MANAGEMENT SCIENCE
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0025-1909 | EISSN 1526-5501 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Punish Underperformance with Suspension — Optimal Dynamic Contracts in the Presence of Switching Cost

Ping Cao
International Institute of Finance, School of Management, University of Science and Technology of China, Hefei 230026, China, pcao@ustc.edu.cn

Peng Sun
The Fuqua School of Business, Duke University, Durham, North Carolina 27708, USA, psun@duke.edu

Feng Tian
Faculty of Business and Economics, University of Hong Kong, Pofulam Road, Hong Kong, fengtian@hku.hk

This paper studies a dynamic principal-agent setting in which the principal needs to dynamically schedule an agent to work or to be suspended. When the agent is directed to work and exert effort, the arrival rate of a Poisson process is increased, which increases the principal's payoff. Suspension, on the other hand, serves as a threat to the agent by delaying future payments. A key feature of our setting is a switching cost whenever the suspension stops and the work starts again. We formulate the problem as an optimal control model with switching, and fully characterize the optimal control policies/contract structures under different parameter settings. Our analysis shows that when the switching cost is not too high, the optimal contract demonstrates a generalized control-band structure. The length of each suspension episode, on the other hand, is fixed. Overall, the optimal contract is easy to describe, compute, and implement.

Key words: dynamic contract; jump process; optimal control; switching cost.

1. Introduction

Designing dynamic contracts to manage incentives is an important and challenging problem. It often involves carefully scheduling "carrots and sticks" over time. In an environment where outcome is stochastically determined by an agent's unobservable effort, it is intuitive that rewards ("carrots"), often in the form of monetary payments, follow good performance. If the agent is cash constrained or has limited liability, however, the principal cannot charge the agent money for bad performance. Therefore, it may not be obvious how to design penalties ("sticks") when performance is bad and leverage them to achieve better contracts. Due to analytical challenges, many dynamic contract design models restrict the focus on contracts that induce agents to always exert effort (see, for example, Demarzo and Sannikov 2006, Sannikov 2008, Biais et al. 2010, Myerson 2015, Sun and Tian 2018). In these situations, the principal with commitment power can use potential contract termination as a form of penalty. That is, the principal can terminate the agent, which stops all future payment opportunities, if the outcome has been bad for a long enough period of time. The threat of termination helps the principal to induce effort while saving costly rewards. However, termination itself may also be quite costly to the principal, especially in situations where a replacement agent is hard to find. In this paper, we focus on an alternative approach to penalize for bad outcomes: temporarily suspending work and pay to the agent. Note that contract termination can be perceived as a special case of suspension, one that never ends.

Temporarily suspending an agent for a period of time in response to poor performance is common practice in certain industries. Real estate agencies in Hong Kong, for example, often temporarily suspend a sales representative's work and pay when the performance has not been up to standard for a period of time. The suspension stops when the situation changes, for example, when new business arises, and the representative is still available to come back. Although this form of temporary suspension is rarely written formally into employment contracts, it is a very common practice, according to our conversations with an agent who worked at the Centaline Property, one of the largest property agencies in Hong Kong.[1]

More broadly, businesses and government agencies sometimes use *furloughs* to temporarily relieve an employee of job responsibilities and pay for a fixed period of time, with the promise of continuing the employment after the furlough period ends. The recent and ongoing pandemic has witnessed more entities using suspensions with no pay or pay reduction, or unpaid leaves, in response to economic difficulties, for example in Hong Kong (Hong Kong Business Times 2020, Hong Kong Economic Times 2020) and Europe (Eurofound 2020, 2021). It is fair to say that furloughs have been used mostly to ease an employer's financial challenges (DerStandard 2009), rather than in response to employees' job performance. However, with these temporary suspension control mechanisms already in existence, one may naturally ask why we do not use them for incentive management as well.

In case it is unclear why the principal can use suspension as a punishment, here is the intuition. In order to motivate a rational agent to exert effort, which is a private action, the principal needs to pay rent, either in the form of an immediate payment or as a promise to be delivered later. During the suspension period, the agent loses this rent income. More precisely, future payments are delayed by the suspension. Payment delay is particularly painful to an agent who is *less patient* than the principal, and therefore serves as a threat. Such a threat can be used to ensure that the agent is willing to exert effort whenever asked to. It is worth noting that during suspension, the principal also cannot enjoy good outcomes brought by the agent's effort. However, compared with losing the agent forever due to termination, it is often less costly for the principal to endure a short period of time without the agent's effort.

Even though the intuition may be clear at this point as to *why* we may use suspension to punish underperformance, deciding *how* to schedule suspension episodes in a dynamic environment remains a challenge. For example, when should suspension start? How long should suspension last? Should the lengths of suspension periods be the same, or should they vary depending on what triggers the suspension? In this paper, we study the optimal scheduling of payment and suspension in a basic dynamic contract design model.

Specifically, we consider a continuous-time optimal contract design problem, in which a principal tries to incentivize an agent to increase the arrival rate of a Poisson process. We may think of the principal as a firm, say the aforementioned real estate agency in Hong Kong, the agent as a sales representative, and each arrival as a successful sale. Whenever the agent exerts effort, the instantaneous arrival *rate* increases. However, the effort/arrival rate is unobservable to the principal and costly to the agent. Therefore, frequent arrivals are associated with good performance, and no arrival for a long period of time is bad. In this dynamic setting, when the two players' discount rates are the same, it is optimal for the principal, who has commitment power, to never temporarily suspend the agent (see Sun and Tian 2018, for the corresponding optimal contract). In practice, their time discount rates are often different. In particular, the principal (employer) often possesses more financial resources, and therefore is more patient than the agent (employee). In this case, suspending the agent once in a while may be beneficial.

1.1. Optimal Contract Structure

Generally speaking, our analysis reveals that optimal contract structures demonstrate three possibilities depending on model parameters, as illustrated in the three regions of Figure 2 later in the paper. First, if the switching cost is rather high, it is not worth paying for the switching cost to start working. In this case, intuitively, it is optimal for the principal not to hire the agent. The second possibility is when the switching cost is not that high but the revenue from each arrival is high enough. In this case, it is optimal to motivate the agent to always work and never to suspend or terminate the agent. This is also intuitive, because high revenue per arrival means that the principal does not want to suspend the agent and forfeit the higher arrival rate. The third possibility is when neither the switching cost nor the revenue is too high. In this case, the optimal contract demonstrates intricate and rich structures.

Specifically, the optimal contract demonstrates a "control-band" policy structure, characterized by a *lower threshold* $\underline{\theta}$ and an *upper threshold* $\bar{\theta}$ of the agent's total future utility, also called the *promised utility* (Spear and Srivastava 1987, Abreu et al. 1990). As illustrated in Figure 1 later in the paper, when directed to work, the agent should continue exerting effort as long as the promised utility is above the lower threshold $\underline{\theta}$. While working, the promised utility takes a fixed upward jump upon each arrival, and continuously decreases between arrivals. If an arrival does not occur for too long a period of time (bad performance) despite the agent's effort, the promised utility decreases to the lower threshold $\underline{\theta}$. At this point, the principal suspends the agent for a *fixed period of time*. At the end of the suspension period, the agent's promised utility is reset to the upper threshold $\bar{\theta}$, when work is switched on again. Overall, Figure 1 illustrates the general promised utility dynamics, which also involve an *upper bound* $\widehat{w}$ and a *lower bound* $\breve{w}$ for the promised utility, related to payments and potential random switching. In Section 4.3, we provide a complete characterization of the contract parameters $\underline{\theta}$, $\bar{\theta}$, $\widehat{w}$, and $\breve{w}$. Furthermore, Section EC.1.2 explains how to easily compute these four contract parameters.

1.2. Contribution

Our paper makes the following contributions. First, we propose that temporary suspension, often used in practice for various purposes, can be used in dynamic contracts for incentive management. In particular, the principal can use suspension as a threat when performance has been undesirable, which motivates the agent to work hard. Allowing temporary suspension helps the principal to reduce contract costs, compared with always incentivizing the agent to work, or using contract termination as a threat. A numerical example reported in Section 5.1 shows that this benefit can be non-trivial.

Second, our results show that the optimal dynamic contract takes a simple form, which makes it easy to implement in practice. The four parameters that characterize the contract sturcture are easy to compute. Using these four parameters, the principal can easily manage the contract over time. As explained in Remark 2 later in the paper, implementation of this dynamic contract is quite simple. Following each arrival as well as after each suspension period ends, the principal only needs to announce a deadline before which the agent needs to bring in an arrival to prevent suspension. If suspension happens, it always lasts for a fixed period of time.

Designing optimal contracts that can endogenously decide temporary suspensions is a technically hard problem, which explains why many dynamic contract design models generally search for contracts that always induce effort from the agent (see, for example, Demarzo and Sannikov 2006, Biais et al. 2010).[2] Zhu (2013) and Grochulski and Zhang (2023) are exceptions, and study contract design allowing shirking. They focus on settings where uncertain outcomes follow a Brownian motion, instead of a jump process as in our paper. Their optimal contract structures involve controls that constantly switch between working and shirking (a "sticky process"). Although these are nice mathematical results, from a managerial point of view, such a control/contract is not practical, because constantly switching between working and shirking must be quite costly in real life. Therefore, we include a fixed cost whenever the principal switches the agent back to work from suspension. The switching cost poses significant analytical challenges. Specifically, in addition to the promised utility, we have to introduce another binary system state that indicates whether the agent has been working or under suspension. This further implies that different from the aforementioned papers, we need to work with two value functions connected through the optimality condition.

Compared with other papers that study dynamic contracts under Poisson arrivals, such as Sun and Tian (2018), Cao et al. (2022), and Tian et al. (2021), our analysis relies on a set of quasi-variational inequalities to represent the optimality conditions. The other papers, on the other hand, directly verify optimality based on proposed optimal contract structures and the corresponding value functions, without the intricate approach based on quasi-variational inequalities. All these papers follow the "guess-and-verify approach", which is logically clean and clear. However, the "verify" step in our proof is much more intricate compared with that in the aforementioned papers, even with the correct "guess" of how to construct the value functions. Overall, our analytical approach could be applicable to other complex dynamic contracting problems that involve state transitions.

1.3. Literature Review

The dynamic moral hazard problem has been a subject of recent management science studies. In particular, Zorc et al. (2019) study a delegated search problem in a discrete-time dynamic environment. A key distinction of that paper is that the agent is risk averse and can borrow from a bank to pay the principal. In comparison, we assume that the risk-neutral agent is cash constrained and therefore, payment only goes from the principal to the agent. Gupta et al. (2022) study "limited-term" non-monetary reward contracts in order to induce agents' effort in the long run. Their model focuses on designing near-optimal "limited-term" stationary policies.

Recent decision analysis literature also includes studies of continuous-time games. The stream of papers Kwon et al. (2016), Kwon (2022), and Georgiadis et al. (2022) study continuous-time stochastic games of stopping-time decisions that are based on Brownian motion uncertainties. Continuous-time games studied in Zorc and Tsetlin (2020) and Hu and Tang (2021) do not include Brownian motion uncertainties, but rather consider richer decision spaces for the players. Unlike our paper, these game-theoretic papers do not focus on dynamic moral hazard issues.

Methodological breakthroughs for continuous-time moral hazard problems started from Demarzo and Sannikov (2006) in the finance/economics literature. Earlier studies often used Brownian motion processes to model the underlying dynamics (see, for example, Demarzo and Sannikov 2006, Sannikov 2008, Cvitanić et al. 2016). To our knowledge, the work of Biais et al. (2010) is the first to model underlying uncertainties as a jump process to capture "large risks." Myerson (2015) studies a similar model in a political-economy setting with agent replacement. Contracts in both these papers try to reduce the arrival rate, instead of increasing the arrival rate as in our paper, and do not consider temporary suspension.

Compared with models based on Brownian motion, the optimal contract structure for jump processes is much easier to describe and implement. This is because the promised utility often takes discrete jumps at arrivals, and otherwise changes deterministically. (In contrast, under Brownian motion uncertainties, the promised utility evolves stochastically all the time.) This simplicity in the optimal contract structure makes the model based on jump processes appealing from the practical and managerial perspectives. Sun and Tian (2018) and Cao et al. (2022) study optimal contracts that induce effort from an agent to increase the unobservable arrival rate of a point process. In particular, Cao et al. (2022) correctly identify the optimal contract within the restrictive class of contracts that motivate continued effort before termination when the two players' discount rates are different. Tian et al. (2021) further extend the model to a two-state setting, where the agent exerts effort to either maintain or repair a machine, depending on which state the machine is in. All these papers consider contracts that always induce effort before contract termination, without temporary suspension.

Also focusing on a point process, but to decrease the arrival rate, Chen et al. (2020) study optimal schedules to monitor (as well as pay) the agent. The end of that paper points out a connection between monitoring and shirking. That is, with a proper transformation, monitoring episodes in their optimal schedule correspond to shirking episodes in a corresponding model (without monitoring) that allows shirking. We believe that our results also speak to optimal contracts with monitoring for the case of increasing the arrival rate. In comparison, the optimal contracts in our paper demonstrate very different structures compared with those in Chen et al. (2020). We also need to model a fixed cost to be practical, as mentioned earlier. Tackling our problem requires different analysis, for example, using quasi-variational-inequality-based optimality condition, which does not arise in Chen et al. (2020). Also trying to decrease the arrival rate of a Poisson process in a bank-monitoring setting, Hernandez Santibanez et al. (2020) extend Pages (2013) and Pages and Possamai (2014), and study a model that involves both adverse selection and moral hazard while allowing shirking.

Another relevant literature is stochastic optimal control in the presence of switching cost, but not about contract design (see, for example, Brekke and Oksendal 1994, Duckworth and Zervos 2001, Vath and Pham 2007, Vath et al. 2008). Our work has two main differences from this literature. First, we consider a jump process, while the aforementioned papers are all based on diffusion processes. Second, the strategic interactions between the two players make our design and analysis more challenging than standard single-decision-maker control problems.

The remainder of the paper is organized as follows. We first introduce the model in Section 2. We then describe the optimal contract structure and the overall results of the paper in Section 3. Section 4 contains detailed analysis of the optimal contract structures under different model parameters and how to prove their optimality. Next, in Section 5, we consider several extensions of the model. In particular, in Section 5.1, we let the switching cost approach zero, which allows us to quantify the potential benefit of considering the suspension option. We conclude the paper in Section 6. Further discussions, as well as proofs for all the results, are presented in the e-companion.

2. Model

Consider a continuous-time principal-agent model. The principal faces a Poisson process of arrivals, each of which brings a revenue R to the principal. Without the agent's effort, the base arrival rate is $\underline{\mu}$. The agent is able to bring the arrival rate up to $\mu > \underline{\mu}$ if exerting effort, which costs the agent b per unit of time. (For simplicity, we consider binary effort levels, consistent with Biais et al. 2010). In order to enjoy the high arrival rate, the principal needs to direct the agent to work and provide the working environment, such as offering office spaces, research labs, production equipment, or supporting personnel. There is a fixed cost $K > 0$ for the principal to set up the environment for the agent to start working, either at the very beginning, or after a suspension period ends. Think about this as the fixed cost related to restarting the lease for office spaces, reopening the lab, resetting production equipment, or recruiting personnel. (We will briefly discuss the case when stopping working also incurs a cost in Section 5.2.) Following standard assumptions, the agent has limited liability and is cash constrained. Therefore, the principal needs to pay the agent a cost b whenever directing the agent to work. While directed to work, the agent can either exert effort, or shirk. Effort is not observable to the principal, who needs to design a contract to motivate the agent's effort. If the agent is directed to work but shirks, the agent effectively receives a shirking benefit b. Directing the agent to work may also involve additional costs to the principal, such as rents, maintenance fees, or personnel salaries. We denote c to represent the principal's total cost rate whenever directing the agent to work, including the payment for the agent's effort cost, that is, $c \geqslant b$.

Let $\mathcal{E}_t \in \{\mathsf{I}, \emptyset\}$ denote the working/suspension state at time $t \geqslant 0$. In particular, state I ("on") represents that the agent is directed to work, while state $\emptyset$ ("off") means that the agent is being suspended. We assume that the principal needs to pay the fixed cost K in the very beginning in order to hire the agent to start working. If the agent is directed to work at time t ($\mathcal{E}_t = \mathsf{I}$), further denote ν_t to represent the agent's effort level, such that $\nu_t = \mu$ and $\nu_t = \underline{\mu}$ represent that the agent exerts effort or not, respectively, at time t.

We define $\Delta\mu = \mu - \underline{\mu}$ and make the following assumption.

Assumption 1. *$R\Delta\mu > c$.*

This is a standard assumption (see, for example, Equation (2) in Sun and Tian 2018), which ensures that exerting effort is socially optimal when the state is I.

Both the principal and the agent are risk neutral and discount future cash flows. Discount rates are r and ρ for the principal and the agent, respectively, such that $0 < r \leqslant \rho$. That is, the principal is at least as patient as the agent. This paper is mostly focused on the case of $r < \rho$. (In Section EC.1.3, we provide a rigorous proof for the claim in Sun and Tian (2018) that when $r = \rho$, it is optimal to motivate continued effort.)

Denote right-continuous point processes $N := \{N_t\}_{t\geqslant 0}$ and $S := \{S_t\}_{t\geqslant 0}$ to record the total number of arrivals and switchings, respectively, from time 0 to t. Define a filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\geqslant 0}$ to capture all relevant public information up to any time t, such that $\mathcal{F}_t = \sigma(N_s, S_s : 0 \leqslant s \leqslant t)$. We need to include state-switching information in the filtration because some of this is random, as we explain next.

The principal has the commitment power to issue a long-term *dynamic contract* $\Gamma = (L, D, q)$. Specifically, the contract needs to specify a payment schedule, which is captured by the cumulative payment process L. It also needs to specify when to start and stop suspension. We use a counting process D to capture "deterministic" switchings between the working and suspension states. Finally, we also need to include a random switching intensity process q. These notations are formally defined as the following.

1. $L = \{L_t\}_{t\geqslant 0}$ is an $\mathcal{F}$-adapted process that tracks the principal's total payment to the agent from time 0 to time t. In particular, at any time t, the payment can be an instantaneous payment ΔL_t, or a flow with rate ℓ_t, such that $\mathrm{d}L_t = \Delta L_t + \ell_t \mathrm{d}t$.[3] Note that it is assumed that the agent is cash constrained and has limited liability, that is, $\Delta L_t \geqslant 0$ and $\ell_t \geqslant 0$ for all $t \geqslant 0$.
2. $D = \{D_t\}_{t\geqslant 0}$ is an $\mathcal{F}$-adapted counting process that records the total number of switchings between "working" and "suspension" up to time t. That is, these switchings are "*deterministic*" with respect to $\mathcal{F}_t$. In order to have a rich enough class of control policies such that optimal contract value is attainable, we also need to allow random switchings as well, which come next.
3. $q = \{q_t\}_{t\geqslant 0}$ is an $\mathcal{F}$-predictable switching intensity process, such that the probability of switching during a short time interval $[t, t+\delta]$ is $q_t\delta + o(\delta)$. In order to establish our optimality results, we need the following technical condition on the switching intensity:

$$\mathbb{E}\left[\int_0^\infty q_t e^{-rt}\mathrm{d}t\right] < \infty. \tag{1}$$

Let $Q = \{Q_t\}_{t \geqslant 0}$ be the corresponding counting process that records the cumulative number of all the random switchings up to time t. Therefore, the total number of switchings by time t is $S_t = D_t + Q_t$. Knowing that the system starts with state $\emptyset$, the total number of switchings S_t identifies the state at any time $t \geqslant 0$. These notations allows us to more rigorously define the $\mathcal{F}$-adapted state process $\mathcal{E} := \{\mathcal{E}_t\}_{t \geqslant 0}$, such that $\mathcal{E}_t \neq \mathcal{E}_{t-}$ if and only if $\mathrm{d}S_t = 1$. Furthermore, state switchings may also include the possible termination of contract, which is the last time that the principal changes the state from working (I) to (permanent) suspension ($\emptyset$), either deterministically or randomly.

Note that only relying on history-dependent (i.e., $\mathcal{F}$-adapted) switching control D is not sufficient. Generally speaking, a restrictive class of control policies without random switching may not contain the optimal one that achieves the optimal contract value. Random switching according to the intensity process q does occur in the optimal policy that we present later in the paper.

Due to limited liability, we need the following constraint for our contract Γ, which states that effort cost b needs to be reimbursed in real time:

$$\ell_t \geqslant b \mathbb{1}_{\mathcal{E}_t = \mathsf{I}}, \ \forall t \geqslant 0. \tag{LL}$$

Further denote a right-continuous process $\nu = \{\nu_t\}_{t \geqslant 0}$ to represent the agent's effort level over time. Under a general contract, the agent may not follow the effort process directed by the principal. In fact, it is easy to make sure that the agent follows the direction to stop working under suspension, by setting $\ell_t = 0$ when $\nu_t = \underline{\mu}$. In this case, the agent cannot afford to work when directed not to. Therefore, any effort process ν that is *admissible to contract* Γ must satisfy $\nu_t = \underline{\mu}$ whenever $\mathcal{E}_t = \emptyset$.

2.1. The Agent's Utility and Incentive-Compatible Contracts

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

$$u(\Gamma, \nu) = \mathbb{E}^{\nu, q} \left[\int_0^\infty e^{-\rho t} (\mathrm{d}L_t - b \mathbb{1}_{\nu_t = \mu} \mathrm{d}t) \right], \tag{2}$$

in which $\mathbb{E}^{\nu,q}$ represents expectation taken with respect to the switching intensity process q in Γ, and arrival rates induced by the effort process ν. For simplicity of notations, when there is no ambiguity, we omit this superscript.

A designed contract needs to induce the agent to follow directions on when to work. Formally, define a "complying effort process" $\bar{\nu}(\Gamma) = \{\bar{\nu}_t\}_{t \geqslant 0}$ for contract Γ, such that $\bar{\nu}_t = \mu$ if $\mathcal{E}_t = \mathsf{I}$, and $\bar{\nu}_t = \underline{\mu}$ if $\mathcal{E}_t = \emptyset$, at any time t. A contract Γ is said to be *incentive compatible* (IC) if

$$u(\Gamma, \bar{\nu}(\Gamma)) \geqslant u(\Gamma, \nu) \text{ for any effort process } \nu \text{ admissible to } \Gamma. \tag{3}$$

That is, under IC contracts, the agent has the incentive to exert effort whenever directed to do so.

Further define the agent's continuation utility at any time $t\in[0,\infty)$ conditional on $\mathcal{F}_t$ as[4]

$$W_t(\Gamma,\nu)=\mathbb{E}\left[\int_{t+}^{\infty}e^{-\rho(s-t)}(\mathrm{d}L_s-b\mathbb{1}_{\nu_s=\mu}\mathrm{d}s)\,\middle|\,\mathcal{F}_t\right]. \tag{4}$$

Therefore, $W_t\big(\Gamma,\bar{\nu}(\Gamma)\big)$ is the agent's continuation utility at time t following the principal's directions, which is often referred to as the *promised utility* (see, for example, Biais et al. 2010). It is convenient to introduce the notation $W_{t-}(\Gamma,\nu)=\lim_{s\uparrow t}W_s(\Gamma,\nu)$. That is, $W_t(\Gamma,\nu)$ is the agent's continuation utility after observing either an arrival or a random switching that occurs at time t, while $W_{t-}(\Gamma,\nu)$ is the continuation utility evaluated before obtaining this knowledge. In a similar vein, we define $W_{0-}(\Gamma,\nu):=u(\Gamma,\nu)$.

Following standard contract theory assumptions, the agent is not required to stay in the contract. Hence, assuming the agent's outside option is normalized to value 0, we impose the following participation (also called the *individual rationality*, IR) constraint:

$$W_t(\Gamma,\nu)\geqslant 0,\ \forall t\geqslant 0. \tag{IR}$$

Furthermore, we assume that for any contract Γ under our consideration, the agent's promised utility W_t is upper bounded. That is, there exists a large enough $\bar{W}$ such that

$$W_t(\Gamma,\nu)\leqslant \bar{W}<\infty,\ \forall t\geqslant 0. \tag{WU}$$

This constraint essentially captures the reality that the principal cannot keep delaying payments while pushing the agent's promised utility to infinity.[5]

The following proposition provides the evolution of the agent's continuation utility process $W_t(\Gamma,\nu)$, which is often called the *promise keeping* (PK) condition in the dynamic contract literature (see, for example, Equation (B.8) of Sun and Tian 2018). The proposition also contains an equivalent recursive representation of incentive compatibility, following Biais et al. (2010).

PROPOSITION 1. (i) *For any contract Γ and agent's effort process ν, there exist $\mathcal{F}$-predictable processes $H(\Gamma,\nu)$ and $H^q(\Gamma,\nu)$ such that*[6]

$$\begin{aligned}\mathrm{d}W_t(\Gamma,\nu)=&[\rho W_{t-}(\Gamma,\nu)+b\mathbb{1}_{\nu_t=\mu}-H_t(\Gamma,\nu)\nu_t+q_tH_t^q(\Gamma,\nu)]\mathrm{d}t\\&-\mathrm{d}L_t+H_t(\Gamma,\nu)\mathrm{d}N_t-H_t^q(\Gamma,\nu)\mathrm{d}Q_t,\quad t\geqslant 0.\end{aligned} \tag{PK}$$

Furthermore, (IR) *implies that*

$$H_t(\Gamma,\nu)\geqslant -W_{t-}(\Gamma,\nu)\ \text{and}\ H_t^q(\Gamma,\nu)\leqslant W_{t-}(\Gamma,\nu),\ \forall t\geqslant 0. \tag{5}$$

(ii) *Define $\beta:=b/\Delta\mu$. Contract Γ being incentive compatible is equivalent to*

$$H_t\big(\Gamma,\bar{\nu}(\Gamma)\big)\geqslant\beta\ \text{if}\ \mathcal{E}_t=\mathsf{I}. \tag{IC}$$

For notational convenience, we omit (Γ, ν) from processes W_t, H_t, and H_t^q when ν is the complying effort process $\bar{\nu}(\Gamma)$. Part (i) of Proposition 1 specifies the dynamics of the agent's promised utility over time. In particular, H_t, if positive, is the magnitude of an *upward* jump at time t if there is an arrival at that time. If it is negative, then the jump is downward. In contrast, H_t^q, if positive, is the magnitude of a *downward* jump at time t if there is a random switching. Condition (5) ensures that W_t remains nonnegative after all these jumps. The reason why we set H_t to capture upward jumps is that increasing the promised utility with an upward jump after an arrival serves as a reward to induce effort. Although we allow H_t to be negative, later we show that it is always nonnegative under the optimal contract. In comparison, a random switching to suspension (or termination) is a punishment, which is associated with a downward jump of promised utility with magnitude H_t^q. Finally, the (IC) condition is only required for state I, because in state $\emptyset$ the principal can induce compliance by simply setting payment to zero.

Denote $\mathfrak{C}$ to represent the set of contracts that satisfy (LL) and yield a promised utility process $\{W_t\}_{t\geqslant 0}$ that satisfies (PK), (IC), (IR), and (WU). Our contract design problem maximizes the principal's utility over the set $\mathfrak{C}$ of contracts. Therefore, we introduce the principal's utility next.

2.2. Principal's Utility

The principal's utility under any contract $\Gamma \in \mathfrak{C}$ is

$$U(\Gamma) = \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\infty e^{-rt}\left[R\mathrm{d}N_t - \mathrm{d}L_t - (c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\mathrm{d}t\right] - \sum_{0\leqslant t\leqslant\infty} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t)\right], \tag{6}$$

where we introduce notation $\kappa(\mathcal{E}_{t-}, \mathcal{E}_t)$ to represent the switching cost when the principal changes the working/suspension state from $\mathcal{E}_{t-}$ to $\mathcal{E}_t$, such that $\kappa(\emptyset, \mathsf{I}) = K$, and $\kappa(\mathsf{I}, \emptyset) = \kappa(\emptyset, \emptyset) = \kappa(\mathsf{I}, \mathsf{I}) = 0$. Within the integral, the term $R\mathrm{d}N_t$ represents the revenue from arrivals; $\mathrm{d}L_t$ is the payment cost, which satisfies (LL); and $(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}$ captures the cost rate of directing the agent to work, in addition to reimbursing the effort cost b already included in the payment term $\mathrm{d}L_t$.

Our optimal contract design problem can be succinctly formulated as the following optimization problem:

$$\mathcal{Z} := \max_{\Gamma\in\mathfrak{C}} \ U(\Gamma). \tag{7}$$

If the agent is ever terminated, the principal's total expected utility after the termination is

$$\underline{v} := \frac{\mu R}{r}, \tag{8}$$

which is also the baseline total expected revenue that the principal collects without hiring the agent.

We first consider a special contract $\bar{\Gamma}$ as an example of feasible contracts in $\mathfrak{C}$, which directs the agent to always work and pays the agent β for each arrival. It is clear that contract $\bar{\Gamma}$ satisfies all the aforementioned constraints for $\mathfrak{C}$. Under such a contract, the agent's promised utility W_t stays as a constant

$$\bar{w} := \frac{\beta\mu}{\rho}. \tag{9}$$

Furthermore, it is easy to verify that the principal's utility under contract $\bar{\Gamma}$ is

$$U(\bar{\Gamma}) = \frac{(R-\beta)\mu - c}{r} - K.$$

We define the corresponding *societal utility*, which is the total utilities of the principal and the agent, after paying for the fixed cost K, as

$$\bar{V} := U(\bar{\Gamma}) + K + \bar{w} = \frac{R\mu - c - (\rho - r)\bar{w}}{r}. \tag{10}$$

Later in the paper, we show that contract $\bar{\Gamma}$ is optimal when the revenue per arrival R is high enough.

3. Main Results

In this section, we summarize the main results of this paper and leave the analysis to the next section. First, we introduce a general contract structure in Section 3.1. Later in the paper, we show that the optimal contracts demonstrate this structure. Then, in Section 3.2, we present the optimality condition, and the main theorem of this paper, which summarizes the optimal contract for different model parameter settings.

3.1. An Overview of General Optimal Contract Structures

We now define a general class of dynamic contracts, which involves a control-band structure. Specifically, we have the following definition.

Definition 1. For any four parameters $\underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}$, such that $0 \leqslant \underline{\theta} \leqslant \breve{w} \leqslant \bar{\theta} \leqslant \widehat{w} \leqslant \bar{w}$, and an initial promised utility $w_0 \in \{0\} \cup [\breve{w}, \widehat{w}]$,[7] define contract $\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}) = (L^*, D^*, q^*)$ as follows.

(i) The dynamics of the agent's promised utility W_t follow $W_0 = w_0$ and

$$\begin{aligned} \mathrm{d}W_t = &\left\{-\rho\left(\bar{w} - W_{t-}\right)\mathrm{d}t \mathbb{1}_{W_{t-}\in(\breve{w},\widehat{w}]} - (\breve{w} - \underline{\theta})\,\mathrm{d}Q_t + \left[(\widehat{w} - W_{t-}) \wedge \beta\right]\mathrm{d}N_t\right\}\mathbb{1}_{\mathcal{E}_{t-}=\mathsf{I}} \\ &+ \rho W_{t-}\mathrm{d}t \mathbb{1}_{\mathcal{E}_{t-}=\emptyset}, \end{aligned} \tag{11}$$

in which we use notation $a \wedge b$ to represent $\min\{a, b\}$ for any $a, b \in \mathbb{R}$, and the point process $\{Q_t\}_{t\geqslant 0}$ represents the total number of random switchings up to time t, following intensity

$$q_t^* = \frac{\rho(\bar{w} - \breve{w})}{\breve{w} - \underline{\theta}} \mathbb{1}_{W_{t-}=\breve{w}}, \text{ if } \breve{w} > \underline{\theta}. \tag{12}$$

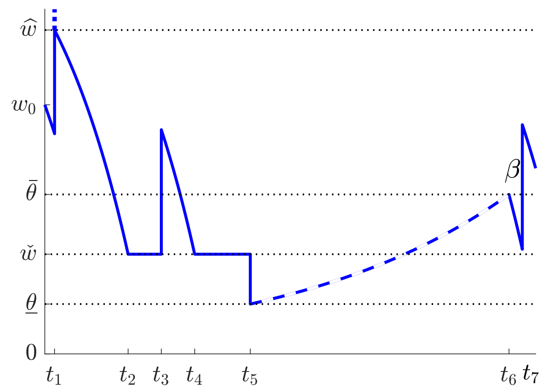
(ii) The payment to the agent follows $\mathrm{d}L_0^* = 0$ and

$$\mathrm{d}L_t^* = \left[(W_{t-} + \beta - \widehat{w})^+ \mathrm{d}N_t + b\mathrm{d}t\right]\mathbb{1}_{\mathcal{E}_{t-}=\mathsf{I}}. \tag{13}$$

(iii) The "deterministic" switching D^* follows $\mathrm{d}D_0^* = 1$ (start working) if and only if $w_0 > 0$, and for $t > 0$,

$$\mathrm{d}D_t^* = \mathbb{1}_{W_{t-}=\bar{\theta},\mathcal{E}_{t-}=\emptyset} + \mathbb{1}_{W_{t-}=\underline{\theta},\mathcal{E}_{t-}=\mathsf{I}}. \tag{14}$$

Figure 1 A Sample Trajectory for the Agent's Promised Utility Following a General Contract of Definition 1



Notes. In this figure, $\rho = 0.5$, $r = 0.2$, $R = 2$, $\mu = 2$, $\Delta\mu = 1.2$, and $c = b = 0.3$, where we set $w_0 = 0.5$, $\underline{\theta} = 0.1$, $\breve{w} = 0.2$, $\bar{\theta} = 0.32$, and $\widehat{w} = 0.65$. The dotted vertical line at time t_1 depicts the payment.

Figure 1 depicts the dynamics of the promised utility following a general contract of Definition 1. As we can see, the agent is working at time 0, and the promised utility starts at w_0 and gradually decreases until the first arrival at time t_1. At this point in time, an upward jump of β would take the promised utility above $\widehat{w}$. Therefore, the promised utility instead jumps to $\widehat{w}$, and the principal pays the agent $W_{t_1-} + \beta - \widehat{w}$ (depicted by the dotted line at t_1). No further arrival occurs until time t_2, when the promised utility reaches $\breve{w}$. From this point on, the promised utility stays the same at $\breve{w}$, while a random switching occurs with rate $q^* = \rho(\bar{w} - \breve{w})/\breve{w}$. At time t_3, there is an arrival before a random switching occurs, which causes the promised utility to jump up by β. The promised utility decreases to $\breve{w}$ again at t_4, and the random switching occurs at time t_5, which brings the promised utility to $\underline{\theta}$. At this point the agent is suspended, until the promised utility increases to $\bar{\theta}$ at t_6. There may be arrivals between t_5 and t_6 if $\underline{\mu} > 0$, but these arrivals do not affect the dynamics of the contract. At time t_6, the principal (deterministically) switches the agent to working again. Time epoch t_7 sees another arrival, which triggers the promised utility to jump up by β.

More generally, according to Definition 1, all the components L^*, D^*, and q^* of the contract Γ^*, as well as the promised utility process $\{W_t\}_{t\geqslant 0}$, are completely determined by parameters $(\underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w})$ starting from w_0. In particular, (11) indicates that the promised utility generally decreases with a slope $\rho(\bar{w} - W_{t-})$ when $W_{t-} \in (\breve{w}, \widehat{w}]$ and the agent is directed to work, except when there is an arrival ($\mathrm{d}N_t = 1$), which triggers an upward jump of magnitude $(\widehat{w} - W_{t-}) \wedge \beta$. This implies that the promised utility is never above $\widehat{w}$. When the promised utility decreases to $\breve{w}$, it stays at that level until either an arrival ($\mathrm{d}N_t = 1$) or a random switching of state ($\mathrm{d}Q_t = 1$) occurs. According to (12), the random switching only occurs if $\breve{w} > \underline{\theta}$ and when the promised utility is at $\breve{w}$. When random switching happens, (11) further indicates that the promised utility takes a downward jump from $\breve{w}$ to $\underline{\theta}$. Furthermore, the last term of (11) indicates that when the agent is under suspension, the promised utility keeps increasing at rate ρW_{t-} regardless of whether there are arrivals. The increasing rate corresponds to accrued interests if we consider the promised utility as a bank account balance.

According to (13), payment only occurs when the agent is directed to work. Besides reimbursing the effort cost ($b\mathrm{d}t$), the principal only pays the agent when an arrival occurs and the current promised utility is within β below $\widehat{w}$. The instantaneous payment, $W_{t-} + \beta - \widehat{w}$, plus the corresponding upward jump in (11), $\widehat{w} - W_{t-}$, is exactly β.

Finally, (14) implies that the principal directs the agent to stop working when the promised utility decreases to $\underline{\theta}$ and to start working again when the promised utility increases to $\bar{\theta}$. Therefore, if $\underline{\theta} = \breve{w}$, there is no random switching, and the switching policy is similar to the traditional "control-band" policy between the two thresholds $\underline{\theta}$ and $\bar{\theta}$.

Remark 1. Note that suspending the agent serves as a type of punishment. Before the promised utility decreases to the threshold $\underline{\theta}$, any arrival brings an upward jump in the promised utility, which makes the agent closer to getting paid, if not already being paid. However, as soon as the agent is suspended, it takes a *fixed period of time* with length $(\ln\bar{\theta} - \ln\underline{\theta})/\rho$ for work to resume. Because the effort cost is reimbursed, from the agent's point of view, the only difference between working and suspension is that working brings potential rent payment, while suspension delays future rent payments for a period of time. Therefore, suspension serves as a threat to the agent, who is less patient than the principal ($\rho > r$). If the lower threshold $\underline{\theta}$ is 0, the length of the suspension period becomes infinity, that is, the suspension is permanent, which is equivalent to contract termination. □

Following Definition 1, if $\widehat{w} = \bar{w}$, upon reaching $\bar{w}$, the promised utility does not decrease any more, and the agent is paid β for each future arrival. (Figure 1, on the other hand, depicts the case that $\widehat{w} < \bar{w}$, and $\breve{w} > \underline{\theta} > 0$.) Therefore, after reaching $\bar{w}$ the contract becomes the contract $\bar{\Gamma}$ defined in the end of the last section. In fact, contract $\bar{\Gamma}$ can be expressed as a special case of $\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w})$, such that

$$\bar{\Gamma} = \Gamma^*(\bar{w}; 0, 0, \bar{w}, \bar{w}). \tag{15}$$

If $\underline{\theta} = \breve{w} > 0$, there is no random switching, and contract $\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w})$ demonstrates a "control-band" structure, where the promised utility is moving between $\underline{\theta}$ and $\bar{\theta}$, unless an arrival triggers an upward jump to carry the promised utility to $(\bar{\theta}, \widehat{w}]$. In this case, the agent is never terminated, as long as $w_0 > 0$.

If $\underline{\theta} = 0$, then following contract $\Gamma^*(w_0; 0, \breve{w}, \bar{\theta}, \widehat{w})$, whenever the state switches to $\emptyset$, the promised utility must have hit $\underline{\theta} = 0$. At this point, the contract is terminated.

Another special case is not to hire the agent from the beginning, or,

$$\underline{\Gamma} := \Gamma^*(0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}). \tag{16}$$

In this case, the agent's promised utility starts at $w_0 = 0$, and never climbs to be positive according to (11). Therefore, the specific values of $\underline{\theta}$, $\breve{w}$, $\bar{\theta}$, and $\widehat{w}$ do not matter.

Later in this section, we see that contracts $\bar{\Gamma}$ and $\underline{\Gamma}$, and other special cases of the general contract structure $\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w})$ could be optimal under different model parameter settings.

Before we close this section, we have the following result, which implies that if the contract Γ^* starts the continuation utility at w_0, then it delivers the agent the total utility w_0.

LEMMA 1. *For any $\underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}$ such that $0 \leqslant \underline{\theta} \leqslant \breve{w} \leqslant \bar{\theta} \leqslant \widehat{w} \leqslant \bar{w}$, we have*

$$u\left(\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}), \bar{\nu}(\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}))\right) = w_0, \ \forall w_0 \in \{0\} \cup [\breve{w}, \widehat{w}]. \tag{17}$$

In order to specify a particular contract, including the optimal one, we need to identify not only the four parameters $\underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}$, but also the initial promised utility w_0.

3.2. Optimal Contracts

In this section, we summarize the main result of the paper. In particular, we present optimal contract structures under different model parameter settings. We leave the detailed optimality analysis to the next section.

We first provide an optimality condition, in the form of quasi-variational inequalities, and show that any function that satisfies these conditions must yield an upper bound of the optimal value $\mathscr{Z}$ defined in (7). In order to prove that the dynamic contracts that we will specify later in this section are indeed optimal, we show that our dynamic contracts achieve the upper bound.

We claim that the optimal value function is concave, although it may not be differentiable on its entire domain. Therefore, denote $\mathbb{C}$ to be the set of all continuous concave functions defined on $\mathbb{R}_+$. It is worth noting that any continuous concave function is differentiable except on a countable set of points. If a function $f \in \mathbb{C}$ is not differentiable at point $w \geqslant 0$, we abuse notation and use $f'(w)$ to represent its left derivative at w. Define operators $\mathcal{A}_{\mathsf{I}}$ and $\mathcal{A}_{\emptyset}$ that map a function $f \in \mathbb{C}$ to functions $\mathcal{A}_{\mathsf{I}} f$ and $\mathcal{A}_{\emptyset} f$, respectively, such that for all $w \geqslant 0$,

$$(\mathcal{A}_{\mathsf{I}} f)(w) := (\mu + r) f(w) - \mu f(w + \beta) + \rho(\bar{w} - w) f'(w) - (\mu R - c) + (\rho - r) w, \text{ and} \tag{18}$$

$$(\mathcal{A}_{\emptyset} f)(w) := r f(w) - \rho w f'(w) + (\rho - r) w - R\underline{\mu}. \tag{19}$$

Equipped with these notations, we are ready to present the following Verification Theorem.[8]

THEOREM 1. *Suppose there exists a pair of nondecreasing functions V_{I} and $V_{\emptyset}$ in $\mathbb{C}$, such that*

$$(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0, \ (\mathcal{A}_{\emptyset} V_{\emptyset})(w) \geqslant 0, \tag{20}$$

$$0 \leqslant V_{\mathsf{I}}(w) - V_{\emptyset}(w) \leqslant K, \ \textit{and} \tag{21}$$

$$V_{\mathsf{I}}(0) \geqslant \underline{v}, \ V_{\emptyset}(0) \geqslant \underline{v}, \tag{22}$$

for any $w \in \mathbb{R}_+$. Then, for any contract $\Gamma \in \mathfrak{C}$ and value $w \in [0, \infty)$ such that $u(\Gamma, \bar{\nu}(\Gamma)) = w$, we have

$$U(\Gamma) \leqslant V_{\emptyset}(w) - w.$$

Therefore, we have

$$\max_{w \in [0,\infty)} \{V_{\emptyset}(w) - w\} \geqslant \mathcal{Z}.$$

Theorem 1 indicates that $V_{\emptyset}(w) - w$ is an upper bound for the principal's utility under any contract in $\mathfrak{C}$ that yields an agent's utility w. Therefore, we can interpret the function $V_{\emptyset}(w)$ as an upper bound for the societal value function that contains both the principal and the agent's utilities. In fact, function $V_{\mathsf{I}}(w)$ can also be perceived as an upper bound for the societal value function if the system starts from state I instead of $\emptyset$, as if the agent has been working when the contract starts. The quasi-variational-inequality-based optimality condition (20)–(22) may not appear intuitive. Therefore, in Section EC.1.1, we provide a heuristic derivation, which reveals how we obtain these conditions. In a nutshell, the condition (20) describes the shape of the value functions; the condition (21) reflects that switching from suspension to working costs K; and the condition (22) captures the intuition that without the agent (the agent's promised utility $w = 0$), the societal value is $\underline{v}$ as defined in (8).

If both functions V_{I} and $V_{\emptyset}$ are differentiable on $\mathbb{R}_+$, then this result is a classic verification theorem, which is extensively used in the optimal control literature. In fact, a typical method of obtaining an optimal control policy, which is called the "guess-and-verify" approach, consists of two main steps. In the first step, we guess a candidate control (contract) structure and its related value function. In the second step, we use the Verification Theorem to establish that this function indeed provides an upper bound of the optimal value function, and is achievable under the guessed contract. It is worth mentioning that the control space in our dynamic contract problem includes potentially randomized control, which is rich enough to achieve the corresponding upper bound. We now present our main result in the next theorem, and leave its justification to the next section of the paper.

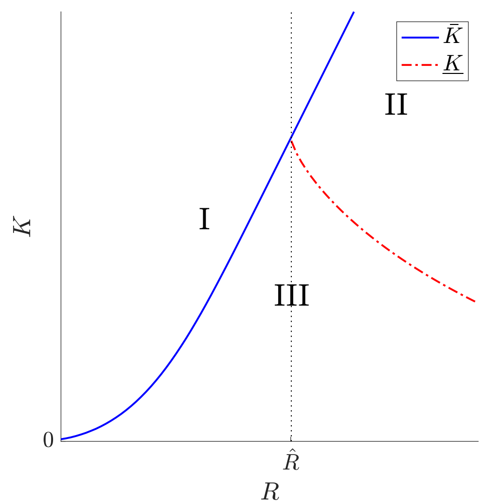
THEOREM 2. *Given model parameters r, ρ, b, c, μ, and $\underline{\mu}$, the optimal contract demonstrates the following three possible structures.*

(i) *If the fixed cost* $K > \bar{K}(R)$*, in which* $\bar{K}(R)$ *is an increasing function of* R *to be defined in Section 4.1, the optimal contract is* $\underline{\Gamma}$*. That is, it is optimal not to hire the agent at all.*

(ii) *If* $\rho - r > \mu$*, there exists a value* $\hat{R}$ *and a non-increasing function* $\underline{K}(R)$ *for* $R \geqslant \hat{R}$*, to be specified in Equations (28), (33), and (34) later in the paper, such that for* $R \geqslant \hat{R}$ *and* $K \in [\underline{K}(R), \bar{K}(R)]$*, contract* $\bar{\Gamma}$ *is optimal. That is, it is optimal to hire the agent and offer payment* β *to each arrival from the beginning.*

(iii) *If* K*,* R*,* ρ*,* r*, and* μ *do not satisfy either condition above, then there exist four parameters* $(\underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w})$ *such that* $0 < \underline{\theta} \leqslant \breve{w} < \bar{\theta} < \widehat{w} \leqslant \bar{w}$*, and an initial promised utility* w_0*, such that the contract* $\Gamma^*\left(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w}\right)$ *is optimal.*

Figure 2 Partition of the (R, K) Plane Based on Optimal Contract Structures



Note. In this figure, $r = 0.2$, $\rho = 1.5$, $c = b = 0.2$, $\Delta\mu = 0.7$, and $\mu = 1$.

Figure 2 demonstrates the optimal contract structures summarized in Theorem 2. First, if the switching cost is above $\bar{K}(R)$ (Region I), it is optimal for the principal not to hire the agent at all. This is intuitive because if the fixed cost is too high to set up the operation, it is not worth hiring the agent. Second, if the switching cost K is lower than the threshold $\bar{K}(R)$ but above $\underline{K}(R)$, then it is optimal for the principal to hire the agent and start paying β for each arrival. Note that this case can be equivalently expressed as that the revenue R is higher than a K-dependent threshold (Region II). This is also intuitive, because if the revenue R is high enough, it is not worth suspending or terminating the agent, which would forfeit the higher arrival rate to receive the revenue. Finally, the third case in Theorem 2 corresponds to the following possibilities: (i) $\rho - r < \mu$ and $K \leqslant \bar{K}(R)$; (ii) $\rho - r \geqslant \mu$, $R < \hat{R}$, and $K \leqslant \bar{K}(R)$; and (iii) $\rho - r \geqslant \mu$ and $K \leqslant \underline{K}(R)$. That is, both K and R are lower than their respective thresholds (Region III). In this case, the optimal contract takes the general form of $\Gamma^*\left(w_0;\underline{\theta},\breve{w},\bar{\theta},\widehat{w}\right)$ following Definition 1. Figure 2 depicts model parameters such that $\rho - r > \mu$. If $\rho - r \leqslant \mu$, on the other hand, $\underline{K}(R)$ is not well defined, and Region II no longer exists in the figure.

Remark 2. It is worth discussing how to implement the richest structured contract $\Gamma^*\left(w_0;\underline{\theta},\breve{w},\bar{\theta},\widehat{w}\right)$ in practice. Although the dynamic of the promised utility may look intricate, the contract implementation can be quite simple. At any point in time, the principal just needs to show when the agent needs to bring in another arrival in order to prevent suspension from kicking in. For example, let us first consider the situation that $\breve{w} = \underline{\theta}$, that is, there is no random starting of suspension. Following the dynamics (11), in particular, $\mathrm{d}W_t = -\rho(\bar{w} - W_{t-})\mathrm{d}t$, the promised utility evolves according to $W_s = \bar{w} - (\bar{w} - W_t)e^{\rho(s-t)}$ for $s > t$ without an arrival. Consequently, if the promised utility is W_t at time t, it takes a period of time with length

$$\mathbf{d}(W_t) := \frac{1}{\rho}\ln\left(\frac{\bar{w} - \breve{w}}{\bar{w} - W_t}\right), \tag{23}$$

for the promised utility to decrease to $\underline{\theta}$ without an arrival.

Therefore, in the very beginning and right after each arrival at time t while the agent has been working, the principal just announces a "deadline" at $\mathbf{d}(W_t) + t$, before which the agent needs to bring in an arrival to prevent suspension. This deadline does not change over time unless an arrival occurs. If an arrival does occur at time t before the deadline when the promised utility takes value W_{t-} right before the arrival, the suspension deadline is postponed to a new epoch following (23), in which W_t takes value $(W_{t-} + \beta) \wedge \widehat{w}$. (As mentioned before, if $W_{t-} + \beta > \widehat{w}$, the agent is paid $W_{t-} + \beta - \widehat{w}$ for this arrival.) In case there is no arrival before the deadline, the agent is suspended for a fixed period of time, as mentioned in Remark 1. Therefore, the contract can be implemented as a sequence of changing suspension deadlines announced after each arrival and at the end of each suspension episode.

If $\breve{w} > \underline{\theta}$, on the other hand, the principal also needs to randomize suspension. In this case, if there is no arrival after the deadline $\mathbf{d}(W_t) + t$, then the principal and agent can use some commonly observable randomization device (for example, the last two digits of a stock index) to implement the random start of suspension. □

Before we close this section, it is worth looking at an example that numerically demonstrates the optimal contract described in Theorem 2.

Example 1. Consider the real estate agency example mentioned in the introduction. On average, each rental sales representative brings in between 1 and 2 new tenants per month. Therefore we set $\mu = 1$ per month. Assume that without the representative's effort, no tenant arrives, or $\underline{\mu} = 0$ per month. Following our conversations with Hong Kong real estate agencies, it is reasonable to set $R = 20000$ Hong Kong dollars (HKDs), which is about one month's rent that the agency charges for an average apartment. Following a representative's base salary and sales bonus levels, we set $c = b = 4000$HKDs/month, implying $\beta = 4000$HKDs. We also set the time discount rates to be $r = 0.0088$ and $\rho = 0.0297$ per month, corresponding to annual discounts of 10% and 30% for the principal and the agent, respectively (Dohmen et al. 2012). Consider a very modest switching cost $K = 100$HKDs. Following Theorem 2, the optimal contract is $\Gamma^*(w_0; \underline{\theta}, \breve{w}, \bar{\theta}, \widehat{w})$, in which $w_0 = 39502.07$, $\underline{\theta} = 10990$, $\breve{w} = 10990$, $\bar{\theta} = 15506.26$, and $\widehat{w} = 51557.05$, all in terms of HKDs. That is, if the promised utility reaches $\widehat{w}$, the sales representative has 12.95 months to bring in a new customer to avoid suspension. Each suspension episode lasts 4.56 months.

In the next section, we provide the proof analysis for Theorem 2, relying on Theorem 1 and construction of value functions for different model parameter settings.

4. Analysis for Theorem 2

In this section, we present the main steps to prove Theorem 2 based on the "guess-and-verify" approach described in the last section. First, we follow Theorem 1 to study potential forms of the value functions. Many details of the proofs are presented in the e-companion.

In order to identify optimal value functions, it is worth considering functions that satisfy conditions (20) and (22) with equality. First, consider the suspension state $\emptyset$. A function $V(w)$ that satisfies $V(0) = \underline{v}$ and the ordinary differential equation $(\mathcal{A}_\emptyset V)(w) = 0$ must have the form

$$\mathsf{V}_c(w) = \underline{v} + w + c w^{r/\rho}, \tag{24}$$

for some constant c. Later in the paper, we show that the suspension state's value function indeed takes this form under certain model parameter settings and for certain w values.

Next, we consider the working state I. In particular, consider a generic function $V_{\widetilde{w}}$ that is differentiable on $[0, \widetilde{w}]$ for some $\widetilde{w} \leqslant \bar{w}$, takes a constant value for $w \geqslant \widetilde{w}$, and satisfies (20) with $(\mathcal{A}_{\mathsf{I}} V)(w) = 0$. That is, $V_{\widetilde{w}}$ satisfies the differential equation

$$(\mu + r)V_{\widetilde{w}}(w) - \mu V_{\widetilde{w}}\big((w+\beta) \wedge \widetilde{w}\big) + \rho(\bar{w} - w)V'_{\widetilde{w}}(w) - (\mu R - c) + (\rho - r)w = 0, \tag{25}$$

with boundary condition

$$V_{\widetilde{w}}(w) = \bar{V}(\widetilde{w}), \ \forall w \geqslant \widetilde{w}, \tag{26}$$

in which $\bar{V}(\cdot)$ is defined as

$$\bar{V}(w) := \frac{R\mu - c - (\rho - r)w}{r}, \tag{27}$$

such that $\bar{V}(\bar{w}) = \bar{V}$. Lemma EC.2 in Section EC.3.1 establishes the existence and uniqueness of function $V_{\widetilde{w}}(w)$, and summarizes its key properties.

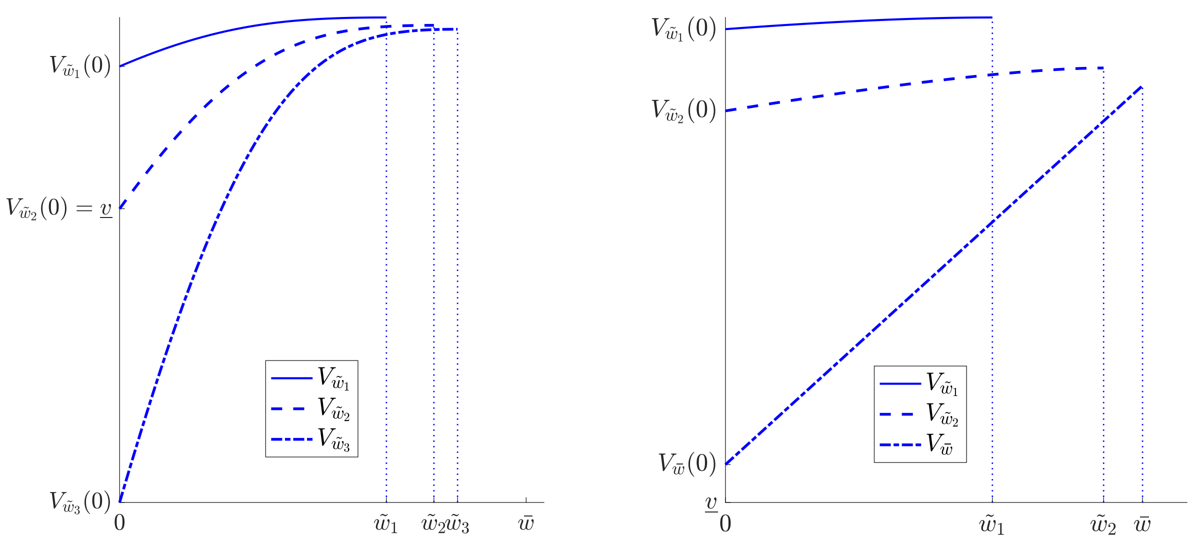
We also desire the boundary condition $V_{\widetilde{w}}(0) = \underline{v}$. Figure 4(a) demonstrates an example in which we can find a $\widetilde{w}$ value such that this boundary condition holds. As we can see, if we increase $\widetilde{w}$ to take three different values, $\tilde{w}_1$, $\tilde{w}_2$, and $\tilde{w}_3$, the entire function decreases as $\widetilde{w}$ increases, consistent with Lemma EC.2. This implies that at $w = 0$, we have $V_{\tilde{w}_1}(0) > V_{\tilde{w}_2}(0) > V_{\tilde{w}_3}(0)$. In particular, we can identify a particular $\widetilde{w} = \tilde{w}_2$ such that $V_{\tilde{w}_2}(0) = \underline{v}$. Following Lemma EC.2, this situation corresponds to model parameters that satisfy the following condition.

CONDITION 1.

$$\mu \geqslant \rho - r \text{ or } R < \hat{R} := \left[\frac{c}{b} + \frac{(\rho - r)\mu}{\Delta\mu(\rho - r - \mu)} \cdot \frac{\rho - \mu}{\rho}\right]\beta. \tag{28}$$

Condition 1 means that either the time discount (patience) levels between the principal and the agent are not too different (their difference is not more than the arrival rate under the agent's effort) or the revenue per arrival is low enough.

Figure 3 Solutions to Differential Equation (25) as $\widetilde{w}$ Varies



(a) Value Functions under Condition 1 (b) Value Functions under Condition 2

Notes. (i) For the left panel, $r = 0.2$, $\rho = 0.5$, $c = 0.2$, $R = 2$, $\Delta\mu = 0.7$, and $\mu = 2$. In this case, $\bar{w} = 1.14$, and we let $\tilde{w}_1 = 0.6$, $\tilde{w}_2 = 0.9$, and $\tilde{w}_3 = 1$. (ii) For the right panel, $r = 0.2$, $\rho = 1.2$, $c = b = 1$, $R = 5$, $\Delta\mu = 0.8$, and $\mu = 0.9$. In this case, $\bar{w} = 0.94$, and we let $\tilde{w}_1 = 0.6$, $\tilde{w}_2 = 0.85$, and $\tilde{w}_3 = \bar{w}$.

However, in general, we may not be able to find a value $\widetilde{w} \leqslant \bar{w}$ to satisfy the boundary condition $V_{\widetilde{w}}(0) = \underline{v}$. Figure 4(b), for example, depicts another model parameter setting such that as we increase $\widetilde{w}$ to approach $\bar{w}$, the corresponding limiting value $V_{\bar{w}}(0)$ is always higher than $\underline{v}$. In this case, we cannot use (25)–(26) to determine the optimal value function. This situation corresponds to model parameters that follow the next condition, opposite to Condition 1.[9]

Condition 2.

$$\mu < \rho - r \text{ and } R \geqslant \hat{R}.$$

Condition 2 states that the principal is much more patient than the agent (their discount rates different by more than the arrival rate under the agent's effort) and that the revenue for each arrival is high enough.

Generally speaking, the solution $V_{\tilde{w}}$ that satisfies (25) with boundary condition (26) may not be concave; see Cao et al. (2022) for such an example. Therefore, we may need to construct a concave value function according to the following result.

Lemma 2. *For any $\tilde{w} \in (0, \bar{w})$, consider the function $V_{\tilde{w}}$ that uniquely solves (25)–(26).*

(i) *There exists a $\tilde{w}$-dependent threshold $\breve{w}(\tilde{w}) \in [0, \tilde{w})$, such that $V''_{\tilde{w}}(w) < 0$ over $w \in (\breve{w}(\tilde{w}), \tilde{w})$ and $V''_{\tilde{w}}(w) > 0$ over $w \in [0, \breve{w}(\tilde{w}))$. Moreover, we have $\breve{w}(\tilde{w}) \leqslant (1 - r/\rho)\beta$.*

(ii) *Define function*

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

Function $\mathcal{V}_{\tilde{w}}(w)$ is increasing and concave in w on $[0, \tilde{w})$.

(iii) *Fixing any $\tilde{w}_1$ and $\tilde{w}_2$ with $0 < \tilde{w}_1 < \tilde{w}_2 < \bar{w}$, we have*

$$\mathcal{V}_{\tilde{w}_1}(w) > \mathcal{V}_{\tilde{w}_2}(w), \text{ and } \mathcal{V}'_{\tilde{w}_1}(w) < \mathcal{V}'_{\tilde{w}_2}(w), \forall w \in [0, \tilde{w}_1).$$

Therefore, if function $V_{\tilde{w}}$ is not concave, we construct a concave function $\mathcal{V}_{\tilde{w}}$ by attaching a linear piece on $\big[0, \breve{w}(\tilde{w})\big)$ to the concave part of function $V_{\tilde{w}}(w)$ for $w \geqslant \breve{w}(\tilde{w})$. This function is closely related to the optimal value function when the state is I, as we show in the next two subsections.

4.1. High Switching Cost K (Area I of Figure 2)

In order to properly define the threshold $\bar{K}(R)$ in Theorem 2, we need to consider Conditions 1 and 2 seperately. For notational brevity, we drop the dependency on R for $\bar{K}(R)$ and $\underline{K}(R)$ defined in Theorem 2 in this section.

4.1.1. Condition 1. Under Condition 1, the optimal value function for state I relies on the following result.

Lemma 3. *Under Condition 1, there exists a unique $\hat{\mathrm{w}}$ in $[0, \bar{w})$ such that $\mathcal{V}_{\hat{\mathrm{w}}}(0) = \underline{v}$, in which the concave function $\mathcal{V}_{\hat{\mathrm{w}}}$ is defined in Lemma 2 with $\hat{\mathrm{w}}$ replacing $\tilde{w}$. Furthermore, if $\breve{w}(\hat{\mathrm{w}}) > 0$, then $\mathcal{V}'_{\hat{\mathrm{w}}}(0) = \mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}})) > 1$.*

Lemma 3 allows us to uniquely identify an upper bound $\hat{\mathrm{w}}$ and a function $\mathcal{V}_{\hat{\mathrm{w}}}$, which is independent of the switching cost K, along with a lower bound $\breve{w}(\hat{\mathrm{w}})$, as defined in Lemma 2. If $\breve{w}(\hat{\mathrm{w}}) > 0$, the value function $\mathcal{V}_{\hat{\mathrm{w}}}$ is linear on the interval $[0, \breve{w}(\hat{\mathrm{w}}))$, which is associated with randomized control.

We first consider a very high switching cost.

Proposition 2. *Under Condition 1 and $K \geqslant \bar{V}(\widehat{\mathrm{w}}) - \underline{v}$, in which the upper bound $\widehat{\mathrm{w}}$ is defined according to Lemma 3 and function $\bar{V}(w)$ is defined in (27), functions*

$$V_{\mathsf{I}}(w) = \mathcal{V}_{\widehat{\mathrm{w}}}(w) \text{ and } V_{\emptyset}(w) = \underline{v}$$

satisfy the optimality condition (20)–(22). Furthermore, $U(\underline{\Gamma}) = \underline{v}$, in which contract $\underline{\Gamma}$ is defined in (16).

Following Theorem 1, we know that for any contract Γ such that the agent's utility $u(\Gamma) = w$, we have

$$U(\Gamma) \leqslant \underline{v} - w \leqslant \underline{v} = U\left(\underline{\Gamma}\right),$$

in which the last equality follows from Proposition 2. Therefore, contract $\underline{\Gamma}$ is optimal under Condition 1 and $K \geqslant \bar{V}(\widehat{\mathrm{w}}) - \underline{v}$.

Lemma 4. *Under Condition 1 and $K < \bar{V}(\widehat{\mathrm{w}}) - \underline{v}$, there exist K-dependent values $\bar{\theta}^K \in [\breve{w}(\widehat{\mathrm{w}}), \widehat{\mathrm{w}}]$ and $m^K \in [0, \mathcal{V}'_{\widehat{\mathrm{w}}}(0)]$ such that*

$$\mathcal{V}_{\widehat{\mathrm{w}}}(\bar{\theta}^K) = m^K \bar{\theta}^K + K + \underline{v}, \text{ and } \mathcal{V}'_{\widehat{\mathrm{w}}}(\bar{\theta}^K) = m^K. \tag{29}$$

Furthermore, we have that $\bar{\theta}^K$ is increasing in K, m^K is decreasing in K, and $\lim_{K \downarrow 0} \bar{\theta}^K = \breve{w}(\widehat{\mathrm{w}})$.

With the help of m^K, we can define the following bound for the switching cost K, which corresponds to the bound $\bar{K}$ in Theorem 2 when model parameters satisfy Condition 1:

$$\bar{K}_1 := \inf \left\{ K \in \left(0,\ \bar{V}(\widehat{\mathrm{w}}) - \underline{v}\right] \mid m^K < 1 \right\}. \tag{K1}$$

Geometrically, Lemma 4 implies that the line $m^K w + \underline{v}$ is tangent to the curve $\mathcal{V}_{\widehat{\mathrm{w}}}(w) - K$ at $w = \bar{\theta}^K$. Therefore, we define the following societal value function for state $\emptyset$,

$$V_{\emptyset}(w) = \begin{cases} m^K w + \underline{v}, & w \in \left[0, \bar{\theta}^K\right], \\ \mathcal{V}_{\widehat{\mathrm{w}}}(w) - K, & w \in \left[\bar{\theta}^K, \widehat{\mathrm{w}}\right], \end{cases} \tag{30}$$

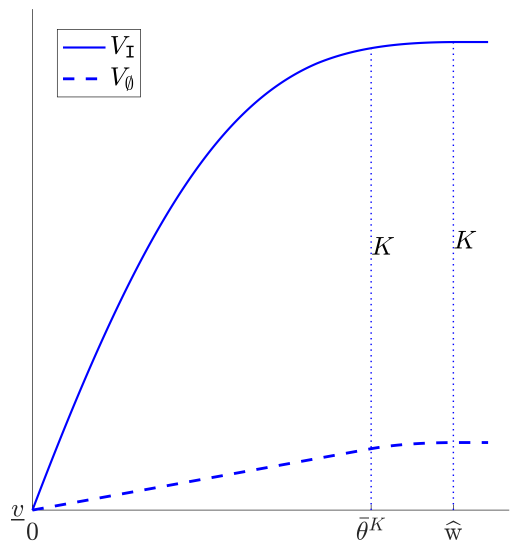
which is clearly concave, and linear on $\left[0, \bar{\theta}^K\right]$ with slope m^K. For any $K > \bar{K}_1$, the slope m^K is less than or equal to 1 according to (K1) and Lemma 4. Therefore, the corresponding principal's utility function, $V_{\emptyset}(w) - w$, is monotonically non-increasing and taking its maximum value at $V_{\emptyset}(0) = \underline{v} = U(\underline{\Gamma})$.

Similar to Proposition 2, we have the following result.

Proposition 3. *Under Condition 1 and $K \in [\bar{K}_1,\ \bar{V}(\widehat{\mathrm{w}}) - \underline{v}]$, functions $V_{\mathsf{I}}(w) = \mathcal{V}_{\widehat{\mathrm{w}}}(w)$ and $V_{\emptyset}(w)$ as defined in (30) satisfy the optimality condition (20)–(22).*

Therefore, the only difference between Propositions 2 and 3 is the value function for state $\emptyset$. The optimal contract is the same $\underline{\Gamma}$, which is not to hire the agent. Figure 4 shows an example of the societal value functions. It is clear that $V_\emptyset(w)$ is linear over the interval $[0, \bar{\theta}^K]$. Furthermore, functions $V_\mathsf{I}(w)$ and $V_\emptyset(w)$ are "parallel" with a difference of K for $w \geqslant \bar{\theta}^K$. At time t, if the state $\mathcal{E}_t = \emptyset$ and the promised utility $W_t > \bar{\theta}^K$ (which would never happen under the optimal contract), it is optimal to switch the agent to work, which explains the difference K between the value functions.

Figure 4 Illustration of Optimal Societal Value Functions for Area I of Figure 2



Notes. In this figure, $r = 0.2$, $\rho = 0.5$, $c = b = 0.2$, $R = 2$, $\Delta\mu = 0.7$, $K = 4$, and $\mu = 2$. Hence, $\bar{w} = 1.14$, $\mathcal{V}_{\widehat{\mathrm{w}}}(\widehat{\mathrm{w}}) = 17.67$, and $\underline{v} = 13$.

4.1.2. Condition 2.

Define a threshold for the switching cost

$$\bar{K}_2 := \bar{V} - \underline{v} - \bar{w}. \tag{K2}$$

Under Condition 2, function $\mathcal{V}_{\widehat{\mathrm{w}}}$ from Lemma 3 no longer exists, because the boundary condition $\mathcal{V}_{\widetilde{w}}(0) = \underline{v}$ does not hold for any $\widetilde{w} \in [0, \bar{w}]$. In this case, the principal needs to set the agent's promised utility at $\bar{w}$ or 0, but never in between. The corresponding value function is linear over the interval $[0, \bar{w}]$, connecting $\underline{v}$ at $w = 0$ and $\bar{V}$ at $w = \bar{w}$, where $\bar{V}$ is defined in (10).

PROPOSITION 4. *Suppose model parameters satisfy Condition 2. If $K > \bar{V} - \underline{v}$, functions*

$$V_\mathsf{I}(w) := \begin{cases} \underline{v} + \dfrac{\bar{V} - \underline{v}}{\bar{w}} \cdot w, & w \in [0, \bar{w}], \\ \bar{V}, & w \geqslant \bar{w}, \end{cases} \tag{31}$$

and $V_\emptyset(w) = \underline{v}$ *satisfy the optimality condition* (20)–(22)*; if* $\bar{K}_2 < K \leqslant \bar{V} - \underline{v}$*, on the other hand, functions* $V_|(w)$ *defined in* (31) *and*

$$V_\emptyset(w) := \begin{cases} \underline{v} + \dfrac{\bar{V} - \underline{v} - K}{\bar{w}} \cdot w, & w \in [0, \bar{w}], \\ \bar{V} - K, & w \geqslant \bar{w}. \end{cases} \tag{32}$$

satisfy (20)–(22).

Note that if $K > \bar{K}_2$, the slope $\dfrac{\bar{V} - \underline{v} - K}{\bar{w}} < 1$, which implies that the function $V_\emptyset(w) - w$ is monotonically decreasing. Therefore, for any contract Γ such that $u(\Gamma) = w$, we have

$$U\big(\underline{\Gamma}\big) = \underline{v} = V_\emptyset(0) \geqslant V_\emptyset(w) - w \geqslant U(\Gamma),$$

which implies the optimality of contract $\underline{\Gamma}$.

To summarize, define $\bar{K}$ to be $\bar{K}_1$ under Condition 1 and $\bar{K}_2$ under Condition 2. The following result corresponds to part (i) of Theorem 2.

