Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Comment on "Optimal Contract to Induce Continued Effort"

Ping Cao, Feng Tian, Peng Sun

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–13
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Comment on "Optimal Contract to Induce Continued Effort"

Ping Cao,[a] Feng Tian,[b] Peng Sun[c]

[a] School of Management, University of Science and Technology of China, Hefei 230026, China; [b] The Stephen M. Ross School of Business, University of Michigan, Ann Arbor, Michigan 48105; [c] The Fuqua School of Business, Duke University, Durham, North Carolina 27708

Contact: pcao@ustc.edu.cn, https://orcid.org/0000-0001-7801-9754 (PC); ftor@umich.edu, https://orcid.org/0000-0002-5211-2384 (FT); psun@duke.edu, https://orcid.org/0000-0002-3892-5936 (PS)

Received: May 1, 2020
Revised: June 30, 2020; November 19, 2020
Accepted: November 23, 2020
Published Online in Articles in Advance: March 25, 2021

https://doi.org/10.1287/mnsc.2020.3927

Abstract. In this comment, we first use a counterexample to demonstrate that the optimal contract structure proposed in section 4 of Sun and Tian (2018) can be wrong when the two players' discount rates are different. We then specify correct optimal contract structures, which involve generalizing the contract space to allow random termination. Numerical study with a wide range of model parameters illustrates that such a random termination only occurs sparingly in optimal contracts. Moreover, the suboptimality gap, measured by the relative improvement of the optimal contract over the best contract without random termination, is extremely small.

History: Accepted by Manel Baucells, decision analysis.
Funding: The work of P. Cao was supported in part by the National Natural Science Foundation of China [Grants 71771202 and 71520107002].

Keywords: dynamic • moral hazard • optimal control • jump process

1. Basic Setup

To ensure this comment is self-contained, we start it by introducing the dynamic moral hazard model and relative notations, first introduced in Sun and Tian (2018). We keep the introduction terse and refer to the original paper for motivation and justification of the model. A principal hires an agent to increase the arrival rate of a Poisson process, which yields the principal a revenue R per arrival over an infinite time horizon. Without the agent or when the agent does not exert effort, the corresponding instantaneous arrival rate is $\underline{\mu}$. Exerting effort allows the agent to increase the instantaneous arrival rate to μ but costs the agent a constant rate c per unit of time. Both the principal and the agent are risk neutral. The principal's discount rate is r, and the agent's discount rate is ρ.

The principal does not observe the effort and is able to commit to a long-term contract that involves history-dependent payments and a random time to terminate the agent. Formally, at any time $t \in [0,\infty)$, we denote $N = \{N_t\}_{t\geq 0}$ to represent the counting process that represents the number of arrivals up to and including time t. We also let $\mathcal{F}$ be the filtration generated by the process N. Denote effort to be an $\mathcal{F}$-predictable process $\nu = \{\nu_t\}_{t\geq 0}$, with $\nu_t \in \{\mu, \underline{\mu}\}$. The principal has commitment power to design and implement long-term contracts, which involve a payment process $L = \{L_t\}_{t\geq 0}$ that is $\mathcal{F}$-adapted, and a termination time τ, an $\mathcal{F}$-random time. Assume the agent has limited liability and is cash constrained, so that $dL_t \geq 0$ for all $t \geq 0$. Payment dL_t at time t can be decomposed into $dL_t = \Delta L_t + \ell_t dt$, in which ΔL_t represents instantaneous and ℓ_t flow payment.

Given a dynamic contract $\Gamma = (L,\tau)$ and an effort process ν, the expected discounted utility of the agent is

$$u(\Gamma,\nu) = \mathbb{E}^{\nu}\left[\int_0^{\tau} e^{-\rho t}\left(dL_t - c\mathbb{1}_{\nu_t=\mu}dt\right)\right], \tag{1}$$

in which the expectation $\mathbb{E}^{\nu}$ is taken with respect to probabilities generated from the effort process ν.

We focus on *Effort-Inducing* (EI) contract,[1] such that $u(\Gamma,\bar{\nu}) \geq u(\Gamma,\nu)$, for all ν, in which $\bar{\nu} := \{\nu_t = \mu\}_{\forall t\in[0,\tau]}$ represents the "always exerting effort before contract termination" strategy by the agent. The principal's utility under an EI contract Γ is

$$U(\Gamma) = \mathbb{E}\left[\int_0^{\tau} e^{-rt}(RdN_t - dL_t) + e^{-r\tau}\underline{v}\right], \text{ in which } \underline{v} := \underline{\mu}R/r. \tag{2}$$

Here, $\underline{v}$ is the principal's baseline total discounted revenue after terminating the agent, and we omit the superscript $\bar{\nu}$ in the expectation under EI contracts. We want to maximize the principal's utility $U(\Gamma)$ over EI contracts Γ. The optimal contract design problem can be formulated as optimal control. In the rest of this comment, we use "contract" and "control policy" interchangeably.

Sun and Tian (2018) correctly identify the optimal contract structure when the discount rate $r = \rho$. However, their section 4, which studies the case of $r < \rho$, contains errors, such that the result is not correct. In this comment, we focus on the case that r is strictly less than ρ: that is, the principal is more patient than the agent.

Before closing this section, we introduce a few more notations, consistent with Sun and Tian (2018). Denote $\Delta\mu := \mu - \underline{\mu} > 0$, $\beta := c/\Delta\mu$, and assume $R \geq \beta$. Consider a simple suboptimal contract, $\bar{\Gamma}$, which pays the agent β for each arrival and never terminates. Denote $\bar{U}$ and $\bar{w}$ as the principal and agent's utilities under contract $\bar{\Gamma}$, respectively, and $\bar{V}$ as the principal's utility if effort is observable. That is,

$$\bar{U} := \frac{\mu(R-\beta)}{r}, \quad \bar{w} := \frac{\mu\beta - c}{\rho} = \frac{\beta\underline{\mu}}{\rho}, \quad \text{and} \quad \bar{V} := \frac{\mu R - c}{r}. \tag{3}$$

2. A Counterexample with a Nonconcave Value Function

Denote the principal and agent's total value function under the optimal contract as $V(w)$, in which state variable w represents the agent's promised utility. In its section 4, Sun and Tian (2018) claims that the following delay differential equation (DDE) uniquely determines this optimal value function V_d along with an upper bound $\hat{w} < \bar{w}$ of the promised utility,

$$0 = (r+\mu)V_d(w) - \mu V_d(w+\beta) + \rho(\bar{w}-w)V_d'(w) + (c - \mu R) + (\rho - r)w, \text{ for } w \in [0, \hat{w}), \tag{4}$$

with boundary conditions

$$V_d(w) = \bar{V}_{\hat{w}} := \bar{V} - \frac{\rho - r}{r}\hat{w}, \text{ for } w \geq \hat{w}, \tag{5}$$

and

$$V_d(0) = \underline{v}. \tag{6}$$

In the proof of optimality, it is crucial to establish that the function V_d is concave. Although DDE (4)–(6) is indeed closely related to the optimal value function, its solution, V_d, in fact, may not be concave, as we demonstrate here using a counterexample.

Following the first part of proposition 4 in Sun and Tian (2018), we know that given any $\tilde{w} \in [0, \bar{w})$, DDE (4) with boundary Condition (5) has a unique solution, $V_{\tilde{w}}(w)$ on $w \in \mathbb{R}_+$. Instead of fixing model parameters and searching for a particular value $\hat{w}$ such that $V_{\hat{w}}$ satisfies the boundary Condition (6), here we argue that for any $\tilde{w}$, there is a revenue parameter R that yields (6). In the following Proposition 1, we write $V_{\tilde{w}}(w; R)$ in place of $V_{\tilde{w}}(w)$, to highlight the dependence on R.

Proposition 1. *For any given $\tilde{w} \in [0, \bar{w}]$, the function $\psi(R, \tilde{w}) := V_{\tilde{w}}(0; R) - \underline{v}$ is strictly increasing and linear in R, satisfying $\psi(\beta, \tilde{w}) < 0$. Therefore, there exists a unique $\hat{R}(\tilde{w}) > \beta$ such that $\psi(\hat{R}(\tilde{w}), \tilde{w}) = 0$.*

Proposition 1 implies that there exists a revenue parameter R such that the solution V_d and $\hat{w}$ to (4)–(6) is such that $\hat{w}$ can be arbitrarily close to $\bar{w}$.

To clearly construct the counterexample, we consider the case that $r + \mu = 2\rho$, and $\underline{\mu} \in (\rho, 2\rho)$. Therefore, $\bar{w} \in (\beta, 2\beta)$. Choose a $\tilde{w}$ such that $\tilde{w} > \beta$, and $R = \hat{R}(\tilde{w})$. In this case, one can verify that $V_{\tilde{w}}$ has the following closed form solution,

$$V_{\tilde{w}}(w) = \frac{\rho - r}{\rho}(\bar{w} - w) + \frac{\mu\bar{V}_{\tilde{w}} + r\bar{V} - (\rho - r)\bar{w}}{r + \mu} - \frac{\rho - r}{2\rho(\bar{w} - \tilde{w})}(\bar{w} - w)^2, \text{ for } w \in [\tilde{w} - \beta, \tilde{w}], \tag{7}$$

which is concave on $[\tilde{w} - \beta, \tilde{w}]$ because $V_{\tilde{w}}''(w) = -\frac{\rho - r}{\rho(\bar{w} - \tilde{w})} < 0$.

Twice differentiating (4) over $[0, \tilde{w})$ yields

$$\rho(\bar{w} - w)V_{\tilde{w}}'''(w) = \mu V_{\tilde{w}}''(w + \beta) + (2\rho - r - \mu)V_{\tilde{w}}''(w) = \mu V_{\tilde{w}}''(w + \beta),$$

in which the last equality follows from $r + \mu = 2\rho$. Hence, for $w \in [0, \tilde{w} - \beta)$, we have

$$\rho(\bar{w} - w)V_{\tilde{w}}'''(w) = \mu V_{\tilde{w}}''(w + \beta) = -\frac{\mu(\rho - r)}{\rho(\bar{w} - \tilde{w})}.$$

Together with the closed form Expression (7), we have, for $w \in [0, \tilde{w} - \beta)$,

$$V_{\tilde{w}}''(w) = \frac{\mu(\rho - r)}{\rho^2(\bar{w} - \tilde{w})}\ln(\bar{w} - w) - \frac{\mu(\rho - r)}{\rho^2(\bar{w} - \tilde{w})}\ln(\bar{w} - w + \beta) - \frac{\mu(\rho - r)}{\rho(\bar{w} - \tilde{w})}.$$

In particular,

$$V_{\tilde{w}}''(0) = \frac{(\rho - r)}{\rho(\bar{w} - \tilde{w})}\left[\frac{\mu}{\rho}\ln\left(\frac{\bar{w}}{\bar{w} - \tilde{w} + \beta}\right) - 1\right]. \tag{8}$$

Take $\rho = 1$, $r = 0.05$, $\mu = 1.95$, $\underline{\mu} = 1.85$, and $c = 0.3$. Therefore, $\beta = c/\Delta\mu = 3$ and $\bar{w} = \underline{\mu}\beta/\rho = 5.55$. Further, take a particular value $\hat{w}$ in place of $\tilde{w}$ such that $\hat{w} = 0.98 \times \bar{w} = 5.439$. (One can calculate that the corresponding $\hat{R}(\hat{w}) = 112.4622$, which does not affect the following calculation.) In this case, Equation (8) becomes

$$V_{\hat{w}}''(0) = \frac{(\rho - r)}{\rho(\bar{w} - \hat{w})}\left[1.95\ln\left(\frac{1.85\beta}{(1.85 \times 0.02 + 1)\beta}\right) - 1\right] > 0,$$

which implies that this $V_{\hat{w}}$ function is not concave near 0. Therefore, function V_d, which is the solution to

DDE (4)–(6) on $[0, \hat{w}]$, in general may not be concave and, therefore, may not be the optimal value function.

In order to construct an optimal value function that is indeed concave, we need to "concavificate" the value function. In fact, section 4 of Sun and Tian (2018) provides a discrete time approximation, which involves obtaining a concave upper envelope of the value function. Specifically, they show that in the discrete time model, concavification only occurs near zero, such that the discrete time optimal value function is linear when the promised utility w is below a threshold. Unfortunately, Sun and Tian (2018) erroneously claimed that the threshold would converge to zero when the discrete time model converges to continuous time, which led to a wrong proposition 5. In fact, even in the continuous time model, under some model parameter settings, we still need to include a linear piece in the optimal value function near zero to ensure its concavity. Also, the optimal value function is the solution to DDE (4) only when w is above a potentially positive threshold. In the control policy space, this corresponds to randomized termination.

3. Optimal Contracts When $\rho > r$

Following the discussion in the previous section, we know that the control policy space in Sun and Tian (2018) is not rich enough to capture the optimal contract. That is, optimality may not be attainable in this continuous time model according to the contract space specified in Section 1. In order to establish an optimal control policy, we need to generalize the contract space or the space of admissible controls.

