Optimal Long-term Contracting with Learning*

Zhiguo He† Bin Wei‡ Jianfeng Yu§

First draft: January 2012

Preliminary and Reference Incomplete

Abstract

This paper introduces profitability uncertainty into an infinite-horizon variation of the classic Holmstrom and Milgrom (1987) model, and studies optimal dynamic contracting with endogenous learning. The agent's potential belief manipulation leads to the hidden information problem, which makes incentive provisions intertemporally linked in the optimal contract. We reduce the contracting problem into a dynamic programming problem with one state variable, and characterize the optimal contract with an ordinary differential equation. In the benchmark case of Holmstrom and Milgrom (1987) without learning, the optimal effort is constant, and the optimal contract is linear. In contrast, in our model with endogenous learning, the optimal effort policy becomes history dependent, and decreases over time on average. Moreover, we show that the optimal contract exhibits an option-like feature in that the incentives rise after good performance shocks.

Keywords: Executive Compensation, Moral Hazard, Bayesian Learning, Hidden Information, Belief Manipulation, Private Savings, Continuous-Time Optimal Contracting.

*This work does not necessarily reflect the views of the Federal Reserve System or its staff. All errors are our own. Jianfeng Yu gratefully acknowledges financial support from the Dean's Small Research Grant from the Carlson School of Management at the University of Minnesota.

†University of Chicago, Booth School of Business, 5807 South Woodlawn Ave., Chicago 60637, Phone: (1)773-834-3769, Email: zhiguo.he@chicagobooth.edu.

‡Board of Governors of the Federal Reserve System, Washington, DC, 20551, Phone: (202)-452-2693; Fax: (202)-728-5887; Email: bin.wei@frb.gov. Also the City University of New York, Zicklin School of Business, Baruch College, 55 Lexington Avenue, New York, NY 10010. Phone: (1)646-312-3469, Email: bin.wei@baruch.cuny.edu.

§University of Minnesota, Carlson School of Management, CSOM 3-122, 321 19th Avenue South, Minneapolis, MN 55455. Phone: (1)612-625-5498. Email: jianfeng@umn.edu.

1 Introduction

Many long-term contractual relationships feature uncertainty and learning. Uncertainy may arise, for instance, when either the project's quality or the agent's ability is unknown in long-term employment contracts. By introducing uncertainty into an infinite-horizon variation of the classic Holmstrom and Milgrom (1987) model, this paper studies optimal long-term contracting with endogenous learning.

Without uncertainty, most existing dynamic principal-agent models study repeated stage agency games with transitory output performance, i.e., the agent's short-term effort plus some noises.[1] In these models, the history dependence of optimal contracting is fully summarized by the agent's continuation payoff, while incentive provisions are essentially independent across periods. Once we allow the output to be also affected by an uncertain profitability component, both the principal and the agent need to learn this profitability component based on the realized path of output performance. The dynamic learning renders an interesting link between incentive provisions across different times over the course of the contractual relationship.

The intertemporally linked incentive provisions are rooted in the following *persistent* "*hidden information*" problem. Although both parties start with a common prior, and thus along the equilibrium path the principal knows as much as the agent does, on off equilibrium paths the agent may enjoy information that is hidden from the principal. This is because only the agent knows his actual effort (i.e., "*hidden action*"), while the principal only knows the recommended effort. Imagine that the agent has followed the recommended effort policy in the past so that both parties share the same belief today. Suppose that the agent shirks today by choosing some lower-than-recommended effort, which lowers today's output on average. However, the principal who anticipates higher effort would mistakenly attribute today's output drop to a lower value of the project's profitability. Thus by shirking today the agent can downward distort the principal's inference about the project's profitability from today on, which is long-lasting (i.e., persistent hidden information). We call this

[1] See, for example, Holmstrom and Milgrom (1987), Spear and Srivastava (1987), Sannikov (2008), etc.

belief manipulation effect. As a result, the agent will be mistakenly rewarded in the future for better realized performance, simply because the agent knows that the project is more profitable than the principal believes. This gives rise to the agent's future information rent, and this future information rent increases with future pay-performance sensitivity. Because this future information rent tarnishes the agent's current working incentives, incentive provisions at all dates are interlinked through this belief-manipulation effect.

It is theoretically challenging to characterize the optimal contract in such a dynamic agency model of both "hidden action" (i.e., effort) and "hidden information" (i.e., potentially diverging beliefs about the project's profitability).[2] Given any compensation contract, we first solve the agent's problem by characterizing the optimality condition that is similar to Prat and Jovanovic (2010). As explained, the belief manipulation effect implies that the agent enjoys information rents, which can be conveniently expressed as a properly discounted future pay-performance sensitivities. The higher the future incentives, the greater the information rents, the lower the agent's current working incentives.

We then solve the principal's problem for the optimal compensation contract and the associated optimal effort policy. Given the stationary nature of learning, there is no time effect, and we reformulate the principal's problem into a dynamic programming problem with two state variables: the agent's continuation payoff and the agent's information rent. Further, the agent's ability to privately save combined with CARA preference (i.e., there is no wealth effect) allow us to separate the agent's continuation payoff from the principal's problem. We thus end up with a unidimensional state variable, which is the agent's information rent. The optimal contract is characterized by solving an ordinary differential equation (ODE) in the spirit of Hamiltanian-Jacobi-Bellman (HJB) equation.

Relative to Prat and Jovanovic (2010) and DeMarzo and Sannikov (2010) who focus on the

[2] Because of the endogenous persistent hidden information, our model differs dramatically from standard *"hidden action"* dynamic agency models where the agent's unobservable shirking has only a *short-lived* effect. In the pure "hidden action" models from each period onwards the principal knows everything (that is payoff-relevant, including belief) regarding the agent, which allows for the simple recursive formulation with the agent's continuation payoff being the only state variable. See, e.g., Spear and Srivastava (1987), and Sannikov (2008).

contract that implements a constant first-best level effort, we make a contribution to the literature by solving for the optimal effort path endogenously jointly with the optimal compensation contract. Recall that in the Holmstrom and Milgrom (1987) benchmark without learning, the optimal effort policy is a constant process. In our model with learning, the optimal effort policy exhibits a *time-decreasing* pattern. Moreover, the optimal effort policy is stochastic, and incentives in the optimal contract rise after good performance shocks, exhibiting an *option-like* feature.

First, the optimal effort policy exhibits a negative drift and thus a time-decreasing pattern. In fact, we solve the deterministic optimal contract (i.e., focusing on deterministic effort policies only) in closed form, and show analytically that the optimal deterministic effort policy is decreasing with the agent's tenure, i.e., the front-loaded effort policy is optimal. This front-loaded pattern comes from the belief-manipulation effect. As mentioned, future incentives increase agent's current information rent for shirking. In other words, future pay-performance sensitivities hurt the agent's working incentives in earlier periods, but not the other way around. Thus, the optimal contract should implement a lower effort in later periods.[3] The force of information rent due to the belief manipulation effect is general and should apply to any dynamic contracting model with endogenous learning.

The second result of effort policy being history-dependent is more surprising. With a standard CARA-Normal setting and stationary learning, the Pareto frontier of the contractual relationship fluctuates with the project profitability. However, without learning but with long-term contracting, the Holmstrom and Milgrom (1987) benchmark shows that a constant effort policy is optimal. Moreover, with learning but without long-term contracting, Holmstrom (1999) derives a deterministic effort policy along the equilibrium path. These two results imply that the history-dependent effort policy is an outcome of the combination of long-term contracting and learning.

In our model, the optimal contract has an option-like feature in that pay-for-performance rises following good performance. The intuition comes from reducing the agent's belief manipulation in

[3]This implication is similar to models with career concerns Gibbons and Murphy (1992), but through a distinct mechanism.

a long-term relationship. As explained, the agent's current shirking, by downward distorting the principal's belief persistently, leads to information rent in the future. For a risk-averse agent, the amount of information rents not only depends on future pay-performance sensitivities, but also the agent's marginal utilities at these future states. Raising incentives after good performance introduces a negative correlation between incentive slope and marginal utility, i.e., allocating greater belief manipulation benefits in states where the agent cares less. Hence, the option-like compensation contract lowers the agent's information rent standing today.

The topic of optimal contracting with endogenous learning belongs to the recent literature studying optimal long-term contract with hidden information (Williams, 2009, 2011; Sannikov, 2007; Zhang, 2009). As mentioned earlier, our paper is closet to Prat and Jovanovic (2010) and DeMarzo and Sannikov (2010); both restricting attention to the optimal contract that implements a constant first-best level of effort. In contrast, we solve for the optimal effort policy endogenously joint with the optimal long-term compensation contract, and emphasize the general economic mechanisms that shape the optimal effort policy in long-term optimal contracting. Learning-wise, DeMarzo and Sannikov (2010), like ours, consider a framework with stationary learning, while Prat and Jovanovic (2010) study "parameter uncertainty" which implies that the project profitability is an unknown constant rather than a moving target. Theoretically speaking, the exact learning technology is less important, as both learning technology lead to long-lasting belief manipulation effect which is the major challenge in solving for the optimal contract.[4,5]

The rest of the paper is organized as follows. Section 2 lays out the model and solves the agent's

[4] On the assumption of the agent's preference, the agent in DeMarzo and Sannikov (2010) is risk neutral. Prat and Jovanovic (2010) cover general utility functions for the agent's problem, but then specialize exponential utility in deriving the optimal contract.

[5] There are other papers that are related to learning but do not deal with the belief manipulation effect. Adrian and Westerfield (2009) focus on the disagreement about the agent's ability between the principal and the agent; in that paper, the agent is dogmatic about his belief (i.e., never updates his posterior belief about profitability from past performance), which eliminates the belief manipulation effect. There, although the agent could distort the principal's belief by shirking, the dogmatic agent (who does not realize that the firm's profitability is in fact above the one perceived by the principal) will not gain anything from this channel, and as a result there is no belief distortion effect. And, a recent paper by Cosimano, Speight, and Yun (2011) study the long-term contracting problem with binary unobservable productivity states, and show that the optimal contract tends to be sticky. They assume that the agent's effort is observable but not contractible, and hence both the principal and the agent always have the same information set, both on- and off- equilibrium paths.

problem. Section 3 reformulates the principal's problem which allows for dynamic programming. Section 4 solves for the optimal contract and discusses the implications. Section 5 concludes.

2 The Model

Consider a continuous-time infinite-horizon principal-agent model with a common constant discount rate $r > 0$. The project generates a cumulative output Y_t up to time t, which evolves according to

$$dY_t = (\mu_t + \theta_t)\, dt + \sigma dB_t, \tag{1}$$

where $\{B_t\}$ is a standard Brownian motion on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, μ_t is the agent's unobservable effort level, and σ is the volatility or risk in cash flows.

Relative to Holmstrom and Milgrom (1987), the novelty is that we introduce the project profitability θ_t into the output process (1). Equivalently, we can also interpret θ_t as the agent's unknown ability. The project profitability θ_t is unobservable and thus calls for learning. If the project profitability θ_t is perfectly observable (or without θ_t), there is no learning and the model can be reduced to Holmstrom and Milgrom (1987). For tractability reasons, we assume that the unknown productivity affects the output dY_t additively.

We assume that the profitability $\{\theta_t\}$ follows a martingale process whose innovations are modeled by another Brownian motion $\{B_{\theta,t}\}$ that is independent to $\{B_t\}$, i.e.,

$$d\theta_t = \phi\sigma dB_{\theta,t},$$

where ϕ is a positive constant. When $\phi = 0$, our model is reduced to the benchmark Holmstrom and Milgrom (1987). Because shocks to θ_t are permanent, learning θ_t is important in designing long-term optimal contract for both parties. At time 0, the principal and the agent share the common normal prior that $\theta_0 \sim \mathcal{N}\left(m_0, \Sigma_0^\theta\right)$ where $\mathcal{N}$ indicates normal distribution. Without loss of generality, we set $m_0 = 0$. For learning to be stationary, we assume the prior uncertainty

$$\Sigma_0^\theta = \sigma^2\phi,$$

Specifically, under $\Sigma_0^\theta = \sigma^2\phi$, the posterior variance $\Sigma_t^\theta = \Sigma_0^\theta$ for all t, and thus Bayesian updating does not depend on time t.

Both parties can commit to the long-term relationship. The risk-neutral principal (*she* later on) offers the risk-averse agent (*he* later on) a contract $\{c_t, \mu_t\}$, so that the agent is recommended to take the effort policy $\mu = \{\mu_t\}$, and is compensated by the wage process $c = \{c_t\}$. Both elements are measurable to $\mathcal{Y}_t \equiv \mathcal{F}\{Y_s : 0 \leq s \leq t\}$ which is the filtration generated by the output history.

As standard in moral hazard problems, the agent's actual effort policy is not observable. Also, we assume that the agent can privately save (i.e., saving is unobservable), thus the wage c_t can differ from the agent's actual consumption. Denote the agent's actual consumption by $\widehat{c}_t$ and actual effort by $\widehat{\mu}_t$. Following Holmstrom and Milgrom (1987), we assume that the agent has an exponential or CARA (constant-absolute-risk-aversion) preference, so that the agent's instantaneous utility from consumption $\widehat{c}_t$ and effort $\widehat{\mu}_t$ is

$$u\left(\widehat{c}_t, \widehat{\mu}_t\right) = -\frac{1}{a}\exp\left[-a\left(\widehat{c}_t - g\left(\widehat{\mu}_t\right)\right)\right],$$

where a is the agent's absolute risk aversion coefficient, and $g\left(\widehat{\mu}_t\right) = \frac{1}{2}\widehat{\mu}_t^2$ is the instantaneous quadratic monetary cost for effort $\widehat{\mu}_t$. The agent without any personal wealth has an exogenous reservation utility of v_0 when both parties sign the contract.

We emphasize that our model allows for the agent's private saving (or hidden saving). It is a celebrated result that CARA preference does not have wealth effect, and the issue of private saving can be easily dealt with (for instance, Fudenberg, Holmstrom, and Milgrom (1990), Williams (2009), He (2011)). In fact, private saving allows us to separate the agent's continuation payoff from another state variable in deriving the optimal contract, which renders the tractability of our problem.

2.1 Bayesian Learning and Effort

Recall that at time 0, the principal and the agent share the common prior: $\theta_0 \sim \mathcal{N}\left(m_0, \Sigma_0^\theta\right)$. Both parties update their beliefs based on their own information sets respectively. Recall that $\mathcal{Y}_t = \mathcal{F}\{Y_s : 0 \leq s \leq t\}$ is the augmented filtration generated by output path Y. Given any contract $\{c_t, \mu_t\}$, the principal's information set at time t can be written as $\mathcal{P}_t = \mathcal{F}\{Y_s, \mu_s : 0 \leq s \leq t\}$, as the principal knows the recommended effort policy μ. However, the agent's information set also

includes his actual effort policy $\widehat{\mu}$, i.e., $\mathcal{A}_t = \mathcal{F}\{Y_s, \mu_s, \widehat{\mu}_s : 0 \leq s \leq t\}$. Intuitively, relative to the principal, the agent knows (weakly) more because he knows his actual past effort choices $\widehat{\mu}$ (unobservable to the principal), and they might be different from the recommended effort policy $\mu \equiv \{\mu_t\}$. This distinction is important for our analysis.

Under the assumption that the agent indeed follows the recommended effort policy μ, the principal's posterior belief about θ_t is fully summarized by the first two moments:

$$m_t^{\mu} \equiv \mathbb{E}\left[\theta_t \,|\mathcal{Y}_t, \mu\right] \text{ and } \Sigma_t^{\theta,\mu} \equiv \mathbb{E}\left[(\theta_t - m_t)^2 \,|\mathcal{Y}_t, \mu\right].$$

Standard filtering argument (e.g., Theorem 12.2 in Liptser and Shiryayev (1977)) implies that

$$dm_t^{\mu} = \Sigma_t^{\theta,\mu} \frac{dY_t - (\mu_t + m_t^{\mu})\, dt}{\sigma^2} = \sigma\phi dB_t^{\mu}, \tag{2}$$

$$\Sigma_t^{\theta,\mu} = \sigma^2\phi \text{ for all } t \tag{3}$$

where B_t^{μ} is a standard Brownian motion under the measure induced by the effort policy μ:

$$dB_t^{\mu} = \frac{dY_t - (\mu_t + m_t^{\mu})\, dt}{\sigma}. \tag{4}$$

Conditional on the actual effort policy $\{\widehat{\mu}_t\}$, the agent forms his posterior belief as

$$m_t^{\widehat{\mu}} \equiv \mathbb{E}\left[\theta_t \,|\mathcal{Y}_t, \widehat{\mu}\right] \text{ and } \Sigma_t^{\theta,\widehat{\mu}} \equiv \mathbb{E}\left[\left(\theta_t - m_t^{\widehat{\mu}}\right)^2 |\mathcal{Y}_t, \widehat{\mu}\right].$$

The superscript $\widehat{\mu}$ is used here to emphasize the dependence on the agent's actual effort policy $\widehat{\mu}$ (which the principal does not know). Similar argument as above yields

$$dm_t^{\widehat{\mu}} = \Sigma_t^{\theta,\widehat{\mu}} \frac{dY_t - \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right) dt}{\sigma^2} = \sigma\phi dB_t^{\widehat{\mu}}, \tag{5}$$

$$\Sigma_t^{\theta,\widehat{\mu}} = \sigma^2\phi \text{ for all } t \tag{6}$$

where $B_t^{\widehat{\mu}}$ is a standard Brownian motion under the measure induced by the actual effort policy $\widehat{\mu}$:

$$dB_t^{\widehat{\mu}} \equiv \frac{1}{\sigma}\left(dY_t - \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right) dt\right). \tag{7}$$

Hence, from the agent's perspective, the output follows $dY_t = \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right) dt + \sigma dB_t^{\widehat{\mu}}$; while from the principal's perspective the evolution is $dY_t = (\mu_t + m_t^{\mu})\, dt + \sigma dB_t^{\mu}$. The potential belief divergence $m_t^{\widehat{\mu}} - m_t^{\mu}$ gives rise to the agent's incentive to manipulate principal's belief, which plays an important role in understanding the information rents enjoyed by the agent in our model.

2.2 Effort and Belief Distortion

The main difficulty of introducing learning into the dynamic moral hazard problem is not learning per se. Rather, the challenge is to deal with the issue of belief manipulation: the agent, simply by shirking from the recommended effort today, can distort downward the principal's future beliefs about the project profitability θ. Consider the following thought experiment which is helpful in understanding later results. Suppose that at time t the agent chooses an effort level $\widehat{\mu}_t$ below the recommended effort μ_t, and thus the output is lower than that is expected by the principal. With learning, the principal mistakenly attributes the lower output to a lower value of profitability θ_t, leading to downward belief manipulation. In words, by shirking, the agent makes the principal (mistakenly) believe that the project profitability is lower in the future. This belief manipulation is beneficial to the agent — when future output turns out to be high, the agent gets rewarded for high profitability (based on the agent's correct information set) rather than his effort.

