agent's continuation value hits its lower bound is surprising. By qualifying this result, our analysis brings the lessons from dynamic optimal contracting models in continuous time closer to the lessons obtained previously in discrete time.

Continuous-time methods applied to incentive problems by Sannikov (2008), however, 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.

Organization Section 2 describes the contracting problem we study, which is exactly the same as the problem studied as in Sannikov (2008). Section 3 recalls the optimal contracts without reflection derived in Sannikov (2008). Section 4 provides an informal description of how fully optimal contracts, i.e., those allowing for reflection, can be constructed by combining low- and high-action ODEs in a fashion similar to 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 allows for Inada conditions 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). The cumulative output X_t produced by the agent up to date 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 $\bar{A} > 0$.[1] 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 principal evaluate the contract according to, respectively,

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

and

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

where $r > 0$. The agent's utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ is C^2 with $u' > 0$, $u'' < 0$, $\lim_{c\to 0} u'(c) = 0$, and $u(0) = 0$. The function $h : \mathcal{A} \to \mathbb{R}_+$ representing the agent's disutility from effort is increasing and convex with $h(0) = 0$. In addition, in this section we follow Sannikov (2008) in

[1] That $\mathcal{A}$ 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 a compact subset of $\mathbb{R}$ with the smallest element 0.

assuming that there exists $\gamma_0 > 0$ such that $h(a) \geq \gamma_0 a$ for all $a \in \mathcal{A}$. Note that this assumption means that if h is differentiable, it does not satisfy the usual Inada condition $\lim_{a \to 0} h'(a) = 0$.

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

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

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

$$dW_t = r(W_t - u(C_t) + h(A_t))dt + rY_t(dX_t - A_t dt), \tag{1}$$

where $dX_t - A_t dt$ is the agent's performance relative to the benchmark $A_t dt$. 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

$$A_t \in \arg\max_{\tilde{a}} Y_t \tilde{a} - h(\tilde{a}). \tag{2}$$

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

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

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

$$W_t \geq B$$

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 non-negative profit to the firm as of $t = 0$.

3 Optimal contracts without reflection

3.1 Optimal contract with $B = 0$

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 $F_0(W) = -u^{-1}(W)$. In Figure 1, $F_{\max}$ and F_0 are depicted as, respectively, the upper and the lower bound functions.

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

$$dW_t = r(W_t - u(c(W_t)) + h(a(W_t)))dt + rY(W_t)(dX_t - a(W_t)dt) \tag{4}$$

for $0 \leq t < \tau$. 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.[2]

More precisely, Sannikov (2008) solves the HJB equation (3) in two steps. In the first step, an additional restriction is imposed in the HJB equation requiring that the controls a and Y satisfy $a > 0$ and $Y \geq \gamma_0$ at all W such that $F(W) > F_0(W)$. Building on the intuition of Zhu (2013), we will refer to such a restricted HJB equation as the high-action ODE, because the zero-effort action $a = 0$ is explicitly excluded unless $F(W) = F_0(W)$. In the second step, it is verified that the above restriction is without loss of efficiency, i.e., that the firm could not make a profit higher than $F_{\max}$ by using the low-action $a = 0$ at some W such that $F_{\max}(W) > F_0(W)$. The optimal contract is absorbed when W_t hits 0 because $c = a = Y = 0$ is the unique set of controls that are feasible at this point, given that $W_t \geq 0$ at all t. The optimal contract is also absorbed when W_t hits W_{gp} because, with $F''_{\max}(W_{gp}) < 0$, $a = Y = 0$ and $c = u^{-1}(W_{gp})$ is the efficient policy at W_{gp}.

3.2 Optimal contract with $B > 0$ without reflection

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 unless $F(W) = F_0(W)$, at which point the agent is retired or fired, so, effectively, the zero-effort action $a = 0$ is not used prior to contract termination. In particular, this solution maintains that the agent is fired when $W_t = B$, from which the firm makes the profit of zero. Under this assumption,

[2]In other words, at 0 and W_{gp} the contract is constructed from the policy functions that attain the retirement profit F_0.

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 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 and thus the contract does not terminate at $B > 0$. The process W_t moves up from B and the contract continues.

We proceed in a way similar to Zhu (2013). We split the HJB equation into two ODEs: the high-action ODE with the restriction that $a > 0$, which is the case studied in detail in Sannikov (2008), and a low-action ODE, where it is assumed that the recommended action is $a = 0$.

4.1 The low-action ODE

If $a = 0$ is the recommend action, the minimal volatility of W_t that makes this action incentive compatible is $Y = 0$. With these two controls fixed, the HJB equation (3) reduces to

$$F(W) = \max_{c \geq 0} -c + F'(W)(W - u(c)). \tag{5}$$

The optimal c is zero if $F' \geq 0$ and solves $-1 - F'u'(c) = 0$ if $F' < 0$. Consider the following cases.

Case 1a. Suppose $F' > 0$. Then $c = 0$ and the ODE reduces to $F = F'(W - u(0)) = F'W$. Solving by separation of variables we have $\frac{dF}{F} = \frac{dW}{W}$ or $d\ln(F) = d\ln(W)$, so $F = We^{C_1}$. Because $e^{C_1} > 0$, we have $F' = e^{C_1} > 0$ consistent with the assumption. Thus, solutions with $F' > 0$ are straight lines $F = \alpha W$ for $\alpha > 0$. These lines go though the origin of the phase plain (W, F). Since we do not consider negative W, our solutions are rays out of the origin with a strictly positive slope.

Case 1b. If $F' = 0$, then optimal $c = 0$, and the ODE is $F = F'(W - u(0)) = 0\,(W - 0) = 0$. This solution coincides with the horizontal axis, i.e., it takes the same form as the solutions in Case 1a.

Case 2 Suppose $F' < 0$. The optimal $c > 0$ solves $-1 - F'u'(c) = 0$. The ODE is $F = -c + F'(W - u(c))$, where c satisfies this FOC. It is not hard to show that solutions of this form are a) the retirement profit function F_0 and b) all tangent lines to F_0. These solutions are of secondary interest 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.

Let us now consider the dynamics of W_t along any Case-1 solution. From (1) we get

$$dW_t = rW_t dt,$$

i.e., W_t grows deterministically at the rate r.

4.2 Combining high- and low-action solutions

Let us now discuss informally how the solutions to the two ODEs can be combined to improve the contract relative to the optimal contract without reflection. Figure 2 replicates again the optimum without reflection, $\tilde{F}$, for the example with $B = 0.1$ presented in Figure 3 (left panel) of Sannikov (2008). As we see, there exists a unique Case-1 solution to the low-action ODE tangent to $\tilde{F}$ at some $W^s > B$. Let us denote this solution by $L(W)$ and 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 (with $a = 0$) 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.

Note that because the optimal controls obtained from the solution L to the low-action ODE are $(c, a, Y) = (0, 0, 0)$, the process W_t 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 hits W^s, it never returns into $[B, W^s)$, i.e., W_t reflects off of W^s and stays in $[W^s, W_{gp}]$ thereafter.

Note also in Figure 2 that the second derivatives of L and $\tilde{F}$ are not equal at W^s. Because $W^s > B$ 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 has a higher slope, while the high-action ODE is everywhere above $\tilde{F}$. As

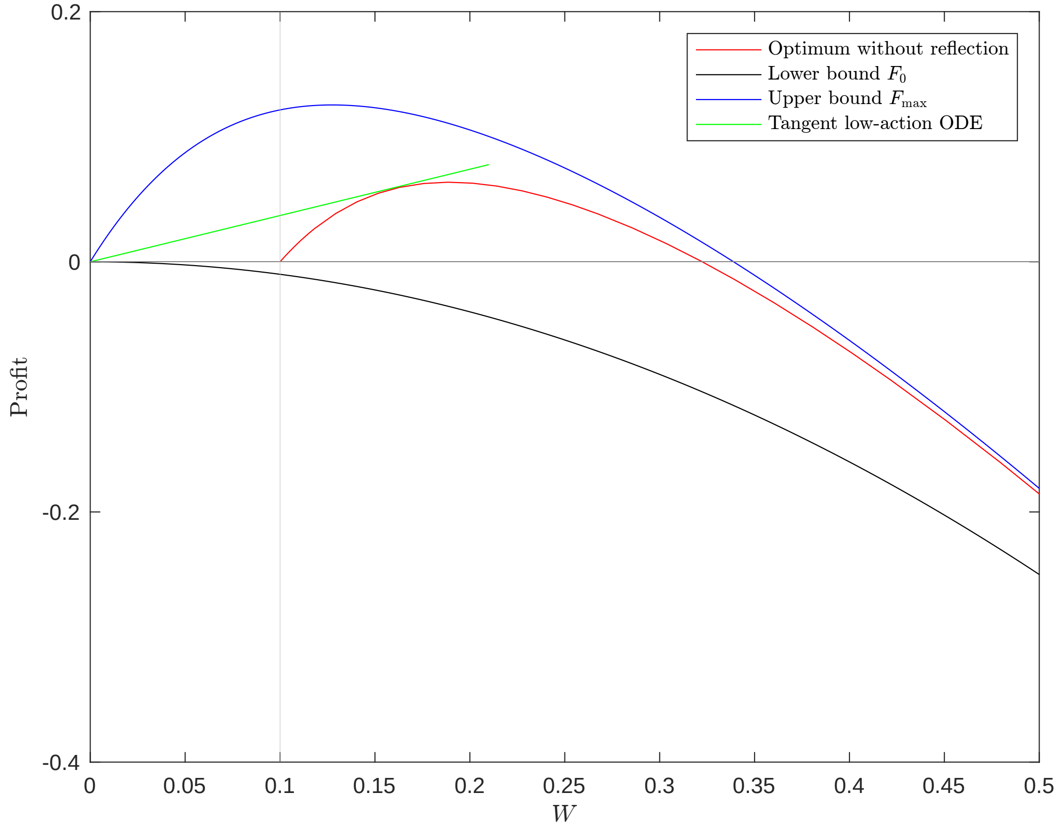

Figure 2: Combining the high- and low-action solutions. The splicing point is $W^s = 0.17 > 0.1 = B$. Positive effort is exerted to the right of W^s. Zero effort is exerted at and to the left of W^s.

before, the two ODEs are spliced at a point where the smooth pasting conditions are satisfied, thus resulting in a consistent contract over the whole domain $[B, W_{gp}]$.[3]

The intuition for why this contract is better has to do with the "endogenous" support $[W^s, W_{gp}]$. The larger this support, the better insurance the firm can provide to the agent without shutting down the agent's effort, i.e., a higher solution curve F of the high-action ODE can be attained. As we show formally next, the optimal contract obtains when the splicing point W^s coincides with the exogenous lower bound B, which case is 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 further enlarged.