In particular, we need to allow the principal to randomly terminate the agent, according to a rate q_t. That is, the agent is terminated in a short time interval $(t, t+\delta]$ with probability $q_t\delta + o(\delta)$. As a part of the control, we only require that

$$\mathbb{E}\left[\int_0^\tau e^{-rt} q_t dt\right] < \infty. \quad (9)$$

Associated with the rate process $q = \{q_t\}_{0 \le t \le \tau}$, there is a counting process $\{Q_t\}_{t \ge 0}$ with intensity rate $q_t \mathbb{1}_{Q_{t-}=0}$. Obviously, Q_t is binary valued, and once $Q_t = 1$, then $Q_{t'} = 1$ for any $t' > t$. In our setting, Q_t indicates whether or not the contract is terminated at time t. Therefore, we represent a contract as $\Gamma = (L, \tau, q)$, such that the termination rate $q = \{q_t\}_{t \ge 0}$ is also $\mathcal{F}$-predictable.

Given a contract Γ and the agent's effort process ν, define the agent's continuation utility for the agent at time t conditional on information up to time t as

$$W_t(\Gamma, \nu) = \mathbb{E}^\nu\left[\int_{t+}^\tau e^{-\rho(s-t)}\left(dL_s - c\mathbb{1}_{\nu_s=\mu}\right)\middle|\mathcal{F}_t\right]\mathbb{1}_{t<\tau}. \quad (10)$$

It is convenient to introduce the notation $W_{t-}(\Gamma, \nu) = \lim_{s \uparrow t} W_s(\Gamma, \nu)$ to denote the left-hand limit of the process $W(\Gamma, \nu)$ at $t \ge 0$. In the sequel, we omit W_t and W_{t-}'s dependence on Γ and ν when there is no confusion.

In this comment, we assume that for any contract Γ under our consideration, W_t is upper bounded by a large-enough $\bar{W}$. That is,

$$W_t(\Gamma, \nu) \le \bar{W} < \infty, \forall t \in [0, \infty) \cup \{0-\}, \Gamma, \nu. \quad \text{(WU)}$$

This is a technical assumption that allows us to establish that a process related to $W_t(\Gamma, \nu)$ is a martingale in the proof of Lemma B.2. The specific value of $\bar{W}$ is not important. As we show in this comment, as long as $\bar{W}$ is high enough, Constraint (WU) is not binding at optimality.

The following lemma extends lemma 6 of Sun and Tian (2018) into our setting.

Lemma 1. *For any contract Γ, there exist $\mathcal{F}_t$-predictable processes H_t and H_t^q such that*

$$\begin{aligned} dW_t = \big[&\left(\rho W_{t-} + c\mathbb{1}_{\nu_t=\mu} - \nu_t H_t + q_t H_t^q\right)dt \\ &+ H_t dN_t - H_t^q dQ_t - dL_t\big]\mathbb{1}_{W_{t-}>0}. \end{aligned} \quad \text{(PK)}$$

Furthermore, we have $W_t = 0$ when $Q_t = 1$, which implies that

$$\rho W_{t-} - H_t \nu_t + H_t^q q_t + c\mathbb{1}_{\nu_t=\mu} - \ell_t = 0, \quad (11)$$

$$\text{and } H_t^q \le W_{t-}, \quad (12)$$

for any time t such that $q_t > 0$.

Finally, contract Γ is EI if and only if

$$H_t \ge \beta. \quad \text{(IC)}$$

The (PK) condition is standard (see, for example, expression 13 in Biais et al. 2010). Conditions (11) and (12) stem from the fact that after a random termination ($dQ_t = 1$), the agent's promised utility needs to be set to zero ($dW_t = -W_t$).

Based on Lemma 1, we next present the optimal contract structure when $\rho > r$ in two cases.

3.1. Optimality of $\bar{\Gamma}$ Contract

In this section, we present conditions under which the simple contract $\bar{\Gamma}$ is optimal. Recall that under contract $\bar{\Gamma}$, the principal pays the agent cash β for each arrival starting from the very beginning while keeping the agent's promised utility at $\bar{w}$. That is, $dL_t = \beta dN_t$, $q_t = 0$, $H_t = \beta$, and $\tau = \infty$. The condition is

$$\frac{\bar{U} - \underline{v}}{\bar{w}} \ge \frac{\mu}{\rho - r - \mu} > 0, \quad (13)$$

in which $\bar{U}$ and $\bar{w}$ are defined in (3) and $\underline{v}$ in (2). Condition (13) is equivalent to $\mu < \rho - r$ and

$$R \ge \beta\left(1 + \frac{(\rho - r)(\rho - \mu)\underline{\mu}}{(\rho - r - \mu)\rho\Delta\mu}\right).$$

That is, the arrival rate μ is sufficiently small, and the revenue R is sufficiently large.

We have the following optimality result.

Proposition 2. *Under Condition* (13), *for any EI contract* Γ, *we have* $U(\Gamma) \le \bar{U}$.

Recall that $\bar{U}$ is the principal's utility under contract $\bar{\Gamma}$. Therefore, Proposition 2 implies that $\bar{\Gamma}$ is the optimal EI contract under Condition (13).

Note that the second part of proposition 4 in Sun and Tian (2018) is wrong under Condition (13). That is, under (13), one cannot find a $\hat{w} < \bar{w}$ together with a function V_d that satisfies (4)–(6). Under this condition, for any V_d and $\hat{w}$ that satisfy (4) and (5), we always have $V_d(0) > \underline{v}$ for all $\hat{w} \le \bar{w}$, violating (6). In fact, although the first part of proposition 4 in Sun and Tian (2018) is correct, the second part holds if and only if Condition (13) does not hold.

3.2. More General Optimal Contract

Now, we consider model parameters that do not satisfy Condition (13). In this case, the corresponding optimal function is closely related to DDE (4). First, we present the following result.

Lemma 2. *Suppose* $\rho > r$, *but* (13) *does not hold. Consider any* $\tilde{w} \in (0, \bar{w})$ *and a corresponding function* $V_{\tilde{w}}$, *such that the function* $V_d = V_{\tilde{w}}$ *together with the value* $\hat{w} = \tilde{w}$ *uniquely solves* (4) *with boundary Condition* (5). *There exists a threshold* $\breve{w}(\tilde{w}) \in [0, \tilde{w})$, *such that* $V''_{\tilde{w}}(w) < 0$ *for* $w \in (\breve{w}(\tilde{w}), \tilde{w}]$ *and* $V''_{\tilde{w}} > 0$ *for* $w \in [0, \breve{w}(\tilde{w}))$.

Remark 1. Following the proof of Lemma 2, if $\breve{w}(\hat{w}) > 0$, then we must have $r + \Delta\mu < \rho < \underline{\mu}$. This suggests that nonconcavity of V_d can only occur in a rather limited parameter regime.

Based on Lemma 2, we define value function

$$\mathcal{V}_{\tilde{w}}(w) = \begin{cases} V_{\tilde{w}}(\breve{w}(\tilde{w})) + V'_{\tilde{w}}(\breve{w}(\tilde{w})) \cdot (w - \breve{w}(\tilde{w})), \\ \qquad\qquad w \in [0, \breve{w}(\tilde{w})), \\ V_{\tilde{w}}(w \wedge \tilde{w}), \\ \qquad\qquad w \in [\breve{w}(\tilde{w}), \infty). \end{cases} \tag{14}$$

That is, function $\mathcal{V}_{\tilde{w}}(w)$ is the solution of DDE (4) only for $w \ge \breve{w}(\tilde{w})$. For $w < \breve{w}(\tilde{w})$, function $\mathcal{V}_{\tilde{w}}(w)$ is linear in w, the slope of which ensures that the left and right derivatives at $\breve{w}(\tilde{w})$ are the same (smooth pasting). This construction ensures that function $\mathcal{V}_{\tilde{w}}(w)$ is indeed concave.

The following lemma further indicates that we can specify a unique value $\hat{w}$, such that function $\mathcal{V}_{\hat{w}}$, as a special case of $\mathcal{V}_{\tilde{w}}$, satisfies an additional boundary condition analogous to (6).

Lemma 3. *There exists a unique* $\hat{w} \in [0, \bar{w})$ *such that* $\mathcal{V}_{\hat{w}}(0) = \underline{v}$. *Additionally, function* $\mathcal{V}_{\hat{w}}(w)$ *is strictly increasing and concave on* $[0, \tilde{w})$. *Furthermore, if* $\breve{w}(\hat{w}) > 0$, *then* $\mathcal{V}'_{\hat{w}}(0) = \mathcal{V}'_{\hat{w}}(\breve{w}(\hat{w})) > 1$.

Therefore, based on the solution to DDE (4) with the right boundary Condition (5), we construct a value function according to (14) and identify an upper threshold $\hat{w}$ such that the left boundary Condition (6) is also satisfied. In the sequel, we use $\breve{w}$ in place of $\breve{w}(\hat{w})$ for notational brevity. The lower threshold $\breve{w}$, if positive, represents the value of the promised utility at which the contract may be terminated after a random time. As we will show later in this section, using the same model parameters as the counterexample in Section 2, that $\breve{w}$ can indeed be positive.

Figure 1 plots the value function $\mathcal{V}_{\hat{w}}$ for the counterexample proposed in Section 2. Note that the function is linear on $[0, \breve{w}]$. Plotting function $V_{\hat{w}}$ on the same figure is uninformative because the difference between the two functions is so small that they are not visually distinguishable.

Based on the construction, we are ready to present a class of EI contracts that contains the optimal one.

Definition 1. For any $w \ge \breve{w}$, define contract $\Gamma^*(w) = (L^*, q^*, \tau^*)$ as follows:

1. Set $W_0 = w$ and $L_0^* = (W_0 - \hat{w})^+$.
2. For $t \ge 0$, set $dL_t^* = (W_{t-} + \beta - \hat{w})^+ dN_t$, $H_t = \beta$, and $q_t^* = 0$ if $\breve{w} = 0$ and $q_t^* = q^* \mathbb{1}_{W_{t-} = \breve{w}}$ if $\breve{w} > 0$, in which

$$q^* := \frac{\rho(\bar{w} - \breve{w})}{\breve{w}},$$

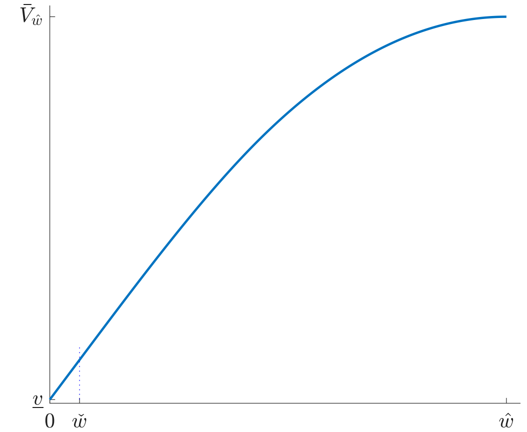
such that

$$dW_t = \rho(W_{t-} - \bar{w})dt \cdot \mathbb{1}_{W_{t-} > \breve{w}} - \breve{w} dQ_t + (\beta, \wedge \hat{w} - W_{t-})dN_t. \tag{15}$$

3. The termination time is $\tau^* = \min\{t : W_t = 0\}$.

When $\breve{w} = 0$, contract $\Gamma^*(w)$ is identical to the contract Γ_d^* proposed in Sun and Tian (2018). When $\breve{w} > 0$,

Figure 1. (Color online) The Value Function $\mathcal{V}_{\hat{w}}$ for the Counterexample in Section 2, with Γ^* with $\rho = 1$, $r = 0.05$, $R = 112.4622$, $\mu = 1.95$, and $c = 0.3$



Note. In this case, a binary search procedure identifies $\hat{w} = 5.438978$, indistinguishable from the 5.439 value proposed in Section 2.

however, the promised utility always stays at or above $\breve{w}$ before termination starting from $W_0 = w \ge \breve{w}$ according to $\Gamma^*(w)$. When $W_{t-} = \breve{w}$, the agent is terminated after an exponentially distributed random time with rate q^*, if no additional arrival occurs during this time period. The termination rate q^* is determined from (11) with $\ell_t = 0$ and $H_t^q = W_{t-}$ (binding (12)).

Note that we only focus on defining the contract for the starting promised utility w at or above the threshold $\breve{w}$. This is because following the dynamic (15), the promised utility never falls below this threshold before the contract is terminated. In fact, if the initial promised utility, W_{0-}, is less than $\breve{w}$, the optimal contract would randomly set the continued promised utility, W_0, to either $\breve{w}$ or zero such that the expectation is kept at W_{0-}. However, rigorously representing this rather simple idea involves more complex notations, which we deem unnecessary.

Define the principal's value function $F_{\hat{w}}(w) := \mathcal{V}_{\hat{w}}(w) - w$. The next proposition states that the principal's utility following contract $\Gamma^*(w)$ is indeed $F_{\hat{w}}(w)$.