We now formally characterize this effect. When the agent deviates from the recommended effort path μ by choosing effort $\widehat{\mu}$, the principal's posterior mean estimate about θ is distorted, and we denote this distortion by

$$\Delta_t \equiv m_t^{\widehat{\mu}} - m_t^{\mu}.$$

In the Gaussian framework with stationary learning, the variances of the posterior beliefs of both parties coincide to be a constant $\phi\sigma^2$. According to (2) and (5), the increment of Δ_t is given by

$$\begin{aligned} d\Delta_t &= dm_t^{\widehat{\mu}} - dm_t^{\mu} = \phi\left(dY_t - \left(\mu_t + m_t^{\mu}\right)dt\right) - \phi\left(dY_t - \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right)dt\right) \\ &= \phi\left(\mu_t - \widehat{\mu}_t - \Delta_t\right)dt, \end{aligned}$$

which allows us to solve for Δ_t to be

$$\Delta_t = \phi\int_0^t e^{\phi(s-t)}\left(\mu_s - \widehat{\mu}_s\right)ds. \tag{8}$$

Here, we have used the fact that both the principal and the agent share the same belief at the beginning, i.e., $\Delta_0 = 0$ because $m_0^{\mu} = m_0^{\widehat{\mu}} = m_0$. Intuitively, (8) says that the current belief

distortion is the properly-weighted cumulative effort deviation in the past. When $\phi = 0$, there will be no belief divergence, and the issue of belief manipulation is absent.

2.3 Formulating the Optimal Contracting Problem

We now formally state the agent's problem. Denote by S_t the balance of the agent's savings account, which earns interest at the constant rate r. Given the contract $c = \{c_t, \mu_t\}$ the agent's problem is

$$\begin{aligned} &\max_{\{\widehat{c},\widehat{\mu}\}} \mathbb{E}^{\widehat{\mu}} \left[\int_0^\infty e^{-rt} u\left(\widehat{c}_t, \widehat{\mu}_t\right) dt \right] \\ \text{s.t. } &dY_t = \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right) dt + \sigma dB_t^{\widehat{\mu}}, \\ &dS_t = rS_t dt + c_t dt - \widehat{c}_t dt \text{ with } S_0 = 0, \end{aligned} \tag{9}$$

with the transversality condition $\lim_{T\to\infty} \mathbb{E}^{\widehat{\mu}}\left[e^{-rT} S_T\right] = 0$. Here, $\mathbb{E}^{\widehat{\mu}}[\cdot]$ indicates that the probability measure is induced by the agent's effort policy $\{\widehat{\mu}_t\}$, and $\{\widehat{c}_t\}$ is the agent's actual consumption policy. Denote the optimal solution to (9) by $\{c_t^\star, \mu_t^\star\}$.

We call the contract $\{c_t, \mu_t\}$ *incentive-compatible and no-savings* if, given the contract $\{c_t, \mu_t\}$, the solution to the agent's problem (9) is $c_t^\star = c_t$ and $\mu_t^\star = \mu_t$, which further implies $S_t = 0$ for any t (i.e., no private savings at any time). In other words, the agent finds it optimal to consume his wages and work as recommended. As typical in the literature, the following lemma shows that it is without loss of generality to restrict our attention to incentive-compatible and no-savings contracts. The basic idea is similar to revelation principle, as the principal who can fully commit to the contract and knows the agent's actual effort policy will perform correct Bayesian updating based on that policy, and thus can save for the agent.

Lemma 1 *It is without loss of generality to focus on contracts that are incentive-compatible and no-savings.*

Proof. *See Appendix A.* ■

The optimal contract solves the principal's problem:

$$\max_{\{c_t,\mu_t\} \text{ is incentive-compatible and no-savings}} \quad \mathbb{E}_t^{\mu}\left[\int_0^{\infty} e^{-rt}\left(dY_t - c_t\right)dt\right] \tag{10}$$

$$\text{s.t.} \quad dY_t = \left(\mu_t + m_t^{\mu}\right)dt + \sigma dB_t^{\mu}; \tag{11}$$

$$\mathbb{E}_0^{\mu}\left[\int_0^{\infty} e^{-rt} u\left(c_t, \mu_t\right)dt\right] = v_0. \tag{12}$$

The first equation (11) describes the evolution of output performance under the principal's belief. Importantly, the principal's Bayesian updating of m_t^{μ} and dB_t^{μ} are based on the correct optimal effort policy $\{\mu_t\}$ taken by the agent in the incentive-compatible contract.

The constraint (12) is the participation constraint, as the contract must attract the agent with a reservation value v_0 at $t = 0$. As standard in the optimal contracting literature with CARA agents, this participation constraint must bind, and is only enforced at time 0 when the contract is signed.

2.4 The Agent's Problem

In this section we analyze the agent's problem when facing a compensation contract $\{c_t, \mu_t\}$, and gives the necessary and sufficient conditions for the contract $\{c_t, \mu_t\}$ to be incentive-compatible and no-savings.

2.4.1 Continuation Value and Incentive Slopes

Given the compensation contract $\{c_t, \mu_t\}$, denote the agent's continuation value from the contract by v_t, which is defined as

$$v_t \equiv \mathbb{E}_t^{\mu}\left[\int_t^{\infty} e^{-r(s-t)} u\left(c_s, \mu_s\right)ds\right]. \tag{13}$$

According to the standard martingale representation argument (e.g., Sannikov, 2008), we know that there exists some progressively measurable process $\{\beta_t\}$ so that the following holds:

$$\begin{aligned} dv_t &= rv_t dt - u\left(c_t, \mu_t\right)dt + \beta_t\left(-arv_t\right)\left(dY_t - m_t^{\mu}dt\right) \\ &= rv_t dt - u\left(c_t, \mu_t\right)dt + \beta_t\left(-arv_t\right)\sigma dB_t^{\mu}. \end{aligned} \tag{14}$$

Intuitively, $\beta_t(-arv_t)$ can be interpreted the incentive loading (in utils) on the agent's unexpected performance $dY_t - m_t^{\mu}dt$. As we show shortly, $-arv_t > 0$ is the marginal utility from consumption at time t. Therefore, by converting utils to dollars, β_t can be interpreted as dollar incentives on the agent's unexpected performance. Later we call $\{\beta_t\}$ the incentive slopes, pay-performance sensitivities, or simply incentives.

2.4.2 No Savings

We now show that due to the property of translation-invariance (or, absence of wealth effect) of CARA preference, under no-savings conditions the following holds:

$$rv_t = u(c_t, \mu_t) = -\frac{1}{a}\exp\left[-a\left(c_t - g(\mu_t)\right)\right]. \tag{15}$$

The argument is similar to He (2011). To see this, we first present the following lemma.

Lemma 2 *At any time $t \geq 0$, consider a deviating agent who has some arbitrary savings S_t and faces the continuation contract Π_t. Denote by $v_t(S;\Pi)$ his deviation continuation value, then we have*

$$v_t(S;\Pi) = v_t(0;\Pi)\cdot e^{-arS} = v_t \cdot e^{-arS}, \tag{16}$$

where we have used the fact that $v_t(0;\Pi)$ is the agent's continuation value v_t along the no-savings path defined in (13).

Proof. *See Appendix A.* ■

The driving force behind this result is simple. For CARA preference, the agent's problem is translation-invariant to his underlying wealth level, as suggested by $u(c_s + rS, \mu_s) = e^{-arS}u(c_s, \mu_s)$. Therefore, for a CARA agent, given the extra savings S, his new optimal policy is to take the optimal consumption-effort-learning policy without savings, but consume an extra rS more for all future dates $s \geq t$. Therefore with savings S, his deviation value is just the original value, multiplying by the adjusting factor e^{-arS} as a result of extra consumption.

Now by the optimality of the agent's consumption-savings policy in problem (9), his marginal utility from consumption must equal his marginal value of wealth:

$$u_c(c_t, \mu_t) = \frac{\partial}{\partial S} v_t(S; \Pi)|_{S=0} = -arv_t. \tag{17}$$

Because $au = -u_c$ under CARA preference, (15) follows immediately from (17).

Plugging (15) into (14), we find that in equilibrium v_t follows an exponential martingale,

$$\begin{aligned} & dv_t = \beta_t(-arv_t)\sigma dB_t^\mu \\ \Leftrightarrow \quad & v_s = v_t \exp\left(-\int_t^s ar\beta_u \sigma dB_u^\mu - \frac{1}{2}\int_t^s a^2 r^2 \beta_u^2 \sigma^2 du\right) \text{ for } s > t. \end{aligned} \tag{18}$$

And, from (15), the equilibrium consumption must satisfy $c_t = \frac{1}{2}\mu_t^2 - \frac{\ln(-arv_t)}{a}$.

2.4.3 Incentive Compatibility Constraint

Now we present the first key proposition that characterizes the agent's incentive compatibility constraint in (19), along with the equilibrium consumption and continuation value heuristically derived above. In the Appendix we provide rigorous proofs for (19), (20) and (21) based on the Pontryagin maximum principle. There, we also establish a sufficient condition so that the policy (c, μ) satisfying the necessary conditions is indeed optimal.

Proposition 1 *For the contract $\{c_t, \mu_t\}$ to be incentive-compatible and no-savings, the incentive slopes $\{\beta_t\}$ must satisfy*

$$\begin{aligned} \mu_t &= \underbrace{\beta_t}_{\textit{Instantaneous Incentive}} \underbrace{- \frac{1}{v_t}\mathbb{E}_t^\mu\left[\int_t^\infty \phi e^{-(\phi+r)(s-t)}\beta_s v_s ds\right]}_{\textit{Information Rent}} \\ &= \beta_t - \mathbb{E}_t^\mu\left[\int_t^\infty \phi e^{-(\phi+r)(s-t)}\beta_s \exp\left(-\int_t^s ar\beta_u\sigma dB_u^\mu - \frac{1}{2}\int_t^s a^2r^2\beta_u^2\sigma^2 du\right) ds\right]. \end{aligned} \tag{19}$$

And, the consumption follows

$$c_t = \frac{1}{2}\mu_t^2 - \frac{\ln(-arv_t)}{a}, \tag{20}$$

and the continuation payoff from the contract is

$$v_t = v_0 \exp\left(-\int_0^t ar\beta_s\sigma dB_s^\mu - \frac{1}{2}\int_0^t a^2r^2\beta_s^2\sigma^2 ds\right). \tag{21}$$

Proof. *See Appendix A.* ■

We devote the rest of this section to explain the intuition behind the effort optimality condition (19). In a standard dynamic agency problem, $\mu_t = \beta_t$, i.e., the agent's effort μ_t at time t should only depend on the time-t incentive slope β_t offered by the contract. With learning and associated belief-manipulation, the agent's effort decisions across periods are interlinked, as suggested by the forward-looking nature of the second downward adjustment term in (19).

This forward looking downward adjustment term represents the *information rent* to the agent, which is expected (properly) discounted future incentives $\{\beta_s : s > t\}$.[6] The discount rate is $\phi + r$ due to the time discounting and stationary learning discounting; and as explained later we also adjust for marginal utilities (which is proportional to v_s) at future states.

The incentive-compatibility condition in (19) is similar to Proposition 1 in Prat and Jovanovic (2010). Using our notation, and adjusting for the stationary learning, Proposition 1 in Prat and Jovanovic (2010) states that the agent's incentive compatibility constraint is

$$\begin{aligned} 0 &= u_\mu(c_t, \mu_t) + \beta_t u_c(c_t, \mu_t) - \mathbb{E}_t^\mu \left[\int_t^\infty \phi e^{-(\phi+r)(s-t)} \beta_s u_c(c_s, \mu_s)\, ds \right] \\ &= u_\mu(c_t, \mu_t) + \beta_t(-arv_t) - \mathbb{E}_t^\mu \left[\int_t^\infty \phi e^{-(\phi+r)(s-t)} \beta_s(-arv_s)\, ds \right]. \end{aligned} \tag{22}$$

This result can be heuristically understood as follows. Consider the agent who reduces his effort to slightly below the recommended effort level μ_t, say $\mu_t - \epsilon$, only at the time interval $[t, t+dt]$. In other words, given the recommended policy $\{\mu_s\}$, the deviation effort policy is

$$\mu^\epsilon = \begin{cases} \mu_s & \text{for } s \notin [t, t+dt] \\ \mu_s - \epsilon & \text{otherwise} \end{cases}.$$

What is the impact on the agent's total payoff onwards?

Shirking affects the agent's instantaneous utility $u(c_t, \mu_t)$ directly. Moreover, from the continuation payoff part, the agent receives his baseline continuation payoff v_t, plus the impact of shirking on his future changes of continuation payoffs. More specifically, we write the agent's total payoff

[6] Technically speaking, this term captures the *marginal* rent that the agent may enjoy by deviating from the recommended effort slightly, rather than the rent that the agent actually enjoys in equilibrium. It is because in equilibrium the principal knows the agent's actual effort and hence the agent does not know more than the principal does. However, the marginal deviation benefit (marginal rent) is important in characterizing the agent's incentive compatibility condition.

from time-t onwards as follows (recall (18)):[7]

$$u\left(c_t, \mu_t - \epsilon\right) dt + v_t + \mathbb{E}_t^{\mu}\left[\int_t^{\infty} e^{-r(s-t)} dv_s\right] \tag{23}$$

$$= \underbrace{u\left(c_t, \mu_t - \epsilon\right) dt}_{\text{Saving effort cost instantaneously}} + v_t + \mathbb{E}_t^{\mu}\left[\underbrace{\beta_t\left(-arv_t\right)\left(dY_t\left(\mu_t - \epsilon\right) - m_t^{\mu} dt\right)}_{\text{Hurting performance instantaneously}} + \int_{t+dt}^{\infty} e^{-r(s-t)} \underbrace{\beta_s\left(-arv_s\right)\left(dY_s\left(\mu_s\right) - m_s^{\mu^{\epsilon}} dt\right)}_{\text{Creating belief divergence persistently}}\right] \tag{24}$$

As the second line in (24) shows, there are two channels that shirking at time t affects future changes in the agent's continuation payoffs. The first channel is the instantaneous effect. Write the performance $dY_t\left(\mu_t\right)$ over $[t, t+dt]$ as a function of time-t effort μ_t. Then, exerting effort $\mu_t - \epsilon$ will hurt the short-term performance over $[t, t+dt]$ relative to the recommended level, because

$$dY_t\left(\mu_t - \epsilon\right) = \left(\mu_t - \epsilon\right) dt + m_t^{\mu} dt + \sigma dB_t^{\mu} = dY_t\left(\mu_t\right) - \epsilon dt.$$

The second effect is the persistent effect due to belief manipulation. We have seen that in (8) that once shirking creates a positive time-t belief divergence between the principal and the agent, this belief divergence will persist in the future contractual relation. This effect brings about benefit to the agent, because it lowers the principal's profitability estimate $m_s^{\mu^{\epsilon}}$ for future performance in (24). Intuitively, if the principal mistakenly believes that the project is less profitable than it should be, the agent's normal performance will be considered superb and enjoys rent from this private information.

Now we recover (22) based on (24). The first term $u_{\mu}\left(c_t, \mu_t\right)$ in (22) is the agent's instantaneous marginal utility from shirking at t, which corresponds to the first effect in (24), i.e., "reduce effort cost instantaneously." The second term in (22) is the marginal cost from hurting his continuation payoff by delivering a lower time-t performance, and it is coming from the second effect in (24), i.e., "hurting performance instantaneously." In standard models (e.g., Sannikov (2008)) without belief

[7] To see this, define $G(t) \equiv \int_0^t e^{-rs} u_s ds + e^{-rt} v_t$, and $G(\infty)$ is the agent's total payoff. Due to private savings, $u_s = rv_s$, and we have $dG(t) = e^{-rt} dv_t$. Therefore the total payoff (inflated by e^{rt}) is $\mathbb{E}_t^{\mu}\left[e^{rt} G(\infty)\right] = e^{rt} G(t) + \mathbb{E}_t^{\mu}\left[\int_t^{\infty} e^{-r(s-t)} dv_s\right]$, which is $u\left(c_t, \mu_t\right) dt + v_t + \mathbb{E}_t^{\mu}\left[\int_t^{\infty} e^{-r(s-t)} dv_s\right]$ by ignoring utilities occuring before t. Under equilibrium effort, dv_s is martingale increment and $\mathbb{E}_t^{\mu}\left[\int_t^{\infty} e^{-r(s-t)} dv_s\right] = 0$.

manipulation, these two forces fully determine the agent's trade-off in choosing his optimal effort at t.

The third term in (22) is due to the agent's belief manipulation, which corresponds to the third effect in (24), i.e., "creating belief divergence persistently." According to (8), because of shirking at t, the total belief divergence at any future time $s > t$ is

$$\Delta_s = \phi \int_t^s e^{\phi(u-s)} \left(\mu_u - \mu_u^{\epsilon}\right) du = \phi e^{-\phi(s-t)} \left(\epsilon dt\right). \tag{25}$$

Intuitively, the belief divergence will be persistent but decay over time, as new information keeps flowing in. The particular stationary setting assumed in this paper implies that the belief divergence decays over time exponentially.

From (24) and (25), at date $s \in (t, \infty)$, given $\beta_s \left(-arv_s\right)$ which is the marginal utility sensitivity to the agent's performance, the marginal benefit of distorting the belief m_s is

$$\beta_s \left(-arv_s\right) \Delta_s ds = \beta_s \left(-arv_s\right) \phi e^{-\phi(s-t)} \left(\epsilon dt\right) ds.$$

Thus, a higher future pay-performance sensitivity β_s gives rise to a greater information rent, because the contract is more likely to reward the agent due to downward distorted project profitability. Also, the information rent depends on the agent's future marginal utility from consumption $\left(-arv_s\right)$ as well; as shown in Section 4.2.3, this is the driver for a history-dependent contract to be optimal. Taking into account of the time-discounting at the rate of r, the total information rent due to belief manipulation is

$$\mathbb{E}_t^{\mu} \left[\int_t^{\infty} \phi e^{-(\phi+r)(s-t)} \beta_s \left(-arv_s\right) ds \right] \epsilon dt.$$

This expression coincides with the third term on the right hand side of (22).

Finally, under CARA preference we have $u_{\mu}\left(c_t, \mu_t\right) = u_c\left(c_t, \mu_t\right) \mu_t = \left(-arv_t\right) \mu_t$. After dividing both sides by the time-t marginal utility $-arv_t$ in (22) and using (21), we reach the key incentive-compatibility condition (19).

3 The Principal's Problem

For ease of notation, from now on we use dB_t to denote the Brownian incremental dB_t^{μ}, and write $\mathbb{E}^{\mu}$ as $\mathbb{E}$ because we now focus on incentive-compatible contracts such that both parties will have the same information set.