Theorem 3. *If* $K \geqslant \bar{K}$*, it is optimal for the principal not to hire the agent.*

4.2. Medium Switching Cost K or High Revenue R (Area II of Figure 2)

We first provide the following expression for the threshold $\underline{K}(R)$, which is well defined under Condition 2:

$$\underline{K} = \underline{K}(R) := \frac{1}{r}\left\{R\Delta\mu - c - \beta\mu - \frac{\mu\rho}{\rho - r - \mu}\left[(R\Delta\mu - c)\,\frac{\rho - r - \mu}{\mu(\rho - r)} - \frac{r}{\rho}\beta\right]^{1 - r/\rho} \cdot \bar{w}^{r/\rho}\right\}\mathbb{1}_{R \in [\hat{R}, \bar{R}]} \tag{33}$$

$$\text{in which } \bar{R} := \left[\frac{c}{b} + \frac{(\rho - r)\mu}{\Delta\mu(\rho - r - \mu)}\right]\beta. \tag{34}$$

The following result helps us identify the value function $V_\emptyset$.

Lemma 5. *Under Condition 2 and* $K \in [\underline{K}, \bar{K}_2)$*, there exists a set of* K*-dependent parameters* $(c_K, m_K, \underline{\theta}_K)$ *with*

$$c_K > 0 \quad \textit{and} \quad m_K \in \left[\frac{\rho - r}{\rho - r - \mu}, \frac{\bar{V} - \underline{v}}{\bar{w}}\right), \tag{35}$$

such that

$$\mathsf{V}_{c_K}(\underline{\theta}_K) = \bar{V} + m_K(w - \bar{w}), \tag{36}$$

$$\mathsf{V}'_{c_K}(\underline{\theta}_K) = m_K, \ \textit{and} \tag{37}$$

$$\mathsf{V}_{c_K}(\bar{w}) = \bar{V} - K, \tag{38}$$

in which V_{c_K} *is defined in* (24) *with* c_K *replacing* c*. Furthermore, we have that* c_K *is decreasing in* K*,* m_K *is increasing in* K*, and* $\underline{\theta}_K$ *is decreasing in* K*.*

We first define function $V_\emptyset(w)$ by extending function V_{c_K} to include $w > \bar{w}$:

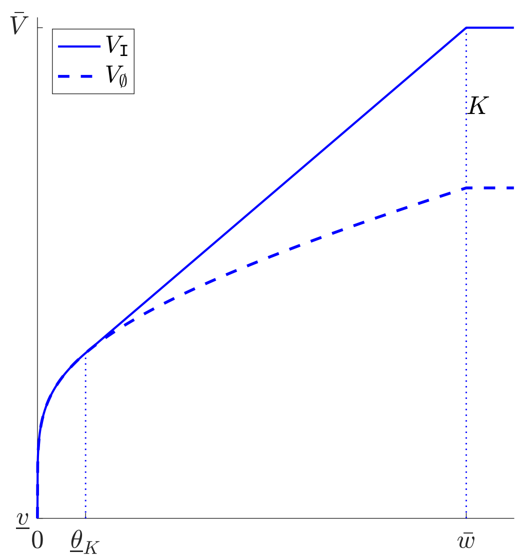
$$V_\emptyset(w) := \begin{cases} \mathsf{V}_{c_K}(w), & w \in [0, \bar{w}], \\ \bar{V} - K, & w > \bar{w}. \end{cases} \tag{39}$$

Futher define the following function for state I, which is obtained by smooth-pasting between the function V_{c_K} and a linear piece for the interval $[\underline{\theta}_K, \bar{w}]$:

$$V_\mathsf{I}(w) := \begin{cases} \mathsf{V}_{c_K}(w), & w \in [0, \underline{\theta}_K], \\ \bar{V} + m_K(w - \bar{w}), & w \in [\underline{\theta}_K, \bar{w}], \\ \bar{V}, & w > \bar{w}. \end{cases} \tag{40}$$

Figure 5 gives an example of the societal value functions $V_\mathsf{I}(w)$ and $V_\emptyset(w)$, as defined in (40) and (39), respectively. As we can see, the two functions are the same for $w < \underline{\theta}_K$, and have the same derivative at $w = \underline{\theta}_K$. The two functions then diverge for $w > \underline{\theta}_K$, with function $V_\mathsf{I}(w)$ being piece-wise linear in this interval. For $w \geqslant \bar{w}$, both functions become constant, and differ by exactly K.

Figure 5 Illustration of Optimal Societal Value Functions for Area II of Figure 2



Notes. In this figure, $r = 0.2$, $\rho = 0.5$, $c = b = 0.3$, $R = 10$, $\Delta\mu = 0.2$, $K = 1.6$, and $\mu = 0.6$. Hence, $\bar{w} = 0.9$, $\underline{\theta}_K = 0.1$, $\bar{V} = 24.9$, and $\underline{v} = 20$.

PROPOSITION 5. *Under Condition 2 and $K \in [\underline{K}, \bar{K}_2)$, functions $V_\emptyset(w)$ and $V_\mathsf{I}(w)$ as defined in (39) and (40), respectively, satisfy the optimality condition (20)–(22).*

Furthermore, for any $w \geqslant 0$, we have

$$U\left(\bar{\Gamma}\right) = \bar{V} - \bar{w} - K = V_\emptyset(\bar{w}) - \bar{w}, \tag{41}$$

in which $\bar{\Gamma}$ is defined in (15).

The fact that $\mathsf{V}_c'(w) > 1$ for any c and w implies that the derivative of $V_\emptyset(w)$ is higher than 1 for any $w \in [0, \bar{w})$. Therefore, function $V_\emptyset(w) - w$ is maximized at $\bar{w}$. Hence, for any contract Γ that yields the agent's utility $u(\Gamma) = w$, we must have

$$U(\Gamma) \leqslant V_\emptyset(w) - w \leqslant V_\emptyset(\bar{w}) - \bar{w} \leqslant U\left(\bar{\Gamma}\right),$$

implying the optimality of contract $\bar{\Gamma}$. To summarize, we have the following result, which corresponds to part (ii) of Theorem 2.

THEOREM 4. *Under Condition 2 and $K \in [\underline{K}, \bar{K}_2)$, it is optimal to hire the agent and pay β for each arrival.*

4.3. Low K and R (Area III of Figure 2)

More interesting and richer structure occurs under the following condition.

CONDITION 3. The model parameters satisfy either Condition 1 and $K < \bar{K}_1$ or Condition 2 and $K < \underline{K}$.

In this case, the value function for state I is no longer $\mathcal{V}_{\hat{\mathrm{w}}}(w)$. Recall the function $\mathcal{V}_{\widetilde{w}}$ defined in Lemma 2 for any $\widetilde{w} \in (0, \bar{w})$. When the switching cost K is low, the additional boundary condition that allows us to identify a particular $\widetilde{w}$ to obtain a value function is no longer at $w = 0$. Instead, we need to identify the threshold $\underline{\theta}$, at which point the value functions for states $\emptyset$ and I are connected. In particular, when the promised utility is below $\underline{\theta}$, the principal should suspend the agent, which allows the promised utility to increase, according to (11). For state $\emptyset$, we use function $\mathsf{V}_c(w)$ defined in (24) as the value function. The next result allows us to identify all the parameters, including the constant c in (24).

PROPOSITION 6. *Under Condition 3, there exists a set of parameters $(\mathsf{c}, \hat{\mathbf{w}}, \bar{\vartheta}, \underline{\vartheta})$ with $\mathsf{c} > 0$ and $\underline{\vartheta} < \bar{\vartheta} < \hat{\mathbf{w}} < \hat{\mathrm{w}}$, in which $\hat{\mathrm{w}}$ is defined in Lemma 3, such that*

$$\mathcal{V}_{\hat{\mathbf{w}}}(\underline{\vartheta}) = \mathsf{V}_{\mathsf{c}}(\underline{\vartheta}), \tag{42}$$

$$\mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = \mathsf{V}'_{\mathsf{c}}(\underline{\vartheta}), \tag{43}$$

$$\mathcal{V}_{\hat{\mathbf{w}}}(\bar{\vartheta}) = \mathsf{V}_{\mathsf{c}}(\bar{\vartheta}) + K, \ \textit{and} \tag{44}$$

$$\mathcal{V}'_{\hat{\mathbf{w}}}(\bar{\vartheta}) = \mathsf{V}'_{\mathsf{c}}(\bar{\vartheta}) > 1, \tag{45}$$

in which $\mathcal{V}_{\hat{\mathbf{w}}}$ is defined in Lemma 2 with $\hat{\mathbf{w}}$ replacing $\widetilde{w}$, and V_{c} is defined in (24) with c replacing c. Moreover, we have $\bar{\vartheta} > \breve{w}(\hat{\mathbf{w}})$, in which $\breve{w}(\hat{\mathbf{w}})$ is defined in Lemma 2(i), with $\hat{\mathbf{w}}$ replacing $\widetilde{w}$.

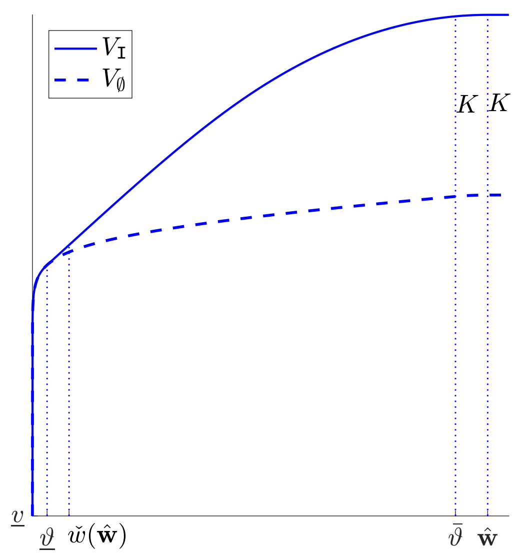
Equations (42) and (43) are called *value-matching* and *smooth-pasting* conditions, respectively, in the optimal control literature, and these occur at the promised utility threshold $\underline{\vartheta}$ when the state is switched from I to $\emptyset$. Similarly, (44) and (45) specify the value-matching and smooth-pasting conditions when the state is switched from $\emptyset$ to I at promised utility $\bar{\vartheta}$, except that the values between the two states differ by K. The proof of Proposition 6 is rather intricate and takes four key steps, as shown in Section EC.3.9. Relying on these key steps, we illustrate how to compute $\bar{\vartheta}$ and $\underline{\vartheta}$ in Section EC.1.2.

Equipped with Proposition 6, we define the following optimal value functions:

$$V_{\mathsf{I}}(w) := \begin{cases} \mathsf{V}_{\mathsf{c}}(w), & w \in [0, \underline{\vartheta}), \\ \mathcal{V}_{\hat{\mathbf{w}}}(w), & w \geqslant \underline{\vartheta}, \end{cases} \quad \text{and} \quad V_{\emptyset}(w) := \begin{cases} \mathsf{V}_{\mathsf{c}}(w), & w \in [0, \bar{\vartheta}), \\ \mathcal{V}_{\hat{\mathbf{w}}}(w) - K, & w \geqslant \bar{\vartheta}. \end{cases} \tag{46}$$

Figure 6 depicts functions V_{I} and $V_{\emptyset}$ defined in (46). In particular, functions $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ are identical for $w \leqslant \underline{\vartheta}$. Furthermore, function $V_{\mathsf{I}}(w)$ is linear in the interval $[\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}})]$, while $V_{\emptyset}(w)$ remains to be $\mathsf{V}_{\mathsf{c}}(w)$ for $w \leqslant \bar{\vartheta}$. For higher w such that $w \geqslant \bar{\vartheta}$, however, function $V_{\emptyset}(w)$ is a parallel shift of $V_{\mathsf{I}}(w)$, where the two functions differ by K.

Figure 6 Illustration of Optimal Societal Value Functions for Area III of Figure 2



Notes. In this figure, $r = 0.05$, $\rho = 1$, $c = b = 0.3$, $R = 112$, $\Delta\mu = 0.1$, $K = 40$, and $\mu = 1.95$. Hence, $\bar{w} = 5.85$, $\underline{\vartheta} = 0.18$, $\breve{w}(\hat{\mathbf{w}}) = 0.45$, $\bar{\vartheta} = 5.22$, and $\hat{\mathbf{w}} = 5.62$.

Note that by expression (45) of Proposition 6, the slope of $V_{\emptyset}(w)$ at $w = \bar{\vartheta}$ is larger than 1. (Figure 6 does not appear this way because of different scales of the x and y-axes. — The value of $\hat{\mathbf{w}}$ is around 5, while the difference $V_{\emptyset}(\hat{\mathbf{w}}) - V_{\emptyset}(0)$ is around a few thousand.) This implies that the principal's utility function $V_{\emptyset}(w) - w$ for state $\emptyset$ is maximized at a point higher than $\bar{\vartheta}$.

Now, we specify the optimal contract structure as follows. For any $w \in [\underline{\vartheta} \vee \breve{w}(\hat{\mathbf{w}}), \hat{\mathbf{w}}]$, define contract

$$\hat{\Gamma}(w) := \Gamma^* \left(w;\ \underline{\vartheta},\ (\underline{\vartheta} \vee \breve{w}(\hat{\mathbf{w}})),\ \bar{\vartheta},\ \hat{\mathbf{w}}\right), \tag{47}$$

where we use notation $a \vee b$ to denote $\max\{a, b\}$ for any $a, b \in \mathbb{R}$. The term $\underline{\vartheta} \vee \breve{w}(\hat{\mathbf{w}})$ as the threshold $\breve{w}$ of Definition 1 implies that random switching from I to $\emptyset$ occurs under this contract if and only if $\underline{\vartheta} < \breve{w}(\hat{\mathbf{w}})$. If $\underline{\vartheta} \geqslant \breve{w}(\hat{\mathbf{w}})$, on the other hand, contract $\hat{\Gamma}(w)$ demonstrates the typical control-band structure.

The following result implies that these contracts are indeed related to the optimal ones.

Proposition 7. *Under Condition 3, functions V_{I} and $V_{\emptyset}$ defined in (46) satisfy the optimality condition (20)–(22). Furthermore, for any $w \in [\underline{\vartheta} \vee \breve{w}(\hat{\mathbf{w}}), \hat{\mathbf{w}}]$, we have*

$$U\left(\hat{\Gamma}(w)\right) = V_{\emptyset}(w) - w. \tag{48}$$

Note that it is quite involved to verify that $V_{\mathsf{I}}(w)$ satisfies $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}} \geqslant 0$ in condition (20) for $w \in [0, \underline{\vartheta}]$. As one can imagine, we need to show that the function $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}}$ is always monotone in this interval. However, this function is not convex. In the proof presented in Section EC.3.12, we have to establish that either $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}}$'s first-order derivative is negative, or its second-order derivative is positive, throughout this interval. Together with the fact that the function $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}}$ takes a non-negative value and negative derivative at $\underline{\vartheta}$, this guarantees $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}} \geqslant 0$. Corresponding proofs in the existing literature, such as Duckworth and Zervos (2001) and Vath and Pham (2007), are much simpler in comparison. In particular, Vath and Pham (2007) rely on showing convexity/concavity to verify quasi-variational inequalities.

The following theorem corresponds to part (iii) of Theorem 2.

Theorem 5. *Under Condition 3, the contract $\hat{\Gamma}(\mathsf{w}_0^*)$ is optimal, in which $\mathsf{w}_0^* \in [\bar{\vartheta}, \hat{\mathbf{w}}]$ is a maximizer of the function $V_{\emptyset}(w) - w$, where function $V_{\emptyset}$ is defined in (46).*

5. Discussions

In this section, we discuss three extensions to the basic model. First, we let the switching cost K approach 0 in Section 5.1. Then, we study the cases in which there is also a fixed cost to switch from state I to $\emptyset$ in Section 5.2. Finally, we allow an extra cost per unit of time during suspension in Section 5.3.

5.1. Switching Cost K Approaching Zero

In this subsection, we discuss impacts of the switching cost K on the optimal contract, especially when K approaches zero. Note that when $K = 0$, if Condition 3 does not hold, the principal should either not hire the agent or always motivate the agent to work.

Proposition 8. *Under Condition 3, thresholds $\underline{\vartheta}$ and $\bar{\vartheta}$ defined in Proposition 6 are decreasing and increasing in K, respectively. Furthermore, these two values converge to the same value as K approaches 0, or equivalently,*

$$\theta_0 := \lim_{K\downarrow 0} \underline{\vartheta} = \lim_{K\downarrow 0} \bar{\vartheta}. \tag{49}$$

The monotonicity of $\underline{\vartheta}$ and $\bar{\vartheta}$ implies that the limit θ_0 is an upper or lower bound for these thresholds. In Section EC.1.2, we demonstrate an algorithm to compute the optimal contract for general K values, in which computing θ_0 is the first step.

Proposition 8 also implies that, as K approaches 0, the control band between $\underline{\vartheta}$ and $\bar{\vartheta}$ diminishes. Consequently, switching occurs more and more frequently. In the limit as K becomes zero, whenever the promised utility reaches the threshold θ_0, with a positive probability, the number of switchings will approach infinity in a finite time period, and thus, the promised utility will oscillate around θ_0. A similar, although not identical, phenomenon in the optimal contract structures arises in the Brownian motion uncertainty case, as demonstrated in Zhu (2013), where the promised utility becomes "sticky" when the promised utility reaches a threshold.

Intuitively, a high-switching-frequency control policy appears impractical. Therefore, it is instructive to reflect on basic modeling choices. If the switching cost is fairly low, it is often a good practice to ignore it when building the first model. However, if the corresponding optimal switching frequency is extremely high, any cost associated with switching cannot be ignored any more.

Although the optimal control is not practical if $K = 0$, we can still study the corresponding optimal value function, which sheds lights on the value of the suspension option, compared with always inducing the agent to work until potential termination. By Proposition 8, we have the following result, which allows us to construct the optimal value function for $K = 0$.

Theorem 6. *Under Condition 3, the following quantities are well defined:*

$$\hat{\mathbf{w}}_0 := \lim_{K\downarrow 0} \hat{\mathbf{w}}, \quad \textit{and} \quad \mathsf{c}_0 := \lim_{K\downarrow 0} \mathsf{c}, \tag{50}$$

in which $\hat{\mathbf{w}}$ and c are defined according to Proposition 6. Further define function

$$\mathbf{V}_{\theta_0}(w) := \begin{cases} \mathsf{V}_{\mathsf{c}_0}(w), & w \in [0, \theta_0], \\ \mathcal{V}_{\hat{\mathbf{w}}_0}(w), & w > \theta_0. \end{cases}$$

Functions $V_{\mathsf{I}} = V_{\emptyset} = \mathfrak{V}_{\theta_0}$ satisfy (20)–(22) in which we set $K = 0$.

Table 1 **Parameters of the Cases with Relative Difference that is Greater than 10%**

r	μ	$\Delta\mu/\mu$	$c/(R\Delta\mu)$	Relative Difference
0.01	1.9	0.9	0.5	68.74%
0.01	1.45	0.9	0.5	58.05%
0.01	1	0.9	0.5	42.14%
0.1	1.9	0.9	0.5	30.06%
0.1	1.45	0.9	0.5	28.36%
0.1	1	0.9	0.5	24.47%
0.01	0.55	0.9	0.5	20.40%
0.1	0.55	0.9	0.5	14.61%

Theorem 6 implies that as K approaches zero, the optimal value functions for positive K values converge to a value function $\mathfrak{V}_{\theta_0}$, which is an upper bound of the optimal value function for $K=0$. Therefore, function $\mathfrak{V}_{\theta_0}$ serves as a benchmark for potential benefits of the switching option. Proposition EC.1 in Section EC.1.2 describes how to compute the function $\mathfrak{V}_{\theta_0}$ directly, rather than treating it as the limit of a sequence of functions.

Following Theorem 6, we define the optimal principal's utility under $K=0$ as

$$\bar{U} := \max_{w\geqslant 0}\left\{\mathfrak{V}_{\theta_0}(w) - w\right\}.$$

It is worth comparing this value with the principal's utility without the suspension option, which is obtained in Cao et al. (2022), defined as

$$\underline{U} := \begin{cases} \max_{w\geqslant 0}\left\{\mathcal{V}_{\hat{\mathrm{w}}}(w) - w\right\}, & \text{under Condition 1,} \\ \bar{V} - \bar{w}, & \text{under Condition 2.} \end{cases} \tag{51}$$

Therefore, it is clear that if model parameters do not satisfy Condition 3, the switching option does not bring any value to the principal. Under Condition 3, we conduct a numerical test to compute the relative difference, $(\bar{U} - \underline{U})/\underline{U}$.

In particular, we consider the following model parameters. Fix $\rho = 1$, $R = 10$ and $c = b$. Take r from the set $\{0.01, 0.1, 0.5, 0.9, 0.99\}$, μ from $\{0.1, 0.55, 1.1, 1.45, 1.9\}$, $\Delta\mu/\mu$ from $\{0.1, 0.5, 0.9\}$, and $c/(R\Delta\mu)$ from $\{0.1, 0.5, 0.9\}$, so that model parameters satisfy Assumption 1 and $r < \rho$. Among these 225 cases, 85 of them satisfy Condition 3. The mean of the relative differences among these 85 cases is 3.71%. However, in 8 cases, the relative difference exceeds 10%. We list the parameters of these 8 cases in Table 1. As we can see, these cases correspond to r being very low (taking values 0.01 and 0.1), μ being not too low (no lower than 0.55), $\Delta\mu$ close to μ (ratio being 0.9), and $c/(R\Delta\mu)$ neither close to 0 nor close to 1. The maximum improvement of considering the switching option can be as high as 68.74%.

5.2. Positive Switching Cost From On to Off

Now, we briefly discuss a generalization of our basic model, which involves a fixed cost, call it $\mathcal{K}$, for the principal to direct the agent to stop working, including terminating the contract. Instead of providing a comprehensive summary of all results, we provide the key ideas and leave some details for the reader to fill in.

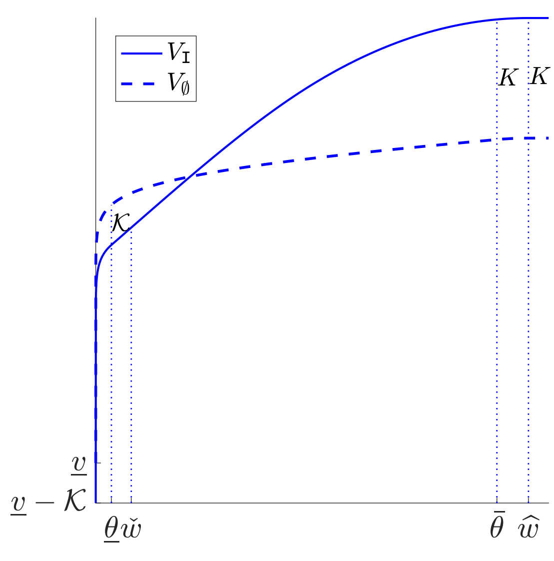
The general contract structure, Γ^* of Definition 1, remains optimal. In order to identify the specific parameters of the policy structure, we describe the optimal value functions.

First of all, in the verification theorem, condition (21) is revised to $-\mathcal{K} \leqslant V_{\mathsf{I}} - V_{\emptyset} \leqslant K$, and the second inequality in (22) changes to $V_{\emptyset}(0) \geqslant \underline{v} - \mathcal{K}$. The key idea for constructing the value functions is that when $w < \underline{\theta}$, function V_{I} is a downward parallel shift of $V_{\emptyset}$ by $\mathcal{K}$. Accordingly, the value-matching and smooth-pasting conditions of Proposition 6 become

$$V_{\mathsf{I}}(\underline{\theta}) = V_{\emptyset}(\underline{\theta}) - \mathcal{K}, \quad V_{\mathsf{I}}'(\underline{\theta}) = V_{\emptyset}'(\underline{\theta}),$$
$$V_{\mathsf{I}}(\bar{\theta}) = V_{\emptyset}(\bar{\theta}) + K, \quad \text{and} \quad V_{\mathsf{I}}'(\bar{\theta}) = V_{\emptyset}'(\bar{\theta}).$$

Figure 7 depicts the value functions. Similar to Figure 6, function $V_{\mathsf{I}}(w)$ is linear in the interval $w \in [\underline{\theta}, \breve{w}]$. Furthermore, for $w \leqslant \underline{\theta}$, function $V_{\mathsf{I}}(w)$ is a downward parallel shift from $V_{\emptyset}(w)$ by $\mathcal{K}$, while for $w \geqslant \bar{\theta}$, function $V_{\emptyset}(w)$ is a downward parallel shift from $V_{\mathsf{I}}(w)$ by K.

Figure 7 Illustration of Optimal Societal Value Functions with Positive Switching Cost from On to Off



Notes. In this figure, $r = 0.05$, $\rho = 1$, $c = b = 0.3$, $R = 112$, $\Delta\mu = 0.1$, $\mathcal{K} = 10$, $K = 30$, and $\mu = 1.95$. Hence, $\bar{w} = 5.85$, $\underline{\theta} = 0.2$, $\breve{w} = 0.46$, $\bar{\theta} = 5.21$, and $\widehat{w} = 5.62$.

5.3. Extra Cost Per Unit of Time During Suspension

In practice, the principal may suffer an exogenous opportunity cost per unit of time whenever the agent is under suspension. For example, when the firm is under competitive pressure, not having access to the agent's work may cost the firm more than the lost revenue. Hence, it would be reasonable to incorporate a cost rate c_s to the principal over the duration of suspension, in addition to the fixed cost of restarting the agent that has been considered previously. Consequently, the principal's total expected utility after the termination, $\underline{v}$, becomes $(\underline{\mu}R - c_s)/r$, and the operator $\mathcal{A}_\emptyset$ is changed to:

$$(\mathcal{A}_\emptyset f)(w) := rf(w) - \rho w f'(w) + (\rho - r)w - R\underline{\mu} + c_s.$$

Therefore, the solution to $(\mathcal{A}_\emptyset V)(w) = 0$ with boundary condition $V(0) = \underline{v}$ still has the form $\underline{v} + w + cw^{r/\rho}$. That is, the introduction of the suspension cost rate c_s may change the value of $\underline{v}$, but not the form of the optimality condition. Going through the proofs, we find that the value of $\underline{v}$ does not change our main results, although some derived quantities, such as $\hat{R}$, need to be modified accordingly.

6. Concluding Remarks

We have fully solved the optimal contract design problem that dynamically schedules an agent to work and temporarily suspend work over time, depending on past arrival times. Our main result shows that when the fixed cost to start working is high enough, the principal should not hire the agent. Otherwise, if the revenue per arrival is high enough, the principal should never suspend the agent and pays a fixed amount for each arrival from the beginning. If neither the fixed cost nor the revenue per arrial is too high, the contract demonstrates a rich but also easy-to-implement structure. In particular, the principal only needs to announce a deadline before which the agent needs to bring in an arrival to prevent suspension after each arrival and at the end of each suspension episode. If suspension happens, it lasts for a fixed period of time. It is interesting to see that such an easy-to-implement contract structure turns out to be optimal.

The fundamental uncertainty in our model is the arrival process. Keeping the per-unit-time revenue the same, if we scale the system such that the arrival rate becomes very large, uncertainty essentially diminishes, and the system achieves its first best. In contrast, if the arrival rate is very small, each arrival becomes so valuable such that the simple $\bar{\Gamma}$ contract should be optimal. In Section EC.1.4, we formally investigate these insights.

It is worth noting that there may be good reasons why in practice we rarely observe firms using suspension as an incentive tool. For example, our model does not capture all the hidden cost of maintaining an employee under suspension, who may become disgruntled, or leave the firm voluntarily. Otherwise, if the cost of replacing an agent is low, the firm may prefer to terminate the contract and replace the agent immediately, instead of keeping the focal one under suspension. As for any models, factors that are not captured deserve caution.

There are natural extensions to this model. For example, our model assumes that the principal undertakes the fixed switching cost. There could be settings where this cost is incurred to the agent and not observable to the principal. In such a setting, even if the principal reimburses this cost, the contract needs to mitigate the incentive for the agent to divert this fund for other purposes instead of switching on effort. Such a model poses additional challenges, and is left to future investigation.

Acknowledgement

The authors would like to thank the department editor Professor Ilia Tsetlin, an anonymous associate editor, and three referees for many constructive comments and suggestions. The authors contribute equally to this study and are ordered alphabetically.

Funding

The work of Ping Cao is financially supported by the National Natural Science Foundation of China [Grants 72122019 and 71771202]. The work of Feng Tian is financially supported by the Hong Kong RGC funding [RGC Early Career Scheme 2022/23 Number: 27500822].

Endnotes

1. For its website, see https://hk.centanet.com/info/en/index.

2. A standard approach in this literature is to first focus on the class of contracts that only motivate the agent to always work. After obtaining the optimal contract in this restrictive class, the authors provide a sufficient condition on model parameters under which the optimal contract indeed falls into this restricted class (see, for example, Demarzo and Sannikov 2006, Biais et al. 2010).

3. Here, we implicitly assume that the continuous part of L, L^c is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}_+$.

4. Notation $W_t(\Gamma,\nu)$ represents the agent's continuation utility after observing either an arrival or a random switching that occurs at time t, which may trigger an instantaneous payment at time t. Hence, in its definition (4), we use the Lebesgue-Stieltjes integral $\int_{t+}^{\infty}$ to exclude the possible instantaneous payment at time t.

5. The specific choice of the upper bound $\bar{W}$ is not important, as long as it is high enough such that constraint (WU) is not binding at optimality. Technically, we need this constraint to establish that a process related to $W_t(\Gamma,\nu)$ is a martingale in the proof of Theorem 1 that comes later in the paper.

6. Technically speaking, all time indices in the dt term in (PK) should be $t-$. However, it does not make any difference as there is no jump in the dt term. This kind of confusion also appears in other places, causing no harm to the results.

7. If the initial promised utility w_0 lies in $(0, \breve{w})$, a public randomization is required in the optimal contract, and if $w_0 > \widehat{w}$, there will be an initial instantaneous payment. It cannot be optimal for the principal to set the value of w_0 in these two intervals. Hence, to reduce technical complexity, we impose that $w_0 \in \{0\} \cup [\breve{w}, \widehat{w}]$ without loss of optimality for the proposed contract.

8. In this paper, we use terminologies such as "increasing" and "decreasing" in the strict sense, and specify "nondecreasing" and "nonincreasing" accordingly.

9. Condition 2 corresponds, but is not identical, to Equation (13) of Cao et al. (2022). The difference is due to model assumptions. The model in Cao et al. (2022) assumes that the effort cost is not immediately reimbursed, while it is in our model.

References

Abreu, D., D. Pearce, and E. Staccheti (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica 58*(5), 1041–1063.

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

Brekke, K. A. and B. Oksendal (1994). Optimal switching in an economic activity under uncertainty. *SIAM Journal on Control and Optimization 32*(4), 1021–1036.

Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer Series in Statistics. Springer.

Cao, P., F. Tian, and P. Sun (2022). Comment on "optimal contract to induce continued effort". *Management Science 68*(1), 796–808.

Chen, M., P. Sun, and Y. Xiao (2020). Optimal monitoring schedule in dynamic contracts. *Operations Research 68*(5), 1285–1624.

Cvitanić, J., D. Possamaï, and N. Touzi (2016). Moral hazard in dynamic risk management. *Management Science 63*(10), 3328–3346.

Demarzo, P. M. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance 61*(6), 2681–2724.

DerStandard (2009). Companies are increasingly using suspension agreements (August 6). https://www.derstandard.at/story/1246543901120/krise-durchtauchen-betriebe-nutzen-vermehrt-aussetzungsvertrag.

Dohmen, T. J., A. Falk, D. Huffman, and U. Sunde (2012). Interpreting time horizon effects in inter-temporal choice. *Available at SSRN 2016024*.

Duckworth, K. and M. Zervos (2001). A model for investment decisions with switching costs. *Annals of Applied Probability 11*(1), 239–260.

Eurofound (2020). Temporary suspension of employment with re-employment commitments (November 17). https://www.eurofound.europa.eu/observatories/emcc/erm/support-instrument/temporary-suspension-of-employment-and-reemployment-commitments.

Eurofound (2021). Temporary reduction of working hours/suspension of employment contracts (August 3). https://www.eurofound.europa.eu/observatories/emcc/erm/support-instrument/temporary-reduction-of-working-hourssuspension-of-employment-contr.

Georgiadis, G., Y. Kim, and H. D. Kwon (2022). The absence of attrition in the war of attrition under complete information. *Games and Economic Behavior 131*, 171–185.

Grochulski, B. and Y. Zhang (2023). Termination as an incentive device. *Theoretical Economics 18*(1), 381–419.

Gupta, S., W. Chen, M. Dawande, and G. Janakiraman (2022). Three years, two papers, one course off: Optimal nonmonetary reward policies. *Management Science*. Forthcoming.

Hernandez Santibanez, N., D. Possamai, and C. Zhou (2020). Bank monitoring incentives under moral hazard and adverse selection. *Journal of Optimization Theory and Applications 184*(3), 988–1035.

Hong Kong Business Times (2020). Companies have tried all the ways to reduce costs in the epidemic market, lay off without pay, reduce pay, take unpaid leave (October 29). https://www.businesstimes.com.hk/articles/130487.

Hong Kong Economic Times (2020). Does the Employment Ordinance allow for "suspension without pay"? How should migrant workers protect themselves? (November 2). https://wealth.hket.com/article/2790377.

Hu, Z. and W. Tang (2021). Size matters, so does duration: The interplay between offer size and offer deadline. *Management Science 67*(8), 4935–4960.

Kwon, H. D. (2022). Game of variable contributions to the common good under uncertainty. *Operations Research 70*(3), 1359–1370.

Kwon, H. D., W. Xu, A. Agrawal, and S. Muthulingam (2016). Impact of Bayesian learning and externalities on strategic investment. *Management Science 62*(2), 550–570.

Myerson, R. B. (2015). Moral hazard in high office and the dynamics of aristocracy. *Econometrica 83*(6), 2083–2126.

Pages, H. (2013). Bank monitoring incentives and optimal ABS. *Journal of Financial Intermediation 22*(1), 30–54.

Pages, H. and D. Possamai (2014). A mathematical treatment of bank monitoring incentives. *Finance and Stochastics 18*(1), 39–73.

Protter, P. E. (2003). *Stochastic Integration and Differential Equations* (second edition ed.), Volume 21 of *Stochastic Modelling and Applied Probability.* Springer.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies 75*(3), 957–984.

Spear, S. and S. Srivastava (1987). On repeated moral hazard with discounting. *Review of Economic Studies 54*(4), 599–617.

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

Tian, F., P. Sun, and I. Duenyas (2021). Optimal contract for machine repair and maintenance. *Operations Research 69*(3), 916–949.

Vath, V. L. and H. Pham (2007). Explicit solution to an optimal switching problem in the two-regime case. *SIAM Journal on Control and Optimization 46*(2), 395–426.

Vath, V. L., H. Pham, and S. Villeneuve (2008). A mixed singular/switching control problem for a dividend policy with reversible technology investment. *Annals of Applied Probability 18*(3), 1164–1200.

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

Zorc, S. and I. Tsetlin (2020). Deadlines, offer timing, and the search for alternatives. *Operations Research 68*(3), 927–948.

Zorc, S., I. Tsetlin, and S. Hasija (2019). The when and how of delegated search. Working paper.

E-Companion for "Punish Underperformance with Suspension — Optimal Dynamic Contracts in the Presence of Switching Cost"

In this e-companion, we present some further discussions in Section EC.1, and provide all the proofs that are omitted from the main paper in Sections EC.2–EC.4.

EC.1. Further Discussions

This section contains four parts. Section EC.1.1 gives a heuristic derivation of the optimality condition (20)–(22) for the optimal value functions V_{I} and $V_{\emptyset}$, which appears in Section 3.2. Section EC.1.2 demonstrates how to compute the optimal contract parameters. Furthermore, we consider a special case of equal time discount in Section EC.1.3 and investigate the effect of arrival rate under fixed revenue rate in Section EC.1.4.

EC.1.1. A Heuristic Derivation of the Optimality Condition (20)–(22)

In this section, we provide a heuristic derivation of the optimality condition for the principal's utility functions and of the main features of the optimal contract, whose main idea follows from Section 4.1 in Biais et al. (2010). However, our arguments are not exactly the same, due to the presence of switching and randomization. Let $F_{\mathsf{I}}(w)$ and $F_{\emptyset}(w)$ be the principal's optimal utility function that yields an agent's utility w when the initial state is I and $\emptyset$, respectively.

For any $t \geqslant 0$, let us first characterize the evolution of the principal's utility function $F_{\mathcal{E}_{t-}}(W_{t-})$. Since the principal discounts the future utility flow at rate r, his expected flow rate of utility at time t is $rF_{\mathcal{E}_{t-}}(W_{t-})$. This must be equal to the sum of expected cash flow, the (possible) switching cost, and the expected rate of change in his continuation utility over $(t-\mathrm{d}t, t]$. Hence, we have

$$rF_{\mathcal{E}_{t-}}(W_{t-})\mathrm{d}t = [\bar{\nu}_t R - (c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}]\mathrm{d}t - \mathrm{d}L_t + \mathbb{E}_{t-}[-\kappa(\mathcal{E}_{t-}, \mathcal{E}_t) + \mathrm{d}F_{\mathcal{E}_t}(W_t)], \tag{EC.1}$$

where $\mathbb{E}_{t-}[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_{t-}]$.

Following the discussions in Section 3.2, we assume that for any $\varepsilon \in \{\mathsf{I}, \emptyset\}$, $F_{\varepsilon}(\cdot)$ is concave and differentiable on $\mathbb{R}_+$. The actual value function might not be differentiable on the entire domain $\mathbb{R}_+$, which is an issue frequently arising in the optimal control literature, and often addressed by the viscosity solution approach. Since this section is devoted to a heuristic derivation of the optimality equation for the optimal utility function F_{ε}, we assume that F_{ε} is smooth enough temporarily.

Recall that $\mathrm{d}L_t = \ell_t \mathrm{d}t + \Delta L_t$. Note that under any admissible IC contract, $\mathbb{1}_{\nu_t=\underline{\mu}} = \mathbb{1}_{\mathcal{E}_t=\emptyset}$ and $\mathbb{1}_{\nu_t=\mu} = \mathbb{1}_{\mathcal{E}_t=\mathsf{I}}$. Using (PK) and regarding $F_{\varepsilon}(w)$ as a function of (w, ε), we apply calculus of point processes to the process $(W, \mathcal{E})$ to obtain

$$\mathrm{d}F_{\mathcal{E}_t}(W_t) = \big(\rho W_{t-} + b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}} - H_t\bar{\nu}_t + q_t H_t^q - \ell_t\big) F'_{\mathcal{E}_{t-}}(W_{t-})\mathrm{d}t$$

$$+\big[F_{\mathcal{E}_{t-}}(W_{t-}-\Delta L_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big]+\big[F_{\mathcal{E}_{t-}}(W_{t-}+H_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big]\mathrm{d}N_t$$
$$+\big[F_{\mathcal{E}_{t-}}(W_{t-}-H_t^q)-F_{\mathcal{E}_{t-}}(W_{t-})\big]\mathrm{d}Q_t+\big[F_{\mathcal{E}_t}(W_t)-F_{\mathcal{E}_{t-}}(W_t)\big].$$

Plugging the above formula into (EC.1) and using $\mathbb{E}_{t-}\mathrm{d}N_t=\bar{\nu}_t\mathrm{d}t$ as well as $\mathbb{E}_{t-}\mathrm{d}Q_t=q_t\mathrm{d}t$, we have

$$\begin{aligned} rF_{\mathcal{E}_{t-}}(W_{t-})\mathrm{d}t=&\Big[R\bar{\nu}_t-(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-\ell_t+\big(\rho W_{t-}+b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-H_t\bar{\nu}_t+q_tH_t^q-\ell_t\big)F'_{\mathcal{E}_{t-}}(W_{t-})\\ &+\big(F_{\mathcal{E}_{t-}}(W_{t-}+H_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big)\bar{\nu}_t+\big(F_{\mathcal{E}_{t-}}(W_{t-}-H_t^q)-F_{\mathcal{E}_{t-}}(W_{t-})\big)q_t\Big]\mathrm{d}t\\ &-\Delta L_t+F_{\mathcal{E}_{t-}}(W_{t-}-\Delta L_t)-F_{\mathcal{E}_{t-}}(W_{t-})+\mathbb{E}_{t-}\big[-\kappa(\mathcal{E}_{t-},\mathcal{E}_t)+F_{\mathcal{E}_t}(W_t)-F_{\mathcal{E}_{t-}}(W_t)\big]. \end{aligned} \tag{EC.2}$$

Here, ℓ_t, ΔL_t, H_t, H_t^q, q_t, and $\mathcal{E}_t$ are all control variables. Besides, the contract might be terminated at time t by paying off the promised utility to the agent instantaneously. Hence, we have $F_{\mathcal{E}_t}(W_t)\geqslant \underline{v}-W_t$. That is, $F_\varepsilon(w)\geqslant\underline{v}-w$ for any $w\in\mathbb{R}_+$ and $\varepsilon\in\{\mathsf{I},\emptyset\}$.

We first optimize the constant-order terms on the right-hand side in (EC.2). Considering that the optimized constant-order terms should be zero, we have

$$\max_{\Delta L_t\geqslant 0}\big\{-\Delta L_t+F_{\mathcal{E}_{t-}}(W_{t-}-\Delta L_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big\}=0,\ \text{and} \tag{EC.3}$$
$$\max_{\mathcal{E}_t\in\{\mathsf{I},\emptyset\}}\big\{-\kappa(\mathcal{E}_{t-},\mathcal{E}_t)+F_{\mathcal{E}_t}(W_t)-F_{\mathcal{E}_{t-}}(W_t)\big\}=0. \tag{EC.4}$$

Equation (EC.3) yields that $F'_\varepsilon(w)\geqslant -1$ for any $w\in\mathbb{R}_+$ and $\varepsilon\in\{\mathsf{I},\emptyset\}$. Let $\widehat{w}_\varepsilon=\inf\{w\geqslant 0\mid F'_\varepsilon(w)=-1\}$. The concavity of $F_\varepsilon(\cdot)$ implies that at any time instant t, it is optimal for the principal to pay $\Delta L_t=\max\{W_{t-}-\widehat{w}_{\mathcal{E}_{t-}},0\}$ instantaneously to the agent.

Equation (EC.4) yields that $F_{\mathsf{I}}(w)\geqslant F_\emptyset(w)$ and $F_\emptyset(w)\geqslant F_{\mathsf{I}}(w)-K$ for any $w\in\mathbb{R}_+$. Besides, $\mathcal{E}_t\neq\mathcal{E}_{t-}$ only if $-\kappa(\mathcal{E}_{t-},\mathcal{E}_{t-}^c)+F_{\mathcal{E}_{t-}^c}(W_t)-F_{\mathcal{E}_{t-}}(W_t)=0$, where ε^c is I if $\varepsilon=\emptyset$ and is $\emptyset$ if $\varepsilon=\mathsf{I}$.

Next, we consider the controls such that $\Delta L_t=0$ and $\mathcal{E}_t=\mathcal{E}_{t-}$. If we plug these values into (EC.2), the symbol "=" should be replaced by "⩽" due to the suboptimality of these controls. Comparing the dt-order terms on both sides of the resulting inequality yields

$$\begin{aligned} rF_{\mathcal{E}_{t-}}(W_{t-})\geqslant\max\Big\{&R\bar{\nu}_t-(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-\ell_t+\big(\rho W_{t-}+b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-H_t\bar{\nu}_t+H_t^qq_t-\ell_t\big)F'_{\mathcal{E}_{t-}}(W_{t-})\\ &+\big(F_{\mathcal{E}_{t-}}(W_{t-}+H_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big)\bar{\nu}_t+\big(F_{\mathcal{E}_{t-}}(W_{t-}-H_t^q)-F_{\mathcal{E}_{t-}}(W_{t-})\big)q_t\Big\}, \end{aligned} \tag{EC.5}$$

where the maximization is taken over the set of controls (ℓ_t,H_t,H_t^q,q_t) that satisfies $\ell_t\geqslant b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}$, the IR constraint (5), and the IC constraint (IC).

Inequality (EC.5) can be written as two inequalities, for working and suspension states. If $\mathcal{E}_{t-}=\mathsf{I}$, by omitting the time index, (EC.5) becomes

$$\begin{aligned} rF_{\mathsf{I}}(w)\geqslant\ &R\mu-(c-b)+(\rho w+b)F'_{\mathsf{I}}(w)\\ &+\max\Big\{-\ell-(\ell+\mu h-qh^q)F'_{\mathsf{I}}(w)+\mu(F_{\mathsf{I}}(w+h)-F_{\mathsf{I}}(w))+(F_{\mathsf{I}}(w-h^q)-F_{\mathsf{I}}(w))q\Big\}, \end{aligned} \tag{EC.6}$$

where the maximization is taken over the set of (ℓ, h, h^q, q) that satisfies

$$\ell \geqslant b, \;\; h \geqslant \beta, \;\; h^q \leqslant w, \;\; q \geqslant 0. \tag{EC.7}$$

If $\mathcal{E}_{t-} = \emptyset$, then (EC.5) becomes

$$\begin{aligned} rF_\emptyset(w) \geqslant R\underline{\mu} + \rho w F_\emptyset'(w) + \max\Big\{ & -\ell - (\ell + \underline{\mu} h - q h^q) F_\emptyset'(w) + \underline{\mu}\big(F_\emptyset(w+h) - F_\emptyset(w)\big) \\ & + \big(F_\emptyset(w-h^q) - F_\emptyset(w)\big) q \Big\}, \end{aligned} \tag{EC.8}$$

where the maximization is taken over the set of (ℓ, h, h^q, q) that satisfies

$$\ell \geqslant 0, \;\; h \geqslant -w, \;\; h^q \leqslant w, \;\; q \geqslant 0. \tag{EC.9}$$

Recall that $V_\mathsf{I}(w) = F_\mathsf{I}(w) + w$ and $V_\emptyset(w) = F_\emptyset(w) + w$. Then, based on the above discussions, we have the following basic properties of V_I and $V_\emptyset$:

1. $V_\mathsf{I}(w) \geqslant \underline{v}$ and $V_\emptyset(w) \geqslant \underline{v}$ for any $w \in \mathbb{R}_+$.
2. $V_\mathsf{I}'(w) \geqslant 0$ and $V_\emptyset'(w) \geqslant 0$ for any $w \in \mathbb{R}_+$ (this follows from the fact that $F_\mathsf{I}'(w) \geqslant -1$ and $F_\emptyset'(w) \geqslant -1$).
3. Both V_I and $V_\emptyset$ are concave on $\mathbb{R}_+$.
4. V_I (resp. $V_\emptyset$) will take constant value on $[\widehat{w}_\mathsf{I}, \infty)$ (resp. $[\widehat{w}_\emptyset, \infty)$).
5. $V_\mathsf{I}(w) \geqslant V_\emptyset(w)$ and $V_\emptyset(w) \geqslant V_\mathsf{I}(w) - K$ for any $w \in \mathbb{R}_+$.

We proceed to analyze (EC.6), which can be rewritten as follows in terms of V_I:

$$\begin{aligned} rV_\mathsf{I}(w) \geqslant R\mu - c - (\rho - r)w + (\rho w + b)V_\mathsf{I}'(w) + \max\Big\{ & -\ell V_\mathsf{I}'(w) + \big(V_\mathsf{I}(w+h) - V_\mathsf{I}(w) - hV_\mathsf{I}'(w)\big)\mu \\ & + \big(V_\mathsf{I}(w-h^q) - V_\mathsf{I}(w) + h^q V_\mathsf{I}'(w)\big) q \Big\}, \end{aligned} \tag{EC.10}$$

where the maximization is taken over the constraints (EC.7).

Optimizing the right-hand side of (EC.10) with respect to ℓ, we have $\ell^* = \arg\max_{\ell \geqslant b}\{-\ell V_\mathsf{I}'(w)\} = b$ if $w \in [0, \widehat{w}_\mathsf{I})$, where we use the fact that $V_\mathsf{I}'(w) > 0$ for $w \in [0, \widehat{w}_\mathsf{I})$.

Optimizing the right-hand side of (EC.10) with respect to h, we have $h^* = \arg\max_{h \geqslant \beta}\{V_\mathsf{I}(w + h) - V_\mathsf{I}'(w)h\} = \beta$, by noting that $V_\mathsf{I}(w+h) - V_\mathsf{I}'(w)h$ is decreasing in h on $[0, \infty)$, since $V_\mathsf{I}'(w+h) - V_\mathsf{I}'(w) \leqslant 0$ for any $h \geqslant 0$ due to the concavity of V_I.

Note that $\max_{h^q \leqslant w}\{V_\mathsf{I}(w - h^q) - V_\mathsf{I}(w) + h^q V_\mathsf{I}'(w)\} = 0$. Hence, (EC.10) reduces to

$$rV_\mathsf{I}(w) \geqslant R\mu - c - (\rho - r)w - \rho(\bar{w} - w)V_\mathsf{I}'(w) + \mu(V_\mathsf{I}(w+\beta) - V_\mathsf{I}(w)), \tag{EC.11}$$

for $w \in \mathbb{R}_+$, which can be rewritten as $(\mathcal{A}_\mathsf{I} V_\mathsf{I})(w) \geqslant 0$ by using the operator $\mathcal{A}_\mathsf{I}$ defined in (18).

We next analyze (EC.8), which can be rewritten as follows in terms of $V_\emptyset$:

$$rV_\emptyset(w) \geqslant R\underline{\mu} - (\rho - r)w + \rho w V_\emptyset'(w) + \max\Big\{ -\ell V_\emptyset'(w) + \underline{\mu}\big(V_\emptyset(w+h) - V_\emptyset(w) - hV_\emptyset'(w)\big) + \big(V_\emptyset(w-h^q) - V_\emptyset(w) + h^q V_\emptyset'(w)\big)q\Big\}, \tag{EC.12}$$

where the maximization is taken over the constraint set (EC.9).

Optimizing the right-hand side of (EC.12) with respect to ℓ, we have $\ell^* = \arg\max_{\ell\geqslant 0}\{-\ell V_\emptyset'(w)\} = 0$ if $w \in [0, \widehat{w}_\emptyset)$. Optimizing the right-hand side of (EC.12) with respect to h, we have $h^* = \arg\max_{h\geqslant -w}\{-V_\emptyset'(w)h + V_\emptyset(w+h)\} = 0$, by noting that $-V_\emptyset'(w)h + V_\emptyset(w+h)$ is increasing in h for $h < 0$ and decreasing in h for $h > 0$ due to the concavity of $V_\emptyset$. Additionally, we have $\max_{h^q\leqslant w}\{V_\emptyset(w-h^q) - V_\emptyset(w) + h^q V_\emptyset'(w)\} = 0$. Consequently, (EC.12) can further reduce to

$$rV_\emptyset(w) \geqslant R\underline{\mu} - (\rho - r)w + \rho w V_\emptyset'(w), \tag{EC.13}$$

which can be rewritten as $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$.

Summarizing the above discussions yields the optimality condition (20)–(22).

EC.1.2. Computing Contract Parameters

For $K = 0$, we have the following results. Since these results have been established in the second part of the proof of Proposition 8, we omit its proof.

PROPOSITION EC.1. (i) *Under Condition 1 and $\bar{K}_1 > 0$, we have $\theta_0 = \underline{\theta}^0$, where θ_0 and $\underline{\theta}^0$ are defined in Proposition 8 and Lemma EC.5, respectively. Correspondingly, we have $\hat{\mathbf{w}}_0 = \widetilde{w}(\theta_0)$ and $\mathsf{c}_0 = C(\theta_0)$, in which functions $\widetilde{w}(\cdot)$ and $C(\cdot)$ are defined in Lemma EC.4.*

(ii) *Under Condition 2 and $\underline{K} > 0$, define a lower bound*

$$\breve{\underline{\theta}} := \frac{(\bar{V} - \underline{v})(\rho - r - \mu) - (\rho - r)\bar{w}}{\mu(\rho/r - 1)}.$$

Similar to Lemmas EC.4 and EC.5, for any $\underline{\theta} \in (\breve{\underline{\theta}}, \bar{w})$, there exist unique values $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ and $C(\underline{\theta})$, such that if we set $\hat{\mathbf{w}} = \widetilde{w}(\underline{\theta})$, $\mathsf{c} = C(\underline{\theta})$, and $\underline{\vartheta} = \underline{\theta}$, the value-matching and smooth-pasting conditions (42) and (43) are satisfied. Furthermore, value $\underline{\theta}^0 := \inf\{\underline{\theta} \in (\breve{\underline{\theta}}, \bar{w}) \mid \widetilde{w}'(\underline{\theta}) \geqslant 0\}$ is well defined, and we have $\theta_0 = \underline{\theta}^0$, $\hat{\mathbf{w}}_0 = \widetilde{w}(\theta_0)$, and $\mathsf{c}_0 = C(\theta_0)$.

For any $\underline{\theta} \in (0, \bar{w})$, function $h(\widetilde{w}, \underline{\theta})$, as defined in (EC.71), is decreasing in $\widetilde{w}$ with $h(\widetilde{w}(\underline{\theta}), \underline{\theta}) = 0$. Hence, $\widetilde{w}(\underline{\theta})$ can be efficiently found by a binary search procedure, starting from lower bound $\underline{\theta}$ and upper bound $\bar{w}$. Consequently, $C(\underline{\theta})$ can also be immediately computed as $C_1(\widetilde{w}(\underline{\theta}), \underline{\theta})$, with $C_1(\widetilde{w}, \underline{\theta})$ defined in (EC.70). Therefore, following Proposition EC.1, in order to determine the optimal contract parameters for $K = 0$ under Condition 3, we only need to find $\underline{\theta}^0$. Based on the definition of $\underline{\theta}^0$ (see part (ii) of Proposition EC.1), this value can be determined by a line search to check at which point $\tilde{w}(\underline{\theta})$ is no longer increasing, starting from 0 under Condition 1 and $\bar{K}_1 > 0$, or from $\breve{\underline{\theta}}$ under Condition 2 and $\underline{K} > 0$.

Computation of the optimal contract parameters for $K > 0$ is more complex. We only demonstrate how to compute the control-band parameters $(\mathsf{c}, \hat{\mathbf{w}}, \bar{\vartheta}, \underline{\vartheta})$ under Condition 1 and $K < \bar{K}_1$ or under Condition 2 and $K < \underline{K}$, as the optimal contract in other cases takes a simpler form. Take the case under Condition 1 and $K < \bar{K}_1$ for illustration. Note that for any $\underline{\theta} \in (0, \underline{\theta}^0)$, the value $\bar{\theta}(\underline{\theta})$ can be determined by (EC.74) using a line search procedure. Hence, function $\psi(\underline{\theta})$, as defined in (EC.75), can be readily computed for each $\underline{\theta} \in (0, \underline{\theta}^0)$. Since, by Lemma EC.7, function $\psi(\underline{\theta})$ is decreasing in $\underline{\theta}$ with $\psi(\underline{\vartheta}) = K$, the quantity $\underline{\vartheta}$ can be efficiently found by a binary search procedure, starting from lower bound 0 and upper bound $\underline{\theta}^0$. The three other parameters, c, $\hat{\mathbf{w}}$, and $\bar{\vartheta}$, are thus immediately computed as $C(\underline{\vartheta})$, $\tilde{w}(\underline{\vartheta})$, and $\bar{\theta}(\underline{\vartheta})$. For the case under Condition 2 and $K < \underline{K}$, the only difference is that the initial lower bound for the binary search is $\breve{\underline{\theta}}$.

The above procedure can be summarized by the following four subroutines.

`Subroutine 1`. Given $\underline{\theta} \in (0, \bar{w})$, compute $\tilde{w}(\underline{\theta})$: Binary search on $[\underline{\theta}, \tilde{w}]$ to determine $\tilde{w}(\underline{\theta})$ according to $h(\tilde{w}(\underline{\theta}), \theta) = 0$ where function $h(\tilde{w}, \underline{\theta})$ is defined in (EC.71).

`Subroutine 2`. Given $\underline{\theta} \in (0, \bar{w})$, compute $C(\underline{\theta})$: Following `Subroutine 1`, we obtain $\tilde{w}(\underline{\theta})$. Then, $C(\underline{\theta}) = C_1(\tilde{w}(\underline{\theta}), \underline{\theta})$ with $C_1(\tilde{w}, \underline{\theta})$ defined in (EC.70).

`Subroutine 3`. Given $\underline{\theta} \in (0, \underline{\theta}^0)$, compute $\bar{\theta}(\underline{\theta})$: Following `Subroutines 1` and `2`, we obtain $\tilde{w}(\underline{\theta})$ and $C(\underline{\theta})$. Then, we calculate $\bar{\theta}(\underline{\theta})$ by (EC.74) using a line search procedure.

`Subroutine 4`. Given $\underline{\theta} \in (0, \underline{\theta}^0)$, compute $\psi(\underline{\theta})$: Following `Subroutines 1-3`, we obtain $\tilde{w}(\underline{\theta})$, $C(\underline{\theta})$, and $\bar{\theta}(\underline{\theta})$. Then, we compute $\psi(\underline{\theta})$, as defined in (EC.75).

With the above four steps, the optimal control-band parameters can be computed by Algorithm 1 below.

Algorithm 1 Compute $(\mathsf{c}, \hat{\mathbf{w}}, \bar{\vartheta}, \underline{\vartheta})$.

1: Line search to determine $\underline{\theta}^0$ according to $\tilde{w}'(\theta) = 0$, in which function $\tilde{w}(\theta)$ is computed according to `Subroutine 1`.

2: Binary search to determine $\underline{\vartheta}$ according to $\psi(\underline{\vartheta}) = K$, where $\psi(\underline{\vartheta})$ can be computed using `Subroutine 4`.

3: Following `Subroutines 1-3`, we obtain $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta})$, $\mathsf{c} = C(\underline{\vartheta})$, and $\underline{\vartheta} = \bar{\theta}(\underline{\vartheta})$, respectively.

EC.1.3. Equal Discount Rate

In the study of dynamic contracts without the switching option, Sun and Tian (2018) claimed, without a formal proof, that under equal discount rates, it is optimal for the principal to always induce the agent to work before contract termination. In our context with switching, this claim corresponds to never switching the agent to suspension and then working again. Here, we provide a formal proof that validates this claim for any $K \geqslant 0$.

When the two players' discount rates are the same, that is, $r = \rho$, various expressions in the main body of the paper become simpler. For example, the value $\bar{V}$ defined in (10) becomes

$$\bar{V}_e := \frac{\mu R - c}{r}, \tag{EC.14}$$

and the differential equation (25), which plays an essential role in deciding the optimal value functions, becomes

$$(\mu + r)V_e(w) - \mu V_e((w+\beta) \wedge \bar{w}) + r(\bar{w} - w)V_e'(w) - (\mu R - c) = 0. \tag{EC.15}$$

According to Lemma 3 of Sun and Tian (2018), differential equation (EC.15) with boundary condition $V_e(0) = \underline{v}$ has a unique solution V_e on $[0, \bar{w}]$, which is increasing and strictly concave, with $V_e(w) = \bar{V}_e$ for all $w \geqslant \bar{w}$. Theorem 1 still holds, in which the operators $\mathcal{A}_{\mathsf{I}}$ and $\mathcal{A}_{\emptyset}$ are simplified to

$$(\mathcal{A}_{\mathsf{I}} f)(w) = (\mu + r)f(w) - \mu f(w+\beta) + r(\bar{w} - w)f'(w) - (\mu R - c), \text{ and}$$
$$(\mathcal{A}_{\emptyset} f)(w) = r f(w) - r w f'(w) - R\underline{\mu},$$

respectively, for differentiable function f.

Furthermore, when $r = \rho$, effectively Condition 1 holds. In particular, we will show that the value function for state I is V_e defined above. Furthermore, the upper threshold $\bar{V}(\hat{\mathrm{w}}) - \underline{v}$ in Proposition 2 becomes

$$\bar{K}_e := \bar{V}_e - \underline{v}. \tag{EC.16}$$

In order to define the lower threshold for the switching cost, we need to define the value function for state $\emptyset$. Note that when $r = \rho$, function $\mathcal{V}_{\hat{\mathrm{w}}}$ becomes V_e, with $\hat{\mathrm{w}}$ being $\bar{w}$ and $\breve{w}(\hat{\mathrm{w}})$ being 0. Hence, following Lemma 4, if $K < \bar{K}_e$, there exist K-dependent values $\bar{\theta}^K \in [0, \bar{w}]$ and $m^K \in [0, V_e'(0)]$ such that

$$V_e(\bar{\theta}^K) = m^K \bar{\theta}^K + K + \underline{v}, \text{ and } V_e'(\bar{\theta}^K) = m^K.$$

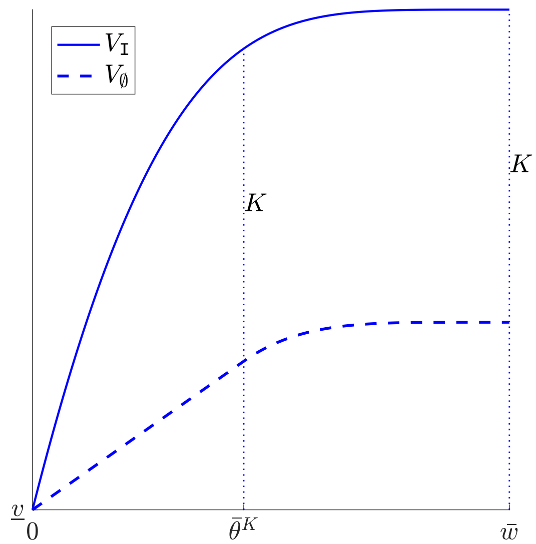
Then, similar to (30), we define the following societal value function for the suspension state:

$$V_{\emptyset}(w) = \begin{cases} m^K w + \underline{v}, & w \in \left[0, \bar{\theta}^K\right], \\ V_e(w) - K, & w \in \left[\bar{\theta}^K, \bar{w}\right]. \end{cases} \tag{EC.17}$$

Figure EC.1 depicts the value functions. It is clear that $V_{\emptyset}$ is linear over the interval $[0, \bar{\theta}^K]$. Furthermore, $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ are "parallel" with a difference of K for $w \geqslant \bar{\theta}^K$.

The following theorem summarizes the optimality results.

Figure EC.1 Illustration of Optimal Societal Value Functions with Equal Discount Rates



Notes. In this figure, $r = 0.5$, $\rho = 0.5$, $c = b = 0.2$, $R = 2$, $\Delta\mu = 0.7$, $K = 1.5$, and $\mu = 2$. Hence, $\bar{\theta}^K = 0.51$, $\bar{w} = 1.14$, $\bar{V}_e = 7.6$, and $\underline{v} = 5.2$.

THEOREM EC.1. *Consider $r = \rho$. For any $w \geqslant 0$, we have*

$$U(\underline{\Gamma}, \emptyset) = \underline{v}.$$

If $K \geqslant \bar{K}_e$, functions $V_{\mathsf{I}} = V_e$ and $V_{\emptyset} = \underline{v}$ satisfy (20)–(22).

If $K < \bar{K}_e$, on the other hand, functions $V_{\mathsf{I}} = V_e$ and $V_{\emptyset}$ as defined in (EC.17) satisfy (20)–(22). Furthermore, if $V_e'(\bar{\theta}^K) > 1$, for any $w \geqslant \bar{\theta}^K$, we have

$$U(\Gamma^*(w; 0, 0, \bar{w}, \bar{w})) = V_{\emptyset}(w) - w.$$

Proof. Using a similar argument as that in the proof of Proposition 2, we can show that (i) $U(\underline{\Gamma}, \emptyset) = \underline{v}$, and (ii) under the condition that $K < \bar{K}_e$ and $m^K > 1$, $U\big(\Gamma^*(w; 0, 0, \bar{w}, \bar{w})\big) = V_{\emptyset}(w) - w$ for any $w \geqslant \bar{\theta}^K$ with $V_{\emptyset}$ as defined in (EC.17).

Next, we show that under condition $K \geqslant \bar{K}_e$, functions $V_{\mathsf{I}} = V_e$ and $V_{\emptyset} = \underline{v}$ satisfy (20)–(22). By the definition of V_e, it is clear that $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}} = 0$. Moreover, $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) = r\underline{v} - \underline{\mu} R = 0$ for any $w \geqslant 0$. Hence, (20) holds.

Note that V_e is increasing on $[0, \bar{w}]$ (see Lemma 3 of Sun and Tian 2018). Hence, for any $w \geqslant 0$, we have $V_{\mathsf{I}}(w) - V_{\emptyset}(w) \geqslant V_e(0) - \underline{v} = 0$ and $V_{\mathsf{I}}(w) - V_{\emptyset}(w) \leqslant \bar{V}_e - \underline{v} = \bar{K}_e \leqslant K$. Therefore, (21) holds. Finally, it is evident that (22) holds.

It remains to show that under condition $K < \bar{K}_e$, functions $V_{\mathsf{I}} = V_e$ and $V_\emptyset$ as defined in (EC.17) satisfy (20)–(22). Obviously, $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}} = 0$. Moreover, we have

$$(\mathcal{A}_\emptyset V_\emptyset)(w) = rV_\emptyset(w) - rwV_\emptyset'(w) - \underline{\mu}R = rw\left(\frac{V_\emptyset(w) - V_\emptyset(0)}{w} - V_\emptyset'(w)\right) \geqslant 0,$$

where the equality follows from $V_\emptyset(0) = \underline{v}$, and the inequality follows from the concavity of $V_\emptyset$. Hence, (20) holds.

If $w \geqslant \bar{\theta}^K$, then $V_{\mathsf{I}}(w) - V_\emptyset(w) = K$. If $w \in [0, \bar{\theta}^K]$, then $V_{\mathsf{I}}'(w) - V_\emptyset'(w) = V_e'(w) - V_e'(\bar{\theta}^K) \geqslant 0$ due to the concavity of V_e, which implies that $V_{\mathsf{I}}(w) - V_\emptyset(w) \geqslant V_{\mathsf{I}}(0) - V_\emptyset(0) = 0$ and $V_{\mathsf{I}}(w) - V_\emptyset(w) \leqslant V_{\mathsf{I}}(\bar{\theta}^K) - V_\emptyset(\bar{\theta}^K) = K$. Hence, (21) holds. It is straightforward to see that (22) holds. □

Therefore, in the equal discount case, contract $\Gamma^*(w_e^*; 0, 0, \bar{w}, \bar{w})$ is optimal if $K < \bar{K}_e$ and $m^K > 1$, in which $w_e^* \in [0, \bar{w}]$ is the unique maximizer of function V_e such that $w_e^* > \bar{\theta}^K$. Otherwise, it is optimal for the principal not to hire the agent at all. Note that because the threshold $\underline{\theta}$ in contract $\Gamma^*(w; 0, 0, \bar{w}, \bar{w})$ is zero, the principal does not direct the agent to stop working until the promised utility has reached 0. At this point, the promised utility cannot become positive again, and the contract is terminated. Therefore, in all these cases, it is never optimal for the principal to direct the agent to stop working and restart later.

EC.1.4. Effect of Arrival Rate Under Fixed Revenue Rate

In this section, we investigate the effect of arrival uncertainty on the optimal contract. In particular, we fix the revenue rates per unit of time ($R\mu$ and $R\underline{\mu}$), the cost rates (c and b), and the switching cost (K), and see how the optimal contract changes with the revenue R. In particular, when R approaches zero, the arrival rate effectively approaches infinity, and the system behaves more like a deterministic one. In this case, mitigating uncertainty effectively removes the rent that the agent is able to obtain. The system should become efficient. On the flip side, if R approaches infinity, the system is extremely uncertain.

For this purpose, we fix $A := R\Delta\mu$ and $B := R\underline{\mu}$ and let them be fixed input parameters. In this setup, we can write all results as well as relevant quantities appeared in the paper in terms of A and B, with $\underline{\mu}$ and $\Delta\mu$ replaced as $\underline{\mu} = B/R$ and $\Delta\mu = A/R$, respectively. It is easy to check that quantities $\underline{v}$, $\bar{w}$, $\bar{V}(\cdot)$, and $\bar{V}$ are all independent of R. Hence, the results in Theorem 2 still hold. Moreover, we have the following result, which explores two extreme cases, the case of extreme uncertainty (i.e., $R \uparrow \infty$) and that of no uncertainty (i.e., $R \downarrow 0$). Note that the first-best societal utility, by considering whether or not to hire the agent, is

$$V^{\mathsf{FB}} := \underline{v} + \left[\frac{R\Delta\mu - c}{r} - K\right]^+.$$

Proposition EC.2. *Fix model parameters A, B, c, b, and K.*

(i) *As $R\uparrow\infty$, it is optimal for the principal to not hire the agent if $K > \bar{V} - \underline{v} - \bar{w}$, and to hire the agent and offer contract $\bar{\Gamma}$ (paying $\beta = bR/A$ to each arrival) otherwise.*

(ii) *As $R\downarrow 0$, it is optimal for the principal not to hire the agent if $K \geqslant (A-c)/r$. If $K < (A-c)/r$, on the other hand, the principal will hire the agent and implement the contract $\Gamma^{*R}\left(\mathsf{w}_0^{*R};\ \underline{\vartheta}^R,\ (\underline{\vartheta}^R \vee \breve{w}(\hat{\mathbf{w}}^R)),\ \bar{\vartheta}^R,\ \hat{\mathbf{w}}^R\right)$ as defined in (47) and Theorem 5, in which the superscript R highlights the parameters' dependence on R. Furthermore, we have*

$$\lim_{R\downarrow 0} \mathsf{w}_0^{*R} = \lim_{R\downarrow 0} \underline{\vartheta}^R = \lim_{R\downarrow 0} \breve{w}(\hat{\mathbf{w}}^R) = \lim_{R\downarrow 0} \bar{\vartheta}^R = \lim_{R\downarrow 0} \hat{\mathbf{w}}^R = 0, \tag{EC.18}$$

and

$$\lim_{R\downarrow 0} U^R\Big(\Gamma^{*R}\left(\mathsf{w}_0^{*R};\ \underline{\vartheta}^R,\ (\underline{\vartheta}^R \vee \breve{w}(\hat{\mathbf{w}}^R)),\ \bar{\vartheta}^R,\ \hat{\mathbf{w}}^R\right)\Big) = \frac{A+B-c}{r} - K = V^{FB}. \tag{EC.19}$$

In either case, the optimal contract yields the first-best societal utility asymptotically.

Proof. First, we show part (i). Note that $R \geqslant \hat{R}$ if and only if

$$R \geqslant \frac{(A+B)[(A-c)A\rho - b(\rho-r)(A+B)]}{\rho(\rho-r)(A^2 - cA - bA - bB)} =: \breve{R}.$$

Hence, Condition 2 holds as $R\uparrow\infty$. Consequently, we have $\lim_{R\uparrow\infty} \bar{K} = \lim_{R\uparrow\infty} \bar{K}_2 = \bar{V} - \underline{v} - \bar{w}$ and $\lim_{R\uparrow\infty} \underline{K} = 0$ by (34). The result stated in part (i) follows immediately from Theorem 2.