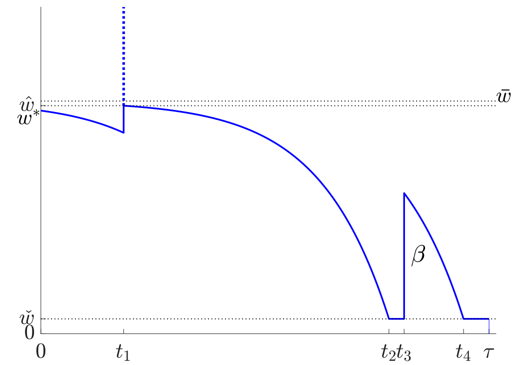
Proposition 3. *For any $w \ge \breve{w}$, we have $U(\Gamma^*(w)) = F_{\hat{w}}(w)$. Furthermore, if $\breve{w} > 0$, then we have $F_{\hat{w}}(\breve{w}) > \underline{v}$.*

Following Lemma 3, define w^* to be the unique maximizer of $F_{\hat{w}}(w)$. Proposition 3 implies that $w^* = 0$ only if $\breve{w} = 0$, and if $\breve{w} > 0$, then $w^* > \breve{w}$. The following result further implies that contract $\Gamma^*(w^*)$ is the optimal EI contract.

Proposition 4. *If $\rho > r$ but (13) does not hold, then for any EI contract Γ, we have $F_{\hat{w}}(w^*) \ge U(\Gamma)$.*

Figure 2 depicts a sample trajectory following the optimal contract Γ^* for the counterexample presented in Section 2. As we can see, the promised utility starts from w^* and gradually decreases until the first arrival, at time t_1. At this point in time, an upward jump of $H_t = \beta$ would take the promised utility above $\hat{w}$. Therefore, the promised utility jumps to the upper bound $\hat{w}$, and the principal pays the agent $W_{t_1-} + \beta - \hat{w}$. No further arrival occurs until time t_2, when the promised utility reaches the lower threshold $\breve{w}$. At this point, the agent would be terminated randomly according to a constant rate q^*, whereas the promised utility stays at $\breve{w}$. At time t_3, there is an arrival before a random termination occurs, which pushes the promised utility up by β. The promised utility decreases to $\breve{w}$ again at t_4. The agent is eventually terminated at time τ, following the same random termination rate q^* after t_4. Note that here the value $\hat{w} = 5.438978$, which is slightly lower than the (exact) 5.439 value in the counterexample of Section 2. Following the monotonicity property identified in proposition 4 of Sun and Tian (2018), this difference implies that the value function $\mathcal{V}_{\hat{w}}$ is indeed higher than function V_d identified in Section 2.

Figure 2. (Color online) A Sample Trajectory for the Agent's Promised Utility According to Γ^* for the Counterexample in Section 2, with $\rho = 1$, $r = 0.05$, $R = 112.4622$, $\mu = 1.95$, $\underline{\mu} = 1.85$, and $c = 0.3$



Notes. In this case, $\beta = 3$, $\bar{w} = 5.55$, $\hat{w} = 5.438978$, $\breve{w} = 0.354$, and $w^* = 5.322$. The solid curve depicts a sample trajectory of $\{W_t\}_{0 \le t \le \tau}$, and the dotted line depicts the payment.

So far, we have fully specified the optimal contract structure when $\rho > r$. The contract structure Γ^* is more general than the contract Γ_d^* specified in Sun and Tian (2018), involving a lower threshold $\breve{w}$ and random termination when the promised utility reaches this point. Arguably, such a structure when $\breve{w} > 0$ is also more complex than Γ_d^*. Therefore, we naturally wonder how often $\breve{w}$ is indeed positive and when it is, how much Γ_d^* is suboptimal.

To answer these questions, we conduct a computational study. We pick the following parameters: $\rho = 1, r \in \{0.05, 0.1, \ldots, 0.5\}, R \in \{5, 20, \ldots, 995\}, c \in \{0.3, 0.9, \ldots, 2.7\}$, $\mu \in \{1.1, 1.5, \ldots, 2.7\}$, and $\underline{\mu} \in \{\rho + 0.05, \rho + 0.35, .., \mu - 0.05\}$. Among all these parameter combinations, we have 33465 cases with $R > \beta$ and $r + \Delta\mu < \rho < \underline{\mu}$. (Recall from Remark 1 that only if $r + \Delta\mu < \rho < \underline{\mu}$, we may have $\breve{w} > 0$.) Among these 33,465 cases, only 1774 (5.3%) cases yield $\breve{w} > 0$. Among these 1774 cases, we investigate the relative improvement of Γ^* over Γ_d^*, calculated as $[U(\Gamma^*) - U(\Gamma_d^*)]/U(\Gamma_d^*)$. We find that the maximum relative improvement is 1.6×10^{-7}, and the mean improvement is only 1.0×10^{-8}. This suggests that although Γ_d^* may be suboptimal in some cases, the suboptimality gap is extremely small. Note that Γ_d^* is easier to implement than the optimal Γ^* in practice. Hence, the principal may still prefer to use Γ_d^* rather than Γ^*.

Appendix A. Proofs in Section 2

A.1. Proof of Proposition 1

Note that $V'_{\tilde{w}}(\tilde{w}-) = 0$. Hence, $V'_{\tilde{w}} \in C^1(\mathbb{R}_+)$. Therefore, differentiating (4) over $[0, \tilde{w})$ yields

$$\rho(\bar{w} - w)V''_{\tilde{w}}(w) = \mu\big(V'_{\tilde{w}}(w + \beta) - V'_{\tilde{w}}(w)\big) + (\rho - r)\big(V'_{\tilde{w}}(w) - 1\big), \quad \text{for } w \in [0, \tilde{w}), \tag{A.1}$$

with boundary condition $V'_{\tilde{w}}(w) = 0$ for $w \geq \tilde{w}$, which is a DDE for $V'_{\tilde{w}}$. DDE (A.1) does not involve parameter R. Therefore, its unique solution $V'_{\tilde{w}}$ is independent of R. Furthermore, we have

$$\psi(R, \tilde{w}) = V_{\tilde{w}}(0; R) - \underline{v} = V_{\tilde{w}}(\tilde{w}; R) - \int_0^{\tilde{w}} V'_{\tilde{w}}(y)dy - \underline{v}$$
$$= \frac{\Delta\mu(R-\beta) - (\rho - r)\tilde{w}}{r} - \int_0^{\tilde{w}} V'_{\tilde{w}}(y)dy, \qquad \text{(A.2)}$$

which is linear and strictly increasing in R. Next, we present a technical lemma to help to complete the proof. The proof is provided at the end of this section. □

Lemma A.1. *$V_{\tilde{w}}(w)$ is strictly increasing in w on $[0, \tilde{w})$ for any $\tilde{w} \in (0, \bar{w})$.*

Lemma A.1 immediately implies that $\psi(\beta, \hat{w}) < 0$ in view of (A.2).

Proof of Lemma A.1. If Lemma A.1 fails to hold, then $w^p = \sup\{w \in [0, \tilde{w}) : V'_{\tilde{w}}(w) \geq 0\}$ is well defined. Because $V_{\tilde{w}} \in C^1([0, \tilde{w}))$ and $V''_{\tilde{w}}(\tilde{w}-) < 0$, we have $w^p < \hat{w}$, $V'_{\tilde{w}}(w^p) = 0$, and $V'_{\tilde{w}} > 0$ over $(w^p, \tilde{w})$. It follows from (4) at w^p that

$$rV_{\tilde{w}}(w^p) = \mu R - c - (\rho - r)w^p + \mu\big(V_{\tilde{w}}(w^p + \beta) - V_{\tilde{w}}(w^p)\big)$$
$$> \mu R - c - (\rho - r)\tilde{w} = rV_{\tilde{w}}(\tilde{w}),$$

where the inequality follows from $w^p < \tilde{w}$ and $V_{\tilde{w}}(w^p + \beta) > V_{\tilde{w}}(w^p)$ by noting that $V'_{\tilde{w}}(w) > 0$ over $(w^p, \tilde{w})$. This contradicts with $V_{\tilde{w}}(w^p) < V_{\tilde{w}}(\tilde{w})$. □

Appendix B. Proofs in Section 3

B.1. Proof of Lemma 1

Define agent's total expected discounted utility conditional on $\mathcal{F}_t$ as

$$u_t(\Gamma, \nu) := \mathbb{E}^{\nu}\left[\int_0^{\tau} e^{-\rho s}\big(dL_s - c\mathbb{1}_{\nu_s=\mu}ds\big) \,|\, \mathcal{F}_t\right]$$
$$= \int_0^{(t\wedge\tau)} e^{-\rho s}\big(dL_s - c\mathbb{1}_{\nu_s=\mu}ds\big) + e^{-\rho t}W_t(\Gamma, \nu). \qquad \text{(B.1)}$$

To ease notation, we omit (Γ, ν) from all relevant quantities.

Given an effect process ν, we use $\mathcal{I}^N_{[t_1,t_2]}$ to denote the set of arrival time epochs during $[t_1, t_2]$. Moreover, we denote $\mathcal{I}^N_t := \mathcal{I}^N_{[0,t]}$ and $\mathcal{I}^N := \mathcal{I}^N_{[0,\infty)}$. We use $\mathcal{I}^Q_{[t_1,t_2]}$ to denote the set of randomized termination time epochs during $[t_1, t_2]$ under the randomized termination policy $\{q_t\}_{t\geq 0}$. Moreover, we denote $\mathcal{I}^Q_t := \mathcal{I}^Q_{[0,t]}$ and $\mathcal{I}^Q := \mathcal{I}^Q_{[0,\infty)}$.

At any time $\zeta-$, $W_{\zeta-}$ can jump to W^N_ζ triggered by an arrival at time ζ, jump to W^Q_ζ triggered by a randomized termination, or jump to W^L_ζ triggered by an instantaneous payment. Thus, we can decompose W_ζ (for $\zeta > t$) into its discrete part

$$\sum_{t\leq\xi\leq\zeta}\Big[(W^N_\xi - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}^N_{[t,\zeta]}} + \big(W^Q_\xi - W_{\xi-}\big)\mathbb{1}_{\xi\in\mathcal{I}^Q_{[t,\zeta]}}$$
$$+ (W^L_\xi - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}^L_{[t,\zeta]}}\Big]$$

and its absolutely continuous part

$$W^c_\zeta := W_\zeta - \sum_{t\leq\xi\leq\zeta}\Big[(W^N_\xi - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}^N_{[t,\zeta]}}\cdot$$
$$+ \big(W^Q_\xi - W_{\xi-}\big)\mathbb{1}_{\xi\in\mathcal{I}^Q_{[t,\zeta]}} + (W^L_\xi - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}^L_{[t,\zeta]}}\Big],$$

where we use $\mathcal{I}^L_{[t,\zeta]}$ to denote the set of time epochs in $[t, \zeta]$ such that a positive instantaneous payment occurs. Hence, $\xi \in \mathcal{I}^L_{[t,\zeta]}$ if $\Delta L_\xi > 0$ and $\xi \in [t, \zeta]$.

According to the definition of admissible contract, we know that both W^N_t and W^Q_t are $\mathcal{F}_t$-predictable. However, W^L_t also depends on dN_t and dQ_t. Hence, W^L_t is $\mathcal{F}_t$-adaptive.

Fix any $t' > t$. By calculus of point process, we have

$$e^{-\rho t'}W_{t'} - e^{-\rho t}W_t = \int_t^{t'} e^{-\rho\zeta}\big(-\rho W_\zeta d\zeta + dW^c_\zeta\big)$$
$$+ \sum_{\zeta\in(t,t']} e^{-\rho\zeta}\Big[(W^N_\zeta - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}^N_{(t,t']}}\cdot$$
$$+ \big(W^Q_\zeta - W_{\zeta-}\big)\mathbb{1}_{\zeta\in\mathcal{I}^Q_{(t,t']}} + (W^L_\zeta - W_{\zeta-})$$
$$\times \mathbb{1}_{\zeta\in\mathcal{I}^L_{(t,t']}}\Big]. \qquad \text{(B.2)}$$