3.1 Rewriting the Principal's Problem

We first rewrite the principal's problem in (10) in light of Proposition 1. Essentially, Proposition 1 establishes an important link between the recommended effort $\{\mu_t\}$ and the incentives $\{\beta_t\}$ in any incentive-compatible contracts, which allows the principal to choose $\{\beta_t\}$ (in turn choose $\{\mu_t\}$) to maximize her value. Once we find the optimal incentives, denoted by $\{\beta_t^*\}$, the corresponding optimal consumption process $\{c_t^*\}$ and the optimal effort policy $\{\mu_t^*\}$ are determined by (20) and (19), respectively. Therefore we can rewrite the principal's problem (10) as

$$\max_{\{\beta_t\}} \mathbb{E}\left[\int_0^{\infty} e^{-rt}\left(dY_t - c_t dt\right)\right]$$

$$s.t. \quad dY_t = (\mu_t + m_t)dt + \sigma dB_t \text{ and } dm_t = \phi\sigma dB_t, \tag{26}$$

$$c_t = g\left(\mu_t\right) - \frac{\ln\left(-arv_t\right)}{a}, \text{ where } g\left(\mu_t\right) = \frac{1}{2}\mu_t^2 \tag{27}$$

$$dv_t = \beta_t\left(-arv_t\right)\sigma dB_t, \text{ given } v_0; \tag{28}$$

$$\mu_t = \beta_t - \mathbb{E}_t\left[\int_t^{\infty} \phi\beta_s e^{-(\phi+r)(s-t)} \exp\left(-\int_t^s ar\beta_u\sigma dB_u - \frac{1}{2}\int_t^s a^2r^2\beta_u^2\sigma^2 du\right) ds\right]. \tag{29}$$

Here, (26) is from Bayesian updating; (27) is from (20); (28) is from (21), with the binding participation constraint. Finally, (29) are from (19) in Propositions 1.

We will use dynamic programming technique to solve the above problem. The formulation above suggests that we need to keep track of two state variables: one is the agent's continuation payoff (i.e., promised utility in Williams (2011)); and the other is the information rent promised to the agent, i.e., the second term in the right hand side of (29), which corresponds to the promised marginal utility in Williams (2011).

We devote the rest of the section to show that intrinsically there is only one relevant state variable in our model. One property of (29) is crucial, that is, the information rent is independent

of the agent's continuation value v_t. As we will show now, this property allows us to formulate the principal's problem in dynamic programming with only one state variable, which is the information rent promised to the agent.

We can plug (26) and (27) to rewrite the principal's objective in a more transparent way. First of all, on outputs from the project, we have

$$\mathbb{E}\left[\int_0^\infty e^{-rt}dY_t\right] = \mathbb{E}\left[\int_0^\infty e^{-rt}(\mu_t + m_t)dt\right] = \mathbb{E}\left[\int_0^\infty e^{-rt}\mu_t dt\right],$$

where we use the fact that m_t is a martingale with a starting value $m_0 = 0$. Second, for wage costs, using integration by parts, $\mathbb{E}\left[\int_0^\infty -e^{-rt}\frac{1}{a}\ln\left(-arv_t\right)dt\right] = \mathbb{E}\left[\int_0^\infty \frac{\ln(-arv_t)}{ar}d\left(e^{-rt}\right)\right]$ can be rewritten as

$$-\frac{\ln\left(-arv_0\right)}{ar} - \mathbb{E}\left[\int_0^\infty e^{-rt}\frac{d\ln\left(-v_t\right)}{ar}\right] = -\frac{\ln\left(-arv_0\right)}{ar} + \mathbb{E}\left[\int_0^\infty e^{-rt}\frac{1}{2}ar\sigma^2\beta_t^2 dt\right],$$

where the last equality we use Eq. (28). Hence, the total expected wage cost is

$$\begin{aligned}\mathbb{E}\left[e^{-rt}c_t dt\right] &= \mathbb{E}\left[\int_0^\infty e^{-rt}\left(g\left(\mu_t\right) - \frac{1}{a}\ln\left(-arv_t\right)\right)dt\right] \\ &= \underbrace{-\frac{\ln\left(-arv_0\right)}{ar}}_{\text{certainty equivalent of outside option } v_0} + \mathbb{E}\left[\int_0^\infty e^{-rt}\left(\underbrace{g\left(\mu_t\right)}_{\text{effort cost}} + \underbrace{\frac{1}{2}ar\sigma^2\beta_t^2}_{\text{risk compensation}}\right)dt\right]\end{aligned}$$

Conveniently, the total compensation cost is the certainty equivalent $-\frac{\ln(-arv_0)}{ar}$ of delivering the agent's outside option v_0, plus the total monetary effort cost $g\left(\mu_t\right) = \mu_t^2/2$ and the total discounted risk compensation due to incentive provisions. Thus, the certainty equivalent $-\frac{\ln(-arv_0)}{ar}$ separates from the problem, and the optimal solution $\{\beta_t^*\}$ will be independent of the agent's initial outside option v_0. This result comes from the translation invariance of CARA preference.

Therefore, the principal's problem can be simplified to:

$$\begin{aligned}&\max_{\{\beta_t\}}\mathbb{E}\left[\int_0^\infty e^{-rt}\left(\mu_t - \frac{1}{2}\mu_t^2 - \frac{1}{2}ar\sigma^2\beta_t^2\right)dt\right] \qquad (30)\\ s.t.\ \ \mu_t &= \beta_t - \mathbb{E}_t\left[\int_t^\infty \phi\beta_s e^{-(\phi+r)(s-t)}\exp\left(-\int_t^s ar\beta_u\sigma dB_u - \frac{1}{2}\int_t^s a^2r^2\beta_u^2\sigma^2 du\right)ds\right], \\ \beta_t &\in [-M, M] \text{ for some sufficiently large } M.\end{aligned}$$

We impose the last constraint as the transversality condition on our infinite horizon problem. Essentially, we restrict the feasible incentive slopes $\{\beta_t\}$ to be bounded, i.e., there exists some

sufficiently large constant M so that $\beta_t \in [-M, M]$. Importantly, as we will prove later that the derived optimal policy is independent of M when M is sufficiently high. We devote the rest of the paper to solve the problem in (30).

3.2 Recursive Formulation of the Problem

3.2.1 Information rent as the unidimensional state variable

In light of the incentive-compatibility constraint in (29), define the information rent p_t as

$$p_t \equiv \mathbb{E}_t \left[\int_t^\infty \phi \beta_s e^{-(\phi+r)(s-t)} \exp\left(- \int_t^s ar\beta_u \sigma dB_u - \int_t^s \frac{1}{2} \left(ar\beta_u \sigma \right)^2 du \right) ds \right]. \tag{31}$$

As shown shortly, p_t serves as the only state variable that summarizes all payoff-relevant information we need to determine the principal's value going forward beyond t. In words, the principal only needs to keep track of the promised information rent in designing the continuation contract.

The following lemma shows that the information rent p_t is bounded if incentives β_t is bounded. Moreover, the boundary for p_t is absorbing; and once the absorbing boundary is hit, all future incentives have to take the same extreme values always. In the next section we use this result to pin down boundary conditions for the principal's value function.

Lemma 3 *Suppose that $\beta_t \in [-M, M]$ where M is fixed. Then the state variable p_t reaches $\pm M_p \equiv \pm \frac{\phi M}{\phi + r}$ if and only if $\beta_s = \pm M, \forall s \geq t$, which implies that $\pm M_p$ are absorbing states for p.*

Proof. *See Appendix A.* ■

When the state variable $p_t \in (-M_p, M_p)$ is away from boundaries, Appendix A.5 shows that the dynamics of p_t can be characterized as follows

$$dp_t = \left[(\phi + r) p_t + \beta_t \left(ar\sigma\sigma_t^p - \phi \right) \right] dt + \sigma_t^p dB_t, \tag{32}$$

where σ_t^p is some progressively measurable process derived from the martingale representation. The policies σ_t^p and β_t for $p_t \in (-M_p, M_p)$, combined with the boundary condition that $\beta_s = \pm M, s > t$ whenever p_t hits $\pm M_p$, give the full history of $\{\beta_t : t \geq 0\}$. We later solve for β_t and σ_t^p as part of the optimal contract.

Now we are able to write problem (30) in the standard dynamic programming language. Substituting $\mu_t = \beta_t - p_t$ in the principal's objective, we have

$$\begin{aligned} & \max_{\{\beta_t, \sigma_t^p\}} \mathbb{E}\left[\int_0^\infty e^{-rt}\left((\beta_t - p_t) - \frac{1}{2}(\beta_t - p_t)^2 - \frac{1}{2}ar\sigma^2\beta_t^2\right)dt\right] \\ s.t.\ dp_t &= \left[(\phi + r)\, p_t + \beta_t\left(ar\sigma\sigma_t^p - \phi\right)\right]dt + \sigma_t^p dB_t \text{ for all } t > 0, \text{ and } p_0 = p \\ \beta_s &\in [-M, M],\ p_t \in [-M_p, +M_p], \text{ and } p_t = \pm M_p \text{ are absorbing.} \end{aligned}$$

Here, we are after the optimal policy $\left\{\beta_t^*, \sigma_t^{p,*}\right\}$ which jointly controls the drift and volatility of the state p_t in (32).

3.2.2 Relaxed problem

Our road map is as follows. We first consider the principal's relaxed maximization problem given M (and hence $M_p = \frac{\phi M}{\phi + r}$), with the value function $V(p; M)$, as

$$\begin{aligned} V(p; M) \equiv & \max_{\{\beta_t, \sigma_t^p\}} \mathbb{E}\left[\int_0^\infty e^{-rt}\left(\mu_t - \frac{1}{2}\mu_t^2 - \frac{1}{2}ar\sigma^2\beta_t^2\right)dt\right] \\ s.t.\ dp_t &= \left[(\phi + r)\, p_t + \beta_t\left(ar\sigma\sigma_t^p - \phi\right)\right]dt + \sigma_t^p dB_t \text{ for all } t > 0, \text{ and } p_0 = p \\ & p_t \text{ is absorbing at } \pm M_p, \\ & \beta_t \text{ can exceed } M \text{ (but remains finite) when } p_t \in (-M_p, M_p). \end{aligned} \tag{33}$$

Note that this problem in (33) is a relaxed version of the principal's original problem (30). Essentially, given M, in the original problem (30) we require $\beta_t \in [-M, M]$ always independent of p_t; while in problem (33) we only require $\beta_t \in [-M, M]$ whenever p_t hits the boundaries $\pm M_p$ in light of Lemma 3. As later we show that for sufficiently large M the value achieved in the relaxed problem is the same as the value from the original problem, the solution to the relaxed problem is also the solution to the original problem.

The absorbing property at $p = \pm M_p$ gives a simple boundary condition for $V(p; M)$. Because once p_t hits $\pm M_p$ the future incentives β_s are absorbing at $\pm M$ for $s > t$, in the proof of Lemma 3 we show that $V(\pm M_p; M)$ is quadratic in M:

$$V(\pm M_p; M) = \pm\frac{M}{\phi + r} - \frac{r}{2}\left[\frac{1}{(\phi + r)^2} + a^2\sigma^2\right]M^2. \tag{34}$$

The Hamilton-Jacobi-Bellman (HJB) equation for (33) is straightforward:

$$rV = \max_{\beta,\sigma^P} (\beta - p) - \frac{1}{2}(\beta - p)^2 - \frac{ar\sigma^2}{2}\beta^2 + V_p\left[(\phi + r)p + \beta\left(ar\sigma\sigma^P - \phi\right)\right] + \frac{1}{2}V_{pp}\left(\sigma^P\right)^2. \quad (35)$$

Under the assumption that $1 + ar\sigma^2 + a^2r^2\sigma^2\frac{(V_p)^2}{V_{pp}} > 0$ and $V_{pp} < 0$ that we will verify later in Proposition 4, the first-order optimality conditions are given by:

$$\beta = \frac{1 + p - \phi V_p}{1 + ar\sigma^2 + a^2r^2\sigma^2\frac{(V_p)^2}{V_{pp}}} \text{ and } \sigma^p = -ar\sigma\beta\frac{V_p}{V_{pp}}. \quad (36)$$

Plugging them back into the HJB equation (35) and dropping M in $V(p; M)$ without confusion, we have

$$rV = \frac{1}{2}\frac{(1 + p - \phi V_p)^2}{1 + ar\sigma^2 + a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} - p - \frac{1}{2}p^2 + V_p(\phi + r)p, \quad (37)$$

with the boundary condition in (34). We will solve the problem in (33) by analyzing the above ordinary differential equation in (37).

3.3 Observable Profitability Benchmark

Before we move on to solve the optimal contract, we consider the benchmark case where the profitability θ_t is observable. This case is almost the same as the standard Holmstrom and Milgrom (1987) model, except that the optimal contract always benchmarks the agent's performance to θ_t. Using the incentive constraint $\mu_t = \beta_t$, the optimal solution is

$$\mu_t^{HM} = \frac{1}{1 + ar\sigma^2},$$

and the principal's value is $V^{HM} = 1/\left(2r\left(1 + ar\sigma^2\right)\right)$. The optimal contract can be implemented by a constant equity share $1/\left(1 + ar\sigma^2\right)$ (with proper benchmarking). Because the principal can ignore the information constraint, V^{HM} serves as an upper bound for our value function, i.e., we have $V(p) \leq 1/\left(2r\left(1 + ar\sigma^2\right)\right)$.

4 Optimal Contracting

We study optimal contracting in this section. As a benchmark, we first solve the optimal contract in closed form with the restriction of deterministic policies. We then study the general optimal

contract by characterizing the solution to the ODE (37). Relative to the existing literature that focuses on implementing a constant effort policy, two interesting results emerge. First, due to the forward looking information rent, the optimal effort policy is decreasing over time. Second, to reduce the information rent, the optimal contract induces a negative correlation between the agent's marginal utility and pay-performance sensitivity. This gives rise to an option-like feature in the optimal contract in the sense that the agent's incentives rises after favorable shocks.

4.1 Optimal Deterministic Contract

Before we characterize the optimal contract fully, to gain better intuition we first constrain the incentive slopes $\{\beta\}$ to be deterministic. Within the class of deterministic contracts, we solve the optimal contract in closed form. Moreover, the result in this section allows us to compare our paper to Prat and Jovanovic (2010) where the optimal contract implements a deterministic (constant) effort policy.

4.1.1 Time-decreasing optimal effort policy

When $\{\beta\}$ is deterministic, we can move the conditional expectation in (31) inside the integral, so that the information rent p_t is simply[8]

$$p_t = \phi \int_t^\infty e^{-(\phi+r)(s-t)} \beta_s ds.$$

Therefore, deterministic incentives imply deterministic information rents, i.e., the volatility σ^p is constrained to zero.

Denote by $V^d(p)$ the value function with deterministic policies. Plugging $\sigma^p = 0$ into (35), we have $\beta^d(p) = \left(1 + p - \phi V_p^d\right) / \left(1 + ar\sigma^2\right)$, with the HJB equation as:

$$rV^d(p) = \frac{1}{2} \frac{\left(1 + p - \phi V^d(p)\right)^2}{1 + ar\sigma^2} - p - \frac{1}{2} p^2 + V_p^d(p)(\phi + r) p.$$

This ODE admits a closed-form solution

$$V^d(p) = -\frac{1}{2} A^d p^2 + B^d p, \tag{38}$$

[8] It is because $\exp\left(-\int_t^s ar\beta_v \sigma dB_v^\mu - \frac{1}{2}\int_t^s a^2 r^2 \beta_v^2 \sigma^2 dv\right)$ is a martingale and independent of β_s (which is deterministic).

where $B^d = 1/\phi$, and

$$A^d = \frac{1}{2\phi^2}\left((2\phi + r)\, ar\sigma^2 + r + \sqrt{(2\phi + r)^2 a^2 r^2 \sigma^4 + 2ar\sigma^2\left[(\phi + r)^2 + \phi^2\right] + r^2}\right) > 0.$$

We have the following proposition.

Proposition 2 *Within the class of deterministic policies, the value function for the optimal contract is given by (38), and it is optimal to set time-0 information rent to be*

$$p_0^d = \overline{p}^d \equiv \frac{B^d}{A^d} > 0, \tag{39}$$

where $\overline{p}^d$ satisfies $V_p^d\left(\overline{p}^d\right) = 0$. The evolution of information rent p is given by

$$p_t^d = \overline{p}^d \exp\left(-\lambda t\right), \tag{40}$$

where $\lambda \equiv -\phi - r + \frac{1+A^d\phi}{1+ar\sigma^2}\phi > 0$. In the deterministic optimal contract, the incentive slope is given by

$$\beta_t^d\left(p\right) = \frac{1 + A^d\phi}{1 + ar\sigma^2} p_t^d = \frac{1 + A^d\phi}{1 + ar\sigma^2}\overline{p}^d \exp\left(-\lambda t\right),$$

and the optimal effort policy is

$$\mu_t^d = \beta_t^d - p_t^d = \frac{A^d\phi - ar\sigma^2}{1 + ar\sigma^2} p_t^d = \frac{A^d\phi - ar\sigma^2}{1 + ar\sigma^2}\overline{p}^d \exp\left(-\lambda t\right).$$

Proof. *See Appendix A.* ■

The above proposition shows that in the optimal deterministic contract, the information rent p_t^d, the incentive slope β_t^d, and the optimal effort μ_t^d all follow some exponentially decaying paths (toward zero), with the same rate of $-\lambda$. Moreover, at $t = 0$,

$$\mu_0^d = \frac{A^d\phi - ar\sigma^2}{1 + ar\sigma^2} p_0^d = \frac{1 - ar\sigma^2\overline{p}^d}{1 + ar\sigma^2} < \frac{1}{1 + ar\sigma^2} = \mu_0^{HM}.$$

In sum, in the optimal deterministic policy the initial effort level is below the observable profitability benchmark, and the optimal effort path is decreasing over time.

The reason for the front-loaded effort policies being optimal is because of the forward looking nature of information rent. From the agent's incentive-compatibility condition in (19), the belief

manipulation effect implies that giving incentives later will give the agent incentives to shirk earlier, but not the other way around. This implies that later incentives are more costly than early ones, and consequently the optimal contract implements higher effort in earlier periods.

4.1.2 Comparison to Prat and Jovanovic (2010)

Now we compare our results to that of Prat and Jovanovic (2010), who focus on contracts that implement a constant first-best effort level 1 which is by definition deterministic.[9] In this regard, the optimal *deterministic* (rather than stochastic) policy derived in Section 4.1.1 allows for a better comparison between our model and Prat and Jovanovic (2010).

Information rents and front-loaded effort policy Similar to the findings in Prat and Jovanovic (2010), front-loaded incentives are also optimal in our model. Because Prat and Jovanovic (2010) implement a constant effort, the forward looking nature of information rents implies that the compensation contract has to offer front-loaded incentives. Our model allows the optimal contract to adjust the implemented optimal effort policy (under the restriction of being deterministic). As the agent enjoys information rents from future pay-performance sensitivities, cheaper incentive provisions in earlier periods naturally push the optimal contract to implement a front-loaded effort profile.