Next, we prove part (ii). Fix any contract $\Gamma \in \mathfrak{C}$. Define $\sigma := \inf\{t \geqslant 0 \mid \mathcal{E}_t = \mathsf{I}\}$, which will take value ∞ if the principal does not hire the agent under contract Γ. We have

$$\begin{aligned}
U(\Gamma) &\leqslant \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\infty e^{-rt}\big(R\mathrm{d}N_t - c\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\mathrm{d}t\big) - \sum_{0\leqslant t\leqslant\infty} e^{-rt}\kappa(\mathcal{E}_{t-},\mathcal{E}_t)\right] \\
&= \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\infty e^{-rt}\Big(R(\mu\mathbb{1}_{\mathcal{E}_t=\mathsf{I}} + \underline{\mu}\mathbb{1}_{\mathcal{E}_t=\emptyset}) - c\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\Big)\mathrm{d}t - \sum_{0\leqslant t\leqslant\infty} e^{-rt}\kappa(\mathcal{E}_{t-},\mathcal{E}_t)\right] \\
&\leqslant \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\sigma e^{-rt}R\underline{\mu}\mathrm{d}t + \int_\sigma^\infty e^{-rt}(R\mu - c)\mathrm{d}t - e^{-r\sigma}K\right] \\
&= \frac{R\underline{\mu}}{r} + \mathbb{E}^{\bar{\nu}(\Gamma)}\left[e^{-r\sigma}\right]\left(\frac{R\Delta\mu - c}{r} - K\right) \\
&= \frac{R\underline{\mu}}{r} + \left(\frac{R\Delta\mu - c}{r} - K\right)^+,
\end{aligned}$$

where the first inequality follows by plugging (LL) into (6), and the second inequality follows from Assumption 1. Therefore, if $K \geqslant (R\Delta\mu - c)/r = (A-c)/r$, then we have $U(\Gamma) \leqslant R\underline{\mu}/r = U(\underline{\Gamma})$, which demonstrates that it is optimal for the principal not to hire the agent. (We point out this result does not depend on the value of R.)

If $K < (A-c)/r$, then we have

$$U(\Gamma) \leqslant \bar{V}(0) - K = (R\mu - c)/r - K = (A+B-c)/r - K. \tag{EC.20}$$

Denote the set of positive R's that satisfy Condition 1 as $\mathcal{R}$. Clearly, $R \in \mathcal{R}$ when it is sufficiently small. For any $R \in \mathcal{R}$, Lemma 3 holds, which demonstrates that $\widehat{\mathrm{w}}^R$ is well defined.

To show the second assertion in part (ii), we need the following limiting result:

$$\lim_{R \downarrow 0} \widehat{\mathrm{w}}^R = 0, \tag{EC.21}$$

which will be proved later using a contradictory argument. This result further implies that $\lim_{R\downarrow 0} \bar{K}_1^R = \bar{V}(0) - \underline{v} = (A-c)/r$. In fact, note that the line $w + \underline{v} + \bar{K}_1^R$ (as a function of w) is above the curve $\mathcal{V}^R_{\widehat{\mathrm{w}}^R}(w)$ for any $R \in \mathcal{R}$. Hence, we have $\bar{K}_1^R \geqslant \mathcal{V}^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - \widehat{\mathrm{w}}^R - \underline{v} = \bar{V}(\widehat{\mathrm{w}}^R) - \widehat{\mathrm{w}}^R - \underline{v}$. In addition, we have $\bar{K}_1^R \leqslant \bar{V}(\widehat{\mathrm{w}}^R) - \underline{v}$. Sending R to zero and using (EC.21), we obtain that $\lim_{R\downarrow 0} \bar{K}_1^R = (A-c)/r$. Consequently, Condition 3 holds as $R \downarrow 0$ if $K < (A-c)/r$, which demonstrates that the contract $\Gamma^{*R}\left(\mathsf{w}_0^{*R};\ \underline{\vartheta}^R,\ (\underline{\vartheta}^R \vee \breve{w}(\hat{\mathbf{w}}^R)),\ \bar{\vartheta}^R,\ \hat{\mathbf{w}}^R\right)$ as defined in (47) and Theorem 5 is well defined, establishing the second assertion in part (ii).

The limiting result (EC.18) follows immediately by noting that $\hat{\mathbf{w}}^R < \widehat{\mathrm{w}}^R$ from Proposition 6 and using (EC.21). Applying Proposition 7 and Theorem 5, we obtain that

$$\begin{aligned} U^R\left(\hat{\Gamma}^R(\mathsf{w}_0^{*R})\right) &= \mathcal{V}^R_{\hat{\mathbf{w}}^R}(\mathsf{w}_0^{*R}) - \mathsf{w}_0^{*R} - K = \max_{w \geqslant 0}\left\{\mathcal{V}^R_{\hat{\mathbf{w}}^R}(w) - w\right\} - K \\ &\geqslant \mathcal{V}^R_{\hat{\mathbf{w}}^R}(\hat{\mathbf{w}}^R) - \hat{\mathbf{w}}^R - K = \bar{V}(\hat{\mathbf{w}}^R) - \hat{\mathbf{w}}^R - K, \end{aligned}$$

which further implies that

$$\liminf_{R\downarrow 0} U^R\left(\hat{\Gamma}^R(\mathsf{w}_0^{*R})\right) \geqslant \lim_{R\downarrow 0}\left\{\bar{V}(\hat{\mathbf{w}}^R) - \hat{\mathbf{w}}^R\right\} - K = \bar{V}(0) - K$$

by (EC.21). This, combining with (EC.20), establishes (EC.19).

It remains to show (EC.21). Note that $\widehat{\mathrm{w}}^R \in [0, \bar{w})$ for any $R \in \mathcal{R}$. Hence, $\{\widehat{\mathrm{w}}^R\}_{R\in\mathcal{R}}$ is a bounded sequence. If $\lim_{R\downarrow 0} \widehat{\mathrm{w}}^R = 0$ fails to hold, according to the Bolzano–Weierstrass theorem, there exists a sequence $\{R_n\}_{n\in\mathbb{N}}$ with $R_n \in \mathcal{R}$ and $\lim_{n\to\infty} R_n = 0$, and a number $w^\ddagger \in (0, \bar{w}]$, such that $\lim_{n\to\infty} \widehat{\mathrm{w}}^{R_n} = w^\ddagger$. Then, we show that

$$\lim_{n\to\infty} V^{R_n}_{\widehat{\mathrm{w}}^{R_n}}(w) = -\infty \tag{EC.22}$$

for any $w \in [0, w^\ddagger)$. Suppose, to the contradictory, that (EC.22) fails to hold for some $w^\dagger \in [0, w^\ddagger)$. Then, we have a subsequence $\{R_{n'}\}_{n'\in\mathbb{N}}$ with $\lim_{n\to\infty} R_{n'} = 0$ such that $\lim_{n'\to\infty} V^{R_{n'}}_{\widehat{\mathrm{w}}^{R_{n'}}}(w^\dagger)$ exists and is finite. Recall from Lemma 2 that $V^R_{\widehat{\mathrm{w}}^R}(\cdot)$ is continuous and increasing. Using a diagonalization argument, we can show that there exists a further subsequence $\{R_{n''}\} \subset \{R_{n'}\}_{n'\in\mathbb{N}}$ and a finite-valued continuous function $v(\cdot)$ defined on $[w^\dagger, w^\ddagger]$ such that

$$\lim_{n'\to\infty} V^{R_{n''}}_{\widehat{\mathrm{w}}^{R_{n''}}}(w) = v(w) \tag{EC.23}$$

for any $w \in [w^\dagger, w^\ddagger]$. (First, we establish the weakly convergence of these functions at all rational numbers on $[w^\dagger, w^\ddagger]$; then we use these functions' continuity and monotonicity to show the weakly convergence on the entire interval $[w^\dagger, w^\ddagger]$.) Moreover, $v(\cdot)$ is nondecreasing on $[w^\dagger, w^\ddagger]$, with $v(w^\ddagger) = \bar{V}(w^\ddagger)$.

Rewriting (25) in terms of A, B with μ and β replaced, we obtain

$$\rho\,(\bar{w}-w)\,(V^R_{\widehat{\mathrm{w}}^R})'(w) + rV^R_{\widehat{\mathrm{w}}^R}(w) - (A+B-c) + (\rho-r)w = \frac{A+B}{R}\left[V^R_{\widehat{\mathrm{w}}^R}\Big(\big(w+\frac{bR}{A}\big)\wedge \widehat{\mathrm{w}}^R\Big) - V^R_{\widehat{\mathrm{w}}^R}(w)\right],$$

or equivalently,

$$\begin{aligned}&\rho\frac{\mathrm{d}}{\mathrm{d}w}\big[(\bar{w}-w)V^R_{\widehat{\mathrm{w}}^R}(w)\big]\\ &\quad=\frac{A+B}{R}\left[V^R_{\widehat{\mathrm{w}}^R}\Big(\big(w+\frac{bR}{A}\big)\wedge \widehat{\mathrm{w}}^R\Big) - V^R_{\widehat{\mathrm{w}}^R}(w)\right] - (\rho+r)V^R_{\widehat{\mathrm{w}}^R}(w) + (A+B-c) - (\rho-r)w.\end{aligned}$$

Integrating the above equation from w to $\widehat{\mathrm{w}}^R$ yields

$$\begin{aligned}&\rho\Big[(\bar{w}-\widehat{\mathrm{w}}^R)V^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - (\bar{w}-w)V^R_{\widehat{\mathrm{w}}^R}(w)\Big]\\ &\quad=\int_w^{\widehat{\mathrm{w}}^R}\left\{\frac{A+B}{R}\left[V^R_{\widehat{\mathrm{w}}^R}\Big(\big(u+\frac{bR}{A}\big)\wedge \widehat{\mathrm{w}}^R\Big) - V^R_{\widehat{\mathrm{w}}^R}(u)\right] - (\rho+r)V^R_{\widehat{\mathrm{w}}^R}(u) + (A+B-c) - (\rho-r)u\right\}\mathrm{d}u.\end{aligned} \tag{EC.24}$$

Note that

$$\int_w^{\widehat{\mathrm{w}}^R}\left[V^R_{\widehat{\mathrm{w}}^R}\left(\big(u+\frac{bR}{A}\big)\wedge \widehat{\mathrm{w}}^R\right) - V^R_{\widehat{\mathrm{w}}^R}(u)\right]\mathrm{d}u = \begin{cases}\int_w^{\widehat{\mathrm{w}}^R}(V^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - V^R_{\widehat{\mathrm{w}}^R}(u))\mathrm{d}u, & w\in(\widehat{\mathrm{w}}^R-\frac{bR}{A},\widehat{\mathrm{w}}^R],\\ \frac{bR}{A}V^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - \int_w^{w+\frac{bR}{A}} V^R_{\widehat{\mathrm{w}}^R}(u)\mathrm{d}u, & w\in[0,\widehat{\mathrm{w}}^R-\frac{bR}{A}].\end{cases}$$

Now consider Equation (EC.24) for the subsequence $\{R_{n''}\}$ and for any $w\in[w^\dagger,w^\ddagger)$. As $w\in[0,\widehat{\mathrm{w}}^R-\frac{bR}{A}]$ for sufficiently small R, by L'Hopital's rule and (EC.23), we have

$$\lim_{n''\to\infty}\frac{\int_w^{\widehat{\mathrm{w}}^{R_{n''}}}\left[V^{R_{n''}}_{\widehat{\mathrm{w}}^{R_{n''}}}\left(\big(u+\frac{bR}{A}\big)\wedge\widehat{\mathrm{w}}^{R_{n''}}\right) - V^{R_{n''}}_{\widehat{\mathrm{w}}^{R_{n''}}}(u)\right]\mathrm{d}u}{R_{n''}} = \frac{b}{A}\bar{V}(w^\ddagger) - \frac{b}{A}v(w).$$

Therefore, letting $n''\to\infty$ in (EC.24) and applying (EC.23), we obtain

$$\begin{aligned}&\rho\big[(\bar{w}-w^\ddagger)\bar{V}(w^\ddagger) - (\bar{w}-w)v(w)\big]\\ &\quad=\frac{b(A+B)}{A}(\bar{V}(w^\ddagger)-v(w)) + \int_w^{w^\ddagger}\left[-(\rho+r)v(u) + (A+B-c) - (\rho-r)u\right]\mathrm{d}u.\end{aligned}$$

Therefore, v is differentiable, and thus the above equality can be written as

$$\rho w v'(w) = rv(w) - (A+B-c) + (\rho - r)w,$$

by noting that $\bar{w} = b(A+B)/(\rho A)$.

Using the boundary condition $v(w^\ddagger) = \bar{V}(w^\ddagger)$, we have

$$v(w) = \bar{V}(w^\ddagger) + w - w^\ddagger + \frac{\rho}{r} w^\ddagger \left[1 - \left(\frac{w}{w^\ddagger}\right)^{r/\rho}\right] \text{ for } w \in [w^\dagger, w^\ddagger],$$

which is decreasing on $[w^\dagger, w^\ddagger]$, reaching a contradiction with the fact that $v(\cdot)$ is nondecreasing on $[w^\dagger, w^\ddagger]$. Hence, (EC.22) holds.

Furthermore, we have $\lim_{n\to\infty} \mathcal{V}^{R_n}_{\hat{\mathrm{w}}^{R_n}}(0) = -\infty$ by noting that $\mathcal{V}_{\tilde{w}} \leqslant V_{\tilde{w}}$ for any $\tilde{w} \in (0, \bar{w})$. This contradicts $\mathcal{V}^{R}_{\hat{\mathrm{w}}^{R}}(0) = \underline{v}$. The proof of (EC.21) is complete. □

As R approaches infinity, the arrival stream is extremely uncertain, and thus it is hard for the principal to distinguish whether the agent exerts effort or not. Hence, it is expected that the promised utility plays little role in the incentive and thus payment should be made completely based on whether an arrival occurs or not. Part (i) of Proposition EC.2 validates this intuition.

Part (ii) of Proposition EC.2 states the result for another extreme case. As R approaches zero, there is essentially no arrival uncertainty. In the absence of information asymmetry (in term of the agent's effort rate), the system's first best can be achieved. In fact, the first-best societal utility is $V^{\mathtt{FB}}$, which indeed is asymptotically achieved under the proposed contract.

EC.2. Proofs of the Results in Sections 2 and 3

EC.2.1. Proof of Proposition 1

The proof of part (i) is exactly the same as that of Proposition 1 in Cao et al. (2022), in which random termination instead of random switching may take place. The proof of part (ii) is similar to that of Lemma 6 in Sun and Tian (2018). To keep this paper self-contained, we provide a complete proof here.

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

$$\begin{aligned} u_t(\Gamma, \nu) &:= \mathbb{E}^{\nu, q}\left[\int_0^\infty e^{-\rho s}\left(\mathrm{d}L_s - b\mathbb{1}_{\nu_s = \mu}\mathrm{d}s\right) \middle| \mathcal{F}_t\right] \\ &= \int_0^t e^{-\rho s}\left(\mathrm{d}L_s - b\mathbb{1}_{\mu_s = \mu}\mathrm{d}s\right) + e^{-\rho t} W_t(\Gamma, \nu). \end{aligned} \tag{EC.25}$$

In what follows, we omit (Γ, ν) from all relevant quantities for the sake of easing notation. 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 $\mathcal{I}^N := \mathcal{I}^N_{[0,\infty)}$. Similarly, we use $\mathcal{I}^Q_{[t_1,t_2]}$ to denote the set of randomized switching time epochs during $[t_1, t_2]$ under the switching intensity process $\{q_t\}_{t\geqslant 0}$. Moreover, we denote $\mathcal{I}_t^Q := \mathcal{I}_{[0,t]}^Q$ and $\mathcal{I}^Q := \mathcal{I}_{[0,\infty)}^Q$.

At any time instant $\zeta-$, $W_{\zeta-}$ can jump to W_ζ^N triggered by an arrival at time ζ, or jump to W_ζ^Q triggered by a randomized switching, or jump to W_ζ^L triggered by an instantaneous payment. (Here, the agent's promised utility will not jump caused by a deterministic switching.) Therefore, we can decompose W_ζ (for $\zeta > t$) into its discrete part

$$\sum_{t\leqslant\xi\leqslant\zeta}\left[(W_\xi^N - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^N} + (W_\xi^Q - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^Q} + (W_\xi^L - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^L}\right]$$

and its absolutely continuous part

$$W_\zeta^c := W_\zeta - \sum_{t\leqslant\xi\leqslant\zeta}\left[(W_\xi^N - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^N} + (W_\xi^Q - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^Q} + (W_\xi^L - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^L}\right],$$

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

According to the definition of admissible contract, we know that both W_t^N and W_t^Q is $\mathcal{F}_t$-predictable. However, W_t^L can also depend on $\mathrm{d}N_t$ and $\mathrm{d}Q_t$, that is, W_t^L is $\mathcal{F}_t$-adaptive.

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

$$\begin{aligned} e^{-\rho t'}W_{t'} - e^{-\rho t}W_t = &\int_t^{t'} e^{-\rho\zeta}\big(-\rho W_\zeta \mathrm{d}\zeta + \mathrm{d}W_\zeta^c\big) \\ &+ \sum_{\zeta\in(t,t']} e^{-\rho\zeta}\left[(W_\zeta^N - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^N} + (W_\zeta^Q - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^Q} + (W_\zeta^L - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^L}\right]. \end{aligned} \tag{EC.26}$$

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

$$\begin{aligned} 0 =& \mathbb{E}_t[u_{t'}] - u_t \\ =& \mathbb{E}_t[e^{-\rho t'}W_{t'} - e^{-\rho t}W_t] + \mathbb{E}_t\left[\int_{t+}^{t'} e^{-\rho\zeta}\big(\mathrm{d}L_\zeta - b\mathbb{1}_{\nu_\zeta=\mu}\mathrm{d}\zeta\big)\right] \\ =& \mathbb{E}_t\left[\int_t^{t'} e^{-\rho\zeta}\big(-\rho W_\zeta\mathrm{d}\zeta + \mathrm{d}W_\zeta^c\big)\right] \\ &+ \mathbb{E}_t\left\{\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\left[(W_\zeta^N - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^N} + (W_\zeta^Q - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^Q} + (W_\zeta^L - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^L}\right]\right\} \\ &+ \mathbb{E}_t\left[\int_{t+}^{t'} e^{-\rho\zeta}\big(\mathrm{d}L_s - b\mathbb{1}_{\nu_\zeta=\mu}\mathrm{d}\zeta\big)\right] \\ =& \mathbb{E}_t\left\{\int_t^{t'} e^{-\rho\zeta}\Big\{\big[-\rho W_\zeta + (W_\zeta^N - W_{\zeta-})\nu_\zeta + (W_\zeta^Q - W_{\zeta-})q_\zeta\big]\mathrm{d}\zeta + \mathrm{d}W_\zeta^c\Big\}\right. \end{aligned}$$

$$+\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\Big[(W^L_\zeta - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}^L_{(t,t']}}\Big]\Bigg\} + \mathbb{E}_t\left[\int_{t+}^{t'} e^{-\rho\zeta}(\mathrm{d}L_\zeta - b\mathbb{1}_{\nu_\zeta=\mu}\mathrm{d}\zeta)\right],$$

where the second equality follows from (EC.25), and the third from (EC.26). The fourth equality follows from the facts that $\{Q_t\}_{t\geqslant 0}$ is a counting process with intensity q_t, and that N_t is a counting process with intensity ν_t, as well as Lemma L3 in Chapter II of Brémaud (1981), noting that

$$\mathbb{E}_t\int_t^{t'} e^{-\rho\zeta}|(W^N_\zeta - W_{\zeta-})\nu_\zeta|\mathrm{d}\zeta \leqslant \bar{W}\mu\int_t^{t'} e^{-\rho\zeta}\mathrm{d}\zeta < \infty, \text{ and} \tag{EC.27}$$

$$\mathbb{E}_t\int_t^{t'} e^{-\rho\zeta}|(W^Q_\zeta - W_{\zeta-})q_\zeta|\mathrm{d}\zeta \leqslant \bar{W}\mathbb{E}_t\int_t^{\infty} e^{-\rho\zeta}q_\zeta\mathrm{d}\zeta \leqslant \bar{W}\mathbb{E}_t\int_t^{\tau} e^{-r\zeta}q_\zeta\mathrm{d}\zeta < \infty, \tag{EC.28}$$

in view of (WU), $\rho > r$, and (1).

Recall that $\mathrm{d}L_t = \ell_t\mathrm{d}t + \Delta L_t$. For any $t < t' < \tau$, the above equality can be stated as

$$\mathbb{E}_t\left\{\int_t^{t'} e^{-\rho\zeta}\big[-\rho W_\zeta + (W^N_\zeta - W_{\zeta-})\nu_\zeta + (W^Q_\zeta - W_\zeta -)q_\zeta - b\mathbb{1}_{\nu_\zeta=\mu} + \ell_\zeta\big]\mathrm{d}\zeta + \mathrm{d}W^c_\zeta\right\}$$
$$+\mathbb{E}_t\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\big[(W^L_\zeta - W_{\zeta-})\mathbb{1}_{\Delta L_\zeta>0} + \Delta L_\zeta\big] = 0. \tag{EC.29}$$

Consider any time t. Letting $t' \downarrow t$ in (EC.29) yields

$$\mathbb{E}_t[(W^L_t - W_{t-})\mathbb{1}_{\Delta L_t>0} + \Delta L_t] = 0, \tag{EC.30}$$

which further implies

$$\mathrm{d}W^c_t = \big[\rho W_{t-} - (W^N_t - W_{t-})\nu_t - (W^Q_t - W_{t-})q_t + b\mathbb{1}_{\nu_t=\mu} - \ell_t\big]\mathrm{d}t,\ t\geqslant 0. \tag{EC.31}$$

Let $H_t := W^N_t - W_{t-}$ and $H^q_t := -W^Q_t + W_{t-}$. Then, both H_t and H^q_t are $\mathcal{F}_t$-predictable. Besides, since W^L_t is $\mathcal{F}_t$-adaptive, (EC.30) in fact is equivalent to

$$(W^L_t - W_{t-})\mathbb{1}_{\Delta L_t>0} + \Delta L_t = 0. \tag{EC.32}$$

We also have

$$\mathrm{d}W_t = \mathrm{d}W^c_t + (W^N_t - W_{t-})\mathrm{d}N_t + (W^Q_t - W_{t-})\mathrm{d}Q_t + (W^L_t - W_{t-})\mathbb{1}_{\Delta L_t>0}. \tag{EC.33}$$

Combining (EC.31)–(EC.33), we obtain (PK).

Relationship (5) follows immediately by noting $W^N_t \geqslant 0$ and $W^Q_t \geqslant 0$ for all $t \geqslant 0$.

(ii) Let $\widetilde{u}_t(\Gamma,\nu',\nu)$ denote the agent's total expected discounted utility conditional on $\mathcal{F}_t$ under contract Γ, when he follows effort process $\nu' = \{\nu'_t\}_{t\geqslant 0}$ before time t and then effort process ν after time t:

$$\widetilde{u}_t(\Gamma,\nu',\nu) = \int_0^t e^{-\rho s}\big(\mathrm{d}L_s - b\mathbb{1}_{\nu'_s=\mu}\mathrm{d}s\big) + e^{-\rho t}W_t(\Gamma,\nu). \tag{EC.34}$$

Here, $\widetilde{u}_{0-}(\Gamma,\nu',\nu)$ can be interpreted in a similar vein as that for $W_{0-}(\Gamma,\nu)$. In fact, we have $\widetilde{u}_{0-}(\Gamma,\nu',\nu)=W_{0-}(\Gamma,\nu)=u(\Gamma,\nu)$. In what follows, we write $\bar{\nu}$ instead of $\bar{\nu}(\Gamma)$ to ease notation. By the above definition, we have

$$\widetilde{u}_t(\Gamma,\nu,\bar{\nu})=u_t(\Gamma,\bar{\nu})+\int_0^t e^{-\rho s}b(\mathbb{1}_{\bar{\nu}_s=\mu}-\mathbb{1}_{\nu_s=\mu})\mathrm{d}s. \tag{EC.35}$$

Besides, by (PK) and (EC.25), we obtain that

$$\begin{aligned}\mathrm{d}u_t(\Gamma,\nu)&=e^{-\rho t}\big(\mathrm{d}L_t-b\mathbb{1}_{\mu_t=\mu}\mathrm{d}t\big)+e^{-\rho t}\big(\mathrm{d}W_t(\Gamma,\nu)-\rho W_t(\Gamma,\mu)\mathrm{d}t\big)\\&=e^{-\rho t}\big[H_t(\Gamma,\nu)(\mathrm{d}N_t-\nu_t\mathrm{d}t)-H^q_t(\Gamma,\nu)(\mathrm{d}Q_t-q_t\mathrm{d}t)\big].\end{aligned} \tag{EC.36}$$

Therefore, for any time points $t<t'$, we have (below, we add superscript ν in some expectation operators, to indicate that the related random variables are induced by the effort process ν)

$$\begin{aligned}\mathbb{E}_t[\widetilde{u}_{t'}(\Gamma,\nu,\bar{\nu})]-\widetilde{u}_t(\Gamma,\nu,\bar{\nu})&=\mathbb{E}_t[u_{t'}(\Gamma,\bar{\nu})]-u_t(\Gamma,\bar{\nu})+\mathbb{E}^{\nu}_t\left[\int_t^{t'}e^{-\rho s}b(\mathbb{1}_{\bar{\nu}_s=\mu}-\mathbb{1}_{\nu_s=\mu})\mathrm{d}s\right]\\&=\mathbb{E}^{\nu}_t\left[\int_{t+}^{t'}e^{-\rho s}\big(H_s(\Gamma,\bar{\nu})(\mathrm{d}N_s-\bar{\nu}_t\mathrm{d}s)-H^q_s(\Gamma,\bar{\nu})(\mathrm{d}Q_s-q_s\mathrm{d}t)+b(\mathbb{1}_{\bar{\nu}_s=\mu}-\mathbb{1}_{\nu_s=\mu})\mathrm{d}s\big)\right]\\&=\mathbb{E}^{\nu}_t\left[\int_t^{t'}e^{-\rho s}\Big(H_s(\Gamma,\bar{\nu})(\nu_s-\bar{\nu}_s)+b(\mathbb{1}_{\bar{\nu}_s=\mu}-\mathbb{1}_{\nu_s=\mu})\Big)\mathrm{d}s\right],\end{aligned} \tag{EC.37}$$

where the first equality follows from (EC.35) and the second equality follows from (EC.36). The last equalities uses the fact that conditional on $\mathcal{F}_t$ and under effort process ν, $\{N_s\}_{s\in(t,t']}$ and $\{Q_s\}_{s\in(t,t']}$ are counting processes with intensities ν_s and q_s respectively, which follows by applying Lemma L3 in Chapter II of Brémaud (1981) with the aid of (EC.27) and (EC.28).

Since both ν and $\bar{\nu}$ are admissible, we have $\nu_t=\bar{\nu}_t=\underline{\mu}$ whenever $\mathcal{E}_t=\emptyset$. Hence, we have

$$H_s(\Gamma,\bar{\nu})(\nu_s-\bar{\nu}_s)+b(\mathbb{1}_{\bar{\nu}_s=\mu}-\mathbb{1}_{\nu_s=\mu})=-(H_s(\Gamma,\bar{\nu})-\beta)\Delta\mu\mathbb{1}_{\mathcal{E}_s=\mathsf{I},\nu_s=\underline{\mu}} \tag{EC.38}$$

for any $s\geqslant 0$. Therefore, if (IC) holds, then we have $\mathbb{E}_t[\widetilde{u}_{t'}(\Gamma,\nu,\bar{\nu})]\leqslant\widetilde{u}_t(\Gamma,\nu,\bar{\nu})$ by (EC.37) and (EC.38), which implies that $\{\widetilde{u}_t(\Gamma,\nu,\bar{\nu})\}_{t\geqslant 0}$ is an $\mathcal{F}$-supermatingale. By (IR) and (WU), we can add $\widetilde{u}_{\infty}(\Gamma,\nu,\bar{\nu}):=\int_0^{\infty}e^{-\rho s}\big(\mathrm{d}L_s-b\mathbb{1}_{\nu_s=\mu}\mathrm{d}s\big)$ as the last element of this supermartingale. Therefore,

$$u(\Gamma,\bar{\nu})=\widetilde{u}_{0-}(\Gamma,\nu,\bar{\nu})\geqslant\mathbb{E}[\widetilde{u}_{\infty}(\Gamma,\nu,\bar{\nu})]=u(\Gamma,\nu),$$

implying the incentive compatibility of $\bar{\nu}$ under contract Γ.

If (IC) fails to hold, then we consider an effort process ν such that $\nu_t=\mu$ if and only if $H_t(\Gamma,\bar{\nu})\geqslant\beta$ and $\mathcal{E}_t=\mathsf{I}$. Clearly, ν is admissible, and the expression in (EC.38) becomes $-(H_s(\Gamma,\bar{\nu})-\beta)\Delta\mu\mathbb{1}_{\mathcal{E}_s=\mathsf{I},H_s(\Gamma,\bar{\nu})<\beta}$, which is always non-negative and positive on a set of positive measure. Thus, by (EC.37), there exists a time $t>0$ such that $\mathbb{E}_{0-}[\widetilde{u}_t(\Gamma,\nu,\bar{\nu})]>\widetilde{u}_{0-}(\Gamma,\nu,\bar{\nu})=u(\Gamma,\bar{\nu})$. Define another effort process ν' which follows ν until time t and then switches to $\bar{\nu}$, which is also admissible. Moreover, we have $u(\Gamma,\nu')=\mathbb{E}_{0-}[\widetilde{u}_t(\Gamma,\nu,\bar{\nu})]$, which indicates $u(\Gamma,\nu')>u(\Gamma,\bar{\nu})$, contradicting (3). The proof is complete.

EC.2.2. Proof of Lemma 1

If we can show that (PK) holds under contract $\Gamma^*(w_0;\underline{\theta},\breve{w},\bar{\theta},\hat{w})$, then (17) follows immediately from (2) and (4) with $t=0$. In fact, (PK) holds by setting $H_t=\beta\mathbb{1}_{\mathcal{E}_{t-}=\mathsf{I}}$ and $H_t^q=(\breve{w}-\underline{\theta})\,\mathbb{1}_{W_{t-}=\breve{w},\mathcal{E}_{t-}=\mathsf{I}}$.

EC.2.3. Proof of Theorem 1

Fix any contract $\Gamma\in\mathfrak{C}$. The agent's promised utility follows a process W with its dynamics described by (PK) with $\nu_t=\mu$ for $\mathcal{E}_t=\mathsf{I}$ and $\nu_t=\underline{\mu}$ for $\mathcal{E}_t=\emptyset$.

Recall that $\mathrm{d}L_t=\ell_t\mathrm{d}t+\Delta L_t$. Write $\phi(w,\varepsilon)=V_\varepsilon(w)-w$ for any $w\in\mathbb{R}_+$ and $\varepsilon\in\{\mathsf{I},\emptyset\}$. Applying the change-of-variable formula (see, for example, Theorem 70 of Chapter IV in Protter 2003, pp. 214) for processes of locally bounded variation to the process $(W,\mathcal{E})$ and using (PK), we have

$$
\begin{aligned}
e^{-rT}\phi(W_T,\mathcal{E}_T)=\phi(W_{0-},\mathcal{E}_{0-})+\int_{0+}^{T}e^{-rt}\Big[&(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t+q_tH_t^q-\ell_t)\cdot D_{t-}\\
&-rV_{\mathcal{E}_{t-}}(W_{t-})\Big]\mathrm{d}t+\sum_{0\leqslant t\leqslant T}e^{-rt}\Delta\phi(W_t,\mathcal{E}_t)
\end{aligned}
$$

for any $T\geqslant 0$, where D_{t-} is the left derivative of $\phi(w,\mathcal{E}_{t-})$ with respect to w at W_{t-}, that is, $D_{t-}=V'_{\mathcal{E}_{t-}}(W_{t-})-1$, by recalling that we use $f'(w)$ to represent the left derivative of f at w for any absolutely continuous function defined on $\mathbb{R}_+$. Besides, we have

$$
\begin{aligned}
\Delta\phi(W_t,\mathcal{E}_t)\;&=\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})\\
&\quad+\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-})-\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})\\
&\quad+\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\ \text{for } t>0,
\end{aligned}
$$

and

$$
\Delta\phi(W_0,\mathcal{E}_0)=\phi(W_0,\mathcal{E}_0)-\phi(W_0,\mathcal{E}_{0-})+\phi(W_0,\mathcal{E}_{0-})-\phi(W_{0-},\mathcal{E}_{0-})
$$

by noting that $\mathrm{d}N_0=\mathrm{d}Q_0=0$ with probability 1.

Define $M^N=\{M_t^N\}_{t\geqslant 0}$ and $M^Q=\{M_t^Q\}_{t\geqslant 0}$ by

$$
M_t^N=N_t-\int_0^t\nu_s\mathrm{d}s\quad\text{and}\quad M_t^Q=Q_t-\int_0^t q_s\mathrm{d}s.
$$

Note that

$$
\begin{aligned}
&\sum_{0<t\leqslant T}\big[\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\\
&=\int_{0+}^{T}e^{-rt}\Big\{[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}N_t+[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}Q_t\Big\}\\
&=\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\mathrm{d}M_t^N+\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\nu_t\mathrm{d}t\\
&\quad+\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\mathrm{d}M_t^Q+\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]q_t\mathrm{d}t,
\end{aligned}
$$

where the first equality uses the fact that $\{t \in [0,T] \mid \mathrm{d}N_t = \mathrm{d}Q_t = 1\}$ has a Lebesgue measure 0 with probability 1. Summarizing the above formulas, we obtain

$$\begin{aligned} e^{-rT}\phi(W_T,\mathcal{E}_T) \;&= \phi(W_{0-},\mathcal{E}_{0-}) + \int_{0+}^{T} e^{-rt}[\phi(W_{t-}+H_t,\mathcal{E}_{t-}) - \phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}M_t^N + \\ &\qquad + \int_{0+}^{T} e^{-rt}[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-}) - \phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}M_t^Q + A_1 + A_2 + A_3 + A_4 + A_5, \end{aligned} \tag{EC.39}$$

where

$$\begin{aligned} A_1 &:= \int_{0+}^{T} e^{-rt}\Big\{(\rho W_{t-} + b\mathbb{1}_{\nu_t=\mu} - H_t\nu_t - \ell_{t-})\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1\big) - r\phi(W_{t-},\mathcal{E}_{t-}) \\ &\qquad + [\phi(W_{t-}+H_t,\mathcal{E}_{t-}) - \phi(W_{t-},\mathcal{E}_{t-})]\nu_t\Big\}\mathrm{d}t, \\ A_2 &:= \sum_{0<t\leqslant T} e^{-rt}\Big[\phi\big(W_{t-} + H_t\mathrm{d}N_t - H_t^q\mathrm{d}Q_t - \Delta L_t, \mathcal{E}_{t-}\big) - \phi\big(W_{t-} + H_t\mathrm{d}N_t - H_t^q\mathrm{d}Q_t, \mathcal{E}_{t-}\big)\Big], \\ A_3 &:= \sum_{0\leqslant t\leqslant T} e^{-rt}[\phi(W_t,\mathcal{E}_t) - \phi(W_t,\mathcal{E}_{t-})], \\ A_4 &:= \int_{0+}^{T} e^{-rt}q_t\Big\{H_t^q\big(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1\big) + \phi(W_{t-}-H_t^q,\mathcal{E}_{t-}) - \phi(W_{t-},\mathcal{E}_{t-})\Big\}\mathrm{d}t, \\ A_5 &:= \phi(W_0,\mathcal{E}_{0-}) - \phi(W_{0-},\mathcal{E}_{0-}). \end{aligned}$$

Below we treat each term separately.

Consider first A_1. If $\mathcal{E}_{t-} = \mathsf{I}$, then $\nu_{t-} = \mu$ and $\phi(W_{t-},\mathcal{E}_{t-}) = V_{\mathsf{I}}(W_{t-}) - W_{t-}$. Since the contract Γ is incentive compatible, we have $H_t \geqslant \beta$ by Proposition 1(ii). Consequently, we have

$$\begin{aligned} &(\rho W_{t-} + b\mathbb{1}_{\nu_t=\mu} - H_t\nu_t - \ell_t)\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1\big) - r\phi(W_{t-},\mathcal{E}_{t-}) + [\phi(W_{t-}+H_t,\mathcal{E}_{t-}) - \phi(W_{t-},\mathcal{E}_{t-})]\nu_t \\ &= (\rho W_{t-} + b - H_t\mu - \ell_t)\cdot\big(V_{\mathsf{I}}'(W_{t-}) - 1\big) - r\cdot(V_{\mathsf{I}}(W_{t-}) - W_{t-}) + [V_{\mathsf{I}}(W_{t-}+H_t) - V_{\mathsf{I}}(W_{t-}) - H_t]\cdot\mu \\ &= \rho W_{t-}\cdot\big(V_{\mathsf{I}}'(W_{t-}) - 1\big) - r\cdot(V_{\mathsf{I}}(W_{t-}) - W_t) - (\ell_t - b)\cdot\big(V_{\mathsf{I}}'(W_{t-}) - 1\big) \\ &\quad + \big[V_{\mathsf{I}}(W_{t-}+H_t) - V_{\mathsf{I}}(W_{t-}) - V_{\mathsf{I}}'(W_{t-})H_t\big]\cdot\mu \\ &\leqslant \rho W_{t-}\cdot(V_{\mathsf{I}}'(W_{t-}) - 1) - r\cdot(V_{\mathsf{I}}(W_{t-}) - W_{t-}) + \ell_t - b + [V_{\mathsf{I}}(W_{t-}+\beta) - V_{\mathsf{I}}(W_{t-}) - V_{\mathsf{I}}'(W_{t-})\beta]\cdot\mu \\ &= -\Big[(\mu+r)V_{\mathsf{I}}(W_{t-}) - \mu V_{\mathsf{I}}(W_{t-}+\beta) + \rho(\bar{w} - W_{t-})V_{\mathsf{I}}'(W_{t-}) - (\mu R - c) + (\rho - r)W_{t-}\Big] + \ell_t - [R\mu - (c-b)] \\ &= -(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(W_{t-}) + \ell_t - [R\mu - (c-b)] \\ &\leqslant \ell_t - [R\mu - (c-b)]. \end{aligned}$$

Here, the first inequality follows from (i) $V_{\mathsf{I}}(W_{t-}) \geqslant 0$ (this follows from the fact that V_{I} is nondecreasing) and (ii) $H_t \geqslant \beta$, and $\beta = \arg\max_{h\geqslant\beta}\{V_{\mathsf{I}}(w+h) - V_{\mathsf{I}}(w) - V_{\mathsf{I}}'(w)\cdot h\}$ due to the concavity of V_{I}, and the last inequality follows from (20).

If $\mathcal{E}_{t-}=\emptyset$, then $\nu_{t-}=\underline{\mu}$. It follows from (5) that $H_t\geqslant -W_{t-}$. Therefore, we have

$$\begin{aligned}
&(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t\\
=&(\rho W_{t-}-H_t\underline{\mu}-\ell_t)\cdot\big(V'_{\emptyset}(W_{t-})-1\big)-r\cdot(V_{\emptyset}(W_{t-})-W_{t-})+\big[V_{\emptyset}(W_{t-}+H_t)-V_{\emptyset}(W_{t-})-H_t\big]\cdot\underline{\mu}\\
=&\rho W_{t-}\cdot\big(V'_{\emptyset}(W_{t-})-1\big)-r\cdot(V_{\emptyset}(W_{t-})-W_{t-})-\ell_t\cdot\big(V'_{\emptyset}(W_{t-})-1\big)\\
&+\big[V_{\emptyset}(W_{t-}+H_t)-V_{\emptyset}(W_{t-})-V'_{\emptyset}(W_{t-})H_t\big]\cdot\underline{\mu}\\
\leqslant&\rho W_{t-}\cdot\big(V'_{\emptyset}(W_{t-})-1\big)-r\cdot(V_{\emptyset}(W_{t-})-W_{t-})+\ell_t\\
=&-\Big[rV_{\emptyset}(W_{t-})-\rho W_{t-}\cdot V'_{\emptyset}(W_{t-})+(\rho-r)W_{t-}-R\underline{\mu}\Big]+\ell_t-R\underline{\mu}\\
=&-(\mathcal{A}_{\emptyset}V_{\emptyset})(W_{t-})+\ell_t-R\underline{\mu}\\
\leqslant&\ell_t-R\underline{\mu},
\end{aligned}$$

where the first inequality follows from (i) $V'_{\emptyset}(W_{t-})\geqslant 0$ (this follows from the fact that $V_{\emptyset}$ is nondecreasing) and (ii) $H_t\geqslant -W_{t-}$, and $0=\arg\max_{h\geqslant -w}\{V_{\emptyset}(w+h)-V_{\emptyset}(w)-V'_{\emptyset}(w)\cdot h\}$ due to the concavity of $V_{\emptyset}$, and the last inequality follows from (20).

Combining the above two cases yields

$$\begin{aligned}
&(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t\\
\leqslant&\ \ell_t-[R\nu_t-(c-b)\mathbb{1}_{\nu_t=\mu}]
\end{aligned}\tag{EC.40}$$

for any $t>0$.

Consider next A_2. We have

$$\begin{aligned}
&\phi(W_{t-}+H_t\mathrm{d}N_t-H^q_t\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-})-\phi(W_{t-}+H_t\mathrm{d}N_t-H^q_t\mathrm{d}Q_t,\mathcal{E}_{t-})\\
=&\ V_{\mathcal{E}_{t-}}(W_{t-}+H_t\mathrm{d}N_t-H^q_t\mathrm{d}Q_t-\Delta L_t)-V_{\mathcal{E}_{t-}}(W_{t-}-H^q_t\mathrm{d}Q_t+H_t\mathrm{d}N_t)+\Delta L_t\\
\leqslant&\ \Delta L_t,\ \forall t>0,
\end{aligned}\tag{EC.41}$$

where the inequality follows from the facts that $\Delta L_t\geqslant 0$ and that V_{ε} is nondecreasing for any $\varepsilon\in\{\mathsf{I},\emptyset\}$.

Consider now A_3. By considering four possible value combinations of $(\mathcal{E}_{t-},\mathcal{E}_t)$ and using (21), we have

$$\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})=V_{\mathcal{E}_t}(W_t)-V_{\mathcal{E}_{t-}}(W_t)\leqslant\kappa(\mathcal{E}_{t-},\mathcal{E}_t).\tag{EC.42}$$

Consider next A_4. We have

$$\begin{aligned}
&H^q_t\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)+\phi(W_{t-}-H^q_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\\
=&\ H^q_tV'_{\mathcal{E}_{t-}}(W_{t-})+V_{\mathcal{E}_{t-}}(W_{t-}-H^q_t)-V_{\mathcal{E}_{t-}}(W_{t-})\leqslant 0,
\end{aligned}$$

where the inequality follows from the concavity of V_ε for any $\varepsilon \in \{\mathsf{I}, \emptyset\}$. This, together with $q_t \geqslant 0$, yields

$$A_4 = \int_{0+}^{T} e^{-rt} q_t \Big\{ H_t^q \big(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1 \big) + \phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-}) \Big\} \mathrm{d}t \leqslant 0. \tag{EC.43}$$

Consider finally A_5. It follows from (2) and (4) with $t = 0$ that $\mathbb{E}[W_0 + \Delta L_0] = W_{0-}$. Therefore, we have

$$\begin{aligned} &\mathbb{E}[\phi(W_0, \mathcal{E}_{0-})] - \phi(W_{0-}, \mathcal{E}_{0-}) = \mathbb{E}[V_{\mathcal{E}_{0-}}(W_0)] - V_{\mathcal{E}_{0-}}(W_{0-}) - \big(\mathbb{E}[W_{0-}] - W_{0-}\big) \\ &\leqslant V_{\mathcal{E}_{0-}}\big(\mathbb{E}[W_0]\big) - V_{\mathcal{E}_{0-}}(W_{0-}) + \mathbb{E}[\Delta L_0] \leqslant \mathbb{E}[\Delta L_0], \end{aligned} \tag{EC.44}$$

where the first inequality follows from the concavity of V_ε for any $\varepsilon \in \{\mathsf{I}, \emptyset\}$ and the Jensen's inequality, and the second inequality follows from the facts that V_ε is nondecreasing and that $W_{0-} = \mathbb{E}[W_0 + L_0] \geqslant \mathbb{E}[W_0]$.

Combining (EC.39)–(EC.43), we have

$$\begin{aligned} e^{-rT}\phi(W_T, \mathcal{E}_T) \leqslant & \phi(W_{0-}, \mathcal{E}_{0-}) + \int_{0+}^{T} e^{-rt}[\phi(W_{t-} + H_t, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})]\mathrm{d}M_t^N \\ &+ \int_{0+}^{T} e^{-rt}[\phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})]\mathrm{d}M_t^Q \\ &+ \int_{0+}^{T} e^{-rt}[\ell_t - (R\nu_t - (c-b)\mathbb{1}_{\nu_t = \mu})]\mathrm{d}t + \sum_{0<t\leqslant T} e^{-rt}\Delta L_t \\ &+ \sum_{0\leqslant t\leqslant T} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t) + \phi(W_0, \mathcal{E}_{0-}) - \phi(W_{0-}, \mathcal{E}_{0-}) \end{aligned}$$

for any $T > 0$, which can be displayed as

$$\begin{aligned} \phi(W_{0-}, \mathcal{E}_{0-}) \geqslant & e^{-rT}\phi(W_T, \mathcal{E}_t) - \int_{0}^{T} e^{-rt}\big[\phi(W_{t-} + H_t, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big]\mathrm{d}M_t^N \\ &- \int_{0+}^{T} e^{-rt}\big[\phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big]\mathrm{d}M_t^Q \\ &+ \int_{0+}^{T} e^{-rt}\big(R\mathrm{d}N_t - \mathrm{d}L_t - (c-b)\mathbb{1}_{\mathcal{E}_t = \mathsf{I}}\mathrm{d}t\big) - \sum_{0\leqslant t\leqslant T} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t) \\ &+ \phi(W_{0-}, \mathcal{E}_{0-}) - \phi(W_0, \mathcal{E}_{0-}). \end{aligned}$$

Taking expectation in the above inequality yields

$$\begin{aligned} \phi(W_{0-}, \mathcal{E}_{0-}) \geqslant & \mathbb{E}[e^{-rT}\phi(W_T, \mathcal{E}_T)] - \mathbb{E}\left[\int_{0+}^{T} e^{-rt}\big(\phi(W_{t-} + H_t, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big)\mathrm{d}M_t^N\right] \\ &- \mathbb{E}\left[\int_{0+}^{T} e^{-rt}\big(\phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big)\mathrm{d}M_t^Q\right] \\ &+ \mathbb{E}\left[\int_{0+}^{T} e^{-rt}\big(R\mathrm{d}N_t - \mathrm{d}L_t - (c-b)\mathbb{1}_{\mathcal{E}_t = \mathsf{I}}\mathrm{d}t\big) - \sum_{0\leqslant t\leqslant T} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t)\right] \end{aligned}$$

$$
\begin{aligned}
&+\phi(W_{0-},\mathcal{E}_{0-})-\mathbb{E}\phi(W_0,\mathcal{E}_{0-})\\
\geqslant&\mathbb{E}[e^{-rT}\phi(W_T,\mathcal{E}_T)]-\mathbb{E}\left[\int_{0+}^{T}e^{-rt}\big(\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big)\mathrm{d}M_t^N\right]\\
&-\mathbb{E}\left[\int_{0+}^{T}e^{-rt}\big(\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big)\mathrm{d}M_t^Q\right]\\
&+\mathbb{E}\left[\int_{0}^{T}e^{-rt}\big(R\mathrm{d}N_t-\mathrm{d}L_t-(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\mathrm{d}t\big)-\sum_{0\leqslant t\leqslant T}e^{-rt}\kappa(\mathcal{E}_{t-},\mathcal{E}_t)\right] \qquad \text{(EC.45)}
\end{aligned}
$$

for any $T>0$, where the last inequality follows from (EC.44).

We claim that it suffices to consider the case that

$$
\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]<\infty. \qquad \text{(EC.46)}
$$

Otherwise, we have $\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]=\infty$. It follows from (PK) and (WU) that $\mathrm{d}L_t\geqslant(H_t-\bar{W})^+\mathrm{d}N_t$ for $t>0$. Hence, we have

$$
\begin{aligned}
&\mathbb{E}\left[\int_{0}^{\infty}e^{-rt}\mathrm{d}L_t\right]\geqslant\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}(H_t-\bar{W})^+\mathrm{d}N_t\right]=\mathbb{E}\left[\int_{0+}^{\tau}e^{-rt}(H_t-\bar{W})^+\nu_t\mathrm{d}t\right]\\
&\geqslant\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}(|H_t|-\bar{W})\nu_t\mathrm{d}t\right]\geqslant\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]-\frac{\bar{W}\mu}{r}=\infty,
\end{aligned}
$$

where the first equality follows from Equation (2.3) in Chapter II of Brémaud (1981), the second inequality follows from $H_t\geqslant -W_{t-}\geqslant -R\mu/r$ in view of (5) and (WU), and the third inequality follows from $\nu_t\leqslant\mu$. Then, we have

$$
U(\Gamma)\leqslant\mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_{0}^{\infty}e^{-rt}\big(R\mathrm{d}N_t-\mathrm{d}L_t\big)\right]\leqslant\frac{R\mu}{r}-\mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_{0}^{\infty}e^{-rt}\mathrm{d}L_t\right]=-\infty,
$$

and thus the desired result follows immediately.

Given (EC.46), we have

$$
\begin{aligned}
&\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})|\nu_t\right]\mathrm{d}t\\
&\leqslant\max_{w>0,\varepsilon\in\{\mathsf{I},\emptyset\}}\{|V_\varepsilon'(w)-1|\}\cdot\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]<\infty,
\end{aligned}
$$

where $\max_{w>0,\varepsilon\in\{\mathsf{I},\emptyset\}}\{|V_\varepsilon'(w)-1|\}<\infty$ follows from the concavity of V_ε and the fact that $V_\varepsilon'\geqslant 0$. It follows from Lemma L3, Chapter II in Brémaud (1981) that $\widetilde{M}=\{\widetilde{M}_t\}_{t\geqslant 0}$, defined by

$$
\widetilde{M}_t:=\int_{0+}^{t}e^{-rs}[\phi(W_{s-}+H_s,\mathcal{E}_{s-})-\phi(W_{s-},\mathcal{E}_{s-})]\mathrm{d}M_s^N,
$$

is an $\mathcal{F}$-martingale. Hence, $\mathbb{E}[\widetilde{M}_T]=\mathbb{E}[\widetilde{M}_0]=0$, that is,

$$
\mathbb{E}\left[\int_{0+}^{T}e^{-rt}[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}M_t^N\right]=0.
$$

Similarly, using (1), we can show that

$$\mathbb{E}\left[\int_{0+}^{T} e^{-rt}\Big(\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\Big)\mathrm{d}M_t^Q\right]=0.$$

It follows from (22) and the fact that both V_{I} and $V_{\emptyset}$ are nondecreasing that $\phi(w,\varepsilon)\geqslant \underline{v}-w$ for any $\varepsilon\in\{\mathsf{I},\emptyset\}$. Letting $T\to\infty$ in (EC.45) and using (WU), we have $\phi(W_{0-},\emptyset)\geqslant U(\Gamma)$ with $W_{0-}=u(\Gamma,\bar{\nu}(\Gamma))$. Hence, the desired result is obtained.

A byproduct of the proof of Theorem 1 is the following result. In the remaining of this e-companion, whenever we need to prove that certain contract achieves the upper bound, we will use this result together with Lemma 1.

Proposition EC.3. *Suppose that the conditions stated in Theorem 1 hold. Furthermore, suppose that there exists a contract $\Gamma^{\diamond}\in\mathfrak{C}$ such that the corresponding agent's promised utility W_t satisfies*

$$(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t$$
$$=\ell_t-[R\nu_t-(c-b)\mathbb{1}_{\nu_t=\mu}], \quad \text{(EC.47)}$$

$$\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-})-\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})=\Delta L_t, \quad \text{(EC.48)}$$

$$\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})=\kappa(\mathcal{E}_{t-},\mathcal{E}_t), \quad \text{(EC.49)}$$

$$q_t\Big\{H_t^q\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)+\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\Big\}=0, \quad \text{(EC.50)}$$

for any $t>0$ and

$$\mathbb{E}[\phi(W_0,\mathcal{E}_{0-})]-\phi(W_{0-},\mathcal{E}_{0-})=\mathbb{E}[\Delta L_0]. \quad \text{(EC.51)}$$

Then, for any value $w\in[0,\infty)$ such that $u(\Gamma^{\diamond},\bar{\nu}(\Gamma^{\diamond}))=w$, we have

$$U(\Gamma^{\diamond})=V_{\emptyset}(w)-w.$$

Proof. Equalities (EC.47)–(EC.51) demonstrate that all the inequalities in the proof of Theorem 1, (EC.40)–(EC.44), hold with equalities under contract $\Gamma^{\diamond}$. The desired result can be shown by going through the proof of Theorem 1, with all inequalities replaced by equalities. ◻

EC.2.4. Proof of Theorem 2

In view of Theorems 3–5, it remains to show that $\bar{K}(R)$ is increasing in R and $\underline{K}(R)$ is decreasing in R on $(c/\Delta\mu,\bar{R})$.

First, under Condition 2, we have $\bar{K}=\bar{K}_2=\bar{V}-\underline{v}-\bar{w}=\dfrac{\Delta\mu\cdot R-c-\mu\beta}{r}$, which is clearly increasing in R.

Second, under Condition 1, we have $\bar{K} = \bar{K}_1$. By the definition of $\bar{K}_1$, we only need to show that m^K is decreasing in R, since m^K is decreasing in K by Lemma 4. Observe that m^K has the following characterization:

$$m^K = \max_{x>0} \left\{ \frac{\mathcal{V}_{\hat{\mathrm{w}}}(x) - \underline{v} - K}{x} \right\}.$$

Hence, it suffices to show that $\mathcal{V}_{\hat{\mathrm{w}}}(x) - \underline{v}$ is decreasing in R for any $x > 0$. Note that for any $x > 0$, we have $\mathcal{V}_{\hat{\mathrm{w}}}(x) - \underline{v} = \int_0^x \mathcal{V}'_{\hat{\mathrm{w}}}(y)\mathrm{d}y$. Hence, the desired result is obtained if we can show that $\mathcal{V}'_{\hat{\mathrm{w}}}(y)$ is decreasing in R for any $y \geqslant 0$, which is exactly Lemma EC.1(iii) below.

LEMMA EC.1. *Let R vary and other model parameters μ, $\underline{\mu}$, c, b, and K be fixed.*

(i) *Both $V'_{\tilde{w}}(w)$ and $\mathcal{V}'_{\tilde{w}}(w)$ do not depend on R for any $w \geqslant 0$.*

(ii) *$\hat{\mathrm{w}}$ is decreasing in R.*

(iii) *$\mathcal{V}'_{\hat{\mathrm{w}}}(w)$ is decreasing in R for any $w \geqslant 0$.*

Proof. (i) Note that $V'_{\tilde{w}}$ satisfies (EC.54) on $[0, \tilde{w}]$, with the boundary condition that $V'_{\tilde{w}}(w) = 0$ for all $w \geqslant \tilde{w}$. As (EC.54) does not involve parameter R, its unique solution $V'_{\tilde{w}}$ is also independent of R, which also implies the independence of $V''(\tilde{w})$ on R. Therefore, $\breve{w}(\tilde{w})$ is also independent of R, by Lemma 2(i), which in turn concludes the independence of $\mathcal{V}'_{\tilde{w}}$ on R by Lemma 2(ii).

(ii) This part can be shown using a similar line as that in the proof of Proposition 1 in Cao et al. (2022). Specifically, we define $\psi(R, \tilde{w}) := \mathcal{V}_{\tilde{w}}(0; R) - \underline{v}$, where we write $\mathcal{V}_{\tilde{w}}(\cdot; R)$ instead of $\mathcal{V}_{\tilde{w}}(\cdot)$ to highlight the dependence on R. (Note that we adopt a different notation from those in Section EC.1.4 as the parameter settings are different.) Then, we have

$$\psi(R, \tilde{w}) = \mathcal{V}_{\tilde{w}}(\tilde{w}; R) - \int_0^{\tilde{w}} \mathcal{V}'_{\tilde{w}}(y)\mathrm{d}y - \underline{v} = \frac{\Delta\mu(R - \beta) - (\rho - r)\tilde{w}}{r} - \int_0^{\tilde{w}} \mathcal{V}'_{\tilde{w}}(y)\mathrm{d}y,$$

which is linear and increasing in R for any $\tilde{w} \in [0, \bar{w}]$, where have we used part (i) in this lemma. Hence, the desired result is obtained by noting that $\psi(R, \hat{\mathrm{w}}) = 0$.

(iii) This result follows immediately from parts (i) and (ii), combining with the monotonicity of $\mathcal{V}'_{\tilde{w}}(y)$ in $\tilde{w}$ (see Lemma 2(iii)). ◻

It remains to show that $\underline{K}$ is non-increasing in R. This can be obtained by taking the first-order derivative of $\underline{K}$ with respect to R in (33), investigating its sign and noting from Assumption 1 that $R > c/\Delta\mu$.

EC.3. Proofs of the Results in Section 4

EC.3.1. Proof of Lemma 2

This result follows almost the same logic as that for the proof of Lemmas 2 and 3 in Cao et al. (2022) and uses Lemma EC.2 below. However, there are minor differences as both β and $\bar{w}$ in Cao et al. (2022) take different values from ours. Hence, to make this paper self-contained, we provide a complete proof here.

We first present Lemma EC.2 below because it will be frequently used in the subsequent analysis.

Lemma EC.2. *For any $\tilde{w} \in [0, \bar{w})$, there exists a unique function $V_{\tilde{w}}$ in $C^1([0, \tilde{w}])$ that solves the differential equation (25) on $[0, \tilde{w}]$ with boundary condition (26). We further extend the domain of $V_{\tilde{w}}$ to $\mathbb{R}_+$ by letting $V_{\tilde{w}}(w) = V_{\tilde{w}}(\tilde{w})$ for all $w > \tilde{w}$. Then, function $V_{\tilde{w}}(w)$ has the following properties.*

(i) $V_{\tilde{w}}(\cdot) \in C^1(\mathbb{R}_+) \cap C^2(\mathbb{R}_+ \backslash \{\tilde{w}\}) \cap C^3(\mathbb{R}_+ \backslash \{\tilde{w}, \tilde{w} - \beta\})$.

(ii) *For any given $w \geqslant 0$, define function $v(\tilde{w}) := V_{\tilde{w}}(w)$. We have $v(\cdot) \in C^1([0, \bar{w}))$.*

(iii) *Function $V_{\tilde{w}}(w)$ is increasing in w on $[0, \tilde{w}]$.*

(iv) *For any $\tilde{w}_1$ and $\tilde{w}_2$ such that $0 < \tilde{w}_1 < \tilde{w}_2 < \bar{w}$, we have $V_{\tilde{w}_1}(w) > V_{\tilde{w}_2}(w)$ and $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ for $w \in [0, \tilde{w}_1)$.*

(v) *If $\rho \leqslant r + \mu$, then for any $w \in [0, \bar{w})$, $V_{\tilde{w}}(w)$ approaches negative infinity as $\tilde{w}$ approaches $\bar{w}$ from below.*

(vi) *If $\rho > r + \mu$, then for any $w \in [0, \bar{w}]$, we have*

$$\lim_{\tilde{w} \uparrow \bar{w}} V_{\tilde{w}}(w) = \bar{V} - \frac{\rho - r}{\rho - r - \mu}(\bar{w} - w),$$

where $\bar{V}$ is defined in (10). Furthermore, $\bar{V} - \dfrac{\rho - r}{\rho - r - \mu}\bar{w} \geqslant \underline{v}$ is equivalent to $R \geqslant \hat{R}$.

Proof. Step 1 in the proof of Proposition 4 in Sun and Tian (2018) has already shown the existence and uniqueness of a function satisfying (25) with boundary condition (26). Here, we adopt their idea with argument slightly modified. First, we observe that (25) reduces to an ordinal differential equation (ODE) on the interval $[(\tilde{w} - \beta)^+, \tilde{w}]$, as $V_{\tilde{w}}(w + \beta) = \bar{V}(\tilde{w})$ for all $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$. Therefore, this problem can be "backwardly" treated as an initial value problem, which satisfies the conditions stated in Cauchy–Lipschitz theorem and thus admits a unique continuously differential solution on $[(\tilde{w} - \beta)^+, \tilde{w}]$. In fact, we have

$$V_{\tilde{w}}(w) = \begin{cases} \bar{V}(\tilde{w}) + \dfrac{\rho - r}{r + \mu - \rho}(\tilde{w} - w) + b_{\tilde{w}}\left((\bar{w} - w)^{\frac{r+\mu}{\rho}} - (\bar{w} - \tilde{w})^{\frac{r+\mu}{\rho}}\right), & \rho \neq r + \mu \\ \bar{V}(\tilde{w}) + \dfrac{\rho - r}{\rho}(\tilde{w} - w) - \dfrac{(\rho - r)(\bar{w} - w)}{\rho} \ln\left(\dfrac{\bar{w} - w}{\bar{w} - \tilde{w}}\right), & \rho = r + \mu \end{cases} \tag{EC.52}$$

for $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$, where

$$b_{\tilde{w}} := \frac{r - \rho}{r + \mu - \rho} \cdot \frac{\rho}{r + \mu}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}}. \tag{EC.53}$$

In general, for any $k \in \mathbb{N}$, given that the values of $V_{\tilde{w}}$ on $w \in [(\tilde{w} - k\beta)^+, \tilde{w}]$ all determined, (25) is an ODE on the interval $[(\tilde{w} - (k+1)\beta)^+, (\tilde{w} - k\beta)^+]$, whose unique solution can be shown by verifying the conditions in Cauchy–Lipschitz Theorem. By induction on k, we can extend the solution to (25) to the entire interval $[0,\widetilde{w}]$, as desired.

Next, we show that such a function $V_{\widetilde{w}}$ possesses properties (i)–(vi).

(i) It follows from (25) and the boundary condition at $\widetilde{w}$ that $V_{\widetilde{w}}(\cdot)\in C^1(\mathbb{R}_+)$. Taking derivative in (25) with respect to w and noting that $V_{\widetilde{w}}(w)=\bar{V}(\widetilde{w})$ for all $w\geqslant\widetilde{w}$, we have

$$(\mu+r)V'_{\widetilde{w}}(w)-\mu V'_{\widetilde{w}}(w+\beta)+\rho(\bar{w}-w)V''_{\widetilde{w}}(w)-\rho V'_{\widetilde{w}}(w)+\rho-r=0 \tag{EC.54}$$

for $w\in[0,\widetilde{w})$, which implies that $V_{\widetilde{w}}(\cdot)\in C^2([0,\widetilde{w}))$. Moreover, $V''_{\widetilde{w}}(\widetilde{w}-)=-(\rho-r)/\big(\rho(\bar{w}-\widetilde{w})\big)<0$. Also, by the definition of $V_{\widetilde{w}}$ on $(\widetilde{w},\infty)$, we have $V''_{\widetilde{w}}(w)=0$ for $w>\widetilde{w}$ and thus $V''_{\widetilde{w}}(\widetilde{w}+)=0$. Therefore, $V_{\widetilde{w}}(\cdot)\in C^2(\mathbb{R}_+\backslash\{\widetilde{w}\})$.

Similarly, taking derivative in (EC.54) with respect to w yields

$$\rho(\bar{w}-w)V'''_{\widetilde{w}}(w)=\mu\big(V''_{\widetilde{w}}(w+\beta)-V''_{\widetilde{w}}(w)\big)+(2\rho-r)V''_{\widetilde{w}}(w), \tag{EC.55}$$

for $w\in[0,\widetilde{w})$. Note that $V''_{\widetilde{w}}$ does not exist only at $\widetilde{w}$, which demonstrates that $V'''_{\widetilde{w}}$ does not exist at $\widetilde{w}$ and $\widetilde{w}-\beta$ (if it is nonnegative). That is, $V_{\widetilde{w}}(\cdot)\in C^3(\mathbb{R}_+\backslash\{\widetilde{w},\widetilde{w}-\beta\})$.

(ii) Fix any $w\geqslant 0$. If $\widetilde{w}\leqslant w$, then $v(\widetilde{w})=V_{\widetilde{w}}(w)=\bar{V}(\widetilde{w})$, which implies that $v(\cdot)\in C^1([0,w\wedge\bar{w}))$. Hence, the desired property is obtained if $w\geqslant\bar{w}$.

Now suppose that $w<\bar{w}$ and $\widetilde{w}\in(w,\bar{w})$. By the above discussion, we have $v(\cdot)\in C^1([0,w])$. For any $w'\in[w,\widetilde{w}]$, it follows from (25) that

$$\rho V'_{\widetilde{w}}(w')=-\frac{(\mu+r)V_{\widetilde{w}}(w')}{\bar{w}-w'}+\frac{\mu V_{\widetilde{w}}\big((w'+\beta)\wedge\widetilde{w}\big)}{\bar{w}-w'}+\frac{(\mu R-c)-(\rho-r)w'}{\bar{w}-w'}.$$

Integrating the above equation with respect to w' from w to $\widetilde{w}$ yields

$$\begin{aligned}\rho\big(V_{\widetilde{w}}(\widetilde{w})-V_{\widetilde{w}}(w)\big)=&-(\mu+r)\int_w^{\widetilde{w}}\frac{V_{\widetilde{w}}(w')}{\bar{w}-w'}\mathrm{d}w'+\mu\int_w^{\widetilde{w}}\frac{V_{\widetilde{w}}\big((w'+\beta)\wedge\widetilde{w}\big)}{\bar{w}-w'}\mathrm{d}w'\\&+\int_w^{\widetilde{w}}\frac{(\mu R-c)-(\rho-r)w'}{\bar{w}-w'}\mathrm{d}w'.\end{aligned}$$

First, using the above equality, we can obtain that $v(\widetilde{w})=V_{\widetilde{w}}(w)$ is continuous in $\widetilde{w}$ on $[w,\bar{w})$. Then, again using this equality, we conclude that $V_{\widetilde{w}}(w)$ is continuously differentiable in $\widetilde{w}$ on $[w,\bar{w})$, which, combining with $v(\cdot)\in C^1([0,w])$, yields that $v(\cdot)\in C^1\big([0,\bar{w})\big)$.

(iii) We show this result by a contradictory argument. Suppose that $w^p:=\sup\{w\in\mathbb{R}_+\mid V'_{\widetilde{w}}(w)<0\}$ exists. Recall from the proof for part (ii) of this lemma that $V''_{\widetilde{w}}(\widetilde{w}-)<0$. Hence, we have $w^p\in[0,\widetilde{w})$, $V'_{\widetilde{w}}(w^p)=0$, and $V'_{\widetilde{w}}>0$ on $(w^p,\widetilde{w})$. Evaluating (25) at w^p gives

$$\begin{aligned}rV_{\widetilde{w}}(w^p)&=\mu R-c-(\rho-r)w^p+\mu\Big[V_{\widetilde{w}}\big((w^p+\beta)\wedge\widetilde{w}\big)-V_{\widetilde{w}}(w^p)\Big]\\&>\mu R-c-(\rho-r)w^p>rV_{\widetilde{w}}(\widetilde{w}),\end{aligned}$$

where the first inequality uses $V'_{\tilde{w}} > 0$ on $(w^p, \tilde{w})$. This reaches a contradiction with $V_{\tilde{w}}(w^p) < V_{\tilde{w}}(\tilde{w})$.

(iv) We first show the second claim. Suppose it fails to hold. Since $V'_{\tilde{w}_1}(\tilde{w}_1) = 0$ and $V'_{\tilde{w}_2}(\tilde{w}_1) > 0$, the quantity $w^\dagger := \sup\{w \in [0, \tilde{w}_1) \mid V'_{\tilde{w}_1}(w) \geqslant V'_{\tilde{w}_2}(w)\}$ is well defined, and satisfies $V'_{\tilde{w}_1}(w^\dagger) = V'_{\tilde{w}_2}(w^\dagger)$ by part (ii) of this lemma. Evaluating (25) at $w^\dagger$ for both $\tilde{w}_1$ and $\tilde{w}_2$, we obtain

$$\mu(V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta)) = (r + \mu)(V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger)).$$

Hence, we have

$$\begin{aligned} V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta) &= V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger) + \int_0^\beta \big(V'_{\tilde{w}_1}(w^\dagger + y) - V'_{\tilde{w}_2}(w^\dagger + y)\big)\mathrm{d}y \\ &< V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger) = \frac{\mu}{r+\mu}\big(V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta)\big), \end{aligned}$$

which indicates that both $V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta)$ and $V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger)$ are negative. By the definition of $w^\dagger$, we have $V'_{\tilde{w}_1} < V'_{\tilde{w}_2}$ on $(w^\dagger, \tilde{w}_1]$, which implies that

$$\begin{aligned} V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger) &= V_{\tilde{w}_1}(\tilde{w}_1) - V_{\tilde{w}_2}(\tilde{w}_1) - \int_{w^\dagger}^{\tilde{w}_1} \big(V'_{\tilde{w}_1}(y) - V'_{\tilde{w}_2}(y)\big)\mathrm{d}y \\ &> V_{\tilde{w}_1}(\tilde{w}_1) - V_{\tilde{w}_2}(\tilde{w}_1) > \bar{V}(\tilde{w}_1) - \bar{V}(\tilde{w}_2) > 0. \end{aligned} \tag{EC.56}$$

This contradiction indicates the correctness of the second claim. The first claim follows by replacing $w^\dagger$ by any $w \in [0, \tilde{w}_1)$ in (EC.56).

(v) For any $w \in [(\bar{w} - \beta)^+, \bar{w})$, we have $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$ when $\tilde{w}$ is close to $\bar{w}$ from below and thus (EC.52) is valid. Letting $\tilde{w} \uparrow \bar{w}$ in (EC.52), we obtain that $\lim_{\tilde{w}\uparrow\bar{w}} V_{\tilde{w}}(w) = -\infty$.

If $w \in [0, [\bar{w} - \beta)^+, \bar{w}))$, then using the fact that $V_{\tilde{w}}$ is nondecreasing on $\mathbb{R}_+$, we obtain $V_{\tilde{w}}(w) \leqslant V_{\tilde{w}}\big((\bar{w} - \beta)^+\big)$, which yields that $\limsup_{\tilde{w}\uparrow\bar{w}} V_{\tilde{w}}(w) \leqslant \lim_{\tilde{w}\uparrow\bar{w}} V_{\tilde{w}}\big((\bar{w} - \beta)^+\big) = -\infty$, concluding the desired result.