It is clear that the process $\{u_t\}_{t\geq 0}$ is an $\mathcal{F}$-martingale. Hence, for any time $t' > t$, we have $u_t = \mathbb{E}_t[u_{t'}]$, where $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_t]$. Consequently, we have

$$0 = \mathbb{E}_t[u_{t'}] - u_t$$
$$= \mathbb{E}_t\big[e^{-\rho t'}W_{t'} - e^{-\rho t}W_t\big]$$
$$+ \mathbb{E}_t\left[\int_{(t\wedge\tau)+}^{(t'\wedge\tau)} e^{-\rho\zeta}\big(dL_\zeta - c\mathbb{1}_{\nu_\zeta=\mu}d\zeta\big)\right]$$
$$= \mathbb{E}_t\left[\int_t^{t'} e^{-\rho\zeta}\big(-\rho W_\zeta d\zeta + dW^c_\zeta\big)\right]$$
$$+ \mathbb{E}_t\Bigg\{\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\Big[(W^N_\zeta - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}^N_{(t,t']}}$$
$$+ \big(W^Q_\zeta - W_{\zeta-}\big)\mathbb{1}_{\zeta\in\mathcal{I}^Q_{(t,t']}} + (W^L_\zeta - W_{\zeta-})\ \mathbb{1}_{\zeta\in\mathcal{I}^L_{(t,t']}}\Big]\Bigg\}$$
$$+ \mathbb{E}_t\left[\int_{(t\wedge\tau)+}^{(t'\wedge\tau)} e^{-\rho\zeta}\big(dL_s - c\mathbb{1}_{\nu_\zeta=\mu}d\zeta\big)\right]$$
$$= \mathbb{E}_t\Bigg\{\int_t^{t'} e^{-\rho\zeta}\{[-\rho W_\zeta + (W^N_\zeta - W_{\zeta-})\nu_\zeta$$
$$+ \big(W^Q_\zeta - W_{\zeta-}\big)q_\zeta\mathbb{1}_{Q_{\zeta-}=0}\Big]d\zeta + dW^c_\zeta\Bigg\}$$
$$+ \sum_{\zeta\in(t,t']} e^{-\rho\zeta}\Big[(W^L_\zeta - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}^L_{(t,t']}}\Big]\Bigg\}$$
$$+ \mathbb{E}_t\left[\int_{(t\wedge\tau)+}^{(t'\wedge\tau)} e^{-\rho\zeta}\big(dL_\zeta - c\mathbb{1}_{\nu_\zeta=\mu}d\zeta\big)\right],$$

where the second equality follows from (B.1) and the third one from (B.2). The fourth equality follows from the fact that $\{Q_t\}_{t\geq 0}$ is a counting process with intensity $q_t\mathbb{1}_{Q_{t-}=0}$ and that

N_t is a counting process with intensity ν_t, as well as lemma L3 in chapter II of Brémaud (1981), while noting that

$$\begin{aligned}&\mathbb{E}_t\int_t^{t'} e^{-\rho\zeta}\left|\left(W_\zeta^N - W_{\zeta-}\right)\nu_\zeta\right|d\zeta\\ &\quad\le R\mu/r\cdot\mu\int_t^{t'} e^{-\rho\zeta}d\zeta<\infty \text{ and}\\ &\mathbb{E}_t\int_t^{t'} e^{-\rho\zeta}\left|\left(W_\zeta^Q - W_{\zeta-}\right)q_\zeta\mathbb{1}_{Q_{\zeta-}=0}\right|d\zeta\\ &\quad\le R\mu/r\cdot\mathbb{E}_t\int_t^{\tau} e^{-\rho\zeta}q_\zeta d\zeta\cdot\mathbb{1}_{t<\tau}\\ &\quad\le R\mu/r\cdot\mathbb{E}_t\int_t^{\tau} e^{-r\zeta}q_\zeta d\zeta\cdot\mathbb{1}_{t<\tau}<\infty,\end{aligned}$$

in view of (WU), $\rho > r$, and (9).

Recall that $dL_t = \ell_t dt + \Delta L_t$. For any $t < t' < \tau$, the equality can be stated as

$$\begin{aligned}&\mathbb{E}_t\left\{\int_t^{t'} e^{-\rho\zeta}\Big[-\rho W_\zeta + \left(W_\zeta^N - W_{\zeta-}\right)\nu_\zeta\right.\\ &\quad\left.+\left(W_\zeta^Q - W_{\zeta-}\right)q_\zeta - c\mathbb{1}_{\nu_\zeta=\mu}\Big]d\zeta + dW_\zeta^c + \ell_\zeta d\zeta\right\}\\ &\quad+\mathbb{E}_t\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\left[\left(W_\zeta^L - W_{\zeta-}\right)\mathbb{1}_{\Delta L_\zeta>0}+\Delta L_\zeta\right]\\ &\quad=0. \qquad\qquad (B.3)\end{aligned}$$

Consider any time $t < \tau$. Letting $t' \downarrow t$ in (B.3) yields

$$\mathbb{E}_t\left[\left(W_t^L - W_{t-}\right)\mathbb{1}_{\Delta L_t>0}+\Delta L_t\right]=0, \qquad (B.4)$$

which further implies

$$\begin{aligned}dW_t^c &= \Big[\rho W_{t-} - \left(W_t^N - W_{t-}\right)\nu_t - \left(W_t^Q - W_{t-}\right)\\ &\quad\times q_t + c\mathbb{1}_{\nu_t=\mu}\Big]dt - \ell_t dt,\ t\in[0,\tau]. \qquad (B.5)\end{aligned}$$

Let $H_t := W_t^N - W_{t-}$ and $H_t^q := -W_t^Q + W_{t-}$. Then, both H_t and H_t^q are $\mathcal{F}_t$-predictable. Also, because W_t^L is $\mathcal{F}_t$-adaptive, (B.4) in fact reduces to

$$(W_t^L - W_{t-})\mathbb{1}_{\Delta L_t>0}+\Delta L_t = 0. \qquad (B.6)$$

Recall that

$$\begin{aligned}dW_t &= dW_t^c + (W_t^N - W_{t-})dN_t + \left(W_t^Q - W_{t-}\right)\\ &\quad\times dQ_t + (W_t^L - W_{t-})\mathbb{1}_{\Delta L_t>0}. \qquad (B.7)\end{aligned}$$

It follows from (B.5)–(B.7) that

$$\begin{aligned}dW_t &= \left(\rho W_{t-} + c\mathbb{1}_{\nu_t=\mu} - \nu_t H_t + q_t H_t^q\right)dt + H_t dN_t\\ &\quad - H_t^q dQ_t - dL_t,\end{aligned}$$

for $t\in[0,\tau]$, which gives (PK).

Under randomized termination at time t (i.e., $dQ_t = 1$), we have $dW_t = -W_{t-}$, which yields (i) $dW_t^c = 0$ and (ii) $\Delta W_t = -W_{t-}$.

Condition (i), combining with (B.5), gives

$$\rho W_{t-} - H_t\nu_t + H_t^q q_t + c\mathbb{1}_{\nu_t=\mu} - \ell_t = 0. \qquad (B.8)$$

Note that at any time t such that $q_t > 0$ and $Q_{t-} = 0$, there is a positive probability that $dQ_t = 1$. Moreover, all terms on the left-hand side of (B.8) are $\mathcal{F}_t$-predictable. Hence, (B.8) holds at any time t such that $q_t > 0$. This gives (11).

Condition (ii), combining with (B.6) and (B.7), yields

$$H_t dN_t - H_t^q dQ_t - \Delta L_t = -W_{t-}.$$

Note that the probability that both $dN_t = 1$ and $dQ_t = 1$ happen is zero. Hence, given that $dQ_t = 1$, we must have $\Delta L_t = W_{t-} - H_t^q \ge 0$. Applying a similar argument as that for (11), we know that $H_t^q \le W_{t-}$ for any time t such that $q_t > 0$. This gives (12). Here, we mention that it does not necessarily hold that $H_t^q = W_{t-}$. In fact, for any time t such that $q_t > 0$, if $dQ_t = 1$, we can set $\Delta L_t = W_{t-} - H_t^q$ to make W_t be zero.

The proof of (IC) can be easily adapted from that for lemma 6 in Sun and Tian (2018), which is omitted here for brevity. □

B.2. An Optimality Condition

In this section, we present an optimality condition, which will help us prove the optimality of contracts in the later sections.

Lemma B.1. *Suppose $F(w)$ is a differentiable, concave, and upper-bounded function, with $F(0) = \underline{v}$ and $F'(w) \ge -1$. Consider any EI contract Γ, which yields the agent's expected utility $w = W_{0-}$, followed by the promised utility process $\{W_t\}_{t\ge0}$ according to (PK). Define a stochastic process $\{\Psi_t\}_{t\ge0}$, where*

$$\begin{aligned}\Psi_t &:= F'(W_{t-})(\rho W_{t-} + c) + q_t\big[F'(W_{t-})H_t^q\\ &\quad + F\big(W_{t-} - H_t^q\big) - F(W_{t-})\big] - rF(W_{t-})\\ &\quad + \mu[R - F'(W_{t-})H_t + F(W_{t-} + H_t)\\ &\quad - F(W_{t-})]. \qquad (B.9)\end{aligned}$$

If the process $\{\Psi_t\}_{t\ge0}$ is nonpositive almost surely, then we have $F(w) \ge U(\Gamma)$.

Proof. Following Ito's formula for jump processes (see, for example, theorem 17.5 of Bass 2011) and considering (PK), we have

$$\begin{aligned}&e^{-r(T\wedge\tau)}F(W_{T\wedge\tau})\\ &\quad = F(W_{0-}) + \int_0^{T\wedge\tau}\big[e^{-rt}dF(W_{t-}) - re^{-rt}\\ &\qquad\times F(W_{t-})dt\big]\\ &\quad = F(W_{0-}) + \int_0^{T\wedge\tau} e^{-rt}(-RdN_t + dL_t)\\ &\qquad + \int_0^{T\wedge\tau} e^{-rt}\mathcal{A}_t, \qquad (B.10)\end{aligned}$$

where

$$\begin{aligned}\mathcal{A}_t &:= dF(W_{t-}) - rF(W_{t-})dt + RdN_t - dL_t\\ &= F'(W_{t-})\big[\rho W_{t-} + c - \mu H_t + q_t H_t^q - \ell_t\big]dt\\ &\quad + F\big(W_{t-} + H_t dN_t - H_t^q dQ_t - \Delta L_t\big)\\ &\quad - F\ (W_{t-}) - rF(W_{t-})dt\\ &\quad + RdN_t - dL_t.\end{aligned}$$

Further define

$$\begin{aligned}\mathcal{B}_t &:= \left[F\left(W_{t-}-H_t^q\right)-F(W_{t-})\right]\left(dQ_t-q_tdt\right)\\ &\quad+[F(W_{t-}+H_t)-F(W_{t-})]\left(dN_t-\mu dt\right)\\ &\quad+R\left(dN_t-\mu dt\right).\end{aligned}$$

Because function $F(w)$ is concave and $F'(w)\geq -1$, we have

$$\begin{aligned}\mathcal{A}_t \leq& F'(W_{t-})\left(\rho W_{t-}+c-\mu H_t+q_tH_t^q\right)dt\\ &+F\left(W_{t-}+H_tdN_t-H_t^qdQ_t\right)-F'(W_{t-})\ell_tdt\\ &-F'\left(W_{t-}+H_tdN_t-H_t^qdQ_t\right)\Delta L_t\\ &-F(W_{t-})-rF(W_{t-})dt+RdN_t-dL_t\\ \leq& F'(W_{t-})\left(\rho W_{t-}+c-\mu H_t+q_tH_t^q\right)dt\\ &-rF(W_{t-})dt+RdN_t\\ &+F\left(W_{t-}+H_tdN_t-H_t^qdQ_t\right)-F(W_{t-})\\ =& F'(W_{t-})\left(\rho W_{t-}+c-\mu H_t+q_tH_t^q\right)dt\\ &-rF(W_{t-})dt+[F(W_{t-}+H_t)-F(W_{t-})]dN_t\\ &+\left[F\left(W_{t-}-H_t^q\right)-F(W_{t-})\right]dQ_t+\mu Rdt\\ =&\mathcal{B}_t+\Psi_tdt. \end{aligned}\tag{B.11}$$

Next, we present a technical lemma, the proof of which is given later in this section.

Lemma B.2. *If $U(\Gamma)>-\infty$, then $\mathbb{E}[\int_{0+}^{T\wedge\tau}e^{-rt}\mathcal{B}_t]=0$.*

To show Lemma B.1, it suffices to consider the case that $U(\Gamma)>-\infty$, so that Lemma B.2 holds. Taking the expectation on both sides of (B.10) and letting $T\longrightarrow\infty$, we have

$$\begin{aligned}F(W_{0-}) &\geq \mathbb{E}\left[e^{-r\tau}F(W_\tau)+\int_0^\tau e^{-rt}(RdN_t-dL_t)\right.\\ &\qquad\left.-\int_0^\tau e^{-rt}\mathcal{B}_t-\int_0^\tau e^{-rt}\Psi_tdt\right]\\ &\geq \mathbb{E}\left[e^{-r\tau}F(W_\tau)+\int_0^\tau e^{-rt}(RdN_t\right.\\ &\qquad\left.-dL_t)\right]=U(\Gamma),\end{aligned}\tag{B.12}$$

where the first inequality follows from (B.11), the second inequality follows from $\Psi_t\leq 0$ and Lemma B.2, and the last equality follows from $F(W_\tau)=F(0)=\underline{v}$. □

In the end of this section, we will present the proof of Lemma B.2 to complete the proof of Lemma B.1. Additionally, we will also present a simplification of Ψ_t, which will be used in the subsequent analysis.