The role of non-stationary learning Another interesting difference between our result and that of Prat and Jovanovic (2010) is due to the assumption of stationary learning. Prat and Jovanovic (2010) adopt the setting with non-stationary learning: the underlying project profitability θ never changes, and as time passes both parties get to learn the true project profitability eventually. More specifically, in Prat and Jovanovic (2010), the posterior updating equations corresponding to

[9] In their model, the effort cost is assumed to be linear over the feasible interval $[0, 1]$; a similar structure is assumed in DeMarzo and Sannikov (2010)).

(2) and (3) are

$$dm_t = \Sigma_t^\theta \frac{dY_t - (\mu_t + m_t)\,dt}{\sigma^2} \equiv \frac{\Sigma_t^\theta}{\sigma} dB_t^\mu, \tag{41}$$

$$\Sigma_t^\theta = \frac{\sigma^2 \Sigma_0^\theta}{\sigma^2 + \Sigma_0^\theta t}, \tag{42}$$

The most salient difference is between (3)-(42). In the stationary case that we are studying, the posterior variance is constant over time, while in the stationary case, the posterior variance decreases over time.

Under non-stationary learning, Prat and Jovanovic (2010) find that the Pareto frontier increases over time.[10] Although there is no flexibility in adjusting effort in the model of Prat and Jovanovic (2010), one tends to extrapolate to a statement that the optimal effort (if allowed to change) will *increase* over time. Indeed, non-stationary learning implies that the profitability θ becomes perfectly known eventually. Given this, one might think that if a continuum of effort level was allowed (as in this paper), the contract would have implemented higher and higher effort over time, approaching to either the first best level 1 or the Holmstrom and Milgrom (1987) level $1/\left(1 + ar\sigma^2\right)$.

This reasoning is potentially flawed. Although both parties learn the profitability θ perfectly when $t \to \infty$ so that we are indeed in the Holmstrom and Milgrom (1987) setting, the optimal contract might not implement a high effort in the distant future, simply because it could leave too much information rents to the agent in earlier periods.

Nevertheless, because the information quality about profitability improves over time, non-stationary learning does push the optimal effort policy to be back-loaded, an opposite force against the information rent effect which favors front-loaded effort policies. Which force is stronger, the information rent effect or non-stationary learning effect? To address this issue, we consider a variation of our model with non-stationary learning where the posterior variance follows (42), as assumed in Prat and Jovanovic (2010). We have the following analytical result.

Proposition 3 *Restrict attention to the deterministic policies. With non-stationary learning, the*

[10] This result is in contrast to the spot-market solution (Holmstrom, 1999) where the Pareto frontier shrinks over time. Because only spot wages (no within-job incentive contracts) are allowed in Holmstrom (1999), the equilibrium effort goes to zero when θ becomes close to perfectly known.

optimal effort policy is decreasing over time.

Proof. *See Appendix B.* ■

Thus, under deterministic policies, the optimal effort policy is still decreasing with time, suggesting that the information rent effect dominates the information quality effect (due to non-stationary learning). We leave future research to investigate the general optimal contract with non-stationary learning.

4.2 Optimal Contract

4.2.1 Characterizing the value function

We now solve the optimal contract without the restriction of incentives $\{\beta_t\}$ being deterministic. Based on dynamic programming technique, we analyze the ODE (37) with boundary conditions (34), which gives the solution to the constrained problem in (33).

Impose the following parametric condition:

$$\frac{\phi}{r} > a\left[2\left(\frac{\phi}{r}+1\right)^3 - \phi\sigma^2\right]. \tag{43}$$

The following proposition characterizes the properties of the value function under the optimal policy $\{\beta^*, \sigma^{p,*}\}$:

Proposition 4 *When M is sufficiently large, we have the following properties for $V(p;M) \in \mathbb{C}^2$.*

1. *$V(0;M) = 0$ and $V_p(0;M) = 1/\phi$.*
2. *$V(p;M)$ is strictly concave over the interval $\left[-\frac{\phi M}{\phi+r}, \frac{\phi M}{\phi+r}\right]$, and $1 + ar\sigma^2 + a^2r^2\sigma^2\frac{(V_p)^2}{V_{pp}} > 0$.*
3. *There exists $\overline{p} \in \left(0, \overline{p}^d\right)$ so that $V_p(\overline{p};M) = 0$. Under the optimal policy, $\overline{p}$ is an upper entrance-no-exit boundary, and 0 is a lower absorbing boundary 0 with $V(0;M) = 0$. This implies that under the optimal policy the endogenous state variable p never exits the interval $[0, \overline{p}]$.*
4. *The endogenous upper boundary $\overline{p}$ is independent of M.*

5. *Thus, when M is sufficiently large, the value function $V(p; M) = V(p)$ and associated optimal policies $\{\beta^*, \sigma^{p,*}\}$ are independent of M.*

Proof. *See Appendix A.* ■

In the optimal contract, at date 0 the principal sets $p_0^* = \overline{p}$; afterwards the state variable p_t^* evolves according to (32). The optimal control is characterized by (36). Interestingly, property (3) states that the information rent p_t^* will never wander out of an endogenous interval $[0, \overline{p}]$. Intuitively, because the information rent p_t^* is the (properly) discounted promised future incentives as in (31), it is suboptimal to promise too much or too little future incentives in optimal contracting. This result is reminiscent of Holmstrom and Milgrom (1987): without learning, the problem is essentially isolated each period, and the optimal incentives $\{\beta_t^*\}$ remains constant. In our model with learning, incentive provisions are necessarily connected across periods, and the optimal incentives $\{\beta_t^*\}$ become stochastic. Nevertheless, since the model primitives are stationary (CARA-Normal setting, additive technology in (26), and stationary learning), the (properly) discounted information rents $\{p_t^*\}$ remain endogenously bounded in the optimal contract.

The next theorem shows that for sufficiently large M the value achieved in the relaxed problem is the same as the value achieved in the original problem. Hence, the solution to the relaxed problem is also the solution to the original problem, which is the optimal contract that we are after.

Theorem 1 *For sufficiently large M, the solution to relaxed problem in (33) is the solution to the original problem in (30). Hence the solution in Proposition 4 characterizes the optimal contract.*

Proof. *See Appendix A.* ■

4.2.2 Asymptotic analysis for a risk-tolerant agent

To achieve more analytical tractability, we perform an asymptotic analysis for agents who are sufficiently risk tolerant (relatively small a). We first establish the result for risk-neutral agent, i.e., $a = 0$. In this case, the optimal deterministic policy obtained in Section 4.1 achieves first-best (and hence is optimal among all possible contracts).

Lemma 4 *When $a = 0$, the deterministic policy in Proposition 2 is optimal with the first-best effort level $\mu^o = 1$. The corresponding value function, denoted by $V^o(p)$, is given by*

$$V^o(p) = -\frac{1}{2}A^o p^2 + B^o p, \tag{44}$$

where $A^o \equiv r/\phi^2$ and $B^o \equiv 1/\phi$, and the optimal information rent $\overline{p}^o$ satisfies $V_p^o(\overline{p}^o) = 0$ so that

$$\overline{p}_t^o = \overline{p}^o = \frac{B^o}{A^o} = \frac{\phi}{r}. \tag{45}$$

Proof. *Because the first-best level is achieved by the proposed contract, the result is trivial.* ■

When the agent is risk averse (i.e, $a > 0$), we consider asymptotic expansions around the benchmark case $a = 0$. More specifically, denote $a_r \equiv ar$, and we solve for the expansions in the following form:

$$\begin{aligned} V(p;a) &= Q_0(p) + \textstyle\sum_{i=1}^{I} a_r^i Q_i(p) + o\left(a_r^I\right), & (46)\\ \overline{p} &= q_0^* - \textstyle\sum_{j=1}^{J} a_r^j q_j^* + o\left(a_r^J\right), & (47) \end{aligned}$$

where the highest expansion orders I and J are positive integers to be chosen later. We have $Q_0(p) = V^o(p)$ as given in (44), and $q_0^* = \overline{p}^o$ as given in (45). We solve for $Q_i(p)$ and $q_i^*(p)$ in closed-form (i.e., they are not asymptotic expansions). Also, given $V(p,a)$ it is easy to derive the optimal control pair $\{\beta(p), \sigma^p(p)\}$ using (36).

The optimal deterministic contract studied in Proposition 2 is helpful in deriving Q_i's in (46) and q_i^*'s in (47). Indeed, by expanding both $V^d(p;a)$ and $\overline{p}^d = B^d/A^d$ in the forms of (46), we can compare the optimal deterministic policies to those optimal stochastic ones. We choose the expansion orders I and J to be the lowest orders so that these two expansions start to differ. For instance, we choose $J = 3$ because the asymptotic analysis reveals that the upper boundary $\overline{p}$ in the stochastic optimal contract starts to differ from the deterministic counterpart $\overline{p}^d$ in the third order a_r^3. The next proposition summarizes our results.

Proposition 5 *When a is relatively small, we have the following approximations:*

$$\overline{p} = \overline{p}^d - a_r^3 \frac{\phi^5 \sigma^4}{r^6} \left(1 + \frac{r}{\phi}\right)^3 + o\left(a_r^3\right), \tag{48}$$

$$V\left(\overline{p}\right) = V^d\left(\overline{p}^d\right) + a_r^4 \left[\frac{\phi^6 \sigma^6}{r^8} \left(1 + \frac{r}{\phi}\right)^5 + \frac{\phi^5 \sigma^6}{r^7} \left(1 + \frac{r}{\phi}\right)^3\right] + o\left(a_r^4\right). \tag{49}$$

Furthermore, we show that

$$\beta\left(p\right) = \beta^d\left(p\right) - a_r^2 \frac{\sigma^2}{\phi} \left(1 + \frac{r}{\phi}\right) p\left(p - \overline{p}^o\right) \left[\left(p - \overline{p}^o\right) + \left(1 + \frac{r}{\phi}\right) p\right] + o\left(a_r^2\right), \tag{50}$$

$$\sigma^p\left(p\right) = a_r \sigma \left(1 + \frac{r}{\phi}\right) \left(1 + a_r \frac{\phi \sigma^2}{r}\right) p\left(\overline{p} - p\right) + o\left(a_r^2\right). \tag{51}$$

Proof. *See Appendix C.* ■

4.2.3 Implications of optimal contracting

We study the implication of the optimal contract now. As from now on we are always referring to the optimal policies, without risk of confusion we omit the superscript "*".

Value function and optimal policies Figure 1 plots the value function $V\left(p\right)$, the optimal control $\left\{\beta\left(p\right), \sigma^p\left(p\right)\right\}$ based on (50)-(51), and the associated optimal policy $\mu_t\left(p\right) = \beta_t\left(p\right) - p$ in solid lines. For comparison, in each panel we also plot the corresponding deterministic counterparts in dashed lines, and the Holmstrom and Milgrom (1987) benchmark in dotted lines.

The value delivered by the optimal stochastic contract must exceed the one under the deterministic counterpart, as shown in the top-left panel in Figure 1. The top-right panel plots the volatility of the agent's information rent, σ^p, which is zero by definition when the contract is restricted to be deterministic. A positive σ^p in the optimal stochastic contract implies that information rents rise after a good performance shock, an interesting property that will be discussed shortly.

It is the more efficient effort policy that drives the stochastic contract to be superior. In the bottom-left panel, we show that the incentive slope $\beta\left(p\right)$ sits above the deterministic counterpart for almost the entire range (except for low p's that are close to zero). A similar pattern holds for the implemented effort $\mu\left(p\right) = \beta\left(p\right) - p$ in the right-bottom panel of Figure 1. Overall, relative to

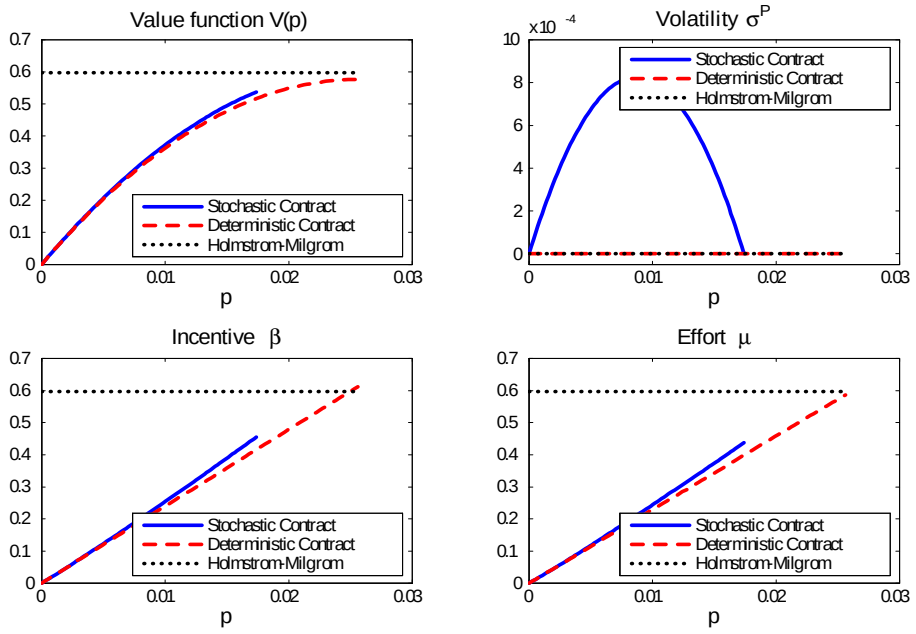

Figure 1: Value function and optimal policies in the optimal contract. Solid lines correspond to the optimal stochastic contract, dashed lines correspond to the optimal deterministic contract, and dotted lines correspond to the Holmstrom-Milgrom (1987) benchmark. The parameters are $r = 0.5, a = 0.6, \sigma = 1.5, \phi = 0.022$.

the optimal deterministic contract, the implemented effort is getting closer to the Holmstrom and Milgrom (1987) benchmark level.

Interestingly, though not evident in Figure 1, when p is close to zero both the incentive slope $\beta(p)$ and the effort $\mu(p)$ go below their deterministic counterparts. This result is a robust feature of the model. Indeed, with the aid of asymptotic analysis in (50), one can analytically verify that the adjustment term between the deterministic and stochastic contracts is negative by setting $p \simeq 0$. The seemingly counter-intuitive result is rooted in the "option-like" feature in the optimal contract, to which we turn next.

Option-like incentives in optimal contracting In our model, it is optimal to implement a history dependent effort policy. With a standard CARA-Normal setting and stationary learning, the Pareto frontier of the contractual relationship moves stochastically with the project profitability. However, without learning but with long-term contracting, the Holmstrom and Milgrom (1987)

benchmark suggests that a constant effort policy is optimal. Moreover, with learning but without long-term contracting, Holmstrom (1999) derives a deterministic effort policy along the equilibrium path. These two facts imply that the history-dependent effort policy is an outcome of the combination of long-term contracting and learning.

To understand the economic mechanism, we study how the history-dependent effort policies improve over the deterministic policies. To this end, we investigate the response of pay-performance sensitivity β and the implemented effort μ to unexpected shocks, which is captured by the diffusion term of $d\beta(p_t)$ (or, $d\mu(p_t)$), i.e., $\beta'(p_t)\sigma^p dB_t$ (or, $\left(\beta'(p_t)-1\right)\sigma^p dB_t$). As shown in the top panels in Figure 2, these diffusion terms are positive. Thus, in contrast to Holmstrom and Milgrom (1987) benchmark where the optimal contract features a constant equity share, with learning the optimal contract has an option-like feature. More specifically, a positive loading of incentives on output shocks implies that the pay-for-performance rises following good performance, i.e., the optimal contract is convex in output. The same pattern holds for the optimal effort policy.

The optimality of this option-like feature is a result of reducing the agent's information rents in a long-term relation, a mechanism requires both learning and long-term contracting. As explained in Section 2.4.3, the thrust of endogenous learning in dynamic contracting is that the agent can manipulate the principal's future belief downward by shirking today, and thus enjoy the potential information rents:

$$p_t = \frac{1}{u_c(c_t,\mu_t)}\mathbb{E}_t\left[\int_t^\infty \phi e^{-(\phi+r)(s-t)}\beta_s u_c(c_s,\mu_s)\,ds\right].$$

These information rents are additional future rewards to the agent because the principal mistakenly attributes the higher-profitability-driven good performance to the agent's effort. This explains that the future pay-performance sensitivities $\{\beta_s\}$ enter the agent's information rent. Equally important, for a risk-averse agent, the amount of information rents also depends on his marginal utilities $u_c(c_s,\mu_s)$ at these future states. Following a positive output shock, the agent becomes wealthier, which drives down his marginal utility $u_{c,s} = -arv_s$.[11] Then, by making the optimal

[11] Formally, we have the evolution of marginal utility to be $d(-arv_t) = -ar\beta_t(-arv_t)\,dB_t$, which has a negative diffusion on the performance shock.

contract option-like, the principal raises incentives after good performance and thus impose a negative correlation between pay-performance-sensitivity and marginal utility. Hence by allocating greater belief manipulation benefits in states where the agent cares less, the option-like feature lowers the agent's information rent standing today.

This option-like feature also explains the seemingly counter-intuitive result in the bottom-left panel of Figure 1, i.e., when p is close to zero the optimal stochastic contract implements a lower effort than the deterministic one. Here is the intuition. A positive diffusion of incentive β (effort μ) implies that the optimal contract allocates lower incentives in the states with poor historical performance (and hence high marginal utility). Because the information rents are positively correlated with performance as indicated by the top panels in Figure 2, the optimal contract impose lower incentives in the states with p close to zero.

Time-decreasing effort policies We also plot the drift of incentives β and effort μ in two bottom panels in Figure 2. As emphasized in Section 4.1.1, because of the effect of forward-looking information rents, the optimal effort policy is decreasing over time in the optimal deterministic contract. Graphically, this result is reflected by the negative drifts (dotted lines for deterministic contracts) in the bottom panels in Figure 2. The pattern of front-loaded efforts is a robust result in the optimal contract, as the bottom-right panel in Figure 2 shows that the effort policy in the optimal stochastic contract exhibits a negative drift as well.