(vi) Note that $\rho > r + \mu$ implies $\bar{w} < \beta$. Hence, (EC.52) is valid for all $w \in [0, \tilde{w}]$. Therefore, the first claim follows by letting $\tilde{w} \uparrow \bar{w}$ in (EC.52) and using $\lim_{\tilde{w}\uparrow\bar{w}} b_{\tilde{w}} = 0$. The second claim is trivial by the definition of $\hat{R}$. □

We proceed to prove Lemma 2 as follows.

Proof of Lemma 2. (i) From (EC.52), we have

$$V''_{\tilde{w}}(w) = -\frac{\rho - r}{\rho}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}}(\bar{w} - w)^{\frac{-2\rho + r + \mu}{\rho}} < 0$$

for any $w \in [(\tilde{w} - \beta)^+, \tilde{w})$. (The above expression also holds if $\mu + r = \rho$.) Hence, the desired result holds with $\breve{w}(\tilde{w}) = 0$ if $\tilde{w} \leqslant \beta$.

Now consider the case that $\tilde{w} > \beta$. In this case, we have $\bar{w} > \beta$, which gives $\rho < \mu$. Define $w^c := \inf\{w \in [0, \tilde{w}) \mid V''_{\tilde{w}}(w) \geqslant 0\}$. If the set is empty, we set $w^c = 0$. By Lemma EC.2(i), we have $V''_{\tilde{w}} < 0$ on $(w^c, \tilde{w})$. Hence, the desired result holds with $\breve{w}(\tilde{w}) = 0$ if $w^c = 0$.

Next, we suppose that $w^c > 0$. Since $V_{\tilde{w}}$ is strictly concave on $[\tilde{w}-\beta, \tilde{w})$, we have $w^c < \tilde{w}-\beta$. According to Lemma EC.2(i), we have $V''_{\tilde{w}}(w^c) = 0$ and $V''_{\tilde{w}} < 0$ on $(w^c, \tilde{w})$.

It follows from (EC.54) 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

$$V'_{\tilde{w}}(w^c+\beta) = \frac{(\mu-\rho+r)V'_{\tilde{w}}(w^c) + (\rho - r)}{\mu}. \tag{EC.57}$$

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

$$V'_{\tilde{w}}(w^c) > 1, \tag{EC.58}$$

in view of (EC.57) and $\rho < \mu$. Evaluating (25) at w^c gives

$$\begin{aligned} rV_{\tilde{w}}(w^c) &= \mu R - c - (\rho - r)w^c - \rho(\bar{w} - w^c)V'_{\tilde{w}}(w^c) + \mu\big(V_{\tilde{w}}(w^c+\beta) - V_{\tilde{w}}(w^c)\big) \\ &> \mu R - c - (\rho - r)w^c - \rho(\bar{w} - w^c)V'_{\tilde{w}}(w^c) + \mu\beta V'_{\tilde{w}}(w^c+\beta) \\ &= \mu R - c - (\rho - r)(w^c - \beta) + \big[\rho(w^c-\beta) + r\beta\big]V'_{\tilde{w}}(w^c), \end{aligned} \tag{EC.59}$$

where the inequality follows from the strict concavity of $V_{\tilde{w}}$ on $(w^c, w^c+\beta)$, and the last equality uses (EC.57) and $\rho\bar{w} = \mu\beta$.

Below we distinguish two cases.

Case 1: $\rho(w^c-\beta) + r\beta \geqslant 0$. It follows from (EC.58) and (EC.59) that

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

which contradicts Lemma EC.2(iii).

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

$$0 < w^c < \frac{(\rho - r)\beta}{\rho}. \tag{EC.60}$$

Below, we will show that

$$V''_{\tilde{w}} > 0 \text{ on } [0, w^c). \tag{EC.61}$$

Evaluating (EC.55) at w^c and using $V''_{\tilde{w}}(w^c) = 0$, we obtain

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

which implies that $V''_{\tilde{w}} > 0$ on $(w^c - \epsilon, w^c)$ for some $\epsilon > 0$. If (EC.61) fails to hold, then $w^d := \sup\{w \in [0, w^c) \mid V''_{\tilde{w}}(w) \leqslant 0\}$ is well defined, satisfying $w^d \in [0, w^c)$. Moreover, we have $V''_{\tilde{w}}(w^d) = 0$ and $V'''_{\tilde{w}}(w^d) \geqslant 0$. (Note that $w^c < \tilde{w} - \beta$, indicating that $V'''_{\tilde{w}}(w^d)$ exists by Lemma EC.2(i).) Hence, evaluating (EC.55) at w^d gives $\rho(\bar{w} - w^d)V'''_{\tilde{w}}(w^d) = \mu V''_{\tilde{w}}(w^d + \beta) \geqslant 0$. By the definition of w^c, we have $w^d + \beta \leqslant w^c$. Consequently, it follows from (EC.60) that $w^d \leqslant w^c - \beta < 0$, which is impossible. Therefore, (EC.61) holds. Letting $\breve{w}(\tilde{w}) = w^c$, we obtain the proof of the first claim in part (i). The second claim in part (ii) follows immediately by (EC.60).

(ii) This claim holds trivially by the first claim in part (i) and Lemma EC.2(iii).

(iii) To ease notation, we write $\breve{w}(\tilde{w}_1)$ and $\breve{w}(\tilde{w}_2)$ as $\breve{w}_1$ and $\breve{w}_2$ respectively. First, we show the second claim, that is, $\mathcal{V}'_{\tilde{w}_1}(w) < \mathcal{V}'_{\tilde{w}_2}(w)$ for any $w \in [0, \tilde{w}_1]$, by considering the following two cases.

Case 1: $\breve{w}_1 \leqslant \breve{w}_2$. In this case, we have

$$\mathcal{V}'_{\tilde{w}_1}(w) = V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w) = \mathcal{V}'_{\tilde{w}_2}(w) \tag{EC.62}$$

for any $w \in [\breve{w}_2, \tilde{w}_1]$, where the two equalities follows from the definition of function $\mathcal{V}_{\tilde{w}}$ and $\breve{w}_1 \leqslant \breve{w}_2$, and the inequality follows from Lemma EC.2(iv). If $w \in [0, \breve{w}_2]$, then we can derive the desired inequality as follows:

$$\mathcal{V}'_{\tilde{w}_1}(w) \leqslant \mathcal{V}'_{\tilde{w}_1}(\breve{w}_1) = V'_{\tilde{w}_1}(\breve{w}_1) < V'_{\tilde{w}_2}(\breve{w}_1) < V'_{\tilde{w}_2}(\breve{w}_2) = \mathcal{V}'_{\tilde{w}_2}(\breve{w}_1) = \mathcal{V}'_{\tilde{w}_2}(w).$$

Here, the first inequality uses the definition and the concavity of $\mathcal{V}_{\tilde{w}_1}$, the second inequality uses Lemma EC.2(iv), and the last inequality uses Lemma 2(i).

Case 2: $\breve{w}_1 > \breve{w}_2$. In this case, we have

$$\mathcal{V}'_{\tilde{w}_1}(w) = V'_{\tilde{w}_1}(w \vee \breve{w}_1) < V'_{\tilde{w}_2}(w \vee \breve{w}_1) \leqslant V'_{\tilde{w}_2}(w \vee \breve{w}_2) = \mathcal{V}'_{\tilde{w}_2}(w) \tag{EC.63}$$

for all $w \in [0, \tilde{w}_1]$, where the first inequality uses Lemma EC.2(iv) and the second inequality uses the concavity of $V_{\tilde{w}_2}$ and the fact that $w \vee \breve{w}_1 \geqslant w \vee \breve{w}_2$.

The first claim can be readily obtained using the second claim. In fact, for any $w \in [0, \tilde{w}_1]$, we have

$$\begin{aligned}\mathcal{V}_{\tilde{w}_1}(w) - \mathcal{V}_{\tilde{w}_2}(w) &= \mathcal{V}_{\tilde{w}_1}(\tilde{w}_1) - \mathcal{V}_{\tilde{w}_2}(\tilde{w}_1) - \int_w^{\tilde{w}_1} \big(\mathcal{V}'_{\tilde{w}_1}(y) - \mathcal{V}'_{\tilde{w}_2}(y)\big)\mathrm{d}y \\ &> \mathcal{V}_{\tilde{w}_1}(\tilde{w}_1) - \mathcal{V}_{\tilde{w}_2}(\tilde{w}_1) > \bar{V}(\tilde{w}_1) - \bar{V}(\tilde{w}_2) > 0,\end{aligned}$$

where the second inequality uses part (ii) of this lemma. ▫

EC.3.2. Proof of Lemma 3

It follows from Lemma 2(iii) that $\mathcal{V}_{\tilde{w}}(0)$ is decreasing in $\tilde{w}$ on $(0,\bar{w})$. By Lemma EC.2(v) and (vi), we have that $\lim_{\tilde{w}\uparrow\bar{w}}\mathcal{V}_{\tilde{w}}(0)$ is either $-\infty$ or $\bar{V}-\frac{\rho-r}{\rho-r-\mu}\bar{w}$, which is less than $\underline{v}$ according to Condition 1. Moreover, $\lim_{\tilde{w}\downarrow 0}\mathcal{V}_{\tilde{w}}(0)=\bar{V}(0)>\underline{v}$ by Assumption 1. Therefore, the first claim is obtained by Lemma EC.2(ii).

For the second claim, we observe that by the proof of Lemma 2(i), if $\breve{w}(\tilde{w})>0$, it must be equal to w^c, in which case (EC.58) holds. This immediately concludes the result by the definition of $\mathcal{V}_{\tilde{w}}$.

EC.3.3. Proof of Proposition 2

Clearly, following the definition of $\underline{\Gamma}$ as in (16), $U(\underline{\Gamma})=\underline{v}$ trivially holds. Hence, it remains to show that functions $V_{\mathsf{I}}(w)=\mathcal{V}_{\hat{\mathrm{w}}}(w)$ and $V_{\emptyset}(w)=\underline{v}$ satisfy the optimality condition (20)–(22).

First, we show that $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)\geqslant 0$ for any $w\in\mathbb{R}_+$. If $w\in[\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}})$, by the definition of $\mathcal{V}_{\hat{\mathrm{w}}}$, we have $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=0$. If $w\in[\hat{\mathrm{w}},\infty)$, then we have

$$
\begin{aligned}
(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)&=(\mu+r)\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}})-\mu\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}})-(\mu R-c)+(\rho-r)w\\
&=(\rho-r)(w-\hat{\mathrm{w}})\geqslant 0.
\end{aligned}
$$

If $w\in[0,\breve{w}(\hat{\mathrm{w}}))$ (if we discuss this case, it is implicitly assumed that $\breve{w}(\hat{\mathrm{w}})>0$), then we have $\mathcal{V}_{\hat{\mathrm{w}}}(w)=\underline{v}+\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))w$. Consequently,

$$
(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=(\mu+r)\big(\underline{v}+\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))w\big)-\mu\mathcal{V}_{\hat{\mathrm{w}}}(w+\beta)+\rho(\bar{w}-w)\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-(\mu R-c)+(\rho-r)w.
$$

Let the last expression be $g_{\mathsf{I}}(w)$. Obviously, $g_{\mathsf{I}}(\breve{w}(\hat{\mathrm{w}}))=0$. Moreover, for $w\in[0,\breve{w}(\hat{\mathrm{w}}))$, we have

$$
\begin{aligned}
g'_{\mathsf{I}}(w)&=(\mu+r)\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-\mu\mathcal{V}'_{\hat{\mathrm{w}}}(w+\beta)-\rho\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))+\rho-r\\
&=(\rho-r)\big(1-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))\big)+\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-\mathcal{V}'_{\hat{\mathrm{w}}}(w+\beta)\big)\\
&\leqslant(\rho-r)\big(1-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))\big)+\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}})+\beta)\big)=0,
\end{aligned}
$$

where the inequality follows from the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$, and the last equality uses the facts that $\rho(\bar{w}-\breve{w}(\hat{\mathrm{w}}))\mathcal{V}''_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))=(\rho-r)\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-1\big)+\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}})+\beta)-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))\big)$ by (EC.54) and that $\mathcal{V}''_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))=0$. Consequently, $g_{\mathsf{I}}(w)\geqslant 0$ for all $w\in[0,\breve{w}(\hat{\mathrm{w}}))$.

Therefore, we have $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)\geqslant 0$ for any $w\in\mathbb{R}_+$. Obviously, we have $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)=(\rho-r)w\geqslant 0$. Hence, (20) holds.

It follows from the facts that $V_{\mathsf{I}}(w)\geqslant V_{\mathsf{I}}(0)=\underline{v}=V_{\emptyset}(w)$ and that $V_{\emptyset}(w)=\underline{v}\geqslant\bar{V}(\hat{\mathrm{w}})-K=\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}})-K\geqslant V_{\mathsf{I}}(w)-K$ (due to $K\geqslant\bar{V}(\hat{\mathrm{w}})-\underline{v}$) that both (21) and (22) hold.

EC.3.4. Proof of Lemma 4

Define

$$g(w,K) := \mathcal{V}_{\hat{\mathrm{w}}}(w) - \mathcal{V}'_{\hat{\mathrm{w}}}(w)w - \underline{v} - K. \tag{EC.64}$$

Then, we have

$$g(\breve{w}(\hat{\mathrm{w}}),K) = -K < 0, \tag{EC.65}$$

where the equality uses the linearity of $\mathcal{V}_{\hat{\mathrm{w}}}$ on $[0,\breve{w}(\hat{\mathrm{w}})]$. In addition,

$$g(\hat{\mathrm{w}},K) = \mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) - \underline{v} - K > 0,$$

where the equality follows from $\mathcal{V}'_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) = 0$ and the inequality follows from the condition that $K < \bar{V}(\hat{\mathrm{w}}) - \underline{v}$. Furthermore, we have

$$\frac{\partial g(w,K)}{\partial w} = -\mathcal{V}''_{\hat{\mathrm{w}}}(w)w > 0 \ \text{ for } \ w \in (\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}}),$$

where the inequality follows from the fact that $\mathcal{V}_{\hat{\mathrm{w}}}$ is strictly concave on $(\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}})$. Since $g(w,K)$ is continuous in w (recalling that $\mathcal{V}_{\hat{\mathrm{w}}}$ is continuously differentiable), for any $K > 0$, there exists a unique $\bar{\theta}^K \in (\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}})$ such that $g(\bar{\theta}^K,K) = 0$. Hence, (29) holds if we define $m^K := \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K)$. Furthermore, by the implicit function theorem, we have

$$\frac{\mathrm{d}\bar{\theta}^K}{\mathrm{d}K} = -\frac{\frac{\partial g(w,K)}{\partial K}}{\frac{\partial g(w,K)}{\partial w}} = \frac{1}{\frac{\partial g(w,K)}{\partial w}} > 0,$$

which implies that $\bar{\theta}^K$ is increasing in K. Since $\mathcal{V}'_{\hat{\mathrm{w}}}(w)$ is decreasing in w, we have $m^K = \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K)$ is decreasing in K. Finally, the limiting result $\lim_{K\downarrow 0}\bar{\theta}^K = \breve{w}(\hat{\mathrm{w}})$ is implied by (EC.65).

EC.3.5. Proof of Proposition 3

Obviously, (22) holds since $V_{\mathsf{l}}(0) = V_{\emptyset}(0) = \underline{v}$. Note that it has been shown in the proof of Proposition 2 that $(\mathcal{A}_{\mathsf{l}}V_{\mathsf{l}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$. Hence, it remains to establish the second part of (20), as well as (21), by considering the following three cases.

Case 1: $w \in [0,\bar{\theta}^K)$. In this case, we have $(\mathcal{A}_{\emptyset}V_{\emptyset})(w) = (\rho - r)(1-m^K)w \geqslant 0$, where we have used the fact that $m^K \leqslant 1$, which follows from Lemma 4, the definition of $\bar{K}_1$ as in (K1), and the condition that $K \geqslant \bar{K}_1$. Besides, we have

$$\begin{aligned} V_{\emptyset}(w) &= \left(1 - \frac{w}{\bar{\theta}^K}\right)\underline{v} + \frac{w}{\bar{\theta}^K}\cdot(\mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) - K) \\ &\leqslant \left(1 - \frac{w}{\bar{\theta}^K}\right)\mathcal{V}_{\hat{\mathrm{w}}}(0) + \frac{w}{\bar{\theta}^K}\cdot\mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) \leqslant \mathcal{V}_{\hat{\mathrm{w}}}(w) = V_{\mathsf{l}}(w), \end{aligned}$$

where the second inequality follows from the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$. Finally, we have

$$\begin{aligned} V_\emptyset(w) - V_\mathsf{l}(w) + K &= m^K w + \underline{v} + K - \mathcal{V}_{\hat{\mathrm{w}}}(w) \\ &= \mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) - m^K \cdot (\bar{\theta}^K - w) - \mathcal{V}_{\hat{\mathrm{w}}}(w) = \int_w^{\bar{\theta}^K} \big(\mathcal{V}'_{\hat{\mathrm{w}}}(y) - m^K\big) \mathrm{d}y > 0, \end{aligned}$$

where the second equality follows from the first equality in (29), and the inequality follows from the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$ and $m^K = \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K)$.

Case 2: $w \in [\bar{\theta}^K, \hat{\mathrm{w}})$. First, we show that $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ in this case. Note that $V_\emptyset(w) = \mathcal{V}_{\hat{\mathrm{w}}}(w) - K$ for $w \in [\bar{\theta}^K, \hat{\mathrm{w}})$. Hence, $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ is equivalent to

$$f_1(w) := r(\mathcal{V}_{\hat{\mathrm{w}}}(w) - K) - \rho w \mathcal{V}'_{\hat{\mathrm{w}}}(w) + (\rho - r)w - R\underline{\mu} \geqslant 0.$$

For $w \in [\breve{w}(\hat{\mathrm{w}}), \hat{\mathrm{w}}]$, it holds that $(\mathcal{A}_\mathsf{l} \mathcal{V}_{\hat{\mathrm{w}}})(w) = 0$. That is,

$$f_2(w) := (\mu + r)\mathcal{V}_{\hat{\mathrm{w}}}(w) - \mu \mathcal{V}_{\hat{\mathrm{w}}}(w + \beta) + \rho(\bar{w} - w)\mathcal{V}'_{\hat{\mathrm{w}}}(w) - (\mu R - c) + (\rho - r)w = 0. \tag{EC.66}$$

Recall that $\bar{\theta}^K \in (\breve{w}(\hat{\mathrm{w}}), \hat{\mathrm{w}})$. Hence, it suffices to show that

$$f_3(w) := f_2(w) - f_1(w) = \mu\big(\mathcal{V}_{\hat{\mathrm{w}}}(w) - \mathcal{V}_{\hat{\mathrm{w}}}(w + \beta)\big) + \rho \bar{w} \mathcal{V}'_{\hat{\mathrm{w}}}(w) + rK - (R\Delta\mu - c) < 0$$

for $w \in [\bar{\theta}^K, \hat{\mathrm{w}})$.

It follows from (29) that $f_1(\bar{\theta}^K) = (\rho - r)\bar{\theta}^K(1 - m^K) > 0$. Hence, $f_3(\bar{\theta}^K) < 0$. Hence, it is enough to show that $f_3'(w) \leqslant 0$, or equivalently, $\mu(\mathcal{V}'_{\hat{\mathrm{w}}}(w) - \mathcal{V}'_{\hat{\mathrm{w}}}(w + \beta)) + \rho \bar{w} \mathcal{V}''_{\hat{\mathrm{w}}}(w) \leqslant 0$, for $w \in [\bar{\theta}^K, \hat{\mathrm{w}})$.

Taking derivative with respect to w in (EC.66) yields

$$(\mu + r)\mathcal{V}'_{\hat{\mathrm{w}}}(w) - \mu \mathcal{V}'_{\hat{\mathrm{w}}}(w + \beta) + \rho(\bar{w} - w)\mathcal{V}''_{\hat{\mathrm{w}}}(w) - \rho \mathcal{V}'_{\hat{\mathrm{w}}}(w) + \rho - r = 0$$

for $w \in [\breve{w}(\hat{\mathrm{w}}), \hat{\mathrm{w}}]$. Hence, for $w \in [\bar{\theta}^K, \hat{\mathrm{w}})$, we have

$$\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(w) - \mathcal{V}'_{\hat{\mathrm{w}}}(w + \beta)\big) + \rho \bar{w} \mathcal{V}''_{\hat{\mathrm{w}}}(w) = (\rho - r)(\mathcal{V}'_{\hat{\mathrm{w}}}(w) - 1) + \rho w \mathcal{V}''_{\hat{\mathrm{w}}}(w) \leqslant 0,$$

where the inequality follows from the fact that $\mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K) = m^K \leqslant 1$ and the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$. Hence, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for $w \in [\bar{\theta}^K, \hat{\mathrm{w}})$. Note that $V_\emptyset(w) - V_\mathsf{l}(w) + K = 0$. Therefore, (21) trivially holds.

Case 3: $w \in [\hat{\mathrm{w}}, \infty)$. It is straightforward to see that $(\mathcal{A}_\emptyset V_\emptyset)(w) = r(\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) - K) + (\rho - r)w - R\underline{\mu} > r\underline{v} - R\underline{\mu} = 0$ and $V_\emptyset(w) - V_\mathsf{l}(w) + K = 0$.

EC.3.6. Proof of Proposition 4

First, we show that under Condition 2 and $K > \bar{V} - \underline{v}$, functions $V_{\mathsf{I}}(w)$ as defined in (31) and $V_{\emptyset}(w) = \underline{v}$ satisfy the optimality condition (20)–(22). Note that the first inequality in Condition 2 implies $\bar{w} < \beta$. If $w \in [0, \bar{w}]$, then we have

$$\begin{aligned}(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w) &= (\mu + r)V_{\mathsf{I}}(w) - \mu V_{\mathsf{I}}(w+\beta) + \rho(\bar{w} - w)V_{\mathsf{I}}'(w) - (\mu R - c) + (\rho - r)w \\ &= (\mu + r)\Big(\underline{v} + \frac{\bar{V} - \underline{v}}{\bar{w}}w\Big) - \mu\bar{V} + \rho(\bar{w} - w)\frac{\bar{V} - \underline{v}}{\bar{w}} - (\mu R - c) + (\rho - r)w \\ &= (\bar{w} - w)\left[\frac{\bar{V} - \underline{v}}{\bar{w}}(\rho - r - \mu) - (\rho - r)\right] \geqslant 0.\end{aligned}$$

Here, it is worth pointing out that although at $\bar{w}$, V_{I} is not differentiable, its left derivative exists and is $(\bar{V} - \underline{v})/\bar{w}$.

If $w \in (\bar{w}, \infty)$, then

$$(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w) = (\mu + r)\bar{V} - \mu\bar{V} - (\mu R - c) + (\rho - r)w = (\rho - r)(w - \bar{w}) > 0.$$

Combining the above two cases yields $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$. Besides, for any $w \in \mathbb{R}_+$, $(\mathcal{A}_{\emptyset}V_{\emptyset})(w) = (\rho - r)w \geqslant 0$. Hence, (20) holds.

It is straightforward to see that $V_{\mathsf{I}}(w) - V_{\emptyset}(w) \geqslant 0$, and $V_{\emptyset}(w) = \underline{v} \geqslant \bar{V} - K \geqslant V_{\mathsf{I}}(w) - K$. Hence, (21) holds. Obviously, $V_{\mathsf{I}}(0) = V_{\emptyset}(0) = \underline{v}$, implying (22).

Second, we show that under Condition 2 and $\bar{K}_2 < K \leqslant \bar{V} - \underline{v}$, functions $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ as defined in (31) and (32), respectively, satisfy (20)–(22). According to the proof for the case under Condition 2 and $K > \bar{V} - \underline{v}$, we have $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$. Below, we show that $(\mathcal{A}_{\emptyset}V_{\emptyset})(w) \geqslant 0$ for all $w \in \mathbb{R}_+$.

If $w \in [0, \bar{w}]$, then

$$(\mathcal{A}_{\emptyset}V_{\emptyset})(w) = rV_{\emptyset}(w) - \rho w V_{\emptyset}'(w) + (\rho - r)w - R\underline{\mu} = (\rho - r)\Big(1 - \frac{\bar{V} - \underline{v} - K}{\bar{w}}\Big)w \geqslant 0.$$

Here, we mention that although at $\bar{w}$, $V_{\emptyset}$ is not differentiable, its left derivative exists and is $(\bar{V} - \underline{v} - K)/\bar{w}$.

If $w \in (\bar{w}, \infty)$, then we have

$$\begin{aligned}(\mathcal{A}_{\emptyset}V_{\emptyset})(w) &= rV_{\emptyset}(w) - \rho w V_{\emptyset}'(w) + (\rho - r)w - R\underline{\mu} \\ &= r(\bar{V} - K) + (\rho - r)w - R\underline{\mu} > r(\bar{V} - K - \underline{v}) > 0.\end{aligned}$$

Therefore, (20) holds. Note that $V_{\mathsf{I}}(w) - V_{\emptyset}(w) = K$ for $w \in [\bar{w}, \infty)$ and $V_{\mathsf{I}}(w) - V_{\emptyset}(w) = K/\bar{w} \cdot w$ for $w \in [0, \bar{w})$. Hence, (21) holds. Besides, $V_{\mathsf{I}}(0) = V_{\emptyset}(0) = \underline{v}$ and thus (22) holds.

EC.3.7. Proof of Lemma 5

The results stated in Lemma 5 hold in fact for any $K \in (0, \bar{K}_2)$. Below, we will show this slightly generalized version.

For any $\underline{\theta} \in [0, \bar{w}]$, it is straightforward to verify that functions

$$C^1(\underline{\theta}) := \frac{\bar{V} - \underline{v} - \bar{w}}{r/\rho \cdot \underline{\theta}^{r/\rho - 1}[(\rho/r - 1)\underline{\theta} + \bar{w}]} \text{ and } m(\underline{\theta}) := \frac{(\rho/r - 1)\underline{\theta} + \bar{V} - \underline{v}}{(\rho/r - 1)\underline{\theta} + \bar{w}} \tag{EC.67}$$

satisfy (36) and (37), with $\underline{\theta}$ replacing $\underline{\theta}_K$, $C^1(\underline{\theta})$ replacing c_K, and $m(\underline{\theta})$ replacing m_K. Moreover, it follows from Condition 2 and $K \in (0, \bar{K}_2)$ that $C^1(\underline{\theta}) > 0$. Note that the denominator of $C^1(\underline{\theta})$ is decreasing in $\underline{\theta}$, as its derivative with respect to $\underline{\theta}$ is always negative when $\underline{\theta} \in (0, \bar{w})$. Therefore, $C^1(\underline{\theta})$ is increasing in $\underline{\theta}$ on $[0, \bar{w}]$. That $m(\underline{\theta})$ is strictly decreasing in $\underline{\theta}$ on $[0, \bar{w}]$ is straightforward.

We have the following result, which is stated as a lemma for the ease of reference. Its proof is elementary and thus omitted.

LEMMA EC.3. *Under Condition 2 and $K \in (0, \bar{K}_2)$, function $\psi_1(\underline{\theta})$, defined by*

$$\psi_1(\underline{\theta}) = \bar{V} - \underline{v} - \bar{w} - C^1(\underline{\theta}) \cdot (\bar{w})^{r/\rho},$$

is continuous and decreasing in $\underline{\theta}$ on $[0, \bar{w}]$. Moreover, $\psi_1(\bar{w}) = 0$ and $\psi_1(0) = \bar{V} - \underline{v} - \bar{w} > K$. Consequently, there exists a unique number $\underline{\theta}_K \in (0, \bar{w})$ such that $\psi_1(\underline{\theta}_K) = K$. Furthermore, $\underline{\theta}_K$ is decreasing in K with $\lim_{K \downarrow 0} \underline{\theta}_K = \bar{w}$.

Lemma EC.3 immediately implies that the triple $(\underline{\theta}_K, c_K, m_K)$ with $c_K = C^1(\underline{\theta}_K)$, $m_K = m(\underline{\theta}_K)$ satisfies (36)–(38), which also states the monotonicity of $\underline{\theta}_K$ in K. The monotonicity of c_K and m_K in K follows from that of $C^1(\underline{\theta})$ and $m(\underline{\theta})$ in $\underline{\theta}$.

Finally, we show that under Condition 2, $K \geqslant \underline{K}$ if and only if (35) holds. First, according to the monotonicity of m_K in K, (35) is equivalent to

$$K \geqslant \breve{K}_2, \text{ in which } \breve{K}_2 := \inf\left\{K \in (0, \underline{K}_2] \;\middle|\; m_K \geqslant \frac{\rho - r}{\rho - r - \mu}\right\}. \tag{EC.68}$$

Next, it follows from (EC.67) and $\lim_{K \downarrow 0} \underline{\theta}_K = \bar{w}$ that

$$\lim_{K \downarrow 0} c_K = \frac{\bar{V} - \underline{v} - \bar{w}}{\bar{w}^{-r/\rho}} \text{ and } \lim_{K \downarrow 0} m_K = 1 + \frac{r(\bar{V} - \underline{v} - \bar{w})}{\rho \bar{w}} = \frac{R\Delta\mu - c}{\mu\beta}.$$

It is straightforward to verify that $\lim_{K \downarrow 0} m_K \geqslant (\rho - r)/(\rho - r - \mu)$ if and only if $R \geqslant \bar{R}$, where $\bar{R}$ is defined in (34). Hence, by the definition of $\breve{K}_2$ and the monotonicity of m_K in K, it is clear that $\breve{K}_2 = 0$ if and only if $R \geqslant \bar{R}$.

If $R < \bar{R}$, we have $m_{\breve{K}_2} = (\rho - r)/(\rho - r - \mu)$. Evaluating (EC.67) at $\underline{\theta} = \underline{\theta}_{\breve{K}_2}$ gives

$$\underline{\theta}_{\breve{K}_2} = \frac{\bar{V} - \underline{v} - (\rho - r)/(\rho - r - \mu)\bar{w}}{\mu/(\rho - r - \mu) \cdot (\rho/r - 1)} \text{ and } c_{\breve{K}_2} = \frac{\mu}{\rho - r - \mu} \frac{\rho}{r} \underline{\theta}_{\breve{K}_2}{}^{1 - r/\rho}.$$

Substituting these values into (38) with $K=\breve{K}_2$, we obtain the following closed-form expression of $\breve{K}_2$:

$$\breve{K}_2=\bar{V}-\underline{v}-\bar{w}-\frac{\mu}{\rho-r-\mu}\frac{\rho}{r}\left[\frac{\bar{V}-\underline{v}-(\rho-r)/(\rho-r-\mu)\cdot\bar{w}}{\mu/(\rho-r-\mu)\cdot(\rho/r-1)}\right]^{1-r/\rho}\bar{w}^{r/\rho}. \tag{EC.69}$$

The proof is complete by verifying that $\underline{K}=\breve{K}_2$.

EC.3.8. Proof of Proposition 5

By the definition of contract $\bar{\Gamma}$, it is clear that $U(\bar{\Gamma})=\bar{V}-\bar{w}-K=V_{\emptyset}(\bar{w})-\bar{w}$. Hence, it remains to show that under Condition 2 and $K\in[\underline{K},\bar{K}_2)$, functions $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ as defined in (40) and (39), respectively, satisfy the optimality condition (20)–(22).

Obviously, (22) holds as $V_{\mathsf{I}}(0)=V_{\emptyset}(0)=\underline{v}$. We proceed to verify that $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ satisfy (20) and (21).

We show that $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)\geqslant 0$ for all $w\in\mathbb{R}_+$ by considering the following cases.

Case 1: $w\in[\underline{\theta}_K,\bar{w}]$. We have

$$\begin{aligned}(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)&=(\mu+r)(\bar{V}+m_K\cdot(w-\bar{w}))-\mu\bar{V}+\rho(\bar{w}-w)\cdot m_K-(\mu R-c)+(\rho-r)w\\&=(\bar{w}-w)[m_K\cdot(\rho-r-\mu)-(\rho-r)]\geqslant 0.\end{aligned}$$

Here, we mention that although at $\bar{w}$, V_{I} is not differentiable, its left derivative exists and is m_K.

Case 2: $w\in[0,\underline{\theta}_K)$. In this case, we have

$$\begin{aligned}(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)&=(\mu+r)V_{\mathsf{I}}(w)-\mu V_{\mathsf{I}}(w+\beta)+\rho(\bar{w}-w)V_{\mathsf{I}}'(w)-(\mu R-c)+(\rho-r)w\\&=\mu V_{\mathsf{I}}(w)-\mu V_{\mathsf{I}}(w+\beta)+\rho\bar{w}V_{\mathsf{I}}'(w)-\Delta\mu\cdot R+c\\&=\mu\big(\underline{v}+w+c_Kw^{r/\rho}\big)-\mu\bar{V}+\mu\beta\big(1+c_Kw^{r/\rho-1}r/\rho\big)-\Delta\mu\cdot R+c=:g_{\mathsf{I}}(w),\end{aligned}$$

where the second equality follows from $\mathcal{A}_{\emptyset}V_{\mathsf{I}}=0$ on $[0,\underline{\theta}_K)$, and the third equality follows from $\beta>\mu\beta/\rho=\bar{w}$ due to Condition 2.

Since V_{I} is continuously differentiable on $[0,\bar{w})$, $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)$ is also continuous in w on $[0,\bar{w})$, which implies that $g_{\mathsf{I}}(\underline{\theta}_K)\geqslant 0$. Hence, it suffices to show that $g_{\mathsf{I}}(w)$ is decreasing in w. Using $g_{\mathsf{I}}'(w)=\mu+\mu r/\rho\cdot c_Kw^{r/\rho-2}\big(w+(r-\rho)/\rho\cdot\beta\big)$, we have

$$\begin{aligned}g_{\mathsf{I}}'(\underline{\theta}_K)&=\mu+\mu r/\rho\cdot c_K\cdot(\underline{\theta}_K)^{r/\rho-2}\big(\underline{\theta}_K+(r-\rho)/\rho\cdot\beta\big)\\&=\mu+\mu\cdot\frac{m_K-1}{\underline{\theta}_K}\cdot\big(\underline{\theta}_K+(r-\rho)/\rho\cdot\beta\big)\\&<\mu+\mu\left(\frac{\rho-r}{\rho-r-\mu}-1\right)\left(1+\frac{(r-\rho)\cdot\beta}{\rho\underline{\theta}_K}\right)\\&<\mu+\mu\frac{\mu}{\rho-r-\mu}\left(1+\frac{(r-\rho)\cdot\beta}{\rho\bar{w}}\right)=0,\end{aligned}$$

where the second equality follows from (37), the first inequality follows from (35) and the fact that $\underline{\theta}_K + (r-\rho)/\rho\cdot\beta < \bar{w} + (r-\rho)/\rho\cdot\beta = (\mu+r-\rho)/\rho\cdot\beta < 0$, and the last equality follows from $\bar{w}=\mu\beta/\rho$. Besides, we have

$$\begin{aligned} g_{\mathsf{I}}''(w) &= r/\rho(r/\rho-1)\cdot c_K w^{r/\rho-3}[\mu w + \rho\bar{w}(r/\rho-2)] \\ &\geqslant r/\rho(r/\rho-1)\cdot c_K w^{r/\rho-3}[\mu\bar{w} + \rho\bar{w}(r/\rho-2)] \\ &= r/\rho(r/\rho-1)\cdot c_K w^{r/\rho-3}(\mu+r-2\rho)\bar{w} > 0, \end{aligned}$$

where the last inequality follows from $\rho > r+\mu$. Therefore, $g_{\mathsf{I}}'(w) < 0$ for $w\in[0,\underline{\theta}_K]$.

Case 3: $w\in(\bar{w},\infty)$. We have

$$(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w) = (\mu+r)\bar{V} - \mu\bar{V} - (\mu R - c) + (\rho-r)w = (\rho-r)(w-\bar{w}) > 0.$$

Combining the above three cases yields $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)\geqslant 0$ for any $w\in\mathbb{R}_+$.

Next, we establish $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)\geqslant 0$ for all $w\in\mathbb{R}_+$. Obviously, we have $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)=0$ for $w\in[0,\bar{w}]$. (Again, although $V_{\emptyset}$ is not differentiable at $\bar{w}$, its left derivative exists.) If $w\in(\bar{w},\infty)$, then

$$(\mathcal{A}_{\emptyset}V_{\emptyset})(w) = r(\bar{V}-K) + (\rho-r)w - R\underline{\mu} > r(\bar{V}-K) - R\underline{\mu} = r(\bar{V}-K-\underline{v}) > 0,$$

proving (20).

Below we establish (21). If $w\in[0,\underline{\theta}_K]$, we have $V_{\mathsf{I}}(w)-V_{\emptyset}(w)=0$, and if $w\in[\bar{w},\infty)$, we have $V_{\mathsf{I}}(w)-V_{\emptyset}(w)=K$. If $w\in(\underline{\theta}_K,\bar{w})$, we have

$$V_{\mathsf{I}}'(w) - V_{\emptyset}'(w) = m_K - V_{\emptyset}'(w) \geqslant m_K - V_{\emptyset}'(\underline{\theta}_K) = 0,$$

which implies that $V_{\mathsf{I}}-V_{\emptyset}$ is increasing on $[\underline{\theta}_K,\bar{w}]$. Consequently, we have $0\leqslant V_{\mathsf{I}}(w)-V_{\emptyset}(w)\leqslant K$ for $w\in(\underline{\theta}_K,\bar{w})$.

EC.3.9. Proof of Proposition 6

The proof of Proposition 6 is rather intricate, which takes a total of four key steps. These steps illustrate how to identify thresholds $\bar{\vartheta}$ and $\underline{\vartheta}$ in computation. Furthermore, these steps help us establish $\underline{\theta}^0$ in Proposition EC.1. As Condition 3 contains two cases, we consider these cases separately below.

EC.3.9.1. Condition 1 and $K<\bar{K}_1$. In Step 1, fixing any $\underline{\theta}$, we identify bound $\hat{\mathbf{w}}$ and slope c as functions of $\underline{\theta}$ to satisfy (42) and (43).

Lemma EC.4. *For any $\underline{\theta}\in(0,\bar{w})$, there exist unique values $\widetilde{w}(\underline{\theta})\in(\underline{\theta},\bar{w})$ and $C(\underline{\theta})$, in place of $\hat{\mathbf{w}}$ and c, respectively, such that value-matching and smooth-pasting conditions (42) and (43) are satisfied at $\underline{\vartheta}=\underline{\theta}$.*

Proof. For any $\widetilde{w} \in [\underline{\theta}, \bar{w})$, define

$$C_1(\widetilde{w}, \underline{\theta}) = (\mathcal{V}_{\widetilde{w}}(\underline{\theta}) - \underline{v} - \underline{\theta})\underline{\theta}^{-r/\rho} \text{ and } C_2(\widetilde{w}, \underline{\theta}) = \rho/r \cdot (\mathcal{V}'_{\widetilde{w}}(\underline{\theta}) - 1)\underline{\theta}^{1-r/\rho}. \tag{EC.70}$$

It follows from Lemma 2(iii) that $C_1(\widetilde{w}, \underline{\theta})$ is decreasing in $\widetilde{w}$ and $C_2(\widetilde{w}, \underline{\theta})$ is increasing in $\widetilde{w}$ on $[\underline{\theta}, \bar{w})$. Note that

$$\begin{aligned} C_1(\underline{\theta}, \underline{\theta}) &= (\mathcal{V}_{\underline{\theta}}(\underline{\theta}) - \underline{v} - \underline{\theta})\underline{\theta}^{-r/\rho} \\ &= \left(\frac{\mu R - c - (\rho - r)\underline{\theta}}{r} - \underline{v} - \underline{\theta}\right) \cdot \underline{\theta}^{-r/\rho} > -\frac{\rho}{r}\underline{\theta}^{1-r/\rho} = C_2(\underline{\theta}, \underline{\theta}), \end{aligned}$$

where the second and the third equalities follow from the boundary conditions at $\underline{\theta}$ (see Lemma EC.2), and the inequality follows from Assumption 1. Besides, Lemma EC.2(v) demonstrates that $C_1(\widetilde{w}, \underline{\theta}) \to -\infty$ and $C_2(\widetilde{w}, \underline{\theta}) \to \infty$ as $\widetilde{w} \uparrow \bar{w}$. In view of Lemma EC.2(i), both $C_1(\widetilde{w}, \underline{\theta})$ and $C_2(\widetilde{w}, \underline{\theta})$ are continuous in $\widetilde{w}$. Hence, there exists a unique $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ such that $C_1(\widetilde{w}(\underline{\theta}), \underline{\theta}) = C_2(\widetilde{w}(\underline{\theta}), \underline{\theta})$. Let $C(\underline{\theta}) := C_1(\widetilde{w}(\underline{\theta}), \underline{\theta})$. Then, $\widetilde{w}(\underline{\theta})$ and $C(\underline{\theta})$ satisfy (42)–(43), as desired. ◻

Step 2 determines an interval to further identify $\underline{\theta}$.

Lemma EC.5. *Value* $\underline{\theta}^0 := \inf\{\underline{\theta} \in (0, \bar{w}) \mid \widetilde{w}'(\underline{\theta}) \geqslant 0\}$ *is well defined. Furthermore, we have that* $\widetilde{w}(\underline{\theta})$ *is decreasing and* $C(\underline{\theta})$ *is increasing for* $\underline{\theta} \in (0, \underline{\theta}^0)$*, with* $\widetilde{w}'(\underline{\theta}^0) = 0$*. Moreover,* $C(\underline{\theta}) > 0$ *for any* $\underline{\theta} \in (0, \underline{\theta}^0)$*.*

Proof. Define

$$h(\widetilde{w}, \underline{\theta}) := \mathcal{V}_{\widetilde{w}}(\underline{\theta}) - \underline{v} - \underline{\theta} - \rho/r \cdot (\mathcal{V}'_{\widetilde{w}}(\underline{\theta}) - 1)\underline{\theta}. \tag{EC.71}$$

By Lemma 2(iii), $h(\widetilde{w}, \underline{\theta})$ is decreasing in $\widetilde{w}$. Besides, $h(\widetilde{w}(\underline{\theta}), \underline{\theta}) = 0$. Note that $h(\widetilde{w}, \underline{\theta})$ is continuously differentiable in $\widetilde{w}$ and $\underline{\theta}$ by Lemma EC.2(i) and (ii). Hence, $\widetilde{w}(\underline{\theta})$ is continuously differentiable in $\underline{\theta}$.

Since $h(\widetilde{w}, 0) = \mathcal{V}_{\widetilde{w}}(0) - \underline{v}$, we have $\widetilde{w}(0) = \hat{\mathrm{w}} > 0$. Besides, it follows from $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ that $\lim_{\underline{\theta} \uparrow \bar{w}} \widetilde{w}(\underline{\theta}) = \bar{w} > \hat{\mathrm{w}}$.

Write $h_1(\widetilde{w}, \underline{\theta}) = \partial h(\widetilde{w}, \underline{\theta})/\partial \widetilde{w}$ and $h_2(\widetilde{w}, \underline{\theta}) = \partial h(\widetilde{w}, \underline{\theta})/\partial \underline{\theta}$. Then, we have $h_1(\widetilde{w}, \underline{\theta}) < 0$,

$$h_2(\widetilde{w}, \underline{\theta}) = \frac{\rho - r - (\rho - r)\mathcal{V}'_{\widetilde{w}}(\underline{\theta}) - \rho\underline{\theta}\mathcal{V}''_{\widetilde{w}}(\underline{\theta})}{r}, \text{ and } \widetilde{w}'(\underline{\theta}) = -\frac{h_2(\widetilde{w}(\underline{\theta}), \underline{\theta})}{h_1(\widetilde{w}(\underline{\theta}), \underline{\theta})}.$$

It follows from $K < \underline{K}_1$ and Lemma 4 that $m^K = \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K) > 1$, which implies that $\mathcal{V}'_{\hat{\mathrm{w}}}(0) > 1$ by the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$. Therefore, we have $h_2(\widetilde{w}(0), 0) = (\rho - r - (\rho - r)\mathcal{V}'_{\hat{\mathrm{w}}}(0))/r < 0$, which in turn gives $\widetilde{w}'(0) < 0$. Therefore, $\widetilde{w}(\underline{\theta})$ is strictly decreasing in $\underline{\theta}$ when $\underline{\theta}$ is near 0.

It follows from $\lim_{\underline{\theta}\uparrow\bar{w}} \widetilde{w}(\underline{\theta}) = \bar{w} > \hat{\mathrm{w}} = \widetilde{w}(0)$ and the continuity of $\widetilde{w}'(\underline{\theta})$ in $\underline{\theta}$ that value $\underline{\theta}^0 := \inf\{\underline{\theta} \in (0,\bar{w}) \mid \widetilde{w}'(\underline{\theta}) \geqslant 0\}$ is well defined, satisfying $\widetilde{w}'(\underline{\theta}) < 0$ for any $\underline{\theta} \in [0,\underline{\theta}^0)$ and $\widetilde{w}'(\underline{\theta}^0) = 0$. Consequently, we have

$$\rho - r - (\rho - r)\mathcal{V}'_{\widetilde{w}(w)}(w) - \rho w \mathcal{V}''_{\widetilde{w}(w)}(w) < 0 \text{ for any } w \in [0,\underline{\theta}^0), \text{ and} \tag{EC.72}$$

$$\rho - r - (\rho - r)\mathcal{V}'_{\widetilde{w}(\underline{\theta}^0)}(\underline{\theta}^0) - \rho\underline{\theta}^0 \mathcal{V}''_{\widetilde{w}(\underline{\theta}^0)}(\underline{\theta}^0) = 0. \tag{EC.73}$$

Note that $\widetilde{w}(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$. We claim that $C(\underline{\theta})$ is increasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$. In fact, for any $\underline{\theta} \in (0,\underline{\theta}^0)$, we have

$$\begin{aligned} C'(\underline{\theta}) = C_1'(\widetilde{w}(\underline{\theta}),\underline{\theta}) &= \frac{\mathrm{d}}{\mathrm{d}\underline{\theta}}\left[\left(\mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - \underline{v} - \underline{\theta}\right)\underline{\theta}^{-r/\rho}\right] \\ &= \left(\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) + \widetilde{w}'(\underline{\theta}) \cdot \frac{\partial \mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta})}{\partial \widetilde{w}(\underline{\theta})} - 1\right)\underline{\theta}^{-r/\rho} - \frac{r}{\rho}\left(\mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - \underline{v} - \underline{\theta}\right)\underline{\theta}^{-r/\rho-1} \\ &= \widetilde{w}'(\underline{\theta}) \cdot \frac{\partial \mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta})}{\partial \widetilde{w}(\underline{\theta})} > 0, \end{aligned}$$

where the last equality follows from $C_1(\widetilde{w}(\underline{\theta}),\underline{\theta}) = C_2(\widetilde{w}(\underline{\theta}),\underline{\theta})$, and the inequality follows from $\widetilde{w}'(\underline{\theta}) < 0$ and $\partial\mathcal{V}_{\widetilde{w}}(w)/\partial\widetilde{w} < 0$ due to Lemma 2(iii).

Note that $\mathcal{V}'_{\hat{\mathrm{w}}}(0) > 1$. Hence, by the continuity of $\widetilde{w}(\underline{\theta})$ in $\underline{\theta}$ and Lemma EC.2(i) and (ii), there exists a number $\epsilon > 0$ such that $\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) > 1$ for any $\underline{\theta} \in (0,\epsilon)$. Consequently, $C(\underline{\theta}) = C_2(\widetilde{w}(\underline{\theta}),\underline{\theta}) > 0$ for any $\underline{\theta} \in (0,\epsilon)$, which implies that $C(\underline{\theta}) > 0$ for any $\underline{\theta} \in (0,\underline{\theta}^0)$ by noting that $C(\underline{\theta})$ is strictly increasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$. □

Next, in Step 3, we define the upper threshold $\bar{\theta}$ as a function of $\underline{\theta}$, such that smooth pasting condition (45) is satisfied.

Lemma EC.6. *We have:*

(i) *for any $\underline{\theta} \in (0,\underline{\theta}^0)$, the threshold*

$$\bar{\theta}(\underline{\theta}) := \inf\left\{w > \underline{\theta} \mid \mathcal{V}'_{\widetilde{w}(\underline{\theta})}(w) \leqslant 1 + C(\underline{\theta})r/\rho \cdot w^{r/\rho-1}\right\} \tag{EC.74}$$

is well defined;

(ii) *as a function of $\underline{\theta}$, threshold $\bar{\theta}(\underline{\theta})$ is decreasing and continuous in $\underline{\theta}$ on $[0,\underline{\theta}^0)$; and*

(iii) $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \bar{\theta}(\underline{\theta}) = \underline{\theta}^0$.

Proof. (i) Define $\Psi(w,\underline{\theta}) = \mathcal{V}'_{\widetilde{w}(\underline{\theta})}(w) - 1 - C(\underline{\theta})r/\rho \cdot w^{r/\rho-1}$. It follows from Lemma EC.5 and Lemma 2(iii) that $\Psi(w,\underline{\theta})$ is decreasing in $\underline{\theta}$ on $(0,\underline{\theta}^0)$. Therefore, for any $w \in (0,\underline{\theta})$, we have

$\Psi(w,\underline{\theta}) < \Psi(w,w) = 0$, which implies that $\underline{\theta} = \inf\{w \geqslant 0 \mid \Psi(w,\underline{\theta}) = 0\}$ as $\Psi(\underline{\theta},\underline{\theta}) = 0$. Moreover, we have

$$\begin{aligned}\frac{\partial \Psi}{\partial w}(\underline{\theta},\underline{\theta}) &= \mathcal{V}''_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - C(\underline{\theta})r/\rho\cdot(r/\rho-1)\underline{\theta}^{r/\rho-2}\\ &= \mathcal{V}''_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - C_2(\widetilde{w}(\underline{\theta}),\underline{\theta})r/\rho\cdot(r/\rho-1)\underline{\theta}^{r/\rho-2}\\ &= \mathcal{V}''_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - \rho/r\cdot(\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\underline{\theta})-1)\underline{\theta}^{1-r/\rho}r/\rho\cdot(r/\rho-1)\underline{\theta}^{r/\rho-2}\\ &= \mathcal{V}''_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - (\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\underline{\theta})-1)\cdot(r/\rho-1)/\underline{\theta}\\ &> 0,\end{aligned}$$

where the last inequality follows from (EC.72). This implies that $\Psi(w,\underline{\theta}) > 0$ for $w \in (\underline{\theta},\underline{\theta}+\epsilon)$ with some $\epsilon > 0$. According to Lemma EC.2(i), $\Psi(w,\underline{\theta})$ is continuous in w. Besides, we have

$$\Psi(\widetilde{w}(\underline{\theta}),\underline{\theta}) = \mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\widetilde{w}(\underline{\theta})) - (1 + C(\underline{\theta})r/\rho\cdot\widetilde{w}(\underline{\theta})^{r/\rho-1}) = -\left(1 + C(\underline{\theta})r/\rho\cdot\widetilde{w}(\underline{\theta})^{r/\rho-1}\right) < 0$$

and $\Psi(\underline{\theta},\underline{\theta}) = 0$. Hence, $\bar{\theta}(\underline{\theta}) = \inf\{w > \underline{\theta} \mid \Psi(w,\underline{\theta}) \leqslant 0\} = \inf\left\{w > \underline{\theta} \mid \mathcal{V}'_{\widetilde{w}(\underline{\theta})}(w) \leqslant 1 + C(\underline{\theta})r/\rho\cdot w^{r/\rho-1}\right\}$ is well defined and satisfies $\bar{\theta}(\underline{\theta}) < \widetilde{w}(\underline{\theta})$.

(ii) This follows immediately by noting that $\Psi(w,\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$ and is continuous in w and $\underline{\theta}$.

(iii) According to (EC.73), we have $\frac{\partial\Psi}{\partial w}(\underline{\theta}^0,\underline{\theta}^0) = 0$, which implies that $\Psi(\cdot,\underline{\theta}^0)$ attains its local maximum at $\underline{\theta}^0$. Hence, we have $\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\bar{\theta}(\underline{\theta}) = \underline{\theta}^0$ by using $\bar{\theta}(\underline{\theta}) \geqslant \underline{\theta}$. ◻

Finally, in Step 4, we find an appropriate $\underline{\vartheta}$ to satisfy (44), and define $(\mathsf{c},\hat{\mathbf{w}},\bar{\vartheta})$ as $(C(\underline{\vartheta}),\widetilde{w}(\underline{\vartheta}),\bar{\theta}(\underline{\vartheta}))$. To this end, we define function

$$\psi(\underline{\theta}) := \mathcal{V}_{\widetilde{w}(\underline{\theta})}(\bar{\theta}(\underline{\theta})) - \left[\underline{v} + \bar{\theta}(\underline{\theta}) + C(\underline{\theta})\left(\bar{\theta}(\underline{\theta})\right)^{r/\rho}\right]. \tag{EC.75}$$

In order to satisfy (44), we hope to identify the value $\underline{\vartheta}$ such that $\psi(\underline{\vartheta}) = K$, whose existence is guaranteed by the following result.

Lemma EC.7. *Function $\psi(\underline{\theta})$ is continuous and decreasing in $\underline{\theta}$ on $(0,\underline{\theta}^0)$, and satisfies*

$$\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\psi(\underline{\theta}) = 0, \text{ and } \lim_{\underline{\theta}\downarrow 0}\psi(\underline{\theta}) > K.$$

Consequently, there exists a unique number $\underline{\vartheta} \in (0,\underline{\theta}^0)$ such that $\psi(\underline{\vartheta}) = K$.

Proof. Note that

$$\psi(\underline{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}(\underline{\theta})}[\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(y) - (1 + C(\underline{\theta})r/\rho\cdot y^{r/\rho-1})]\mathrm{d}y = \int_{\underline{\theta}}^{\bar{\theta}(\underline{\theta})}\Psi(y,\underline{\theta})\mathrm{d}y.$$

Fix any $\underline{\theta}^1 < \underline{\theta}^2$ in $(0, \underline{\theta}^0)$. We have

$$\psi(\underline{\theta}^1) = \int_{\underline{\theta}^1}^{\bar{\theta}(\underline{\theta}^1)} \Psi(y, \underline{\theta}^1)\mathrm{d}y > \int_{\underline{\theta}^2}^{\bar{\theta}(\underline{\theta}^2)} \Psi(y, \underline{\theta}^1)\mathrm{d}y > \int_{\underline{\theta}^2}^{\bar{\theta}(\underline{\theta}^2)} \Psi(y, \underline{\theta}^2)\mathrm{d}y = \psi(\underline{\theta}^2),$$

where the first inequality follows from $\Psi(y, \underline{\theta}^1) > 0$ for $y \in \left(\underline{\theta}^1, \underline{\theta}^2\right) \cup \left(\bar{\theta}(\underline{\theta}^2), \bar{\theta}(\underline{\theta}^1)\right)$, and the second inequality uses the fact that $\Psi(y, \underline{\theta})$ is decreasing in $\underline{\theta}$ on $(0, \underline{\theta}^0)$. Hence, $\psi(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $(0, \underline{\theta}^0)$. The continuity of $\psi(\underline{\theta})$ follows from Lemma EC.2(i) and (ii).

Since $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \bar{\theta}(\underline{\theta}) = \underline{\theta}^0$, we have $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \psi(\underline{\theta}) = 0$. Note that $\tilde{w}(0) = \hat{\mathrm{w}}$ and $C_2(\tilde{w}, 0) = 0$ for any $\tilde{w} \in [0, \bar{w})$. Hence, we have $\lim_{\underline{\theta}\downarrow 0} C(\underline{\theta}) = 0$ and thus $\lim_{\underline{\theta}\downarrow 0} \bar{\theta}(\underline{\theta}) = \inf\{w > 0 \mid \mathcal{V}'_{\hat{\mathrm{w}}}(w) = 1\}$. This yields

$$\begin{aligned}\lim_{\underline{\theta}\downarrow 0} \psi(\underline{\theta}) = \int_0^{\infty} (\mathcal{V}'_{\hat{\mathrm{w}}}(y) - 1)^+ dy > \int_0^{\bar{\theta}^K} (\mathcal{V}'_{\hat{\mathrm{w}}}(y) - 1) dy \\ = \mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) - \underline{v} - \bar{\theta}^K = K + (m^K - 1)\bar{\theta}^K > K,\end{aligned}$$

where the first inequality follows from the facts that $\mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K) = m^K > 1$ and that $\mathcal{V}'_{\hat{\mathrm{w}}}$ is non-increasing, and the last inequality holds due to $m^K > 1$. Consequently, it follows from the continuity of $\psi(\cdot)$ that there exists a unique $\underline{\vartheta} \in (0, \underline{\theta}^0)$ such that $\psi(\underline{\vartheta}) = K$. □

According to these results, the quadruple $(\hat{\mathbf{w}}, \mathsf{c}, \underline{\vartheta}, \bar{\vartheta})$ defined by $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta})$, $\mathsf{c} = C(\underline{\vartheta})$ and $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ satisfies (42)–(45). Besides, it follows from $\bar{\theta}(\underline{\theta}) < \tilde{w}(\underline{\theta})$ for $\underline{\theta} \in [0, \underline{\theta}^0)$ that $\bar{\vartheta} < \hat{\mathbf{w}}$, which further implies $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta}) < \tilde{w}(0) = \hat{\mathrm{w}}$ by noting that $\tilde{w}(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[0, \underline{\theta}^0)$. To complete the proof, we need to show that $\bar{\vartheta} > \breve{w}(\hat{\mathbf{w}})$. If it fails to hold, then we have $\mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = \mathcal{V}'_{\hat{\mathbf{w}}}(\bar{\vartheta}) = \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))$. On the other side, it follows from $\mathsf{c} > 0$ and $\underline{\vartheta} < \bar{\vartheta}$ that $\mathsf{V}'_{\mathsf{c}}(\underline{\vartheta}) > \mathsf{V}'_{\mathsf{c}}(\bar{\vartheta})$. This contradicts (43) and (45).

EC.3.9.2. Condition 2 and $K < \underline{K}$. Since most arguments are exactly the same as those as in the previous case, we only provide a sketch here. To start, we observe that $\breve{\underline{\theta}} := \frac{(\bar{V} - \underline{v})(\rho - r - \mu) - (\rho - r)\bar{w}}{\mu(\rho/r - 1)} \in (0, \underline{\theta}_K)$ satisfies $m(\breve{\underline{\theta}}) = \frac{\rho - r}{\rho - r - \mu}$ by (EC.67). Moreover, we have

$$\bar{V} - \underline{v} - \bar{w} - C^1(\breve{\underline{\theta}}) \cdot \bar{w}^{r/\rho - 1} > K \tag{EC.76}$$

since $\psi_1(\breve{\underline{\theta}}) > \psi_1(\underline{\theta}_K) = 0$.

Note that under Condition 2, $\breve{w}(\tilde{w}) = 0$ by Lemma 2(i). Hence, we will use $V_{\tilde{w}}$ instead of $\mathcal{V}_{\tilde{w}}$ in the proof. Next, we will show the desired result by the following four lemmas, which parallel Lemmas EC.4–EC.7 in Section EC.3.9.1.

Lemma EC.8. *For any $\underline{\theta} \in (\breve{\underline{\theta}}, \bar{w})$, there exists unique values $\tilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ and $C(\underline{\theta})$, in place of $\tilde{\mathrm{w}}$ and c, such that (42)–(43) are satisfied at $\underline{\vartheta} = \underline{\theta}$.*

Proof. We will use functions $C_1(\widetilde{w},\underline{\theta})$ and $C_2(\widetilde{w},\underline{\theta})$ defined as in the proof of Lemma EC.4 to obtain the desired result. In the proof of Lemma EC.4, we have established that $C_1(\widetilde{w},\underline{\theta})\to-\infty$ and $C_2(\widetilde{w},\underline{\theta})\to\infty$ as $\widetilde{w}\uparrow\bar{w}$ and thus

$$\lim_{\widetilde{w}\uparrow\bar{w}} C_1(\widetilde{w},\underline{\theta}) < \lim_{\widetilde{w}\uparrow\bar{w}} C_2(\widetilde{w},\underline{\theta}) \tag{EC.77}$$

for any $\underline{\theta}\in(0,\bar{w})$. Now, we claim that (EC.77) still holds under Condition 2 and $K<\underline{K}$ for any $\underline{\theta}\in(\breve{\underline{\theta}},\bar{w})$. In fact, it follows from Lemma EC.2(vi) that

$$\lim_{\widetilde{w}\uparrow\bar{w}} C_1(\widetilde{w},\underline{\theta}) = \Big(\bar{V}-\frac{\rho-r}{\rho-r-\mu}\bar{w}-\underline{v}+\frac{\mu}{\rho-r-\mu}\underline{\theta}\Big)\underline{\theta}^{-r/\rho} \text{ and}$$
$$\lim_{\widetilde{w}\uparrow\bar{w}} C_2(\widetilde{w},\underline{\theta}) = \frac{\rho}{r}\frac{\mu}{\rho-r-\mu}\cdot\underline{\theta}^{1-r/\rho}.$$

It is clear that

$$\frac{\lim_{\widetilde{w}\uparrow\bar{w}} C_1(\widetilde{w},\underline{\theta})}{\lim_{\widetilde{w}\uparrow\bar{w}} C_2(\widetilde{w},\underline{\theta})} = \frac{\bar{V}-\frac{\rho-r}{\rho-r-\mu}\bar{w}-\underline{v}+\frac{\mu}{\rho-r-\mu}\underline{\theta}}{\frac{\rho}{r}\frac{\mu}{\rho-r-\mu}\underline{\theta}}$$

is decreasing in $\underline{\theta}$ and takes value 1 at $\breve{\underline{\theta}}$. Hence, (EC.77) holds for any $\underline{\theta}\in(\breve{\underline{\theta}},\bar{w})$. The remaining argument is exactly the same as that for Lemma EC.4 and thus omitted. Moreover, we have the following byproduct:

$$\widetilde{w}(\breve{\underline{\theta}}) := \lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}} \widetilde{w}(\underline{\theta}) = \bar{w} \text{ and } C(\breve{\underline{\theta}}) = \frac{\rho\mu}{r(\rho-r-\mu)}(\breve{\underline{\theta}})^{1-r/\rho}, \tag{EC.78}$$

which will be used in the subsequent analysis. ▫

Lemma EC.9. *Value* $\underline{\theta}^0 := \inf\{\underline{\theta}\in(\breve{\underline{\theta}},\bar{w}) \mid \widetilde{w}'(\underline{\theta}^0)\geqslant 0\}$ *is well defined. We have* $\widetilde{w}(\underline{\theta})$ *is strictly decreasing in* $\underline{\theta}$*, and* $C(\underline{\theta})$ *is strictly increasing in* $\underline{\theta}$ *on* $(\breve{\underline{\theta}},\underline{\theta}^0)$ *with* $\widetilde{w}'(\underline{\theta}^0)=0$*. Moreover,* $C(\underline{\theta})>0$ *for any* $\underline{\theta}\in(\breve{\underline{\theta}},\underline{\theta}^0)$*.*

Proof. We only point out the differences between this proof and that of Lemma EC.5. First, we show $\widetilde{w}'(\breve{\underline{\theta}})<0$ instead of $\widetilde{w}'(0)<0$. This holds by noting that $h_2(\widetilde{w}(\breve{\underline{\theta}}),\breve{\underline{\theta}}) = \lim_{\widetilde{w}\uparrow\bar{w}} h_2(\widetilde{w},\breve{\underline{\theta}}) = (\rho-r-(\rho-r)\cdot(\rho-r)/(\rho-r-\mu))/r<0$. Second, we use the result $\lim_{\underline{\theta}\uparrow\bar{w}}\widetilde{w}(\underline{\theta}) = \lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\widetilde{w}(\underline{\theta}) = \bar{w}$ instead of $\lim_{\underline{\theta}\uparrow\bar{w}}\widetilde{w}(\underline{\theta}) = \bar{w} > \hat{\mathrm{w}} = \widetilde{w}(0)$ to establish the existence of $\underline{\theta}^0$. Finally, we use $V'_{\widetilde{w}(\breve{\underline{\theta}})}(\breve{\underline{\theta}}) = \lim_{\widetilde{w}\uparrow\bar{w}} V'_{\widetilde{w}}(\breve{\underline{\theta}}) = (\rho-r)/(\rho-r-\mu)>1$ to characterize the monotonicity of $C(\cdot)$ near $\breve{\underline{\theta}}$, instead of using $\mathcal{V}'_{\hat{\mathrm{w}}}(0)>1$ to characterize the monotonicity of $C(\cdot)$ near 0. ▫

Lemma EC.10. *For any* $\underline{\theta}\in(\breve{\underline{\theta}},\underline{\theta}^0)$*, the threshold* $\bar{\theta}(\underline{\theta})$

$$\bar{\theta}(\underline{\theta}) := \inf\left\{w>\underline{\theta} \mid V'_{\widetilde{w}(\underline{\theta})}(w)\leqslant 1+C(\underline{\theta})r/\rho\cdot w^{r/\rho-1}\right\}$$

is well defined. As a function of $\underline{\theta}$*, threshold* $\bar{\theta}(\underline{\theta})$ *is decreasing in* $\underline{\theta}$ *on* $[0,\underline{\theta}^0)$*, satisfying* $\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\bar{\theta}(\underline{\theta})=\underline{\theta}^0$ *and* $\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\bar{\theta}(\underline{\theta})=\bar{w}$*.*

Proof. The proof is the same as that for Lemma EC.6, with the range of $\underline{\theta}$ changed from $(0,\underline{\theta}^0)$ to $(\breve{\underline{\theta}},\underline{\theta}^0)$. One exception is that we need to show $\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\bar{\theta}(\underline{\theta})=\bar{w}$. To show this, we first note that for any $w\in(\breve{\underline{\theta}},\bar{w})$, we have

$$\begin{aligned}\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\Psi(w,\underline{\theta})&=\lim_{\tilde{w}\uparrow\bar{w}}\left\{V'_{\tilde{w}}(w)-1-C(\breve{\underline{\theta}})\cdot r/\rho\cdot w^{r/\rho-1}\right\}\\&=\frac{\rho-r}{\rho-r-\mu}-1-\frac{\rho\mu}{r(\rho-r-\mu)}(\breve{\underline{\theta}})^{1-r/\rho}\cdot r/\rho\cdot w^{r/\rho-1}\\&=\frac{\mu}{\rho-r-\mu}\Big[1-\big(\frac{w}{\breve{\underline{\theta}}}\big)^{r/\rho-1}\Big]>0,\end{aligned}$$

where the first equality follows from (EC.78) and Lemma EC.2(ii), and the second equality follows from Lemma EC.2(vi). This inequality, together with $\bar{\theta}(\underline{\theta})<\tilde{w}(\underline{\theta})$, yields that $\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\bar{\theta}(\underline{\theta})=\bar{w}$. ◻

Lemma EC.11. *Function $\psi(\underline{\theta})$, as defined in (EC.75), is continuous and decreasing in $\underline{\theta}$ on $(\breve{\underline{\theta}},\underline{\theta}^0)$, and satisfies*

$$\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\psi(\underline{\theta})=0\ \text{and}\ \lim_{\underline{\theta}\downarrow 0}\psi(\underline{\theta})>K.$$

Consequently, there exists a unique value $\underline{\vartheta}\in(\breve{\underline{\theta}},\underline{\theta}^0)$ such that $\psi(\underline{\vartheta})=K$.

Proof. The proof is exactly the same as that for Lemma EC.7, with the range of $\underline{\theta}$ changed from $(0,\underline{\theta}^0)$ to $(\breve{\underline{\theta}},\underline{\theta}^0)$, except that we will show $\lim_{\underline{\theta}\downarrow\underline{\theta}^0}\psi(\underline{\theta})>K$ rather than $\lim_{\underline{\theta}\downarrow 0}\psi(\underline{\theta})>K$. In fact, we have

$$\lim_{\underline{\theta}\downarrow\underline{\theta}^0}\psi(\underline{\theta})=\lim_{\tilde{w}\uparrow\bar{w}}V_{\tilde{w}}(\bar{w})-[\underline{v}+\bar{w}+C(\breve{\underline{\theta}})\,(\bar{w})^{r/\rho}]=\bar{V}-\underline{v}-\bar{w}-C(\breve{\underline{\theta}})\,(\bar{w})^{r/\rho}>K,$$

where the second equality uses Lemma EC.2(vi), and the inequality follows from (EC.76). ◻

Now we are ready to complete the proof of Proposition 6 under Condition 2 and $K<\underline{K}$. According to Lemmas EC.8–EC.11, $(\hat{\mathbf{w}},\mathsf{c},\underline{\vartheta},\bar{\vartheta})$ defined by $\hat{\mathbf{w}}=\tilde{w}(\underline{\vartheta})$, $\mathsf{c}=C(\underline{\vartheta})$ and $\bar{\vartheta}=\bar{\theta}(\underline{\vartheta})$ satisfies (42)–(45). Besides, it follows from $\bar{\theta}(\underline{\theta})<\tilde{w}(\underline{\theta})$ for $\underline{\theta}\in[\breve{\underline{\theta}},\underline{\theta}^0)$ that $\bar{\vartheta}<\hat{\mathbf{w}}$, which implies $\hat{\mathbf{w}}=\tilde{w}(\underline{\vartheta})<\tilde{w}(\underline{\theta}^0)=\bar{w}$ by using that $\tilde{w}(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[\breve{\underline{\theta}},\underline{\theta}^0)$.

EC.3.10. Proof of Proposition 7

We only consider the case in which Condition 1 and $K<\bar{K}_1$ hold, because the proof is the same for the case in which Condition 2 and $K<\underline{K}$ hold.

First, we prove (48). Similar to the proof of Proposition 2, we apply Lemma 1 and Proposition EC.3 by verifying that (EC.47)–(EC.51) all hold.

Equality (EC.47) holds by noting that (i) $\ell_t=b\mathbb{1}_{\nu_t=\mu}$; (ii) for any $t>0$, $W_{t-}\in[\underline{\vartheta},\hat{\mathbf{w}}]$ if $\mathcal{E}_{t-}=\mathsf{I}$ and $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=0$ if $w\in[\underline{\vartheta},\hat{\mathbf{w}}]$; and (iii) for any $t>0$, $W_{t-}\in(0,\bar{\vartheta}]$ if $\mathcal{E}_{t-}=0$ and $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)=0$ if $w\in(0,\bar{\vartheta}]$.

Equality (EC.48) holds by noting that $\Delta L_t > 0$ only if $\mathcal{E}_{t-} = 1$ and $W_t + H_t \mathrm{d}N_t - H_t^q \mathrm{d}Q_t > \hat{\mathbf{w}}$, as well as that $V_{\mathsf{I}}(w) = V_{\mathsf{I}}(\hat{\mathbf{w}})$ for any $w \geqslant \hat{\mathbf{w}}$.

Equality (EC.49) holds by noting that for any $t \geqslant 0$, (i) $\mathcal{E}_t = 1 - \mathcal{E}_{t-} = 1$ only if $W_{t-} \in [\bar{\vartheta}, \hat{\mathbf{w}}]$ and $V_{\mathsf{I}}(w) - V_{\emptyset}(w) = K$ if $w \in [\bar{\vartheta}, \hat{\mathbf{w}}]$; and (ii) $\mathcal{E}_t = 1 - \mathcal{E}_{t-} = 0$ only if $W_{t-} \in (0, \underline{\vartheta}]$ and $V_{\emptyset}(w) - V_{\mathsf{I}}(w) = 0$ if $w \in (0, \underline{\vartheta}]$.