Proof of Lemma B.2. First, we show that

$$\mathbb{E}\left[\int_{0+}^{T\wedge\tau}e^{-rt}[F(W_{t-}+H_t)-F(W_{t-})]\left(dN_t-\mu dt\right)\right]=0.\tag{B.13}$$

If $\mathbb{E}[\int_{0+}^{\tau}e^{-rt}|H_t|dt]<\infty$, then

$$\begin{aligned}&\mathbb{E}\left[\int_{0+}^{T\wedge\tau}e^{-rt}|F(W_{t-}+H_t)-F(W_{t-})|\mu dt\right]\\ &\leq \max_{w\geq 0}\{|F'(w)|\}\cdot\mu\mathbb{E}\left[\int_{0+}^{\tau}e^{-rt}|H_t|dt\right]<\infty,\end{aligned}$$

where $\max_{w\geq 0}\{|F'(w)|\}<\infty$ follows from the concavity of F and that $F'\geq -1$. Define $\mathcal{F}^\tau=\{\mathcal{F}_{t\wedge\tau}\}_{t\geq 0}$. It follows from lemma L3 in chapter II in Brémaud (1981) that $\tilde{M}=\{\tilde{M}_t\}_{t\geq 0}$, defined by

$$\tilde{M}_t=\int_{0+}^{t\wedge\tau}e^{-rs}[F(W_{s-}+H_s)-F(W_{s-})](dN_t-\mu dt),$$

is an $\mathcal{F}^\tau$-martingale. Hence, $\mathbb{E}\tilde{M}_T=\mathbb{E}\tilde{M}_0=0$: that is, $\mathbb{E}[\int_0^{T\wedge\tau}e^{-rt}[F(W_{t-}+H_t)-F(W_{t-})](dN_t-\mu dt)]=0$.

Now, suppose that $\mathbb{E}[\int_{0+}^{\tau}e^{-rt}|H_t|dt]=\infty$. It follows from (PK) and (WU) that $dL_t\geq(H_t-\bar{W})^+dN_t$, for $t\in(0,\tau)$. Hence, we have

$$\begin{aligned}\mathbb{E}\left[\int_0^\tau e^{-rt}dL_t\right] &\geq \mathbb{E}\left[\int_0^\tau e^{-rt}\left(H_t-\bar{W}\right)^+dN_t\right]\\ &=\mathbb{E}\left[\int_0^\tau e^{-rt}\left(H_t-\bar{W}\right)^+\mu dt\right]\\ &\geq \mathbb{E}\left[\int_0^\tau e^{-rt}\left(|H_t|-\bar{W}\right)\mu dt\right]\\ &\geq \mathbb{E}\left[\int_0^\tau e^{-rt}|H_t|dt\right]-\frac{\mu\bar{W}}{r}=\infty,\end{aligned}$$

where, the equality follows from equation (2.3) in chapter II in Brémaud (1981) and the second inequality follows from $H_t\geq -W_t\geq -\bar{W}$ in view of Lemma 1 and (WU). This, combining with (2), yields

$$\begin{aligned}U(\Gamma) &\leq \mathbb{E}\left[\int_0^\tau e^{-rt}R\mu dt+e^{-r\tau}\underline{v}\right]\\ &\quad-\mathbb{E}\left[\int_0^\tau e^{-rt}dL_t\right]\leq\frac{R\mu}{r}-\mathbb{E}\left[\int_0^\tau e^{-rt}dL_t\right]\\ &=-\infty,\end{aligned}$$

which reaches a contradiction. Therefore, we have $\mathbb{E}[\int_0^{T\wedge\tau}e^{-rt}[F(W_{t-}+H_t)-F(W_{t-})](dN_t-\mu dt)dt]=0$.

Next, we show that

$$\begin{aligned}&\mathbb{E}\left[\int_{0+}^{T\wedge\tau}e^{-rt}\left[F\left(W_{t-}-H_t^q\right)-F(W_{t-})\right]\right.\\ &\qquad\left.\times\left(dQ_t-q_tdt\right)\right]=0.\end{aligned}\tag{B.14}$$

If $\mathbb{E}[\int_{0+}^{\tau}e^{-rt}|H_t^q|q_tdt]<\infty$, then we have

$$\begin{aligned}&\mathbb{E}\left[\int_{0+}^{T\wedge\tau}e^{-rt}|F\left(W_{t-}-H_t^q\right)-F(W_{t-})|q_tdt\right]\\ &\leq\max_{w\geq 0}\{|F'(w)|\}\cdot\mathbb{E}\left[\int_{0+}^{\tau}e^{-rt}|H_t^q|q_tdt\right]<\infty,\end{aligned}$$

which yields (B.14) by using a similar argument for (B.13) and applying lemma L3 in chapter II of Brémaud (1981).

Now we show that $\mathbb{E}[\int_{0+}^{\tau}e^{-rt}|H_t^q|q_tdt]<\infty$. If $H_t^q\geq 0$, then it follows from (12) that

$$|H_t^q|q_t\leq W_{t-}q_t\leq\bar{W}\cdot q_t.$$

If $H_t^q<0$, then it follows from (11) that

$$\begin{aligned}|H_t^q|q_t&=-H_t^qq_t=\rho W_{t-}-H_t\nu_t-c\mathbb{1}_{\nu_t=\mu}-\ell_t\\ &\leq\rho\bar{W}-\beta\underline{\mu}.\end{aligned}$$

Therefore, we have

$$\mathbb{E}\left[\int_{0+}^{\tau} e^{-rt}|H_t^q|q_t dt\right] \le \mathbb{E}\left[\int_{0+}^{\tau} e^{-rt}\left(\bar{W}\cdot q_t + \rho\bar{W} - \beta\underline{\mu}\right)dt\right] < \infty,$$

where the last inequality follows from (9).

Finally, we have $\mathbb{E}[\int_{0+}^{T\wedge\tau} e^{-rt}R(dN_t - \mu dt)dt] = 0$, which follows immediately from lemma L3 in chapter II of Brémaud (1981). Hence, $\mathbb{E}[\int_{0+}^{T\wedge\tau} e^{-rt}\mathcal{B}_t] = 0$. □

B.2.1. A Simplification of Ψ_t.

Following (B.9), if we define $V(w) := F(w) + w$, then

$$\begin{aligned}\Psi_t &= V'(W_{t-})(\rho W_{t-} + c) - \rho W_{t-} - c + q_t\\ &\quad\times\left[V'(W_{t-})H_t^q + V(W_{t-} - H_t^q) - V(W_{t-})\right]\\ &\quad - rV(W_{t-}) + rW_{t-}\\ &\quad + \mu[R - V'(W_{t-})H_t + V(W_{t-} + H_t)\\ &\quad - V(W_{t-})]\\ &\le V'(W_{t-})(\rho W_{t-} + c) - \rho W_{t-} - c - rV(W_{t-})\\ &\quad + rW_{t-} + \mu[R - V'(W_{t-})H_t\\ &\quad + V(W_{t-} + H_t) - V(W_{t-})]\\ &\le V'(W_{t-})(\rho W_{t-} + c) - \rho W_{t-} - c - rV(W_{t-})\\ &\quad + rW_{t-} + \mu[R - V'(W_{t-})\beta + V(W_{t-} + \beta)\\ &\quad - V(W_{t-})],\end{aligned} \tag{B.15}$$

where the first inequality follows from $\max_{H_t^q}\{V'(W_{t-})H_t^q + V(W_{t-} - H_t^q) - V(W_{t-})\} = 0$ and the second inequality follows from $\beta = \arg\max_{H_t\ge\beta}\{-V'(W_{t-})H_t + V(W_{t-} + H_t)\}$.

B.3. Proof of Proposition 2

Define

$$F_{\bar{w}}(w) := \begin{cases}\underline{v} + aw, & w\in[0,\bar{w}),\\ \bar{U} - (w - \bar{w}), & w\in[\bar{w},\infty),\end{cases}$$

where $a := (\bar{U} - \underline{v})/\bar{w}$. It is clear that $F_{\bar{w}}(\bar{w}) = \bar{U}$. Following Condition (13), we have $F_{\bar{w}}(\bar{w}) \ge F_{\bar{w}}(w)$ for any w. Hence, to complete the proof, we will verify that $F_{\bar{w}}(w) \ge U(\Gamma)$ where $u(\Gamma, \nu^*) = w$. Because $F_{\bar{w}}(w)$ is not differentiable at $w = \bar{w}$, we cannot directly apply Lemma B.1. To address this issue, we construct a sequence of functions of $C^1(\mathbb{R}_+)$, which converges to $F_{\bar{w}}$ to help us complete the proof. For any $\varepsilon\in(0,\bar{w})$, define

$$F_\varepsilon(w) := \begin{cases}\underline{v} + aw + \frac{a\varepsilon}{2}, & w\in[0,\bar{w}-\varepsilon),\\ -\frac{a}{2\varepsilon}(w-\bar{w})^2 + \bar{U}, & w\in\left[\bar{w}-\varepsilon,\bar{w}+\frac{\varepsilon}{a}\right),\\ \bar{U} + \bar{w} + \frac{\varepsilon}{2a} - w, & w\in\left[\bar{w}+\frac{\varepsilon}{a},\infty\right).\end{cases}$$

It is straightforward to check that $F_\varepsilon\in C^1(\mathbb{R}_+)$, and it is concave. Moreover, we can easily see that $F_\varepsilon \ge F_{\bar{w}}$ and $\lim_{\varepsilon\to 0}F_\varepsilon = F_{\bar{w}}$.

Choose ε sufficiently small such that $\beta > \varepsilon + \varepsilon/a$. Next, we show that for any EI contract Γ, $F_\varepsilon(w) \ge U(\Gamma(w)) - K\varepsilon$ for some $K > 0$. Define

$$\begin{aligned}G(w) &:= V'(w)(\rho w + c) - \rho w - c - rV(w) + rw\\ &\quad + \mu\left[R - V'(w)\beta + V(w+\beta) - V(w)\right],\end{aligned}$$

for $w\in\mathbb{R}_+$, where $V(w) := F_\varepsilon(w) + w$.

For $w\in[\bar{w}-\varepsilon,\bar{w}+\varepsilon/a]$, we have

$$\begin{aligned}G(w) &= \left[-\frac{a}{\varepsilon}(w-\bar{w}) + 1\right](\rho w + c) - \rho w - c - r\\ &\quad\times\left[-\frac{a}{2\varepsilon}(w-\bar{w})^2 + \bar{U}\right]\\ &\quad + \mu\Bigg[R - \left(-\frac{a}{\varepsilon}(w-\bar{w}) + 1\right)\beta + \bar{w} + \frac{\varepsilon}{2}\\ &\quad\times\left(\frac{1}{a} - a\right) + \frac{a}{2\varepsilon}(w-\bar{w})^2 + \frac{a\varepsilon}{2} - w\Bigg]\\ &= -\frac{a}{\varepsilon}(w-\bar{w})(\rho w + c) - r\left[-\frac{a}{2\varepsilon}(w-\bar{w})^2 + \bar{U}\right]\\ &\quad + \mu\Bigg[R - \left(-\frac{a}{\varepsilon}(w-\bar{w}) + 1\right)\beta + \frac{\varepsilon}{2a}\\ &\quad + \frac{a}{2\varepsilon}(w-\bar{w})^2 - (w-\bar{w})\Bigg]\\ &= -\frac{a}{2\varepsilon}(w-\bar{w})\Bigg[2(\rho w + c) - r(w-\bar{w})\\ &\quad - 2\mu\beta - \mu(w-\bar{w}) + \frac{2\mu\varepsilon}{a}\Bigg] - r\bar{U} + \mu\left(R - \beta + \frac{\varepsilon}{2a}\right)\\ &= -\frac{a}{2\varepsilon}(w-\bar{w})\left[(2\rho - r - \mu)(w-\bar{w}) + \frac{2\mu\varepsilon}{a}\right] + \mu\left(\frac{\varepsilon}{2a}\right)\\ &\le \frac{\mu^2\varepsilon}{2(2\rho - \mu - r)a} + \frac{\mu\varepsilon}{2a},\end{aligned}$$

where the last equality uses $r\bar{U} = \mu(R-\beta)$ and the inequality follows from that $G(w)$ is a quadratic function whose maximizer is $\bar{w} - \dfrac{\mu\varepsilon}{a(2\rho - r - \mu)} > \bar{w} - \varepsilon$.

For $w\in[0,\bar{w}-\varepsilon]$,

$$\begin{aligned}G(w) &= (a+1)(\rho w + c) - \rho w - c - r(\underline{v} + (a+1)w) + rw + \mu\\ &\quad\times\left[R - a\beta + \bar{V}_{\bar{w}} + \frac{\varepsilon}{2a} - aw - \underline{v}\right]\\ &= \left[(a+1)(\rho - r - \mu) - (\rho - r)\right](w-\bar{w}) + \frac{\mu\varepsilon}{2}\frac{1}{a} \le \frac{\mu\varepsilon}{2}\frac{1}{a},\end{aligned}$$

where the inequality follows from $a > \dfrac{\mu}{\rho - r - \mu}$.