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

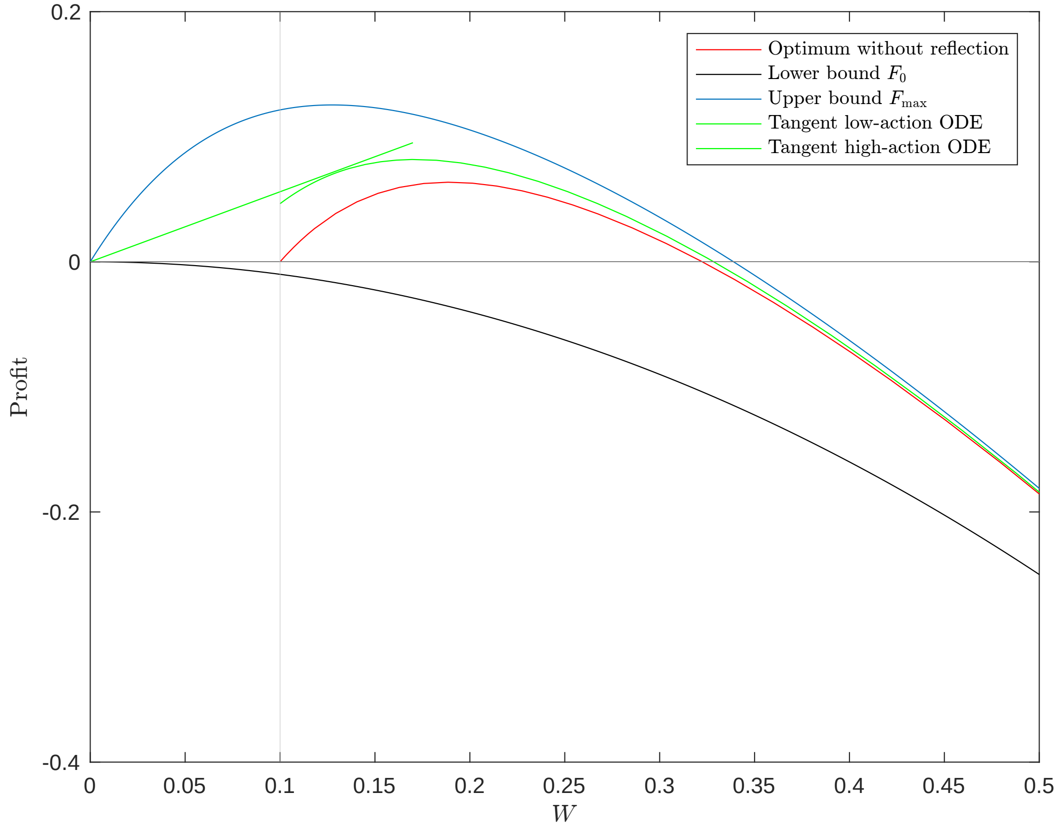

Figure 3: Combining the high- and low-action solutions at the splicing point $W^s = 0.13$ leads to a better contract than $\tilde{F}$ at all W.

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 $F(B) = y$, $F'(B) = y'$.[4]

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

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

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) there exists* $W_{gp} \in [\bar{B}, W^*_{gp}]$ *such that* $F_{(\bar{B},0,0)}(W_{gp}) = F_0(W_{gp})$ *and* $F'_{(\bar{B},0,0)}(W_{gp}) = F'_0(W_{gp})$*.*

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 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 $W_t \geq B$ 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 under which the agent's quitting constraint is satisfied and which 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 is obtained. For any $B \in (0, \bar{B}]$, for some $y > 0$ and $y' > 0$, a high-action ODE solution $F_{(B,y,y')}$ satisfying conditions a) and b) can be pasted at $W = B$ with a strictly positively sloped low-action ODE solution (i.e., a solution of Case 1a). The requirement of smooth pasting between $F_{(B,y,y')}$ and a straight ray coming out of the origin pins down the initial slope y' for given B and y. We must have $y' = \frac{y}{B}$.

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) there exists* $W_{gp} \in [B, W^*_{gp}]$ *such that* $F_{(B,y,\frac{y}{B})}(W_{gp}) = F_0(W_{gp})$ *and* $F'_{(B,y,\frac{y}{B})}(W_{gp}) = F'_0(W_{gp})$*. 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 $V : [B, W_{gp}(B)] \to \mathbb{R}$ 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

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

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 $\{W_t, 0 \le t < \infty\}$ 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 $W_t = W_{gp}(B)$.*

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 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 of B. The reflection is slow, as in Zhu (2013). That is, the process W_t spends a positive expected amount of time at the lower bound.

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, $W_0(B) := \operatorname{argmax}_{W \ge B} V(W)$, 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.*

We can interpret $W_0(B) - B$ as the agent's 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

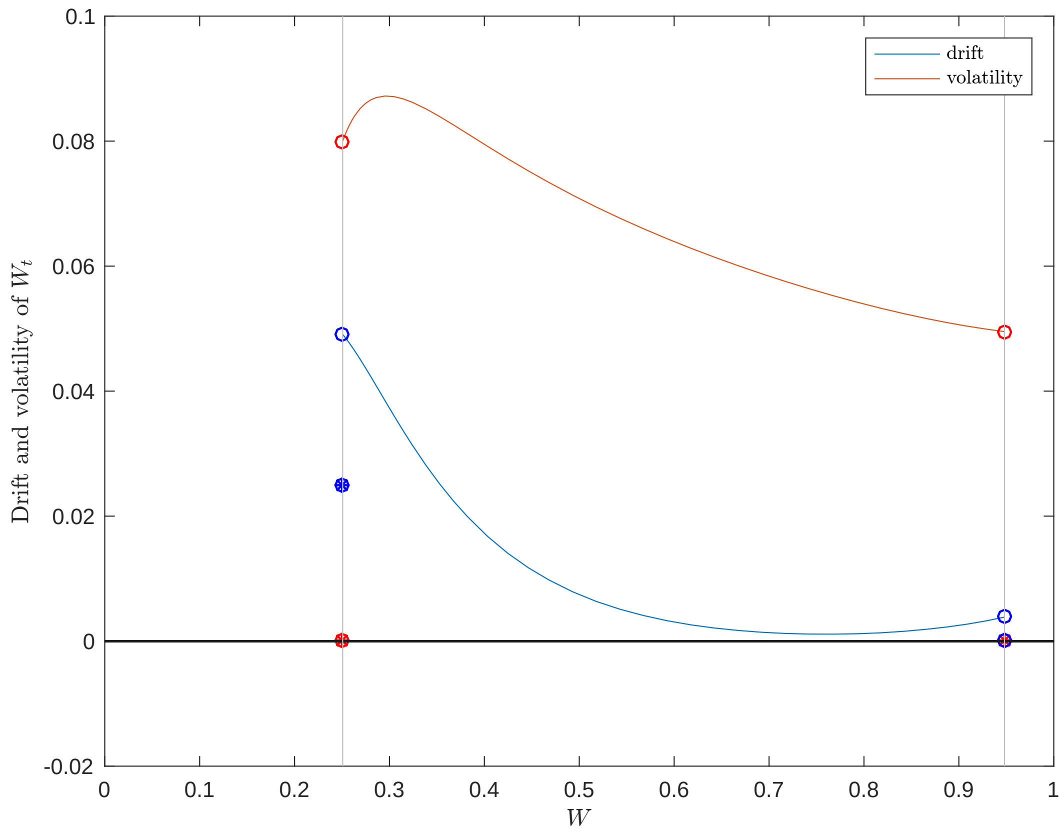

Figure 4: The dynamics of W_t. The two vertical lines mark B and W_{gp}. Volatility of W_t is discontinuous jumping down to zero at both boundaries. Drift jumps as well but not all the way to zero at B.

obtains increases in absolute terms. But the contracting relationship becomes less profitable, as the firm has less room to provide incentives. As a result, both the firm's surplus and the worker's surplus relative to his outside option are reduced. Figure 5 shows the inverses relationship between W_0 and $V(W_0(B))$ graphically.