Note that $q_t > 0$ only if $\breve{w}(\hat{\mathbf{w}}) > 0$, $W_{t-} = \breve{w}(\hat{\mathbf{w}})$ and $\mathcal{E}_{t-} = \mathsf{I}$. Hence, if $q_t > 0$, then we have

$$\begin{aligned} &H_t^q \cdot \big(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1\big) + \phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-}) \\ =&(\breve{w}(\hat{\mathbf{w}}) - \underline{\vartheta})\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) + \mathcal{V}_{\hat{\mathbf{w}}}(\underline{\vartheta}) - \mathcal{V}_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = 0, \end{aligned}$$

where the last equality follows from Lemma 2(ii). Hence, (EC.50) holds.

Finally, (EC.51) holds by noting that (i) if $W_{0-} \leqslant \hat{\mathbf{w}}$, then $\mathbb{E}[\phi(W_0, \mathcal{E}_{0-})] - \phi(W_{0-}, \mathcal{E}_{0-}) - \mathbb{E}\Delta L_0 = \phi(W_{0-}, \mathcal{E}_{0-}) - \phi(W_{0-}, \mathcal{E}_{0-}) = 0$; and (ii) if $W_{0-} > \hat{\mathbf{w}}$, then $\mathbb{E}[\phi(W_0, \mathcal{E}_{0-})] - \phi(W_{0-}, \mathcal{E}_{0-}) - \mathbb{E}\Delta L_0 = \phi(\hat{\mathbf{w}}, \mathcal{E}_{0-}) - \phi(W_{0-}, \mathcal{E}_{0-}) - (W_{0-} - \hat{\mathbf{w}}) = 0$.

Next, we show that functions $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ as defined in (46) satisfy the optimality condition (20)–(22). First, (22) holds by noting that $V_{\mathsf{I}}(0) = V_{\emptyset}(0) = \underline{v}$. To verify (20) and (21), we consider the following three cases separately: $w \in [0, \underline{\vartheta})$, $w \in [\underline{\vartheta}, \hat{\mathbf{w}})$, and $w \in [\hat{\mathbf{w}}, \infty)$. We will study the case of $w \in [\underline{\vartheta}, \hat{\mathbf{w}})$ before $w \in [0, \underline{\vartheta})$.

Case 1: $w \in [\underline{\vartheta}, \hat{\mathbf{w}})$. First, we prove that

$$(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0 \text{ on } [\underline{\vartheta}, \hat{\mathbf{w}}). \tag{EC.79}$$

Obviously, we have $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) = 0$ for $w \in [\breve{w}(\hat{\mathbf{w}}), \hat{\mathbf{w}})$. It remains to show that (EC.79) holds for $w \in [\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}}))$ if $\underline{\vartheta} < \breve{w}(\hat{\mathbf{w}}) < \hat{\mathbf{w}}$. For $w \in [\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}}))$, function $\mathcal{V}_{\hat{\mathbf{w}}}$ is linear and thus we have

$$\begin{aligned} (\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) &= (\mu + r)\mathcal{V}_{\hat{\mathbf{w}}}(w) - \mu \mathcal{V}_{\hat{\mathbf{w}}}(w + \beta) + \rho(\bar{w} - w)\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - (\mu R - c) + (\rho - r)w \\ &=: g_{\mathsf{I}}(w). \end{aligned}$$

Note that

$$\begin{aligned} g'_{\mathsf{I}}(w) &= (\mu + r)\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mu \mathcal{V}'_{\hat{\mathbf{w}}}(w + \beta) - \rho \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) + \rho - r \\ &= (\rho - r)\big(1 - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))\big) + \mu\big(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mathcal{V}'_{\hat{\mathbf{w}}}(w + \beta)\big) \\ &\leqslant (\rho - r)\big(1 - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))\big) + \mu\big(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}) + \beta)\big) = 0, \end{aligned}$$

where the inequality uses the concavity of $\mathcal{V}_{\hat{\mathbf{w}}}$ and the last equality uses

$$0 = \rho(\bar{w} - \breve{w}(\hat{\mathbf{w}}))\mathcal{V}''_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = (\rho - r)\big(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - 1\big) + \mu\big(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}) + \beta) - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))\big).$$

Consequently, $g_{\mathsf{I}}(w) \geqslant 0$ for all $w \in [\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}}))$, which yields (EC.79). Note that $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) = 0$ if $w \in [\underline{\vartheta}, \bar{\vartheta}]$. Hence, (20) holds by the following result, whose proof is relegated to Section EC.3.11.

Lemma EC.12. *Under the conditions stated in Proposition 7, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for $w \in [\bar{\vartheta}, \hat{\mathbf{w}})$.*

If $w \in [\bar{\vartheta}, \hat{\mathbf{w}})$, then $V_{\mathsf{I}}(w) - V_\emptyset(w) = K > 0$. To establish (21), we need to show that $0 \leqslant V_{\mathsf{I}}(w) - V_\emptyset(w) \leqslant K$ if $w \in [\underline{\vartheta}, \bar{\vartheta}]$.

Let $\Phi(w) := V_{\mathsf{I}}(w) - V_\emptyset(w)$ and $\chi(w) := \mathcal{V}'_{\hat{\mathbf{w}}}(w) - 1 - \mathsf{c} \cdot r/\rho \cdot w^{r/\rho - 1}$. Obviously, we have $\Phi(\underline{\vartheta}) = 0$ and $\chi(\underline{\vartheta}) = \chi(\bar{\vartheta}) = 0$. It follows from the proof of Lemma EC.6(i) that $\Phi'(w) = \chi(w) > 0$ for any $w \in (\underline{\vartheta}, \bar{\vartheta})$. Hence, for any $w \in [\underline{\vartheta}, \bar{\vartheta})$, we have $\Phi(w) \geqslant \Phi(\underline{\vartheta}) = 0$ and $\Phi(w) \leqslant \Phi(\bar{\vartheta}) = K$.

Case 2: $w \in [0, \underline{\vartheta})$. We claim that:

Lemma EC.13. *Under the conditions stated in Proposition 7, we have $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0$ for $w \in [0, \underline{\vartheta})$.*

Its proof is rather involved, which is relegated to Section EC.3.12. Obviously, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) = 0$ on $[0, \underline{\vartheta})$. Hence, (20) holds. Inequality (21) also holds by noting that $V_{\mathsf{I}}(w) = V_\emptyset(w)$ in this case.

Case 3: $w \in [\hat{\mathbf{w}}, \infty)$. Using the boundary condition $\mathcal{V}_{\hat{\mathbf{w}}}(\hat{\mathbf{w}}) = (\mu R - c - (\rho - r)\hat{\mathbf{w}})/r$, we have

$$(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) = r\mathcal{V}_{\hat{\mathbf{w}}}(\hat{\mathbf{w}}) - (\mu R - c) + (\rho - r)w = (\rho - r)(w - \hat{\mathbf{w}}) \geqslant 0,$$

and

$$\begin{aligned}(\mathcal{A}_\emptyset V_\emptyset)(w) &= r(\mathcal{V}_{\hat{\mathbf{w}}}(\hat{\mathbf{w}}) - K) + (\rho - r)w - R\underline{\mu} \\ &= \mu R - c - (\rho - r)\hat{\mathbf{w}} + (\rho - r)w - R\underline{\mu} - rK \\ &= R\Delta\mu - c + (\rho - r)(w - \hat{\mathbf{w}}) - rK \geqslant 0,\end{aligned}$$

where the last inequality follows as $K < \bar{V}(\hat{\mathrm{w}}) - \underline{v} = (\mu R - c - (\rho - r)\hat{\mathrm{w}})/r - R\underline{\mu}/r < (R\Delta\mu - c)/r$. Hence, (20) holds. Inequality (21) also holds since $V_{\mathsf{I}}(w) - V_\emptyset(w) = K$.

EC.3.11. Proof of Lemma EC.12

For $w \in [\bar{\vartheta}, \hat{\mathbf{w}})$, it holds that $V_\emptyset(w) = \mathcal{V}_{\hat{\mathbf{w}}}(w) - K$. Define $\varpi_0 := \inf\{w > 0 \mid \mathcal{V}'_{\hat{\mathrm{w}}}(w) = 1\}$, which is well defined by noting that $\mathcal{V}'_{\hat{\mathrm{w}}}(0) > 1$ and $\mathcal{V}'_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) = 0$. For any $w \in [\varpi_0, \hat{\mathbf{w}}]$, let

$$(\mathcal{A}_\emptyset V_\emptyset)(w) = r(\mathcal{V}_{\hat{\mathbf{w}}}(w) - K) - \rho w \mathcal{V}'_{\hat{\mathbf{w}}}(w) + (\rho - r)w - \underline{\mu}R =: g_\emptyset(w).$$

It follows from Lemma 2(iii) and $\hat{\mathbf{w}} < \hat{\mathrm{w}}$ that $\mathcal{V}'_{\hat{\mathbf{w}}}(w) < \mathcal{V}'_{\hat{\mathrm{w}}}(w)$ and $\mathcal{V}_{\hat{\mathbf{w}}}(w) > \mathcal{V}_{\hat{\mathrm{w}}}(w)$ for $w \in [0, \hat{\mathbf{w}}]$. Hence, for $w \in [\varpi_0, \hat{\mathbf{w}}]$, we have

$$g'_\emptyset(w) = r\mathcal{V}'_{\hat{\mathbf{w}}}(w) - \rho\mathcal{V}'_{\hat{\mathbf{w}}}(w) - \rho w \mathcal{V}''_{\hat{\mathbf{w}}}(w) + \rho - r = (\rho - r)\big(1 - \mathcal{V}'_{\hat{\mathbf{w}}}(w)\big) - \rho w \mathcal{V}''_{\hat{\mathbf{w}}}(w) \geqslant 0,$$

where the last inequality follows from the fact that $\mathcal{V}'_{\hat{\mathbf{w}}}(w) \leqslant \mathcal{V}'_{\hat{\mathrm{w}}}(w) \leqslant 1$ for $w \geqslant \varpi_0$ and the concavity of $\mathcal{V}_{\hat{\mathbf{w}}}$. In addition, we have

$$\begin{aligned} g_\emptyset(\varpi_0) &> r(\mathcal{V}_{\hat{\mathbf{w}}}(\varpi_0) - K) - \rho\varpi_0 + (\rho - r)\varpi_0 - \underline{\mu}R \\ &> r(\underline{v} + \varpi_0 + K - K) - \rho\varpi_0 + (\rho - r)\varpi_0 - \underline{\mu}R = 0, \end{aligned}$$

where the first inequality uses $\mathcal{V}'_{\hat{\mathbf{w}}}(\varpi_0) < \mathcal{V}'_{\hat{\mathrm{w}}}(\varpi_0) = 1$, and the second inequality follows from the fact that $\mathcal{V}_{\hat{\mathbf{w}}}(\varpi_0) > \mathcal{V}_{\hat{\mathrm{w}}}(\varpi_0) > \underline{v} + \varpi_0 + K$ (the last inequality holds because of $m^K > 1$). As a result, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for all $w \in [\varpi_0, \hat{\mathbf{w}}]$.

Next, we prove that $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for $w \in [\bar{\vartheta}, \varpi_0)$ by a contradictory argument.

Suppose, to the contradictory, that there exists a number $\varpi \in (\bar{\vartheta}, \varpi_0)$ such that $(\mathcal{A}_\emptyset V_\emptyset)(\varpi) < 0$. Then, we have $(\mathcal{A}_\emptyset V_\emptyset)(\varpi) = r(\mathcal{V}_{\hat{\mathbf{w}}}(\varpi) - K) - \rho\varpi \cdot \mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) + (\rho - r)\varpi - \underline{\mu}R < 0$, and thus

$$\mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) > \frac{(\rho - r)\varpi + r\big(\mathcal{V}_{\hat{\mathbf{w}}}(\varpi) - K - \underline{v}\big)}{\rho\varpi}. \tag{EC.80}$$

It follows from (EC.74) that $\lim_{\underline{\theta}\downarrow 0} \bar{\theta}(\underline{\theta}) = \inf\{w > 0 \mid \mathcal{V}'_{\hat{\mathrm{w}}}(w) = 1\} = \varpi_0$. Note that $\varpi > \bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ and $\varpi < \varpi_0$. Hence, it follows from Lemma EC.6 that there exists a number $\underline{\theta}' \in (0, \underline{\vartheta})$ such that $\bar{\theta}(\underline{\theta}') = \varpi$. Using Lemmas EC.5 and EC.7, we have $\tilde{w}(\underline{\theta}') > \tilde{w}(\underline{\vartheta}) = \hat{\mathbf{w}}$, $C(\underline{\theta}') < C(\underline{\vartheta}) = \mathsf{c}$, and

$$\psi(\underline{\theta}') = \mathcal{V}_{\tilde{w}(\underline{\theta}')}(\varpi) - \left[\underline{v} + \varpi + C(\underline{\theta}')\varpi^{r/\rho}\right] > \psi(\underline{\vartheta}) = K. \tag{EC.81}$$

Moreover, according to 2(iii), we have

$$\mathcal{V}_{\hat{\mathbf{w}}}(\varpi) > \mathcal{V}_{\tilde{w}(\underline{\theta}')}(\varpi) \text{ and } \mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) < \mathcal{V}'_{\tilde{w}(\underline{\theta}')}(\varpi). \tag{EC.82}$$

Consequently,

$$\begin{aligned} \mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) &> \frac{(\rho - r)\varpi + r\big(\mathcal{V}_{\tilde{w}(\underline{\theta}')}(\varpi) - K - \underline{v}\big)}{\rho\varpi} \\ &> \frac{(\rho - r)\varpi + r \cdot \left[(\underline{v} + \varpi + C(\underline{\theta}')\varpi^{r/\rho}) - \underline{v}\right]}{\rho\varpi} \\ &= 1 + r/\rho \cdot C(\underline{\theta}')\varpi^{r/\rho - 1} = \mathcal{V}'_{\tilde{w}(\underline{\theta}')}(\varpi), \end{aligned}$$

where the first inequality uses (EC.80) and (EC.82), the second inequality uses (EC.81), and the last equality follows from $\varpi = \bar{\theta}(\underline{\theta}')$ and the definition of $\bar{\theta}(\cdot)$. This reaches a contradiction with (EC.82).

EC.3.12. Proof of Lemma EC.13

The proof of Lemma EC.13 is probably the most complex proof in the paper. As mentioned in the paragraph below Proposition 7, the key step is to establish Lemma EC.16 below, which states that either $\mathcal{A}_{\mathsf{l}} V_{\mathsf{l}}$'s first-order derivative is negative, or its second-order derivative is positive on $(0, \underline{\vartheta})$. This crucial result is obtained by studying a total of four cases, which are summarized as Lemmas EC.17–EC.20.

Following from $V_{\mathsf{l}}(w) = \underline{v} + w + \mathsf{c}w^{r/\rho}$ for $w \in [0, \underline{\vartheta}]$ and

$$(\mathcal{A}_{\mathsf{l}} f - \mathcal{A}_{\emptyset} f)(w) = \mu(f(w) - f(w+\beta)) + \rho\bar{w} f'(w) - (R\Delta\mu - c),$$

we define

$$\begin{aligned} g_{\mathsf{l}}(w) &:= (\mathcal{A}_{\mathsf{l}} V_{\mathsf{l}})(w) = \mu(V_{\mathsf{l}}(w) - V_{\mathsf{l}}(w+\beta)) + \rho\bar{w} V_{\mathsf{l}}'(w) - (R\Delta\mu - c) \\ &= \mu\big(\underline{v} + w + \mathsf{c}w^{r/\rho} - V_{\mathsf{l}}(w+\beta)\big) + \rho\bar{w}(1 + \mathsf{c}r/\rho \cdot w^{r/\rho - 1}) - (R\Delta\mu - c) \end{aligned}$$

for $w > 0$.

Using the same argument as that in Lemma EC.2(i) and by the definition of $\mathcal{V}_{\tilde{w}}$ as stated in Lemma 2(ii), we can obtain the following result. Its proof is omitted for brevity.

LEMMA EC.14. *For any $\tilde{w} \in (0, \bar{w})$, we have $\mathcal{V}_{\tilde{w}} \in C^1(\mathbb{R}_+) \cap C^2(\mathbb{R}_+ \backslash \{\tilde{w}\}) \cap C^3(\mathbb{R}_+ \backslash \{\tilde{w}, \tilde{w} - \beta, \breve{w}(\tilde{w})\}) \cap C^4(\mathbb{R}_+ \backslash \{\tilde{w}, \tilde{w} - \beta, \tilde{w} - 2\beta, \breve{w}(\tilde{w})\})$.*

By the definition of V_{l} as in (46) and the smooth-pasting condition at $\underline{\vartheta}$, V_{l} is differentiable at $\underline{\vartheta}$, but may not be twice differentiable at $\underline{\vartheta}$. Therefore, $V_{\mathsf{l}} \in C^1(\mathbb{R}_{++}) \cap C^2(\mathbb{R}_{++} \backslash \{\hat{\mathbf{w}}, \underline{\vartheta}\})$, which implies that $g_{\mathsf{l}} \in C^1(\mathbb{R}_{++}) \cap C^2(\mathbb{R}_{++} \backslash \{\hat{\mathbf{w}} - \beta, \underline{\vartheta} - \beta\})$. (Here, we use $\mathbb{R}_{++}$ to denote the set of all positive numbers.) Besides, it holds that

$$g_{\mathsf{l}}'(w) = \mu\big(1 + \mathsf{c}r/\rho \cdot w^{r/\rho - 1} - V_{\mathsf{l}}'(w+\beta)\big) + \rho\bar{w} \cdot \mathsf{c}r/\rho \cdot (r/\rho - 1)w^{r/\rho - 2} \text{ and} \tag{EC.83}$$

$$g_{\mathsf{l}}''(w) = \mathsf{c}r/\rho \cdot (r/\rho - 1)w^{r/\rho - 3}[\mu w + \rho\bar{w}(r/\rho - 2)] - \mu V_{\mathsf{l}}''(w+\beta). \tag{EC.84}$$

Here, g_{l}'' may not exist at $\hat{\mathbf{w}} - \beta$ and $\underline{\vartheta} - \beta$. In this case, we follow the convention to use g_{l}'' to represent the *left*-second-order derivative of the function g_{l}'' at such a point. Similarly, we also use $\mathcal{V}_{\tilde{w}}'''(w)$ and $\mathcal{V}_{\tilde{w}}''''(w)$ to represent the left-third-order derivative and the left-fourth-order derivative of the function $\mathcal{V}_{\tilde{w}}$ at w (if needed) in the subsequent analysis.

Lemma EC.13 is equivalent to $g_{\mathsf{l}}(w) \geqslant 0$ for $w \in (0, \underline{\vartheta}]$. From (EC.79) at $\underline{\vartheta}$, we have $g_{\mathsf{l}}(\underline{\vartheta}) \geqslant 0$. Moreover, the following holds.

LEMMA EC.15. *We have $g_{\mathsf{l}}'(\underline{\vartheta}) < 0$.*

Proof. First, we consider the case that $\underline{\vartheta} \leqslant \breve{w}(\hat{\mathbf{w}})$. Evaluating (EC.54) at $\breve{w}(\hat{\mathbf{w}})$ (with $\hat{\mathbf{w}}$ replacing $\widetilde{w}$) yields

$$(\mu+r)V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mu V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})+\beta) + \rho(\bar{w}-\breve{w}(\hat{\mathbf{w}}))V''_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \rho V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) + \rho - r = 0.$$

Clearly, $V''_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = 0$. Since $\underline{\vartheta} \leqslant \breve{w}(\hat{\mathbf{w}})$, we have $V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = \mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = \mathsf{V}'_{\mathsf{c}}(\underline{\vartheta}) = 1 + \mathsf{c}r/\rho \cdot \underline{\vartheta}^{r/\rho-1}$, which, together with the above expression, implies that

$$\begin{aligned} V'_{\mathsf{l}}(\underline{\vartheta}+\beta) &= \mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}+\beta) \geqslant \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})+\beta) = V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})+\beta) \\ &= \big[(\mu+r-\rho)(1+\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1}) + \rho - r\big]/\mu. \end{aligned}$$

In the above, the first inequality follows from the convexity of $\mathcal{V}_{\hat{\mathbf{w}}}$ and the fact that $\underline{\vartheta} \leqslant \breve{w}(\hat{\mathbf{w}})$. Substituting the above inequality into (EC.83) at $\underline{\vartheta}$ yields

$$g'_{\mathsf{l}}(\underline{\vartheta}) \leqslant (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} = (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-2}(\underline{\vartheta}-\bar{w}) < 0.$$

Next, we consider the case that $\underline{\vartheta} > \breve{w}(\hat{\mathbf{w}})$. Evaluating (EC.54) at $\underline{\vartheta}$ gives

$$(\mu+r)V'_{\hat{\mathbf{w}}}(\underline{\vartheta}) - \mu V'_{\hat{\mathbf{w}}}(\underline{\vartheta}+\beta) + \rho(\bar{w}-\underline{\vartheta})V''_{\hat{\mathbf{w}}}(\underline{\vartheta}) - \rho V'_{\hat{\mathbf{w}}}(\underline{\vartheta}) + \rho - r = 0.$$

Note that $V'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = 1 + \mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1}$ and $V'_{\mathsf{l}}(\underline{\vartheta}+\beta) = V'_{\hat{\mathbf{w}}}(\underline{\vartheta}+\beta)$. Hence, we have

$$\begin{aligned} g'_{\mathsf{l}}(\underline{\vartheta}) &= \mu\big(1+\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} - V'_{\mathsf{l}}(\underline{\vartheta}+\beta)\big) + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &= (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} - \rho(\bar{w}-\underline{\vartheta})V''_{\hat{\mathbf{w}}}(\underline{\vartheta}) + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &< (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} - \frac{(\bar{w}-\underline{\vartheta})(\rho-r)(1-V'_{\hat{\mathbf{w}}}(\underline{\vartheta}))}{\underline{\vartheta}} + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &= (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} + \frac{(\bar{w}-\underline{\vartheta})(\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1}}{\underline{\vartheta}} + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &= 0, \end{aligned}$$

where the inequality follows from (EC.72) at $\underline{\vartheta}$ and $\hat{\mathbf{w}} = \widetilde{w}(\underline{\vartheta})$. □

Next, we show the following crucial result.

LEMMA EC.16. *For any $w \in (0,\underline{\vartheta})$, we have either $g'_{\mathsf{l}}(w) \leqslant 0$ or $g''_{\mathsf{l}}(w) \geqslant 0$.*

The above result, combining with Lemma EC.15, yields that $g'_{\mathsf{l}}(w) \leqslant 0$ for any $w \in (0,\underline{\vartheta}]$, which immediately concludes the result stated in Lemma EC.13. In fact, if it fails to hold, $w^\dagger := \sup\{w \in (0,\underline{\vartheta}) \mid g'_{\mathsf{l}}(w) > 0\}$ is well defined, which further implies that $g''_{\mathsf{l}}(w^\dagger) < 0$. This contradicts Lemma EC.16.

Lemma EC.16 follows immediately from Lemmas EC.17–EC.20 below.

LEMMA EC.17. *For any* $w \in [0, \bar{\vartheta} - \beta]$, *we have* $g_{\mathsf{l}}'(w) < 0$.

LEMMA EC.18. *For any* $w \in (\bar{\vartheta} - \beta, \underline{\vartheta})$ *such that* $\mathcal{V}_{\hat{\mathbf{w}}}'''(w+\beta) \leqslant 0$, *we have* $g_{\mathsf{l}}''(w) > 0$.

LEMMA EC.19. *For any* $w \in \big(0, (2 - r/\rho)\beta \wedge \underline{\vartheta}\big)$, *we have* $g_{\mathsf{l}}''(w) \geqslant 0$.

LEMMA EC.20. *For any* $w \in [(1 - r/\rho)\beta, \underline{\vartheta})$ *such that* $\mathcal{V}_{\hat{\mathbf{w}}}'''(w+\beta) > 0$, *we have* $g_{\mathsf{l}}'(w) \leqslant 0$.

In the proofs of Lemmas EC.18 and EC.20, we also need the following technical result.

LEMMA EC.21. *For any* $\widetilde{w} \in [0, \bar{w})$, *the following results hold:*

(i) *If* $2\rho < r + \mu$, *then there exists a number* $\varsigma \in [\breve{w}(\widetilde{w}), \widetilde{w})$, *such that* $\mathcal{V}_{\widetilde{w}}''' > 0$ *on* $(\varsigma, \widetilde{w})$ *and* $\mathcal{V}_{\widetilde{w}}''' < 0$ *on* $(\breve{w}(\widetilde{w}), \varsigma]$;

(ii) *Otherwise,* $\mathcal{V}_{\widetilde{w}}''' \leqslant 0$ *on* $(\breve{w}(\widetilde{w}), \widetilde{w})$.

The remaining part of this subsection is devoted to the proofs of Lemmas EC.17–EC.21. To proceed, we need some preliminary results of V_{c}. Using the explicit expression of V_{c} in (24), we obtain that V_{c} is strictly concave on $\mathbb{R}_{++}$, i.e., $\mathsf{V}_{\mathsf{c}}'' < 0$ on $\mathbb{R}_{++}$,

$$\mathsf{V}_{\mathsf{c}}'''(w) = \mathsf{c}r/\rho \cdot (r/\rho - 1)(r/\rho - 2) w^{r/\rho - 3} > 0 \quad \text{and} \tag{EC.85}$$

$$\mathsf{V}_{\mathsf{c}}''''(w) = \mathsf{c}r/\rho \cdot (r/\rho - 1)(r/\rho - 2)(r/\rho - 3) w^{r/\rho - 4} < 0 \tag{EC.86}$$

for all $w \in \mathbb{R}_{++}$. Hence, V_{c}' is strictly convex and $\mathsf{V}_{\mathsf{c}}''$ is strictly concave, which further implies that

$$\big(\mathsf{V}_{\mathsf{c}}'(w) - \mathsf{V}_{\mathsf{c}}'(w+\beta)\big) + \beta \mathsf{V}_{\mathsf{c}}''(w) < 0 \text{ and} \tag{EC.87}$$

$$\big(\mathsf{V}_{\mathsf{c}}''(w) - \mathsf{V}_{\mathsf{c}}''(w+\beta)\big) + \beta \mathsf{V}_{\mathsf{c}}'''(w) > 0 \tag{EC.88}$$

for any $w \in \mathbb{R}_{++}$.

Proof of Lemma EC.17. Using $V_{\mathsf{l}}(w) = \mathsf{V}_{\mathsf{c}}(w)$ for $w \in [0, \underline{\vartheta}]$, we have

$$g_{\mathsf{l}}'(w) = \mu\big(V_{\mathsf{l}}'(w) - V_{\mathsf{l}}'(w+\beta)\big) + \mu\beta \mathsf{V}_{\mathsf{c}}''(w) \leqslant \mu\big(\mathsf{V}_{\mathsf{c}}'(w) - \mathsf{V}_{\mathsf{c}}'(w+\beta)\big) + \mu\beta \mathsf{V}_{\mathsf{c}}''(w) < 0,$$

where the first inequality uses $V_{\mathsf{l}}'(w+\beta) \geqslant \mathsf{V}_{\mathsf{c}}'(w+\beta)$ because of $w + \beta \leqslant \bar{\vartheta}$ and (EC.74) with $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$, and the second inequality follows from (EC.87). ▫

Proof of Lemma EC.18. Define $\phi(w) := \mathcal{V}_{\hat{\mathbf{w}}}'(w) - \mathsf{V}_{\mathsf{c}}'(w)$. Since $\bar{\vartheta} = \inf\{w > \underline{\vartheta} \mid \phi(w) = 0\}$ and $\phi > 0$ over $(\underline{\vartheta}, \bar{\vartheta})$, we have $\phi'(\bar{\vartheta}) \leqslant 0$. It follows from $\mathcal{V}_{\hat{\mathbf{w}}}'''(w+\beta) \leqslant 0$ and Lemma EC.21 with $\widetilde{w}$ replaced by $\hat{\mathbf{w}}$ that $\mathcal{V}_{\hat{\mathbf{w}}}''' < 0$ on $(\breve{w}(\hat{\mathbf{w}}), w+\beta)$.

For any $w \in (\bar{\vartheta} - \beta, \underline{\vartheta})$, we have

$$\begin{aligned} g_{\mathsf{l}}''(w) &= \mu\big(\mathsf{V}_{\mathsf{c}}''(w) - V_{\mathsf{l}}''(w+\beta)\big) + \mu\beta \mathsf{V}_{\mathsf{c}}'''(w) \\ &= \mu\big(\mathsf{V}_{\mathsf{c}}''(w) - \mathcal{V}_{\hat{\mathbf{w}}}''(w+\beta)\big) + \mu\beta \mathsf{V}_{\mathsf{c}}'''(w) \\ &= \mu\big[(\mathsf{V}_{\mathsf{c}}''(w) - \mathsf{V}_{\mathsf{c}}''(w+\beta)) + \beta \mathsf{V}_{\mathsf{c}}'''(w)\big] + \mu\big(\mathsf{V}_{\mathsf{c}}''(w+\beta) - \mathcal{V}_{\hat{\mathbf{w}}}''(w+\beta)\big) \\ &> \mu\big(\mathsf{V}_{\mathsf{c}}''(w+\beta) - \mathcal{V}_{\hat{\mathbf{w}}}''(w+\beta)\big) \\ &= \mu\big(\mathsf{V}_{\mathsf{c}}''(\bar{\vartheta}) - \mathcal{V}_{\hat{\mathbf{w}}}''(\bar{\vartheta})\big) + \mu \int_{\bar{\vartheta}}^{w+\beta} \big(\mathsf{V}_{\mathsf{c}}'''(y) - \mathcal{V}_{\hat{\mathbf{w}}}'''(y)\big)\mathrm{d}y \\ &> 0, \end{aligned}$$

where the first inequality uses (EC.88), and the last inequality follows from $\mathsf{V}_{\mathsf{c}}''(\bar{\vartheta}) - \mathcal{V}_{\hat{\mathbf{w}}}''(\bar{\vartheta}) = -\phi'(\bar{\vartheta}) \geqslant 0$, $w + \beta > \bar{\vartheta}$, $\mathsf{V}_{\mathsf{c}}''' > 0$ (see (EC.85)) and $\mathcal{V}_{\hat{\mathbf{w}}}''' < 0$ on $(\breve{w}(\hat{\mathbf{w}}), w+\beta)$. □

Proof of Lemma EC.19. According to (EC.84), we have

$$\begin{aligned} g_{\mathsf{l}}''(w) &= \mathsf{c}r/\rho \cdot (r/\rho - 1) w^{r/\rho - 3} \mu[w + \beta(r/\rho - 2)] - \mu V_{\mathsf{l}}''(w+\beta) \\ &\geqslant -\mu V_{\mathsf{l}}''(w+\beta) \geqslant 0, \end{aligned}$$

where the first inequality follows from $w \leqslant (2 - r/\rho)\beta$ and the last inequality follows from the concavity of V_{l}. □

Proof of Lemma EC.20. Suppose, to the contradictory, that there exists a $w^\dagger \in [(1 - r/\rho)\beta, \underline{\vartheta})$ such that $\mathcal{V}_{\hat{\mathbf{w}}}'''(w^\dagger + \beta) > 0$ and $g_{\mathsf{l}}'(w^\dagger) > 0$. According to Lemma EC.15, there must exist a number $w^\ddagger \in (w^\dagger, \underline{\vartheta})$ such that

$$g_{\mathsf{l}}'(w^\ddagger) = 0 \text{ and } g_{\mathsf{l}}''(w^\ddagger) \leqslant 0. \tag{EC.89}$$

First, we claim that $w^\ddagger + \beta < \hat{\mathbf{w}}$. Otherwise, we have $V_{\mathsf{l}}'(w^\ddagger + \beta) = V_{\hat{\mathbf{w}}}'(w^\ddagger + \beta) = 0$ and thus

$$g_{\mathsf{l}}'(w^\ddagger) = \mu + \mathsf{c}r/\rho \cdot (w^\ddagger)^{r/\rho - 2} \mu\big[w^\ddagger - \beta(1 - r/\rho)\big] > \mu > 0,$$

where the first inequality holds as $w^\ddagger > w^\dagger \geqslant (1 - r/\rho)\beta$, leading to a contradiction.

Hence, by Lemma EC.21, we have $\mathcal{V}_{\hat{\mathbf{w}}}'''(w^\ddagger + \beta) > 0$. Furthermore, evaluating (EC.54) at $w^\ddagger + \beta$ (with $\hat{\mathbf{w}}$ replacing $\widetilde{w}$) gives

$$(\mu + r)\mathcal{V}_{\hat{\mathbf{w}}}'(w^\ddagger + \beta) - \mu \mathcal{V}_{\hat{\mathbf{w}}}'(w^\ddagger + 2\beta) + \rho(\bar{w} - w^\ddagger - \beta)\mathcal{V}_{\hat{\mathbf{w}}}''(w^\ddagger + \beta) - \rho \mathcal{V}_{\hat{\mathbf{w}}}'(w^\ddagger + \beta) + \rho - r = 0.$$

Since $\mathcal{V}_{\hat{\mathbf{w}}}'''(w^\ddagger + \beta) > 0$, we have $\mathcal{V}_{\hat{\mathbf{w}}}''' > 0$ on $[w^\ddagger + \beta, \hat{\mathbf{w}})$ by Lemma EC.21. That is, $\mathcal{V}_{\hat{\mathbf{w}}}'$ is strictly convex on $[w^\ddagger + \beta, \hat{\mathbf{w}})$, which, together with the concavity of $\mathcal{V}_{\hat{\mathbf{w}}}$, yields that

$$\begin{aligned} \mathcal{V}_{\hat{\mathbf{w}}}'(w^\ddagger + 2\beta) - \mathcal{V}_{\hat{\mathbf{w}}}'(w^\ddagger + \beta) &= \mathcal{V}_{\hat{\mathbf{w}}}'\big((w^\ddagger + 2\beta) \wedge \hat{\mathbf{w}}\big) - \mathcal{V}_{\hat{\mathbf{w}}}'(w^\ddagger + \beta) \\ &> \big(\beta \wedge (\hat{\mathbf{w}} - w^\ddagger + \beta)\big) \cdot \mathcal{V}_{\hat{\mathbf{w}}}''(w^\ddagger + \beta) \geqslant \beta \mathcal{V}_{\hat{\mathbf{w}}}''(w^\ddagger + \beta). \end{aligned}$$

Therefore, we have

$$\begin{aligned}\rho(\bar{w}-w^{\ddagger}-\beta)\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta) &= (\rho-r)\big(\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)-1\big)+\mu\big(\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+2\beta)-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\big)\\ &> (\rho-r)\big(\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)-1\big)+\mu\beta\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta),\end{aligned}$$

which, along with $\rho\bar{w}=\mu\beta$, can be rewritten as

$$(\rho-r)\big(1-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\big)>\rho(w^{\ddagger}+\beta)\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta). \tag{EC.90}$$

Since $g'_{\mathsf{l}}(w^{\ddagger})=0$, using (EC.83) we have

$$1-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)=-\mathsf{c}r/\rho\cdot(w^{\ddagger})^{r/\rho-2}\left[w^{\ddagger}-(1-r/\rho)\beta\right]. \tag{EC.91}$$

Evaluating (EC.84) at $w^{\ddagger}$ yields

$$\begin{aligned}g''_{\mathsf{l}}(w^{\ddagger})/\mu &= \mathsf{c}r/\rho\cdot(r/\rho-1)(w^{\ddagger})^{r/\rho-3}[w^{\ddagger}+\beta(r/\rho-2)]-\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\\ &= \frac{(\rho-r)(1-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta))}{w^{\ddagger}-(1-r/\rho)\beta}\cdot\frac{w^{\ddagger}+\beta(r/\rho-2)}{\rho w^{\ddagger}}-\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\\ &> \left[\frac{\rho(w^{\ddagger}+\beta)[w^{\ddagger}+\beta(r/\rho-2)]}{(w^{\ddagger}-(1-r/\rho)\beta)\cdot\rho w^{\ddagger}}-1\right]\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\\ &= -\frac{(2\rho-r)\beta^2}{(w^{\ddagger}-(1-r/\rho)\beta)\cdot\rho w^{\ddagger}}\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)>0,\end{aligned}$$

where the second equality follows from (EC.91), and the first inequality follows from (EC.90) and uses the fact that $w^{\ddagger}>w^{\dagger}\geqslant(1-r/\rho)\beta$. This reaches a contradiction with (EC.89). ▫

Proof of Lemma EC.21. According to (EC.52) and the definition of $\mathcal{V}_{\tilde{w}}$, we have

$$\mathcal{V}'''_{\tilde{w}}(w)=-\frac{(\rho-r)(2\rho-r-\mu)}{\rho^2}(\bar{w}-\tilde{w})^{\frac{\rho-r-\mu}{\rho}}(\bar{w}-w)^{\frac{-3\rho+r+\mu}{\rho}}=:\zeta(w) \tag{EC.92}$$

for $w\in\big((\tilde{w}-\beta)^+\vee\breve{w}(\tilde{w}),\tilde{w}\big)$.

In the case of $(\tilde{w}-\beta)^+\leqslant\breve{w}(\tilde{w})$, then the result stated in the lemma holds, with $\varsigma=\breve{w}(\tilde{w})$ if $2\rho<r+\mu$. Below, we consider the case of $(\tilde{w}-\beta)^+>\breve{w}(\tilde{w})$, or equivalently, $\tilde{w}-\beta>\breve{w}(\tilde{w})$.

Besides, for $w\in(\breve{w}(\tilde{w}),\tilde{w})$, we have

$$\rho(\bar{w}-w)\mathcal{V}'''_{\tilde{w}}(w)=\mu\big(\mathcal{V}''_{\tilde{w}}(w+\beta)-\mathcal{V}''_{\tilde{w}}(w)\big)+(2\rho-r)\mathcal{V}''_{\tilde{w}}(w), \tag{EC.93}$$

and thus

$$\rho(\bar{w}-w)\mathcal{V}''''_{\tilde{w}}(w)=\mu\big(\mathcal{V}'''_{\tilde{w}}(w+\beta)-\mathcal{V}'''_{\tilde{w}}(w)\big)+(3\rho-r)\mathcal{V}'''_{\tilde{w}}(w). \tag{EC.94}$$

By Lemma EC.14, $\mathcal{V}'''_{\tilde{w}}$ may not exist at $\tilde{w}-\beta$. In fact, by (EC.92), we have $\mathcal{V}'''_{\tilde{w}}((\tilde{w}-\beta)+)=\zeta(\tilde{w}-\beta)$. Evaluating (EC.93) at $(\tilde{w}-\beta)-$ (to be precise, we consider an increasing sequence of

$\{w_n\}_{n\in\mathbb{N}}$ near $\widetilde{w}-\beta$ which tends to $\widetilde{w}-\beta$ from below, evaluate (EC.93) at these w_n's and then let $n\to\infty$), we obtain

$$\rho(\bar{w}-\widetilde{w}+\beta)\mathcal{V}'''_{\widetilde{w}}(\widetilde{w}-\beta)=\mu\big(\mathcal{V}''_{\widetilde{w}}(\widetilde{w}-)-\mathcal{V}''_{\widetilde{w}}(\widetilde{w}-\beta)\big)+(2\rho-r)\mathcal{V}''_{\widetilde{w}}(\widetilde{w}-\beta).$$

In the above, as mentioned earlier, we adopt the convention to use $\mathcal{V}'''_{\widetilde{w}}(\widetilde{w}-\beta)$ to denote the left-third-order derivative of function $\mathcal{V}_{\widetilde{w}}$ at $\widetilde{w}-\beta$.

In a similar vein, evaluating (EC.93) at $(\widetilde{w}-\beta)+$ yields

$$\rho(\bar{w}-\widetilde{w}+\beta)\mathcal{V}'''_{\widetilde{w}}\big((\widetilde{w}-\beta)+\big)=\mu\big(\mathcal{V}''_{\widetilde{w}}(\widetilde{w}+)-\mathcal{V}''_{\widetilde{w}}(\widetilde{w}-\beta)\big)+(2\rho-r)\mathcal{V}''_{\widetilde{w}}(\widetilde{w}-\beta).$$

Combining the above two equations and using $\mathcal{V}''_{\widetilde{w}}(\widetilde{w}+)=0$, we have

$$\begin{aligned}\mathcal{V}'''_{\widetilde{w}}(\widetilde{w}-\beta)&=\mathcal{V}'''_{\widetilde{w}}\big((\widetilde{w}-\beta)+\big)+\frac{\mu\mathcal{V}''_{\widetilde{w}}(\widetilde{w}-)}{\rho(\bar{w}-\widetilde{w}+\beta)}\\&=\zeta(\widetilde{w}-\beta)-\frac{(\rho-r)\mu}{\rho^2(\bar{w}-\widetilde{w}+\beta)(\bar{w}-\widetilde{w})}<\zeta(\widetilde{w}-\beta).\end{aligned}\tag{EC.95}$$

We break the proof of the lemma into two cases.

(i) Suppose that $2\rho<r+\mu$, in which case (EC.92) implies that $\mathcal{V}'''_{\widetilde{w}}>0$ on $(\widetilde{w}-\beta,\widetilde{w})$. From (EC.95), $\mathcal{V}'''_{\widetilde{w}}(\widetilde{w}-\beta)$ may not be larger than 0, and thus we consider the following two cases.

Case 1: $\mathcal{V}'''_{\widetilde{w}}(\widetilde{w}-\beta)>0$. Let $w_1:=\sup\{w\in(\breve{w}(\widetilde{w}),\widetilde{w}-\beta)\mid\mathcal{V}'''_{\widetilde{w}}(w)\leqslant 0\}$. If the set is empty, we have $\mathcal{V}'''_{\widetilde{w}}>0$ on $(\breve{w}(\widetilde{w}),\widetilde{w})$. Therefore, the result stated in (i) is obtained by letting $\varsigma=\breve{w}(\widetilde{w})$.

If the above set is nonempty, then $\breve{w}(\widetilde{w})<w_1<\widetilde{w}-\beta$. Since $\mathcal{V}_{\widetilde{w}}\in C^3((\breve{w}(\widetilde{w}),\widetilde{w}-\beta))$, we have $\mathcal{V}''''_{\widetilde{w}}$ exists on $(\breve{w}(\widetilde{w}),\widetilde{w}-\beta)$. In addition, $\mathcal{V}'''_{\widetilde{w}}(w_1)=0$ and $\mathcal{V}'''_{\widetilde{w}}>0$ on $(w_1,\widetilde{w})$. Now, we prove that

$$\mathcal{V}'''_{\widetilde{w}}<0\text{ on }(\breve{w}(\widetilde{w}),w_1).\tag{EC.96}$$

Evaluating (EC.94) at w_1 and using $\mathcal{V}'''_{\widetilde{w}}(w_1)=0$, we obtain $\rho(\bar{w}-w_1)\mathcal{V}''''_{\widetilde{w}}(w_1)=\mu\mathcal{V}'''_{\widetilde{w}}(w_1+\beta)>0$ (here if $w_1=\widetilde{w}-2\beta$, the left derivatives are used), which implies that $\mathcal{V}'''_{\widetilde{w}}<0$ on $(w_1-\epsilon,w_1)$ for some $\epsilon>0$.

If (EC.96) fails to hold, then $w_2:=\sup\{w\in(\breve{w}(\widetilde{w}),w_1)\mid\mathcal{V}'''_{\widetilde{w}}(w)\geqslant 0\}$ is well defined and $w_2\in(\breve{w}(\widetilde{w}),w_1)$. Hence, we have $\mathcal{V}'''_{\widetilde{w}}(w_2)=0$ and $\mathcal{V}''''_{\widetilde{w}}(w_2+)<0$. Evaluating (EC.94) at w_2+ yields $\rho(\bar{w}-w_2)\mathcal{V}''''_{\widetilde{w}}(w_2+)=\mu\mathcal{V}'''_{\widetilde{w}}((w_2+\beta)+)$, implying that $\mathcal{V}'''_{\widetilde{w}}((w_2+\beta)+)<0$. If $w_2=\widetilde{w}-2\beta$, then $\mathcal{V}'''_{\widetilde{w}}((w_2+\beta)+)=\zeta(\widetilde{w}-\beta)>0$, a contradiction. Otherwise, $\mathcal{V}'''_{\widetilde{w}}$ exists at $w_2+\beta$, and thus $\mathcal{V}'''_{\widetilde{w}}(w_2+\beta)<0$. By the definition of w_1, we have $w_2+\beta<w_1$. Hence, $\mathcal{V}'''_{\widetilde{w}}<0$ on $(w_2,w_2+\beta]$ and thus $\mathcal{V}''_{\widetilde{w}}(w_2+\beta)<\mathcal{V}''_{\widetilde{w}}(w_2)$. Evaluating (EC.93) at w_2, together with $\mathcal{V}'''_{\widetilde{w}}(w_2)=0$, yields

$$\mu\mathcal{V}''_{\widetilde{w}}(w_2+\beta)=(\mu+r-2\rho)\mathcal{V}''_{\widetilde{w}}(w_2)<\mu\mathcal{V}''_{\widetilde{w}}(w_2),$$

which gives $\mathcal{V}''_{\tilde{w}}(w_2) > 0$. This makes a contradiction with the concavity of $\mathcal{V}_{\tilde{w}}$. Hence, (EC.96) holds, indicating that the result stated in (i) is obtained by letting $\varsigma = w_1$.

Case 2: $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) \leqslant 0$. We show that

$$\mathcal{V}'''_{\tilde{w}} < 0 \text{ on } (\breve{w}(\tilde{w}), \tilde{w}-\beta). \tag{EC.97}$$

If $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) = 0$, then by evaluating (EC.94) at $(\tilde{w}-\beta)-$ and using $\mathcal{V}'''_{\tilde{w}}(\tilde{w}) = \zeta(\tilde{w}) > 0$, we obtain that $\mathcal{V}''''_{\tilde{w}}(\tilde{w}-\beta) > 0$. (Again, the left derivatives are used.) Hence, we have $\mathcal{V}'''_{\tilde{w}} < 0$ on $(\tilde{w}-\beta-\epsilon, \tilde{w}-\beta)$ for some $\epsilon > 0$. If, on the other side, $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) < 0$, then the assertion that $\mathcal{V}'''_{\tilde{w}} < 0$ on $(\tilde{w}-\beta-\epsilon, \tilde{w}-\beta)$ for some $\epsilon > 0$ still holds by Lemma EC.14.

If (EC.97) fails to hold, then $w_3 := \sup\{w \in [\breve{w}(\tilde{w}), \tilde{w}-\beta) \mid \mathcal{V}'''_{\tilde{w}}(w) \geqslant 0\}$ is well defined, satisfying $w_3 \in [\breve{w}(\tilde{w}), \tilde{w}-\beta)$. Moreover, we have $\mathcal{V}'''_{\tilde{w}}(w_3) = 0$ and $\mathcal{V}''''_{\tilde{w}}(w_3+) < 0$. With exactly the same argument as that in the previous case for treating w_2, a contradiction can be reached. Therefore, (EC.97) holds. The result stated in (i) is obtained by using (EC.92) and letting $\varsigma = \tilde{w}-\beta$.

(ii) Next, we turn to study the case that $2\rho \geqslant r+\mu$. In this case, (EC.92) implies that $\mathcal{V}'''_{\tilde{w}} \leqslant 0$ over $(\tilde{w}-\beta, \tilde{w})$. Hence, we must have $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) < 0$ in view of (EC.95). Therefore, the argument for the second case above is valid, which leads us to the desired result. ◽

EC.4. Proofs of the Results in Sections 5

EC.4.1. Proof of Proposition 8

We only consider the case in which both Condition 1 and $K < \bar{K}_1$ hold, since the case in which both Condition 2 and $K < \underline{K}$ hold can be treated similarly. It follows from Lemma EC.7 that $\psi(\underline{\vartheta}) = K$ with ψ being decreasing on $(0, \underline{\theta}^0)$. Therefore, $\underline{\vartheta}$ is decreasing in K. Recalling $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ and using Lemma EC.6(ii), we obtain that $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ is increasing in K.

For the last assertion, we first note that under Condition 1 and $\bar{K}_1 > 0$, $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \psi(\underline{\theta}) = 0$ by Lemma EC.7, which implies $\lim_{K\downarrow 0} \underline{\vartheta} = \underline{\theta}^0$. Then, using Lemma EC.6(iii), we obtain $\lim_{K\downarrow 0} \bar{\vartheta} = \underline{\theta}^0$. Under Condition 2 and $\underline{K} > 0$, we also have $\lim_{K\downarrow 0} \underline{\vartheta} = \underline{\theta}^0 = \lim_{K\downarrow 0} \bar{\vartheta}$, by a similar argument and Lemmas EC.9 and EC.11. Hence, the desired result holds with $\theta_0 = \underline{\theta}^0$.

EC.4.2. Proof of Theorem 6

First, we consider the case in which Condition 1 and $\bar{K}_1 > 0$ hold. It follows from Proposition 8 that $\lim_{K\downarrow 0} \underline{\vartheta} = \lim_{K\downarrow 0} \bar{\vartheta} = \underline{\theta}^0$. Recall from the proof of Proposition 6 that $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta})$ and $\mathsf{c} = C(\underline{\vartheta})$. Hence, we have $\lim_{K\downarrow 0} \hat{\mathbf{w}} = \lim_{K\downarrow 0} \tilde{w}(\underline{\vartheta}) = \tilde{w}(\underline{\theta}^0)$ and $\lim_{K\downarrow 0} \mathsf{c} = \lim_{K\downarrow 0} C(\underline{\vartheta}) = C(\underline{\theta}^0)$. The result (50) is obtained by setting $\theta_0 = \underline{\theta}^0$, $\hat{\mathbf{w}}_0 = \tilde{w}(\underline{\theta}^0)$, and $\mathsf{c}_0 = C(\underline{\theta}^0)$. Moreover, $\mathcal{V}_{\hat{\mathbf{w}}}$ and V_{c} converge uniformly to $\mathcal{V}_{\hat{\mathbf{w}}_0}$ and $\mathsf{V}_{\mathsf{c}_0}$, respectively, as K approaches 0. Consequently, both value functions as defined in (46) converge to $\mathfrak{V}_{\theta_0}$ uniformly as K approaches 0. Using Proposition 7 and sending K

to zero, we conclude that functions $V_{\mathsf{I}} = V_{\emptyset} = \mathfrak{V}_{\theta_0}$ satisfy the optimality conditions (20)–(22) for $K = 0$.

The argument for the case in which Condition 2 and $\underline{K} > 0$ hold is exactly the same, and thus is omitted.

References

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

Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer Series in Statistics. Springer.

Cao, P., F. Tian, and P. Sun (2022). Comment on “optimal contract to induce continued effort”. *Management Science 68*(1), 796–808.

Protter, P. E. (2003). *Stochastic Integration and Differential Equations* (second edition ed.), Volume 21 of *Stochastic Modelling and Applied Probability*. Springer.

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

E-Companion for "Punish Underperformance with Suspension — Optimal Dynamic Contracts in the Presence of Switching Cost"

In this e-companion, we present some further discussions in Section EC.1, and provide all the proofs that are omitted from the main paper in Sections EC.2–EC.4.

EC.1. Further Discussions

This section contains four parts. Section EC.1.1 gives a heuristic derivation of the optimality condition (20)–(22) for the optimal value functions V_{I} and $V_{\emptyset}$, which appears in Section 3.2. Section EC.1.2 demonstrates how to compute the optimal contract parameters. Furthermore, we consider a special case of equal time discount in Section EC.1.3 and investigate the effect of arrival rate under fixed revenue rate in Section EC.1.4.

EC.1.1. A Heuristic Derivation of the Optimality Condition (20)–(22)

In this section, we provide a heuristic derivation of the optimality condition for the principal's utility functions and of the main features of the optimal contract, whose main idea follows from Section 4.1 in Biais et al. (2010). However, our arguments are not exactly the same, due to the presence of switching and randomization. Let $F_{\mathsf{I}}(w)$ and $F_{\emptyset}(w)$ be the principal's optimal utility function that yields an agent's utility w when the initial state is I and $\emptyset$, respectively.

For any $t \geqslant 0$, let us first characterize the evolution of the principal's utility function $F_{\mathcal{E}_{t-}}(W_{t-})$. Since the principal discounts the future utility flow at rate r, his expected flow rate of utility at time t is $rF_{\mathcal{E}_{t-}}(W_{t-})$. This must be equal to the sum of expected cash flow, the (possible) switching cost, and the expected rate of change in his continuation utility over $(t-\mathrm{d}t, t]$. Hence, we have

$$rF_{\mathcal{E}_{t-}}(W_{t-})\mathrm{d}t = [\bar{\nu}_t R - (c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}]\mathrm{d}t - \mathrm{d}L_t + \mathbb{E}_{t-}[-\kappa(\mathcal{E}_{t-}, \mathcal{E}_t) + \mathrm{d}F_{\mathcal{E}_t}(W_t)], \tag{EC.1}$$

where $\mathbb{E}_{t-}[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_{t-}]$.

Following the discussions in Section 3.2, we assume that for any $\varepsilon \in \{\mathsf{I}, \emptyset\}$, $F_\varepsilon(\cdot)$ is concave and differentiable on $\mathbb{R}_+$. The actual value function might not be differentiable on the entire domain $\mathbb{R}_+$, which is an issue frequently arising in the optimal control literature, and often addressed by the viscosity solution approach. Since this section is devoted to a heuristic derivation of the optimality equation for the optimal utility function F_ε, we assume that F_ε is smooth enough temporarily.

Recall that $\mathrm{d}L_t = \ell_t \mathrm{d}t + \Delta L_t$. Note that under any admissible IC contract, $\mathbb{1}_{\nu_t=\underline{\mu}} = \mathbb{1}_{\mathcal{E}_t=\emptyset}$ and $\mathbb{1}_{\nu_t=\mu} = \mathbb{1}_{\mathcal{E}_t=\mathsf{I}}$. Using (PK) and regarding $F_\varepsilon(w)$ as a function of (w, ε), we apply calculus of point processes to the process $(W, \mathcal{E})$ to obtain

$$\mathrm{d}F_{\mathcal{E}_t}(W_t) = \big(\rho W_{t-} + b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}} - H_t\bar{\nu}_t + q_t H_t^q - \ell_t\big) F'_{\mathcal{E}_{t-}}(W_{t-})\mathrm{d}t$$

$$+\big[F_{\mathcal{E}_{t-}}(W_{t-}-\Delta L_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big]+\big[F_{\mathcal{E}_{t-}}(W_{t-}+H_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big]\mathrm{d}N_t$$
$$+\big[F_{\mathcal{E}_{t-}}(W_{t-}-H_t^q)-F_{\mathcal{E}_{t-}}(W_{t-})\big]\mathrm{d}Q_t+\big[F_{\mathcal{E}_t}(W_t)-F_{\mathcal{E}_{t-}}(W_t)\big].$$

Plugging the above formula into (EC.1) and using $\mathbb{E}_{t-}\mathrm{d}N_t=\bar{\nu}_t\mathrm{d}t$ as well as $\mathbb{E}_{t-}\mathrm{d}Q_t=q_t\mathrm{d}t$, we have

$$\begin{aligned}rF_{\mathcal{E}_{t-}}(W_{t-})\mathrm{d}t=&\Big[R\bar{\nu}_t-(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-\ell_t+\big(\rho W_{t-}+b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-H_t\bar{\nu}_t+q_tH_t^q-\ell_t\big)F'_{\mathcal{E}_{t-}}(W_{t-})\\&+\big(F_{\mathcal{E}_{t-}}(W_{t-}+H_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big)\bar{\nu}_t+\big(F_{\mathcal{E}_{t-}}(W_{t-}-H_t^q)-F_{\mathcal{E}_{t-}}(W_{t-})\big)q_t\Big]\mathrm{d}t\\&-\Delta L_t+F_{\mathcal{E}_{t-}}(W_{t-}-\Delta L_t)-F_{\mathcal{E}_{t-}}(W_{t-})+\mathbb{E}_{t-}\big[-\kappa(\mathcal{E}_{t-},\mathcal{E}_t)+F_{\mathcal{E}_t}(W_t)-F_{\mathcal{E}_{t-}}(W_t)\big].\end{aligned}\tag{EC.2}$$

Here, ℓ_t, ΔL_t, H_t, H_t^q, q_t, and $\mathcal{E}_t$ are all control variables. Besides, the contract might be terminated at time t by paying off the promised utility to the agent instantaneously. Hence, we have $F_{\mathcal{E}_t}(W_t)\geqslant\underline{v}-W_t$. That is, $F_\varepsilon(w)\geqslant\underline{v}-w$ for any $w\in\mathbb{R}_+$ and $\varepsilon\in\{\mathsf{I},\emptyset\}$.

We first optimize the constant-order terms on the right-hand side in (EC.2). Considering that the optimized constant-order terms should be zero, we have

$$\max_{\Delta L_t\geqslant0}\big\{-\Delta L_t+F_{\mathcal{E}_{t-}}(W_{t-}-\Delta L_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big\}=0,\ \text{and}\tag{EC.3}$$
$$\max_{\mathcal{E}_t\in\{\mathsf{I},\emptyset\}}\big\{-\kappa(\mathcal{E}_{t-},\mathcal{E}_t)+F_{\mathcal{E}_t}(W_t)-F_{\mathcal{E}_{t-}}(W_t)\big\}=0.\tag{EC.4}$$

Equation (EC.3) yields that $F'_\varepsilon(w)\geqslant-1$ for any $w\in\mathbb{R}_+$ and $\varepsilon\in\{\mathsf{I},\emptyset\}$. Let $\widehat{w}_\varepsilon=\inf\{w\geqslant0\mid F'_\varepsilon(w)=-1\}$. The concavity of $F_\varepsilon(\cdot)$ implies that at any time instant t, it is optimal for the principal to pay $\Delta L_t=\max\{W_{t-}-\widehat{w}_{\mathcal{E}_{t-}},0\}$ instantaneously to the agent.

Equation (EC.4) yields that $F_{\mathsf{I}}(w)\geqslant F_\emptyset(w)$ and $F_\emptyset(w)\geqslant F_{\mathsf{I}}(w)-K$ for any $w\in\mathbb{R}_+$. Besides, $\mathcal{E}_t\neq\mathcal{E}_{t-}$ only if $-\kappa(\mathcal{E}_{t-},\mathcal{E}_{t-}^c)+F_{\mathcal{E}_{t-}^c}(W_t)-F_{\mathcal{E}_{t-}}(W_t)=0$, where ε^c is I if $\varepsilon=\emptyset$ and is $\emptyset$ if $\varepsilon=\mathsf{I}$.

Next, we consider the controls such that $\Delta L_t=0$ and $\mathcal{E}_t=\mathcal{E}_{t-}$. If we plug these values into (EC.2), the symbol "=" should be replaced by "$\leqslant$" due to the suboptimality of these controls. Comparing the dt-order terms on both sides of the resulting inequality yields

$$\begin{aligned}rF_{\mathcal{E}_{t-}}(W_{t-})\geqslant\max\Big\{&R\bar{\nu}_t-(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-\ell_t+\big(\rho W_{t-}+b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}-H_t\bar{\nu}_t+H_t^qq_t-\ell_t\big)F'_{\mathcal{E}_{t-}}(W_{t-})\\&+\big(F_{\mathcal{E}_{t-}}(W_{t-}+H_t)-F_{\mathcal{E}_{t-}}(W_{t-})\big)\bar{\nu}_t+\big(F_{\mathcal{E}_{t-}}(W_{t-}-H_t^q)-F_{\mathcal{E}_{t-}}(W_{t-})\big)q_t\Big\},\end{aligned}\tag{EC.5}$$

where the maximization is taken over the set of controls (ℓ_t,H_t,H_t^q,q_t) that satisfies $\ell_t\geqslant b\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}$, the IR constraint (5), and the IC constraint (IC).

Inequality (EC.5) can be written as two inequalities, for working and suspension states. If $\mathcal{E}_{t-}=\mathsf{I}$, by omitting the time index, (EC.5) becomes

$$\begin{aligned}rF_{\mathsf{I}}(w)\geqslant&R\mu-(c-b)+(\rho w+b)F'_{\mathsf{I}}(w)\\&+\max\Big\{-\ell-(\ell+\mu h-qh^q)F'_{\mathsf{I}}(w)+\mu(F_{\mathsf{I}}(w+h)-F_{\mathsf{I}}(w))+(F_{\mathsf{I}}(w-h^q)-F_{\mathsf{I}}(w))q\Big\},\end{aligned}\tag{EC.6}$$

where the maximization is taken over the set of (ℓ, h, h^q, q) that satisfies

$$\ell \geqslant b,\ h \geqslant \beta,\ h^q \leqslant w,\ q \geqslant 0. \tag{EC.7}$$

If $\mathcal{E}_{t-} = \emptyset$, then (EC.5) becomes

$$\begin{aligned} rF_\emptyset(w) \geqslant R\underline{\mu} + \rho w F_\emptyset'(w) + \max\Big\{ & -\ell - (\ell + \underline{\mu} h - q h^q) F_\emptyset'(w) + \underline{\mu}\big(F_\emptyset(w+h) - F_\emptyset(w)\big) \\ & + \big(F_\emptyset(w-h^q) - F_\emptyset(w)\big) q \Big\}, \end{aligned} \tag{EC.8}$$

where the maximization is taken over the set of (ℓ, h, h^q, q) that satisfies

$$\ell \geqslant 0,\ h \geqslant -w,\ h^q \leqslant w,\ q \geqslant 0. \tag{EC.9}$$

Recall that $V_{\mathsf{I}}(w) = F_{\mathsf{I}}(w) + w$ and $V_\emptyset(w) = F_\emptyset(w) + w$. Then, based on the above discussions, we have the following basic properties of V_{I} and $V_\emptyset$:

1. $V_{\mathsf{I}}(w) \geqslant \underline{v}$ and $V_\emptyset(w) \geqslant \underline{v}$ for any $w \in \mathbb{R}_+$.
2. $V_{\mathsf{I}}'(w) \geqslant 0$ and $V_\emptyset'(w) \geqslant 0$ for any $w \in \mathbb{R}_+$ (this follows from the fact that $F_{\mathsf{I}}'(w) \geqslant -1$ and $F_\emptyset'(w) \geqslant -1$).
3. Both V_{I} and $V_\emptyset$ are concave on $\mathbb{R}_+$.
4. V_{I} (resp. $V_\emptyset$) will take constant value on $[\widehat{w}_{\mathsf{I}}, \infty)$ (resp. $[\widehat{w}_\emptyset, \infty)$).
5. $V_{\mathsf{I}}(w) \geqslant V_\emptyset(w)$ and $V_\emptyset(w) \geqslant V_{\mathsf{I}}(w) - K$ for any $w \in \mathbb{R}_+$.

We proceed to analyze (EC.6), which can be rewritten as follows in terms of V_{I}:

$$\begin{aligned} rV_{\mathsf{I}}(w) \geqslant R\mu - c - (\rho - r)w + (\rho w + b)V_{\mathsf{I}}'(w) + \max\Big\{ & -\ell V_{\mathsf{I}}'(w) + \big(V_{\mathsf{I}}(w+h) - V_{\mathsf{I}}(w) - hV_{\mathsf{I}}'(w)\big)\mu \\ & + \big(V_{\mathsf{I}}(w-h^q) - V_{\mathsf{I}}(w) + h^q V_{\mathsf{I}}'(w)\big) q \Big\}, \end{aligned} \tag{EC.10}$$

where the maximization is taken over the constraints (EC.7).

Optimizing the right-hand side of (EC.10) with respect to ℓ, we have $\ell^* = \arg\max_{\ell \geqslant b}\{-\ell V_{\mathsf{I}}'(w)\} = b$ if $w \in [0, \widehat{w}_{\mathsf{I}})$, where we use the fact that $V_{\mathsf{I}}'(w) > 0$ for $w \in [0, \widehat{w}_{\mathsf{I}})$.

Optimizing the right-hand side of (EC.10) with respect to h, we have $h^* = \arg\max_{h \geqslant \beta}\{V_{\mathsf{I}}(w + h) - V_{\mathsf{I}}'(w)h\} = \beta$, by noting that $V_{\mathsf{I}}(w+h) - V_{\mathsf{I}}'(w)h$ is decreasing in h on $[0, \infty)$, since $V_{\mathsf{I}}'(w+h) - V_{\mathsf{I}}'(w) \leqslant 0$ for any $h \geqslant 0$ due to the concavity of V_{I}.

Note that $\max_{h^q \leqslant w}\{V_{\mathsf{I}}(w - h^q) - V_{\mathsf{I}}(w) + h^q V_{\mathsf{I}}'(w)\} = 0$. Hence, (EC.10) reduces to

$$rV_{\mathsf{I}}(w) \geqslant R\mu - c - (\rho - r)w - \rho(\bar{w} - w)V_{\mathsf{I}}'(w) + \mu(V_{\mathsf{I}}(w+\beta) - V_{\mathsf{I}}(w)), \tag{EC.11}$$

for $w \in \mathbb{R}_+$, which can be rewritten as $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0$ by using the operator $\mathcal{A}_{\mathsf{I}}$ defined in (18).

We next analyze (EC.8), which can be rewritten as follows in terms of $V_\emptyset$:

$$rV_\emptyset(w) \geqslant R\underline{\mu} - (\rho - r)w + \rho w V_\emptyset'(w) + \max\Big\{ -\ell V_\emptyset'(w) + \underline{\mu}\big(V_\emptyset(w+h) - V_\emptyset(w) - hV_\emptyset'(w)\big) + \big(V_\emptyset(w - h^q) - V_\emptyset(w) + h^q V_\emptyset'(w)\big)q\Big\}, \tag{EC.12}$$

where the maximization is taken over the constraint set (EC.9).

Optimizing the right-hand side of (EC.12) with respect to ℓ, we have $\ell^* = \arg\max_{\ell\geqslant 0}\{-\ell V_\emptyset'(w)\} = 0$ if $w \in [0, \widehat{w}_\emptyset)$. Optimizing the right-hand side of (EC.12) with respect to h, we have $h^* = \arg\max_{h\geqslant -w}\{-V_\emptyset'(w)h + V_\emptyset(w+h)\} = 0$, by noting that $-V_\emptyset'(w)h + V_\emptyset(w+h)$ is increasing in h for $h < 0$ and decreasing in h for $h > 0$ due to the concavity of $V_\emptyset$. Additionally, we have $\max_{h^q\leqslant w}\{V_\emptyset(w-h^q) - V_\emptyset(w) + h^q V_\emptyset'(w)\} = 0$. Consequently, (EC.12) can further reduce to

$$rV_\emptyset(w) \geqslant R\underline{\mu} - (\rho - r)w + \rho w V_\emptyset'(w), \tag{EC.13}$$

which can be rewritten as $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$.

Summarizing the above discussions yields the optimality condition (20)–(22).

EC.1.2. Computing Contract Parameters

For $K = 0$, we have the following results. Since these results have been established in the second part of the proof of Proposition 8, we omit its proof.

Proposition EC.1. (i) *Under Condition 1 and $\bar{K}_1 > 0$, we have $\theta_0 = \underline{\theta}^0$, where θ_0 and $\underline{\theta}^0$ are defined in Proposition 8 and Lemma EC.5, respectively. Correspondingly, we have $\hat{\mathbf{w}}_0 = \widetilde{w}(\theta_0)$ and $\mathsf{c}_0 = C(\theta_0)$, in which functions $\widetilde{w}(\cdot)$ and $C(\cdot)$ are defined in Lemma EC.4.*

(ii) *Under Condition 2 and $\underline{K} > 0$, define a lower bound*

$$\breve{\underline{\theta}} := \frac{(\bar{V} - \underline{v})(\rho - r - \mu) - (\rho - r)\bar{w}}{\mu(\rho/r - 1)}.$$

Similar to Lemmas EC.4 and EC.5, for any $\underline{\theta} \in (\breve{\underline{\theta}}, \bar{w})$, there exist unique values $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ and $C(\underline{\theta})$, such that if we set $\hat{\mathbf{w}} = \widetilde{w}(\underline{\theta})$, $\mathsf{c} = C(\underline{\theta})$, and $\underline{\vartheta} = \underline{\theta}$, the value-matching and smooth-pasting conditions (42) and (43) are satisfied. Furthermore, value $\underline{\theta}^0 := \inf\{\underline{\theta} \in (\breve{\underline{\theta}}, \bar{w}) \mid \widetilde{w}'(\underline{\theta}) \geqslant 0\}$ is well defined, and we have $\theta_0 = \underline{\theta}^0$, $\hat{\mathbf{w}}_0 = \widetilde{w}(\theta_0)$, and $\mathsf{c}_0 = C(\theta_0)$.

For any $\underline{\theta} \in (0, \bar{w})$, function $h(\widetilde{w}, \underline{\theta})$, as defined in (EC.71), is decreasing in $\widetilde{w}$ with $h(\widetilde{w}(\underline{\theta}), \underline{\theta}) = 0$. Hence, $\widetilde{w}(\underline{\theta})$ can be efficiently found by a binary search procedure, starting from lower bound $\underline{\theta}$ and upper bound $\bar{w}$. Consequently, $C(\underline{\theta})$ can also be immediately computed as $C_1(\widetilde{w}(\underline{\theta}), \underline{\theta})$, with $C_1(\widetilde{w}, \underline{\theta})$ defined in (EC.70). Therefore, following Proposition EC.1, in order to determine the optimal contract parameters for $K = 0$ under Condition 3, we only need to find $\underline{\theta}^0$. Based on the definition of $\underline{\theta}^0$ (see part (ii) of Proposition EC.1), this value can be determined by a line search to check at which point $\tilde{w}(\underline{\theta})$ is no longer increasing, starting from 0 under Condition 1 and $\bar{K}_1 > 0$, or from $\breve{\underline{\theta}}$ under Condition 2 and $\underline{K} > 0$.

Computation of the optimal contract parameters for $K > 0$ is more complex. We only demonstrate how to compute the control-band parameters $(\mathsf{c}, \hat{\mathbf{w}}, \bar{\vartheta}, \underline{\vartheta})$ under Condition 1 and $K < \bar{K}_1$ or under Condition 2 and $K < \underline{K}$, as the optimal contract in other cases takes a simpler form. Take the case under Condition 1 and $K < \bar{K}_1$ for illustration. Note that for any $\underline{\theta} \in (0, \underline{\theta}^0)$, the value $\bar{\theta}(\underline{\theta})$ can be determined by (EC.74) using a line search procedure. Hence, function $\psi(\underline{\theta})$, as defined in (EC.75), can be readily computed for each $\underline{\theta} \in (0, \underline{\theta}^0)$. Since, by Lemma EC.7, function $\psi(\underline{\theta})$ is decreasing in $\underline{\theta}$ with $\psi(\underline{\vartheta}) = K$, the quantity $\underline{\vartheta}$ can be efficiently found by a binary search procedure, starting from lower bound 0 and upper bound $\underline{\theta}^0$. The three other parameters, c, $\hat{\mathbf{w}}$, and $\bar{\vartheta}$, are thus immediately computed as $C(\underline{\vartheta})$, $\tilde{w}(\underline{\vartheta})$, and $\bar{\theta}(\underline{\vartheta})$. For the case under Condition 2 and $K < \underline{K}$, the only difference is that the initial lower bound for the binary search is $\breve{\underline{\theta}}$.