For $w \ge \bar{w} + \varepsilon/a$,

$$\begin{aligned}G(w) &= -\rho w - c - r\left(\bar{V}_{\bar{w}} + \frac{\varepsilon}{2a}\right) + rw + \mu R\\ &= r(\bar{V} - \bar{V}_{\bar{w}}) - (\rho - r)w - \frac{r\varepsilon}{2a} = (\rho - r)(\bar{w} - W_{t-}) \le 0.\end{aligned}$$

Hence, it follows from (B.15) that

$$\Psi_t \le G(W_{t-}) \le K_0\varepsilon,\ \text{where } K_0 := \frac{\mu^2}{2(2\rho - \mu - r)a} + \frac{\mu}{2a}.$$

Although we cannot directly apply Lemma B.1, we can follow the first inequality of (B.12) in the proof of Lemma B.1 to show that

$$\begin{aligned}F_\varepsilon(W_{0-}) &\ge \mathbb{E}\Bigg[e^{-r\tau}F_\varepsilon(W_\tau) + \int_0^\tau e^{-rt}(RdN_t\\ &\quad - dL_t) - \int_0^\tau e^{-rt}\mathcal{B}_t - \int_0^\tau e^{-rt}\Psi_t dt\Bigg]\\ &\ge \mathbb{E}\Bigg[e^{-r\tau}F_\varepsilon(W_\tau) + \int_0^\tau e^{-rt}(RdN_t\\ &\quad - dL_t)\Bigg] - \frac{K_0}{r}\varepsilon \ge U(\Gamma) - K\varepsilon,\end{aligned}$$

where $K := K_0/r$. Hence, letting $\varepsilon \to 0$ in the inequality yields $F(w) \geq U(\Gamma(w))$. $\square$

B.4. Proof of Lemmas 2 and 3

To prove Lemmas 2 and 3, we start from a simple case that $\mu + r \leq \rho$ and then turn to analyze the more complex case that $\mu + r > \rho$ later. The main difference between these two cases is that $b_{\tilde{w}}$ defined in (B.16) will take a different limit as $\tilde{w}$ tends to $\bar{w}$, which makes further analysis rather different.

B.4.1. The Case That $\mu + r \leq \rho$. We have $\rho > \underline{\mu}$, which further implies that $\bar{w} < \beta$. For any $\hat{w} = \tilde{w} \in (0, \bar{w})$, differential Equation (4) becomes an ordinary differential equation (ODE), which can be solved in closed form. We further distinguish two subcases: $\mu + r < \rho$ versus $\mu + r = \rho$.
Subcase 1. $\mu + r < \rho$. We have

$$V_{\tilde{w}}(w) = \frac{\rho - r}{r + \mu - \rho}(\bar{w} - w) + \frac{\mu V_{\tilde{w}}(\tilde{w}) + r\bar{V} + (r - \rho)\bar{w}}{r + \mu} + b_{\tilde{w}} \times (\bar{w} - w)^{\frac{r+\mu}{\rho}} \quad \text{for} \quad w \in [0, \tilde{w}],$$

where

$$b_{\tilde{w}} := \frac{r - \rho}{r + \mu - \rho}\frac{\rho}{r + \mu}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}} > 0. \tag{B.16}$$

Also, we have

$$V''_{\tilde{w}}(w) = -\frac{\rho - r}{\rho}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}}(\bar{w} - w)^{\frac{-2\rho + r + \mu}{\rho}} < 0 \quad \text{if} \quad w \in [0, \tilde{w}].$$

Hence, Lemma 2 holds with $\check{w}(\tilde{w}) = 0$, and thus, $\mathcal{V}_{\tilde{w}}(w) = V_{\tilde{w}}(w)$.

If $\tilde{w}_1 < \tilde{w}_2$, then we have $b_{\tilde{w}_1} > b_{\tilde{w}_2}$, $\mathcal{V}'_{\tilde{w}_1}(w) < \mathcal{V}'_{\tilde{w}_2}(w)$, and $\mathcal{V}_{\tilde{w}_1}(w) > \mathcal{V}_{\tilde{w}_2}(w)$ for $w \in [0, \tilde{w}_1]$. Hence, $\mathcal{V}_{\tilde{w}}(0)$ is strictly decreasing in $\tilde{w}$.

As $\tilde{w} \to \bar{w}$, we have $b_{\tilde{w}} \to 0$, and thus,

$$\mathcal{V}_{\tilde{w}}(0) \to \frac{\rho - r}{r + \mu - \rho}\bar{w} + \bar{V} - \frac{(\rho - r)\bar{w}}{r} = \frac{\rho - r}{r + \mu - \rho}\bar{w} + \bar{U} + \bar{w} = \bar{U} + \frac{\mu}{r + \mu - \rho}\bar{w} < \underline{v},$$

where the inequality follows from that Condition (13) is not satisfied. Additionally, if $\tilde{w} = 0$, we have $\mathcal{V}_{\tilde{w}}(0) = \bar{V} > \underline{v}$. It is trivial to see that $\mathcal{V}_{\tilde{w}}(0)$ is continuous in $\tilde{w}$. Hence, there exists a unique $\hat{w} \in (0, \bar{w})$ such that $\mathcal{V}_{\hat{w}}(0) = \underline{v}$. This concludes Lemma 3.
Subcase 2. $\mu + r = \rho$. Now, we have

$$V_{\tilde{w}}(w) = \frac{(\rho - r)(\bar{w} - w)}{\rho} - \frac{(\rho - r)(\bar{w} - w)}{\rho}\ln\left(\frac{\bar{w} - w}{\bar{w} - \tilde{w}}\right) + \frac{r\bar{V} - (\rho - r)\bar{w} + \mu V_{\tilde{w}}(\tilde{w})}{\rho} \quad \text{for} \quad w \in [0, \tilde{w}].$$

Further, we have

$$V'_{\tilde{w}}(w) = \frac{\rho - r}{\rho}\ln\left(\frac{\bar{w} - w}{\bar{w} - \tilde{w}}\right) \geq 0 \text{ and}$$
$$V''_{\tilde{w}}(w) = -\frac{\rho - r}{\rho(\bar{w} - w)} < 0, \quad w \in [0, \tilde{w}].$$

Hence, in this case, Lemma 2 holds with $\check{w}(\tilde{w}) = 0$, and thus, $\mathcal{V}_{\tilde{w}}(w) = V_{\tilde{w}}(w)$.

If $\tilde{w}_1 < \tilde{w}_2$, then we have $\mathcal{V}'_{\tilde{w}_1}(w) < \mathcal{V}'_{\tilde{w}_2}(w)$ and $\mathcal{V}_{\tilde{w}_1}(w) > \mathcal{V}_{\tilde{w}_2}(w)$ for $w \in [0, \tilde{w}_1]$. Hence, $\mathcal{V}_{\tilde{w}}(0)$ is strictly decreasing in $\tilde{w}$. As $\tilde{w} \to \bar{w}$, $\mathcal{V}'_{\tilde{w}}(w) \to -\infty$ for $w \in [0, \tilde{w})$, which implies that $\lim_{\tilde{w} \to \bar{w}} \mathcal{V}_{\tilde{w}}(0) = -\infty$. Additionally, if $\tilde{w} = 0$, then $\mathcal{V}_{\tilde{w}}(0) = \bar{V} > \underline{v}$. Hence, there exists $\hat{w}$ such that $\mathcal{V}_{\hat{w}}(0) = \underline{v}$. Lemma 3 is also obtained.

B.4.2. The Case That $\mu + r > \rho$. For this case, we prove Lemmas 2 and 3 in the following steps.

Step 1. For any $\tilde{w} < \bar{w}$, there exists a unique continuously differentiable function $V_{\tilde{w}}$ that satisfies (4) with the boundary condition at $\tilde{w}$. Following Lemma A.1, we have that $V_{\tilde{w}}(w)$ is strictly increasing in w on $[0, \tilde{w})$.

Step 2. There exists a $\check{w}(\tilde{w}) \in [0, \tilde{w})$, such that $V''_{\tilde{w}}(w) < 0$ for $w \in (\check{w}(\tilde{w}), \tilde{w}]$ and $V''_{\tilde{w}}(w) > 0$ for $w \in (\check{w}(\tilde{w}), \tilde{w}]$. Moreover, if $\check{w}(\tilde{w}) > 0$, then $V'_{\tilde{w}}(\check{w}(\tilde{w})) > 1$.

Step 3. There exists a unique $\hat{w} \in [0, \bar{w})$ such that $\mathcal{V}_{\hat{w}}(0) = \underline{v}$. Additionally, $\mathcal{V}_{\hat{w}}(w)$ is strictly increasing and concave in w on $[0, \tilde{w})$. Finally, if $\check{w}(\hat{w}) > 0$, then $\mathcal{V}'_{\hat{w}}(0) = \mathcal{V}'_{\hat{w}}(\check{w}(\hat{w})) > 1$.

Steps 1 and 3 are similar to the proof of proposition 4 in Sun and Tian (2018). However, we still provide a detailed argument here to ensure the self-containment of this comment.
Step 1. For $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$, differential Equation (4) becomes an ODE, and thus,

$$V_{\tilde{w}}(w) = \frac{\rho - r}{r + \mu - \rho}(\bar{w} - w) + \frac{\mu V_{\tilde{w}}(\tilde{w}) + r\bar{V} + (r - \rho)\bar{w}}{r + \mu} + b_{\tilde{w}} \times (\bar{w} - w)^{\frac{r+\mu}{\rho}} \quad \text{if} \quad w \in \left[(\tilde{w} - \beta)^+, \tilde{w}\right], \tag{B.17}$$

with $b_{\tilde{w}} = \frac{r - \rho}{r + \mu - \rho}\frac{\rho}{r + \mu}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}} < 0$ from the boundary condition.

It is straightforward to verify that $V_{\tilde{w}}(w)$ is increasing and strictly concave in w on $[(\tilde{w} - \beta)^+, \tilde{w}]$. Moreover, we have $V'_{\tilde{w}}(\tilde{w}-) = 0 = V'_{\tilde{w}'}(\tilde{w}+) = 0$. Therefore, Equation (4) is reduced to a sequence of initial value problems over the intervals $[(\tilde{w} - (k+1)\beta)^+, (\tilde{w} - k\beta)^+)$, $k \in \mathbb{N} \setminus \{0\}$, which satisfy the assumptions of the Cauchy–Lipschitz theorem and therefore, admit unique continuously differentiable solutions.
Step 2. Differentiating (4) at w yields

$$\rho(\bar{w} - w)V''_{\tilde{w}}(w) = \mu\left(V'_{\tilde{w}}(w + \beta) - V'_{\tilde{w}}(w)\right) + (\rho - r)\left(V'_{\tilde{w}}(w) - 1\right). \tag{B.18}$$

Differentiating (B.18) at w yields

$$\rho(\bar{w} - w)V'''_{\tilde{w}}(w) = \mu\left(V''_{\tilde{w}}(w + \beta) - V''_{\tilde{w}}(w)\right) + (2\rho - r)V''_{\tilde{w}}(w). \tag{B.19}$$

If $\tilde{w} \leq \beta$, then $V_{\tilde{w}}$ has the close form Expression (B.17) for $w \in [0, \tilde{w}]$, which easily concludes that $V''_{\tilde{w}} < 0$ over $[0, \tilde{w})$. Hence, Lemma 2 holds with $\check{w}(\tilde{w}) = 0$.

Now, we consider the case that $\tilde{w} > \beta$. Let $w^c := \inf\{w \in [0, \tilde{w}) : V''_{\tilde{w}}(w) \geq 0\}$. If the set is empty, we set $w^c = 0$. Obviously, $V''_{\tilde{w}}(w) < 0$ over $(w^c, \tilde{w})$. Hence, Lemma 2 holds with $\check{w}(\tilde{w}) = w^c$ if $w^c = 0$.

Next, we consider the case that $w^c > 0$. Because $V_{\tilde{w}}(w)$ is strictly concave in w on $[\tilde{w} - \beta, \tilde{w})$, we have $w^c < \tilde{w} - \beta$.

Because $V_{\tilde{w}} \in C^2([0,\tilde{w}))$, we have $V''_{\tilde{w}}(w^c) = 0$ and $V''_{\tilde{w}}(w) < 0$ over $(w^c, \tilde{w})$. It follows from (B.18) at w^c that

$$\mu\big(V'_{\tilde{w}}(w^c+\beta) - V'_{\tilde{w}}(w^c)\big) = (\rho - r)\big(1 - V'_{\tilde{w}}(w^c)\big),$$

which implies

$$V'_{\tilde{w}}(w^c+\beta) = \frac{(\mu-\rho+r)V'_{\tilde{w}}(w^c) + (\rho-r)}{\mu}. \tag{B.20}$$

Moreover, because $V'_{\tilde{w}}$ decreases over $(w^c, \tilde{w})$, we have $V'_{\tilde{w}}(w^c+\beta) < V'_{\tilde{w}}(w^c)$, which yields

$$V'_{\tilde{w}}(w^c) > 1, \tag{B.21}$$

in view of (B.20).