6 Endogenous outside option

In this section, we endogenize the agent's outside option B by embedding the contacting 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 proposition. 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

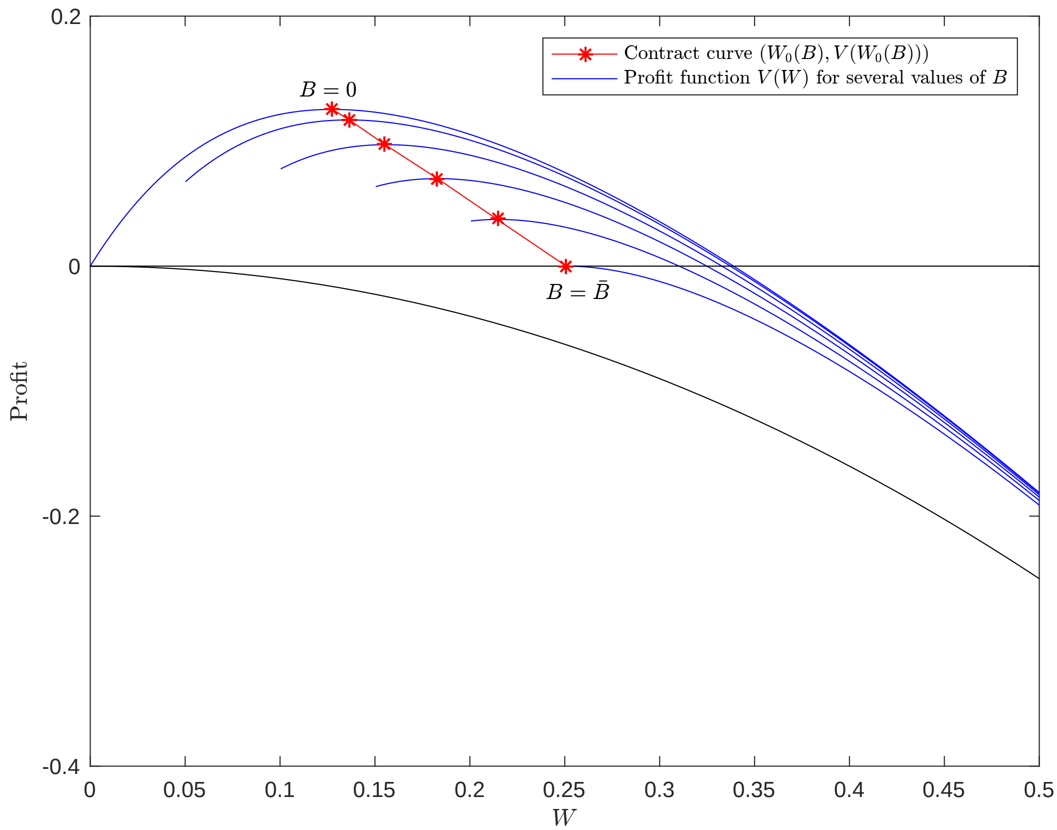

Figure 5: Comparative statics with respect to B.

agent consumes zero and exerts zero effort.[5] 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

$$B = \int_0^\infty \lambda e^{-\lambda t} e^{-rt} W_0 \, dt = \frac{\lambda}{r+\lambda} W_0. \tag{6}$$

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

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*

[5]This normalizing assumption can easily be relaxed.

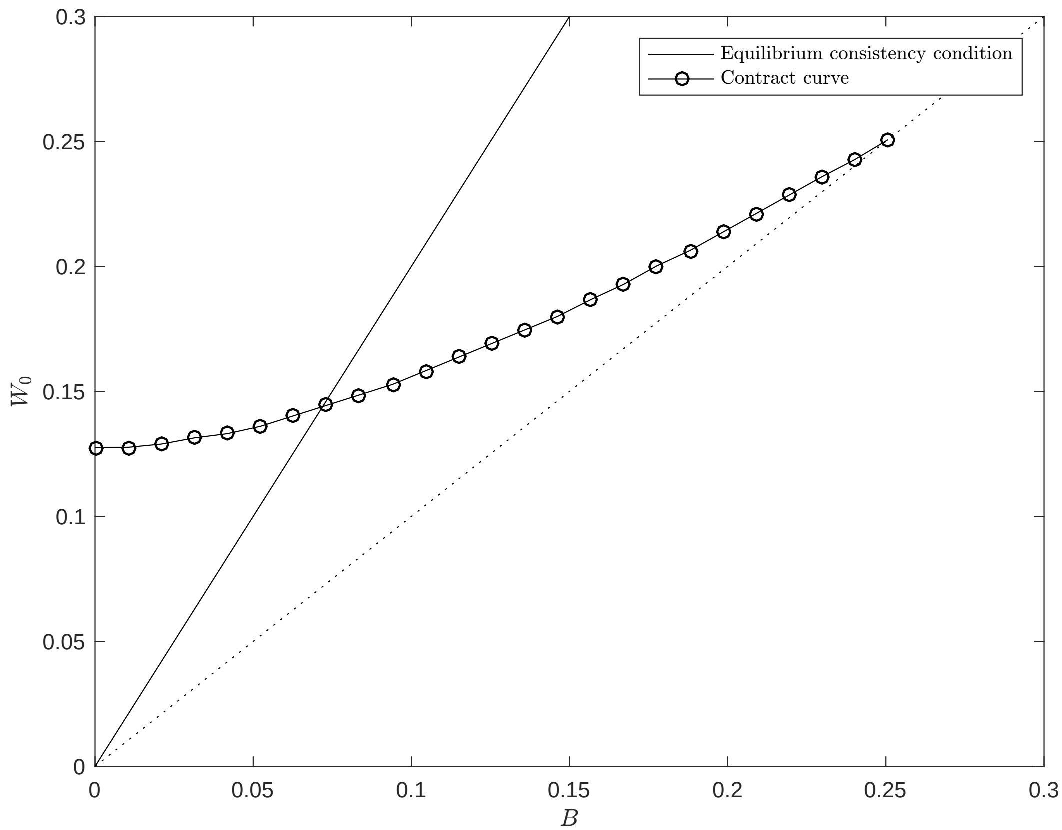

Figure 6: Equilibrium conditions.

function given the agent's outside option B, and 2) the consistency condition (6) *holds with* $W_0 \in \operatorname{argmax}_{W \geq B} V(W)$.

Proposition 2 *For every $\lambda \in [0, \infty]$, there exists a unique competitive equilibrium (W_0, B, V). The mapping from the expected search time $1/\lambda$ to the equilibrium values W_0, B, and $V(W_0)$ is monotonic, with W_0 and B strictly decreasing and $V(W_0)$ strictly increasing in $1/\lambda$. 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 $V(W_0) = \max F_{\max}(W)$.*

Proof of Proposition 2 follows from Figure 6. In that figure, the circled 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 $W_0(B) := \operatorname{argmax}_{W \geq B} V(W)$ 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 6 also shows how equilibrium depends on the new match arrival rate λ or on the expected search time $1/\lambda$. If $\lambda = \infty$, i.e., the agent can get a new match immediately after quitting, the consistency condition coincides with the 45-degrees line and equilibrium lower bound B is determined at the highest possible level $\bar{B}$. When search cost $1/\lambda$ becomes positive, both B and $W_0(B)$ decline.[6] The lower bound B reaches zero only when search cost $1/\lambda$ becomes infinite. This means that the equilibrium contract features termination of output at the lower bound only if the agent's search cost is infinite. For all finite values of the search cost $1/\lambda$, the lower bound B determined in equilibrium is strictly positive, which implies that the process W_t is not absorbed at B but rather reflects off of it in equilibrium.

7 Fast reflection

In this section, we assume that h has a bounded second derivative and that it satisfies the standard Inada condition $\lim_{a \to 0} h'(a) = 0$. 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

As before, we will construct solutions to the HJB equation by splicing the low- and high-action ODEs at an exogenous lower bound B for the agent's continuation value process W_t. Let L be a solution to the low-action ODE and F a solution to the high-action ODE. The value-matching and smooth-pasting conditions, $F(B) = L(B)$ and $F'(B) = L'(B)$, imply

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

If $\lim_{a \to 0} h'(a) = 0$, this condition cannot be met with a finite $F''(B)$. Indeed, 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 $a = 0$:

[6] By Proposition 1, the firm's value $V(W_0(B))$ increases.

$1 + L'h'(a) + r\sigma^2 F'' h'(a)h''(a)|_{a=0} = 1 > 0$. In order to satisfy the above condition, therefore, we must allow for $F''(B) = -\infty$, in which case the optimal action in the high-action ODE is $a = 0$ at $W = B$. [7] [8]

In sum, in order to satisfy the value-matching and smooth-pasting conditions at B, a solution F to the HJB equation for $W \geq B$ must satisfy $F(B) = L(B)$, $F'(B) = L'(B)$, and $F''(B) = -\infty$, i.e., F must be singular at B. In the Appendix, we present a change-of-variable technique that allows for solving of the HJB equation forward from B despite this singularity.

Because the firm and the agent discount at the same rate, the optimal contract under full information (observable a) 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 at the first best, even with very high W.

Note that $F_0 < F_{\text{fb}}$ because

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

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

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

The next two lemmas are analogs of Lemma 1 and Lemma 2.

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

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

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*

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

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

[7] Note that this problem does not arise if $h'(a) \geq \gamma_0 > 0$ for all a. Indeed, $F''h'(a)^2 \leq F''\gamma_0^2$ and $\max_{a\in\mathcal{A}} a + L'h(a) \leq M$ for some M imply $a + L'h(a) + \frac{1}{2}r\sigma^2 F''h'(a)^2 \leq M + \frac{1}{2}r\sigma^2 F''\gamma_0^2$ for all a, so a finite number $F'' < 0$ exists with which the value-matching and smooth-pasting conditions can be met.

[8] The high-action ODE is in this case the unrestricted ODE, where $a = 0$ is allowed.

7.2 Optimal contract

For each $B \in (0, \bar{B}]$, let us define a function $V : [B, \infty) \to \mathbb{R}$ 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

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

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 $\{W_t, 0 \le t < \infty\}$ 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 $W > B$.*

7.3 The dynamics of reflection

Under the optimal contract, the process W_t is not Sticky Brownian Motion studied in Zhu (2013). See Figure 7. As in the non-Inada case shown earlier in Figure 4, drift of W_t is strictly positive at the lower bound $B > 0$ and in a neighborhood of B, while volatility is zero at B and strictly positive in a neighborhood of B. In the non-Inada case, the volatility of W_t was discontinuous at B. In the Inada case, however, the volatility of W_t converges to zero as W goes to the lower bound B, i.e., it is right-continuous at B.

This means that when W_t gets close to B, its drift remains bounded away from zero 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. As a result, the (Lebesgue) measure of time that W_t spends at B is zero with probability one. Thus, the reflection of W_t off of B is faster in the Inada case than in the non-Inada case.