5 Conclusion

We introduce profitability uncertainty into Holmstrom and Milgrom (1987) and study the optimal long-term contracting with endogenous learning. Although along the equilibrium path the principal and the agent hold the same belief about the project profitability, the agent's potential deviation by exerting effort below the recommended level leads to potential long-lasting belief divergence between both parties, leading to the "hidden information" problem. Utilize the convenient property of CARA preference, we show that the optimal contracting can be reformulated to a dynamic pro-

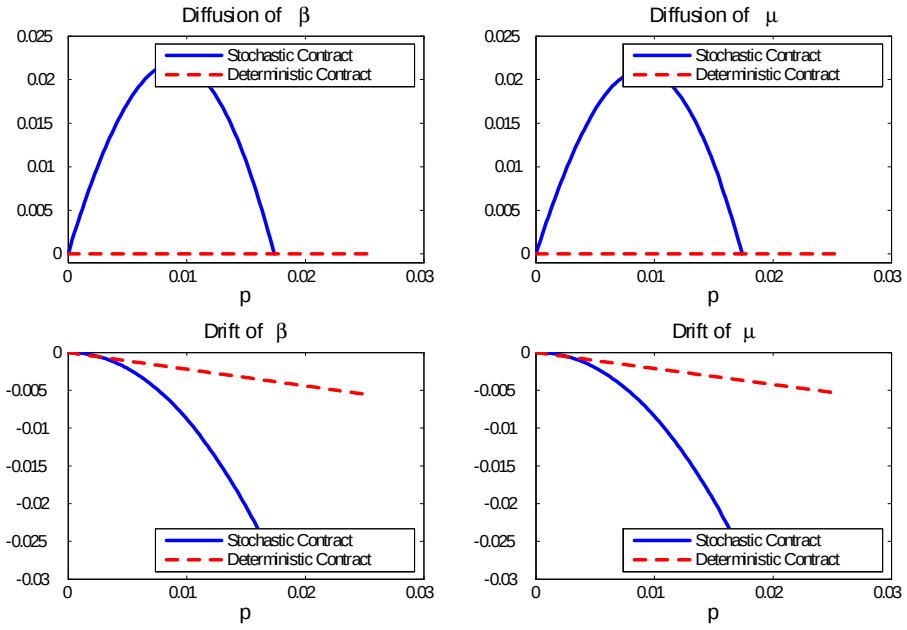

Figure 2: Diffusion and drift for incentives and effort in the optimal contract. Solid lines correspond to the optimal stochastic contract, and dashed lines correspond to the optimal deterministic contract. The parameters are $r = 0.5, a = 0.6, \sigma = 1.5, \phi = 0.022$.

gramming problem with only one state variable, which allows us to characterize the optimal contract by the solution to an Ordinary Differential Equation. Relative to the benchmark Holmstrom and Milgrom (1987) model with a constant effort level being optimal, we show that with learning, the optimal effort decreases with tenure on average. More importantly, the optimal contract exhibits an option-like feature, in the sense that incentives/effort rise after positive performance shocks.

References

Adrian, Tobias, and Mark M. Westerfield, 2009, Disagreement and learning in a dynamic contracting model, *Review of Financial Studies* 22-10, 3873–3906.

Cosimano, Thomas., Adam Speight, and Hayong Yun, 2011, Executive compensation in a changing environment: cautious investors and sticky contracts, working paper, University of Notre Dame.

Cole, H., and N. Kocherlakota, 2001, Efficient allocations with hidden income and hidden storage, *Review of Economic Studies* 68, 523-542.

DeMarzo, Peter M., and Yuliy Sannikov, 2008, Learning in dynamic incentive contracts, working paper, Stanford University and Princeton University.

Fleming, W. and R. Rishel, 1975, *Deterministic and stochastic optimal control*, Springer-Verlag, New York.

Fudenberg, D., Holmstrom, B., Milgrom, P., 1990. Short-term contracts and long-term agency relationships. *Journal of Economic Theory* 51, 1–31.

Gibbons, Robert and Kevin J. Murphy. 1992, Optimal incentive contracts in the presence of career concerns: theory and evidence." *Journal of Political Economy* 100(3): 468-505.

He, Zhiguo, 2011, A model of dynamic compensation and capital structure, *Journal of Financial Economics* 100, 351-366

He, Zhiguo, 2012, Dynamic compensation contracts with private savings, forthcoming in *Review of Financial Studies*.

Holmstrom, Bengt, 1999, Managerial incentive problems: a dynamic perspective, *Review of Economic Studies* 66, 169–182.

Holmstrom, Bengt, and Paul Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica* 55-2, 303–328.

Liptser, Robert S., and Albert N. Shiryaev, 1977, *Statistics of random processes II: Applications*, Volume 6 in Applications of Mathematics, Springer-Verlag, New York.

Prat, Julien, and Boyan Jovanovic, 2010, Dynamic contracts when agent's quality is unknown, working paper, NYU.

Sannikov, Yuliy, 2007, Agency problems, screening and increasing credit lines, working paper, Princeton.

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

Spear, Stephen E., and Sanjay Srivastava, 1987, On repeated moral hazard with discounting, *Review of Economic Studies* 54-4, 599–617.

Williams, Noah, 2009, On dynamic principal-agent problems in continuous time, working paper, University of Wisconsin-Madison.

Williams, Noah, 2011, Persistent private information, forthcoming in *Econometrica*.

Zhang, Y., 2009, Dynamic contracting with persistent shocks, *Journal of Economic Theory* 144, 635-675.

A Appendix A: Proofs

A.1 Proof for Lemma 1

The argument is similar to Cole and Kocherlakota (2000) and He (2011). Consider any contract $\Pi = \{c_t, \mu_t\}$ which induces an optimal policy $\{c_t^\star, \mu_t^\star\}$ from the agent with a value $v_0^\star$. The principal knows the resulting optimal effort policy $\{\mu_t^\star\}$ and she updates her belief according to $\{\mu_t^\star\}$ rather than the recommended effort policy $\{\mu_t\}$. From the agent's budget equation, we have

$$S_t = \int_0^t e^{r(t-s)} \left(c_s - c_s^\star\right) ds,$$

which gives the agent's optimal savings path, with the transversality condition $\lim_{T\to\infty} \mathbb{E}\left[e^{-rT} S_T\right] = 0$. Note that the transversality condition holds for all measures induced by any feasible effort policies.

By invoking the replication argument similar to revelation principle, we consider giving the agent a direct contract $\Pi^\star = \{c_t^\star, \mu_t^\star\}$; notice that 1) the principal will not change his Bayesian updating because she just changed her "recommended" effort policy which has no effect on the contract at all; and 2) the principal has the same cost to deliver this contract as $\Pi = \{c_t, \mu_t\}$. Clearly taking consumption-effort policy $\{c_t^\star, \mu_t^\star\}$ is feasible for the agent with no private-savings. Therefore we only need to show that $\{c_t^\star, \mu_t^\star\}$ is indeed optimal for the agent, because if it is true, then the principal still correctly knows that the agent's actual optimal effort policy $\{\mu_t^\star\}$, and her payoff is the same as that under the contract $\Pi = \{c_t, \mu_t\}$.

Suppose, conterfactually, that given the contract $\Pi^\star$, the agent finds that $\{c_t', \mu_t'\}$ yields a strictly higher payoff $v_0' > v_0^\star$ in her problem, with associated savings path

$$S_t' = \int_0^t e^{r(t-s)} \left(c_s^\star - c_s'\right) ds,$$

which satisfies transversality condition. However, consider the following saving path

$$S_t'' = S_t + S_t' = \int_0^t e^{r(t-s)} \left(c_s - c_s'\right) ds,$$

which also satisfies the transversality condition. But this implies that given the original contract Π, the saving rule $\{S_t''\}$ supports $\{c_t', \mu_t'\}$ but delivering a strictly higher payoff v_0'. This contradicts with the optimality of $\{c_t^\star, \mu_t^\star\}$ under the contract Π.

A.2 Proof for Lemma 2

Fix any constant S. Given any savings $S_t = S$ and a contract $\Pi = \{c\}$, from time-t on the agent's problem is

$$\begin{aligned} \max_{\{\widehat{c}_s\},\{\widehat{\mu}_s\}} \quad & \mathbb{E}^{\widehat{\mu}}\left[\int_t^\infty -\frac{1}{a} e^{-a\left(\widehat{c}_s - \frac{1}{2}\widehat{\mu}_s^2\right) - r(s-t)} ds\right] \\ s.t. \quad & dS_s = rS_s ds + c_s ds - \widehat{c}_s ds,\ S_t = S,\ s > t \\ & dY_s = \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right) dt + \sigma dB_s^{\widehat{\mu}}, \end{aligned} \tag{52}$$

given his information set. Note that the agent will learn actively. Denote by $\{c_s^\star, \mu_s^\star\}$ the solution to the above problem, and by $v_t\left(S;\Pi\right)$ the resulting agent's value.

Now consider the problem with $S = 0$, which is the continuation payoff along the equilibrium path:

$$\begin{aligned} \max_{\{\widehat{c}_s\},\{\widehat{a}_s\}} \quad & \mathbb{E}^{\widehat{\mu}}\left[\int_t^\infty -\frac{1}{a} e^{-a\left(\widehat{c}_s - \frac{1}{2}\widehat{\mu}_s^2\right) - r(s-t)} ds\right] \\ s.t. \quad & dS_s = rS_s ds + c_s ds - \widehat{c}_s ds,\ S_t = S,\ s > t \\ & dY_s = \left(\widehat{\mu}_t + m_t^{\widehat{\mu}}\right) dt + \sigma dB_s^{\widehat{\mu}}, \end{aligned}$$

We claim that the solution to this problem is $\{c_s^\star - rS, \mu_s^\star\}$, and therefore the value is $v_t\left(0;\Pi\right) = e^{arS} v_t\left(S;\Pi\right)$. There are two steps to show this. First, this solution is feasible. Second, suppose that there exists another policy $\left\{\widehat{c}_s', \widehat{\mu}_s'\right\}$ that is superior to $\{c_s^\star - rS, \mu_s^\star\}$, so that the associated value $v_t'\left(0;\Pi\right) > e^{-arS} v_t\left(S;\Pi\right)$. Consider $\left\{\widehat{c}_s' + rS, \widehat{\mu}_s'\right\}$, which is feasible to the problem in (52). But under this plan the agent's objective is

$$e^{-arS} \cdot \max \mathbb{E}_t^{\widehat{\mu}}\left[\int_t^\infty -\frac{1}{a} e^{-\gamma\left(\widehat{c}_s' - \frac{1}{2}\widehat{\mu}_s'^2\right) - r(s-t)} ds\right] = e^{-arS} v_t'\left(0;\Pi\right) > v_t\left(S;\Pi\right),$$

which contradicts with the optimality of $\{c_s^\star, \mu_s^\star\}$. As a result, $v_t\left(S;\Pi\right) = e^{-arS} v_t\left(0;\Pi\right)$.

A.3 Proof for Proposition 1

We apply the Pontryagin's Maximum Principle (MP). For this reason, we first consider the following agent's problem in finite horizon with terminal date T; we then will let T tend to infinity and impose transversality conditions. To preserve stationarity, we assume that the agent's utility function at

$$U\left(C_T\right) = -\frac{1}{ar}\exp\left(-arC_T\right) \tag{53}$$

which is as if the infinitely-lived agent is retired at T and consume rC_T afterwards. Clearly this treatment is innocuous as we take T to infinity. We have the following steps.

Step 1. Describing the optimal policy

We take a guess-and-verify approach. Consider the *proposed* policy $\{c^\star, \mu^\star\}$ with the following properties. First, define the agent's continuation payoff as

$$v_t^\star \equiv \mathbb{E}_t^{\mu^\star}\left[\int_t^T e^{-r(s-t)}\left(u\left(c_s^\star, \mu_s^\star\right)\right)ds + e^{-r(T-t)}U\left(C_T^\star\right)\right].$$

Then the consumption $c_t^\star$ satisfies

$$u\left(c_t^\star, \mu_t^\star\right) = rv_t^\star. \tag{54}$$

And, according to the martingale representation theorem, there exists a progressive process $\{\beta^\star\}$ so that

$$dv_t^\star = rv_t^\star - u\left(c_t^\star, \mu_t^\star\right)dt + \beta_t^\star\left(-arv_t\right)dB_t^{\mu^\star} = \beta_t^\star\left(-arv_t\right)dB_t^{\mu^\star}. \tag{55}$$

Given $\beta^\star$, we can rewrite the total reward from the optimal policy in its martingale representation form:

$$I_T^{c^\star,\mu^\star} = \int_0^T e^{-rt}u\left(c_t^\star, \mu_t^\star\right)dt + e^{-rT}U\left(C_T^\star\right) = V^{c^\star,\mu^\star}\left(0\right) + \int_0^T e^{-rt}\beta_t^\star\left(-arv_t^\star\right)\sigma dB_t^{\mu^\star}. \tag{56}$$

Then the effort $\mu_t^\star$ satisfies

$$\mu_t^\star = \beta_t^\star - \mathbb{E}_t^{\mu^\star}\left[\int_t^T e^{-(r+\phi)(s-t)}\beta_s^\star\frac{v_s^\star}{v_t^\star}ds\right]. \tag{57}$$

The rest of steps is trying to show that the above proposed policy is optimal.

Step 2. Rewriting the objective function.

Consider any alternative policy $\{c, \mu\}$, and define the associated reward process

$$I_t^{c,\mu} \equiv \mathbb{E}_t^{\mu}\left[\int_0^T e^{-rs}u\left(c_s, \mu_s\right)ds + e^{-rT}U\left(C_T\right)\right] = \int_0^t e^{-rs}u\left(c_s, \mu_s\right)ds + e^{-rt}v_t$$

where as before $v_t = \mathbb{E}_t^{\mu}\left[\int_t^T e^{-r(s-t)}u\left(c_s, \mu_s\right)ds + e^{-r(T-t)}U\left(C_T\right)\right]$. Define the local effort deviation by $\delta_t \equiv \mu_t - \mu_t^\star$ and recall that $\Delta_t = m_t^\mu - m_t^{\mu^\star} = \phi\int_0^t e^{\phi(s-t)}\left(\mu_s^\star - \mu_s\right)ds$. Let $A_t \equiv \int_0^t e^{\phi(s-t)}\mu_s ds$, $A_t^\star \equiv \int_0^t e^{\phi(s-t)}\mu_s^\star ds$, and $\Delta_t^A \equiv A_t - A_t^\star$. Then from (7) and (4) we know that

$$\sigma dB_t^{\mu^\star} = \sigma dB_t^{\mu} + \left[\delta_t - \phi\Delta_t^A\right]dt. \tag{58}$$

Therefore, the total reward from an arbitrary policy $\{c, \mu\}$ is given by

$$\begin{aligned}
I_T^{c,\mu} &= I_T^{c,\mu} - I_T^{c^\star,\mu^\star} + I_T^{c^\star,\mu^\star} \\
&= \int_0^T e^{-rt}\left[u\left(c_t, \mu_t\right) - u\left(c_t^\star, \mu_t^\star\right)\right]dt + e^{-rT}\left[U\left(C_T\right) - U\left(C_T^\star\right)\right] + I^{c^\star,\mu^\star}\left(T\right) \\
&= \int_0^T e^{-rt}\left[u\left(c_t, \mu_t\right) - u\left(c_t^\star, \mu_t^\star\right)\right]dt + e^{-rT}\left[U\left(C_T\right) - U\left(C_T^\star\right)\right] \\
&\quad + V^{c^\star,\mu^\star}\left(0\right) + \int_0^T e^{-rt}\beta_t^\star\left(-arv_t^\star\right)\left[\delta_t - \phi\Delta_t^A\right]dt + \int_0^T e^{-rt}\beta_t^\star\left(-arv_t^\star\right)\sigma dB_t^{\mu},
\end{aligned} \tag{59}$$

where in the last equality we have used Eq. (56) and Eq. (58).

In Eq. (59), the local deviation of μ_t will have a short-term effect on $\delta_t = \mu_t - \mu_t^\star$ and some long-lasting effect on $\Delta_t^A = \int_0^t e^{\phi(s-t)} (\mu_s - \mu_s^\star) ds$. We aim to rewrite Eq. (59) to tease out the total instantaneous effect of μ_t on $I_T^{c,\mu}$. To this end, define $n_t^\star$ as

$$n_t^\star \equiv \mathbb{E}_t^{\mu^\star} \left[- \int_t^T e^{-(r+\phi)(s-t)} \beta_t^\star (-arv_t^\star) dt \right] \tag{60}$$

According to martingale representation theorem, there exists some progressive process $\{\xi_t^\star\}$ so that

$$- \int_t^T e^{-(r+\phi)s} \beta_s^\star (-arv_s^\star) ds = e^{-(r+\phi)t} n_t^\star + \int_t^T \sigma e^{-(r+\phi)s} \xi_s^\star dB_s^{\mu^\star}; \tag{61}$$

or equivalently $\xi_t^\star$ is defined as the solution to the following equation:

$$dn_t^\star = \left[(r+\phi) n_t^\star + \beta_t^\star (-arv_t^\star) \right] dt + \sigma \xi_t^\star dB_t^{\mu^\star}.$$

It follows that (where we use (61) in the second equality):

$$\begin{aligned} - \int_0^T \beta_t^\star (-arv_t^\star) \phi \Delta_t^A dt &= \int_0^T \delta_t \phi e^{\phi t} \left(- \int_t^T e^{-\phi s} \beta_s^\star (-arv_s^\star) ds \right) dt \\ &= \int_0^T \delta_t \phi e^{\phi t} \left(e^{-(r+\phi)t} n_t^\star + \int_t^T \sigma e^{-(r+\phi)s} \xi_s^\star dB_s^{\mu^\star} \right) dt \\ &= \int_0^T \left(\delta_t \phi e^{-rt} n_t^\star + \delta_t \phi e^{\phi t} \int_t^T \sigma e^{-(r+\phi)s} \xi_s^\star dB_s^{\mu^\star} \right) dt \end{aligned}$$

Hence, we express the long-lasting effect of μ_t on Δ_t^A in terms of instantaneous effect on δ_t. The second term requires a bit more manipulation because $\sigma dB_s^{\mu^\star} = \sigma dB_t^\mu + \left[\delta_t - \phi \Delta_t^A \right] dt$ as in (58):

$$\begin{aligned} & \int_0^T \delta_t \phi e^{\phi t} \left[\int_t^T e^{-(r+\phi)s} \xi_s^\star \left(\sigma dB_s^\mu + \left(\delta_s - \phi \Delta_s^A \right) \right) \right] ds dt \\ = & \int_0^T \delta_t \phi e^{\phi t} \left[\int_t^T e^{-(r+\phi)s} \xi_s^\star \left(\delta_s - \phi \Delta_s^A \right) \right] ds dt + \text{Zero\_mean\_under\_}\mu \\ = & \int_0^T e^{-(r+\phi)t} \xi_t^\star \left(\delta_t - \phi \Delta_t^A \right) \left(\int_0^t \delta_s \phi e^{\phi s} ds \right) dt + \text{Zero\_mean\_under\_}\mu \\ = & \int_0^T e^{-rt} \xi_t^\star \left(\delta_t - \phi \Delta_t^A \right) \Delta_t^A dt + \text{Zero\_mean\_under\_}\mu, \end{aligned}$$

where the third inequality uses integration by parts, and the fourth equality uses the fact that $\Delta_t^A = \int_0^t \delta_s \phi e^{\phi(s-t)} ds$.