The above procedure can be summarized by the following four subroutines.

`Subroutine 1.` Given $\underline{\theta} \in (0, \bar{w})$, compute $\tilde{w}(\underline{\theta})$: Binary search on $[\underline{\theta}, \tilde{w}]$ to determine $\tilde{w}(\underline{\theta})$ according to $h(\tilde{w}(\underline{\theta}), \theta) = 0$ where function $h(\tilde{w}, \underline{\theta})$ is defined in (EC.71).

`Subroutine 2.` Given $\underline{\theta} \in (0, \bar{w})$, compute $C(\underline{\theta})$: Following `Subroutine 1`, we obtain $\tilde{w}(\underline{\theta})$. Then, $C(\underline{\theta}) = C_1(\tilde{w}(\underline{\theta}), \underline{\theta})$ with $C_1(\tilde{w}, \underline{\theta})$ defined in (EC.70).

`Subroutine 3.` Given $\underline{\theta} \in (0, \underline{\theta}^0)$, compute $\bar{\theta}(\underline{\theta})$: Following `Subroutines 1` and `2`, we obtain $\tilde{w}(\underline{\theta})$ and $C(\underline{\theta})$. Then, we calculate $\bar{\theta}(\underline{\theta})$ by (EC.74) using a line search procedure.

`Subroutine 4.` Given $\underline{\theta} \in (0, \underline{\theta}^0)$, compute $\psi(\underline{\theta})$: Following `Subroutines 1-3`, we obtain $\tilde{w}(\underline{\theta})$, $C(\underline{\theta})$, and $\bar{\theta}(\underline{\theta})$. Then, we compute $\psi(\underline{\theta})$, as defined in (EC.75).

With the above four steps, the optimal control-band parameters can be computed by Algorithm 1 below.

Algorithm 1 Compute $(\mathsf{c}, \hat{\mathbf{w}}, \bar{\vartheta}, \underline{\vartheta})$.

1: Line search to determine $\underline{\theta}^0$ according to $\tilde{w}'(\theta) = 0$, in which function $\tilde{w}(\theta)$ is computed according to `Subroutine 1`.

2: Binary search to determine $\underline{\vartheta}$ according to $\psi(\underline{\vartheta}) = K$, where $\psi(\underline{\vartheta})$ can be computed using `Subroutine 4`.

3: Following `Subroutines 1-3`, we obtain $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta})$, $\mathsf{c} = C(\underline{\vartheta})$, and $\underline{\vartheta} = \bar{\theta}(\underline{\vartheta})$, respectively.

EC.1.3. Equal Discount Rate

In the study of dynamic contracts without the switching option, Sun and Tian (2018) claimed, without a formal proof, that under equal discount rates, it is optimal for the principal to always induce the agent to work before contract termination. In our context with switching, this claim corresponds to never switching the agent to suspension and then working again. Here, we provide a formal proof that validates this claim for any $K \geqslant 0$.

When the two players' discount rates are the same, that is, $r = \rho$, various expressions in the main body of the paper become simpler. For example, the value $\bar{V}$ defined in (10) becomes

$$\bar{V}_e := \frac{\mu R - c}{r}, \tag{EC.14}$$

and the differential equation (25), which plays an essential role in deciding the optimal value functions, becomes

$$(\mu + r)V_e(w) - \mu V_e((w + \beta) \wedge \bar{w}) + r(\bar{w} - w)V_e'(w) - (\mu R - c) = 0. \tag{EC.15}$$

According to Lemma 3 of Sun and Tian (2018), differential equation (EC.15) with boundary condition $V_e(0) = \underline{v}$ has a unique solution V_e on $[0, \bar{w}]$, which is increasing and strictly concave, with $V_e(w) = \bar{V}_e$ for all $w \geqslant \bar{w}$. Theorem 1 still holds, in which the operators $\mathcal{A}_{\mathsf{l}}$ and $\mathcal{A}_{\emptyset}$ are simplified to

$$(\mathcal{A}_{\mathsf{l}} f)(w) = (\mu + r)f(w) - \mu f(w + \beta) + r(\bar{w} - w)f'(w) - (\mu R - c), \text{ and}$$
$$(\mathcal{A}_{\emptyset} f)(w) = rf(w) - rwf'(w) - R\underline{\mu},$$

respectively, for differentiable function f.

Furthermore, when $r = \rho$, effectively Condition 1 holds. In particular, we will show that the value function for state l is V_e defined above. Furthermore, the upper threshold $\bar{V}(\hat{\mathrm{w}}) - \underline{v}$ in Proposition 2 becomes

$$\bar{K}_e := \bar{V}_e - \underline{v}. \tag{EC.16}$$

In order to define the lower threshold for the switching cost, we need to define the value function for state $\emptyset$. Note that when $r = \rho$, function $\mathcal{V}_{\hat{\mathrm{w}}}$ becomes V_e, with $\hat{\mathrm{w}}$ being $\bar{w}$ and $\breve{w}(\hat{\mathrm{w}})$ being 0. Hence, following Lemma 4, if $K < \bar{K}_e$, there exist K-dependent values $\bar{\theta}^K \in [0, \bar{w}]$ and $m^K \in [0, V_e'(0)]$ such that

$$V_e(\bar{\theta}^K) = m^K \bar{\theta}^K + K + \underline{v}, \text{ and } V_e'(\bar{\theta}^K) = m^K.$$

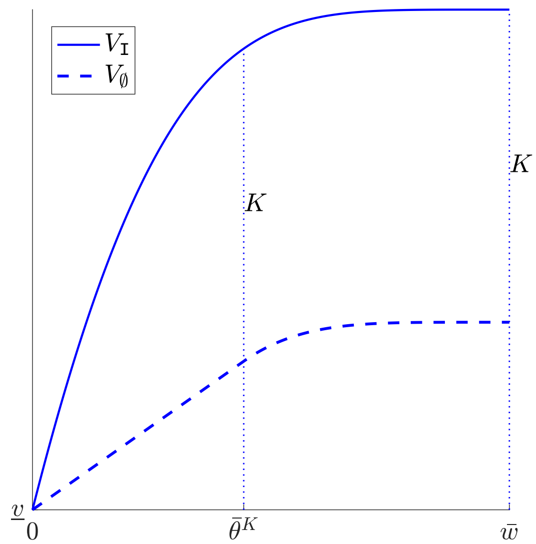
Then, similar to (30), we define the following societal value function for the suspension state:

$$V_{\emptyset}(w) = \begin{cases} m^K w + \underline{v}, & w \in \left[0, \bar{\theta}^K\right], \\ V_e(w) - K, & w \in \left[\bar{\theta}^K, \bar{w}\right]. \end{cases} \tag{EC.17}$$

Figure EC.1 depicts the value functions. It is clear that $V_{\emptyset}$ is linear over the interval $[0, \bar{\theta}^K]$. Furthermore, $V_{\mathsf{l}}(w)$ and $V_{\emptyset}(w)$ are "parallel" with a difference of K for $w \geqslant \bar{\theta}^K$.

The following theorem summarizes the optimality results.

Figure EC.1 Illustration of Optimal Societal Value Functions with Equal Discount Rates



Notes. In this figure, $r=0.5$, $\rho=0.5$, $c=b=0.2$, $R=2$, $\Delta\mu=0.7$, $K=1.5$, and $\mu=2$. Hence, $\bar{\theta}^K=0.51$, $\bar{w}=1.14$, $\bar{V}_e=7.6$, and $\underline{v}=5.2$.

Theorem EC.1. *Consider $r=\rho$. For any $w\geqslant 0$, we have*

$$U(\underline{\Gamma},\emptyset)=\underline{v}.$$

If $K\geqslant \bar{K}_e$, functions $V_{\mathsf{I}}=V_e$ and $V_{\emptyset}=\underline{v}$ satisfy (20)–(22).

If $K<\bar{K}_e$, on the other hand, functions $V_{\mathsf{I}}=V_e$ and $V_{\emptyset}$ as defined in (EC.17) satisfy (20)–(22). Furthermore, if $V_e'(\bar{\theta}^K)>1$, for any $w\geqslant\bar{\theta}^K$, we have

$$U(\Gamma^*(w;0,0,\bar{w},\bar{w}))=V_{\emptyset}(w)-w.$$

Proof. Using a similar argument as that in the proof of Proposition 2, we can show that (i) $U(\underline{\Gamma},\emptyset)=\underline{v}$, and (ii) under the condition that $K<\bar{K}_e$ and $m^K>1$, $U\big(\Gamma^*(w;0,0,\bar{w},\bar{w})\big)=V_{\emptyset}(w)-w$ for any $w\geqslant\bar{\theta}^K$ with $V_{\emptyset}$ as defined in (EC.17).

Next, we show that under condition $K\geqslant\bar{K}_e$, functions $V_{\mathsf{I}}=V_e$ and $V_{\emptyset}=\underline{v}$ satisfy (20)–(22). By the definition of V_e, it is clear that $\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}}=0$. Moreover, $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)=r\underline{v}-\underline{\mu}R=0$ for any $w\geqslant 0$. Hence, (20) holds.

Note that V_e is increasing on $[0,\bar{w}]$ (see Lemma 3 of Sun and Tian 2018). Hence, for any $w\geqslant 0$, we have $V_{\mathsf{I}}(w)-V_{\emptyset}(w)\geqslant V_e(0)-\underline{v}=0$ and $V_{\mathsf{I}}(w)-V_{\emptyset}(w)\leqslant\bar{V}_e-\underline{v}=\bar{K}_e\leqslant K$. Therefore, (21) holds. Finally, it is evident that (22) holds.

It remains to show that under condition $K < \bar{K}_e$, functions $V_{\mathsf{I}} = V_e$ and $V_\emptyset$ as defined in (EC.17) satisfy (20)–(22). Obviously, $\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}} = 0$. Moreover, we have

$$(\mathcal{A}_\emptyset V_\emptyset)(w) = rV_\emptyset(w) - rwV_\emptyset'(w) - \underline{\mu}R = rw\left(\frac{V_\emptyset(w) - V_\emptyset(0)}{w} - V_\emptyset'(w)\right) \geqslant 0,$$

where the equality follows from $V_\emptyset(0) = \underline{v}$, and the inequality follows from the concavity of $V_\emptyset$. Hence, (20) holds.

If $w \geqslant \bar{\theta}^K$, then $V_{\mathsf{I}}(w) - V_\emptyset(w) = K$. If $w \in [0, \bar{\theta}^K]$, then $V_{\mathsf{I}}'(w) - V_\emptyset'(w) = V_e'(w) - V_e'(\bar{\theta}^K) \geqslant 0$ due to the concavity of V_e, which implies that $V_{\mathsf{I}}(w) - V_\emptyset(w) \geqslant V_{\mathsf{I}}(0) - V_\emptyset(0) = 0$ and $V_{\mathsf{I}}(w) - V_\emptyset(w) \leqslant V_{\mathsf{I}}(\bar{\theta}^K) - V_\emptyset(\bar{\theta}^K) = K$. Hence, (21) holds. It is straightforward to see that (22) holds. ◻

Therefore, in the equal discount case, contract $\Gamma^*(w_e^*; 0, 0, \bar{w}, \bar{w})$ is optimal if $K < \bar{K}_e$ and $m^K > 1$, in which $w_e^* \in [0, \bar{w}]$ is the unique maximizer of function V_e such that $w_e^* > \bar{\theta}^K$. Otherwise, it is optimal for the principal not to hire the agent at all. Note that because the threshold $\underline{\theta}$ in contract $\Gamma^*(w; 0, 0, \bar{w}, \bar{w})$ is zero, the principal does not direct the agent to stop working until the promised utility has reached 0. At this point, the promised utility cannot become positive again, and the contract is terminated. Therefore, in all these cases, it is never optimal for the principal to direct the agent to stop working and restart later.

EC.1.4. Effect of Arrival Rate Under Fixed Revenue Rate

In this section, we investigate the effect of arrival uncertainty on the optimal contract. In particular, we fix the revenue rates per unit of time ($R\mu$ and $R\underline{\mu}$), the cost rates (c and b), and the switching cost (K), and see how the optimal contract changes with the revenue R. In particular, when R approaches zero, the arrival rate effectively approaches infinity, and the system behaves more like a deterministic one. In this case, mitigating uncertainty effectively removes the rent that the agent is able to obtain. The system should become efficient. On the flip side, if R approaches infinity, the system is extremely uncertain.

For this purpose, we fix $A := R\Delta\mu$ and $B := R\underline{\mu}$ and let them be fixed input parameters. In this setup, we can write all results as well as relevant quantities appeared in the paper in terms of A and B, with $\underline{\mu}$ and $\Delta\mu$ replaced as $\underline{\mu} = B/R$ and $\Delta\mu = A/R$, respectively. It is easy to check that quantities $\underline{v}$, $\bar{w}$, $\bar{V}(\cdot)$, and $\bar{V}$ are all independent of R. Hence, the results in Theorem 2 still hold. Moreover, we have the following result, which explores two extreme cases, the case of extreme uncertainty (i.e., $R \uparrow \infty$) and that of no uncertainty (i.e., $R \downarrow 0$). Note that the first-best societal utility, by considering whether or not to hire the agent, is

$$V^{\mathsf{FB}} := \underline{v} + \left[\frac{R\Delta\mu - c}{r} - K\right]^+.$$

Proposition EC.2. *Fix model parameters A, B, c, b, and K.*

(i) *As $R\uparrow\infty$, it is optimal for the principal to not hire the agent if $K > \bar{V} - \underline{v} - \bar{w}$, and to hire the agent and offer contract $\bar{\Gamma}$ (paying $\beta = bR/A$ to each arrival) otherwise.*

(ii) *As $R\downarrow 0$, it is optimal for the principal not to hire the agent if $K \geqslant (A-c)/r$. If $K < (A-c)/r$, on the other hand, the principal will hire the agent and implement the contract $\Gamma^{*R}\left(\mathsf{w}_0^{*R};\ \underline{\vartheta}^R,\ (\underline{\vartheta}^R \vee \breve{w}(\hat{\mathbf{w}}^R)),\ \bar{\vartheta}^R,\ \hat{\mathbf{w}}^R\right)$ as defined in (47) and Theorem 5, in which the superscript R highlights the parameters' dependence on R. Furthermore, we have*

$$\lim_{R\downarrow 0} \mathsf{w}_0^{*R} = \lim_{R\downarrow 0} \underline{\vartheta}^R = \lim_{R\downarrow 0} \breve{w}(\hat{\mathbf{w}}^R) = \lim_{R\downarrow 0} \bar{\vartheta}^R = \lim_{R\downarrow 0} \hat{\mathbf{w}}^R = 0, \tag{EC.18}$$

and

$$\lim_{R\downarrow 0} U^R\Big(\Gamma^{*R}\left(\mathsf{w}_0^{*R};\ \underline{\vartheta}^R,\ (\underline{\vartheta}^R \vee \breve{w}(\hat{\mathbf{w}}^R)),\ \bar{\vartheta}^R,\ \hat{\mathbf{w}}^R\right)\Big) = \frac{A+B-c}{r} - K = V^{FB}. \tag{EC.19}$$

In either case, the optimal contract yields the first-best societal utility asymptotically.

Proof. First, we show part (i). Note that $R \geqslant \hat{R}$ if and only if

$$R \geqslant \frac{(A+B)[(A-c)A\rho - b(\rho - r)(A+B)]}{\rho(\rho - r)(A^2 - cA - bA - bB)} =: \breve{R}.$$

Hence, Condition 2 holds as $R\uparrow\infty$. Consequently, we have $\lim_{R\uparrow\infty} \bar{K} = \lim_{R\uparrow\infty} \bar{K}_2 = \bar{V} - \underline{v} - \bar{w}$ and $\lim_{R\uparrow\infty} \underline{K} = 0$ by (34). The result stated in part (i) follows immediately from Theorem 2.

Next, we prove part (ii). Fix any contract $\Gamma \in \mathfrak{C}$. Define $\sigma := \inf\{t \geqslant 0 \mid \mathcal{E}_t = \mathsf{I}\}$, which will take value ∞ if the principal does not hire the agent under contract Γ. We have

$$\begin{aligned}
U(\Gamma) &\leqslant \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\infty e^{-rt}\left(R\mathrm{d}N_t - c\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\mathrm{d}t\right) - \sum_{0\leqslant t\leqslant\infty} e^{-rt}\kappa(\mathcal{E}_{t-},\mathcal{E}_t)\right]\\
&= \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\infty e^{-rt}\Big(R(\mu\mathbb{1}_{\mathcal{E}_t=\mathsf{I}} + \underline{\mu}\mathbb{1}_{\mathcal{E}_t=\emptyset}) - c\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\Big)\mathrm{d}t - \sum_{0\leqslant t\leqslant\infty} e^{-rt}\kappa(\mathcal{E}_{t-},\mathcal{E}_t)\right]\\
&\leqslant \mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_0^\sigma e^{-rt}R\underline{\mu}\mathrm{d}t + \int_\sigma^\infty e^{-rt}(R\mu - c)\mathrm{d}t - e^{-r\sigma}K\right]\\
&= \frac{R\underline{\mu}}{r} + \mathbb{E}^{\bar{\nu}(\Gamma)}\left[e^{-r\sigma}\right]\left(\frac{R\Delta\mu - c}{r} - K\right)\\
&= \frac{R\underline{\mu}}{r} + \left(\frac{R\Delta\mu - c}{r} - K\right)^+,
\end{aligned}$$

where the first inequality follows by plugging (LL) into (6), and the second inequality follows from Assumption 1. Therefore, if $K \geqslant (R\Delta\mu - c)/r = (A-c)/r$, then we have $U(\Gamma) \leqslant R\underline{\mu}/r = U(\underline{\Gamma})$, which demonstrates that it is optimal for the principal not to hire the agent. (We point out this result does not depend on the value of R.)

If $K < (A-c)/r$, then we have

$$U(\Gamma) \leqslant \bar{V}(0) - K = (R\mu - c)/r - K = (A + B - c)/r - K. \tag{EC.20}$$

Denote the set of positive R's that satisfy Condition 1 as $\mathcal{R}$. Clearly, $R \in \mathcal{R}$ when it is sufficiently small. For any $R \in \mathcal{R}$, Lemma 3 holds, which demonstrates that $\widehat{\mathrm{w}}^R$ is well defined.

To show the second assertion in part (ii), we need the following limiting result:

$$\lim_{R \downarrow 0} \widehat{\mathrm{w}}^R = 0, \tag{EC.21}$$

which will be proved later using a contradictory argument. This result further implies that $\lim_{R\downarrow 0} \bar{K}_1^R = \bar{V}(0) - \underline{v} = (A-c)/r$. In fact, note that the line $w + \underline{v} + \bar{K}_1^R$ (as a function of w) is above the curve $\mathcal{V}^R_{\widehat{\mathrm{w}}^R}(w)$ for any $R \in \mathcal{R}$. Hence, we have $\bar{K}_1^R \geqslant \mathcal{V}^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - \widehat{\mathrm{w}}^R - \underline{v} = \bar{V}(\widehat{\mathrm{w}}^R) - \widehat{\mathrm{w}}^R - \underline{v}$. In addition, we have $\bar{K}_1^R \leqslant \bar{V}(\widehat{\mathrm{w}}^R) - \underline{v}$. Sending R to zero and using (EC.21), we obtain that $\lim_{R\downarrow 0} \bar{K}_1^R = (A-c)/r$. Consequently, Condition 3 holds as $R \downarrow 0$ if $K < (A-c)/r$, which demonstrates that the contract $\Gamma^{*R}\left(\mathsf{w}_0^{*R};\ \underline{\vartheta}^R,\ (\underline{\vartheta}^R \vee \breve{w}(\hat{\mathbf{w}}^R)),\ \bar{\vartheta}^R,\ \hat{\mathbf{w}}^R\right)$ as defined in (47) and Theorem 5 is well defined, establishing the second assertion in part (ii).

The limiting result (EC.18) follows immediately by noting that $\hat{\mathbf{w}}^R < \widehat{\mathrm{w}}^R$ from Proposition 6 and using (EC.21). Applying Proposition 7 and Theorem 5, we obtain that

$$\begin{aligned} U^R\left(\hat{\Gamma}^R(\mathsf{w}_0^{*R})\right) &= \mathcal{V}^R_{\hat{\mathbf{w}}^R}(\mathsf{w}_0^{*R}) - \mathsf{w}_0^{*R} - K = \max_{w \geqslant 0}\left\{\mathcal{V}^R_{\hat{\mathbf{w}}^R}(w) - w\right\} - K \\ &\geqslant \mathcal{V}^R_{\hat{\mathbf{w}}^R}(\hat{\mathbf{w}}^R) - \hat{\mathbf{w}}^R - K = \bar{V}(\hat{\mathbf{w}}^R) - \hat{\mathbf{w}}^R - K, \end{aligned}$$

which further implies that

$$\liminf_{R\downarrow 0} U^R\left(\hat{\Gamma}^R(\mathsf{w}_0^{*R})\right) \geqslant \lim_{R\downarrow 0}\left\{\bar{V}(\hat{\mathbf{w}}^R) - \hat{\mathbf{w}}^R\right\} - K = \bar{V}(0) - K$$

by (EC.21). This, combining with (EC.20), establishes (EC.19).

It remains to show (EC.21). Note that $\widehat{\mathrm{w}}^R \in [0, \bar{w})$ for any $R \in \mathcal{R}$. Hence, $\{\widehat{\mathrm{w}}^R\}_{R\in\mathcal{R}}$ is a bounded sequence. If $\lim_{R\downarrow 0} \widehat{\mathrm{w}}^R = 0$ fails to hold, according to the Bolzano–Weierstrass theorem, there exists a sequence $\{R_n\}_{n\in\mathbb{N}}$ with $R_n \in \mathcal{R}$ and $\lim_{n\to\infty} R_n = 0$, and a number $w^\ddagger \in (0, \bar{w}]$, such that $\lim_{n\to\infty} \widehat{\mathrm{w}}^{R_n} = w^\ddagger$. Then, we show that

$$\lim_{n\to\infty} V^{R_n}_{\widehat{\mathrm{w}}^{R_n}}(w) = -\infty \tag{EC.22}$$

for any $w \in [0, w^\ddagger)$. Suppose, to the contradictory, that (EC.22) fails to hold for some $w^\dagger \in [0, w^\ddagger)$. Then, we have a subsequence $\{R_{n'}\}_{n'\in\mathbb{N}}$ with $\lim_{n\to\infty} R_{n'} = 0$ such that $\lim_{n'\to\infty} V^{R_{n'}}_{\widehat{\mathrm{w}}^{R_{n'}}}(w^\dagger)$ exists and is finite. Recall from Lemma 2 that $V^R_{\widehat{\mathrm{w}}^R}(\cdot)$ is continuous and increasing. Using a diagonalization argument, we can show that there exists a further subsequence $\{R_{n''}\} \subset \{R_{n'}\}_{n' \in \mathbb{N}}$ and a finite-valued continuous function $v(\cdot)$ defined on $[w^\dagger, w^\ddagger]$ such that

$$\lim_{n' \to \infty} V^{R_{n''}}_{\widehat{\mathrm{w}}^{R_{n''}}}(w) = v(w) \tag{EC.23}$$

for any $w \in [w^\dagger, w^\ddagger]$. (First, we establish the weakly convergence of these functions at all rational numbers on $[w^\dagger, w^\ddagger]$; then we use these functions' continuity and monotonicity to show the weakly convergence on the entire interval $[w^\dagger, w^\ddagger]$.) Moreover, $v(\cdot)$ is nondecreasing on $[w^\dagger, w^\ddagger]$, with $v(w^\ddagger) = \bar{V}(w^\ddagger)$.

Rewriting (25) in terms of A, B with μ and β replaced, we obtain

$$\rho\,(\bar{w} - w)\,(V^R_{\widehat{\mathrm{w}}^R})'(w) + rV^R_{\widehat{\mathrm{w}}^R}(w) - (A + B - c) + (\rho - r)w = \frac{A+B}{R}\left[V^R_{\widehat{\mathrm{w}}^R}\Big(\big(w + \frac{bR}{A}\big) \wedge \widehat{\mathrm{w}}^R\Big) - V^R_{\widehat{\mathrm{w}}^R}(w)\right],$$

or equivalently,

$$\begin{aligned}&\rho\frac{\mathrm{d}}{\mathrm{d}w}\big[(\bar{w} - w)V^R_{\widehat{\mathrm{w}}^R}(w)\big]\\ &\quad= \frac{A+B}{R}\left[V^R_{\widehat{\mathrm{w}}^R}\Big(\big(w + \frac{bR}{A}\big) \wedge \widehat{\mathrm{w}}^R\Big) - V^R_{\widehat{\mathrm{w}}^R}(w)\right] - (\rho + r)V^R_{\widehat{\mathrm{w}}^R}(w) + (A + B - c) - (\rho - r)w.\end{aligned}$$

Integrating the above equation from w to $\widehat{\mathrm{w}}^R$ yields

$$\begin{aligned}&\rho\Big[(\bar{w} - \widehat{\mathrm{w}}^R)V^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - (\bar{w} - w)V^R_{\widehat{\mathrm{w}}^R}(w)\Big]\\ &\quad= \int_w^{\widehat{\mathrm{w}}^R}\left\{\frac{A+B}{R}\left[V^R_{\widehat{\mathrm{w}}^R}\Big(\big(u + \frac{bR}{A}\big) \wedge \widehat{\mathrm{w}}^R\Big) - V^R_{\widehat{\mathrm{w}}^R}(u)\right] - (\rho + r)V^R_{\widehat{\mathrm{w}}^R}(u) + (A + B - c) - (\rho - r)u\right\}\mathrm{d}u.\end{aligned} \tag{EC.24}$$

Note that

$$\int_w^{\widehat{\mathrm{w}}^R}\left[V^R_{\widehat{\mathrm{w}}^R}\left(\big(u + \frac{bR}{A}\big) \wedge \widehat{\mathrm{w}}^R\right) - V^R_{\widehat{\mathrm{w}}^R}(u)\right]\mathrm{d}u = \begin{cases}\int_w^{\widehat{\mathrm{w}}^R}(V^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - V^R_{\widehat{\mathrm{w}}^R}(u))\mathrm{d}u, & w \in (\widehat{\mathrm{w}}^R - \frac{bR}{A}, \widehat{\mathrm{w}}^R],\\ \frac{bR}{A}V^R_{\widehat{\mathrm{w}}^R}(\widehat{\mathrm{w}}^R) - \int_w^{w+\frac{bR}{A}} V^R_{\widehat{\mathrm{w}}^R}(u)\mathrm{d}u, & w \in [0, \widehat{\mathrm{w}}^R - \frac{bR}{A}].\end{cases}$$

Now consider Equation (EC.24) for the subsequence $\{R_{n''}\}$ and for any $w \in [w^\dagger, w^\ddagger)$. As $w \in [0, \widehat{\mathrm{w}}^R - \frac{bR}{A}]$ for sufficiently small R, by L'Hopital's rule and (EC.23), we have

$$\lim_{n'' \to \infty}\frac{\int_w^{\widehat{\mathrm{w}}^{R_{n''}}}\left[V^{R_{n''}}_{\widehat{\mathrm{w}}^{R_{n''}}}\left(\big(u + \frac{bR}{A}\big) \wedge \widehat{\mathrm{w}}^{R_{n''}}\right) - V^{R_{n''}}_{\widehat{\mathrm{w}}^{R_{n''}}}(u)\right]\mathrm{d}u}{R_{n''}} = \frac{b}{A}\bar{V}(w^\ddagger) - \frac{b}{A}v(w).$$

Therefore, letting $n'' \to \infty$ in (EC.24) and applying (EC.23), we obtain

$$\begin{aligned}&\rho\big[(\bar{w} - w^\ddagger)\bar{V}(w^\ddagger) - (\bar{w} - w)v(w)\big]\\ &\quad= \frac{b(A+B)}{A}(\bar{V}(w^\ddagger) - v(w)) + \int_w^{w^\ddagger}\left[-(\rho + r)v(u) + (A + B - c) - (\rho - r)u\right]\mathrm{d}u.\end{aligned}$$

Therefore, v is differentiable, and thus the above equality can be written as

$$\rho w v'(w) = rv(w) - (A+B-c) + (\rho - r)w,$$

by noting that $\bar{w} = b(A+B)/(\rho A)$.

Using the boundary condition $v(w^{\ddagger}) = \bar{V}(w^{\ddagger})$, we have

$$v(w) = \bar{V}(w^{\ddagger}) + w - w^{\ddagger} + \frac{\rho}{r} w^{\ddagger} \left[1 - \left(\frac{w}{w^{\ddagger}}\right)^{r/\rho}\right] \text{ for } w \in [w^{\dagger}, w^{\ddagger}],$$

which is decreasing on $[w^{\dagger}, w^{\ddagger}]$, reaching a contradiction with the fact that $v(\cdot)$ is nondecreasing on $[w^{\dagger}, w^{\ddagger}]$. Hence, (EC.22) holds.

Furthermore, we have $\lim_{n\to\infty} \mathcal{V}^{R_n}_{\hat{\mathrm{w}}^{R_n}}(0) = -\infty$ by noting that $\mathcal{V}_{\tilde{w}} \leqslant V_{\tilde{w}}$ for any $\tilde{w} \in (0, \bar{w})$. This contradicts $\mathcal{V}^{R}_{\hat{\mathrm{w}}^{R}}(0) = \underline{v}$. The proof of (EC.21) is complete. ◻

As R approaches infinity, the arrival stream is extremely uncertain, and thus it is hard for the principal to distinguish whether the agent exerts effort or not. Hence, it is expected that the promised utility plays little role in the incentive and thus payment should be made completely based on whether an arrival occurs or not. Part (i) of Proposition EC.2 validates this intuition.

Part (ii) of Proposition EC.2 states the result for another extreme case. As R approaches zero, there is essentially no arrival uncertainty. In the absence of information asymmetry (in term of the agent's effort rate), the system's first best can be achieved. In fact, the first-best societal utility is V^{FB}, which indeed is asymptotically achieved under the proposed contract.

EC.2. Proofs of the Results in Sections 2 and 3

EC.2.1. Proof of Proposition 1

The proof of part (i) is exactly the same as that of Proposition 1 in Cao et al. (2022), in which random termination instead of random switching may take place. The proof of part (ii) is similar to that of Lemma 6 in Sun and Tian (2018). To keep this paper self-contained, we provide a complete proof here.

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

$$\begin{aligned} u_t(\Gamma,\nu) &:= \mathbb{E}^{\nu,q}\left[\int_0^\infty e^{-\rho s}\big(\mathrm{d}L_s - b\mathbb{1}_{\nu_s=\mu}\mathrm{d}s\big) \,\Big|\, \mathcal{F}_t\right] \\ &= \int_0^t e^{-\rho s}\big(\mathrm{d}L_s - b\mathbb{1}_{\mu_s=\mu}\mathrm{d}s\big) + e^{-\rho t} W_t(\Gamma,\nu). \end{aligned} \tag{EC.25}$$

In what follows, we omit (Γ,ν) from all relevant quantities for the sake of easing notation. 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 $\mathcal{I}^N := \mathcal{I}^N_{[0,\infty)}$. Similarly, we use $\mathcal{I}^Q_{[t_1,t_2]}$ to denote the set of randomized switching time epochs during $[t_1, t_2]$ under the switching intensity process $\{q_t\}_{t\geqslant 0}$. Moreover, we denote $\mathcal{I}_t^Q := \mathcal{I}_{[0,t]}^Q$ and $\mathcal{I}^Q := \mathcal{I}_{[0,\infty)}^Q$.

At any time instant $\zeta-$, $W_{\zeta-}$ can jump to W_ζ^N triggered by an arrival at time ζ, or jump to W_ζ^Q triggered by a randomized switching, or jump to W_ζ^L triggered by an instantaneous payment. (Here, the agent's promised utility will not jump caused by a deterministic switching.) Therefore, we can decompose W_ζ (for $\zeta > t$) into its discrete part

$$\sum_{t\leqslant\xi\leqslant\zeta}\left[(W_\xi^N - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^N} + (W_\xi^Q - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^Q} + (W_\xi^L - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^L}\right]$$

and its absolutely continuous part

$$W_\zeta^c := W_\zeta - \sum_{t\leqslant\xi\leqslant\zeta}\left[(W_\xi^N - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^N} + (W_\xi^Q - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^Q} + (W_\xi^L - W_{\xi-})\mathbb{1}_{\xi\in\mathcal{I}_{[t,\zeta]}^L}\right],$$

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

According to the definition of admissible contract, we know that both W_t^N and W_t^Q is $\mathcal{F}_t$-predictable. However, W_t^L can also depend on $\mathrm{d}N_t$ and $\mathrm{d}Q_t$, that is, W_t^L is $\mathcal{F}_t$-adaptive.

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

$$\begin{aligned} e^{-\rho t'}W_{t'} - e^{-\rho t}W_t = &\int_t^{t'} e^{-\rho\zeta}\big(-\rho W_\zeta \mathrm{d}\zeta + \mathrm{d}W_\zeta^c\big) \\ &+ \sum_{\zeta\in(t,t']} e^{-\rho\zeta}\left[(W_\zeta^N - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^N} + (W_\zeta^Q - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^Q} + (W_\zeta^L - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^L}\right]. \end{aligned} \tag{EC.26}$$

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

$$\begin{aligned} 0 =& \mathbb{E}_t[u_{t'}] - u_t \\ =& \mathbb{E}_t[e^{-\rho t'}W_{t'} - e^{-\rho t}W_t] + \mathbb{E}_t\left[\int_{t+}^{t'} e^{-\rho\zeta}\big(\mathrm{d}L_\zeta - b\mathbb{1}_{\nu_\zeta=\mu}\mathrm{d}\zeta\big)\right] \\ =& \mathbb{E}_t\left[\int_t^{t'} e^{-\rho\zeta}\big(-\rho W_\zeta\mathrm{d}\zeta + \mathrm{d}W_\zeta^c\big)\right] \\ &+ \mathbb{E}_t\left\{\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\left[(W_\zeta^N - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^N} + (W_\zeta^Q - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^Q} + (W_\zeta^L - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}_{(t,t']}^L}\right]\right\} \\ &+ \mathbb{E}_t\left[\int_{t+}^{t'} e^{-\rho\zeta}\big(\mathrm{d}L_s - b\mathbb{1}_{\nu_\zeta=\mu}\mathrm{d}\zeta\big)\right] \\ =& \mathbb{E}_t\left\{\int_t^{t'} e^{-\rho\zeta}\Big\{\big[-\rho W_\zeta + (W_\zeta^N - W_{\zeta-})\nu_\zeta + (W_\zeta^Q - W_{\zeta-})q_\zeta\big]\mathrm{d}\zeta + \mathrm{d}W_\zeta^c\Big\}\right. \end{aligned}$$

$$+\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\Big[(W^L_\zeta - W_{\zeta-})\mathbb{1}_{\zeta\in\mathcal{I}^L_{(t,t']}}\Big]\Bigg\} + \mathbb{E}_t\left[\int_{t+}^{t'} e^{-\rho\zeta}(\mathrm{d}L_\zeta - b\mathbb{1}_{\nu_\zeta=\mu}\mathrm{d}\zeta)\right],$$

where the second equality follows from (EC.25), and the third from (EC.26). The fourth equality follows from the facts that $\{Q_t\}_{t\geqslant 0}$ is a counting process with intensity q_t, and that N_t is a counting process with intensity ν_t, as well as Lemma L3 in Chapter II of Brémaud (1981), noting that

$$\mathbb{E}_t\int_t^{t'} e^{-\rho\zeta}|(W^N_\zeta - W_{\zeta-})\nu_\zeta|\mathrm{d}\zeta \leqslant \bar{W}\mu\int_t^{t'} e^{-\rho\zeta}\mathrm{d}\zeta < \infty, \text{ and} \tag{EC.27}$$

$$\mathbb{E}_t\int_t^{t'} e^{-\rho\zeta}|(W^Q_\zeta - W_{\zeta-})q_\zeta|\mathrm{d}\zeta \leqslant \bar{W}\mathbb{E}_t\int_t^{\infty} e^{-\rho\zeta}q_\zeta\mathrm{d}\zeta \leqslant \bar{W}\mathbb{E}_t\int_t^{\tau} e^{-r\zeta}q_\zeta\mathrm{d}\zeta < \infty, \tag{EC.28}$$

in view of (WU), $\rho > r$, and (1).

Recall that $\mathrm{d}L_t = \ell_t\mathrm{d}t + \Delta L_t$. For any $t < t' < \tau$, the above equality can be stated as

$$\mathbb{E}_t\left\{\int_t^{t'} e^{-\rho\zeta}\big[-\rho W_\zeta + (W^N_\zeta - W_{\zeta-})\nu_\zeta + (W^Q_\zeta - W_\zeta -)q_\zeta - b\mathbb{1}_{\nu_\zeta=\mu} + \ell_\zeta\big]\mathrm{d}\zeta + \mathrm{d}W^c_\zeta\right\}$$
$$+\mathbb{E}_t\sum_{\zeta\in(t,t']} e^{-\rho\zeta}\big[(W^L_\zeta - W_{\zeta-})\mathbb{1}_{\Delta L_\zeta>0} + \Delta L_\zeta\big] = 0. \tag{EC.29}$$

Consider any time t. Letting $t'\downarrow t$ in (EC.29) yields

$$\mathbb{E}_t[(W^L_t - W_{t-})\mathbb{1}_{\Delta L_t>0} + \Delta L_t] = 0, \tag{EC.30}$$

which further implies

$$\mathrm{d}W^c_t = \big[\rho W_{t-} - (W^N_t - W_{t-})\nu_t - (W^Q_t - W_{t-})q_t + b\mathbb{1}_{\nu_t=\mu} - \ell_t\big]\mathrm{d}t,\ t\geqslant 0. \tag{EC.31}$$

Let $H_t := W^N_t - W_{t-}$ and $H^q_t := -W^Q_t + W_{t-}$. Then, both H_t and H^q_t are $\mathcal{F}_t$-predictable. Besides, since W^L_t is $\mathcal{F}_t$-adaptive, (EC.30) in fact is equivalent to

$$(W^L_t - W_{t-})\mathbb{1}_{\Delta L_t>0} + \Delta L_t = 0. \tag{EC.32}$$

We also have

$$\mathrm{d}W_t = \mathrm{d}W^c_t + (W^N_t - W_{t-})\mathrm{d}N_t + (W^Q_t - W_{t-})\mathrm{d}Q_t + (W^L_t - W_{t-})\mathbb{1}_{\Delta L_t>0}. \tag{EC.33}$$

Combining (EC.31)–(EC.33), we obtain (PK).

Relationship (5) follows immediately by noting $W^N_t \geqslant 0$ and $W^Q_t \geqslant 0$ for all $t\geqslant 0$.

(ii) Let $\widetilde{u}_t(\Gamma,\nu',\nu)$ denote the agent's total expected discounted utility conditional on $\mathcal{F}_t$ under contract Γ, when he follows effort process $\nu' = \{\nu'_t\}_{t\geqslant 0}$ before time t and then effort process ν after time t:

$$\widetilde{u}_t(\Gamma,\nu',\nu) = \int_0^t e^{-\rho s}\big(\mathrm{d}L_s - b\mathbb{1}_{\nu'_s=\mu}\mathrm{d}s\big) + e^{-\rho t}W_t(\Gamma,\nu). \tag{EC.34}$$

Here, $\widetilde{u}_{0-}(\Gamma,\nu',\nu)$ can be interpreted in a similar vein as that for $W_{0-}(\Gamma,\nu)$. In fact, we have $\widetilde{u}_{0-}(\Gamma,\nu',\nu) = W_{0-}(\Gamma,\nu) = u(\Gamma,\nu)$. In what follows, we write $\bar{\nu}$ instead of $\bar{\nu}(\Gamma)$ to ease notation. By the above definition, we have

$$\widetilde{u}_t(\Gamma,\nu,\bar{\nu}) = u_t(\Gamma,\bar{\nu}) + \int_0^t e^{-\rho s} b(\mathbb{1}_{\bar{\nu}_s=\mu} - \mathbb{1}_{\nu_s=\mu})\mathrm{d}s. \tag{EC.35}$$

Besides, by (PK) and (EC.25), we obtain that

$$\begin{aligned}\mathrm{d}u_t(\Gamma,\nu) &= e^{-\rho t}\big(\mathrm{d}L_t - b\mathbb{1}_{\mu_t=\mu}\mathrm{d}t\big) + e^{-\rho t}\big(\mathrm{d}W_t(\Gamma,\nu) - \rho W_t(\Gamma,\mu)\mathrm{d}t\big)\\ &= e^{-\rho t}\big[H_t(\Gamma,\nu)(\mathrm{d}N_t - \nu_t\mathrm{d}t) - H_t^q(\Gamma,\nu)(\mathrm{d}Q_t - q_t\mathrm{d}t)\big].\end{aligned} \tag{EC.36}$$

Therefore, for any time points $t < t'$, we have (below, we add superscript ν in some expectation operators, to indicate that the related random variables are induced by the effort process ν)

$$\begin{aligned}\mathbb{E}_t[\widetilde{u}_{t'}(\Gamma,\nu,\bar{\nu})] - \widetilde{u}_t(\Gamma,\nu,\bar{\nu}) &= \mathbb{E}_t[u_{t'}(\Gamma,\bar{\nu})] - u_t(\Gamma,\bar{\nu}) + \mathbb{E}_t^{\nu}\left[\int_t^{t'} e^{-\rho s} b(\mathbb{1}_{\bar{\nu}_s=\mu} - \mathbb{1}_{\nu_s=\mu})\mathrm{d}s\right]\\ &= \mathbb{E}_t^{\nu}\left[\int_{t+}^{t'} e^{-\rho s}\big(H_s(\Gamma,\bar{\nu})(\mathrm{d}N_s - \bar{\nu}_t\mathrm{d}s) - H_s^q(\Gamma,\bar{\nu})(\mathrm{d}Q_s - q_s\mathrm{d}t) + b(\mathbb{1}_{\bar{\nu}_s=\mu} - \mathbb{1}_{\nu_s=\mu})\mathrm{d}s\big)\right]\\ &= \mathbb{E}_t^{\nu}\left[\int_t^{t'} e^{-\rho s}\Big(H_s(\Gamma,\bar{\nu})(\nu_s - \bar{\nu}_s) + b(\mathbb{1}_{\bar{\nu}_s=\mu} - \mathbb{1}_{\nu_s=\mu})\Big)\mathrm{d}s\right],\end{aligned} \tag{EC.37}$$

where the first equality follows from (EC.35) and the second equality follows from (EC.36). The last equalities uses the fact that conditional on $\mathcal{F}_t$ and under effort process ν, $\{N_s\}_{s\in(t,t']}$ and $\{Q_s\}_{s\in(t,t']}$ are counting processes with intensities ν_s and q_s respectively, which follows by applying Lemma L3 in Chapter II of Brémaud (1981) with the aid of (EC.27) and (EC.28).

Since both ν and $\bar{\nu}$ are admissible, we have $\nu_t = \bar{\nu}_t = \underline{\mu}$ whenever $\mathcal{E}_t = \emptyset$. Hence, we have

$$H_s(\Gamma,\bar{\nu})(\nu_s - \bar{\nu}_s) + b(\mathbb{1}_{\bar{\nu}_s=\mu} - \mathbb{1}_{\nu_s=\mu}) = -(H_s(\Gamma,\bar{\nu}) - \beta)\Delta\mu\mathbb{1}_{\mathcal{E}_s=\mathsf{I},\nu_s=\underline{\mu}} \tag{EC.38}$$

for any $s \geqslant 0$. Therefore, if (IC) holds, then we have $\mathbb{E}_t[\widetilde{u}_{t'}(\Gamma,\nu,\bar{\nu})] \leqslant \widetilde{u}_t(\Gamma,\nu,\bar{\nu})$ by (EC.37) and (EC.38), which implies that $\{\widetilde{u}_t(\Gamma,\nu,\bar{\nu})\}_{t\geqslant 0}$ is an $\mathcal{F}$-supermatingale. By (IR) and (WU), we can add $\widetilde{u}_\infty(\Gamma,\nu,\bar{\nu}) := \int_0^\infty e^{-\rho s}\big(\mathrm{d}L_s - b\mathbb{1}_{\nu_s=\mu}\mathrm{d}s\big)$ as the last element of this supermartingale. Therefore,

$$u(\Gamma,\bar{\nu}) = \widetilde{u}_{0-}(\Gamma,\nu,\bar{\nu}) \geqslant \mathbb{E}[\widetilde{u}_\infty(\Gamma,\nu,\bar{\nu})] = u(\Gamma,\nu),$$

implying the incentive compatibility of $\bar{\nu}$ under contract Γ.

If (IC) fails to hold, then we consider an effort process ν such that $\nu_t = \mu$ if and only if $H_t(\Gamma,\bar{\nu}) \geqslant \beta$ and $\mathcal{E}_t = \mathsf{I}$. Clearly, ν is admissible, and the expression in (EC.38) becomes $-(H_s(\Gamma,\bar{\nu}) - \beta)\Delta\mu\mathbb{1}_{\mathcal{E}_s=\mathsf{I},H_s(\Gamma,\bar{\nu})<\beta}$, which is always non-negative and positive on a set of positive measure. Thus, by (EC.37), there exists a time $t > 0$ such that $\mathbb{E}_{0-}[\widetilde{u}_t(\Gamma,\nu,\bar{\nu})] > \widetilde{u}_{0-}(\Gamma,\nu,\bar{\nu}) = u(\Gamma,\bar{\nu})$. Define another effort process ν' which follows ν until time t and then switches to $\bar{\nu}$, which is also admissible. Moreover, we have $u(\Gamma,\nu') = \mathbb{E}_{0-}[\widetilde{u}_t(\Gamma,\nu,\bar{\nu})]$, which indicates $u(\Gamma,\nu') > u(\Gamma,\bar{\nu})$, contradicting (3). The proof is complete.

EC.2.2. Proof of Lemma 1

If we can show that (PK) holds under contract $\Gamma^*(w_0;\underline{\theta},\breve{w},\bar{\theta},\hat{w})$, then (17) follows immediately from (2) and (4) with $t=0$. In fact, (PK) holds by setting $H_t=\beta\mathbb{1}_{\mathcal{E}_{t-}=\mathsf{I}}$ and $H_t^q=(\breve{w}-\underline{\theta})\,\mathbb{1}_{W_{t-}=\breve{w},\mathcal{E}_{t-}=\mathsf{I}}$.

EC.2.3. Proof of Theorem 1

Fix any contract $\Gamma\in\mathfrak{C}$. The agent's promised utility follows a process W with its dynamics described by (PK) with $\nu_t=\mu$ for $\mathcal{E}_t=\mathsf{I}$ and $\nu_t=\underline{\mu}$ for $\mathcal{E}_t=\emptyset$.

Recall that $\mathrm{d}L_t=\ell_t\mathrm{d}t+\Delta L_t$. Write $\phi(w,\varepsilon)=V_\varepsilon(w)-w$ for any $w\in\mathbb{R}_+$ and $\varepsilon\in\{\mathsf{I},\emptyset\}$. Applying the change-of-variable formula (see, for example, Theorem 70 of Chapter IV in Protter 2003, pp. 214) for processes of locally bounded variation to the process $(W,\mathcal{E})$ and using (PK), we have

$$
\begin{aligned}
e^{-rT}\phi(W_T,\mathcal{E}_T)=\phi(W_{0-},\mathcal{E}_{0-})+\int_{0+}^{T}e^{-rt}\Big[&(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t+q_tH_t^q-\ell_t)\cdot D_{t-}\\
&-rV_{\mathcal{E}_{t-}}(W_{t-})\Big]\mathrm{d}t+\sum_{0\leqslant t\leqslant T}e^{-rt}\Delta\phi(W_t,\mathcal{E}_t)
\end{aligned}
$$

for any $T\geqslant 0$, where D_{t-} is the left derivative of $\phi(w,\mathcal{E}_{t-})$ with respect to w at W_{t-}, that is, $D_{t-}=V'_{\mathcal{E}_{t-}}(W_{t-})-1$, by recalling that we use $f'(w)$ to represent the left derivative of f at w for any absolutely continuous function defined on $\mathbb{R}_+$. Besides, we have

$$
\begin{aligned}
\Delta\phi(W_t,\mathcal{E}_t)\ =\ &\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})\\
&+\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-})-\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})\\
&+\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\ \text{for } t>0,
\end{aligned}
$$

and

$$
\Delta\phi(W_0,\mathcal{E}_0)=\phi(W_0,\mathcal{E}_0)-\phi(W_0,\mathcal{E}_{0-})+\phi(W_0,\mathcal{E}_{0-})-\phi(W_{0-},\mathcal{E}_{0-})
$$

by noting that $\mathrm{d}N_0=\mathrm{d}Q_0=0$ with probability 1.

Define $M^N=\{M_t^N\}_{t\geqslant 0}$ and $M^Q=\{M_t^Q\}_{t\geqslant 0}$ by

$$
M_t^N=N_t-\int_0^t\nu_s\mathrm{d}s\quad\text{and}\quad M_t^Q=Q_t-\int_0^t q_s\mathrm{d}s.
$$

Note that

$$
\begin{aligned}
&\sum_{0<t\leqslant T}\big[\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\\
&=\int_{0+}^{T}e^{-rt}\Big\{[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}N_t+[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}Q_t\Big\}\\
&=\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\mathrm{d}M_t^N+\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\nu_t\mathrm{d}t\\
&\quad+\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]\mathrm{d}M_t^Q+\int_{0+}^{T}e^{-rt}\big[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big]q_t\mathrm{d}t,
\end{aligned}
$$

where the first equality uses the fact that $\{t\in[0,T]\mid \mathrm{d}N_t=\mathrm{d}Q_t=1\}$ has a Lebesgue measure 0 with probability 1. Summarizing the above formulas, we obtain

$$\begin{aligned} e^{-rT}\phi(W_T,\mathcal{E}_T) &= \phi(W_{0-},\mathcal{E}_{0-})+\int_{0+}^{T}e^{-rt}[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}M_t^N+\\ &\quad+\int_{0+}^{T}e^{-rt}[\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}M_t^Q+A_1+A_2+A_3+A_4+A_5, \end{aligned} \tag{EC.39}$$

where

$$\begin{aligned} A_1 &:= \int_{0+}^{T}e^{-rt}\Big\{(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_{t-})\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})\\ &\qquad+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t\Big\}\mathrm{d}t,\\ A_2 &:= \sum_{0<t\leqslant T}e^{-rt}\Big[\phi\big(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-}\big)-\phi\big(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-}\big)\Big],\\ A_3 &:= \sum_{0\leqslant t\leqslant T}e^{-rt}[\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})],\\ A_4 &:= \int_{0+}^{T}e^{-rt}q_t\Big\{H_t^q\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)+\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\Big\}\mathrm{d}t,\\ A_5 &:= \phi(W_0,\mathcal{E}_{0-})-\phi(W_{0-},\mathcal{E}_{0-}). \end{aligned}$$

Below we treat each term separately.

Consider first A_1. If $\mathcal{E}_{t-}=\mathsf{I}$, then $\nu_{t-}=\mu$ and $\phi(W_{t-},\mathcal{E}_{t-})=V_{\mathsf{I}}(W_{t-})-W_{t-}$. Since the contract Γ is incentive compatible, we have $H_t\geqslant\beta$ by Proposition 1(ii). Consequently, we have

$$\begin{aligned} &(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t\\ &=(\rho W_{t-}+b-H_t\mu-\ell_t)\cdot\big(V_{\mathsf{I}}'(W_{t-})-1\big)-r\cdot(V_{\mathsf{I}}(W_{t-})-W_{t-})+[V_{\mathsf{I}}(W_{t-}+H_t)-V_{\mathsf{I}}(W_{t-})-H_t]\cdot\mu\\ &=\rho W_{t-}\cdot\big(V_{\mathsf{I}}'(W_{t-})-1\big)-r\cdot(V_{\mathsf{I}}(W_{t-})-W_t)-(\ell_t-b)\cdot\big(V_{\mathsf{I}}'(W_{t-})-1\big)\\ &\quad+\big[V_{\mathsf{I}}(W_{t-}+H_t)-V_{\mathsf{I}}(W_{t-})-V_{\mathsf{I}}'(W_{t-})H_t\big]\cdot\mu\\ &\leqslant\rho W_{t-}\cdot(V_{\mathsf{I}}'(W_{t-})-1)-r\cdot(V_{\mathsf{I}}(W_{t-})-W_{t-})+\ell_t-b+[V_{\mathsf{I}}(W_{t-}+\beta)-V_{\mathsf{I}}(W_{t-})-V_{\mathsf{I}}'(W_{t-})\beta]\cdot\mu\\ &=-\Big[(\mu+r)V_{\mathsf{I}}(W_{t-})-\mu V_{\mathsf{I}}(W_{t-}+\beta)+\rho(\bar{w}-W_{t-})V_{\mathsf{I}}'(W_{t-})-(\mu R-c)+(\rho-r)W_{t-}\Big]+\ell_t-[R\mu-(c-b)]\\ &=-(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(W_{t-})+\ell_t-[R\mu-(c-b)]\\ &\leqslant\ell_t-[R\mu-(c-b)]. \end{aligned}$$

Here, the first inequality follows from (i) $V_{\mathsf{I}}(W_{t-})\geqslant 0$ (this follows from the fact that V_{I} is nondecreasing) and (ii) $H_t\geqslant\beta$, and $\beta=\arg\max_{h\geqslant\beta}\{V_{\mathsf{I}}(w+h)-V_{\mathsf{I}}(w)-V_{\mathsf{I}}'(w)\cdot h\}$ due to the concavity of V_{I}, and the last inequality follows from (20).

If $\mathcal{E}_{t-}=\emptyset$, then $\nu_{t-}=\underline{\mu}$. It follows from (5) that $H_t\geqslant -W_{t-}$. Therefore, we have

$$\begin{aligned}
&(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t\\
=&(\rho W_{t-}-H_t\underline{\mu}-\ell_t)\cdot\big(V'_\emptyset(W_{t-})-1\big)-r\cdot(V_\emptyset(W_{t-})-W_{t-})+\big[V_\emptyset(W_{t-}+H_t)-V_\emptyset(W_{t-})-H_t\big]\cdot\underline{\mu}\\
=&\rho W_{t-}\cdot\big(V'_\emptyset(W_{t-})-1\big)-r\cdot(V_\emptyset(W_{t-})-W_{t-})-\ell_t\cdot\big(V'_\emptyset(W_{t-})-1\big)\\
&+\big[V_\emptyset(W_{t-}+H_t)-V_\emptyset(W_{t-})-V'_\emptyset(W_{t-})H_t\big]\cdot\underline{\mu}\\
\leqslant&\rho W_{t-}\cdot\big(V'_\emptyset(W_{t-})-1\big)-r\cdot(V_\emptyset(W_{t-})-W_{t-})+\ell_t\\
=&-\Big[rV_\emptyset(W_{t-})-\rho W_{t-}\cdot V'_\emptyset(W_{t-})+(\rho-r)W_{t-}-R\underline{\mu}\Big]+\ell_t-R\underline{\mu}\\
=&-(\mathcal{A}_\emptyset V_\emptyset)(W_{t-})+\ell_t-R\underline{\mu}\\
\leqslant&\ell_t-R\underline{\mu},
\end{aligned}$$

where the first inequality follows from (i) $V'_\emptyset(W_{t-})\geqslant 0$ (this follows from the fact that $V_\emptyset$ is nondecreasing) and (ii) $H_t\geqslant -W_{t-}$, and $0=\arg\max_{h\geqslant -w}\{V_\emptyset(w+h)-V_\emptyset(w)-V'_\emptyset(w)\cdot h\}$ due to the concavity of $V_\emptyset$, and the last inequality follows from (20).

Combining the above two cases yields

$$\begin{aligned}
&(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\cdot\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t\\
\leqslant&\ \ell_t-[R\nu_t-(c-b)\mathbb{1}_{\nu_t=\mu}]
\end{aligned}\tag{EC.40}$$

for any $t>0$.

Consider next A_2. We have

$$\begin{aligned}
&\phi(W_{t-}+H_t\mathrm{d}N_t-H^q_t\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-})-\phi(W_{t-}+H_t\mathrm{d}N_t-H^q_t\mathrm{d}Q_t,\mathcal{E}_{t-})\\
=&\ V_{\mathcal{E}_{t-}}(W_{t-}+H_t\mathrm{d}N_t-H^q_t\mathrm{d}Q_t-\Delta L_t)-V_{\mathcal{E}_{t-}}(W_{t-}-H^q_t\mathrm{d}Q_t+H_t\mathrm{d}N_t)+\Delta L_t\\
\leqslant&\ \Delta L_t,\ \forall t>0,
\end{aligned}\tag{EC.41}$$

where the inequality follows from the facts that $\Delta L_t\geqslant 0$ and that V_ε is nondecreasing for any $\varepsilon\in\{\mathsf{I},\emptyset\}$.

Consider now A_3. By considering four possible value combinations of $(\mathcal{E}_{t-},\mathcal{E}_t)$ and using (21), we have

$$\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})=V_{\mathcal{E}_t}(W_t)-V_{\mathcal{E}_{t-}}(W_t)\leqslant\kappa(\mathcal{E}_{t-},\mathcal{E}_t).\tag{EC.42}$$

Consider next A_4. We have

$$\begin{aligned}
&H^q_t\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)+\phi(W_{t-}-H^q_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\\
=&\ H^q_tV'_{\mathcal{E}_{t-}}(W_{t-})+V_{\mathcal{E}_{t-}}(W_{t-}-H^q_t)-V_{\mathcal{E}_{t-}}(W_{t-})\leqslant 0,
\end{aligned}$$

where the inequality follows from the concavity of V_ε for any $\varepsilon \in \{\mathsf{I}, \emptyset\}$. This, together with $q_t \geqslant 0$, yields

$$A_4 = \int_{0+}^{T} e^{-rt} q_t \Big\{ H_t^q \big(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1\big) + \phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-}) \Big\} \mathrm{d}t \leqslant 0. \tag{EC.43}$$

Consider finally A_5. It follows from (2) and (4) with $t = 0$ that $\mathbb{E}[W_0 + \Delta L_0] = W_{0-}$. Therefore, we have

$$\begin{aligned} &\mathbb{E}[\phi(W_0, \mathcal{E}_{0-})] - \phi(W_{0-}, \mathcal{E}_{0-}) = \mathbb{E}[V_{\mathcal{E}_{0-}}(W_0)] - V_{\mathcal{E}_{0-}}(W_{0-}) - \big(\mathbb{E}[W_{0-}] - W_{0-}\big) \\ &\leqslant V_{\mathcal{E}_{0-}}\big(\mathbb{E}[W_0]\big) - V_{\mathcal{E}_{0-}}(W_{0-}) + \mathbb{E}[\Delta L_0] \leqslant \mathbb{E}[\Delta L_0], \end{aligned} \tag{EC.44}$$

where the first inequality follows from the concavity of V_ε for any $\varepsilon \in \{\mathsf{I}, \emptyset\}$ and the Jensen's inequality, and the second inequality follows from the facts that V_ε is nondecreasing and that $W_{0-} = \mathbb{E}[W_0 + L_0] \geqslant \mathbb{E}[W_0]$.

Combining (EC.39)–(EC.43), we have

$$\begin{aligned} e^{-rT}\phi(W_T, \mathcal{E}_T) \leqslant & \phi(W_{0-}, \mathcal{E}_{0-}) + \int_{0+}^{T} e^{-rt}[\phi(W_{t-} + H_t, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})]\mathrm{d}M_t^N \\ &+ \int_{0+}^{T} e^{-rt}[\phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})]\mathrm{d}M_t^Q \\ &+ \int_{0+}^{T} e^{-rt}[\ell_t - (R\nu_t - (c-b)\mathbb{1}_{\nu_t = \mu})]\mathrm{d}t + \sum_{0<t\leqslant T} e^{-rt}\Delta L_t \\ &+ \sum_{0\leqslant t\leqslant T} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t) + \phi(W_0, \mathcal{E}_{0-}) - \phi(W_{0-}, \mathcal{E}_{0-}) \end{aligned}$$

for any $T > 0$, which can be displayed as

$$\begin{aligned} \phi(W_{0-}, \mathcal{E}_{0-}) \geqslant & e^{-rT}\phi(W_T, \mathcal{E}_t) - \int_{0}^{T} e^{-rt}\big[\phi(W_{t-} + H_t, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big]\mathrm{d}M_t^N \\ &- \int_{0+}^{T} e^{-rt}\big[\phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big]\mathrm{d}M_t^Q \\ &+ \int_{0+}^{T} e^{-rt}\big(R\mathrm{d}N_t - \mathrm{d}L_t - (c-b)\mathbb{1}_{\mathcal{E}_t = \mathsf{I}}\mathrm{d}t\big) - \sum_{0\leqslant t\leqslant T} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t) \\ &+ \phi(W_{0-}, \mathcal{E}_{0-}) - \phi(W_0, \mathcal{E}_{0-}). \end{aligned}$$

Taking expectation in the above inequality yields

$$\begin{aligned} \phi(W_{0-}, \mathcal{E}_{0-}) \geqslant & \mathbb{E}[e^{-rT}\phi(W_T, \mathcal{E}_T)] - \mathbb{E}\left[\int_{0+}^{T} e^{-rt}\big(\phi(W_{t-} + H_t, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big)\mathrm{d}M_t^N\right] \\ &- \mathbb{E}\left[\int_{0+}^{T} e^{-rt}\big(\phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-})\big)\mathrm{d}M_t^Q\right] \\ &+ \mathbb{E}\left[\int_{0+}^{T} e^{-rt}\big(R\mathrm{d}N_t - \mathrm{d}L_t - (c-b)\mathbb{1}_{\mathcal{E}_t = \mathsf{I}}\mathrm{d}t\big) - \sum_{0\leqslant t\leqslant T} e^{-rt}\kappa(\mathcal{E}_{t-}, \mathcal{E}_t)\right] \end{aligned}$$

$$
\begin{aligned}
&+\phi(W_{0-},\mathcal{E}_{0-})-\mathbb{E}\phi(W_0,\mathcal{E}_{0-})\\
\geqslant&\mathbb{E}[e^{-rT}\phi(W_T,\mathcal{E}_T)]-\mathbb{E}\left[\int_{0+}^{T}e^{-rt}\big(\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big)\mathrm{d}M_t^N\right]\\
&-\mathbb{E}\left[\int_{0+}^{T}e^{-rt}\big(\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\big)\mathrm{d}M_t^Q\right]\\
&+\mathbb{E}\left[\int_{0}^{T}e^{-rt}\big(R\mathrm{d}N_t-\mathrm{d}L_t-(c-b)\mathbb{1}_{\mathcal{E}_t=\mathsf{I}}\mathrm{d}t\big)-\sum_{0\leqslant t\leqslant T}e^{-rt}\kappa(\mathcal{E}_{t-},\mathcal{E}_t)\right]
\end{aligned}
\tag{EC.45}
$$

for any $T>0$, where the last inequality follows from (EC.44).

We claim that it suffices to consider the case that

$$
\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]<\infty. \tag{EC.46}
$$

Otherwise, we have $\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]=\infty$. It follows from (PK) and (WU) that $\mathrm{d}L_t\geqslant(H_t-\bar{W})^+\mathrm{d}N_t$ for $t>0$. Hence, we have

$$
\begin{aligned}
&\mathbb{E}\left[\int_{0}^{\infty}e^{-rt}\mathrm{d}L_t\right]\geqslant\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}(H_t-\bar{W})^+\mathrm{d}N_t\right]=\mathbb{E}\left[\int_{0+}^{\tau}e^{-rt}(H_t-\bar{W})^+\nu_t\mathrm{d}t\right]\\
&\geqslant\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}(|H_t|-\bar{W})\nu_t\mathrm{d}t\right]\geqslant\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]-\frac{\bar{W}\mu}{r}=\infty,
\end{aligned}
$$

where the first equality follows from Equation (2.3) in Chapter II of Brémaud (1981), the second inequality follows from $H_t\geqslant-W_{t-}\geqslant-R\mu/r$ in view of (5) and (WU), and the third inequality follows from $\nu_t\leqslant\mu$. Then, we have

$$
U(\Gamma)\leqslant\mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_{0}^{\infty}e^{-rt}\big(R\mathrm{d}N_t-\mathrm{d}L_t\big)\right]\leqslant\frac{R\mu}{r}-\mathbb{E}^{\bar{\nu}(\Gamma)}\left[\int_{0}^{\infty}e^{-rt}\mathrm{d}L_t\right]=-\infty,
$$

and thus the desired result follows immediately.

Given (EC.46), we have

$$
\begin{aligned}
&\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})|\nu_t\right]\mathrm{d}t\\
&\leqslant\max_{w>0,\varepsilon\in\{\mathsf{I},\emptyset\}}\{|V_\varepsilon'(w)-1|\}\cdot\mathbb{E}\left[\int_{0+}^{\infty}e^{-rt}|H_t|\nu_t\mathrm{d}t\right]<\infty,
\end{aligned}
$$

where $\max_{w>0,\varepsilon\in\{\mathsf{I},\emptyset\}}\{|V_\varepsilon'(w)-1|\}<\infty$ follows from the concavity of V_ε and the fact that $V_\varepsilon'\geqslant0$. It follows from Lemma L3, Chapter II in Brémaud (1981) that $\widetilde{M}=\{\widetilde{M}_t\}_{t\geqslant0}$, defined by

$$
\widetilde{M}_t:=\int_{0+}^{t}e^{-rs}[\phi(W_{s-}+H_s,\mathcal{E}_{s-})-\phi(W_{s-},\mathcal{E}_{s-})]\mathrm{d}M_s^N,
$$

is an $\mathcal{F}$-martingale. Hence, $\mathbb{E}[\widetilde{M}_T]=\mathbb{E}[\widetilde{M}_0]=0$, that is,

$$
\mathbb{E}\left[\int_{0+}^{T}e^{-rt}[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\mathrm{d}M_t^N\right]=0.
$$

Similarly, using (1), we can show that

$$\mathbb{E}\left[\int_{0+}^{T} e^{-rt}\Big(\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\Big)\mathrm{d}M_t^Q\right]=0.$$

It follows from (22) and the fact that both V_{I} and $V_{\emptyset}$ are nondecreasing that $\phi(w,\varepsilon)\geqslant \underline{v}-w$ for any $\varepsilon\in\{\mathsf{I},\emptyset\}$. Letting $T\to\infty$ in (EC.45) and using (WU), we have $\phi(W_{0-},\emptyset)\geqslant U(\Gamma)$ with $W_{0-}=u(\Gamma,\bar{\nu}(\Gamma))$. Hence, the desired result is obtained.

A byproduct of the proof of Theorem 1 is the following result. In the remaining of this e-companion, whenever we need to prove that certain contract achieves the upper bound, we will use this result together with Lemma 1.

Proposition EC.3. *Suppose that the conditions stated in Theorem 1 hold. Furthermore, suppose that there exists a contract $\Gamma^\diamond\in\mathfrak{C}$ such that the corresponding agent's promised utility W_t satisfies*

$$(\rho W_{t-}+b\mathbb{1}_{\nu_t=\mu}-H_t\nu_t-\ell_t)\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)-r\phi(W_{t-},\mathcal{E}_{t-})+[\phi(W_{t-}+H_t,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})]\nu_t$$
$$=\ell_t-[R\nu_t-(c-b)\mathbb{1}_{\nu_t=\mu}], \tag{EC.47}$$
$$\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t-\Delta L_t,\mathcal{E}_{t-})-\phi(W_{t-}+H_t\mathrm{d}N_t-H_t^q\mathrm{d}Q_t,\mathcal{E}_{t-})=\Delta L_t, \tag{EC.48}$$
$$\phi(W_t,\mathcal{E}_t)-\phi(W_t,\mathcal{E}_{t-})=\kappa(\mathcal{E}_{t-},\mathcal{E}_t), \tag{EC.49}$$
$$q_t\Big\{H_t^q\big(V'_{\mathcal{E}_{t-}}(W_{t-})-1\big)+\phi(W_{t-}-H_t^q,\mathcal{E}_{t-})-\phi(W_{t-},\mathcal{E}_{t-})\Big\}=0, \tag{EC.50}$$

for any $t>0$ and

$$\mathbb{E}[\phi(W_0,\mathcal{E}_{0-})]-\phi(W_{0-},\mathcal{E}_{0-})=\mathbb{E}[\Delta L_0]. \tag{EC.51}$$

Then, for any value $w\in[0,\infty)$ such that $u(\Gamma^\diamond,\bar{\nu}(\Gamma^\diamond))=w$, we have

$$U(\Gamma^\diamond)=V_\emptyset(w)-w.$$

Proof. Equalities (EC.47)–(EC.51) demonstrate that all the inequalities in the proof of Theorem 1, (EC.40)–(EC.44), hold with equalities under contract $\Gamma^\diamond$. The desired result can be shown by going through the proof of Theorem 1, with all inequalities replaced by equalities. ◻

EC.2.4. Proof of Theorem 2

In view of Theorems 3–5, it remains to show that $\bar{K}(R)$ is increasing in R and $\underline{K}(R)$ is decreasing in R on $(c/\Delta\mu,\bar{R})$.

First, under Condition 2, we have $\bar{K}=\bar{K}_2=\bar{V}-\underline{v}-\bar{w}=\dfrac{\Delta\mu\cdot R-c-\mu\beta}{r}$, which is clearly increasing in R.

Second, under Condition 1, we have $\bar{K} = \bar{K}_1$. By the definition of $\bar{K}_1$, we only need to show that m^K is decreasing in R, since m^K is decreasing in K by Lemma 4. Observe that m^K has the following characterization:

$$m^K = \max_{x>0}\left\{\frac{\mathcal{V}_{\hat{\mathrm{w}}}(x) - \underline{v} - K}{x}\right\}.$$

Hence, it suffices to show that $\mathcal{V}_{\hat{\mathrm{w}}}(x) - \underline{v}$ is decreasing in R for any $x > 0$. Note that for any $x > 0$, we have $\mathcal{V}_{\hat{\mathrm{w}}}(x) - \underline{v} = \int_0^x \mathcal{V}'_{\hat{\mathrm{w}}}(y)\mathrm{d}y$. Hence, the desired result is obtained if we can show that $\mathcal{V}'_{\hat{\mathrm{w}}}(y)$ is decreasing in R for any $y \geqslant 0$, which is exactly Lemma EC.1(iii) below.