Proposition 3 *B is nonsingular and, for any $W_0 \ge B$, we have $P_{W_0}(Leb(\{t \ge 0 : W_t = B\}) = 0) = 1$.*

7.4 Remark on immiserization

In the dynamic contracting literature, 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

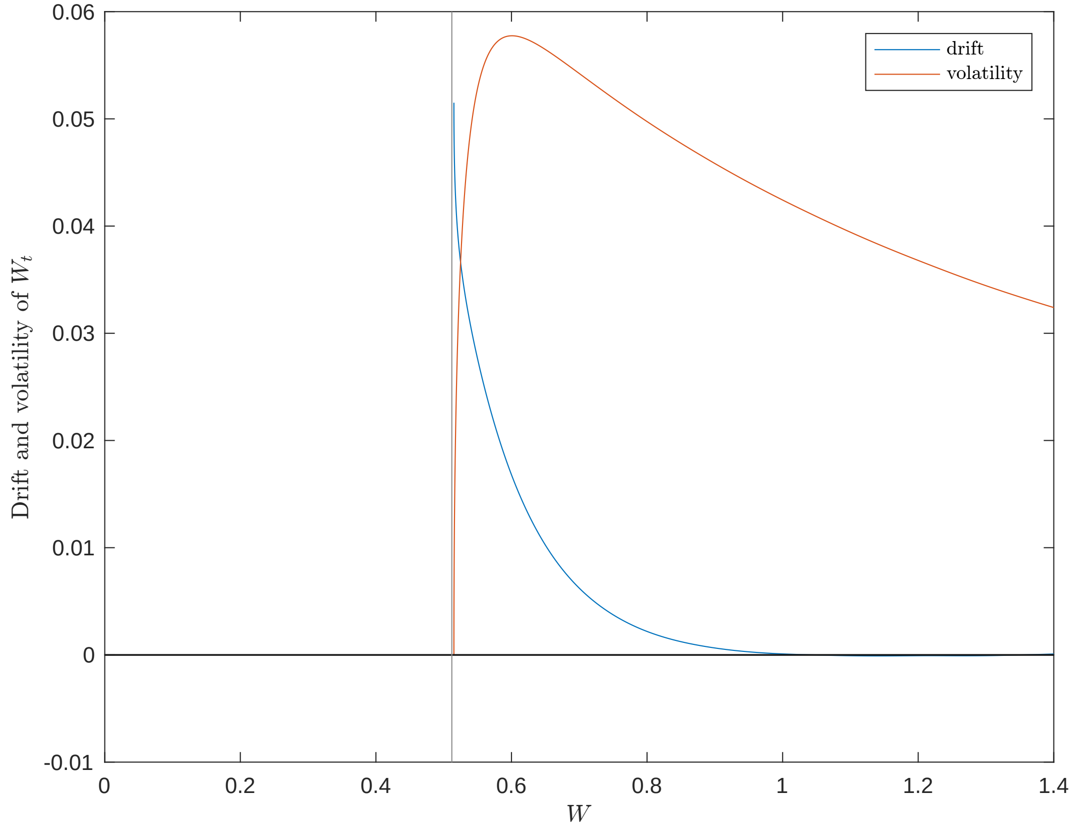

Figure 7: Drift and volatility of W_t with $h(a) = \frac{1}{2}a^2$. The lower bound is the largest equilibrium lower bound $\bar{B}$. Upper bound does not exist. Both drift and volatility are continuous on $[\bar{B}, \infty)$. Reflection is fast.

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 when effort recommendation is nonzero.[9] The fact that $F'(W_t)$ is no longer a martingale when $W_t = B$ is key to eliminating the immiserization result. Since W_t has a drift of $r(W_t - u(C_t)) > 0$ and a volatility of zero at $W_t = B$, $F'(W_t)$ changes deterministically and its drift is $F''(W_t)r(W_t - u(C_t)) < 0$. One might be surprised that nonzero drift at just a single point, B, is strong enough to make the entire process fail to be a martingale. The subtlety here lies in Sticky Brownian Motion: even though B is a singleton, the process W_t spends a positive amount of time at B (see Section 5). In Section 7, where

[9]The 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).

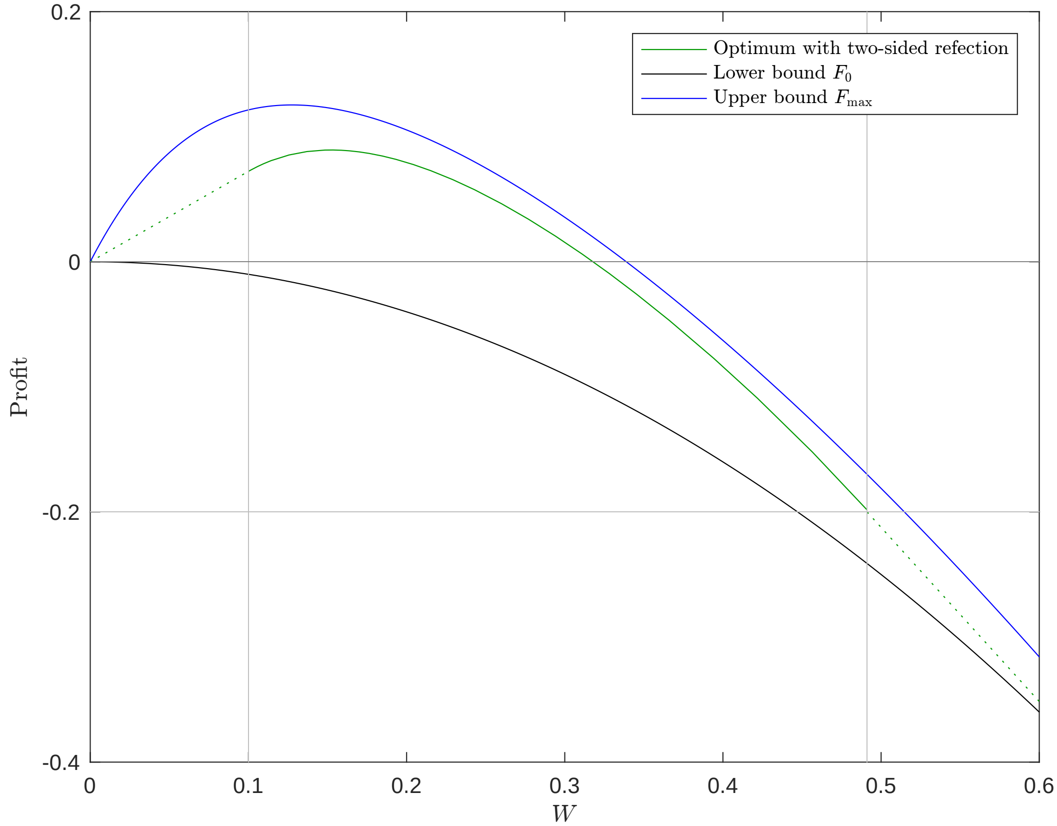

Figure 8: Optimal contract with two-sided limited commitment. The high-action ODE solution curve goes from the vertical line $W = B = 0.1$ to the horizontal line $F = -K = -0.2$. At both ends, it pastes smoothly with low-action ODE solutions. The higher splicing point $D = 0.49$ becomes a reflective upper bound for W_t.

the Inada condition for h holds, W_t spends zero amount of time at B, but the size of the drift $|F''(W_t)r(W_t - u(C_t))|$ is infinity because $F''(B) = -\infty$. An infinite drift over a small time interval (of measure zero) has a nontrivial effect on the dynamics of the process.

8 Extensions

In this section, we extend our analysis to include two significant extensions. First, we allow for limited commitment on the side of the firm. Then, 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 $K \geq 0$.[10] This possibility adds another constraint to the contract design problem:

$$F(W_t) \geq -K \text{ at all } t. \tag{7}$$

That is, a firm cannot credibly promise a contract that allows for 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 the previous section. 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 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 $D > B$.[11] 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 type-2b 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 of D, i.e., D becomes an upper bound for W_t.

Figure 8 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.

[10] That 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.

[11] If $F_{(B,y,\frac{y}{B})}$ 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 (7) 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.

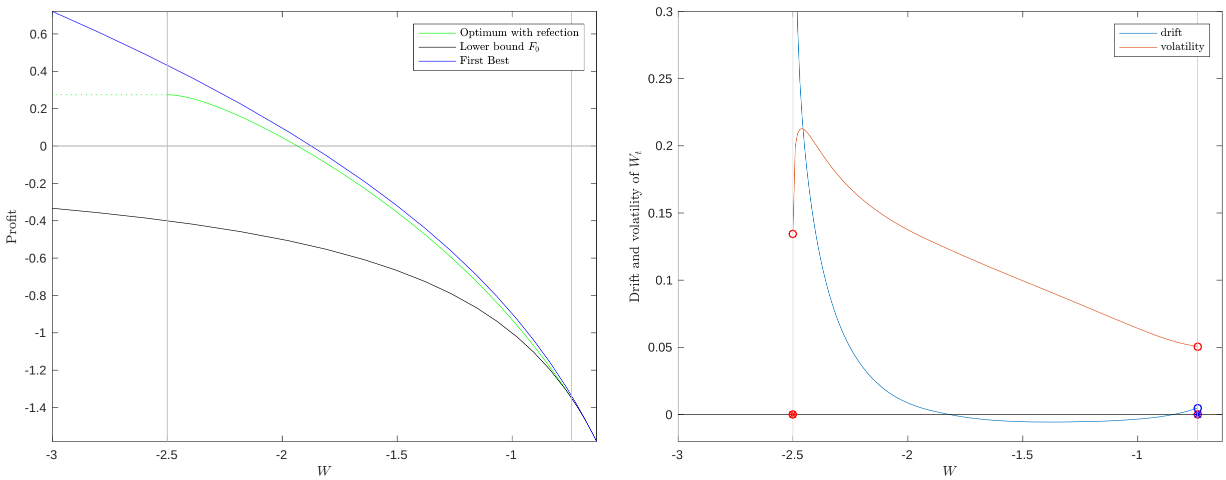

Figure 9: The case of unbounded u. Left panel: the firm's value function. Right panel: dynamics of the optimal contract.

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 often are 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 $\lim_{c \to 0} u(c) \to -\infty$. We show that the optimal contract is reflective at an exogenous lower bound B.[12] 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^*, which restricts the support of W_t to the values above W^*. 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 for W_t and allows the firm to best insure the agent 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

[12] The 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, which is necessary in order to obtain $B > -\infty$.

of B.[13]

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 $F'(B) = 0$.[14] We search for an initial y such that the solution curve $F_{(B,y,0)}$ remains above F_0 touching 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, i.e., compensation c_t is strictly positive everywhere outside of the lower bound $W = B$. As W_t becomes close the lower bound, drift of W_t becomes large. Therefore, reflection off of B is fast. Figure 9 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 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 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 of the lower bound is fast, i.e., the

[13] Informally, we can think of the low-action ODE (5) 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.