Thus, $V_0^{c,\mu} - V_0^{c^\star,\mu^\star}$, which is the difference between the expected reward from alternative policy and the one from proposed policy, is

$$\begin{aligned} & \mathbb{E}_0^\mu \left[I_T^{c,\mu} \right] - V_0^{c^\star,\mu^\star} \\ = & \mathbb{E}_0^\mu \left[\int_0^T e^{-rt} \left[u(c_t, \mu_t) - u(c_t^\star, \mu_t^\star) \right] dt + e^{-rT} \left[U(C_T) - U(C_T^\star) \right] \right] \\ & + \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt} \beta_t^\star (-arv_t) \left[\delta_t - \phi \Delta_t^A \right] dt + \int_0^\infty e^{-rt} \beta_t^\star (-arv_t) \sigma dB_t^\mu \right] \\ = & \mathbb{E}_0^\mu \left[\int_0^T e^{-rt} \left[u(c_t, \mu_t) - u(c_t^\star, \mu_t^\star) + \delta_t \left(\beta_t^\star (-arv_t^\star) + \phi n_t^\star \right) \right] dt + e^{-rT} \left(U(C_T) - U(C_T^\star) \right) \right] \\ & + \mathbb{E}_0^\mu \left[\int_0^T \phi e^{-rt} \xi_t^\star \Delta_t^A \left(\delta_t - \phi \Delta_t^A \right) dt \right] \end{aligned}$$

where the last inequality we use the change-of-measure formula in (58) again.

Step 3. Pontryagin's Maximum Principle (MP)

Our goal is to maximize $V_0^{c,\mu} - V_0^{c^\star,\mu^\star}$ which is equivalent to maximize $V_0^{c,\mu}$ directly. To this end, consider the following optimization problem:

$$\begin{aligned}
\max_{c,\mu} \quad & \mathbb{E}_0^\mu \left\{ \begin{array}{c} \int_0^T e^{-rt}\left[u\left(c_t,\mu_t\right) - u\left(c_t^\star,\mu_t^\star\right) + \delta_t\left(\beta_t^\star\left(-arv_t^\star\right) + \phi n_t^\star\right)\right] \\ + e^{-rT}\left[U\left(C_T^\star + S_T\right) - U\left(C_T^\star\right)\right] dt + \int_0^T \phi e^{-rt}\xi_t^\star \Delta_t^A\left(\delta_t - \phi\Delta_t^A\right) dt \end{array} \right\} \\
s.t. \Delta_t^A &= \int_0^t e^{\phi(s-t)}\left(\mu_s - \mu_s^\star\right) ds \text{ so that } d\Delta_t^A = \left(\mu_t - \mu_t^\star - \phi\Delta_t^A\right) dt, \\
dS_t &= \left(rS_t + c_t^\star - c_t\right) dt.
\end{aligned} \tag{62}$$

Note that here we treat $c_t^\star$ as the agent's compensation income from the contract because of Lemma 1; more importantly, we have imposed the agent's budget constraint because we have set $C_T = C_T^\star + S_T$ into the objective function (i.e., the agent have to clear his private saving account at terminal date T).

In the following steps, we first show that $\{c^\star, \mu^\star\}$ satisfies the maximum condition in the standard approach of Pontryagin's Maximum Principle (MP). We then give sufficient condition that the associated Hamiltonian is concave so that maximum condition is sufficient for optimality.

Denote the vector of state variables by $x_t \equiv \left(\Delta_t^A, S_t\right)^T$ whose evolution is given by

$$\begin{aligned}
dx_t &= \left(\begin{array}{c} \mu_t - \mu_t^\star - \phi\Delta_t^A \\ rS_t + c_t^\star - c_t \end{array} \right) dt \\
&= \left(\begin{pmatrix} -\phi & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \Delta_t^A \\ S_t \end{pmatrix} + \begin{pmatrix} \mu_t - \mu_t^\star \\ -\left(c_t - c_t^\star\right) \end{pmatrix} \right) dt
\end{aligned}$$

Denote the adjoint variables by $(\widehat{y}, \widehat{z})$. Given the optimal trajectory $\left(\overline{\Delta}_t^A, \overline{S}_t\right)$ and the optimal control $(\overline{\mu}_t, \overline{c}_t)$, the evolution for adjoint variables and their terminal conditions are

$$\begin{aligned}
d\begin{pmatrix} \widehat{y} \\ \widehat{z} \end{pmatrix} &= -\left\{ \begin{pmatrix} -\phi & 0 \\ 0 & r \end{pmatrix}^T \begin{pmatrix} \widehat{y} \\ \widehat{z} \end{pmatrix} + \begin{pmatrix} \phi e^{-rt}\xi_t^\star\left(\overline{\mu}_t - \mu_t^\star - 2\phi\overline{\Delta}_t^A\right) \\ 0 \end{pmatrix} \right\} dt + \begin{pmatrix} \widehat{Y} \\ \widehat{Z} \end{pmatrix} \sigma dB_t^\mu \\
\widehat{y}_T &= 0 \\
\widehat{z}_T &= e^{-rT} U'\left(C_T^\star + \overline{S}_T\right)
\end{aligned} \tag{63}$$

The Hamiltonian is defined as

$$\begin{aligned}
H\left(t,(\mu,c),(\widehat{y},\widehat{z})\right) &= e^{-rt}\left[u\left(c_t,\mu_t\right) - u\left(c_t^\star,\mu_t^\star\right) + \delta_t\left(\beta_t^\star\left(-arv_t^\star\right) + \phi n_t^\star\right)\right] \\
&\quad + \phi e^{-rt}\xi_t^\star \Delta_t^A\left(\delta_t - \phi\Delta_t^A\right) + \left(\mu_t - \mu_t^\star - \phi\Delta_t^A\right)\widehat{y}_t + \left(rS_t + c_t^\star - c_t\right)\widehat{z}_t,
\end{aligned} \tag{64}$$

and the maximum conditions are:

$$e^{-rt}\left[u_\mu\left(\overline{c}_t,\overline{\mu}_t\right) + \left(\left(-arv_t^\star\right) + \phi n_t^\star\right)\right] + \phi e^{-rt}\xi_t^\star\overline{\Delta}_t^A + \widehat{y}_t = 0 \tag{65}$$

$$e^{-rt} u_c\left(\overline{c}_t,\overline{\mu}_t\right) - \widehat{z}_t = 0 \tag{66}$$

Now we verify the proposed policy $(\mu^\star, c^\star)$ satisfies the above conditions with $\overline{c}_t = c_t^\star$, $\overline{\mu}_t = \mu_t^\star$, $\overline{\Delta}_t^A = 0$, and $\overline{S}_t = 0$.

For the adjoint variable of $\overline{\Delta}^A$, $\widehat{y}_t = 0$ is the solution to for the first equation in (63); intuitively, we have written the long-lasting effect of μ directly into the objective function so that the marginal value of $\overline{\Delta}^A$ becomes zero. Then, we need to verify that the proposed policy $(\mu^\star, c^\star)$ satisfies the maximum condition (65):

$$\begin{aligned}
& u_\mu\left(c_t^\star,\mu_t^\star\right) + \beta_t^\star\left(-arv_t^\star\right) + \phi n_t^\star \\
\Leftrightarrow \quad & \mu_t^\star\left(arv_t^\star\right) + \beta_t^\star\left(-arv_t^\star\right) + \phi\mathbb{E}_t^{\mu^\star}\left[-\int_t^T e^{-(r+\phi)(s-t)}\beta_t^\star\left(-arv_t^\star\right) dt\right] = 0.
\end{aligned}$$

where we have used the fact that $u_\mu\left(c_t^\star,\mu_t^\star\right) = -\mu_t^\star u_c\left(c_t^\star,\mu_t^\star\right) = \mu_t^\star arv_t^\star$ given CARA preference, and the definition of $n_t^\star$ in (60). But this is exactly the condition in (57).

For the adjoint variable of saving S_t, by taking the transformation $z_t = \widehat{z}e^{rt}$ and $Z = \widehat{Z}e^{rt}$, then we have

$$dz_t = Z_t \sigma dB_t^{\mu^\star} \text{ with } z_T = U'\left(C_T^\star\right), \text{ and } u_c\left(c_t^\star, \mu_t^\star\right) = z_t.$$

where the last condition is from the maximum condition in (66). We claim that

$$z_t = -arv_t^\star \text{ and } Z_t = \frac{1}{\sigma}\beta_t^\star a^2 r^2 v_t^\star dB_t^{\mu^\star}$$

satisfy the above equations. From Eq. (53) we have $z_T = -arv_T^\star = U'\left(C_T^\star\right) = -arU\left(C_T^\star\right)$ if we truncate the problem at T. And, the rest of claim can be easily verified from Eq. (54) and Eq. (55).

Combining the above steps, we have shown that $(\mu^\star, c^\star)$ satisfies the necessary conditions for the agent's problem in (62).

Step 4. Sufficiency

The incentive compatibility condition is only the necessary condition for the agent's problem. We now establish a sufficient condition under which the Hamiltonian in (64) is concave and thus the control $(c^\star, \mu^\star)$ is indeed optimal. It is easy to derive the gradient as

$$H_{x,u} = \begin{bmatrix} H_S \\ H_c \\ H_\mu \\ H_\Delta \end{bmatrix} = \begin{bmatrix} r\widehat{z}_t \\ e^{-rt}u_c\left(c_t, \mu_t\right) - \widehat{z} \\ e^{-rt}\left[u_\mu\left(c_t, \mu_t\right) + \left(\beta_t^\star\left(-arv_t^\star\right) + \phi n_t^\star\right)\right] + \phi e^{-rt}\xi_t^\star \Delta_t^A + \widehat{y} \\ \phi e^{-rt}\xi_t^\star\left(\left(\mu_t - \mu_t^\star\right) - 2\phi\Delta_t^A\right) - \phi\widehat{y} \end{bmatrix}$$

and the Hessian as

$$H_{x,u}^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & e^{-rt}u_{cc}\left(c_t, \mu_t\right) & e^{-rt}u_{\mu c}\left(c_t, \mu_t\right) & 0 \\ 0 & e^{-rt}u_{\mu c}\left(c_t, \mu_t\right) & e^{-rt}u_{\mu\mu}\left(c_t, \mu_t\right) & \phi e^{-rt}\xi_t^\star \\ 0 & 0 & \phi e^{-rt}\xi_t^\star & -2\phi^2 e^{-rt}\xi_t^\star \end{bmatrix}$$

The sufficient condition is one such that $H_{x,u}^2$ is negative semi-definite, or equivalently the following matrix is negative semi-definite:

$$\widetilde{H}_{x,u}^2 = \begin{bmatrix} u_{cc}\left(c_t, \mu_t\right) & 0 & 0 \\ 0 & u_{\mu\mu}\left(c_t, \mu_t\right) - \frac{u_{\mu c}(c_t,\mu_t)^2}{u_{cc}(c_t,\mu_t)} & \phi e\xi_t^\star \\ 0 & \phi\xi_t^\star & -2\phi^2\xi_t^\star \end{bmatrix}$$

Because $u_{cc}\left(c_t, \mu_t\right) < 0$, we only need to derive the condition for the following 2×2 matrix to be negative semi-definite:

$$\begin{bmatrix} u_{\mu\mu}\left(c_t, \mu_t\right) - \frac{u_{\mu c}(c_t,\mu_t)^2}{u_{cc}(c_t,\mu_t)} & \phi\xi_t^\star \\ \phi\xi_t^\star & -2\phi^2\xi_t^\star \end{bmatrix}$$

which is $-2\left[u_{\mu\mu}\left(c_t, \mu_t\right) - \frac{u_{\mu c}(c_t,\mu_t)^2}{u_{cc}(c_t,\mu_t)}\right] \geq \xi_t^\star \geq 0$.

A.4 Proof for Lemma 3

Suppose that the control variable is constrained such that $\beta_t \in [-M, M]$, where $M > 0$ is an arbitrarily large, but fixed constant. Define $M_p \equiv \frac{\phi M}{\phi + r}$. Recall the definition of p_t, and we have

$$\begin{aligned} p &= \mathbb{E}\left[\int_0^\infty \phi\beta_t e^{-(\phi+r)t} e^{\left(-\int_0^t ar\beta_v\sigma dB_v - \frac{1}{2}\int_0^t a^2r^2\beta_v^2\sigma^2 dv\right)} dt\right] \\ &\leq \mathbb{E}\left[\int_0^\infty \phi M e^{-(\phi+r)t} e^{\left(-\int_0^t ar\beta_v\sigma dB_v - \frac{1}{2}\int_0^t a^2r^2\beta_v^2\sigma^2 dv\right)} dt\right] = \frac{\phi M}{\phi + r} = M_p, \end{aligned}$$

where the equality is obtained only if $\beta_t = M$ for all t. Thus, at any time the feasible state variable p is bounded. Similarly, we can show that $p \geq -\frac{\phi M}{\phi + r} = -M_p$. Moreover, this result implies that whenever $p_t = \pm M_p$, we must have that for $\forall s \geq t$,

$$\beta_s = \pm M, p_s = \pm M_p, \text{and } \mu_s = \beta_s - p_s = \pm\frac{rM}{\phi + r}.$$

Therefore, once p_t hits $\pm M_p$, the state p_s will stay there from then on. In this sense $\pm M_p$ are absorbing boundaries.

This result helps us in deriving the value function at these boundaries. Since

$$V(p) = \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}\left(\mu_s - \frac{1}{2}\mu_s^2 - \frac{1}{2}a^2r^2\sigma^2\beta_s^2\right)ds\right],$$

we have

$$\begin{aligned} V(M_p) &= \int_0^\infty e^{-rt}\left(\frac{rM}{\phi+r} - \frac{1}{2}\left(\frac{rM}{\phi+r}\right)^2 - \frac{1}{2}a^2r^2\sigma^2M^2\right)ds \\ &= \frac{1}{\phi}M_p - \frac{r}{2\phi^2}M_p^2 - \frac{a^2r\sigma^2(\phi+r)^2}{2\phi^2}M_p^2, \end{aligned}$$

and similarly

$$V(-M_p) = -\frac{1}{\phi}M_p - \frac{r}{2\phi^2}M_p^2 - \frac{a^2r\sigma^2(\phi+r)^2}{2\phi^2}M_p^2.$$

Both of them are concave and quadratic in M. Finally, since M_p (M) can be arbitrarily large, $V(\pm M_p)$ can be arbitrarily small.

A.5 Derivation of Evoluation of p_t in Section 3.2.1

Let

$$P_t \equiv v_0\mathbb{E}_t\left[\int_t^\infty \phi\beta_s e^{-(\phi+r)s}\exp\left(-\int_0^s ar\beta_u\sigma dB_u - \int_0^s \frac{1}{2}(ar\beta_u\sigma)^2 du\right)ds\right],$$

then $p_t = e^{(\phi+r)t}\frac{P_t}{v_t}$. According to martingale representation theorem, there exists some progressively measurable process Φ_t^P so that

$$dP_t = -\phi e^{-(\phi+r)t}\beta_t v_t dt + \Phi_t^P dB_t.$$

Recall that $dv_t = -ar\beta_t\sigma v_t dB_t$; then we have $d\left(\frac{1}{v_t}\right) = \frac{1}{v_t}\left(ar\beta_t\sigma dB_t + (ar\beta_t\sigma)^2 dt\right)$. Hence, according to Ito's lemma,

$$\begin{aligned} d\left(\frac{P_t}{v_t}\right) &= P_t d\left(\frac{1}{v_t}\right) + \frac{1}{v_t}dP_t + d\left(\frac{1}{v_t}, P_t\right) \\ &= \frac{P_t}{v_t}\left(ar\beta_t\sigma dB_t + (ar\beta_t\sigma)^2 dt\right) - \phi e^{-(\phi+r)t}\beta_t dt + \frac{\Phi_t^P}{v_t}dB_t + \frac{ar\beta_t\sigma}{v_t}\Phi_t^P dt. \end{aligned}$$

Therefore, we have

$$\begin{aligned} dp_t &= e^{(\phi+r)t}d\left(\frac{P_t}{v_t}\right) + d\left(e^{(\phi+r)t}\right)\frac{P_t}{v_t} \\ &= e^{(\phi+r)t}\left[(\phi+r)\frac{P_t}{v_t}dt + \frac{P_t}{v_t}\left(ar\beta_t\sigma dB_t + (ar\beta_t\sigma)^2 dt\right) - \phi e^{-(\phi+r)t}\beta_t dt + \frac{\Phi_t^P}{v_t}dB_t + \frac{ar\beta_t\sigma}{v_t}\Phi_t^P dt\right] \\ &= (\phi+r)p_t dt + p_t\left(ar\beta_t\sigma dB_t + (ar\beta_t\sigma)^2 dt\right) - \phi\beta_t dt + e^{(\phi+r)t}\frac{\Phi_t^P}{v_t}dB_t + ar\beta_t\sigma e^{(\phi+r)t}\frac{\Phi_t^P}{v_t}dt \\ &= \left[(\phi+r)p_t - \phi\beta_t + ar\beta_t\sigma\underbrace{\left(ar\beta_t\sigma p_t + e^{(\phi+r)t}\frac{\Phi_t^P}{v_t}\right)}_{\text{defined as } \sigma^p}\right]dt + \underbrace{\left(ar\beta_t\sigma p_t + e^{(\phi+r)t}\frac{\Phi_t^P}{v_t}\right)}_{\text{defined as } \sigma^p}dB_t \\ &= \left[(\phi+r)p_t + \beta_t(ar\sigma\sigma^p - \phi)\right]dt + \sigma^p dB_t, \end{aligned}$$

which is equation (32).

A.6 Proof for Proposition 2

We first conjecture that the value function for the deterministic policy, $V^d(p)$, has the following quadratic form

$$V^d(p) = -\frac{1}{2}A^d p^2 + B^d p + C^d.$$

Plugging the above conjecture into the following ODE for the deterministic value function:

$$rV^d(p) = \frac{1}{2}\frac{\left(1+p-\phi V^d(p)\right)^2}{1+ar\sigma^2} - p - \frac{1}{2}p^2 + V_p^d(p)(\phi+r)p,$$

we can easily show that $B^d = \frac{1}{\phi}$, $C^d = 0$, and A^d satisfies

$$-\frac{1}{2}rA^d p^2 = \frac{1}{2}\frac{\left(1+\phi A^d\right)^2 p^2}{1+ar\sigma^2} - \frac{1}{2}p^2 - A^d(\phi+r)p^2.$$

Rearranging the above equation, we have

$$\phi^2\left(A^d\right)^2 - A^d\phi\left[\frac{r}{\phi}\left(1+ar\sigma^2\right) + 2ar\sigma^2\right] - ar\sigma^2 = 0,$$

which gives the solution for A^d,

$$A^d = \frac{1}{2\phi}\left[\frac{r}{\phi}\left(1+ar\sigma^2\right) + 2ar\sigma^2 + \sqrt{\left[\frac{r}{\phi}\left(1+ar\sigma^2\right)+2ar\sigma^2\right]^2 + 4ar\sigma^2}\right],$$

which is the same as that in the main text. Then, the optimal initial $p_0^d = \frac{B^d}{A^d}$ follows easily from the first order equation.

The incentive slope, as a function of information rent p_t, is

$$\beta_t^d = \frac{1+p_t - V_p^d\phi}{1+ar\sigma^2} = \frac{1+A^d\phi}{1+ar\sigma^2}p_t.$$

Using (32), we can derive the evolution of information rent p_t to be

$$\frac{dp_t^d}{p_t^d} = (\phi+r)\,dt - \frac{\beta_t^d}{p_t}\phi dt = \left(\phi + r - \frac{1+A^d\phi}{1+ar\sigma^2}\phi\right)dt \equiv -\lambda dt.$$

To show that $\lambda = \frac{1+A^d\phi}{1+ar\sigma^2}\phi - (\phi+r) > 0$, it is equivalent to show that $A^d > \left(1+\frac{r}{\phi}\right)\frac{ar\sigma^2}{\phi} + \frac{r}{\phi^2}$ which always holds by following the equation (74) in Lemma 5 (to be shown shortly). Finally, the optimal effort can be calculated as $\mu_t^d = \beta_t^d - p_t^d$.