It follows from (4) at w^c that

$$\begin{aligned} rV_{\tilde{w}}(w^c) &= \mu R - c - (\rho-r)w^c - \rho(\bar{w}-w^c)V'_{\tilde{w}} \times (w^c) \\ &\quad + \mu\big(V_{\tilde{w}}(w^c+\beta) - V_{\tilde{w}}(w^c)\big) \\ &= \mu R - c - (\rho-r)w^c - \rho(\bar{w}-w^c)V'_{\tilde{w}} \times (w^c) \\ &\quad + \mu\beta V'_{\tilde{w}}(w^b) \\ &> \mu R - c - (\rho-r)w^c - \rho(\bar{w}-w^c)V'_{\tilde{w}} \times (w^c) \\ &\quad + \mu\beta V'_{\tilde{w}}(w^c+\beta) \\ &= \mu R - c - (\rho-r)(w^c-\beta) \\ &\quad + \big[\rho(w^c-\beta) + r\beta + c\big]V'_{\tilde{w}}(w^c), \end{aligned} \tag{B.22}$$

where the second equality follows from the mean value theorem with some $w^b \in (w^c, w^c+\beta)$, the inequality follows from $V''_{\tilde{w}}(w) < 0$ over $(w^c, w^c+\beta)$, and the last equality follows from (B.20) and $\rho\bar{w} = \mu\beta - c$.

We consider two cases.

Case 1. $\rho(w^c-\beta) + c + \beta r \ge 0$. It follows from (B.22) and $V'_{\tilde{w}}(w^c) > 1$ that

$$\begin{aligned} rV_{\tilde{w}}(w^c) &> \mu R - c - (\rho-r)(w^c-\beta) \\ &\quad + \big[\rho(w^c-\beta) + r\beta + c\big] \\ &= \mu R - c + r(w^c-\beta) + c + \beta r \\ &> \mu R - c + c > \mu R - c - (\rho-r)\hat{w} \\ &= rV_{\tilde{w}}(\tilde{w}), \end{aligned}$$

which reaches a contradiction with Lemma A.1.

Case 2. $\rho(w^c-\beta) + c + \beta r < 0$. Hence, we have

$$0 < w^c < \frac{-c + (\rho-r)\beta}{\rho}, \tag{B.23}$$

which implies that $(\rho-r)\beta > c$ (i.e., $\rho - r > \Delta\mu$). In this case, we will show that

$$V''_{\tilde{w}} > 0 \text{ over } [0, w^c). \tag{CX}$$

It follows from (B.19) at w^c and $V''_{\tilde{w}}(w^c) = 0$ that

$$\rho(\bar{w}-w^c)V'''_{\tilde{w}}(w^c) = \mu V''_{\tilde{w}}(w^c+\beta) < 0,$$

which implies that $V''_{\tilde{w}} > 0$ over $(w^c-\varepsilon, w^c)$ for some $\varepsilon > 0$. If (CX) fails to hold, then $w^d := \sup\{w \in [0,w^c) : V''_{\tilde{w}}(w) \le 0\}$ is well defined and $w^d \in [0, w^c)$. Moreover, we have $V''_{\tilde{w}}(w^d) = 0$ and $V'''_{\tilde{w}}(w^d) \ge 0$. Hence, it follows from (B.19) at w^d that $\rho(\bar{w}-w^d)V'''_{\tilde{w}}(w^d) = \mu V''_{\tilde{w}}(w^d+\beta) \ge 0$. By the definition of w^c, we have $w^d + \beta \le w^c$. Consequently, it follows from (B.23) that $w^d \le w^c - \beta < 0$, which reaches a contradiction. Hence, (CX) holds. Therefore, letting $\check{w}(\tilde{w}) = w^c$, we complete the proof of Lemma 2. Additionally, we have $V'_{\tilde{w}}(w^c) = V'_{\tilde{w}}(\check{w}(\tilde{w})) > 1$ in view of (B.21).

Step 3. Following the definition (14), we immediately have that $\mathcal{V}_{\tilde{w}}$ is strictly increasing and concave in w on $[0,\tilde{w})$. Next, we prove that $\mathcal{V}_{\tilde{w}}(0)$ is strictly decreasing in $\tilde{w}$.

First, we prove that if $\tilde{w}_1 < \tilde{w}_2 \in (0,\bar{w})$, then $V_{\tilde{w}_1}(w) > V_{\tilde{w}_2}(w)$ and $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ for $w \in [0,\tilde{w}_1]$. An equivalent argument is if $\tilde{w}_1 < \tilde{w}_2 \in (0,\bar{w})$ and $\tilde{w}_2 - \tilde{w}_1 \le \beta/2$, then $V_{\tilde{w}_1}(w) > V_{\tilde{w}_2}(w)$ and $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ for $w \in [0,\tilde{w}_1]$.

Because for any $\tilde{w} \in (0,\bar{w})$, $V_{\tilde{w}}$ is continuously differentiable, $V_{\tilde{w}_1}(w) - V_{\tilde{w}_2}(w)$ must also be continuously differentiable.

In the interval $[\tilde{w}_1 - \beta/2, \tilde{w}_1]$, $V_{\tilde{w}_1}(w) > V_{\tilde{w}_2}(w)$ and $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ because $b_{\tilde{w}_1} > b_{\tilde{w}_2}$ and $V_{\tilde{w}_1}(\tilde{w}_1) > V_{\tilde{w}_2}(\tilde{w}_2) > V_{\tilde{w}_2}(\tilde{w}_1)$. In the interval $[\tilde{w}_1, \tilde{w}_1 + \beta/2]$, on the other hand, $V_{\tilde{w}_1}(w) > V_{\tilde{w}_2}(w)$ and $0 = V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ because $V_{\tilde{w}_1}(\tilde{w}_1) > V_{\tilde{w}_2}(\tilde{w}_2)$.

Now, we claim that $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$, $\forall w \in [0,\tilde{w}_1]$. Otherwise, because $V_{\tilde{w}_1}(w) - V_{\tilde{w}_2}(w)$ is continuously differentiable, there must exist a $\tilde{w}' := \max\{w \in [0,\tilde{w}_1] : V'_{\tilde{w}_1}(w) - V'_{\tilde{w}_2}(w) = 0\}$. Then, we obtain $\mu(V_{\tilde{w}_1}(\tilde{w}'+\beta) - V'_{\tilde{w}_2}(\tilde{w}'+\beta)) = (r+\mu)(V_{\tilde{w}_1}(\tilde{w}') - V_{\tilde{w}_2}(\tilde{w}'))$. However, it contradicts with

$$\begin{aligned} 0 < V_{\tilde{w}_1}(\tilde{w}'+\beta) - V_{\tilde{w}_2}(\tilde{w}'+\beta) &= V_{\tilde{w}_1}(\tilde{w}') - V_{\tilde{w}_2}(\tilde{w}') \\ &\quad + \int_0^\beta \Big[V'_{\tilde{w}_1}(\tilde{w}'+x) - V'_{\tilde{w}_2}(\tilde{w}'+x)\Big]dx. \end{aligned}$$

Then, we must have $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ and $V'_{\tilde{w}_1}(w) > V'_{\tilde{w}_2}(w)$, $\forall w \in [0,\tilde{w}_1]$.

Now, we go back to prove that $\mathcal{V}_{\tilde{w}_1}(0) > \mathcal{V}_{\tilde{w}_2}(0)$. We consider two cases.

Case 1. $\check{w}(\tilde{w}_1) \ge \check{w}(\tilde{w}_2)$. We have $\mathcal{V}_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) = V_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) > V_{\tilde{w}_2}(\check{w}(\tilde{w}_1)) = \mathcal{V}_{\tilde{w}_2}(\check{w}(\tilde{w}_1))$ and $\mathcal{V}'_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) = V'_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) < V'_{\tilde{w}_2}(\check{w}(\tilde{w}_1)) = \mathcal{V}'_{\tilde{w}_2}(\check{w}(\tilde{w}_1))$. Consequently,

$$\begin{aligned} \mathcal{V}_{\tilde{w}_1}(0) &= V_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) - V'_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) \cdot \check{w}(\tilde{w}_1) \\ &> V_{\tilde{w}_2}(\check{w}(\tilde{w}_1)) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_1)) \cdot \check{w}(\tilde{w}_1) \\ &\ge V_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) \cdot \check{w}(\tilde{w}_1) \\ &= \mathcal{V}_{\tilde{w}_2}(0), \end{aligned}$$

where the second inequality follows by noting that $[V_{\tilde{w}_2}(w) - V'_{\tilde{w}_2}(w) \cdot \check{w}(\tilde{w}_1)]' = V'_{\tilde{w}_2}(w) - V''_{\tilde{w}_2}(w) \cdot \check{w}(\tilde{w}_1) > 0$ on $[\check{w}(\tilde{w}_2), \check{w}(\tilde{w}_1)]$ because $V_{\tilde{w}_2}(w)$ is strictly increasing and concave in w on $[\check{w}(\tilde{w}_2), \check{w}(\tilde{w}_1)]$.

Case 2. $\check{w}(\tilde{w}_1) < \check{w}(\tilde{w}_2)$. We have $\mathcal{V}_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) = V_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) > V_{\tilde{w}_2}(\check{w}(\tilde{w}_1))$ and $\mathcal{V}'_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) = V'_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) < V'_{\tilde{w}_2}(\check{w}(\tilde{w}_1))$. Hence,

$$\begin{aligned} \mathcal{V}_{\tilde{w}_1}(0) &= V_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) - V'_{\tilde{w}_1}(\check{w}(\tilde{w}_1)) \cdot \check{w}(\tilde{w}_1) \\ &> V_{\tilde{w}_2}(\check{w}(\tilde{w}_1)) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_1) \cdot \check{w}(\tilde{w}_1) \\ &> V_{\tilde{w}_2}(\check{w}(\tilde{w}_1)) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) \cdot \check{w}(\tilde{w}_1) \\ &> V_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) \cdot \check{w}(\tilde{w}_2) \\ &= \mathcal{V}_{\tilde{w}_2}(0), \end{aligned}$$

where the second inequality follows by noting that $V''_{\tilde{w}_1} > 0$ over $(\check{w}(\tilde{w}_1), \check{w}(\tilde{w}_2))$ and the third inequality follows by noting that $[V_{\tilde{w}_2}(w) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) \cdot w]' = V'_{\tilde{w}_2}(w) - V'_{\tilde{w}_2}(\check{w}(\tilde{w}_2)) < 0$ for $w \in (\check{w}(\tilde{w}_1), \check{w}(\tilde{w}_2))$ because $V_{\tilde{w}_2}(w)$ is convex in $[\check{w}(\tilde{w}_1), \check{w}(\tilde{w}_2))$.

Combining the two cases, we conclude that $\mathcal{V}_{\tilde{w}}(0)$ is strictly decreasing in $\tilde{w}$.

Additionally, if $\tilde{w}=0$, then the boundary condition states that $\mathcal{V}_{\tilde{w}}(0)=\bar{V}>\underline{v}$. If we let $\tilde{w}\to\bar{w}$, then $b_{\tilde{w}}\to-\infty$, $\mathcal{V}_{\tilde{w}}(\tilde{w}-\beta)=V_{\tilde{w}}(\tilde{w}-\beta)\to-\infty$. Because $\mathcal{V}_{\tilde{w}}$ is continuous and increasing, we have $\mathcal{V}_{\tilde{w}}(0)\to-\infty$.

Therefore, there must exist a unique $\hat{w}\in(0,\bar{w})$ that satisfies the additional boundary condition $\mathcal{V}_{\hat{w}}(0)=\underline{v}$. Additionally, recalling (B.21), we have that if $\check{w}(\hat{w})>0$, then $\mathcal{V}'_{\hat{w}}(0)=\mathcal{V}'_{\hat{w}}(\check{w}(\hat{w}))>1$. □

B.5. Proof of Proposition 3

If $\check{w}=0$, the proof is exactly the same as that for part 1 of proposition 6 in Sun and Tian (2018). Hence, we only need to prove it for the case that $\check{w}>0$.