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

continuation value process spends zero total time at the lower bound. Another case of fast-reflection dynamics obtains when the agent's utility function is unbounded below.

Appendix

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

$$F''(W) = -\max_{a\geq 0} \underbrace{\frac{a + F'h(a) - F + \max_c \{F'(W - u(c)) - c\}}{r\sigma^2(h'(a))^2/2}}_{H_a(W,F,F')}. \tag{8}$$

The following lemma, analogous to Lemma 2 in Sannikov (2008), orders the solutions to (8).

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*

$$F'(W') < \tilde{F}'(W'), \quad \forall W' > W. \tag{9}$$

Proof This proof modifies the proof of Lemma 2 in Sannikov (2008). First, we show (9) 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

$$F''(W) \leq -\max_{a\geq 0} H_a(W, F(W), F'(W) < -\max_{a\geq 0} H_a(W, \tilde{F}(W), \tilde{F}'(W)) = \tilde{F}''(W).$$

It follows from $F'(W) = \tilde{F}'(W)$ and $F''(W) < \tilde{F}''(W)$ that (9) holds in a small neighborhood of W.

Second, we show (9) for all $W' > W$ by contradiction. Suppose (9) does not hold, then there exists a smallest $W^* > W$ at which $F'(W^*) = \tilde{F}'(W^*)$. Since $F'(W') < \tilde{F}'(W')$ for all $W' \in (W, W^*)$, we have $F(W^*) < \tilde{F}(W^*)$ and again

$$F''(W^*) \leq -\max_{a\geq 0} H_a(W^*, F(W^*), F'(W^*) < -\max_{a\geq 0} H_a(W^*, \tilde{F}(W^*), \tilde{F}'(W^*)) = \tilde{F}''(W^*).$$

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

Proof of Lemma 1

Uniqueness: By contradiction, suppose for some $B' < B$, both $F_{(B',0,0)}$ and $F_{(B,0,0)}$ satisfy conditions a) and b). Since $F_{(B',0,0)}(B) < 0 = F_{(B,0,0)}(B)$ and $F'_{(B',0,0)}(B) < 0 = F'_{(B,0,0)}(B)$, it follows from Lemma A.1 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 $W \geq B$. This contradicts condition b) for $F_{(B,0,0)}(W_{gp})$.

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 $\{\bar{B} - \frac{1}{n} : n = 1, 2, ...\}$, 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 $\{W_{n_k} : k = 1, 2, ...\}$ of $\{W_n : n = 1, 2, ...\}$, then $F_{(\bar{B},0,0)}(W_{gp}) = \lim_{k\to\infty} F_{(\bar{B}-\frac{1}{n_k},0,0)}(W_{n_k}) \leq \lim_{k\to\infty} F_0(W_{n_k}) = F_0(W_{gp})$. 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. QED

Proof of Lemma 2

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

Uniqueness: By contradiction, suppose 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). Since $F_{(B,y_1,\frac{y_1}{B})}(B) < F_{(B,y_2,\frac{y_2}{B})}(B)$ and $F'_{(B,y_1,\frac{y_1}{B})}(B) < F'_{(B,y_2,\frac{y_2}{B})}(B)$, it follows from Lemma A.1 that $F_{(B,y_1,\frac{y_1}{B})}(W) < F_{(B,y_2,\frac{y_2}{B})}(W)$ for all $W \geq B$. 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 $F_{\max}$ means that $F_{\max}(B) - F'_{\max}(B)B > F_{\max}(0) = 0$, so we have $F'_{\max}(B)B < F_{\max}(B)$. This means $F'_{\max}(B) < F'_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(B)$. Lemma A.1 now implies that $F_{\max}(W) < F_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(W)$ for all $W > B$. Because

$F_{\max}(W) \geq F_0(W)$ for all $W \leq W^*_{gp}$, we have $F_{(B,F_{\max}(B),\frac{F_{\max}(B)}{B})}(W) > F_0(W)$ 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 $B \geq \bar{B}$. 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 $\{y - \frac{1}{n} : n = 1, 2, ...\}$, which converges to y from below. Since $y - \frac{1}{n} \notin \{x \geq 0 : F_{(B,x,\frac{x}{B})}(W) \geq F_0(W), \forall W \in [B, W^*_{gp}]\}$, then 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 $\{W_{n_k} : k = 1, 2, ...\}$ of $\{W_n : n = 1, 2, ...\}$, then $F_{(B,y,\frac{y}{B})}(W_{gp}) = \lim_{k\to\infty} F_{(B,y-\frac{1}{n_k},\frac{y-\frac{1}{n_k}}{B})}(W_{n_k}) \leq \lim_{k\to\infty} F_0(W_{n_k}) = F_0(W_{gp})$. 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 $(F_{(B,y,\frac{y}{B})})'_-(W_{gp}) \leq (F_0)'_-(W_{gp})$. QED

The next auxiliary lemma will 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*

$$\min_{c\geq 0} F(W) + c + F'(W)(u(c) - W) \geq 0 \quad \text{at all } W \geq B. \tag{10}$$

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

$$F(W) + F'(W)(W' - W) \geq F_0(W') \quad \forall W' \geq 0. \tag{11}$$

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 $F_{(B,y^*(B),\frac{y^*(B)}{B})}$ 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 (10) follows now from (11) by changing the variable $W' \in [0, \infty)$ to $u(c) \in [0, \infty)$, where c attains $F_0(W')$. ■

Proof of Theorem 1

First, we show that any incentive compatible contract (C, A) achieves profit 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 the restriction 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

$$G_t := r \int_0^t e^{-rt}(A_s - C_s)ds + e^{-rt}F(W_t). \tag{12}$$

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 $A_t = 0$, then (10) 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 principal's future profit is less than or equal to $F_0(W^*_{gp}) \leq F(W^*_{gp})$. Therefore, the principal'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 (12), 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

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

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 principal's future profit is equal to $F_0(W_{gp}) = F(W_{gp})$. Therefore, the principal'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 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 equilibirum. Precisely:

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

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).

If $W < \tilde{W}$, then suppose $V \leq \tilde{V}$. Then $F'_{(\tilde{W},\tilde{V},0)}(\tilde{W}) = 0 = F_{(W,V,0)}(W) > F_{(W,V,0)}(\tilde{W})$. Again Lemma A.1 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). ■

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

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

Proof First, we show the inequality for small $x > 0$. It is sufficient to show that $\mathscr{F}''_W(W) < \mathscr{F}''_{\tilde{W}}(\tilde{W})$, which follows from (14) and $\mathscr{V}(W) < \mathscr{V}(\tilde{W})$.

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

$$\begin{aligned}&\mathscr{F}_W(W-x) - \mathscr{F}'_W(W-x)(W-x)\\ <\ &\mathscr{F}_{\tilde{W}}(\tilde{W}-x) - \mathscr{F}'_{\tilde{W}}(\tilde{W}-x)(\tilde{W}-x).\end{aligned}$$

It follows from (14) again that $\mathscr{F}''_W(W-x) < \mathscr{F}''_{\tilde{W}}(\tilde{W}-x)$ and therefore $\mathscr{F}'_W(W-x+\epsilon) < \mathscr{F}'_{\tilde{W}}(\tilde{W}-x+\epsilon)$ for all sufficiently small $\epsilon > 0$, a contradiction.

Third, we show the monotonicity of $\mathscr{F}_W(W-x) - \mathscr{F}'_W(W-x)(W-x)$ in W. Because

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

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

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

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

■

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, $\mathscr{F}_{W_0(B)}$ and $\mathscr{F}_{W_0(\tilde{B})}$, satisfy the smooth pasting condition with a low-action ODE solution at their respective lower bounds, B and $\tilde{B}$, we have

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

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 $\mathscr{V}$ is strictly decreasing, which was shown in Lemma A.3.

Proofs for Section 7

Writing the high-action ODE (8) as a system of first-order equations, we have

$$\frac{dF}{dW} = F', \tag{13}$$

$$\frac{dF'}{dW} = -\max_{a\geq 0} \frac{a + F'h(a) - F + \max_c\{F'(W-u(c)) - c\}}{r\sigma^2(h'(a))^2/2}. \tag{14}$$

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 (13)-(14) is therefore singlular at $W = B$. We now use a change-of-variable technique to obtain an alternative system that is equivalent to (13)-(14) and well-behaved in the neighborhood of $W = B$.

Change of variable

Dependent variable S. If we define

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

then the HJB equation (3) can be written as

$$S = \max_{a}\{a + F'(W)h(a) + \frac{1}{2}F''(W)r\sigma^2 h'(a)^2\}. \tag{15}$$

Economically, S represents the firm's surplus from being able to induce positive effort from the agent at W. To see this note that

$$F(W) - \max_{c \geq 0}\{-c + F'(W)\,(W - u(c))\} = F(W) - \max_{U \geq 0}\{F_0(U) + F'(W)\,(W - U)\}.$$

If positive effort could not be used at W, then the firm's profit would be $F_0(U)$ minus the cost $-F'(W)(W - U)$ of optimally adjusting the agent's continuation value from W to U, where $-F'(W)$ is the firm's marginal cost of delivering utility to the agent in state W. Therefore, the difference between $F(W)$ and $\max_{U \geq 0}\{F_0(U) + F'(W)\,(W - U)\}$ is the surplus the firm can generate by inducing positive effort from the agent. In the HJB equation (15), we see that this surplus comes from the expected output due to the agent's effort, a, less the cost of compensating the agent for his disutility of effort $h(a)$, less the firm's cost of having to induce positive volatility $Y = h'(a)$ in the state variable. Because the firm always has the option to ask for zero effort, S is always non-negative. Further, because

$$\begin{aligned} F(W) - \max_{U \geq 0}\{F_0(U) + F'(W)\,(W - U)\} &= F(W) + \min_{U \geq 0}\{-F_0(U) - F'(W)\,(W - U)\} \\ &= \min_{U \geq 0}\{F(W) + F'(W)\,(U - W) - F_0(U)\}, \end{aligned}$$

the geometric interpretation of S is the minimum vertical distance between the line tangent to the solution curve F at W and F_0. Because F is concave, the tangent line is always above F, which again shows that S is non-negative. For future reference, note that with S we can also write the HJB equation as

$$F'' = -\max_{a \geq 0} \frac{a + F'h(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}. \tag{16}$$

Inependent 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 $S = F(W) - \max_{U\geq 0}\{F_0(U) - X\,(W-U)\}$ as dependent variables.