A.7 Proof for Proposition 4

Before we solve the relaxed problem and derive the properties of the value function, we first define two useful functions, which prove to be useful in saving notations.

$$\begin{aligned} H_1(p) &\equiv \frac{1}{\phi}p - \frac{1}{2}\frac{r}{\phi^2}p^2 - \frac{a\sigma^2(\phi+r)^2}{2\phi^2}p^2, \\ H_2(p) &\equiv \frac{1}{r}\left[\frac{1}{2}\frac{(1+p)^2}{1+ar\sigma^2} - p - \frac{1}{2}p^2\right]. \end{aligned}$$

We state assumptions required for the proof below.

Assumption 1: The parameters satisfy the following condition:

$$a\left[2\left(\frac{\phi}{r}+1\right)^3 - \phi\sigma^2\right] < \frac{\phi}{r}. \tag{67}$$

Assumption 2: The feasible policy space for incentives is bounded given state p, i.e., $\beta(p) < \infty$ (though $\beta(p)$ may exceed M given M).

To solve the relaxed problem, we first focus on the following key ODE below:

$$\begin{aligned} rV(p) &= \frac{1}{2}\frac{(1+p-\phi V_p(p))^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{(V_p(p))^2}{V_{pp}}} - p - \frac{1}{2}p^2 + V_p(p)(\phi+r)p, \\ V(\pm M_p) &= H_1(\pm M_p). \end{aligned} \tag{68}$$

We proceed our proof with the following four steps.

Step 1: We first focus on ODE in equation (68). We show that at $p=0$, we have $V(0)=0$ and $V_p(0)=\frac{1}{\phi}$.

Step 2: Under assumption 2, this ODE in (68) satisfies concavity and positivity of the denominator condition.

(a), Prove the concavity and positivity of the denominator at $p=0$.

(b), Prove the concavity and positivity of the denominator for general $p>0$.

(c), Prove the concavity and positivity of the denominator for $p<0$.

(d), Prove that the lower boundary 0 is absorbing and the upper boundary $\bar{p}$ is entrance-no-exit.

Step 3: From Steps 2.a-2.d, it follows from a standard verification theorem that the solution to the ODE equation (68) is the value function of the principal's relaxed problem. Furthermore, it is never optimal to run outside the region $[0,\bar{p}]$ for the relaxed problem.

Step 4: Show that the value function for the relaxed problem is also the value function for the principal's original problem and $\bar{p}$ is independent of M for sufficiently large M.

A.7.1 Step 1: $V(0)=0$ and $V_p(0)=\frac{1}{\phi}$.

We proceed to show $V(0)=0$ and $V_p(0)=\frac{1}{\phi}$ by following the next three-step scheme.

Step 1.a: We first show that there must exist a solution to the following equation:

$$T(p) \equiv 1+p-\phi V_p = 0.$$

We know that $V(0)\geq 0$ which is the value of deterministic policy. We also know that

$$V(-M_p) = H_1(-M_p) < 0 \text{ and } V(M_p) = H_1(M_p) < 0.$$

Then according to the intermediate value theorem, there exists $p_1>0$ so that

$$T(p_1) = 1+p_1-\phi V_p(p_1) = 1+p_1-\phi\frac{V(M_p)-V(0)}{M_p} > 1,$$

and there also exists $p_2<0$ such that

$$T(p_2) = 1+p_2-\phi V_p(p_2) = 1+p_2-\phi\frac{V(0)-V(-M_p)}{M_p} < 0,$$

for sufficiently high M_p. Therefore, we can find a point $\underline{p}$ such that

$$1+\underline{p}-\phi V_p(\underline{p}) = 0. \tag{69}$$

Step 1.b: Suppose that $1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2(\underline{p})}{V_{pp}(\underline{p})} \neq 0$. We aim to show that $\underline{p}=0$ so that equations (69) and (68) imply $V(0)=0$ and $V_p(0)=\frac{1}{\phi}$. Differentiating the HJB in equation (68) with respect to p at $\underline{p}$, we have

$$rV_p(\underline{p}) = -1-\underline{p}+V_{pp}(\underline{p})(\phi+r)\underline{p}+V_p(\underline{p})(\phi+r),$$

which, together with equation (69), imply that

$$V_{pp}(\underline{p})\underline{p} = 0.$$

Therefore, either $\underline{p}=0$ or $V_{pp}\left(\underline{p}\right)=0$. We first rule out the case of $V_{pp}\left(\underline{p}\right)=0$. Note that this case implies the denominator of the first term in the right hand side of equation (68) is infinite, so that V must satisfy

$$rV\left(\underline{p}\right)=-\underline{p}-\frac{1}{2}\underline{p}^{2}+\left(1+\underline{p}\right)\left(1+\frac{r}{\phi}\right)\underline{p}, \tag{70}$$

Further, it follows from Taylor expansion that

$$\begin{aligned} V\left(\underline{p}+\epsilon\right) &= V\left(\underline{p}\right)+\frac{1}{\phi}\left(1+\underline{p}\right)\epsilon+o\left(\epsilon^{2}\right) \\ V_{p}\left(\underline{p}+\epsilon\right) &= \frac{1}{\phi}\left(1+\underline{p}\right)+o\left(\epsilon\right). \end{aligned} \tag{71}$$

Thus, evaluating the HJB equation (68) at $\underline{p}+\epsilon$, we have

$$\begin{aligned} rV\left(\underline{p}+\epsilon\right) &= \frac{\frac{1}{2}\left(\epsilon-o\left(\epsilon\right)\right)^{2}}{1+ar\sigma^{2}+a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}\left(\underline{p}+\epsilon\right)}{V_{pp}\left(\underline{p}+\epsilon\right)}}-\underline{p}-\epsilon-\frac{\left(\underline{p}+\epsilon\right)^{2}}{2}+\frac{1}{\phi}\left(1+\underline{p}\right)\left(\phi+r\right)\left(\underline{p}+\epsilon\right)+o\left(\epsilon^{2}\right) \\ &= -\underline{p}-\frac{1}{2}\underline{p}^{2}+\left(1+\underline{p}\right)\left(1+\frac{r}{\phi}\right)\underline{p}-\epsilon-\underline{p}\epsilon-\frac{1}{2}\epsilon^{2}+\left(1+\underline{p}\right)\left(1+\frac{r}{\phi}\right)\epsilon+o\left(\epsilon^{2}\right) \\ &= rV\left(\underline{p}\right)+\frac{r}{\phi}\left(1+\underline{p}\right)\epsilon-\frac{1}{2}\epsilon^{2}+o\left(\epsilon^{2}\right), \end{aligned}$$

where the second equality uses the fact $\frac{V_{p}^{2}\left(\underline{p}+\epsilon\right)}{V_{pp}\left(\underline{p}+\epsilon\right)}$ goes to infinity as ϵ goes to zero because the continuity of $V_{pp}\left(\underline{p}\right)$ implies that $V_{pp}\left(\underline{p}+\epsilon\right)$ is at the order of $o\left(1\right)$, and the last equality follows from equation (70). But this is in contradiction with (71) since they do not match at the second order ϵ^{2}. As a result, $\underline{p}=0$ and thus $V_{p}\left(0\right)=\frac{1}{\phi}$.

Step 1.c: Now suppose that $1+ar\sigma^{2}+a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}\left(\underline{p}\right)}{V_{pp}\left(\underline{p}\right)}=0$, i.e.,

$$V_{pp}\left(\underline{p}\right)=-\frac{a^{2}r^{2}\sigma^{2}V_{p}^{2}\left(\underline{p}\right)}{1+ar\sigma^{2}}=-\frac{a^{2}r^{2}\sigma^{2}}{1+ar\sigma^{2}}\frac{1}{\phi^{2}}\left(1+\underline{p}\right)^{2}, \tag{72}$$

which implies that we cannot ignore the term with $1+p-\phi V_{p}$. Due to L'Hopital's rule,

$$\frac{\left(1+p-\phi V_{p}\right)^{2}}{1+ar\sigma^{2}+a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}}{V_{pp}}}=\frac{2\left(1+p-\phi V_{p}\right)\left(1-\phi V_{pp}\right)}{a^{2}r^{2}\sigma^{2}\frac{2V_{p}V_{pp}^{2}-V_{p}^{2}V_{ppp}}{V_{pp}^{2}}}. \tag{73}$$

Differentiate the HJB equation (68), we have,

$$\begin{aligned} rV_{p} &= \frac{\left(1+\underline{p}-\phi V_{p}\right)\left(1-\phi V_{pp}\right)}{1+ar\sigma^{2}+a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}}{V_{pp}}}-\frac{1}{2}\frac{\left(1+\underline{p}-\phi V_{p}\right)^{2}}{\left(1+ar\sigma^{2}+a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}}{V_{pp}}\right)^{2}}\left[a^{2}r^{2}\sigma^{2}\frac{2V_{p}V_{pp}^{2}-V_{p}^{2}V_{ppp}}{V_{pp}^{2}}\right] \\ &\quad -1-\underline{p}+V_{p}\left(\phi+r\right)+V_{pp}\left(\phi+r\right)\underline{p}. \end{aligned}$$

Plugging in equation (73) into the above equation, we find that the two terms in the first line cancel each other, and

$$0=-1-\underline{p}+V_{p}\phi+V_{pp}\left(\phi+r\right)\underline{p}=V_{pp}\left(\phi+r\right)\underline{p},$$

which is the same as before. Therefore, either we have

$$\underline{p}=0,\text{ and }V_{p}\left(0\right)=\frac{1}{\phi},$$

or we have $V_{pp}\left(\underline{p}\right)=0$ and $\underline{p}=-1$ due to equation (72). In the second case, we further have

$$V_{pp}\left(-1\right)=0,V_{p}\left(-1\right)=0,\text{ and }\frac{V_{p}^{2}\left(-1\right)}{V_{pp}\left(-1\right)}=-\frac{1+ar\sigma^{2}}{a^{2}r^{2}\sigma^{2}}$$

In Lemma 7 below,[12] we will show that for any p, if

$$V_{pp}(p) = V_p(p) = 0 \text{ and } \lim \frac{V_p^2(p)}{V_{pp}(p)} \to -q,$$

then $q = 0$ or $\frac{1}{ar}$. Therefore, the second alternative results in contradiction and we must have $\underline{p} = 0$, and $V_p(0) = \frac{1}{\phi}$.

Now we still need to show that $V(0) = 0$ under the assumption $1 + ar\sigma^2 + a^2r^2\sigma^2\frac{V_p^2(\underline{p})}{V_{pp}(\underline{p})} = 0$. Suppose that $V(0) = v > 0$. First, the HJB in equation (68) implies that

$$\lim_{p\to 0} \frac{(1+p-\phi V_p)^2}{1 + ar\sigma^2 + a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} = 2rv > 0.$$

Thus, it follows from equation (36) that the policy function $\beta(p)$ at $p = 0$ has value

$$\begin{aligned}\beta(0) &= \lim_{p\to 0} \frac{(1+p-\phi V_p)^2}{1 + ar\sigma^2 + a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} \cdot \frac{1}{1+p-\phi V_p} \\ &= \lim_{p\to 0} 2rv\frac{1}{1+p-\phi V_p} = \infty.\end{aligned}$$

This contradicts with Assumption 2 that our policy space is restricted to be bounded at $p = 0$.

A.7.2 Step 2: Prove the concavity and positivity

Step 2.a: Concavity & Positivity of the Denominator at $p = 0$: We want to show that under assumption 1-2, the value function at $p = 0$ is concave, $V_{pp}(0) < 0$, and $1 + ar\sigma^2 + a^2r^2\sigma^2\frac{V_p(0)^2}{V_{pp}(0)} > 0$. Before we proceed to the main proof, we first show the following two lemmas, which are needed for our main proof.

Lemma 5 *Denote*

$$\chi^o \equiv \frac{A^o}{1/\phi + A^o} \text{ and } \chi \equiv \frac{A^d - \frac{a^2r^2\sigma^2(1/\phi)^2}{1+ar\sigma^2}}{1/\phi + A^d}.$$

Under the condition in equation (43), for $a > 0$, the following equalities hold:

$$\begin{aligned}A^d - A^o &> \frac{ar\sigma^2}{\phi}\left(1+\frac{r}{\phi}\right) > \frac{2a^2\sigma^2\phi}{r(1+ar\sigma^2)}\left(1+\frac{r}{\phi}\right)^4, \\ \chi^o &< \chi.\end{aligned} \tag{74}$$

Proof. Recall the definition of A^d and A^o, we have $A^o = \frac{r}{\phi^2}$, and

$$\begin{aligned}A^d &= \frac{1}{2\phi^2}\left((2\phi+r)ar\sigma^2 + r + \sqrt{(2\phi+r)^2a^2r^2\sigma^4 + 2\sigma^2\left[(\phi+r)^2+\phi^2\right]ar + r^2}\right) \\ &= \frac{1}{\phi}\left(\frac{r}{2\phi}\left(1+ar\sigma^2\right) + ar\sigma^2 + \sqrt{\left[\frac{r}{2\phi}(1+ar\sigma^2) + ar\sigma^2\right]^2 + ar\sigma^2}\right).\end{aligned}$$

Because $\sqrt{\left(\frac{r}{2\phi}(1+ar\sigma^2) + ar\sigma^2\right)^2 + ar\sigma^2} > \frac{r}{2\phi}\left(1+ar\sigma^2\right) + ar\sigma^2$, it follows

$$A^d - A^o > \frac{1}{\phi}\left(\frac{r}{\phi}\left(1+ar\sigma^2\right) + 2ar\sigma^2 - \frac{r}{\phi}\right) > \frac{ar\sigma^2}{\phi}\left(1+\frac{r}{\phi}\right). \tag{75}$$

[12] It is important to point out that the proof for Lemma 7 below does not use any results from step 1 here. Thus, there is no circular argument.

Furthermore, assumption 1 (i.e., equation (43)) implies that

$$\begin{aligned}
& \frac{\phi}{r}+a\phi\sigma^2>2a\left(\frac{\phi}{r}+1\right)^3 \Leftrightarrow \left(\frac{r}{\phi}\right)^3\left(\frac{\phi}{r}+a\sigma^2\phi\right)>2a\left(1+\frac{r}{\phi}\right)^3 \\
\Leftrightarrow \quad & \left(\frac{r}{\phi}\right)^2\left(1+ar\sigma^2\right)>2a\left(1+\frac{r}{\phi}\right)^3 \\
\Leftrightarrow \quad & \frac{ar\sigma^2}{\phi}\left(1+\frac{r}{\phi}\right)>\frac{2a^2\sigma^2\phi}{r\left(1+ar\sigma^2\right)}\left(1+\frac{r}{\phi}\right)^4. \qquad (76)
\end{aligned}$$

Thus, combining equation (75) and equation (76), we complete the proof for equation (74).

Lastly, we want to show $\chi^o<\chi$. First, the equality $2\left(1+\frac{r}{\phi}\right)^3>\left(\frac{r}{\phi}\right)^3$ implies that

$$\frac{2a^2\sigma^2\phi}{r\left(1+ar\sigma^2\right)}\left(1+\frac{r}{\phi}\right)^4>\left(1+\frac{r}{\phi}\right)\frac{a^2r^2\sigma^2\left(1/\phi\right)^2}{1+ar\sigma^2}.$$

Combining with the equality in equation (74), we have

$$A^d>A^o+\left(1+\frac{r}{\phi}\right)\frac{a^2r^2\sigma^2\left(1/\phi\right)^2}{1+ar\sigma^2}.$$

Further, notice that

$$\begin{aligned}
\chi^o-1 \quad &< \quad \chi-1\Leftrightarrow -\frac{1}{1+\phi A^o}<\frac{-1-\frac{a^2r^2\sigma^2(1/\phi)}{1+ar\sigma^2}}{1+\phi A^d} \\
A^d \quad &> \quad A^o+\frac{a^2r^2\sigma^2\left(1/\phi\right)^2}{1+ar\sigma^2}\frac{\left(1/\phi+A^o\right)}{1/\phi}\Leftrightarrow A^d>A^o+\frac{a^2r^2\sigma^2\left(1/\phi\right)^2}{1+ar\sigma^2}\left(1+\frac{r}{\phi}\right).
\end{aligned}$$

Therefore, the inequality $\chi^o<\chi$ follows immediately. ■

Lemma 6 *Under Condition equation (43) and for $a>0$, there exists a unique root $x_0\in(0,\chi^o)$ to the following cubic polynomial:*

$$\begin{aligned}
G\left(x\right) \quad \equiv \quad & \left(A^d-x\left(\frac{1}{\phi}+A^d\right)\right)(1-x)^2+\frac{2x}{1+\phi A^d}\left(\left(A^d-x\left(\frac{1}{\phi}+A^d\right)\right)\left(1+ar\sigma^2\right)-a^2r^2\frac{\sigma^2}{\phi^2}\right) \\
& -\left(A^d-x\left(\frac{1}{\phi}+A^d\right)\right)+\frac{a^2r^2}{1+ar\sigma^2}\frac{\sigma^2}{\phi^2}. \qquad (77)
\end{aligned}$$

Proof. Firstly, $G\left(0\right)=\frac{a^2r^2}{1+ar\sigma^2}\frac{\sigma^2}{\phi^2}>0$. Secondly, notice that

$$\begin{aligned}
G'\left(x\right) \quad = \quad & -\left(\frac{1}{\phi}+A^d\right)(1-x)^2-2\left(A^d-x\left(\frac{1}{\phi}+A^d\right)\right)(1-x) \\
& +\frac{2}{1+\phi A^d}\left(\left(A^d-x\left(\frac{1}{\phi}+A^d\right)\right)\left(1+ar\sigma^2\right)-a^2r^2\frac{\sigma^2}{\phi^2}\right) \\
& -\frac{2x}{1+\phi A^d}\left(\frac{1}{\phi}+A^d\right)\left(1+ar\sigma^2\right)+\left(\frac{1}{\phi}+A^d\right).
\end{aligned}$$

Evaluate the above equation at $p=0$, we have

$$G'\left(0\right)=\frac{2A^d}{1+\phi A^d}\left(-\phi A^d+ar\sigma^2-\frac{a^2r^2\sigma^2}{A^d\phi^2}\right)-\frac{\frac{1}{\phi}+2A^d+\phi\left(A^d\right)^2}{1+\phi A^d}<0,$$

where the last inequality is due to the following fact

$$\phi A^d=\frac{r}{2\phi}\left(1+ar\sigma^2\right)+ar\sigma^2+\sqrt{\left[\frac{r}{2\phi}\left(1+ar\sigma^2\right)+ar\sigma^2\right]^2+ar\sigma^2}>ar\sigma^2.$$

Thirdly, we proceed to show that $G''(x) > 0$ for $x \leq \chi^o$. Note that the second derivative of $G(x)$ can be simplified to the following:

$$G''(x) = 4\left(\frac{1}{\phi} + A^d\right)\left(1 - x - \frac{1 + ar\sigma^2}{1 + \phi A^d}\right) + 2\left(A^d - x\left(\frac{1}{\phi} + A^d\right)\right). \tag{78}$$

From Lemma 5, we have $\chi^o < \chi$. Note that $\chi < \frac{A^d}{1/\phi + A^d}$, it follows that for $x \leq \chi^o < \chi$,

$$A^d - x\left(\frac{1}{\phi} + A^d\right) > 0.$$

Furthermore, for $x \leq \chi^o$, we have

$$1 - x - \frac{1 + ar\sigma^2}{1 + \phi A^d} \geq 1 - \chi_0 - \frac{1 + ar\sigma^2}{1 + \phi A^d} = \frac{\left(A^d - A^o\right) - \frac{ar\sigma^2}{\phi}\left(1 + \frac{r}{\phi}\right)}{\left(1/\phi + A^o\right)\left(1 + \phi A^d\right)} > 0,$$

where the last inequality is due to equation (74). Thus, both terms in equation (78) are positive, and hence $G''(x) > 0$.