LEMMA EC.1. *Let R vary and other model parameters μ, $\underline{\mu}$, c, b, and K be fixed.*

(i) *Both $V'_{\tilde{w}}(w)$ and $\mathcal{V}'_{\tilde{w}}(w)$ do not depend on R for any $w \geqslant 0$.*

(ii) *$\hat{\mathrm{w}}$ is decreasing in R.*

(iii) *$\mathcal{V}'_{\hat{\mathrm{w}}}(w)$ is decreasing in R for any $w \geqslant 0$.*

Proof. (i) Note that $V'_{\tilde{w}}$ satisfies (EC.54) on $[0, \tilde{w}]$, with the boundary condition that $V'_{\tilde{w}}(w) = 0$ for all $w \geqslant \tilde{w}$. As (EC.54) does not involve parameter R, its unique solution $V'_{\tilde{w}}$ is also independent of R, which also implies the independence of $V''(\tilde{w})$ on R. Therefore, $\breve{w}(\tilde{w})$ is also independent of R, by Lemma 2(i), which in turn concludes the independence of $\mathcal{V}'_{\tilde{w}}$ on R by Lemma 2(ii).

(ii) This part can be shown using a similar line as that in the proof of Proposition 1 in Cao et al. (2022). Specifically, we define $\psi(R, \tilde{w}) := \mathcal{V}_{\tilde{w}}(0; R) - \underline{v}$, where we write $\mathcal{V}_{\tilde{w}}(\cdot; R)$ instead of $\mathcal{V}_{\tilde{w}}(\cdot)$ to highlight the dependence on R. (Note that we adopt a different notation from those in Section EC.1.4 as the parameter settings are different.) Then, we have

$$\psi(R, \tilde{w}) = \mathcal{V}_{\tilde{w}}(\tilde{w}; R) - \int_0^{\tilde{w}} \mathcal{V}'_{\tilde{w}}(y)\mathrm{d}y - \underline{v} = \frac{\Delta\mu(R-\beta) - (\rho - r)\tilde{w}}{r} - \int_0^{\tilde{w}} \mathcal{V}'_{\tilde{w}}(y)\mathrm{d}y,$$

which is linear and increasing in R for any $\tilde{w} \in [0, \bar{w}]$, where have we used part (i) in this lemma. Hence, the desired result is obtained by noting that $\psi(R, \hat{\mathrm{w}}) = 0$.

(iii) This result follows immediately from parts (i) and (ii), combining with the monotonicity of $\mathcal{V}'_{\tilde{w}}(y)$ in $\tilde{w}$ (see Lemma 2(iii)). ◻

It remains to show that $\underline{K}$ is non-increasing in R. This can be obtained by taking the first-order derivative of $\underline{K}$ with respect to R in (33), investigating its sign and noting from Assumption 1 that $R > c/\Delta\mu$.

EC.3. Proofs of the Results in Section 4

EC.3.1. Proof of Lemma 2

This result follows almost the same logic as that for the proof of Lemmas 2 and 3 in Cao et al. (2022) and uses Lemma EC.2 below. However, there are minor differences as both β and $\bar{w}$ in Cao et al. (2022) take different values from ours. Hence, to make this paper self-contained, we provide a complete proof here.

We first present Lemma EC.2 below because it will be frequently used in the subsequent analysis.

LEMMA EC.2. *For any $\tilde{w} \in [0, \bar{w})$, there exists a unique function $V_{\tilde{w}}$ in $C^1([0,\tilde{w}])$ that solves the differential equation (25) on $[0,\tilde{w}]$ with boundary condition (26). We further extend the domain of $V_{\tilde{w}}$ to $\mathbb{R}_+$ by letting $V_{\tilde{w}}(w) = V_{\tilde{w}}(\tilde{w})$ for all $w > \tilde{w}$. Then, function $V_{\tilde{w}}(w)$ has the following properties.*

(i) $V_{\tilde{w}}(\cdot) \in C^1(\mathbb{R}_+) \cap C^2(\mathbb{R}_+ \backslash \{\tilde{w}\}) \cap C^3(\mathbb{R}_+ \backslash \{\tilde{w}, \tilde{w} - \beta\})$.

(ii) *For any given $w \geqslant 0$, define function $v(\tilde{w}) := V_{\tilde{w}}(w)$. We have $v(\cdot) \in C^1([0,\bar{w}))$.*

(iii) *Function $V_{\tilde{w}}(w)$ is increasing in w on $[0,\tilde{w}]$.*

(iv) *For any $\tilde{w}_1$ and $\tilde{w}_2$ such that $0 < \tilde{w}_1 < \tilde{w}_2 < \bar{w}$, we have $V_{\tilde{w}_1}(w) > V_{\tilde{w}_2}(w)$ and $V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w)$ for $w \in [0, \tilde{w}_1)$.*

(v) *If $\rho \leqslant r + \mu$, then for any $w \in [0,\bar{w})$, $V_{\tilde{w}}(w)$ approaches negative infinity as $\tilde{w}$ approaches $\bar{w}$ from below.*

(vi) *If $\rho > r + \mu$, then for any $w \in [0,\bar{w}]$, we have*

$$\lim_{\tilde{w} \uparrow \bar{w}} V_{\tilde{w}}(w) = \bar{V} - \frac{\rho - r}{\rho - r - \mu}(\bar{w} - w),$$

where $\bar{V}$ is defined in (10). Furthermore, $\bar{V} - \dfrac{\rho - r}{\rho - r - \mu}\bar{w} \geqslant \underline{v}$ is equivalent to $R \geqslant \hat{R}$.

Proof. Step 1 in the proof of Proposition 4 in Sun and Tian (2018) has already shown the existence and uniqueness of a function satisfying (25) with boundary condition (26). Here, we adopt their idea with argument slightly modified. First, we observe that (25) reduces to an ordinal differential equation (ODE) on the interval $[(\tilde{w} - \beta)^+, \tilde{w}]$, as $V_{\tilde{w}}(w + \beta) = \bar{V}(\tilde{w})$ for all $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$. Therefore, this problem can be "backwardly" treated as an initial value problem, which satisfies the conditions stated in Cauchy–Lipschitz theorem and thus admits a unique continuously differential solution on $[(\tilde{w} - \beta)^+, \tilde{w}]$. In fact, we have

$$V_{\tilde{w}}(w) = \begin{cases} \bar{V}(\tilde{w}) + \dfrac{\rho - r}{r + \mu - \rho}(\tilde{w} - w) + b_{\tilde{w}}\left((\bar{w} - w)^{\frac{r+\mu}{\rho}} - (\bar{w} - \tilde{w})^{\frac{r+\mu}{\rho}}\right), & \rho \neq r + \mu \\ \bar{V}(\tilde{w}) + \dfrac{\rho - r}{\rho}(\tilde{w} - w) - \dfrac{(\rho - r)(\bar{w} - w)}{\rho} \ln\left(\dfrac{\bar{w} - w}{\bar{w} - \tilde{w}}\right), & \rho = r + \mu \end{cases} \tag{EC.52}$$

for $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$, where

$$b_{\tilde{w}} := \frac{r - \rho}{r + \mu - \rho} \cdot \frac{\rho}{r + \mu}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}}. \tag{EC.53}$$

In general, for any $k \in \mathbb{N}$, given that the values of $V_{\tilde{w}}$ on $w \in [(\tilde{w} - k\beta)^+, \tilde{w}]$ all determined, (25) is an ODE on the interval $[(\tilde{w} - (k+1)\beta)^+, (\tilde{w} - k\beta)^+]$, whose unique solution can be shown by verifying the conditions in Cauchy–Lipschitz Theorem. By induction on k, we can extend the solution to (25) to the entire interval $[0,\widetilde{w}]$, as desired.

Next, we show that such a function $V_{\widetilde{w}}$ possesses properties (i)–(vi).

(i) It follows from (25) and the boundary condition at $\widetilde{w}$ that $V_{\widetilde{w}}(\cdot)\in C^1(\mathbb{R}_+)$. Taking derivative in (25) with respect to w and noting that $V_{\widetilde{w}}(w)=\bar{V}(\widetilde{w})$ for all $w\geqslant\widetilde{w}$, we have

$$(\mu+r)V_{\widetilde{w}}'(w)-\mu V_{\widetilde{w}}'(w+\beta)+\rho(\bar{w}-w)V_{\widetilde{w}}''(w)-\rho V_{\widetilde{w}}'(w)+\rho-r=0 \tag{EC.54}$$

for $w\in[0,\widetilde{w})$, which implies that $V_{\widetilde{w}}(\cdot)\in C^2([0,\widetilde{w}))$. Moreover, $V_{\widetilde{w}}''(\widetilde{w}-)=-(\rho-r)/\big(\rho(\bar{w}-\widetilde{w})\big)<0$. Also, by the definition of $V_{\widetilde{w}}$ on $(\widetilde{w},\infty)$, we have $V_{\widetilde{w}}''(w)=0$ for $w>\widetilde{w}$ and thus $V_{\widetilde{w}}''(\widetilde{w}+)=0$. Therefore, $V_{\widetilde{w}}(\cdot)\in C^2(\mathbb{R}_+\backslash\{\widetilde{w}\})$.

Similarly, taking derivative in (EC.54) with respect to w yields

$$\rho(\bar{w}-w)V_{\widetilde{w}}'''(w)=\mu\big(V_{\widetilde{w}}''(w+\beta)-V_{\widetilde{w}}''(w)\big)+(2\rho-r)V_{\widetilde{w}}''(w), \tag{EC.55}$$

for $w\in[0,\widetilde{w})$. Note that $V_{\widetilde{w}}''$ does not exist only at $\widetilde{w}$, which demonstrates that $V_{\widetilde{w}}'''$ does not exist at $\widetilde{w}$ and $\widetilde{w}-\beta$ (if it is nonnegative). That is, $V_{\widetilde{w}}(\cdot)\in C^3(\mathbb{R}_+\backslash\{\widetilde{w},\widetilde{w}-\beta\})$.

(ii) Fix any $w\geqslant 0$. If $\widetilde{w}\leqslant w$, then $v(\widetilde{w})=V_{\widetilde{w}}(w)=\bar{V}(\widetilde{w})$, which implies that $v(\cdot)\in C^1([0,w\wedge\bar{w}))$. Hence, the desired property is obtained if $w\geqslant\bar{w}$.

Now suppose that $w<\bar{w}$ and $\widetilde{w}\in(w,\bar{w})$. By the above discussion, we have $v(\cdot)\in C^1([0,w])$. For any $w'\in[w,\widetilde{w}]$, it follows from (25) that

$$\rho V_{\widetilde{w}}'(w')=-\frac{(\mu+r)V_{\widetilde{w}}(w')}{\bar{w}-w'}+\frac{\mu V_{\widetilde{w}}\big((w'+\beta)\wedge\widetilde{w}\big)}{\bar{w}-w'}+\frac{(\mu R-c)-(\rho-r)w'}{\bar{w}-w'}.$$

Integrating the above equation with respect to w' from w to $\widetilde{w}$ yields

$$\begin{aligned}\rho\big(V_{\widetilde{w}}(\widetilde{w})-V_{\widetilde{w}}(w)\big)=&-(\mu+r)\int_w^{\widetilde{w}}\frac{V_{\widetilde{w}}(w')}{\bar{w}-w'}\mathrm{d}w'+\mu\int_w^{\widetilde{w}}\frac{V_{\widetilde{w}}\big((w'+\beta)\wedge\widetilde{w}\big)}{\bar{w}-w'}\mathrm{d}w'\\&+\int_w^{\widetilde{w}}\frac{(\mu R-c)-(\rho-r)w'}{\bar{w}-w'}\mathrm{d}w'.\end{aligned}$$

First, using the above equality, we can obtain that $v(\widetilde{w})=V_{\widetilde{w}}(w)$ is continuous in $\widetilde{w}$ on $[w,\bar{w})$. Then, again using this equality, we conclude that $V_{\widetilde{w}}(w)$ is continuously differentiable in $\widetilde{w}$ on $[w,\bar{w})$, which, combining with $v(\cdot)\in C^1([0,w])$, yields that $v(\cdot)\in C^1\big([0,\bar{w})\big)$.

(iii) We show this result by a contradictory argument. Suppose that $w^p:=\sup\{w\in\mathbb{R}_+\mid V_{\widetilde{w}}'(w)<0\}$ exists. Recall from the proof for part (ii) of this lemma that $V_{\widetilde{w}}''(\widetilde{w}-)<0$. Hence, we have $w^p\in[0,\widetilde{w})$, $V_{\widetilde{w}}'(w^p)=0$, and $V_{\widetilde{w}}'>0$ on $(w^p,\widetilde{w})$. Evaluating (25) at w^p gives

$$\begin{aligned}rV_{\widetilde{w}}(w^p)&=\mu R-c-(\rho-r)w^p+\mu\Big[V_{\widetilde{w}}\big((w^p+\beta)\wedge\widetilde{w}\big)-V_{\widetilde{w}}(w^p)\Big]\\&>\mu R-c-(\rho-r)w^p>rV_{\widetilde{w}}(\widetilde{w}),\end{aligned}$$

where the first inequality uses $V'_{\tilde{w}} > 0$ on $(w^p, \tilde{w})$. This reaches a contradiction with $V_{\tilde{w}}(w^p) < V_{\tilde{w}}(\tilde{w})$.

(iv) We first show the second claim. Suppose it fails to hold. Since $V'_{\tilde{w}_1}(\tilde{w}_1) = 0$ and $V'_{\tilde{w}_2}(\tilde{w}_1) > 0$, the quantity $w^\dagger := \sup\{w \in [0, \tilde{w}_1) \mid V'_{\tilde{w}_1}(w) \geqslant V'_{\tilde{w}_2}(w)\}$ is well defined, and satisfies $V'_{\tilde{w}_1}(w^\dagger) = V'_{\tilde{w}_2}(w^\dagger)$ by part (ii) of this lemma. Evaluating (25) at $w^\dagger$ for both $\tilde{w}_1$ and $\tilde{w}_2$, we obtain

$$\mu(V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta)) = (r + \mu)(V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger)).$$

Hence, we have

$$\begin{aligned} V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta) &= V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger) + \int_0^\beta \big(V'_{\tilde{w}_1}(w^\dagger + y) - V'_{\tilde{w}_2}(w^\dagger + y)\big)\mathrm{d}y \\ &< V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger) = \frac{\mu}{r+\mu}\big(V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta)\big), \end{aligned}$$

which indicates that both $V_{\tilde{w}_1}(w^\dagger + \beta) - V_{\tilde{w}_2}(w^\dagger + \beta)$ and $V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger)$ are negative. By the definition of $w^\dagger$, we have $V'_{\tilde{w}_1} < V'_{\tilde{w}_2}$ on $(w^\dagger, \tilde{w}_1]$, which implies that

$$\begin{aligned} V_{\tilde{w}_1}(w^\dagger) - V_{\tilde{w}_2}(w^\dagger) &= V_{\tilde{w}_1}(\tilde{w}_1) - V_{\tilde{w}_2}(\tilde{w}_1) - \int_{w^\dagger}^{\tilde{w}_1} \big(V'_{\tilde{w}_1}(y) - V'_{\tilde{w}_2}(y)\big)\mathrm{d}y \\ &> V_{\tilde{w}_1}(\tilde{w}_1) - V_{\tilde{w}_2}(\tilde{w}_1) > \bar{V}(\tilde{w}_1) - \bar{V}(\tilde{w}_2) > 0. \end{aligned} \tag{EC.56}$$

This contradiction indicates the correctness of the second claim. The first claim follows by replacing $w^\dagger$ by any $w \in [0, \tilde{w}_1)$ in (EC.56).

(v) For any $w \in [(\bar{w} - \beta)^+, \bar{w})$, we have $w \in [(\tilde{w} - \beta)^+, \tilde{w}]$ when $\tilde{w}$ is close to $\bar{w}$ from below and thus (EC.52) is valid. Letting $\tilde{w} \uparrow \bar{w}$ in (EC.52), we obtain that $\lim_{\tilde{w}\uparrow\bar{w}} V_{\tilde{w}}(w) = -\infty$.

If $w \in [0, [\bar{w} - \beta)^+, \bar{w}))$, then using the fact that $V_{\tilde{w}}$ is nondecreasing on $\mathbb{R}_+$, we obtain $V_{\tilde{w}}(w) \leqslant V_{\tilde{w}}\big((\bar{w} - \beta)^+\big)$, which yields that $\limsup_{\tilde{w}\uparrow\bar{w}} V_{\tilde{w}}(w) \leqslant \lim_{\tilde{w}\uparrow\bar{w}} V_{\tilde{w}}\big((\bar{w} - \beta)^+\big) = -\infty$, concluding the desired result.

(vi) Note that $\rho > r + \mu$ implies $\bar{w} < \beta$. Hence, (EC.52) is valid for all $w \in [0, \tilde{w}]$. Therefore, the first claim follows by letting $\tilde{w} \uparrow \bar{w}$ in (EC.52) and using $\lim_{\tilde{w}\uparrow\bar{w}} b_{\tilde{w}} = 0$. The second claim is trivial by the definition of $\hat{R}$. ◻

We proceed to prove Lemma 2 as follows.

Proof of Lemma 2. (i) From (EC.52), we have

$$V''_{\tilde{w}}(w) = -\frac{\rho - r}{\rho}(\bar{w} - \tilde{w})^{\frac{\rho - r - \mu}{\rho}}(\bar{w} - w)^{\frac{-2\rho + r + \mu}{\rho}} < 0$$

for any $w \in [(\tilde{w} - \beta)^+, \tilde{w})$. (The above expression also holds if $\mu + r = \rho$.) Hence, the desired result holds with $\breve{w}(\tilde{w}) = 0$ if $\tilde{w} \leqslant \beta$.

Now consider the case that $\tilde{w} > \beta$. In this case, we have $\bar{w} > \beta$, which gives $\rho < \mu$. Define $w^c := \inf\{w \in [0, \tilde{w}) \mid V''_{\tilde{w}}(w) \geqslant 0\}$. If the set is empty, we set $w^c = 0$. By Lemma EC.2(i), we have $V''_{\tilde{w}} < 0$ on $(w^c, \tilde{w})$. Hence, the desired result holds with $\breve{w}(\tilde{w}) = 0$ if $w^c = 0$.

Next, we suppose that $w^c > 0$. Since $V_{\tilde{w}}$ is strictly concave on $[\tilde{w} - \beta, \tilde{w})$, we have $w^c < \tilde{w} - \beta$. According to Lemma EC.2(i), we have $V''_{\tilde{w}}(w^c) = 0$ and $V''_{\tilde{w}} < 0$ on $(w^c, \tilde{w})$.

It follows from (EC.54) 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

$$V'_{\tilde{w}}(w^c + \beta) = \frac{(\mu - \rho + r)V'_{\tilde{w}}(w^c) + (\rho - r)}{\mu}. \tag{EC.57}$$

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

$$V'_{\tilde{w}}(w^c) > 1, \tag{EC.58}$$

in view of (EC.57) and $\rho < \mu$. Evaluating (25) at w^c gives

$$\begin{aligned} rV_{\tilde{w}}(w^c) &= \mu R - c - (\rho - r)w^c - \rho(\bar{w} - w^c)V'_{\tilde{w}}(w^c) + \mu\big(V_{\tilde{w}}(w^c + \beta) - V_{\tilde{w}}(w^c)\big) \\ &> \mu R - c - (\rho - r)w^c - \rho(\bar{w} - w^c)V'_{\tilde{w}}(w^c) + \mu\beta V'_{\tilde{w}}(w^c + \beta) \\ &= \mu R - c - (\rho - r)(w^c - \beta) + \big[\rho(w^c - \beta) + r\beta\big]V'_{\tilde{w}}(w^c), \end{aligned} \tag{EC.59}$$

where the inequality follows from the strict concavity of $V_{\tilde{w}}$ on $(w^c, w^c + \beta)$, and the last equality uses (EC.57) and $\rho\bar{w} = \mu\beta$.

Below we distinguish two cases.

Case 1: $\rho(w^c - \beta) + r\beta \geqslant 0$. It follows from (EC.58) and (EC.59) that

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

which contradicts Lemma EC.2(iii).

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

$$0 < w^c < \frac{(\rho - r)\beta}{\rho}. \tag{EC.60}$$

Below, we will show that

$$V''_{\tilde{w}} > 0 \text{ on } [0, w^c). \tag{EC.61}$$

Evaluating (EC.55) at w^c and using $V''_{\tilde{w}}(w^c) = 0$, we obtain

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

which implies that $V''_{\tilde{w}} > 0$ on $(w^c - \epsilon, w^c)$ for some $\epsilon > 0$. If (EC.61) fails to hold, then $w^d := \sup\{w \in [0, w^c) \mid V''_{\tilde{w}}(w) \leqslant 0\}$ is well defined, satisfying $w^d \in [0, w^c)$. Moreover, we have $V''_{\tilde{w}}(w^d) = 0$ and $V'''_{\tilde{w}}(w^d) \geqslant 0$. (Note that $w^c < \tilde{w} - \beta$, indicating that $V'''_{\tilde{w}}(w^d)$ exists by Lemma EC.2(i).) Hence, evaluating (EC.55) at w^d gives $\rho(\bar{w} - w^d)V'''_{\tilde{w}}(w^d) = \mu V''_{\tilde{w}}(w^d + \beta) \geqslant 0$. By the definition of w^c, we have $w^d + \beta \leqslant w^c$. Consequently, it follows from (EC.60) that $w^d \leqslant w^c - \beta < 0$, which is impossible. Therefore, (EC.61) holds. Letting $\breve{w}(\tilde{w}) = w^c$, we obtain the proof of the first claim in part (i). The second claim in part (ii) follows immediately by (EC.60).

(ii) This claim holds trivially by the first claim in part (i) and Lemma EC.2(iii).

(iii) To ease notation, we write $\breve{w}(\tilde{w}_1)$ and $\breve{w}(\tilde{w}_2)$ as $\breve{w}_1$ and $\breve{w}_2$ respectively. First, we show the second claim, that is, $\mathcal{V}'_{\tilde{w}_1}(w) < \mathcal{V}'_{\tilde{w}_2}(w)$ for any $w \in [0, \tilde{w}_1]$, by considering the following two cases.

Case 1: $\breve{w}_1 \leqslant \breve{w}_2$. In this case, we have

$$\mathcal{V}'_{\tilde{w}_1}(w) = V'_{\tilde{w}_1}(w) < V'_{\tilde{w}_2}(w) = \mathcal{V}'_{\tilde{w}_2}(w) \tag{EC.62}$$

for any $w \in [\breve{w}_2, \tilde{w}_1]$, where the two equalities follows from the definition of function $\mathcal{V}_{\tilde{w}}$ and $\breve{w}_1 \leqslant \breve{w}_2$, and the inequality follows from Lemma EC.2(iv). If $w \in [0, \breve{w}_2]$, then we can derive the desired inequality as follows:

$$\mathcal{V}'_{\tilde{w}_1}(w) \leqslant \mathcal{V}'_{\tilde{w}_1}(\breve{w}_1) = V'_{\tilde{w}_1}(\breve{w}_1) < V'_{\tilde{w}_2}(\breve{w}_1) < V'_{\tilde{w}_2}(\breve{w}_2) = \mathcal{V}'_{\tilde{w}_2}(\breve{w}_1) = \mathcal{V}'_{\tilde{w}_2}(w).$$

Here, the first inequality uses the definition and the concavity of $\mathcal{V}_{\tilde{w}_1}$, the second inequality uses Lemma EC.2(iv), and the last inequality uses Lemma 2(i).

Case 2: $\breve{w}_1 > \breve{w}_2$. In this case, we have

$$\mathcal{V}'_{\tilde{w}_1}(w) = V'_{\tilde{w}_1}(w \vee \breve{w}_1) < V'_{\tilde{w}_2}(w \vee \breve{w}_1) \leqslant V'_{\tilde{w}_2}(w \vee \breve{w}_2) = \mathcal{V}'_{\tilde{w}_2}(w) \tag{EC.63}$$

for all $w \in [0, \tilde{w}_1]$, where the first inequality uses Lemma EC.2(iv) and the second inequality uses the concavity of $V_{\tilde{w}_2}$ and the fact that $w \vee \breve{w}_1 \geqslant w \vee \breve{w}_2$.

The first claim can be readily obtained using the second claim. In fact, for any $w \in [0, \tilde{w}_1]$, we have

$$\begin{aligned}\mathcal{V}_{\tilde{w}_1}(w) - \mathcal{V}_{\tilde{w}_2}(w) &= \mathcal{V}_{\tilde{w}_1}(\tilde{w}_1) - \mathcal{V}_{\tilde{w}_2}(\tilde{w}_1) - \int_w^{\tilde{w}_1} \big(\mathcal{V}'_{\tilde{w}_1}(y) - \mathcal{V}'_{\tilde{w}_2}(y)\big)\mathrm{d}y \\ &> \mathcal{V}_{\tilde{w}_1}(\tilde{w}_1) - \mathcal{V}_{\tilde{w}_2}(\tilde{w}_1) > \bar{V}(\tilde{w}_1) - \bar{V}(\tilde{w}_2) > 0,\end{aligned}$$

where the second inequality uses part (ii) of this lemma. □

EC.3.2. Proof of Lemma 3

It follows from Lemma 2(iii) that $\mathcal{V}_{\tilde{w}}(0)$ is decreasing in $\tilde{w}$ on $(0,\bar{w})$. By Lemma EC.2(v) and (vi), we have that $\lim_{\tilde{w}\uparrow\bar{w}}\mathcal{V}_{\tilde{w}}(0)$ is either $-\infty$ or $\bar{V}-\frac{\rho-r}{\rho-r-\mu}\bar{w}$, which is less than $\underline{v}$ according to Condition 1. Moreover, $\lim_{\tilde{w}\downarrow 0}\mathcal{V}_{\tilde{w}}(0)=\bar{V}(0)>\underline{v}$ by Assumption 1. Therefore, the first claim is obtained by Lemma EC.2(ii).

For the second claim, we observe that by the proof of Lemma 2(i), if $\breve{w}(\tilde{w})>0$, it must be equal to w^c, in which case (EC.58) holds. This immediately concludes the result by the definition of $\mathcal{V}_{\tilde{w}}$.

EC.3.3. Proof of Proposition 2

Clearly, following the definition of $\underline{\Gamma}$ as in (16), $U(\underline{\Gamma})=\underline{v}$ trivially holds. Hence, it remains to show that functions $V_{\mathsf{I}}(w)=\mathcal{V}_{\hat{\mathrm{w}}}(w)$ and $V_{\emptyset}(w)=\underline{v}$ satisfy the optimality condition (20)–(22).

First, we show that $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)\geqslant 0$ for any $w\in\mathbb{R}_+$. If $w\in[\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}})$, by the definition of $\mathcal{V}_{\hat{\mathrm{w}}}$, we have $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=0$. If $w\in[\hat{\mathrm{w}},\infty)$, then we have

$$\begin{aligned}(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)&=(\mu+r)\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}})-\mu\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}})-(\mu R-c)+(\rho-r)w\\&=(\rho-r)(w-\hat{\mathrm{w}})\geqslant 0.\end{aligned}$$

If $w\in[0,\breve{w}(\hat{\mathrm{w}}))$ (if we discuss this case, it is implicitly assumed that $\breve{w}(\hat{\mathrm{w}})>0$), then we have $\mathcal{V}_{\hat{\mathrm{w}}}(w)=\underline{v}+\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))w$. Consequently,

$$(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=(\mu+r)\big(\underline{v}+\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))w\big)-\mu\mathcal{V}_{\hat{\mathrm{w}}}(w+\beta)+\rho(\bar{w}-w)\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-(\mu R-c)+(\rho-r)w.$$

Let the last expression be $g_{\mathsf{I}}(w)$. Obviously, $g_{\mathsf{I}}(\breve{w}(\hat{\mathrm{w}}))=0$. Moreover, for $w\in[0,\breve{w}(\hat{\mathrm{w}}))$, we have

$$\begin{aligned}g'_{\mathsf{I}}(w)&=(\mu+r)\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-\mu\mathcal{V}'_{\hat{\mathrm{w}}}(w+\beta)-\rho\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))+\rho-r\\&=(\rho-r)\big(1-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))\big)+\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-\mathcal{V}'_{\hat{\mathrm{w}}}(w+\beta)\big)\\&\leqslant(\rho-r)\big(1-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))\big)+\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}})+\beta)\big)=0,\end{aligned}$$

where the inequality follows from the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$, and the last equality uses the facts that $\rho(\bar{w}-\breve{w}(\hat{\mathrm{w}}))\mathcal{V}''_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))=(\rho-r)\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))-1\big)+\mu\big(\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}})+\beta)-\mathcal{V}'_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))\big)$ by (EC.54) and that $\mathcal{V}''_{\hat{\mathrm{w}}}(\breve{w}(\hat{\mathrm{w}}))=0$. Consequently, $g_{\mathsf{I}}(w)\geqslant 0$ for all $w\in[0,\breve{w}(\hat{\mathrm{w}}))$.

Therefore, we have $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)\geqslant 0$ for any $w\in\mathbb{R}_+$. Obviously, we have $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)=(\rho-r)w\geqslant 0$. Hence, (20) holds.

It follows from the facts that $V_{\mathsf{I}}(w)\geqslant V_{\mathsf{I}}(0)=\underline{v}=V_{\emptyset}(w)$ and that $V_{\emptyset}(w)=\underline{v}\geqslant\bar{V}(\hat{\mathrm{w}})-K=\mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}})-K\geqslant V_{\mathsf{I}}(w)-K$ (due to $K\geqslant\bar{V}(\hat{\mathrm{w}})-\underline{v}$) that both (21) and (22) hold.

EC.3.4. Proof of Lemma 4

Define

$$g(w,K) := \mathcal{V}_{\hat{\mathrm{w}}}(w) - \mathcal{V}'_{\hat{\mathrm{w}}}(w)w - \underline{v} - K. \tag{EC.64}$$

Then, we have

$$g(\breve{w}(\hat{\mathrm{w}}),K) = -K < 0, \tag{EC.65}$$

where the equality uses the linearity of $\mathcal{V}_{\hat{\mathrm{w}}}$ on $[0,\breve{w}(\hat{\mathrm{w}})]$. In addition,

$$g(\hat{\mathrm{w}},K) = \mathcal{V}_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) - \underline{v} - K > 0,$$

where the equality follows from $\mathcal{V}'_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) = 0$ and the inequality follows from the condition that $K < \bar{V}(\hat{\mathrm{w}}) - \underline{v}$. Furthermore, we have

$$\frac{\partial g(w,K)}{\partial w} = -\mathcal{V}''_{\hat{\mathrm{w}}}(w)w > 0 \text{ for } w \in (\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}}),$$

where the inequality follows from the fact that $\mathcal{V}_{\hat{\mathrm{w}}}$ is strictly concave on $(\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}})$. Since $g(w,K)$ is continuous in w (recalling that $\mathcal{V}_{\hat{\mathrm{w}}}$ is continuously differentiable), for any $K > 0$, there exists a unique $\bar{\theta}^K \in (\breve{w}(\hat{\mathrm{w}}),\hat{\mathrm{w}})$ such that $g(\bar{\theta}^K,K) = 0$. Hence, (29) holds if we define $m^K := \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K)$. Furthermore, by the implicit function theorem, we have

$$\frac{\mathrm{d}\bar{\theta}^K}{\mathrm{d}K} = -\frac{\frac{\partial g(w,K)}{\partial K}}{\frac{\partial g(w,K)}{\partial w}} = \frac{1}{\frac{\partial g(w,K)}{\partial w}} > 0,$$

which implies that $\bar{\theta}^K$ is increasing in K. Since $\mathcal{V}'_{\hat{\mathrm{w}}}(w)$ is decreasing in w, we have $m^K = \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K)$ is decreasing in K. Finally, the limiting result $\lim_{K\downarrow 0}\bar{\theta}^K = \breve{w}(\hat{\mathrm{w}})$ is implied by (EC.65).

EC.3.5. Proof of Proposition 3

Obviously, (22) holds since $V_{\mathsf{l}}(0) = V_{\emptyset}(0) = \underline{v}$. Note that it has been shown in the proof of Proposition 2 that $(\mathcal{A}_{\mathsf{l}}V_{\mathsf{l}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$. Hence, it remains to establish the second part of (20), as well as (21), by considering the following three cases.

Case 1: $w \in [0,\bar{\theta}^K)$. In this case, we have $(\mathcal{A}_{\emptyset}V_{\emptyset})(w) = (\rho - r)(1 - m^K)w \geqslant 0$, where we have used the fact that $m^K \leqslant 1$, which follows from Lemma 4, the definition of $\bar{K}_1$ as in (K1), and the condition that $K \geqslant \bar{K}_1$. Besides, we have

$$\begin{aligned} V_{\emptyset}(w) &= \left(1 - \frac{w}{\bar{\theta}^K}\right)\underline{v} + \frac{w}{\bar{\theta}^K}\cdot(\mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) - K) \\ &\leqslant \left(1 - \frac{w}{\bar{\theta}^K}\right)\mathcal{V}_{\hat{\mathrm{w}}}(0) + \frac{w}{\bar{\theta}^K}\cdot\mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) \leqslant \mathcal{V}_{\hat{\mathrm{w}}}(w) = V_{\mathsf{l}}(w), \end{aligned}$$

where the second inequality follows from the concavity of $\mathcal{V}_{\hat{w}}$. Finally, we have

$$\begin{aligned} V_{\emptyset}(w) - V_{\mathsf{I}}(w) + K &= m^K w + \underline{v} + K - \mathcal{V}_{\hat{w}}(w) \\ &= \mathcal{V}_{\hat{w}}(\bar{\theta}^K) - m^K \cdot (\bar{\theta}^K - w) - \mathcal{V}_{\hat{w}}(w) = \int_w^{\bar{\theta}^K} \big(\mathcal{V}'_{\hat{w}}(y) - m^K\big)\mathrm{d}y > 0, \end{aligned}$$

where the second equality follows from the first equality in (29), and the inequality follows from the concavity of $\mathcal{V}_{\hat{w}}$ and $m^K = \mathcal{V}'_{\hat{w}}(\bar{\theta}^K)$.

Case 2: $w \in [\bar{\theta}^K, \hat{w})$. First, we show that $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) \geqslant 0$ in this case. Note that $V_{\emptyset}(w) = \mathcal{V}_{\hat{w}}(w) - K$ for $w \in [\bar{\theta}^K, \hat{w})$. Hence, $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) \geqslant 0$ is equivalent to

$$f_1(w) := r(\mathcal{V}_{\hat{w}}(w) - K) - \rho w \mathcal{V}'_{\hat{w}}(w) + (\rho - r)w - R\underline{\mu} \geqslant 0.$$

For $w \in [\breve{w}(\hat{w}), \hat{w}]$, it holds that $(\mathcal{A}_{\mathsf{I}} \mathcal{V}_{\hat{w}})(w) = 0$. That is,

$$f_2(w) := (\mu + r)\mathcal{V}_{\hat{w}}(w) - \mu \mathcal{V}_{\hat{w}}(w + \beta) + \rho(\bar{w} - w)\mathcal{V}'_{\hat{w}}(w) - (\mu R - c) + (\rho - r)w = 0. \tag{EC.66}$$

Recall that $\bar{\theta}^K \in (\breve{w}(\hat{w}), \hat{w})$. Hence, it suffices to show that

$$f_3(w) := f_2(w) - f_1(w) = \mu\big(\mathcal{V}_{\hat{w}}(w) - \mathcal{V}_{\hat{w}}(w + \beta)\big) + \rho \bar{w} \mathcal{V}'_{\hat{w}}(w) + rK - (R\Delta\mu - c) < 0$$

for $w \in [\bar{\theta}^K, \hat{w})$.

It follows from (29) that $f_1(\bar{\theta}^K) = (\rho - r)\bar{\theta}^K(1 - m^K) > 0$. Hence, $f_3(\bar{\theta}^K) < 0$. Hence, it is enough to show that $f_3'(w) \leqslant 0$, or equivalently, $\mu(\mathcal{V}'_{\hat{w}}(w) - \mathcal{V}'_{\hat{w}}(w + \beta)) + \rho \bar{w} \mathcal{V}''_{\hat{w}}(w) \leqslant 0$, for $w \in [\bar{\theta}^K, \hat{w})$.

Taking derivative with respect to w in (EC.66) yields

$$(\mu + r)\mathcal{V}'_{\hat{w}}(w) - \mu \mathcal{V}'_{\hat{w}}(w + \beta) + \rho(\bar{w} - w)\mathcal{V}''_{\hat{w}}(w) - \rho \mathcal{V}'_{\hat{w}}(w) + \rho - r = 0$$

for $w \in [\breve{w}(\hat{w}), \hat{w}]$. Hence, for $w \in [\bar{\theta}^K, \hat{w})$, we have

$$\mu\big(\mathcal{V}'_{\hat{w}}(w) - \mathcal{V}'_{\hat{w}}(w + \beta)\big) + \rho \bar{w} \mathcal{V}''_{\hat{w}}(w) = (\rho - r)(\mathcal{V}'_{\hat{w}}(w) - 1) + \rho w \mathcal{V}''_{\hat{w}}(w) \leqslant 0,$$

where the inequality follows from the fact that $\mathcal{V}'_{\hat{w}}(\bar{\theta}^K) = m^K \leqslant 1$ and the concavity of $\mathcal{V}_{\hat{w}}$. Hence, we have $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) \geqslant 0$ for $w \in [\bar{\theta}^K, \hat{w})$. Note that $V_{\emptyset}(w) - V_{\mathsf{I}}(w) + K = 0$. Therefore, (21) trivially holds.

Case 3: $w \in [\hat{w}, \infty)$. It is straightforward to see that $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) = r(\mathcal{V}_{\hat{w}}(\hat{w}) - K) + (\rho - r)w - R\underline{\mu} > r\underline{v} - R\underline{\mu} = 0$ and $V_{\emptyset}(w) - V_{\mathsf{I}}(w) + K = 0$.

EC.3.6. Proof of Proposition 4

First, we show that under Condition 2 and $K > \bar{V} - \underline{v}$, functions $V_{\mathsf{I}}(w)$ as defined in (31) and $V_{\emptyset}(w) = \underline{v}$ satisfy the optimality condition (20)–(22). Note that the first inequality in Condition 2 implies $\bar{w} < \beta$. If $w \in [0, \bar{w}]$, then we have

$$\begin{aligned}(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) &= (\mu + r) V_{\mathsf{I}}(w) - \mu V_{\mathsf{I}}(w + \beta) + \rho(\bar{w} - w) V_{\mathsf{I}}'(w) - (\mu R - c) + (\rho - r) w \\ &= (\mu + r)\Big(\underline{v} + \frac{\bar{V} - \underline{v}}{\bar{w}} w\Big) - \mu \bar{V} + \rho(\bar{w} - w)\frac{\bar{V} - \underline{v}}{\bar{w}} - (\mu R - c) + (\rho - r) w \\ &= (\bar{w} - w)\left[\frac{\bar{V} - \underline{v}}{\bar{w}}(\rho - r - \mu) - (\rho - r)\right] \geqslant 0.\end{aligned}$$

Here, it is worth pointing out that although at $\bar{w}$, V_{I} is not differentiable, its left derivative exists and is $(\bar{V} - \underline{v})/\bar{w}$.

If $w \in (\bar{w}, \infty)$, then

$$(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) = (\mu + r)\bar{V} - \mu \bar{V} - (\mu R - c) + (\rho - r) w = (\rho - r)(w - \bar{w}) > 0.$$

Combining the above two cases yields $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$. Besides, for any $w \in \mathbb{R}_+$, $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) = (\rho - r) w \geqslant 0$. Hence, (20) holds.

It is straightforward to see that $V_{\mathsf{I}}(w) - V_{\emptyset}(w) \geqslant 0$, and $V_{\emptyset}(w) = \underline{v} \geqslant \bar{V} - K \geqslant V_{\mathsf{I}}(w) - K$. Hence, (21) holds. Obviously, $V_{\mathsf{I}}(0) = V_{\emptyset}(0) = \underline{v}$, implying (22).

Second, we show that under Condition 2 and $\bar{K}_2 < K \leqslant \bar{V} - \underline{v}$, functions $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ as defined in (31) and (32), respectively, satisfy (20)–(22). According to the proof for the case under Condition 2 and $K > \bar{V} - \underline{v}$, we have $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$. Below, we show that $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) \geqslant 0$ for all $w \in \mathbb{R}_+$.

If $w \in [0, \bar{w}]$, then

$$(\mathcal{A}_{\emptyset} V_{\emptyset})(w) = r V_{\emptyset}(w) - \rho w V_{\emptyset}'(w) + (\rho - r) w - R\underline{\mu} = (\rho - r)\Big(1 - \frac{\bar{V} - \underline{v} - K}{\bar{w}}\Big) w \geqslant 0.$$

Here, we mention that although at $\bar{w}$, $V_{\emptyset}$ is not differentiable, its left derivative exists and is $(\bar{V} - \underline{v} - K)/\bar{w}$.

If $w \in (\bar{w}, \infty)$, then we have

$$\begin{aligned}(\mathcal{A}_{\emptyset} V_{\emptyset})(w) &= r V_{\emptyset}(w) - \rho w V_{\emptyset}'(w) + (\rho - r) w - R\underline{\mu} \\ &= r(\bar{V} - K) + (\rho - r) w - R\underline{\mu} > r(\bar{V} - K - \underline{v}) > 0.\end{aligned}$$

Therefore, (20) holds. Note that $V_{\mathsf{I}}(w) - V_{\emptyset}(w) = K$ for $w \in [\bar{w}, \infty)$ and $V_{\mathsf{I}}(w) - V_{\emptyset}(w) = K/\bar{w} \cdot w$ for $w \in [0, \bar{w})$. Hence, (21) holds. Besides, $V_{\mathsf{I}}(0) = V_{\emptyset}(0) = \underline{v}$ and thus (22) holds.

EC.3.7. Proof of Lemma 5

The results stated in Lemma 5 hold in fact for any $K \in (0, \bar{K}_2)$. Below, we will show this slightly generalized version.

For any $\underline{\theta} \in [0, \bar{w}]$, it is straightforward to verify that functions

$$C^1(\underline{\theta}) := \frac{\bar{V} - \underline{v} - \bar{w}}{r/\rho \cdot \underline{\theta}^{r/\rho - 1}[(\rho/r - 1)\underline{\theta} + \bar{w}]} \text{ and } m(\underline{\theta}) := \frac{(\rho/r - 1)\underline{\theta} + \bar{V} - \underline{v}}{(\rho/r - 1)\underline{\theta} + \bar{w}} \tag{EC.67}$$

satisfy (36) and (37), with $\underline{\theta}$ replacing $\underline{\theta}_K$, $C^1(\underline{\theta})$ replacing c_K, and $m(\underline{\theta})$ replacing m_K. Moreover, it follows from Condition 2 and $K \in (0, \bar{K}_2)$ that $C^1(\underline{\theta}) > 0$. Note that the denominator of $C^1(\underline{\theta})$ is decreasing in $\underline{\theta}$, as its derivative with respect to $\underline{\theta}$ is always negative when $\underline{\theta} \in (0, \bar{w})$. Therefore, $C^1(\underline{\theta})$ is increasing in $\underline{\theta}$ on $[0, \bar{w}]$. That $m(\underline{\theta})$ is strictly decreasing in $\underline{\theta}$ on $[0, \bar{w}]$ is straightforward.

We have the following result, which is stated as a lemma for the ease of reference. Its proof is elementary and thus omitted.

LEMMA EC.3. *Under Condition 2 and $K \in (0, \bar{K}_2)$, function $\psi_1(\underline{\theta})$, defined by*

$$\psi_1(\underline{\theta}) = \bar{V} - \underline{v} - \bar{w} - C^1(\underline{\theta}) \cdot (\bar{w})^{r/\rho},$$

is continuous and decreasing in $\underline{\theta}$ on $[0, \bar{w}]$. Moreover, $\psi_1(\bar{w}) = 0$ and $\psi_1(0) = \bar{V} - \underline{v} - \bar{w} > K$. Consequently, there exists a unique number $\underline{\theta}_K \in (0, \bar{w})$ such that $\psi_1(\underline{\theta}_K) = K$. Furthermore, $\underline{\theta}_K$ is decreasing in K with $\lim_{K \downarrow 0} \underline{\theta}_K = \bar{w}$.

Lemma EC.3 immediately implies that the triple $(\underline{\theta}_K, c_K, m_K)$ with $c_K = C^1(\underline{\theta}_K)$, $m_K = m(\underline{\theta}_K)$ satisfies (36)–(38), which also states the monotonicity of $\underline{\theta}_K$ in K. The monotonicity of c_K and m_K in K follows from that of $C^1(\underline{\theta})$ and $m(\underline{\theta})$ in $\underline{\theta}$.

Finally, we show that under Condition 2, $K \geqslant \underline{K}$ if and only if (35) holds. First, according to the monotonicity of m_K in K, (35) is equivalent to

$$K \geqslant \breve{K}_2, \text{ in which } \breve{K}_2 := \inf\left\{K \in (0, \underline{K}_2] \;\middle|\; m_K \geqslant \frac{\rho - r}{\rho - r - \mu}\right\}. \tag{EC.68}$$

Next, it follows from (EC.67) and $\lim_{K \downarrow 0} \underline{\theta}_K = \bar{w}$ that

$$\lim_{K \downarrow 0} c_K = \frac{\bar{V} - \underline{v} - \bar{w}}{\bar{w}^{-r/\rho}} \text{ and } \lim_{K \downarrow 0} m_K = 1 + \frac{r(\bar{V} - \underline{v} - \bar{w})}{\rho \bar{w}} = \frac{R\Delta\mu - c}{\mu\beta}.$$

It is straightforward to verify that $\lim_{K \downarrow 0} m_K \geqslant (\rho - r)/(\rho - r - \mu)$ if and only if $R \geqslant \bar{R}$, where $\bar{R}$ is defined in (34). Hence, by the definition of $\breve{K}_2$ and the monotonicity of m_K in K, it is clear that $\breve{K}_2 = 0$ if and only if $R \geqslant \bar{R}$.

If $R < \bar{R}$, we have $m_{\breve{K}_2} = (\rho - r)/(\rho - r - \mu)$. Evaluating (EC.67) at $\underline{\theta} = \underline{\theta}_{\breve{K}_2}$ gives

$$\underline{\theta}_{\breve{K}_2} = \frac{\bar{V} - \underline{v} - (\rho - r)/(\rho - r - \mu)\bar{w}}{\mu/(\rho - r - \mu) \cdot (\rho/r - 1)} \text{ and } c_{\breve{K}_2} = \frac{\mu}{\rho - r - \mu} \frac{\rho}{r} \underline{\theta}_{\breve{K}_2}{}^{1 - r/\rho}.$$

Substituting these values into (38) with $K = \breve{K}_2$, we obtain the following closed-form expression of $\breve{K}_2$:

$$\breve{K}_2 = \bar{V} - \underline{v} - \bar{w} - \frac{\mu}{\rho - r - \mu}\frac{\rho}{r}\left[\frac{\bar{V} - \underline{v} - (\rho - r)/(\rho - r - \mu)\cdot\bar{w}}{\mu/(\rho - r - \mu)\cdot(\rho/r - 1)}\right]^{1 - r/\rho}\bar{w}^{r/\rho}. \tag{EC.69}$$

The proof is complete by verifying that $\underline{K} = \breve{K}_2$.

EC.3.8. Proof of Proposition 5

By the definition of contract $\bar{\Gamma}$, it is clear that $U(\bar{\Gamma}) = \bar{V} - \bar{w} - K = V_\emptyset(\bar{w}) - \bar{w}$. Hence, it remains to show that under Condition 2 and $K \in [\underline{K}, \bar{K}_2)$, functions $V_\mathsf{I}(w)$ and $V_\emptyset(w)$ as defined in (40) and (39), respectively, satisfy the optimality condition (20)–(22).

Obviously, (22) holds as $V_\mathsf{I}(0) = V_\emptyset(0) = \underline{v}$. We proceed to verify that $V_\mathsf{I}(w)$ and $V_\emptyset(w)$ satisfy (20) and (21).

We show that $(\mathcal{A}_\mathsf{I} V_\mathsf{I})(w) \geqslant 0$ for all $w \in \mathbb{R}_+$ by considering the following cases.

Case 1: $w \in [\underline{\theta}_K, \bar{w}]$. We have

$$\begin{aligned}(\mathcal{A}_\mathsf{I} V_\mathsf{I})(w) &= (\mu + r)(\bar{V} + m_K\cdot(w - \bar{w})) - \mu\bar{V} + \rho(\bar{w} - w)\cdot m_K - (\mu R - c) + (\rho - r)w\\ &= (\bar{w} - w)[m_K\cdot(\rho - r - \mu) - (\rho - r)] \geqslant 0.\end{aligned}$$

Here, we mention that although at $\bar{w}$, V_I is not differentiable, its left derivative exists and is m_K.

Case 2: $w \in [0, \underline{\theta}_K)$. In this case, we have

$$\begin{aligned}(\mathcal{A}_\mathsf{I} V_\mathsf{I})(w) &= (\mu + r)V_\mathsf{I}(w) - \mu V_\mathsf{I}(w + \beta) + \rho(\bar{w} - w)V_\mathsf{I}'(w) - (\mu R - c) + (\rho - r)w\\ &= \mu V_\mathsf{I}(w) - \mu V_\mathsf{I}(w + \beta) + \rho\bar{w}V_\mathsf{I}'(w) - \Delta\mu\cdot R + c\\ &= \mu\big(\underline{v} + w + c_K w^{r/\rho}\big) - \mu\bar{V} + \mu\beta\big(1 + c_K w^{r/\rho - 1} r/\rho\big) - \Delta\mu\cdot R + c =: g_\mathsf{I}(w),\end{aligned}$$

where the second equality follows from $\mathcal{A}_\emptyset V_\mathsf{I} = 0$ on $[0, \underline{\theta}_K)$, and the third equality follows from $\beta > \mu\beta/\rho = \bar{w}$ due to Condition 2.

Since V_I is continuously differentiable on $[0, \bar{w})$, $(\mathcal{A}_\mathsf{I} V_\mathsf{I})(w)$ is also continuous in w on $[0, \bar{w})$, which implies that $g_\mathsf{I}(\underline{\theta}_K) \geqslant 0$. Hence, it suffices to show that $g_\mathsf{I}(w)$ is decreasing in w. Using $g_\mathsf{I}'(w) = \mu + \mu r/\rho\cdot c_K w^{r/\rho - 2}\big(w + (r - \rho)/\rho\cdot\beta\big)$, we have

$$\begin{aligned}g_\mathsf{I}'(\underline{\theta}_K) &= \mu + \mu r/\rho\cdot c_K\cdot(\underline{\theta}_K)^{r/\rho - 2}\big(\underline{\theta}_K + (r - \rho)/\rho\cdot\beta\big)\\ &= \mu + \mu\cdot\frac{m_K - 1}{\underline{\theta}_K}\cdot\big(\underline{\theta}_K + (r - \rho)/\rho\cdot\beta\big)\\ &< \mu + \mu\left(\frac{\rho - r}{\rho - r - \mu} - 1\right)\left(1 + \frac{(r - \rho)\cdot\beta}{\rho\underline{\theta}_K}\right)\\ &< \mu + \mu\frac{\mu}{\rho - r - \mu}\left(1 + \frac{(r - \rho)\cdot\beta}{\rho\bar{w}}\right) = 0,\end{aligned}$$

where the second equality follows from (37), the first inequality follows from (35) and the fact that $\underline{\theta}_K + (r-\rho)/\rho \cdot \beta < \bar{w} + (r-\rho)/\rho \cdot \beta = (\mu + r - \rho)/\rho \cdot \beta < 0$, and the last equality follows from $\bar{w} = \mu\beta/\rho$. Besides, we have

$$\begin{aligned} g_{\mathsf{I}}''(w) &= r/\rho(r/\rho - 1) \cdot c_K w^{r/\rho - 3}[\mu w + \rho\bar{w}(r/\rho - 2)] \\ &\geqslant r/\rho(r/\rho - 1) \cdot c_K w^{r/\rho - 3}[\mu\bar{w} + \rho\bar{w}(r/\rho - 2)] \\ &= r/\rho(r/\rho - 1) \cdot c_K w^{r/\rho - 3}(\mu + r - 2\rho)\bar{w} > 0, \end{aligned}$$

where the last inequality follows from $\rho > r + \mu$. Therefore, $g_{\mathsf{I}}'(w) < 0$ for $w \in [0, \underline{\theta}_K]$.

Case 3: $w \in (\bar{w}, \infty)$. We have

$$(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) = (\mu + r)\bar{V} - \mu\bar{V} - (\mu R - c) + (\rho - r)w = (\rho - r)(w - \bar{w}) > 0.$$

Combining the above three cases yields $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0$ for any $w \in \mathbb{R}_+$.

Next, we establish $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for all $w \in \mathbb{R}_+$. Obviously, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) = 0$ for $w \in [0, \bar{w}]$. (Again, although $V_\emptyset$ is not differentiable at $\bar{w}$, its left derivative exists.) If $w \in (\bar{w}, \infty)$, then

$$(\mathcal{A}_\emptyset V_\emptyset)(w) = r(\bar{V} - K) + (\rho - r)w - R\underline{\mu} > r(\bar{V} - K) - R\underline{\mu} = r(\bar{V} - K - \underline{v}) > 0,$$

proving (20).

Below we establish (21). If $w \in [0, \underline{\theta}_K]$, we have $V_{\mathsf{I}}(w) - V_\emptyset(w) = 0$, and if $w \in [\bar{w}, \infty)$, we have $V_{\mathsf{I}}(w) - V_\emptyset(w) = K$. If $w \in (\underline{\theta}_K, \bar{w})$, we have

$$V_{\mathsf{I}}'(w) - V_\emptyset'(w) = m_K - V_\emptyset'(w) \geqslant m_K - V_\emptyset'(\underline{\theta}_K) = 0,$$

which implies that $V_{\mathsf{I}} - V_\emptyset$ is increasing on $[\underline{\theta}_K, \bar{w}]$. Consequently, we have $0 \leqslant V_{\mathsf{I}}(w) - V_\emptyset(w) \leqslant K$ for $w \in (\underline{\theta}_K, \bar{w})$.

EC.3.9. Proof of Proposition 6

The proof of Proposition 6 is rather intricate, which takes a total of four key steps. These steps illustrate how to identify thresholds $\bar{\vartheta}$ and $\underline{\vartheta}$ in computation. Furthermore, these steps help us establish $\underline{\theta}^0$ in Proposition EC.1. As Condition 3 contains two cases, we consider these cases separately below.

EC.3.9.1. Condition 1 and $K < \bar{K}_1$. In Step 1, fixing any $\underline{\theta}$, we identify bound $\hat{\mathbf{w}}$ and slope c as functions of $\underline{\theta}$ to satisfy (42) and (43).

Lemma EC.4. *For any $\underline{\theta} \in (0, \bar{w})$, there exist unique values $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ and $C(\underline{\theta})$, in place of $\hat{\mathbf{w}}$ and c, respectively, such that value-matching and smooth-pasting conditions (42) and (43) are satisfied at $\underline{\vartheta} = \underline{\theta}$.*

Proof. For any $\widetilde{w} \in [\underline{\theta}, \bar{w})$, define

$$C_1(\widetilde{w}, \underline{\theta}) = (\mathcal{V}_{\widetilde{w}}(\underline{\theta}) - \underline{v} - \underline{\theta})\underline{\theta}^{-r/\rho} \text{ and } C_2(\widetilde{w}, \underline{\theta}) = \rho/r \cdot (\mathcal{V}'_{\widetilde{w}}(\underline{\theta}) - 1)\underline{\theta}^{1-r/\rho}. \tag{EC.70}$$

It follows from Lemma 2(iii) that $C_1(\widetilde{w}, \underline{\theta})$ is decreasing in $\widetilde{w}$ and $C_2(\widetilde{w}, \underline{\theta})$ is increasing in $\widetilde{w}$ on $[\underline{\theta}, \bar{w})$. Note that

$$\begin{aligned} C_1(\underline{\theta}, \underline{\theta}) &= (\mathcal{V}_{\underline{\theta}}(\underline{\theta}) - \underline{v} - \underline{\theta})\underline{\theta}^{-r/\rho} \\ &= \left(\frac{\mu R - c - (\rho - r)\underline{\theta}}{r} - \underline{v} - \underline{\theta}\right) \cdot \underline{\theta}^{-r/\rho} > -\frac{\rho}{r}\underline{\theta}^{1-r/\rho} = C_2(\underline{\theta}, \underline{\theta}), \end{aligned}$$

where the second and the third equalities follow from the boundary conditions at $\underline{\theta}$ (see Lemma EC.2), and the inequality follows from Assumption 1. Besides, Lemma EC.2(v) demonstrates that $C_1(\widetilde{w}, \underline{\theta}) \to -\infty$ and $C_2(\widetilde{w}, \underline{\theta}) \to \infty$ as $\widetilde{w} \uparrow \bar{w}$. In view of Lemma EC.2(i), both $C_1(\widetilde{w}, \underline{\theta})$ and $C_2(\widetilde{w}, \underline{\theta})$ are continuous in $\widetilde{w}$. Hence, there exists a unique $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ such that $C_1(\widetilde{w}(\underline{\theta}), \underline{\theta}) = C_2(\widetilde{w}(\underline{\theta}), \underline{\theta})$. Let $C(\underline{\theta}) := C_1(\widetilde{w}(\underline{\theta}), \underline{\theta})$. Then, $\widetilde{w}(\underline{\theta})$ and $C(\underline{\theta})$ satisfy (42)–(43), as desired. ◻

Step 2 determines an interval to further identify $\underline{\theta}$.

LEMMA EC.5. *Value $\underline{\theta}^0 := \inf\{\underline{\theta} \in (0, \bar{w}) \mid \widetilde{w}'(\underline{\theta}) \geqslant 0\}$ is well defined. Furthermore, we have that $\widetilde{w}(\underline{\theta})$ is decreasing and $C(\underline{\theta})$ is increasing for $\underline{\theta} \in (0, \underline{\theta}^0)$, with $\widetilde{w}'(\underline{\theta}^0) = 0$. Moreover, $C(\underline{\theta}) > 0$ for any $\underline{\theta} \in (0, \underline{\theta}^0)$.*

Proof. Define

$$h(\widetilde{w}, \underline{\theta}) := \mathcal{V}_{\widetilde{w}}(\underline{\theta}) - \underline{v} - \underline{\theta} - \rho/r \cdot (\mathcal{V}'_{\widetilde{w}}(\underline{\theta}) - 1)\underline{\theta}. \tag{EC.71}$$

By Lemma 2(iii), $h(\widetilde{w}, \underline{\theta})$ is decreasing in $\widetilde{w}$. Besides, $h(\widetilde{w}(\underline{\theta}), \underline{\theta}) = 0$. Note that $h(\widetilde{w}, \underline{\theta})$ is continuously differentiable in $\widetilde{w}$ and $\underline{\theta}$ by Lemma EC.2(i) and (ii). Hence, $\widetilde{w}(\underline{\theta})$ is continuously differentiable in $\underline{\theta}$.

Since $h(\widetilde{w}, 0) = \mathcal{V}_{\widetilde{w}}(0) - \underline{v}$, we have $\widetilde{w}(0) = \hat{\mathrm{w}} > 0$. Besides, it follows from $\widetilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ that $\lim_{\underline{\theta} \uparrow \bar{w}} \widetilde{w}(\underline{\theta}) = \bar{w} > \hat{\mathrm{w}}$.

Write $h_1(\widetilde{w}, \underline{\theta}) = \partial h(\widetilde{w}, \underline{\theta})/\partial \widetilde{w}$ and $h_2(\widetilde{w}, \underline{\theta}) = \partial h(\widetilde{w}, \underline{\theta})/\partial \underline{\theta}$. Then, we have $h_1(\widetilde{w}, \underline{\theta}) < 0$,

$$h_2(\widetilde{w}, \underline{\theta}) = \frac{\rho - r - (\rho - r)\mathcal{V}'_{\widetilde{w}}(\underline{\theta}) - \rho\underline{\theta}\mathcal{V}''_{\widetilde{w}}(\underline{\theta})}{r}, \text{ and } \widetilde{w}'(\underline{\theta}) = -\frac{h_2(\widetilde{w}(\underline{\theta}), \underline{\theta})}{h_1(\widetilde{w}(\underline{\theta}), \underline{\theta})}.$$

It follows from $K < \underline{K}_1$ and Lemma 4 that $m^K = \mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K) > 1$, which implies that $\mathcal{V}'_{\hat{\mathrm{w}}}(0) > 1$ by the concavity of $\mathcal{V}_{\hat{\mathrm{w}}}$. Therefore, we have $h_2(\widetilde{w}(0), 0) = (\rho - r - (\rho - r)\mathcal{V}'_{\hat{\mathrm{w}}}(0))/r < 0$, which in turn gives $\widetilde{w}'(0) < 0$. Therefore, $\widetilde{w}(\underline{\theta})$ is strictly decreasing in $\underline{\theta}$ when $\underline{\theta}$ is near 0.

It follows from $\lim_{\underline{\theta}\uparrow\bar{w}} \widetilde{w}(\underline{\theta}) = \bar{w} > \hat{\text{w}} = \widetilde{w}(0)$ and the continuity of $\widetilde{w}'(\underline{\theta})$ in $\underline{\theta}$ that value $\underline{\theta}^0 := \inf\{\underline{\theta} \in (0,\bar{w}) \mid \widetilde{w}'(\underline{\theta}) \geqslant 0\}$ is well defined, satisfying $\widetilde{w}'(\underline{\theta}) < 0$ for any $\underline{\theta} \in [0,\underline{\theta}^0)$ and $\widetilde{w}'(\underline{\theta}^0) = 0$. Consequently, we have

$$\rho - r - (\rho - r)\mathcal{V}'_{\widetilde{w}(w)}(w) - \rho w \mathcal{V}''_{\widetilde{w}(w)}(w) < 0 \text{ for any } w \in [0,\underline{\theta}^0), \text{ and} \tag{EC.72}$$

$$\rho - r - (\rho - r)\mathcal{V}'_{\widetilde{w}(\underline{\theta}^0)}(\underline{\theta}^0) - \rho\underline{\theta}^0 \mathcal{V}''_{\widetilde{w}(\underline{\theta}^0)}(\underline{\theta}^0) = 0. \tag{EC.73}$$

Note that $\widetilde{w}(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$. We claim that $C(\underline{\theta})$ is increasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$. In fact, for any $\underline{\theta} \in (0,\underline{\theta}^0)$, we have

$$\begin{aligned}
C'(\underline{\theta}) = C_1'(\widetilde{w}(\underline{\theta}),\underline{\theta}) &= \frac{\mathrm{d}}{\mathrm{d}\underline{\theta}}\left[\left(\mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - \underline{v} - \underline{\theta}\right)\underline{\theta}^{-r/\rho}\right] \\
&= \left(\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) + \widetilde{w}'(\underline{\theta}) \cdot \frac{\partial \mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta})}{\partial \widetilde{w}(\underline{\theta})} - 1\right)\underline{\theta}^{-r/\rho} - \frac{r}{\rho}\left(\mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) - \underline{v} - \underline{\theta}\right)\underline{\theta}^{-r/\rho-1} \\
&= \widetilde{w}'(\underline{\theta}) \cdot \frac{\partial \mathcal{V}_{\widetilde{w}(\underline{\theta})}(\underline{\theta})}{\partial \widetilde{w}(\underline{\theta})} > 0,
\end{aligned}$$

where the last equality follows from $C_1(\widetilde{w}(\underline{\theta}),\underline{\theta}) = C_2(\widetilde{w}(\underline{\theta}),\underline{\theta})$, and the inequality follows from $\widetilde{w}'(\underline{\theta}) < 0$ and $\partial\mathcal{V}_{\widetilde{w}}(w)/\partial\widetilde{w} < 0$ due to Lemma 2(iii).

Note that $\mathcal{V}'_{\hat{\text{w}}}(0) > 1$. Hence, by the continuity of $\widetilde{w}(\underline{\theta})$ in $\underline{\theta}$ and Lemma EC.2(i) and (ii), there exists a number $\epsilon > 0$ such that $\mathcal{V}'_{\widetilde{w}(\underline{\theta})}(\underline{\theta}) > 1$ for any $\underline{\theta} \in (0,\epsilon)$. Consequently, $C(\underline{\theta}) = C_2(\widetilde{w}(\underline{\theta}),\underline{\theta}) > 0$ for any $\underline{\theta} \in (0,\epsilon)$, which implies that $C(\underline{\theta}) > 0$ for any $\underline{\theta} \in (0,\underline{\theta}^0)$ by noting that $C(\underline{\theta})$ is strictly increasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$. □

Next, in Step 3, we define the upper threshold $\bar{\theta}$ as a function of $\underline{\theta}$, such that smooth pasting condition (45) is satisfied.

Lemma EC.6. *We have:*

(i) *for any $\underline{\theta} \in (0,\underline{\theta}^0)$, the threshold*

$$\bar{\theta}(\underline{\theta}) := \inf\left\{w > \underline{\theta} \mid \mathcal{V}'_{\widetilde{w}(\underline{\theta})}(w) \leqslant 1 + C(\underline{\theta})r/\rho \cdot w^{r/\rho-1}\right\} \tag{EC.74}$$

is well defined;

(ii) *as a function of $\underline{\theta}$, threshold $\bar{\theta}(\underline{\theta})$ is decreasing and continuous in $\underline{\theta}$ on $[0,\underline{\theta}^0)$; and*

(iii) $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \bar{\theta}(\underline{\theta}) = \underline{\theta}^0$.

Proof. (i) Define $\Psi(w,\underline{\theta}) = \mathcal{V}'_{\widetilde{w}(\underline{\theta})}(w) - 1 - C(\underline{\theta})r/\rho \cdot w^{r/\rho-1}$. It follows from Lemma EC.5 and Lemma 2(iii) that $\Psi(w,\underline{\theta})$ is decreasing in $\underline{\theta}$ on $(0,\underline{\theta}^0)$. Therefore, for any $w \in (0,\underline{\theta})$, we have

$\Psi(w,\underline{\theta}) < \Psi(w,w) = 0$, which implies that $\underline{\theta} = \inf\{w \geqslant 0 \mid \Psi(w,\underline{\theta}) = 0\}$ as $\Psi(\underline{\theta},\underline{\theta}) = 0$. Moreover, we have

$$\begin{aligned}\frac{\partial\Psi}{\partial w}(\underline{\theta},\underline{\theta}) &= \mathcal{V}''_{\tilde{w}(\underline{\theta})}(\underline{\theta}) - C(\underline{\theta})r/\rho\cdot(r/\rho-1)\underline{\theta}^{r/\rho-2}\\ &= \mathcal{V}''_{\tilde{w}(\underline{\theta})}(\underline{\theta}) - C_2(\tilde{w}(\underline{\theta}),\underline{\theta})r/\rho\cdot(r/\rho-1)\underline{\theta}^{r/\rho-2}\\ &= \mathcal{V}''_{\tilde{w}(\underline{\theta})}(\underline{\theta}) - \rho/r\cdot(\mathcal{V}'_{\tilde{w}(\underline{\theta})}(\underline{\theta})-1)\underline{\theta}^{1-r/\rho}r/\rho\cdot(r/\rho-1)\underline{\theta}^{r/\rho-2}\\ &= \mathcal{V}''_{\tilde{w}(\underline{\theta})}(\underline{\theta}) - (\mathcal{V}'_{\tilde{w}(\underline{\theta})}(\underline{\theta})-1)\cdot(r/\rho-1)/\underline{\theta}\\ &> 0,\end{aligned}$$

where the last inequality follows from (EC.72). This implies that $\Psi(w,\underline{\theta}) > 0$ for $w \in (\underline{\theta},\underline{\theta}+\epsilon)$ with some $\epsilon > 0$. According to Lemma EC.2(i), $\Psi(w,\underline{\theta})$ is continuous in w. Besides, we have

$$\Psi(\tilde{w}(\underline{\theta}),\underline{\theta}) = \mathcal{V}'_{\tilde{w}(\underline{\theta})}(\tilde{w}(\underline{\theta})) - (1 + C(\underline{\theta})r/\rho\cdot\tilde{w}(\underline{\theta})^{r/\rho-1}) = -\big(1 + C(\underline{\theta})r/\rho\cdot\tilde{w}(\underline{\theta})^{r/\rho-1}\big) < 0$$

and $\Psi(\underline{\theta},\underline{\theta}) = 0$. Hence, $\bar{\theta}(\underline{\theta}) = \inf\{w > \underline{\theta} \mid \Psi(w,\underline{\theta}) \leqslant 0\} = \inf\big\{w > \underline{\theta} \mid \mathcal{V}'_{\tilde{w}(\underline{\theta})}(w) \leqslant 1 + C(\underline{\theta})r/\rho\cdot w^{r/\rho-1}\big\}$ is well defined and satisfies $\bar{\theta}(\underline{\theta}) < \tilde{w}(\underline{\theta})$.

(ii) This follows immediately by noting that $\Psi(w,\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[0,\underline{\theta}^0)$ and is continuous in w and $\underline{\theta}$.

(iii) According to (EC.73), we have $\frac{\partial\Psi}{\partial w}(\underline{\theta}^0,\underline{\theta}^0) = 0$, which implies that $\Psi(\cdot,\underline{\theta}^0)$ attains its local maximum at $\underline{\theta}^0$. Hence, we have $\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\bar{\theta}(\underline{\theta}) = \underline{\theta}^0$ by using $\bar{\theta}(\underline{\theta}) \geqslant \underline{\theta}$. □

Finally, in Step 4, we find an appropriate $\underline{\vartheta}$ to satisfy (44), and define $(\mathsf{c},\hat{\mathbf{w}},\bar{\vartheta})$ as $(C(\underline{\vartheta}),\tilde{w}(\underline{\vartheta}),\bar{\theta}(\underline{\vartheta}))$. To this end, we define function

$$\psi(\underline{\theta}) := \mathcal{V}_{\tilde{w}(\underline{\theta})}(\bar{\theta}(\underline{\theta})) - \big[\underline{v} + \bar{\theta}(\underline{\theta}) + C(\underline{\theta})\,\big(\bar{\theta}(\underline{\theta})\big)^{r/\rho}\big]. \tag{EC.75}$$

In order to satisfy (44), we hope to identify the value $\underline{\vartheta}$ such that $\psi(\underline{\vartheta}) = K$, whose existence is guaranteed by the following result.

Lemma EC.7. *Function $\psi(\underline{\theta})$ is continuous and decreasing in $\underline{\theta}$ on $(0,\underline{\theta}^0)$, and satisfies*

$$\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\psi(\underline{\theta}) = 0,\ \ \textit{and}\ \ \lim_{\underline{\theta}\downarrow 0}\psi(\underline{\theta}) > K.$$

Consequently, there exists a unique number $\underline{\vartheta} \in (0,\underline{\theta}^0)$ such that $\psi(\underline{\vartheta}) = K$.