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

$$\frac{dW}{dX} = \frac{-1}{F''}. \tag{17}$$

From $-X = \frac{dF}{dW} = \frac{dF}{dX}\frac{dX}{dW} = \frac{dF}{dX}(-F'')$ we have $\frac{dF}{dX} = \frac{X}{F''}$. Further,

$$\begin{aligned}\frac{dS}{dX} &= \frac{d}{dX}\left(F + XW - \max_{U\geq 0}\{F_0(U) + XU\}\right) = \frac{dF}{dW}\frac{dW}{dX} + W + X\frac{dW}{dX} - U(X)\\ &= -X\frac{dW}{dX} + W + X\frac{dW}{dX} - U(X) = W - U(X),\end{aligned}$$

where $U(X) = \arg\max_{U\geq 0}\{F_0(U) + XU\}$.[15] Using the HJB equation (16) to eliminate F'' from (17), we get the following system:

$$\frac{dS}{dX} = W - U(X), \tag{18}$$

$$\frac{dW}{dX} = \frac{1}{\max_{a\geq 0}\frac{a - Xh(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}}. \tag{19}$$

In the new variables (X, W, S), the initial conditions at the lower bound B are as follows. We have $W_0 = B > 0$ and $S_0 = F(B) - F'(B)B = F - \frac{F}{B}B = 0$ due to the smooth pasting with the low-action ODE at B. The starting X is unknown. It is some value $X_0 \leq 0$. When B is the highest lower bound $\bar{B}$, then $X_0 = 0$. For $B < \bar{B}$, $X_0 < 0$.

Example For $u(c) = \sqrt{c}$ and $h(a) = \frac{1}{2}a^2$, which are the functions used in Figure 7, we have $u(c(X)) = \frac{1}{2}X^+$, so $\frac{dS}{dX} = W - \frac{1}{2}X^+$ and $\frac{dW}{dX} = (\max_a \frac{a - X\frac{1}{2}a^2 - S}{\frac{1}{2}r\sigma^2a^2})^{-1}$. The FOC from this maximization problem gives us $a = 2S$, so $\frac{dW}{dX} = \frac{\frac{1}{2}r\sigma^2 4S^2}{2S - X\frac{1}{2}4S^2 - S} = \frac{r\sigma^2 S}{\frac{1}{2} - XS}$, which is well-behaved around the initial condition point, where $S = 0$.

We will show that the system (18)-(19) is well-behaved in general. We start with Lipschitz continuity of (18)-(19) around $S = 0$.

Define

$$\psi(X) := \max_{a\geq 0} a - Xh(a) = \begin{cases} \bar{A} - Xh(\bar{A}), & X \leq 0\\ \bar{a} - Xh(\bar{a}), & X > 0\end{cases}$$

[15] $U(X)$ has the interpretation of the continuation value to which the firm would adjust W if effort must be zero and the marginal cost of utility is X. Clearly, $U(X) = 0$ for $X \leq 0$, and $U(X)$ solves $F_0'(U) + X = 0$ for $X > 0$.

where $\bar{a}$ satisfies $1 = Xh'(\bar{a})$. There are four possibilities:

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

Lemma A.5 *Fix a point $(\hat{X}, \hat{W}, \hat{S})$ in the domain of the ODE system* (18)-(19).

1. *If $\hat{S} \in [0, \psi(\hat{X}))$, then the ODE system in* (18)-(19) *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* (13)-(14) *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 let $M(X, S)$ denote $\max_{a\geq 0} \frac{a-Xh(a)-S}{(h'(a))^2}$ to simplify notation. Pick a small $\epsilon > 0$ such that $S < \psi(X)$ for all $X \in [\hat{X} - \epsilon, \hat{X} + \epsilon]$ and $S \in [\hat{S} - \epsilon, \hat{S} + \epsilon]$. Lipschitz continuity of (19) 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 because $\frac{1}{M(X,S)} = 0$ is constant when $S \leq 0$. The rest of this proof relies on the following equation,
$$\frac{1}{M(X, S_2)} - \frac{1}{M(X, S_1)} = \int_{S_1}^{S_2} \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2} dS, \tag{20}$$
where $a^*(X, S)$ is the maximizer in $\max_{a\geq 0} \frac{a-Xh(a)-S}{(h'(a))^2}$. We can understand (20) heuristically by the envelope theorem, because
$$\partial \frac{1}{M(X, S)} / \partial S = \frac{1}{M(X, S)^2} \frac{1}{(h'(a^*(X, S)))^2}.$$
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 (20) at the end of this proof using an generalized envelope theorem from Milgrom and Segal (2002). Here we shall proceed, assuming (20) is correct.

   To show Lipschitz continuity, it is sufficient to show that the integrand in (20) is bounded when $X \in [\hat{X} - \epsilon, \hat{X} + \epsilon]$ and $S \in [\hat{S} - \epsilon, \hat{S} + \epsilon]$. We show the boundedness as follows. Because $\lim_{a\to 0} h'(a) = 0$, we can pick a small $\delta > 0$ such that for all $a \in (0, \delta)$ and $X \in [\hat{X} - \epsilon, \hat{X} + \epsilon]$,
$$1 - Xh'(a) > \frac{1}{2}. \tag{21}$$

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

$$\begin{aligned}\frac{1}{M(X,S)^2}\frac{1}{(h'(a^*(X,S)))^2} &\leq \frac{1}{M(X,S)^2}\frac{1}{(h'(\delta))^2} \\ &\leq \frac{1}{(\max_{a\geq 0}\frac{a-(\hat{X}+\epsilon)h(a)-(\hat{S}+\epsilon)}{(h'(a))^2})^2}\frac{1}{(h'(\delta))^2},\end{aligned}$$

where the first inequality follows from $a^*(X,S) \geq \delta$, the second inequality is because $\max_{a\geq 0}\frac{a-Xh(a)-S}{(h'(a))^2}$ 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 (21). Therefore, the integrand in (20) satisfies

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

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

Finally, we prove (20) 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

$$M(X,S) = \max_{a\geq 0}\frac{a-Xh(a)-S}{(h'(a))^2} = \max_{a\geq a^*(X,S_1)}\frac{a-Xh(a)-S}{(h'(a))^2}, \quad \forall S \geq S_1.$$

Because $\partial\frac{a-Xh(a)-S}{(h'(a))^2}/\partial S = \frac{-1}{(h'(a))^2}$ is bounded and continuous on the compact set $[a^*(X,S_1),A]$, 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., $M(X,S) \geq M(X,S_2) > 0$), function $\frac{1}{M(X,S)}$ is also absolutely continuous in S and its derivative equals $\frac{1}{M(X,S)^2}\frac{1}{(h'(a^*(X,S)))^2}$ almost everywhere. Absolute continuity of $\frac{1}{M(X,S)}$ implies that

$$\frac{1}{M(X,S_2)} - \frac{1}{M(X,S_1)} = \int_{S_1}^{S_2}\frac{1}{M(X,S)^2}\frac{1}{(h'(a^*(X,S)))^2}dS. \tag{22}$$

Second, we consider the case of $S_1 = 0$. Taking limit $S_1 \downarrow 0$ in (22) 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 (14) satisfies the Lipschitz condition, we first show that the optimal a^* 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, $\max_{a\geq 0} H_a(W,F,F') = \max_{a\geq\epsilon} H_a(W,F,F')$ 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 $a^*(\hat{W},\hat{F},\hat{F}')$. 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 $W \in (\hat{W}-\epsilon, \hat{W}+\epsilon)$, $F \in (\hat{F}-\epsilon, \hat{F}+\epsilon)$, and $F' \in (\hat{F}'-\epsilon, \hat{F}'+\epsilon)$, 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, $\max_{a\geq 0} H_a(W,F,F') = \max_{a\geq\epsilon} H_a(W,F,F')$ whenever $W \in (\hat{W}-\epsilon, \hat{W}+\epsilon)$, $F \in (\hat{F}-\epsilon, \hat{F}+\epsilon)$, and $F' \in (\hat{F}'-\epsilon, \hat{F}'+\epsilon)$.

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

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

These derivatives are uniformly bounded over $a \in [\epsilon, A]$, $W \in (\hat{W}-\epsilon, \hat{W}+\epsilon)$, and $F' \in (\hat{F}'-\epsilon, \hat{F}'+\epsilon)$. As Sannikov (2008) argues in his Lemma 1, $\max_{a\geq\epsilon} H_a(W,F,F')$ is Lipschitz continuous in (F,F').

■

If $S=0$, we can solve (19) in its neighborhood; if $S=\psi(X)$, we can solve (14); if $S \in (0,\psi(X))$, we can solve either (14) or (19) because they are equivalent.

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

1. *There is a $\epsilon > 0$ such that the solution to (18)-(19) 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 (18)-(19) 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), $\lim_{X\to\infty} W(X) = \infty$. In case (2), $\lim_{X\to\bar{X}} W(X) = \infty$. In case (3), $W(\bar{X}) < U(\bar{X}) < \infty$. So in the first two cases, function $F(\cdot)$ implied by the solution to (18)-(19) is a global solution, but in case (3), it is not. In case (3), $\frac{dS}{dX}|_{X=\bar{X}} < 0$, which implies that we cannot extend the solution curve to $X > \bar{X}$ because our model requires $S \geq 0$.*

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 $W^* \geq B$ such that $F(W^*) = F_{fb}(W^*)$, then $F'(W) > (F_{fb})'(W^*), \forall W \geq W^*$. Moreover, case (2) in part 2 applies, which implies that $S > 0, \forall W \geq W^*$.