First, it is easy to see that $U(\Gamma_d^*(0))=\underline{v}=F_{\hat{w}}(0)$. Given $W_{0-}=w\in[\check{w},\hat{w}]$, following Ito's formula for jump processes (see, for example, theorem 17.5 of Bass 2011) and considering (15), we have

$$\begin{aligned}e^{-r\tau}F_{\hat{w}}(W_\tau)&=F_{\hat{w}}(W_{0-})+\int_0^\tau\left[e^{-rt}dF_{\hat{w}}(W_{t-})-re^{-rt}F_{\hat{w}}(W_{t-})dt\right]\\&=F_{\hat{w}}(W_{0-})+\int_0^\tau e^{-rt}\left(-RdN_t+dL_t^*\right)+\int_0^\tau e^{-rt}\mathcal{A}_t,\end{aligned}\tag{B.24}$$

where

$$\begin{aligned}\mathcal{A}_t:=&\,dF_{\hat{w}}(W_{t-})-rF_{\hat{w}}(W_{t-})dt+RdN_t-dL_t^*\\=&\,F'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})\mathbb{1}_{W_{t-}>\check{w}(\hat{w})}dt\\&+F_{\hat{w}}\left(W_{t-}+[\beta\wedge(\hat{w}-W_{t-})]dN_t-\check{w}\mathbb{1}_{W_{t-}=\check{w}(\hat{w})}dQ_t\right)\\&-F_{\hat{w}}(W_{t-})\\&+RdN_t-\left(W_{t-}+\beta-\hat{w}\right)^+dN_t-rF_{\hat{w}}(W_{t-})dt\\=&\,F'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})\mathbb{1}_{W_{t-}>\check{w}(\hat{w})}dt\\&+\left[F_{\hat{w}}\left(W_{t-}+[\beta\wedge(\hat{w}-W_{t-})]\right)-F_{\hat{w}}(W_{t-})\right.\\&\left.-\left(W_{t-}+\beta-\hat{w}\right)^+\right]dN_t\\&+[F_{\hat{w}}(0)-F_{\hat{w}}(W_{t-})]dQ_t+RdN_t-rF_{\hat{w}}(W_{t-})dt\\=&\,F'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})\mathbb{1}_{W_{t-}>\check{w}(\hat{w})}dt\\&+\left[F_{\hat{w}}\left(W_{t-}+\beta\right)-F_{\hat{w}}(W_{t-})\right]dN_t\\&+[F_{\hat{w}}(0)-F_{\hat{w}}(\check{w}(\hat{w}))]dQ_t+RdN_t-rF_{\hat{w}}(W_{t-})dt.\end{aligned}$$

Further define

$$\begin{aligned}\mathcal{B}_t:=&\,[F_{\hat{w}}(0)-F_{\hat{w}}(\check{w}(\hat{w}))]\left(dQ_t-q_t^*dt\right)+\left[F_{\hat{w}}\left(W_{t-}+\beta\right)\right.\\&\left.-F_{\hat{w}}(W_{t-})\right]\left(dN_t-\mu dt\right)+R\left(dN_t-\mu dt\right).\end{aligned}\tag{B.25}$$

If $W_{t-}\in(\check{w}(\hat{w}),\hat{w}]$, then

$$\begin{aligned}\mathcal{A}_t=&\,F'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})dt+\left[F_{\hat{w}}\left(W_{t-}+\beta\right)\right.\\&\left.-F_{\hat{w}}(W_{t-})\right]dN_t+RdN_t-rF_{\hat{w}}(W_{t-})dt\\=&\,F'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})dt+\mu\left[F_{\hat{w}}\left(W_{t-}+\beta\right)\right.\\&\left.-F_{\hat{w}}(W_{t-})\right]dt+\mu Rdt-rF_{\hat{w}}(W_{t-})dt+\mathcal{B}_t\\=&\,\{V'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})-\rho(W_{t-}-\bar{w})+\mu\\&\times\left[V_{\hat{w}}\left(W_{t-}+\beta\right)-V_{\hat{w}}(W_{t-})\right]-\mu\beta+\mu R\\&-rV_{\hat{w}}(W_{t-})+rW_{t-}\}dt+\mathcal{B}_t\\=&\,\{V'_{\hat{w}}(W_{t-})\rho(W_{t-}-\bar{w})-\left(\rho-r\right)W_{t-}+\mu\\&\times\left[V_{\hat{w}}\left(W_{t-}+\beta\right)-V_{\hat{w}}(W_{t-})\right]-c+\mu R\\&-rV_{\hat{w}}(W_{t-})\}dt+\mathcal{B}_t\\=&\,\mathcal{B}_t,\end{aligned}$$

where the last equality follows from (4).

If $W_{t-}=\check{w}(\hat{w})$ and $\check{w}(\hat{w})>0$, then

$$\begin{aligned}\mathcal{A}_t=&\left[F_{\hat{w}}\left(\check{w}(\tilde{w})+\beta\right)-F_{\hat{w}}(\check{w}(\tilde{w}))\right]dN_t+RdN_t\\&-rF_{\hat{w}}(\check{w}(\tilde{w}))dt+[F_{\hat{w}}(0)-F_{\hat{w}}(\check{w}(\hat{w}))]dQ_t\\=&\,\mu\left[F_{\hat{w}}\left(\check{w}(\tilde{w})+\beta\right)-F_{\hat{w}}(\check{w}(\tilde{w}))\right]dt+\mu Rdt\\&-rF_{\hat{w}}(\check{w}(\tilde{w}))dt+[F_{\hat{w}}(0)-F_{\hat{w}}(\check{w}(\hat{w}))]\\&\times q_t^*dt+\mathcal{B}_t\\=&\,\{\mu\left[V_{\hat{w}}\left(\check{w}(\tilde{w})+\beta\right)-V_{\hat{w}}(\check{w}(\tilde{w}))\right]-\mu\beta+\mu R\\&-rV_{\hat{w}}(\check{w}(\tilde{w}))+r\check{w}(\tilde{w})+[F_{\hat{w}}(0)\\&-F_{\hat{w}}\left(\check{w}(\hat{w})\right)]q_t^*\}dt+\mathcal{B}_t\\=&\,\{V'_{\hat{w}}(\check{w}(\tilde{w}))\rho(\bar{w}-\check{w}(\hat{w}))+\rho\check{w}+c-\mu\beta\\&+[F_{\hat{w}}(0)-F_{\hat{w}}(\check{w}(\hat{w}))]q_t^*\}dt+\mathcal{B}_t\\=&\left\{F'_{\hat{w}}(\check{w}(\tilde{w}))\rho(\bar{w}-\check{w}(\hat{w}))+[F_{\hat{w}}(0)\right.\\&\left.-F_{\hat{w}}(\check{w}(\hat{w}))]\frac{\rho(\bar{w}-\check{w}(\hat{w}))}{\check{w}(\hat{w})}\right\}dt+\mathcal{B}_t\\=&\,\rho(\bar{w}-\check{w}(\hat{w}))\left[F'_{\hat{w}}(\check{w}(\tilde{w}))\right.\\&\left.-\frac{F_{\hat{w}}(\check{w}(\hat{w}))-F_{\hat{w}}(0)}{\check{w}(\hat{w})}\right]dt+\mathcal{B}_t\\=&\,\rho(\bar{w}-\check{w}(\hat{w}))\left[\mathcal{V}'_{\hat{w}}(\check{w}(\tilde{w}))\right.\\&\left.-\frac{\mathcal{V}_{\hat{w}}(\check{w}(\hat{w}))-\mathcal{V}_{\hat{w}}(0)}{\check{w}(\hat{w})}\right]dt+\mathcal{B}_t=\mathcal{B}_t,\end{aligned}$$

where the forth equality follows from (4) and the seventh equality follows from the definition of $\mathcal{V}_{\hat{w}}(w)$.

In both cases, we have $\mathcal{A}_t=\mathcal{B}_t$. Taking the expectation on both sides of (B.24), we have

$$\begin{aligned}F_{\hat{w}}(w)&=F_{\hat{w}}(W_{0-})\\&=\mathbb{E}\left[e^{-r\tau}F(W_\tau)+\int_0^\tau e^{-rt}\left(RdN_t-dL_t^*\right)\right.\\&\quad\left.-\int_0^\tau e^{-rt}\mathcal{B}_tdt\right]\\&=\mathbb{E}\left[e^{-r\tau}F(W_\tau)+\int_0^\tau e^{-rt}\left(RdN_t-dL_t^*\right)\right]\\&=U\left(\Gamma_d^*(w)\right),\end{aligned}$$

where the third inequality follows from the fact that $\int_0^\tau e^{-rt}\mathcal{B}_t$ is a martingale and the last equality follows from $F(W_\tau)=F(0)=\underline{v}$. Finally, for $W_{0-}=w>\hat{w}$, following Definition 1, we have

$$\begin{aligned}U\left(\Gamma_d^*(w)\right)&=U\left(\Gamma_d^*(\hat{w})\right)-(w-\hat{w})\\&=F_{\hat{w}}(\hat{w})-(w-\hat{w})=F_{\hat{w}}(w).\end{aligned}$$

To conclude, we have $U(\Gamma_d^*(w))=F_{\hat{w}}(w)$ for $w\geq\check{w}$. Following Lemma 3, we have that if $\check{w}>0$, then $w\in[0,\check{w}]$, $F'_{\hat{w}}(w)=V'_{\hat{w}}(w)-1>0$, which further implies that $F_{\hat{w}}(\check{w})>F_{\hat{w}}(0)=\underline{v}$. □

B.6. Proof of Proposition 4

Part 2 in proposition 6 in Sun and Tian (2018) has already presented the result for $\check{w}=0$. We prove it for the case that $\check{w}>0$.

Note that $F_{\hat{w}}(w^*) \geq F_{\hat{w}}(w)$. The proof is complete if we can verify that $F_{\hat{w}}(w) \geq U(\Gamma)$ where $u(\Gamma, \nu^*) = w$. Recall that following Lemmas 2 and 3, we have that $F_{\hat{w}}(w)$ is a differentiable, concave, and upper-bounded function, with $F_{\hat{w}}(0) = \underline{v}$ and $F'_{\hat{w}}(w) \geq -1$. Based on Lemma B.1, to prove $F_{\hat{w}}(w) \geq U(\Gamma)$, we only need to show that $\{\Psi_t\}_{t\geq 0}$ is nonpositive almost surely when we let $F(w) = F_{\hat{w}}(w)$. Following (B.15), we have

$$\begin{aligned}\Psi_t \leq\ & \mathcal{V}'_{\hat{w}}(W_{t-})\big(\rho W_{t-} + c\big) - \rho W_{t-} - c\\ & - r\mathcal{V}_{\hat{w}}(W_{t-}) + rW_{t-} + \mu\big[R - \mathcal{V}'_{\hat{w}}(W_{t-})\beta\\ & + \mathcal{V}_{\hat{w}}\big(W_{t-} + \beta\big) - \mathcal{V}_{\hat{w}}(W_{t-})\big].\end{aligned}$$

If $W_{t-} \geq \check{w}$, then

$$\begin{aligned}\Psi_t \leq\ & \mathcal{V}'_{\hat{w}}(W_{t-})\rho(W_{t-} - \bar{w}) - \rho W_{t-} - c\\ & - r\mathcal{V}_{\hat{w}}(W_{t-}) + rW_{t-} + \mu\big[R + \mathcal{V}_{\hat{w}}\big(W_{t-} + \beta\big)\\ & - \mathcal{V}_{\hat{w}}(W_{t-})\big] = 0, \end{aligned} \tag{B.26}$$

where the equality follows from the fact that for $w \in [\check{w}, \infty]$, $\mathcal{V}_{\hat{w}}(w) = V_{\hat{w}}(w)$ and (4).

If $W_{t-} \in [0, \check{w})$, then $\mathcal{V}'_{\hat{w}}(W_{t-}) = \mathcal{V}'_{\hat{w}}(\check{w})$. Define

$$\begin{aligned}g(W_{t-}) :=\ & \mathcal{V}'_{\hat{w}}(W_{t-})\rho(W_{t-} - \bar{w}) - \rho W_{t-} - c\\ & - r\mathcal{V}_{\hat{w}}(W_{t-}) + rW_{t-} + \mu\big[R.\\ & + \mathcal{V}_{\hat{w}}\left(W_{t-} + \beta\right) - \mathcal{V}_{\hat{w}}(W_{t-})\big].\end{aligned}$$

We have

$$\begin{aligned}g'(W_{t-}) &= \mathcal{V}'_{\hat{w}}(\check{w})\rho - \rho - r\mathcal{V}'_{\hat{w}}(\check{w}) + r + \mu\big[\mathcal{V}'_{\hat{w}}\big(W_{t-} + \beta\big) - \mathcal{V}'_{\hat{w}}(\check{w})\big]\\ &= \big(\rho - r\big)\big(\mathcal{V}'_{\hat{w}}(\check{w}) - 1\big) + \mu\big[\mathcal{V}'_{\hat{w}}\big(W_{t-} + \beta\big) - \mathcal{V}'_{\hat{w}}(\check{w})\big]\\ &\geq \big(\rho - r\big)\big(\mathcal{V}'_{\hat{w}}(\check{w}) - 1\big) + \mu\big[\mathcal{V}'_{\hat{w}}\big(\check{w} + \beta\big) - \mathcal{V}'_{\hat{w}}(\check{w})\big] = 0,\end{aligned} \tag{B.27}$$

where the inequality follows from the concavity of $\mathcal{V}_{\hat{w}}$ and the last equality follows from $\rho(\bar{w} - \check{w})\mathcal{V}''_{\hat{w}}(\check{w}) = (\rho - r)\times(\mathcal{V}'_{\hat{w}}(\check{w}) - 1) + \mu[\mathcal{V}'_{\hat{w}}(\check{w} + \beta) - \mathcal{V}'_{\hat{w}}(\check{w})] = 0$. Therefore,

$$\Psi_t \leq g(W_{t-}) \leq g(\check{w}) = 0. \tag{B.28}$$

To conclude, we have established that $\Psi_t \leq 0$ for $t \geq 0$. This completes the proof. □

Endnote

[1] Sun and Tian (2018) refer to this as *Incentive-Compatible* contract.

References

Bass R (2011) *Stochastic Processes* (Cambridge University Press, Cambridge, UK).

Biais B, Mariotti T, Rochet J-C, Villeneuve S (2010) Large risks, limited liability, and dynamic moral hazard. *Econometrica* 78(1):73–118.

Brémaud P (1981) *Point Processes and Queues* (Springer-Verlag, Berlin).

Sun P, Tian F (2018) Optimal contract to induce continued effort. *Management Sci.* 64(9):4193–4217.