Proof. Note that

$$\psi(\underline{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}(\underline{\theta})}[\mathcal{V}'_{\tilde{w}(\underline{\theta})}(y) - (1 + C(\underline{\theta})r/\rho\cdot y^{r/\rho-1})]\mathrm{d}y = \int_{\underline{\theta}}^{\bar{\theta}(\underline{\theta})}\Psi(y,\underline{\theta})\mathrm{d}y.$$

Fix any $\underline{\theta}^1 < \underline{\theta}^2$ in $(0, \underline{\theta}^0)$. We have

$$\psi(\underline{\theta}^1) = \int_{\underline{\theta}^1}^{\bar{\theta}(\underline{\theta}^1)} \Psi(y, \underline{\theta}^1)\mathrm{d}y > \int_{\underline{\theta}^2}^{\bar{\theta}(\underline{\theta}^2)} \Psi(y, \underline{\theta}^1)\mathrm{d}y > \int_{\underline{\theta}^2}^{\bar{\theta}(\underline{\theta}^2)} \Psi(y, \underline{\theta}^2)\mathrm{d}y = \psi(\underline{\theta}^2),$$

where the first inequality follows from $\Psi(y, \underline{\theta}^1) > 0$ for $y \in \left(\underline{\theta}^1, \underline{\theta}^2\right) \cup \left(\bar{\theta}(\underline{\theta}^2), \bar{\theta}(\underline{\theta}^1)\right)$, and the second inequality uses the fact that $\Psi(y, \underline{\theta})$ is decreasing in $\underline{\theta}$ on $(0, \underline{\theta}^0)$. Hence, $\psi(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $(0, \underline{\theta}^0)$. The continuity of $\psi(\underline{\theta})$ follows from Lemma EC.2(i) and (ii).

Since $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \bar{\theta}(\underline{\theta}) = \underline{\theta}^0$, we have $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \psi(\underline{\theta}) = 0$. Note that $\tilde{w}(0) = \hat{\mathrm{w}}$ and $C_2(\tilde{w}, 0) = 0$ for any $\tilde{w} \in [0, \bar{w})$. Hence, we have $\lim_{\underline{\theta}\downarrow 0} C(\underline{\theta}) = 0$ and thus $\lim_{\underline{\theta}\downarrow 0} \bar{\theta}(\underline{\theta}) = \inf\{w > 0 \mid \mathcal{V}'_{\hat{\mathrm{w}}}(w) = 1\}$. This yields

$$\begin{aligned}\lim_{\underline{\theta}\downarrow 0} \psi(\underline{\theta}) &= \int_0^{\infty} (\mathcal{V}'_{\hat{\mathrm{w}}}(y) - 1)^+ dy > \int_0^{\bar{\theta}^K} (\mathcal{V}'_{\hat{\mathrm{w}}}(y) - 1) dy \\ &= \mathcal{V}_{\hat{\mathrm{w}}}(\bar{\theta}^K) - \underline{v} - \bar{\theta}^K = K + (m^K - 1)\bar{\theta}^K > K,\end{aligned}$$

where the first inequality follows from the facts that $\mathcal{V}'_{\hat{\mathrm{w}}}(\bar{\theta}^K) = m^K > 1$ and that $\mathcal{V}'_{\hat{\mathrm{w}}}$ is non-increasing, and the last inequality holds due to $m^K > 1$. Consequently, it follows from the continuity of $\psi(\cdot)$ that there exists a unique $\underline{\vartheta} \in (0, \underline{\theta}^0)$ such that $\psi(\underline{\vartheta}) = K$. □

According to these results, the quadruple $(\hat{\mathbf{w}}, \mathsf{c}, \underline{\vartheta}, \bar{\vartheta})$ defined by $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta})$, $\mathsf{c} = C(\underline{\vartheta})$ and $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ satisfies (42)–(45). Besides, it follows from $\bar{\theta}(\underline{\theta}) < \tilde{w}(\underline{\theta})$ for $\underline{\theta} \in [0, \underline{\theta}^0)$ that $\bar{\vartheta} < \hat{\mathbf{w}}$, which further implies $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta}) < \tilde{w}(0) = \hat{\mathrm{w}}$ by noting that $\tilde{w}(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[0, \underline{\theta}^0)$. To complete the proof, we need to show that $\bar{\vartheta} > \breve{w}(\hat{\mathbf{w}})$. If it fails to hold, then we have $\mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = \mathcal{V}'_{\hat{\mathbf{w}}}(\bar{\vartheta}) = \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))$. On the other side, it follows from $\mathsf{c} > 0$ and $\underline{\vartheta} < \bar{\vartheta}$ that $\mathsf{V}'_{\mathsf{c}}(\underline{\vartheta}) > \mathsf{V}'_{\mathsf{c}}(\bar{\vartheta})$. This contradicts (43) and (45).

EC.3.9.2. Condition 2 and $K < \underline{K}$. Since most arguments are exactly the same as those as in the previous case, we only provide a sketch here. To start, we observe that $\breve{\underline{\theta}} := \frac{(\bar{V} - \underline{v})(\rho - r - \mu) - (\rho - r)\bar{w}}{\mu(\rho/r - 1)} \in (0, \underline{\theta}_K)$ satisfies $m(\breve{\underline{\theta}}) = \frac{\rho - r}{\rho - r - \mu}$ by (EC.67). Moreover, we have

$$\bar{V} - \underline{v} - \bar{w} - C^1(\breve{\underline{\theta}}) \cdot \bar{w}^{r/\rho - 1} > K \tag{EC.76}$$

since $\psi_1(\breve{\underline{\theta}}) > \psi_1(\underline{\theta}_K) = 0$.

Note that under Condition 2, $\breve{w}(\tilde{w}) = 0$ by Lemma 2(i). Hence, we will use $V_{\tilde{w}}$ instead of $\mathcal{V}_{\tilde{w}}$ in the proof. Next, we will show the desired result by the following four lemmas, which parallel Lemmas EC.4–EC.7 in Section EC.3.9.1.

LEMMA EC.8. *For any $\underline{\theta} \in (\breve{\underline{\theta}}, \bar{w})$, there exists unique values $\tilde{w}(\underline{\theta}) \in (\underline{\theta}, \bar{w})$ and $C(\underline{\theta})$, in place of $\tilde{\mathrm{w}}$ and c, such that (42)–(43) are satisfied at $\underline{\vartheta} = \underline{\theta}$.*

Proof. We will use functions $C_1(\widetilde{w},\underline{\theta})$ and $C_2(\widetilde{w},\underline{\theta})$ defined as in the proof of Lemma EC.4 to obtain the desired result. In the proof of Lemma EC.4, we have established that $C_1(\widetilde{w},\underline{\theta})\to-\infty$ and $C_2(\widetilde{w},\underline{\theta})\to\infty$ as $\widetilde{w}\uparrow\bar{w}$ and thus

$$\lim_{\widetilde{w}\uparrow\bar{w}} C_1(\widetilde{w},\underline{\theta}) < \lim_{\widetilde{w}\uparrow\bar{w}} C_2(\widetilde{w},\underline{\theta}) \tag{EC.77}$$

for any $\underline{\theta}\in(0,\bar{w})$. Now, we claim that (EC.77) still holds under Condition 2 and $K<\underline{K}$ for any $\underline{\theta}\in(\breve{\underline{\theta}},\bar{w})$. In fact, it follows from Lemma EC.2(vi) that

$$\lim_{\widetilde{w}\uparrow\bar{w}} C_1(\widetilde{w},\underline{\theta}) = \left(\bar{V}-\frac{\rho-r}{\rho-r-\mu}\bar{w}-\underline{v}+\frac{\mu}{\rho-r-\mu}\underline{\theta}\right)\underline{\theta}^{-r/\rho} \text{ and}$$
$$\lim_{\widetilde{w}\uparrow\bar{w}} C_2(\widetilde{w},\underline{\theta}) = \frac{\rho}{r}\frac{\mu}{\rho-r-\mu}\cdot\underline{\theta}^{1-r/\rho}.$$

It is clear that

$$\frac{\lim_{\widetilde{w}\uparrow\bar{w}} C_1(\widetilde{w},\underline{\theta})}{\lim_{\widetilde{w}\uparrow\bar{w}} C_2(\widetilde{w},\underline{\theta})} = \frac{\bar{V}-\frac{\rho-r}{\rho-r-\mu}\bar{w}-\underline{v}+\frac{\mu}{\rho-r-\mu}\underline{\theta}}{\frac{\rho}{r}\frac{\mu}{\rho-r-\mu}\underline{\theta}}$$

is decreasing in $\underline{\theta}$ and takes value 1 at $\breve{\underline{\theta}}$. Hence, (EC.77) holds for any $\underline{\theta}\in(\breve{\underline{\theta}},\bar{w})$. The remaining argument is exactly the same as that for Lemma EC.4 and thus omitted. Moreover, we have the following byproduct:

$$\widetilde{w}(\breve{\underline{\theta}}) := \lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}} \widetilde{w}(\underline{\theta}) = \bar{w} \text{ and } C(\breve{\underline{\theta}}) = \frac{\rho\mu}{r(\rho-r-\mu)}(\breve{\underline{\theta}})^{1-r/\rho}, \tag{EC.78}$$

which will be used in the subsequent analysis. ▫

Lemma EC.9. *Value* $\underline{\theta}^0 := \inf\{\underline{\theta}\in(\breve{\underline{\theta}},\bar{w}) \mid \widetilde{w}'(\underline{\theta}^0)\geqslant 0\}$ *is well defined. We have* $\widetilde{w}(\underline{\theta})$ *is strictly decreasing in* $\underline{\theta}$*, and* $C(\underline{\theta})$ *is strictly increasing in* $\underline{\theta}$ *on* $(\breve{\underline{\theta}},\underline{\theta}^0)$ *with* $\widetilde{w}'(\underline{\theta}^0)=0$*. Moreover,* $C(\underline{\theta})>0$ *for any* $\underline{\theta}\in(\breve{\underline{\theta}},\underline{\theta}^0)$*.*

Proof. We only point out the differences between this proof and that of Lemma EC.5. First, we show $\widetilde{w}'(\breve{\underline{\theta}})<0$ instead of $\widetilde{w}'(0)<0$. This holds by noting that $h_2(\widetilde{w}(\breve{\underline{\theta}}),\breve{\underline{\theta}}) = \lim_{\widetilde{w}\uparrow\bar{w}} h_2(\widetilde{w},\breve{\underline{\theta}}) = (\rho-r-(\rho-r)\cdot(\rho-r)/(\rho-r-\mu))/r<0$. Second, we use the result $\lim_{\underline{\theta}\uparrow\bar{w}}\widetilde{w}(\underline{\theta}) = \lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\widetilde{w}(\underline{\theta})=\bar{w}$ instead of $\lim_{\underline{\theta}\uparrow\bar{w}}\widetilde{w}(\underline{\theta})=\bar{w}>\hat{\mathrm{w}}=\widetilde{w}(0)$ to establish the existence of $\underline{\theta}^0$. Finally, we use $V'_{\widetilde{w}(\breve{\underline{\theta}})}(\breve{\underline{\theta}}) = \lim_{\widetilde{w}\uparrow\bar{w}} V'_{\widetilde{w}}(\breve{\underline{\theta}}) = (\rho-r)/(\rho-r-\mu)>1$ to characterize the monotonicity of $C(\cdot)$ near $\breve{\underline{\theta}}$, instead of using $\mathcal{V}'_{\hat{\mathrm{w}}}(0)>1$ to characterize the monotonicity of $C(\cdot)$ near 0. ▫

Lemma EC.10. *For any* $\underline{\theta}\in(\breve{\underline{\theta}},\underline{\theta}^0)$*, the threshold* $\bar{\theta}(\underline{\theta})$

$$\bar{\theta}(\underline{\theta}) := \inf\left\{w>\underline{\theta} \mid V'_{\widetilde{w}(\underline{\theta})}(w)\leqslant 1+C(\underline{\theta})r/\rho\cdot w^{r/\rho-1}\right\}$$

is well defined. As a function of $\underline{\theta}$*, threshold* $\bar{\theta}(\underline{\theta})$ *is decreasing in* $\underline{\theta}$ *on* $[0,\underline{\theta}^0)$*, satisfying* $\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\bar{\theta}(\underline{\theta})=\underline{\theta}^0$ *and* $\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\bar{\theta}(\underline{\theta})=\bar{w}$*.*

Proof. The proof is the same as that for Lemma EC.6, with the range of $\underline{\theta}$ changed from $(0,\underline{\theta}^0)$ to $(\breve{\underline{\theta}},\underline{\theta}^0)$. One exception is that we need to show $\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\bar{\theta}(\underline{\theta})=\bar{w}$. To show this, we first note that for any $w\in(\breve{\underline{\theta}},\bar{w})$, we have

$$\begin{aligned}\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\Psi(w,\underline{\theta})&=\lim_{\tilde{w}\uparrow\bar{w}}\left\{V_{\tilde{w}}'(w)-1-C(\breve{\underline{\theta}})\cdot r/\rho\cdot w^{r/\rho-1}\right\}\\&=\frac{\rho-r}{\rho-r-\mu}-1-\frac{\rho\mu}{r(\rho-r-\mu)}(\breve{\underline{\theta}})^{1-r/\rho}\cdot r/\rho\cdot w^{r/\rho-1}\\&=\frac{\mu}{\rho-r-\mu}\Big[1-\Big(\frac{w}{\breve{\underline{\theta}}}\Big)^{r/\rho-1}\Big]>0,\end{aligned}$$

where the first equality follows from (EC.78) and Lemma EC.2(ii), and the second equality follows from Lemma EC.2(vi). This inequality, together with $\bar{\theta}(\underline{\theta})<\tilde{w}(\underline{\theta})$, yields that $\lim_{\underline{\theta}\downarrow\breve{\underline{\theta}}}\bar{\theta}(\underline{\theta})=\bar{w}$. ▫

Lemma EC.11. *Function $\psi(\underline{\theta})$, as defined in (EC.75), is continuous and decreasing in $\underline{\theta}$ on $(\breve{\underline{\theta}},\underline{\theta}^0)$, and satisfies*

$$\lim_{\underline{\theta}\uparrow\underline{\theta}^0}\psi(\underline{\theta})=0\ \textit{and}\ \lim_{\underline{\theta}\downarrow 0}\psi(\underline{\theta})>K.$$

Consequently, there exists a unique value $\underline{\vartheta}\in(\breve{\underline{\theta}},\underline{\theta}^0)$ such that $\psi(\underline{\vartheta})=K$.

Proof. The proof is exactly the same as that for Lemma EC.7, with the range of $\underline{\theta}$ changed from $(0,\underline{\theta}^0)$ to $(\breve{\underline{\theta}},\underline{\theta}^0)$, except that we will show $\lim_{\underline{\theta}\downarrow\underline{\theta}^0}\psi(\underline{\theta})>K$ rather than $\lim_{\underline{\theta}\downarrow 0}\psi(\underline{\theta})>K$. In fact, we have

$$\lim_{\underline{\theta}\downarrow\underline{\theta}^0}\psi(\underline{\theta})=\lim_{\tilde{w}\uparrow\bar{w}}V_{\tilde{w}}(\bar{w})-[\underline{v}+\bar{w}+C(\breve{\underline{\theta}})\,(\bar{w})^{r/\rho}]=\bar{V}-\underline{v}-\bar{w}-C(\breve{\underline{\theta}})\,(\bar{w})^{r/\rho}>K,$$

where the second equality uses Lemma EC.2(vi), and the inequality follows from (EC.76). ▫

Now we are ready to complete the proof of Proposition 6 under Condition 2 and $K<\underline{K}$. According to Lemmas EC.8–EC.11, $(\hat{\mathbf{w}},\mathsf{c},\underline{\vartheta},\bar{\vartheta})$ defined by $\hat{\mathbf{w}}=\tilde{w}(\underline{\vartheta})$, $\mathsf{c}=C(\underline{\vartheta})$ and $\bar{\vartheta}=\bar{\theta}(\underline{\vartheta})$ satisfies (42)–(45). Besides, it follows from $\bar{\theta}(\underline{\theta})<\tilde{w}(\underline{\theta})$ for $\underline{\theta}\in[\breve{\underline{\theta}},\underline{\theta}^0)$ that $\bar{\vartheta}<\hat{\mathbf{w}}$, which implies $\hat{\mathbf{w}}=\tilde{w}(\underline{\vartheta})<\tilde{w}(\underline{\theta}^0)=\bar{w}$ by using that $\tilde{w}(\underline{\theta})$ is decreasing in $\underline{\theta}$ on $[\breve{\underline{\theta}},\underline{\theta}^0)$.

EC.3.10. Proof of Proposition 7

We only consider the case in which Condition 1 and $K<\bar{K}_1$ hold, because the proof is the same for the case in which Condition 2 and $K<\underline{K}$ hold.

First, we prove (48). Similar to the proof of Proposition 2, we apply Lemma 1 and Proposition EC.3 by verifying that (EC.47)–(EC.51) all hold.

Equality (EC.47) holds by noting that (i) $\ell_t=b\mathbb{1}_{\nu_t=\mu}$; (ii) for any $t>0$, $W_{t-}\in[\underline{\vartheta},\hat{\mathbf{w}}]$ if $\mathcal{E}_{t-}=\mathsf{I}$ and $(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=0$ if $w\in[\underline{\vartheta},\hat{\mathbf{w}}]$; and (iii) for any $t>0$, $W_{t-}\in(0,\bar{\vartheta}]$ if $\mathcal{E}_{t-}=0$ and $(\mathcal{A}_{\emptyset}V_{\emptyset})(w)=0$ if $w\in(0,\bar{\vartheta}]$.

Equality (EC.48) holds by noting that $\Delta L_t > 0$ only if $\mathcal{E}_{t-} = 1$ and $W_t + H_t \mathrm{d}N_t - H_t^q \mathrm{d}Q_t > \hat{\mathbf{w}}$, as well as that $V_{\mathsf{I}}(w) = V_{\mathsf{I}}(\hat{\mathbf{w}})$ for any $w \geqslant \hat{\mathbf{w}}$.

Equality (EC.49) holds by noting that for any $t \geqslant 0$, (i) $\mathcal{E}_t = 1 - \mathcal{E}_{t-} = 1$ only if $W_{t-} \in [\bar{\vartheta}, \hat{\mathbf{w}}]$ and $V_{\mathsf{I}}(w) - V_{\emptyset}(w) = K$ if $w \in [\bar{\vartheta}, \hat{\mathbf{w}}]$; and (ii) $\mathcal{E}_t = 1 - \mathcal{E}_{t-} = 0$ only if $W_{t-} \in (0, \underline{\vartheta}]$ and $V_{\emptyset}(w) - V_{\mathsf{I}}(w) = 0$ if $w \in (0, \underline{\vartheta}]$.

Note that $q_t > 0$ only if $\breve{w}(\hat{\mathbf{w}}) > 0$, $W_{t-} = \breve{w}(\hat{\mathbf{w}})$ and $\mathcal{E}_{t-} = \mathsf{I}$. Hence, if $q_t > 0$, then we have

$$\begin{aligned} &H_t^q \cdot \left(V'_{\mathcal{E}_{t-}}(W_{t-}) - 1\right) + \phi(W_{t-} - H_t^q, \mathcal{E}_{t-}) - \phi(W_{t-}, \mathcal{E}_{t-}) \\ =&(\breve{w}(\hat{\mathbf{w}}) - \underline{\vartheta})\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) + \mathcal{V}_{\hat{\mathbf{w}}}(\underline{\vartheta}) - \mathcal{V}_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = 0, \end{aligned}$$

where the last equality follows from Lemma 2(ii). Hence, (EC.50) holds.

Finally, (EC.51) holds by noting that (i) if $W_{0-} \leqslant \hat{\mathbf{w}}$, then $\mathbb{E}[\phi(W_0, \mathcal{E}_{0-})] - \phi(W_{0-}, \mathcal{E}_{0-}) - \mathbb{E}\Delta L_0 = \phi(W_{0-}, \mathcal{E}_{0-}) - \phi(W_{0-}, \mathcal{E}_{0-}) = 0$; and (ii) if $W_{0-} > \hat{\mathbf{w}}$, then $\mathbb{E}[\phi(W_0, \mathcal{E}_{0-})] - \phi(W_{0-}, \mathcal{E}_{0-}) - \mathbb{E}\Delta L_0 = \phi(\hat{\mathbf{w}}, \mathcal{E}_{0-}) - \phi(W_{0-}, \mathcal{E}_{0-}) - (W_{0-} - \hat{\mathbf{w}}) = 0$.

Next, we show that functions $V_{\mathsf{I}}(w)$ and $V_{\emptyset}(w)$ as defined in (46) satisfy the optimality condition (20)–(22). First, (22) holds by noting that $V_{\mathsf{I}}(0) = V_{\emptyset}(0) = \underline{v}$. To verify (20) and (21), we consider the following three cases separately: $w \in [0, \underline{\vartheta})$, $w \in [\underline{\vartheta}, \hat{\mathbf{w}})$, and $w \in [\hat{\mathbf{w}}, \infty)$. We will study the case of $w \in [\underline{\vartheta}, \hat{\mathbf{w}})$ before $w \in [0, \underline{\vartheta})$.

Case 1: $w \in [\underline{\vartheta}, \hat{\mathbf{w}})$. First, we prove that

$$(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) \geqslant 0 \text{ on } [\underline{\vartheta}, \hat{\mathbf{w}}). \tag{EC.79}$$

Obviously, we have $(\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) = 0$ for $w \in [\breve{w}(\hat{\mathbf{w}}), \hat{\mathbf{w}})$. It remains to show that (EC.79) holds for $w \in [\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}}))$ if $\underline{\vartheta} < \breve{w}(\hat{\mathbf{w}}) < \hat{\mathbf{w}}$. For $w \in [\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}}))$, function $\mathcal{V}_{\hat{\mathbf{w}}}$ is linear and thus we have

$$\begin{aligned} (\mathcal{A}_{\mathsf{I}} V_{\mathsf{I}})(w) &= (\mu + r)\mathcal{V}_{\hat{\mathbf{w}}}(w) - \mu \mathcal{V}_{\hat{\mathbf{w}}}(w + \beta) + \rho(\bar{w} - w)\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - (\mu R - c) + (\rho - r)w \\ &=: g_{\mathsf{I}}(w). \end{aligned}$$

Note that

$$\begin{aligned} g'_{\mathsf{I}}(w) &= (\mu + r)\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mu \mathcal{V}'_{\hat{\mathbf{w}}}(w + \beta) - \rho \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) + \rho - r \\ &= (\rho - r)\left(1 - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))\right) + \mu\left(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mathcal{V}'_{\hat{\mathbf{w}}}(w + \beta)\right) \\ &\leqslant (\rho - r)\left(1 - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))\right) + \mu\left(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}) + \beta)\right) = 0, \end{aligned}$$

where the inequality uses the concavity of $\mathcal{V}_{\hat{\mathbf{w}}}$ and the last equality uses

$$0 = \rho(\bar{w} - \breve{w}(\hat{\mathbf{w}}))\mathcal{V}''_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = (\rho - r)\left(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - 1\right) + \mu\left(\mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}) + \beta) - \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}}))\right).$$

Consequently, $g_{\mathsf{I}}(w) \geqslant 0$ for all $w \in [\underline{\vartheta}, \breve{w}(\hat{\mathbf{w}}))$, which yields (EC.79). Note that $(\mathcal{A}_{\emptyset} V_{\emptyset})(w) = 0$ if $w \in [\underline{\vartheta}, \bar{\vartheta}]$. Hence, (20) holds by the following result, whose proof is relegated to Section EC.3.11.

Lemma EC.12. *Under the conditions stated in Proposition 7, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for $w \in [\bar{\vartheta}, \hat{\mathbf{w}})$.*

If $w \in [\bar{\vartheta}, \hat{\mathbf{w}})$, then $V_\mathsf{l}(w) - V_\emptyset(w) = K > 0$. To establish (21), we need to show that $0 \leqslant V_\mathsf{l}(w) - V_\emptyset(w) \leqslant K$ if $w \in [\underline{\vartheta}, \bar{\vartheta}]$.

Let $\Phi(w) := V_\mathsf{l}(w) - V_\emptyset(w)$ and $\chi(w) := \mathcal{V}'_{\hat{\mathbf{w}}}(w) - 1 - \mathsf{c} \cdot r/\rho \cdot w^{r/\rho - 1}$. Obviously, we have $\Phi(\underline{\vartheta}) = 0$ and $\chi(\underline{\vartheta}) = \chi(\bar{\vartheta}) = 0$. It follows from the proof of Lemma EC.6(i) that $\Phi'(w) = \chi(w) > 0$ for any $w \in (\underline{\vartheta}, \bar{\vartheta})$. Hence, for any $w \in [\underline{\vartheta}, \bar{\vartheta})$, we have $\Phi(w) \geqslant \Phi(\underline{\vartheta}) = 0$ and $\Phi(w) \leqslant \Phi(\bar{\vartheta}) = K$.

Case 2: $w \in [0, \underline{\vartheta})$. We claim that:

Lemma EC.13. *Under the conditions stated in Proposition 7, we have $(\mathcal{A}_\mathsf{l} V_\mathsf{l})(w) \geqslant 0$ for $w \in [0, \underline{\vartheta})$.*

Its proof is rather involved, which is relegated to Section EC.3.12. Obviously, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) = 0$ on $[0, \underline{\vartheta})$. Hence, (20) holds. Inequality (21) also holds by noting that $V_\mathsf{l}(w) = V_\emptyset(w)$ in this case.

Case 3: $w \in [\hat{\mathbf{w}}, \infty)$. Using the boundary condition $\mathcal{V}_{\hat{\mathbf{w}}}(\hat{\mathbf{w}}) = (\mu R - c - (\rho - r)\hat{\mathbf{w}})/r$, we have

$$(\mathcal{A}_\mathsf{l} V_\mathsf{l})(w) = r\mathcal{V}_{\hat{\mathbf{w}}}(\hat{\mathbf{w}}) - (\mu R - c) + (\rho - r)w = (\rho - r)(w - \hat{\mathbf{w}}) \geqslant 0,$$

and

$$\begin{aligned}(\mathcal{A}_\emptyset V_\emptyset)(w) &= r(\mathcal{V}_{\hat{\mathbf{w}}}(\hat{\mathbf{w}}) - K) + (\rho - r)w - R\underline{\mu} \\ &= \mu R - c - (\rho - r)\hat{\mathbf{w}} + (\rho - r)w - R\underline{\mu} - rK \\ &= R\Delta\mu - c + (\rho - r)(w - \hat{\mathbf{w}}) - rK \geqslant 0,\end{aligned}$$

where the last inequality follows as $K < \bar{V}(\hat{\mathrm{w}}) - \underline{v} = (\mu R - c - (\rho - r)\hat{\mathrm{w}})/r - R\underline{\mu}/r < (R\Delta\mu - c)/r$. Hence, (20) holds. Inequality (21) also holds since $V_\mathsf{l}(w) - V_\emptyset(w) = K$.

EC.3.11. Proof of Lemma EC.12

For $w \in [\bar{\vartheta}, \hat{\mathbf{w}})$, it holds that $V_\emptyset(w) = \mathcal{V}_{\hat{\mathbf{w}}}(w) - K$. Define $\varpi_0 := \inf\{w > 0 \mid \mathcal{V}'_{\hat{\mathrm{w}}}(w) = 1\}$, which is well defined by noting that $\mathcal{V}'_{\hat{\mathrm{w}}}(0) > 1$ and $\mathcal{V}'_{\hat{\mathrm{w}}}(\hat{\mathrm{w}}) = 0$. For any $w \in [\varpi_0, \hat{\mathbf{w}}]$, let

$$(\mathcal{A}_\emptyset V_\emptyset)(w) = r(\mathcal{V}_{\hat{\mathbf{w}}}(w) - K) - \rho w \mathcal{V}'_{\hat{\mathbf{w}}}(w) + (\rho - r)w - \underline{\mu}R =: g_\emptyset(w).$$

It follows from Lemma 2(iii) and $\hat{\mathbf{w}} < \hat{\mathrm{w}}$ that $\mathcal{V}'_{\hat{\mathbf{w}}}(w) < \mathcal{V}'_{\hat{\mathrm{w}}}(w)$ and $\mathcal{V}_{\hat{\mathbf{w}}}(w) > \mathcal{V}_{\hat{\mathrm{w}}}(w)$ for $w \in [0, \hat{\mathbf{w}}]$. Hence, for $w \in [\varpi_0, \hat{\mathbf{w}}]$, we have

$$g'_\emptyset(w) = r\mathcal{V}'_{\hat{\mathbf{w}}}(w) - \rho\mathcal{V}'_{\hat{\mathbf{w}}}(w) - \rho w\mathcal{V}''_{\hat{\mathbf{w}}}(w) + \rho - r = (\rho - r)\big(1 - \mathcal{V}'_{\hat{\mathbf{w}}}(w)\big) - \rho w \mathcal{V}''_{\hat{\mathbf{w}}}(w) \geqslant 0,$$

where the last inequality follows from the fact that $\mathcal{V}'_{\hat{\mathbf{w}}}(w) \leqslant \mathcal{V}'_{\hat{\mathrm{w}}}(w) \leqslant 1$ for $w \geqslant \varpi_0$ and the concavity of $\mathcal{V}_{\hat{\mathbf{w}}}$. In addition, we have

$$\begin{aligned} g_\emptyset(\varpi_0) &> r(\mathcal{V}_{\hat{\mathbf{w}}}(\varpi_0) - K) - \rho\varpi_0 + (\rho - r)\varpi_0 - \underline{\mu}R \\ &> r(\underline{v} + \varpi_0 + K - K) - \rho\varpi_0 + (\rho - r)\varpi_0 - \underline{\mu}R = 0, \end{aligned}$$

where the first inequality uses $\mathcal{V}'_{\hat{\mathbf{w}}}(\varpi_0) < \mathcal{V}'_{\hat{\mathrm{w}}}(\varpi_0) = 1$, and the second inequality follows from the fact that $\mathcal{V}_{\hat{\mathbf{w}}}(\varpi_0) > \mathcal{V}_{\hat{\mathrm{w}}}(\varpi_0) > \underline{v} + \varpi_0 + K$ (the last inequality holds because of $m^K > 1$). As a result, we have $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for all $w \in [\varpi_0, \hat{\mathbf{w}}]$.

Next, we prove that $(\mathcal{A}_\emptyset V_\emptyset)(w) \geqslant 0$ for $w \in [\bar{\vartheta}, \varpi_0)$ by a contradictory argument.

Suppose, to the contradictory, that there exists a number $\varpi \in (\bar{\vartheta}, \varpi_0)$ such that $(\mathcal{A}_\emptyset V_\emptyset)(\varpi) < 0$. Then, we have $(\mathcal{A}_\emptyset V_\emptyset)(\varpi) = r(\mathcal{V}_{\hat{\mathbf{w}}}(\varpi) - K) - \rho\varpi \cdot \mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) + (\rho - r)\varpi - \underline{\mu}R < 0$, and thus

$$\mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) > \frac{(\rho - r)\varpi + r\big(\mathcal{V}_{\hat{\mathbf{w}}}(\varpi) - K - \underline{v}\big)}{\rho\varpi}. \tag{EC.80}$$

It follows from (EC.74) that $\lim_{\underline{\theta}\downarrow 0} \bar{\theta}(\underline{\theta}) = \inf\{w > 0 \mid \mathcal{V}'_{\hat{\mathrm{w}}}(w) = 1\} = \varpi_0$. Note that $\varpi > \bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ and $\varpi < \varpi_0$. Hence, it follows from Lemma EC.6 that there exists a number $\underline{\theta}' \in (0, \underline{\vartheta})$ such that $\bar{\theta}(\underline{\theta}') = \varpi$. Using Lemmas EC.5 and EC.7, we have $\tilde{w}(\underline{\theta}') > \tilde{w}(\underline{\vartheta}) = \hat{\mathbf{w}}$, $C(\underline{\theta}') < C(\underline{\vartheta}) = \mathsf{c}$, and

$$\psi(\underline{\theta}') = \mathcal{V}_{\tilde{w}(\underline{\theta}')}(\varpi) - \big[\underline{v} + \varpi + C(\underline{\theta}')\varpi^{r/\rho}\big] > \psi(\underline{\vartheta}) = K. \tag{EC.81}$$

Moreover, according to 2(iii), we have

$$\mathcal{V}_{\hat{\mathbf{w}}}(\varpi) > \mathcal{V}_{\tilde{w}(\underline{\theta}')}(\varpi) \text{ and } \mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) < \mathcal{V}'_{\tilde{w}(\underline{\theta}')}(\varpi). \tag{EC.82}$$

Consequently,

$$\begin{aligned} \mathcal{V}'_{\hat{\mathbf{w}}}(\varpi) &> \frac{(\rho - r)\varpi + r\big(\mathcal{V}_{\tilde{w}(\underline{\theta}')}(\varpi) - K - \underline{v}\big)}{\rho\varpi} \\ &> \frac{(\rho - r)\varpi + r \cdot \big[(\underline{v} + \varpi + C(\underline{\theta}')\varpi^{r/\rho}) - \underline{v}\big]}{\rho\varpi} \\ &= 1 + r/\rho \cdot C(\underline{\theta}')\varpi^{r/\rho - 1} = \mathcal{V}'_{\tilde{w}(\underline{\theta}')}(\varpi), \end{aligned}$$

where the first inequality uses (EC.80) and (EC.82), the second inequality uses (EC.81), and the last equality follows from $\varpi = \bar{\theta}(\underline{\theta}')$ and the definition of $\bar{\theta}(\cdot)$. This reaches a contradiction with (EC.82).

EC.3.12. Proof of Lemma EC.13

The proof of Lemma EC.13 is probably the most complex proof in the paper. As mentioned in the paragraph below Proposition 7, the key step is to establish Lemma EC.16 below, which states that either $\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}}$'s first-order derivative is negative, or its second-order derivative is positive on $(0,\underline{\vartheta})$. This crucial result is obtained by studying a total of four cases, which are summarized as Lemmas EC.17–EC.20.

Following from $V_{\mathsf{I}}(w)=\underline{v}+w+\mathsf{c}w^{r/\rho}$ for $w\in[0,\underline{\vartheta}]$ and

$$(\mathcal{A}_{\mathsf{I}}f-\mathcal{A}_{\emptyset}f)(w)=\mu(f(w)-f(w+\beta))+\rho\bar{w}f'(w)-(R\Delta\mu-c),$$

we define

$$\begin{aligned}g_{\mathsf{I}}(w)&:=(\mathcal{A}_{\mathsf{I}}V_{\mathsf{I}})(w)=\mu(V_{\mathsf{I}}(w)-V_{\mathsf{I}}(w+\beta))+\rho\bar{w}V_{\mathsf{I}}'(w)-(R\Delta\mu-c)\\&=\mu\big(\underline{v}+w+\mathsf{c}w^{r/\rho}-V_{\mathsf{I}}(w+\beta)\big)+\rho\bar{w}(1+\mathsf{c}r/\rho\cdot w^{r/\rho-1})-(R\Delta\mu-c)\end{aligned}$$

for $w>0$.

Using the same argument as that in Lemma EC.2(i) and by the definition of $\mathcal{V}_{\tilde{w}}$ as stated in Lemma 2(ii), we can obtain the following result. Its proof is omitted for brevity.

Lemma EC.14. *For any $\tilde{w}\in(0,\bar{w})$, we have $\mathcal{V}_{\tilde{w}}\in C^1(\mathbb{R}_+)\cap C^2(\mathbb{R}_+\backslash\{\tilde{w}\})\cap C^3(\mathbb{R}_+\backslash\{\tilde{w},\tilde{w}-\beta,\breve{w}(\tilde{w})\})\cap C^4(\mathbb{R}_+\backslash\{\tilde{w},\tilde{w}-\beta,\tilde{w}-2\beta,\breve{w}(\tilde{w})\})$.*

By the definition of V_{I} as in (46) and the smooth-pasting condition at $\underline{\vartheta}$, V_{I} is differentiable at $\underline{\vartheta}$, but may not be twice differentiable at $\underline{\vartheta}$. Therefore, $V_{\mathsf{I}}\in C^1(\mathbb{R}_{++})\cap C^2(\mathbb{R}_{++}\backslash\{\hat{\mathbf{w}},\underline{\vartheta}\})$, which implies that $g_{\mathsf{I}}\in C^1(\mathbb{R}_{++})\cap C^2(\mathbb{R}_{++}\backslash\{\hat{\mathbf{w}}-\beta,\underline{\vartheta}-\beta\})$. (Here, we use $\mathbb{R}_{++}$ to denote the set of all positive numbers.) Besides, it holds that

$$g_{\mathsf{I}}'(w)=\mu\big(1+\mathsf{c}r/\rho\cdot w^{r/\rho-1}-V_{\mathsf{I}}'(w+\beta)\big)+\rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)w^{r/\rho-2}\text{ and}\tag{EC.83}$$

$$g_{\mathsf{I}}''(w)=\mathsf{c}r/\rho\cdot(r/\rho-1)w^{r/\rho-3}[\mu w+\rho\bar{w}(r/\rho-2)]-\mu V_{\mathsf{I}}''(w+\beta).\tag{EC.84}$$

Here, g_{I}'' may not exist at $\hat{\mathbf{w}}-\beta$ and $\underline{\vartheta}-\beta$. In this case, we follow the convention to use g_{I}'' to represent the *left*-second-order derivative of the function g_{I}'' at such a point. Similarly, we also use $\mathcal{V}_{\tilde{w}}'''(w)$ and $\mathcal{V}_{\tilde{w}}''''(w)$ to represent the left-third-order derivative and the left-fourth-order derivative of the function $\mathcal{V}_{\tilde{w}}$ at w (if needed) in the subsequent analysis.

Lemma EC.13 is equivalent to $g_{\mathsf{I}}(w)\geqslant 0$ for $w\in(0,\underline{\vartheta}]$. From (EC.79) at $\underline{\vartheta}$, we have $g_{\mathsf{I}}(\underline{\vartheta})\geqslant 0$. Moreover, the following holds.

Lemma EC.15. *We have $g_{\mathsf{I}}'(\underline{\vartheta})<0$.*

Proof. First, we consider the case that $\underline{\vartheta} \leqslant \breve{w}(\hat{\mathbf{w}})$. Evaluating (EC.54) at $\breve{w}(\hat{\mathbf{w}})$ (with $\hat{\mathbf{w}}$ replacing $\widetilde{w}$) yields

$$(\mu+r)V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \mu V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})+\beta) + \rho(\bar{w}-\breve{w}(\hat{\mathbf{w}}))V''_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) - \rho V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) + \rho - r = 0.$$

Clearly, $V''_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = 0$. Since $\underline{\vartheta} \leqslant \breve{w}(\hat{\mathbf{w}})$, we have $V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})) = \mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = \mathsf{V}'_{\mathsf{c}}(\underline{\vartheta}) = 1 + \mathsf{c}r/\rho \cdot \underline{\vartheta}^{r/\rho-1}$, which, together with the above expression, implies that

$$\begin{aligned} V'_{\mathsf{l}}(\underline{\vartheta}+\beta) &= \mathcal{V}'_{\hat{\mathbf{w}}}(\underline{\vartheta}+\beta) \geqslant \mathcal{V}'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})+\beta) = V'_{\hat{\mathbf{w}}}(\breve{w}(\hat{\mathbf{w}})+\beta) \\ &= \big[(\mu+r-\rho)(1+\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1}) + \rho - r\big]/\mu. \end{aligned}$$

In the above, the first inequality follows from the convexity of $\mathcal{V}_{\hat{\mathbf{w}}}$ and the fact that $\underline{\vartheta} \leqslant \breve{w}(\hat{\mathbf{w}})$. Substituting the above inequality into (EC.83) at $\underline{\vartheta}$ yields

$$g'_{\mathsf{l}}(\underline{\vartheta}) \leqslant (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} = (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-2}(\underline{\vartheta}-\bar{w}) < 0.$$

Next, we consider the case that $\underline{\vartheta} > \breve{w}(\hat{\mathbf{w}})$. Evaluating (EC.54) at $\underline{\vartheta}$ gives

$$(\mu+r)V'_{\hat{\mathbf{w}}}(\underline{\vartheta}) - \mu V'_{\hat{\mathbf{w}}}(\underline{\vartheta}+\beta) + \rho(\bar{w}-\underline{\vartheta})V''_{\hat{\mathbf{w}}}(\underline{\vartheta}) - \rho V'_{\hat{\mathbf{w}}}(\underline{\vartheta}) + \rho - r = 0.$$

Note that $V'_{\hat{\mathbf{w}}}(\underline{\vartheta}) = 1 + \mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1}$ and $V'_{\mathsf{l}}(\underline{\vartheta}+\beta) = V'_{\hat{\mathbf{w}}}(\underline{\vartheta}+\beta)$. Hence, we have

$$\begin{aligned} g'_{\mathsf{l}}(\underline{\vartheta}) &= \mu\big(1+\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} - V'_{\mathsf{l}}(\underline{\vartheta}+\beta)\big) + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &= (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} - \rho(\bar{w}-\underline{\vartheta})V''_{\hat{\mathbf{w}}}(\underline{\vartheta}) + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &< (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} - \frac{(\bar{w}-\underline{\vartheta})(\rho-r)(1-V'_{\hat{\mathbf{w}}}(\underline{\vartheta}))}{\underline{\vartheta}} + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &= (\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1} + \frac{(\bar{w}-\underline{\vartheta})(\rho-r)\mathsf{c}r/\rho\cdot\underline{\vartheta}^{r/\rho-1}}{\underline{\vartheta}} + \rho\bar{w}\cdot\mathsf{c}r/\rho\cdot(r/\rho-1)\underline{\vartheta}^{r/\rho-2} \\ &= 0, \end{aligned}$$

where the inequality follows from (EC.72) at $\underline{\vartheta}$ and $\hat{\mathbf{w}} = \widetilde{w}(\underline{\vartheta})$. ◻

Next, we show the following crucial result.

Lemma EC.16. *For any $w \in (0, \underline{\vartheta})$, we have either $g'_{\mathsf{l}}(w) \leqslant 0$ or $g''_{\mathsf{l}}(w) \geqslant 0$.*

The above result, combining with Lemma EC.15, yields that $g'_{\mathsf{l}}(w) \leqslant 0$ for any $w \in (0, \underline{\vartheta}]$, which immediately concludes the result stated in Lemma EC.13. In fact, if it fails to hold, $w^\dagger := \sup\{w \in (0,\underline{\vartheta}) \mid g'_{\mathsf{l}}(w) > 0\}$ is well defined, which further implies that $g''_{\mathsf{l}}(w^\dagger) < 0$. This contradicts Lemma EC.16.

Lemma EC.16 follows immediately from Lemmas EC.17–EC.20 below.

LEMMA EC.17. *For any* $w \in [0, \bar{\vartheta} - \beta]$, *we have* $g_{\mathsf{l}}'(w) < 0$.

LEMMA EC.18. *For any* $w \in (\bar{\vartheta} - \beta, \underline{\vartheta})$ *such that* $\mathcal{V}_{\hat{\mathbf{w}}}''''(w + \beta) \leqslant 0$, *we have* $g_{\mathsf{l}}''(w) > 0$.

LEMMA EC.19. *For any* $w \in \big(0, (2 - r/\rho)\beta \wedge \underline{\vartheta}\big)$, *we have* $g_{\mathsf{l}}''(w) \geqslant 0$.

LEMMA EC.20. *For any* $w \in [(1 - r/\rho)\beta, \underline{\vartheta})$ *such that* $\mathcal{V}_{\hat{\mathbf{w}}}''''(w + \beta) > 0$, *we have* $g_{\mathsf{l}}'(w) \leqslant 0$.

In the proofs of Lemmas EC.18 and EC.20, we also need the following technical result.

LEMMA EC.21. *For any* $\tilde{w} \in [0, \bar{w})$, *the following results hold:*

(i) *If* $2\rho < r + \mu$, *then there exists a number* $\varsigma \in [\breve{w}(\tilde{w}), \tilde{w})$, *such that* $\mathcal{V}_{\tilde{w}}'''' > 0$ *on* $(\varsigma, \tilde{w})$ *and* $\mathcal{V}_{\tilde{w}}'''' < 0$ *on* $(\breve{w}(\tilde{w}), \varsigma]$;

(ii) *Otherwise,* $\mathcal{V}_{\tilde{w}}'''' \leqslant 0$ *on* $(\breve{w}(\tilde{w}), \tilde{w})$.

The remaining part of this subsection is devoted to the proofs of Lemmas EC.17–EC.21. To proceed, we need some preliminary results of V_{c}. Using the explicit expression of V_{c} in (24), we obtain that V_{c} is strictly concave on $\mathbb{R}_{++}$, i.e., $\mathsf{V}_{\mathsf{c}}'' < 0$ on $\mathbb{R}_{++}$,

$$\mathsf{V}_{\mathsf{c}}'''(w) = \mathsf{c} r/\rho \cdot (r/\rho - 1)(r/\rho - 2) w^{r/\rho - 3} > 0 \quad \text{and} \tag{EC.85}$$

$$\mathsf{V}_{\mathsf{c}}''''(w) = \mathsf{c} r/\rho \cdot (r/\rho - 1)(r/\rho - 2)(r/\rho - 3) w^{r/\rho - 4} < 0 \tag{EC.86}$$

for all $w \in \mathbb{R}_{++}$. Hence, V_{c}' is strictly convex and $\mathsf{V}_{\mathsf{c}}''$ is strictly concave, which further implies that

$$\big(\mathsf{V}_{\mathsf{c}}'(w) - \mathsf{V}_{\mathsf{c}}'(w + \beta)\big) + \beta \mathsf{V}_{\mathsf{c}}''(w) < 0 \text{ and} \tag{EC.87}$$

$$\big(\mathsf{V}_{\mathsf{c}}''(w) - \mathsf{V}_{\mathsf{c}}''(w + \beta)\big) + \beta \mathsf{V}_{\mathsf{c}}'''(w) > 0 \tag{EC.88}$$

for any $w \in \mathbb{R}_{++}$.

Proof of Lemma EC.17. Using $V_{\mathsf{l}}(w) = \mathsf{V}_{\mathsf{c}}(w)$ for $w \in [0, \underline{\vartheta}]$, we have

$$g_{\mathsf{l}}'(w) = \mu\big(V_{\mathsf{l}}'(w) - V_{\mathsf{l}}'(w + \beta)\big) + \mu\beta \mathsf{V}_{\mathsf{c}}''(w) \leqslant \mu\big(\mathsf{V}_{\mathsf{c}}'(w) - \mathsf{V}_{\mathsf{c}}'(w + \beta)\big) + \mu\beta \mathsf{V}_{\mathsf{c}}''(w) < 0,$$

where the first inequality uses $V_{\mathsf{l}}'(w + \beta) \geqslant \mathsf{V}_{\mathsf{c}}'(w + \beta)$ because of $w + \beta \leqslant \bar{\vartheta}$ and (EC.74) with $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$, and the second inequality follows from (EC.87). ▫

Proof of Lemma EC.18. Define $\phi(w) := \mathcal{V}_{\hat{\mathbf{w}}}'(w) - \mathsf{V}_{\mathsf{c}}'(w)$. Since $\bar{\vartheta} = \inf\{w > \underline{\vartheta} \mid \phi(w) = 0\}$ and $\phi > 0$ over $(\underline{\vartheta}, \bar{\vartheta})$, we have $\phi'(\bar{\vartheta}) \leqslant 0$. It follows from $\mathcal{V}_{\hat{\mathbf{w}}}''''(w + \beta) \leqslant 0$ and Lemma EC.21 with $\tilde{w}$ replaced by $\hat{\mathbf{w}}$ that $\mathcal{V}_{\hat{\mathbf{w}}}'''' < 0$ on $(\breve{w}(\hat{\mathbf{w}}), w + \beta)$.

For any $w \in (\bar{\vartheta} - \beta, \underline{\vartheta})$, we have

$$\begin{aligned}
g_{\mathsf{l}}''(w) &= \mu\big(\mathsf{V}_{\mathsf{c}}''(w) - V_{\mathsf{l}}''(w+\beta)\big) + \mu\beta\mathsf{V}_{\mathsf{c}}'''(w) \\
&= \mu\big(\mathsf{V}_{\mathsf{c}}''(w) - \mathcal{V}_{\hat{\mathbf{w}}}''(w+\beta)\big) + \mu\beta\mathsf{V}_{\mathsf{c}}'''(w) \\
&= \mu\big[(\mathsf{V}_{\mathsf{c}}''(w) - \mathsf{V}_{\mathsf{c}}''(w+\beta)) + \beta\mathsf{V}_{\mathsf{c}}'''(w)\big] + \mu\big(\mathsf{V}_{\mathsf{c}}''(w+\beta) - \mathcal{V}_{\hat{\mathbf{w}}}''(w+\beta)\big) \\
&> \mu\big(\mathsf{V}_{\mathsf{c}}''(w+\beta) - \mathcal{V}_{\hat{\mathbf{w}}}''(w+\beta)\big) \\
&= \mu\big(\mathsf{V}_{\mathsf{c}}''(\bar{\vartheta}) - \mathcal{V}_{\hat{\mathbf{w}}}''(\bar{\vartheta})\big) + \mu\int_{\bar{\vartheta}}^{w+\beta}\big(\mathsf{V}_{\mathsf{c}}'''(y) - \mathcal{V}_{\hat{\mathbf{w}}}'''(y)\big)\mathrm{d}y \\
&> 0,
\end{aligned}$$

where the first inequality uses (EC.88), and the last inequality follows from $\mathsf{V}_{\mathsf{c}}''(\bar{\vartheta}) - \mathcal{V}_{\hat{\mathbf{w}}}''(\bar{\vartheta}) = -\phi'(\bar{\vartheta}) \geqslant 0$, $w + \beta > \bar{\vartheta}$, $\mathsf{V}_{\mathsf{c}}''' > 0$ (see (EC.85)) and $\mathcal{V}_{\hat{\mathbf{w}}}''' < 0$ on $(\breve{w}(\hat{\mathbf{w}}), w+\beta)$. ▫

Proof of Lemma EC.19. According to (EC.84), we have

$$\begin{aligned}
g_{\mathsf{l}}''(w) &= \mathsf{c}r/\rho\cdot(r/\rho - 1)w^{r/\rho-3}\mu[w + \beta(r/\rho - 2)] - \mu V_{\mathsf{l}}''(w+\beta) \\
&\geqslant -\mu V_{\mathsf{l}}''(w+\beta) \geqslant 0,
\end{aligned}$$

where the first inequality follows from $w \leqslant (2 - r/\rho)\beta$ and the last inequality follows from the concavity of V_{l}. ▫

Proof of Lemma EC.20. Suppose, to the contradictory, that there exists a $w^{\dagger} \in [(1-r/\rho)\beta, \underline{\vartheta})$ such that $\mathcal{V}_{\hat{\mathbf{w}}}'''(w^{\dagger}+\beta) > 0$ and $g_{\mathsf{l}}'(w^{\dagger}) > 0$. According to Lemma EC.15, there must exist a number $w^{\ddagger} \in (w^{\dagger}, \underline{\vartheta})$ such that

$$g_{\mathsf{l}}'(w^{\ddagger}) = 0 \text{ and } g_{\mathsf{l}}''(w^{\ddagger}) \leqslant 0. \tag{EC.89}$$

First, we claim that $w^{\ddagger} + \beta < \hat{\mathbf{w}}$. Otherwise, we have $V_{\mathsf{l}}'(w^{\ddagger}+\beta) = V_{\hat{\mathbf{w}}}'(w^{\ddagger}+\beta) = 0$ and thus

$$g_{\mathsf{l}}'(w^{\ddagger}) = \mu + \mathsf{c}r/\rho\cdot(w^{\ddagger})^{r/\rho-2}\mu\big[w^{\ddagger} - \beta(1-r/\rho)\big] > \mu > 0,$$

where the first inequality holds as $w^{\ddagger} > w^{\dagger} \geqslant (1-r/\rho)\beta$, leading to a contradiction.

Hence, by Lemma EC.21, we have $\mathcal{V}_{\hat{\mathbf{w}}}'''(w^{\ddagger}+\beta) > 0$. Furthermore, evaluating (EC.54) at $w^{\ddagger}+\beta$ (with $\hat{\mathbf{w}}$ replacing $\widetilde{w}$) gives

$$(\mu + r)\mathcal{V}_{\hat{\mathbf{w}}}'(w^{\ddagger}+\beta) - \mu\mathcal{V}_{\hat{\mathbf{w}}}'(w^{\ddagger}+2\beta) + \rho(\bar{w} - w^{\ddagger} - \beta)\mathcal{V}_{\hat{\mathbf{w}}}''(w^{\ddagger}+\beta) - \rho\mathcal{V}_{\hat{\mathbf{w}}}'(w^{\ddagger}+\beta) + \rho - r = 0.$$

Since $\mathcal{V}_{\hat{\mathbf{w}}}'''(w^{\ddagger}+\beta) > 0$, we have $\mathcal{V}_{\hat{\mathbf{w}}}''' > 0$ on $[w^{\ddagger}+\beta, \hat{\mathbf{w}})$ by Lemma EC.21. That is, $\mathcal{V}_{\hat{\mathbf{w}}}'$ is strictly convex on $[w^{\ddagger}+\beta, \hat{\mathbf{w}})$, which, together with the concavity of $\mathcal{V}_{\hat{\mathbf{w}}}$, yields that

$$\begin{aligned}
\mathcal{V}_{\hat{\mathbf{w}}}'(w^{\ddagger}+2\beta) - \mathcal{V}_{\hat{\mathbf{w}}}'(w^{\ddagger}+\beta) &= \mathcal{V}_{\hat{\mathbf{w}}}'\big((w^{\ddagger}+2\beta)\wedge\hat{\mathbf{w}}\big) - \mathcal{V}_{\hat{\mathbf{w}}}'(w^{\ddagger}+\beta) \\
&> \big(\beta\wedge(\hat{\mathbf{w}} - w^{\ddagger} + \beta)\big)\cdot\mathcal{V}_{\hat{\mathbf{w}}}''(w^{\ddagger}+\beta) \geqslant \beta\mathcal{V}_{\hat{\mathbf{w}}}''(w^{\ddagger}+\beta).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}\rho(\bar{w}-w^{\ddagger}-\beta)\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta) &= (\rho-r)\big(\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)-1\big)+\mu\big(\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+2\beta)-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\big)\\ &> (\rho-r)\big(\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)-1\big)+\mu\beta\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta),\end{aligned}$$

which, along with $\rho\bar{w}=\mu\beta$, can be rewritten as

$$(\rho-r)\big(1-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\big)>\rho(w^{\ddagger}+\beta)\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta). \tag{EC.90}$$

Since $g'_{\mathsf{l}}(w^{\ddagger})=0$, using (EC.83) we have

$$1-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)=-\mathsf{c}r/\rho\cdot(w^{\ddagger})^{r/\rho-2}\left[w^{\ddagger}-(1-r/\rho)\beta\right]. \tag{EC.91}$$

Evaluating (EC.84) at $w^{\ddagger}$ yields

$$\begin{aligned}g''_{\mathsf{l}}(w^{\ddagger})/\mu &= \mathsf{c}r/\rho\cdot(r/\rho-1)(w^{\ddagger})^{r/\rho-3}[w^{\ddagger}+\beta(r/\rho-2)]-\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\\ &= \frac{(\rho-r)(1-\mathcal{V}'_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta))}{w^{\ddagger}-(1-r/\rho)\beta}\cdot\frac{w^{\ddagger}+\beta(r/\rho-2)}{\rho w^{\ddagger}}-\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\\ &> \left[\frac{\rho(w^{\ddagger}+\beta)[w^{\ddagger}+\beta(r/\rho-2)]}{(w^{\ddagger}-(1-r/\rho)\beta)\cdot\rho w^{\ddagger}}-1\right]\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)\\ &= -\frac{(2\rho-r)\beta^2}{(w^{\ddagger}-(1-r/\rho)\beta)\cdot\rho w^{\ddagger}}\mathcal{V}''_{\hat{\mathbf{w}}}(w^{\ddagger}+\beta)>0,\end{aligned}$$

where the second equality follows from (EC.91), and the first inequality follows from (EC.90) and uses the fact that $w^{\ddagger}>w^{\dagger}\geqslant(1-r/\rho)\beta$. This reaches a contradiction with (EC.89). ▫

Proof of Lemma EC.21. According to (EC.52) and the definition of $\mathcal{V}_{\tilde{w}}$, we have

$$\mathcal{V}'''_{\tilde{w}}(w)=-\frac{(\rho-r)(2\rho-r-\mu)}{\rho^2}(\bar{w}-\tilde{w})^{\frac{\rho-r-\mu}{\rho}}(\bar{w}-w)^{\frac{-3\rho+r+\mu}{\rho}}=:\zeta(w) \tag{EC.92}$$

for $w\in\big((\tilde{w}-\beta)^+\vee\breve{w}(\tilde{w}),\tilde{w}\big)$.

In the case of $(\tilde{w}-\beta)^+\leqslant\breve{w}(\tilde{w})$, then the result stated in the lemma holds, with $\varsigma=\breve{w}(\tilde{w})$ if $2\rho<r+\mu$. Below, we consider the case of $(\tilde{w}-\beta)^+>\breve{w}(\tilde{w})$, or equivalently, $\tilde{w}-\beta>\breve{w}(\tilde{w})$.

Besides, for $w\in(\breve{w}(\tilde{w}),\tilde{w})$, we have

$$\rho(\bar{w}-w)\mathcal{V}'''_{\tilde{w}}(w)=\mu\big(\mathcal{V}''_{\tilde{w}}(w+\beta)-\mathcal{V}''_{\tilde{w}}(w)\big)+(2\rho-r)\mathcal{V}''_{\tilde{w}}(w), \tag{EC.93}$$

and thus

$$\rho(\bar{w}-w)\mathcal{V}''''_{\tilde{w}}(w)=\mu\big(\mathcal{V}'''_{\tilde{w}}(w+\beta)-\mathcal{V}'''_{\tilde{w}}(w)\big)+(3\rho-r)\mathcal{V}'''_{\tilde{w}}(w). \tag{EC.94}$$

By Lemma EC.14, $\mathcal{V}'''_{\tilde{w}}$ may not exist at $\tilde{w}-\beta$. In fact, by (EC.92), we have $\mathcal{V}'''_{\tilde{w}}((\tilde{w}-\beta)+)=\zeta(\tilde{w}-\beta)$. Evaluating (EC.93) at $(\tilde{w}-\beta)-$ (to be precise, we consider an increasing sequence of

$\{w_n\}_{n\in\mathbb{N}}$ near $\widetilde{w}-\beta$ which tends to $\widetilde{w}-\beta$ from below, evaluate (EC.93) at these w_n's and then let $n\to\infty$), we obtain

$$\rho(\bar{w}-\widetilde{w}+\beta)\mathcal{V}_{\widetilde{w}}'''(\widetilde{w}-\beta)=\mu\big(\mathcal{V}_{\widetilde{w}}''(\widetilde{w}-)-\mathcal{V}_{\widetilde{w}}''(\widetilde{w}-\beta)\big)+(2\rho-r)\mathcal{V}_{\widetilde{w}}''(\widetilde{w}-\beta).$$

In the above, as mentioned earlier, we adopt the convention to use $\mathcal{V}_{\widetilde{w}}'''(\widetilde{w}-\beta)$ to denote the left-third-order derivative of function $\mathcal{V}_{\widetilde{w}}$ at $\widetilde{w}-\beta$.

In a similar vein, evaluating (EC.93) at $(\widetilde{w}-\beta)+$ yields

$$\rho(\bar{w}-\widetilde{w}+\beta)\mathcal{V}_{\widetilde{w}}'''\big((\widetilde{w}-\beta)+\big)=\mu\big(\mathcal{V}_{\widetilde{w}}''(\widetilde{w}+)-\mathcal{V}_{\widetilde{w}}''(\widetilde{w}-\beta)\big)+(2\rho-r)\mathcal{V}_{\widetilde{w}}''(\widetilde{w}-\beta).$$

Combining the above two equations and using $\mathcal{V}_{\widetilde{w}}''(\widetilde{w}+)=0$, we have

$$\begin{aligned}\mathcal{V}_{\widetilde{w}}'''(\widetilde{w}-\beta)&=\mathcal{V}_{\widetilde{w}}'''\big((\widetilde{w}-\beta)+\big)+\frac{\mu\mathcal{V}_{\widetilde{w}}''(\widetilde{w}-)}{\rho(\bar{w}-\widetilde{w}+\beta)}\\&=\zeta(\widetilde{w}-\beta)-\frac{(\rho-r)\mu}{\rho^2(\bar{w}-\widetilde{w}+\beta)(\bar{w}-\widetilde{w})}<\zeta(\widetilde{w}-\beta).\end{aligned}\tag{EC.95}$$

We break the proof of the lemma into two cases.

(i) Suppose that $2\rho<r+\mu$, in which case (EC.92) implies that $\mathcal{V}_{\widetilde{w}}'''>0$ on $(\widetilde{w}-\beta,\widetilde{w})$. From (EC.95), $\mathcal{V}_{\widetilde{w}}'''(\widetilde{w}-\beta)$ may not be larger than 0, and thus we consider the following two cases.

Case 1: $\mathcal{V}_{\widetilde{w}}'''(\widetilde{w}-\beta)>0$. Let $w_1:=\sup\{w\in(\breve{w}(\widetilde{w}),\widetilde{w}-\beta)\mid\mathcal{V}_{\widetilde{w}}'''(w)\leqslant 0\}$. If the set is empty, we have $\mathcal{V}_{\widetilde{w}}'''>0$ on $(\breve{w}(\widetilde{w}),\widetilde{w})$. Therefore, the result stated in (i) is obtained by letting $\varsigma=\breve{w}(\widetilde{w})$.

If the above set is nonempty, then $\breve{w}(\widetilde{w})<w_1<\widetilde{w}-\beta$. Since $\mathcal{V}_{\widetilde{w}}\in C^3((\breve{w}(\widetilde{w}),\widetilde{w}-\beta))$, we have $\mathcal{V}_{\widetilde{w}}'''$ exists on $(\breve{w}(\widetilde{w}),\widetilde{w}-\beta)$. In addition, $\mathcal{V}_{\widetilde{w}}'''(w_1)=0$ and $\mathcal{V}_{\widetilde{w}}'''>0$ on $(w_1,\widetilde{w})$. Now, we prove that

$$\mathcal{V}_{\widetilde{w}}'''<0 \text{ on } (\breve{w}(\widetilde{w}),w_1).\tag{EC.96}$$

Evaluating (EC.94) at w_1 and using $\mathcal{V}_{\widetilde{w}}'''(w_1)=0$, we obtain $\rho(\bar{w}-w_1)\mathcal{V}_{\widetilde{w}}''''(w_1)=\mu\mathcal{V}_{\widetilde{w}}'''(w_1+\beta)>0$ (here if $w_1=\widetilde{w}-2\beta$, the left derivatives are used), which implies that $\mathcal{V}_{\widetilde{w}}'''<0$ on $(w_1-\epsilon,w_1)$ for some $\epsilon>0$.

If (EC.96) fails to hold, then $w_2:=\sup\{w\in(\breve{w}(\widetilde{w}),w_1)\mid\mathcal{V}_{\widetilde{w}}'''(w)\geqslant 0\}$ is well defined and $w_2\in(\breve{w}(\widetilde{w}),w_1)$. Hence, we have $\mathcal{V}_{\widetilde{w}}'''(w_2)=0$ and $\mathcal{V}_{\widetilde{w}}''''(w_2+)<0$. Evaluating (EC.94) at w_2+ yields $\rho(\bar{w}-w_2)\mathcal{V}_{\widetilde{w}}''''(w_2+)=\mu\mathcal{V}_{\widetilde{w}}'''\left((w_2+\beta)+\right)$, implying that $\mathcal{V}_{\widetilde{w}}'''\left((w_2+\beta)+\right)<0$. If $w_2=\widetilde{w}-2\beta$, then $\mathcal{V}_{\widetilde{w}}'''\left((w_2+\beta)+\right)=\zeta(\widetilde{w}-\beta)>0$, a contradiction. Otherwise, $\mathcal{V}_{\widetilde{w}}'''$ exists at $w_2+\beta$, and thus $\mathcal{V}_{\widetilde{w}}'''(w_2+\beta)<0$. By the definition of w_1, we have $w_2+\beta<w_1$. Hence, $\mathcal{V}_{\widetilde{w}}'''<0$ on $(w_2,w_2+\beta]$ and thus $\mathcal{V}_{\widetilde{w}}''(w_2+\beta)<\mathcal{V}_{\widetilde{w}}''(w_2)$. Evaluating (EC.93) at w_2, together with $\mathcal{V}_{\widetilde{w}}'''(w_2)=0$, yields

$$\mu\mathcal{V}_{\widetilde{w}}''(w_2+\beta)=(\mu+r-2\rho)\mathcal{V}_{\widetilde{w}}''(w_2)<\mu\mathcal{V}_{\widetilde{w}}''(w_2),$$

which gives $\mathcal{V}''_{\tilde{w}}(w_2) > 0$. This makes a contradiction with the concavity of $\mathcal{V}_{\tilde{w}}$. Hence, (EC.96) holds, indicating that the result stated in (i) is obtained by letting $\varsigma = w_1$.

Case 2: $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) \leqslant 0$. We show that

$$\mathcal{V}'''_{\tilde{w}} < 0 \text{ on } (\breve{w}(\tilde{w}), \tilde{w}-\beta). \tag{EC.97}$$

If $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) = 0$, then by evaluating (EC.94) at $(\tilde{w}-\beta)-$ and using $\mathcal{V}'''_{\tilde{w}}(\tilde{w}) = \zeta(\tilde{w}) > 0$, we obtain that $\mathcal{V}''''_{\tilde{w}}(\tilde{w}-\beta) > 0$. (Again, the left derivatives are used.) Hence, we have $\mathcal{V}'''_{\tilde{w}} < 0$ on $(\tilde{w}-\beta-\epsilon, \tilde{w}-\beta)$ for some $\epsilon > 0$. If, on the other side, $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) < 0$, then the assertion that $\mathcal{V}'''_{\tilde{w}} < 0$ on $(\tilde{w}-\beta-\epsilon, \tilde{w}-\beta)$ for some $\epsilon > 0$ still holds by Lemma EC.14.

If (EC.97) fails to hold, then $w_3 := \sup\{w \in [\breve{w}(\tilde{w}), \tilde{w}-\beta) \mid \mathcal{V}'''_{\tilde{w}}(w) \geqslant 0\}$ is well defined, satisfying $w_3 \in [\breve{w}(\tilde{w}), \tilde{w}-\beta)$. Moreover, we have $\mathcal{V}'''_{\tilde{w}}(w_3) = 0$ and $\mathcal{V}''''_{\tilde{w}}(w_3+) < 0$. With exactly the same argument as that in the previous case for treating w_2, a contradiction can be reached. Therefore, (EC.97) holds. The result stated in (i) is obtained by using (EC.92) and letting $\varsigma = \tilde{w}-\beta$.

(ii) Next, we turn to study the case that $2\rho \geqslant r+\mu$. In this case, (EC.92) implies that $\mathcal{V}'''_{\tilde{w}} \leqslant 0$ over $(\tilde{w}-\beta, \tilde{w})$. Hence, we must have $\mathcal{V}'''_{\tilde{w}}(\tilde{w}-\beta) < 0$ in view of (EC.95). Therefore, the argument for the second case above is valid, which leads us to the desired result. ▫

EC.4. Proofs of the Results in Sections 5

EC.4.1. Proof of Proposition 8

We only consider the case in which both Condition 1 and $K < \bar{K}_1$ hold, since the case in which both Condition 2 and $K < \underline{K}$ hold can be treated similarly. It follows from Lemma EC.7 that $\psi(\underline{\vartheta}) = K$ with ψ being decreasing on $(0, \underline{\theta}^0)$. Therefore, $\underline{\vartheta}$ is decreasing in K. Recalling $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ and using Lemma EC.6(ii), we obtain that $\bar{\vartheta} = \bar{\theta}(\underline{\vartheta})$ is increasing in K.

For the last assertion, we first note that under Condition 1 and $\bar{K}_1 > 0$, $\lim_{\underline{\theta}\uparrow\underline{\theta}^0} \psi(\underline{\theta}) = 0$ by Lemma EC.7, which implies $\lim_{K\downarrow 0} \underline{\vartheta} = \underline{\theta}^0$. Then, using Lemma EC.6(iii), we obtain $\lim_{K\downarrow 0} \bar{\vartheta} = \underline{\theta}^0$. Under Condition 2 and $\underline{K} > 0$, we also have $\lim_{K\downarrow 0} \underline{\vartheta} = \underline{\theta}^0 = \lim_{K\downarrow 0} \bar{\vartheta}$, by a similar argument and Lemmas EC.9 and EC.11. Hence, the desired result holds with $\theta_0 = \underline{\theta}^0$.

EC.4.2. Proof of Theorem 6

First, we consider the case in which Condition 1 and $\bar{K}_1 > 0$ hold. It follows from Proposition 8 that $\lim_{K\downarrow 0} \underline{\vartheta} = \lim_{K\downarrow 0} \bar{\vartheta} = \underline{\theta}^0$. Recall from the proof of Proposition 6 that $\hat{\mathbf{w}} = \tilde{w}(\underline{\vartheta})$ and $\mathsf{c} = C(\underline{\vartheta})$. Hence, we have $\lim_{K\downarrow 0} \hat{\mathbf{w}} = \lim_{K\downarrow 0} \tilde{w}(\underline{\vartheta}) = \tilde{w}(\underline{\theta}^0)$ and $\lim_{K\downarrow 0} \mathsf{c} = \lim_{K\downarrow 0} C(\underline{\vartheta}) = C(\underline{\theta}^0)$. The result (50) is obtained by setting $\theta_0 = \underline{\theta}^0$, $\hat{\mathbf{w}}_0 = \tilde{w}(\underline{\theta}^0)$, and $\mathsf{c}_0 = C(\underline{\theta}^0)$. Moreover, $\mathcal{V}_{\hat{\mathbf{w}}}$ and V_{c} converge uniformly to $\mathcal{V}_{\hat{\mathbf{w}}_0}$ and $\mathsf{V}_{\mathsf{c}_0}$, respectively, as K approaches 0. Consequently, both value functions as defined in (46) converge to $\mathfrak{V}_{\theta_0}$ uniformly as K approaches 0. Using Proposition 7 and sending K

to zero, we conclude that functions $V_{\mathsf{I}} = V_{\emptyset} = \mathfrak{V}_{\theta_0}$ satisfy the optimality conditions (20)–(22) for $K = 0$.

The argument for the case in which Condition 2 and $\underline{K} > 0$ hold is exactly the same, and thus is omitted.

References

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

Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer Series in Statistics. Springer.

Cao, P., F. Tian, and P. Sun (2022). Comment on "optimal contract to induce continued effort". *Management Science 68*(1), 796–808.

Protter, P. E. (2003). *Stochastic Integration and Differential Equations* (second edition ed.), Volume 21 of *Stochastic Modelling and Applied Probability*. Springer.

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