Proof

1. $S > 0$ because $\frac{dS}{dX}|_{X=\underline{X}} = W - U(\underline{X}) = B > 0$. If $S \in (0, \psi(X))$, then $\max_{a \geq 0} \frac{a - Xh(a) - S}{\frac{r}{2}\sigma^2 (h'(a))^2} > 0$ is positive and finite. Therefore,

$$\frac{dW}{dX} = \frac{1}{\max_{a\geq 0} \frac{a-Xh(a)-S}{\frac{r}{2}\sigma^2(h'(a))^2}} > 0. \tag{23}$$

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

   In case (2), by contradiction, suppose $\bar{W} := W(\bar{X}) < \infty$. Then (14) 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 (13)-(14). On the other hand, the solution to (18)-(19) 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 (13)-(14) at $(\bar{W}, F(\bar{W}), F'(\bar{W})$, contradicting the result that (13)-(14) satisfies the Lipschitz condition (part 2 of Lemma 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 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

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

   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 $F(W) \leq F_{fb}(W), \forall W \leq W^*$ that $F'(W^*) \geq (F_{fb})'(W^*) \equiv -X$. To show $F'(W) > (F_{fb})'(W^*), \forall W \geq W^*$, suppose by contradiction $W \in [W^*, \infty)$ is the smallest W such that $F'(W) = (F_{fb})'(W^*) = -X$. Since F is concave, $F(W)+F'(W)(W^*-W) \geq F(W^*) = F_{fb}(W^*)$, which implies

$$F(W) + XW \geq F_{fb}(W^*) + XW^* = F_0(U_0(X)) + XU_0(X) + \psi(X),$$

or $S \geq \psi(X)$. The fact that W is finite contradicts case (2) in part 2. Therefore, $F'(W) > (F_{fb})'(W^*), \forall W \geq W^*$ and the solution curve 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 $\bar{X} \geq -F'(W^*)$. Proof of case (3) in part 2 shows that $U(\bar{X}) > W(\bar{X}) \geq W^*$, which and the property $F'(W) > (F_{fb})'(W^*), \forall W \geq W^*$ imply

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

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

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 defined by $F_{\text{fb}}(M) = 0$. Define $\mathcal{U}$ as the set of initial conditions under which the solution crosses the upper bound F_{fb}.

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

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

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

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

   To show $\mathcal{L} \neq \emptyset$, we show that $B \in \mathcal{L}$ when B is sufficiently small. If $B = 0$, then $\frac{dS}{dX}|_{X=0} = W - U(X) = B = 0$, and $\frac{d^2S}{dX^2}|_{X=0} = \frac{dW}{dX} - \frac{dU(X)}{dX} < 0$. This implies that $S(X^*) < 0$ for some small $X^* > 0$. Because the ODE system satisfies the Lipschitz condition, its solution is continuous in the initial condition B. That is, $S(X^*) < 0$ 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 $F_{(B,0,0,-\infty)}(W) > F_{\text{fb}}(W)$ for some $W > B$. Suppose this happens at $W = W(X)$ in the ODE (14). Since the solution to (14) is continuous in its initial conditions, $(W(X,B), S(X,B))$ are continuous in y. Moreover, $F_{\text{fb}}(W)$ is continuous in W and $F_{(B,0,0,-\infty)}(W) = S(X,B) + F_0(U(X)) + X(U(X) - W)$ is continuous in (W, X). Therefore,

   $$\begin{aligned} & F_{(B,0,0,-\infty)}(W(X,B)) - F_{\text{fb}}(W(X,B)) \\ = \; & S(X,B) + F_0(U(X)) + X(U(X) - W(X,B)) - F_{\text{fb}}(W(X,B)) \end{aligned}$$

   is continuous in y. There exists an $\epsilon > 0$ such that for all $\tilde{B} \in (B - \epsilon, B + \epsilon)$, $F_{(\tilde{B},0,0,-\infty)}(W(X,\tilde{B})) - F_{\text{fb}}(W(X,\tilde{B})) > 0$. 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 $S(X') < 0$ if X' is slightly above X. Since the solution to (14) is continuous in its initial conditions, $S(X', B)$ is continuous in B. There exists an $\epsilon > 0$ such that for all $\tilde{B} \in (B - \epsilon, B + \epsilon)$, $S(X', \tilde{B}) < 0$. For each $\tilde{B} \in (B - \epsilon, B + \epsilon)$, $S(X', \tilde{B}) < 0$ and $S(\tilde{X}, \tilde{B}) > 0$ when $\tilde{X}$ is slightly above 0, therefore $S(., \tilde{B})$ reaches 0. Therefore, $(B - \epsilon, B + \epsilon) \subset \mathcal{L}$.

3. Because of the above properties, $\mathcal{U} \cup \mathcal{L} \neq (0, M]$. 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.

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 1 implies that $F_{(\bar{B}^1,0,0,-\infty)}(W) - F_{(\bar{B}^2,0,0,-\infty)}(W)$ is increasing in W. In particular, if

$\bar{B}^2 < W$,

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

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

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

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 $F'_{(\bar{B},0,0,-\infty)}(W) < F'_{(B,0,0,-\infty)}(W) \equiv -X$. 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 $y \in (0, \bar{B})$ is in $\mathcal{U}$, then $F_{(B,0,0,-\infty)}(W) = F_{\text{fb}}(W)$ for some W. Lemma A.1 implies that $F_{(\bar{B},0,0,-\infty)}(W) > F_{(B,0,0,-\infty)}(W) = F_{\text{fb}}(W)$, 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}.

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

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

$\mathcal{L} := \{y \in [0, F_{\text{fb}}(B)]$: there exists a $W > B$ such that $F''(W) = -\infty$, 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 $F_{(B,y,\frac{y}{B},-\infty)}(W) > F_{\text{fb}}(W)$ for some $W > B$. Suppose this happens at $W = W(X)$ in the ODE (14). Since the solution to (14) 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 $F_{(B,y,\frac{y}{B},-\infty)}(W) = S(X,y) + F_0(U(X)) + X(U(X) - W)$ is continuous in (W, X). Therefore,

$$\begin{aligned} & F_{(B,y,\frac{y}{B},-\infty)}(W(X,y)) - F_{\text{fb}}(W(X,y)) \\ = \quad & S(X,y) + F_0(U(X)) + X(U(X) - W(X,y)) - F_{\text{fb}}(W(X,y)) \end{aligned}$$

   is continuous in y. There exists $\epsilon > 0$ such that for all $\tilde{y} \in (y-\epsilon, y+\epsilon)$, $F_{(B,\tilde{y},\frac{\tilde{y}}{B},-\infty)}(W(X,\tilde{y})) - F_{\text{fb}}(W(X,\tilde{y})) > 0$. 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 (14) 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 $-\frac{y}{B}$, therefore $S(\cdot, \tilde{y})$ reaches 0. Therefore, $(y-\epsilon, y+\epsilon) \subset \mathcal{L}$.

3. Because of the above properties, $\mathcal{U} \cup \mathcal{L} \neq [0, F_{\text{fb}}(B)]$. 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 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 $W \to \infty$, we have

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

   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 1 implies that $F_{(B,y^*,\frac{y^*}{B},-\infty)}(W) < F_{(B,y,\frac{y}{B},-\infty)}(W)$ and $F'_{(B,y^*,\frac{y^*}{B},-\infty)}(W) < F'_{(B,y,\frac{y}{B},-\infty)}(W) \equiv -X$. Therefore,

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

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 $F_{(B,y,\frac{y}{B},-\infty)}(W) = F_{\text{fb}}(W)$ for some W. Lemma A.1 implies that $F_{(B,y^*,\frac{y^*}{B},-\infty)}(W) > F_{(B,y,\frac{y}{B},-\infty)}(W) = F_{\text{fb}}(W)$, which means that $y^* \in \mathcal{U}$, a contradiction.

Proof of Theorem 2 (verification in the Inada case)

First, we show that any incentive compatible contract (C, A) achieves profit at most $F_{(B,y^*(B),\frac{y^*(B)}{B},-\infty)}(W_0(C, A))$. To simplify the notation, we will often drop the subscript in $F_{(B,y^*(B),\frac{y^*(B)}{B},-\infty)}$ and refer to this solution simply as F in this section. Let $\epsilon > 0$ be a small number. Since F asymptotically approaches F_{fb} as $W \to \infty$, there is a large $\bar{W} > W_0(C, A)$ such that $F(\bar{W}) + \epsilon \geq F_{\text{fb}}(\bar{W})$. Define

$$G_t := r \int_0^t e^{-rt}(A_s - C_s)ds + e^{-rt}F(W_t). \tag{24}$$

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 (10) holds also for $F = F_{(B,y^*(B),\frac{y^*(B)}{B},-\infty)}$. Inequality (10) 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 $t < \infty$,

$$\begin{aligned}
&\mathbb{E}\left[r\int_0^{\tau'} e^{-rt}(A_t - C_t)dt + e^{-r\tau'}F(\bar{W})\right] \\
= \;& \mathbb{E}\left[G_{t\wedge\tau'} + 1_{t\leq\tau'}\left(r\int_t^{\tau'} e^{-rs}(A_s - C_s)ds + e^{-r\tau'}F(\bar{W}) - e^{-rt}F(W_t)\right)\right] \\
= \;& \mathbb{E}\left[G_{t\wedge\tau'}\right] + e^{-rt}\mathbb{E}\left[1_{t\leq\tau'}\left(r\int_t^{\tau'} e^{-r(s-t)}(A_s - C_s)ds + e^{-r(\tau'-t)}F(\bar{W}) - F(W_t)\right)\right] \\
\leq \;& G_0 + e^{-rt}\left[\bar{A} + 2\max_{W\in[B,\bar{W}]} F(W)\right].
\end{aligned}$$

Taking $t \to \infty$ 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 principal's future profit is less than or equal to $F_{\text{fb}}(\bar{W}) \leq F(\bar{W}) + \epsilon$. Therefore, the principal'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 principal'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 (24), 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

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

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

$$\mathbb{E}\left[r\int_0^t e^{-rs}(A_s - C_s)ds\right] \geq F(W_0) - e^{-rt}\max_W F(W). \tag{25}$$

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 (25) imply

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

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

QED

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

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

Proof Using L-Hopital's rule,

$$\lim_{X\to -F'(B)} \frac{(h'(a))^2}{W - B} = \lim_{X\to -F'(B)} \frac{2h'(a)\frac{dh'(a)}{dX}}{\frac{dW}{dX}}. \tag{26}$$

We need to evaluate $\frac{dh'(a)}{dX}$ and $\frac{dW}{dX}$. The FOC with respect to a taken in the HJB equation (3) is

$$1 + F'(W)h'(a) + F''(W)r\sigma^2 h'(a)h''(a) = 0. \tag{27}$$

The sufficiency of this first-order condition is shown in Lemma A.8. With effort a that satisfies this FO condition, we can drop the max operator in (16) and write

$$F''(W) = -\frac{a + F'(W)h(a) - S}{\frac{1}{2}r\sigma^2(h'(a))^2}, \tag{28}$$

which we use in (17) to obtain

$$\frac{dW}{dX} = \frac{-1}{F''(W)} = \frac{\frac{1}{2}r\sigma^2(h'(a))^2}{a - Xh(a) - S}. \tag{29}$$

Next, we substitute (28) into (27) to obtain

$$1 + F'(W)h'(a) + \frac{S - a - F'(W)h(a)}{\frac{1}{2}h'(a)}h''(a) = 0,$$

or

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

Using (19), we now totally differentiate this equation with respect to X

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

Using $\frac{da}{dX} = \frac{1}{h''}\frac{dh'}{dX}$, $\frac{dh''}{dX} = h'''\frac{da}{dX} = \frac{h'''}{h''}\frac{dh'}{dX}$, we solve for $\frac{dh'(a)}{dX}$ to obtain

$$\frac{dh'(a)}{dX} = \frac{2h(a)h''(a) - (h'(a))^2 + 2(W - u(c(X)))h''(a)}{1 - 2(S - a + xh(a))\frac{h'''(a)}{h''(a)}},$$

which converges to $2(B - u(0))h''(0) = 2Bh''(0) > 0$ at the lower bound. Using this in (26), we have

$$\lim_{X \to -F'(B)} \frac{h'(a)^2}{W - B} = 4Bh''(0) \lim_{X \to -F'(B)} \frac{h'(a)}{\frac{dW}{dX}}.$$

Also, from $\frac{dh'(a)}{dX} \to 2Bh''(0)$ follows that $\frac{da}{dX} = \frac{1}{h''}\frac{dh'}{dX} \to 2B$ at the lower bound. Next, substituting from (29) and simplifying, we obtain

$$\lim_{X \to -F'(B)} \frac{h'(a)^2}{W - B} = \frac{8Bh''(0)}{r\sigma^2} \lim_{X \to -F'(B)} \frac{a - Xh(a) - S}{h'(a)}.$$

Using L-Hopital's rule again,

$$\begin{aligned}
\lim_{X\to -F'(B)} \frac{h'(a)^2}{(W-B)} &= \frac{8Bh''(0)}{r\sigma^2} \lim_{X\to -F'(B)} \frac{\frac{da}{dX} - h(a) - Xh'(a)\frac{da}{dX} - (W-u(c(X)))}{h''(a)\frac{da}{dX}} \\
&= \frac{8Bh''(0)}{r\sigma^2} \frac{2B - 0 + F'(B)02B - (B-u(0))}{h''(a)2B} \\
&= \frac{8Bh''(0)}{r\sigma^2} \frac{B}{h''(0)2B} \\
&= \frac{4B}{r\sigma^2}.
\end{aligned}$$

■

Proof of Proposition 3 (fast reflection)

We consider the diffusion of W_t with no killing measure (i.e., $k \equiv 0$):

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

where Z_t is a standard Brownian motion. Let m be the speed measure of the diffusion process W_t and s be the scale function. Using the formula on page 17 of Borodin and Salminen (2002), we calculate the density function of the speed measure m as follows. Define $\mathcal{B}(x) := \int_C^x \frac{2r(W-u(c)+h(a))}{(rh'(a)\sigma)^2} dW$, where $C \in (B, \infty)$ is a constant. Then

$$\begin{aligned}
m(x) &= \frac{2e^{\mathcal{B}(x)}}{(rh'(a(x))\sigma)^2}, \quad \forall x > B, \\
m((a,z)) &= \int_a^z m(x)dx, \quad \forall z > a > B.
\end{aligned}$$

The scale function s satisfies

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

First, we show that B is a nonsingular point, that is, B is both an exit and an entrance for the diffusion W_t. To do so, 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,

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

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

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

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

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

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

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

Moreover, if $x < C$,

$$\begin{aligned}
\frac{2\left(\frac{x-B}{C-B}\right)^{(\frac{1}{2}+\epsilon)}}{r^2\sigma^2(\frac{4B}{r\sigma^2}+\epsilon)(W-B)} &\leq \frac{2\left(\frac{x-B}{C-B}\right)^{(\frac{1}{2}+\epsilon)}}{(rh'(a(x))\sigma)^2} \\
&\leq m(x) := \frac{2e^{\mathcal{B}(x)}}{(rh'(a(x))\sigma)^2} \\
&\leq \frac{2\left(\frac{x-B}{C-B}\right)^{(\frac{1}{2}-\epsilon)}}{r^2\sigma^2(\frac{4B}{r\sigma^2}-\epsilon)(W-B)}.
\end{aligned}$$

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

$$\lim_{a\downarrow B}\int_a^z m(x)dx \leq \lim_{a\downarrow B}\int_a^z \frac{2\left(\frac{x-B}{C-B}\right)^{(\frac{1}{2}-\epsilon)}}{r^2\sigma^2(\frac{4B}{r\sigma^2}-\epsilon)(W-B)}dx < \infty.$$

B is an exit because

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

B is an entrance because

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

Because the boundary point B is nonsingular, W_t reaches B with positive probability.

Second, we show that $m(\{B\}) = 0$. Pick a bounded and smooth function f, and define

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

Also, define $f^+(x) := \frac{f'(x)}{s'(x)}$. Note that $\lim_{x\downarrow B} s'(x) \geq \lim_{x\downarrow B} \left(\frac{x-B}{C-B}\right)^{-(\frac{1}{2}-\epsilon)} = \infty$. Then equation (c) in page 16 of Borodin and Salminen (2002) implies that

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

Finally, since $m(\{B\}) = 0$, the diffusion process W_t is *reflecting*. A reflecting process has the property that $P_x(Leb(\{t \geq 0 : W_t = B\}) = 0) = 1$ for all $x \geq B$.

QED

The last lemma justifies the first-order condition (27) used in Lemma A.7.

Lemma A.8 *Pick* $\delta > 0$ *such that*

$$\left|\frac{h'(a)h'''(a)}{(h''(a))^2}\right| < \frac{1}{2}, \quad \forall a \in [0, \delta].$$

There exists $\epsilon > 0$*, such that for all* $W \in [B, B+\epsilon]$*, the maximum* $\max_a a + F'(W)h(a) + F''(W)r\sigma^2(h'(a))^2$ *a) is not achieved on* $[\delta, \bar{A}]$*, b) is strictly concave in* $[0, \delta]$ *and c) the maximizer is given by the first-order condition.*

Proof Both $\min_{a\in[\delta,\bar{A}],W\in[B,W_{gp}]} \frac{-a-F'(W)h(a)}{r\sigma^2(h'(a))^2}$ and $\min_{a\in[0,\delta],W\in[B,W_{gp}]} \frac{-2F'(W)}{r\sigma^2 h''(a)}$ are finite. It follows from $\lim_{W\downarrow B} F''(W) = -\infty$ that, there is $\epsilon > 0$ such that for all $W \in [B, B+\epsilon]$,

$$F''(W) < \min\left\{\min_{a\in[\delta,\bar{A}],W\in[B,W_{gp}]} \frac{-a - F'(W)h(a)}{r\sigma^2(h'(a))^2}, \min_{a\in[0,\delta],W\in[B,W_{gp}]} \frac{-2F'(W)}{r\sigma^2 h''(a)}\right\}.$$

Equivalently, if $a \in [\delta, \bar{A}]$ and $W \in [B, B+\epsilon]$, then

$$a + F'(W)h(a) + F''(W)r\sigma^2(h'(a))^2 < 0, \tag{30}$$

and if $a \in [0, \delta]$ and $W \in [B, B+\epsilon]$, then

$$F'(W) + \frac{1}{2}F''(W)r\sigma^2 h''(a) < 0. \tag{31}$$

It follows from (30) that, if $W \in [B, B+\epsilon]$, then $a + F'(W)h(a) + F''(W)r\sigma^2(h'(a))^2$ cannot achieve a maximum in $a \in [\delta, \bar{A}]$.

To show that $a + F'(W)h(a) + F''(W)r\sigma^2(h'(a))^2$ is strictly concave in $[0, \delta]$, we check its second derivative. The second derivative is

$$\begin{aligned} & F'(W)h''(a) + F''(W)r\sigma^2(h''(a)h''(a) + h'(a)h'''(a)) \\ < \; & F'(W)h''(a) + F''(W)r\sigma^2\frac{1}{2}h''(a)h''(a) \\ = \; & (F'(W) + F''(W)r\sigma^2\frac{1}{2}h''(a))h''(a) \\ < \; & 0, \end{aligned}$$

where the second inequality follows from (31).

Finally, we verify that the first-order condition

$$1 + F'(W)h'(a) + F''(W)r\sigma^2 h'(a)h''(a) = 0$$

pins down a unique $a \in (0, \delta)$. If $a = 0$, then $1 + F'(W)h'(a) + F''(W)r\sigma^2 h'(a)h''(a) > 0$, violating the first-order condition.

■

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.

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).

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.

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.

Zhu, J. Y. (2013). Optimal contracts with shirking. *Review of Economic Studies 80*(2), 812–839.