Below we further show that $G(\chi^o) < 0$ holds. After tedious algebra, one can show

$$\begin{aligned} G(\chi^o) &= (1/\phi)(2\chi^o)\left\{\frac{A^d - A^o}{1/\phi + A^o}\left[-1 + \frac{\chi^o}{2} + \frac{1 + ar\sigma^2}{1 + \phi A^d}\right] + \frac{a^2 r^2 \sigma^2 (1/\phi)}{(1 + ar\sigma^2)}\left[\frac{1}{2\chi^o} - \frac{1 + ar\sigma^2}{1 + \phi A^d}\right]\right\} \\ &< (1/\phi)(2\chi^o)\left\{\frac{A^d - A^o}{1/\phi + A^o}\left[-1 + \frac{\chi^o}{2} + \frac{1 + ar\sigma^2}{1 + \phi A^d}\right] + \frac{a^2 r^2 \sigma^2 (1/\phi)}{(1 + ar\sigma^2)}\frac{1}{2\chi^o}\right\}. \end{aligned}$$

It follows from straightforward algebra that

$$\begin{aligned} 1 - \frac{\chi^o}{2} - \frac{1 + ar\sigma^2}{1 + \phi A^d} &= \frac{\frac{r}{2\phi} + \sqrt{\left[\frac{r}{2\phi} + \frac{ar\sigma^2}{\left(1 + ar\sigma^2\right)}\right]^2 + \frac{ar\sigma^2}{\left(1 + ar\sigma^2\right)^2}}}{\left(\frac{r}{2\phi} + 1\right) + \sqrt{\left[\frac{r}{2\phi} + \frac{ar\sigma^2}{\left(1 + ar\sigma^2\right)}\right]^2 + \frac{ar\sigma^2}{\left(1 + ar\sigma^2\right)^2}}} - \frac{\frac{r}{2\phi}}{1 + \frac{r}{\phi}} \\ &> \frac{\frac{r}{2\phi}}{1 + \frac{r}{2\phi}} - \frac{\frac{r}{2\phi}}{1 + \frac{r}{\phi}} > \frac{\left(\frac{r}{2}\right)^2}{(\phi + r)^2}. \end{aligned}$$

Together with equation (74), we obtain

$$\begin{aligned} G(\chi^o) &< (1/\phi)(2\chi^o)\left\{-\frac{A^d - A^o}{1/\phi + A^o}\frac{\left(\frac{r}{2\phi}\right)^2}{\left(1 + \frac{r}{\phi}\right)^2} + \frac{a^2 r^2 \sigma^2 (1/\phi)}{(1 + ar\sigma^2)}\frac{1 + \frac{r}{\phi}}{2\frac{r}{\phi}}\right\} \\ &= -(1/\phi)\frac{\chi^o r^2 \frac{1}{\phi}}{2\left(1 + \frac{r}{\phi}\right)^3}\left\{\left(A^d - A^o\right) - \frac{2a^2\sigma^2\phi}{r\left(1 + ar\sigma^2\right)}\left(1 + \frac{r}{\phi}\right)^4\right\} < 0. \end{aligned}$$

Now we have $G''(x) > 0$ for $x \leq \chi^o$, $G'(0) < 0$, $G(0) > 0$ and $G(\chi^o) > 0$. Hence, there must exist one point $x \in (0, \chi^o)$ such that $G(x) = 0$. We still need to show the uniqueness. We consider two cases: $G'(\chi^o) \leq 0$ and $G'(\chi^o) > 0$. For the first case, since $G''(x) > 0$ for $x \leq \chi^o$, we have $G'(x) < 0$ for all $x \in (0, \chi^o)$, and hence $G(x)$ is monotonically decreasing in $(0, \chi^o)$. Thus, there is a unique x such that $G(x) = 0$. For the second case, since $G''(x) > 0$ for $x \leq \chi^o$, there is a unique $x_1 \in (0, \chi^o)$ such that $G'(x_1) = 0$. For $x \in (0, x_1)$, $G(x)$ is strictly decreasing since $G'(x) < 0$ for $x \in (0, x_1)$. For $x \in (x_1, \chi^o)$, $G(x)$ is strictly increasing since $G'(x) > 0$ for $x \in (x_1, \chi^o)$. Thus, $G(x_1) < 0$ and there exists a unique $x \in (0, x_1) \subset (0, \chi^o)$ such that $G(x) = 0$. We complete our proof. ■

Given the above two lemmas, now we proceed to the main proof for step 2.a. Recall that the key ODE equation (68) is

$$rV = \frac{1}{2}\frac{(1 + p - \phi V_p)^2}{1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{V_p^2}{V_{pp}}} - p - \frac{1}{2}p^2 + V_p(\phi + r)p.$$

The value function in the deterministic case V^d satisfies the following ODE

$$rV^d = \frac{1}{2}\frac{\left(1+p-\phi V_p^d\right)^2}{1+ar\sigma^2} - p - \frac{1}{2}p^2 + V_p^d\left(\phi + r\right)p.$$

Thus, taking the difference of the above two questions, we have

$$V - V^d = \frac{1}{2r}\left[\frac{\left(1+p-\phi V_p\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} - \frac{\left(1+p-\phi V_p^d\right)^2}{1+ar\sigma^2}\right] + \left(\frac{\phi}{r}+1\right)p\left[V_p - V_p^d\right]$$

For $p > 0$, dividing both sides by p and letting p go to zero, we have

$$\lim_{p\downarrow 0}\frac{V-V^d}{p} = \lim_{p\downarrow 0}\frac{1}{2rp}\left[\frac{\left(1+p-\phi V_p\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} - \frac{\left(1+p-\phi V_p^d\right)^2}{1+ar\sigma^2}\right] + \left(\frac{\phi}{r}+1\right)\lim_{p\downarrow 0}\left(V_p - V_p^d\right) \tag{79}$$

Because $V(0) = V^d(0) = 0$, we have

$$\lim_{p\downarrow 0}\frac{V-V^d}{p} = \lim_{p\downarrow 0}\frac{V-0}{p} - \lim_{p\downarrow 0}\frac{V^d-0}{p} = V_p(0) - V_p^d(0) = \lim_{p\downarrow 0}\left(V_p - V_p^d\right). \tag{80}$$

Thus, combining equation (79) and equation (80) yields

$$-\frac{\phi}{r}\lim_{p\downarrow 0}\left(V_p - V_p^d\right) = \lim_{p\downarrow 0}\frac{1}{2rp}\left[\frac{\left(1+p-\phi V_p\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} - \frac{\left(1+p-\phi V_p^d\right)^2}{1+ar\sigma^2}\right]$$

Dividing both sides again by p and letting p go to zero, we have

$$-\frac{\phi}{r}\lim_{p\downarrow 0}\frac{V_p - V_p^d}{p} = \lim_{p\downarrow 0}\frac{1}{2rp^2}\left[\frac{\left(1+p-\phi V_p\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2}{V_{pp}}} - \frac{\left(1+p-\phi V_p^d\right)^2}{1+ar\sigma^2}\right]. \tag{81}$$

Noting that $V_p(0) = V_p^d(0) = \frac{1}{\phi}$, it follows from L' Hopital's rule that

$$\lim_{p\downarrow 0}\frac{V_p - V_p^d}{p} = \lim_{p\downarrow 0}\frac{V_p - (1/\phi)}{p} - \lim_{p\downarrow 0}\frac{V_p^d - (1/\phi)}{p} = V_{pp}(0) - V_{pp}^d(0), \tag{82}$$

and

$$\lim_{p\downarrow 0}\frac{\left(1+p-\phi V_p\right)^2}{p^2} = \lim_{p\downarrow 0}\frac{2\left(1+p-\phi V_p\right)\left(1-\phi V_{pp}\right)}{2p} = \left(1-\phi V_{pp}(0)\right)^2. \tag{83}$$

Further, plugging the expression for $V^d(p)$, we have

$$1+p-\phi V_p^d = 1+p-\phi\left(-A^d p + B^d\right) = \left(1+\phi A^d\right)p. \tag{84}$$

Substituting the above equations (82), (83), and (84) back into equation (81), we have

$$-\phi\left(V_{pp}(0) - V_{pp}^d(0)\right) = \frac{1}{2}\frac{\left(1-\phi V_{pp}(0)\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{(1/\phi)^2}{V_{pp}(0)}} - \frac{\left(1+\phi A^d\right)^2}{2\left(1+ar\sigma^2\right)}$$

Let's define constant $\hat{x}$ as

$$\hat{x} \equiv \frac{V_{pp}(0) - V_{pp}^d(0)}{\frac{1}{\phi}+A^d}.$$

Then, after tedious algebraic manipulation, one can show that $\hat{x}$ is a root of the 3-order polynomial $G(x)$, which is defined in equation (77) in Lemma 6. According to Lemma 6, there exists a unique positive root $x_0 \in (0, \chi^o)$. Thus, $\hat{x} = x_0$ is such a unique root. Further, Lemma 5 implies $\chi^o < \chi < \frac{A^d}{1/\phi+A^d}$, and hence

$$0 < V_{pp}(0) - V_{pp}^d(0) = x_0\left(1/\phi + A^d\right) < \chi\left(1/\phi + A^d\right) < A^d.$$

Therefore,

$$V_{pp}(0) < V_{pp}^{d}(0) + A^{d} = -A^{d} + A^{d} = 0.$$

Furthermore, since $x_0 < \chi$, we have $x_0\left(1/\phi + A^{d}\right) - A^{d} < \chi\left(1/\phi + A^{d}\right) - A^{d} < 0$, and hence,

$$\begin{aligned} 1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(0)}{V_{pp}(0)} &= 1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{(1/\phi)^{2}}{x_0\left(1/\phi + A^{d}\right) - A^{d}} \\ &> 1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{(1/\phi)^{2}}{\chi\left(1/\phi + A^{d}\right) - A^{d}} = 0. \end{aligned}$$

Thus, the denominator at $p = 0$ is always positive.

Step 2.b: Concavity & Positivity of the Denominator at $p > 0$: We want to prove for any $p > 0$, the following must hold:

$$1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(p)}{V_{pp}(p)} > 0 \text{ and } V_{pp}(p) < 0.$$

We prove the above inequalities by contradiction. Suppose that the above two inequalities fail to hold at some points. Denote the smallest point $\widehat{p} > 0$ at which at least one of the two inequalities does not hold. That is,

$$1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(p)}{V_{pp}(p)} > 0 \text{ and } V_{pp}(p) < 0 \text{ for } p \in [0, \widehat{p})$$

while at $\widehat{p}$, one of the following three cases holds:

$$\text{(Case 1) } 1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(\widehat{p})}{V_{pp}(\widehat{p})} = 0 \text{ and } V_{pp}(\widehat{p}) < 0$$

or

$$\text{(Case 2) } 1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(\widehat{p})}{V_{pp}(\widehat{p})} = 0 \text{ and } V_{pp}(\widehat{p}) = 0$$

or

$$\text{(Case 3) } 1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(\widehat{p})}{V_{pp}(\widehat{p})} > 0 \text{ and } V_{pp}(\widehat{p}) = 0.$$

We first show that Case 1 is impossible. Since $V_{pp}(p) < 0$ for $p \in [0, \widehat{p})$, we know that

$$T(p) \equiv 1 + p - \phi V_{p} \tag{85}$$

is strictly increasing in p and positive. To see this,

$$T'(p) = 1 - \phi V_{pp} > 0,\ T(0) = 0.$$

Therefore, $T(\widehat{p}) > 0$ and $T'(\widehat{p}) > 0$. This implies that

$$\frac{\left(1 + \widehat{p} - \phi V_{p}(\widehat{p})\right)^{2}}{1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_{p}^{2}(\widehat{p})}{V_{pp}(\widehat{p})}} = \infty,$$

which contradicts with the finiteness of V. Also we should use the fact that V_p is bounded due to $V \in \mathbb{C}^{2}$. By the same argument, Case 2 is impossible either.

Then we consider Case 3. Note that Case 3 occurs only in the situation that

$$V_{p}(p) \to 0^{+} \text{ and } V_{pp}(p) \to 0^{-} \text{ when } p \uparrow \widehat{p}$$

Otherwise, a non-zero $V_p(\widehat{p})$ implies $\frac{V_p^2(\widehat{p})}{V_{pp}(\widehat{p})} = -\infty$, a contradiction to the fact that $1 + ar\sigma^{2} + a^{2}r^{2}\sigma^{2}\frac{V_p^2(p)}{V_{pp}(p)} > 0$ for all $p < \widehat{p}$. Therefore, we can define a positive finite number $q \geq 0$ such that there exists some subsequence $\{p_n = \hat{p} - \epsilon_n\}$ approaching $\widehat{p}$ from left so that

$$\lim_{p_n \uparrow \widehat{p}} Q(p_n) \equiv \lim_{p_n \uparrow \widehat{p}} \frac{V_{p}^{2}(p_n)}{V_{pp}(p_n)} = -q \leq 0.$$

Moreover, we must have $\lim_{p\downarrow\widehat{p}}\frac{V_p^2(p)}{V_{pp}(p)}=-q\leq 0$ as well because otherwise $V(p)$ will exhibit a jump at $\widehat{p}$. Finally, we must have $1+ar\sigma^2-a^2r^2\sigma^2q>0$.

Taking difference on the key ODE equation (68) at $p=p_n$ and $p=\hat{p}$, for any sequence $\{p_n\}\to\hat{p}$ we have

$$\begin{aligned}&\left(1+\hat{p}-\phi V_p(\hat{p})\right)^2-\left(1+p_n-\phi V_p(p_n)\right)^2\\ =\;& 2\left[rV(\hat{p})+\hat{p}+\frac{1}{2}\hat{p}^2-V_p(\hat{p})(\phi+r)\hat{p}\right]\left(1+ar\sigma^2+a^2r^2\sigma^2\frac{\left(V_p(\hat{p})\right)^2}{V_{pp}(\hat{p})}\right)\\ &-2\left[rV(p_n)+p_n+\frac{1}{2}p_n^2-V_p(p_n)(\phi+r)p_n\right]\left(1+ar\sigma^2+a^2r^2\sigma^2\frac{\left(V_p(p_n)\right)^2}{V_{pp}(p_n)}\right).\end{aligned}$$

Simplifying the above equation further, we obtain

$$\begin{aligned}(1+\hat{p}) \;=\;& (1+\hat{p})\left(1+ar\sigma^2+a^2r^2\sigma^2Q(\hat{p})\right)+o_n(1)\\ &+\left[\frac{1}{2}\frac{\left(1+p_n-\phi V_p(p_n)\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2(p_n)}{V_{pp}(p_n)}}\right]\left(a^2r^2\sigma^2\frac{Q(p)-Q(p_n)}{p-p_n}\right).\end{aligned}$$

Rearranging the above equation and noticing that $1+ar\sigma^2-a^2r^2\sigma^2q>0$, we have

$$arq-1=\frac{1}{2}\frac{1+\hat{p}}{1+ar\sigma^2-a^2r^2\sigma^2q}ar\left(\lim_{n\to\infty}\frac{Q(\hat{p})-Q(p_n)}{\hat{p}-p_n}\right). \tag{86}$$

Thus, $\lim_{n\to\infty}\frac{Q(\hat{p})-Q(p_n)}{\hat{p}-p_n}$ must exist and be finite.

First, suppose $q\neq 0$, then we know $\lim_{n\to\infty}\frac{V_{pp}(p_n)}{V_p^2(p_n)}=-\frac{1}{q}$, which is finite. Thus, it follows from $\frac{V_{pp}(\widehat{p}-\epsilon_n)}{V_p^2(\widehat{p}-\epsilon_n)}=-\frac{1}{q}+o_n(1)$ that

$$\begin{aligned}V_{pp}(p_n) \;=\;& -\frac{1}{q}V_p^2(p_n)+V_p^2(p_n)o_n(1)\\ =\;& -\frac{1}{q}2V_p(\widehat{p}_n)V_{pp}(\widehat{p}_n)\epsilon_n+2V_p(\widehat{p})V_{pp}(\widehat{p}_n)\epsilon_n o_n(1),\end{aligned}$$

where the second equality is due to the mean-value theorem and $\hat{p}\geq\hat{p}_n\geq p_n$. Therefore, for any sequence $\{p_n\}\to\hat{p}$ we have

$$V_{ppp}(\widehat{p})\equiv\lim_{n\to\infty}\frac{V_{pp}(\widehat{p}-\epsilon_n)}{\epsilon_n}=\lim_{n\to\infty}-\frac{1}{q}2V_p(\widehat{p}_n)V_{pp}(\widehat{p}_n)+2V_p(\widehat{p})V_{pp}(\widehat{p}_n)o_n(1)=0.$$

Thus, $V_{ppp}(\widehat{p})$ exists, and hence using $Q(p)=\frac{V_p^2(p_n)}{V_{pp}(p_n)}$ the equation (86) implies that

$$1-arq=\frac{1}{2}\frac{1+\widehat{p}}{1+ar\sigma^2-a^2r^2\sigma^2q}\left[ar\frac{V_p^2(\widehat{p})}{V_{pp}(\widehat{p})}\frac{V_{ppp}(\widehat{p})}{V_{pp}(\widehat{p})}\right]. \tag{87}$$

Below we show that there is a contradiction for Case 3 by the following three lemmas.

Lemma 7 *It must be that either $q=0$ or $q=\frac{1}{ar}$.*

Proof. *Assume that $q\neq 0$. Recall q is defined so that*

$$\lim_{p\uparrow\widehat{p}}\frac{V_p^2}{V_{pp}}=-q<0,$$

which implies that

$$\lim_{p\uparrow\widehat{p}}\frac{V_p^2}{V_{pp}}=\lim_{p\uparrow\widehat{p}}\frac{2V_pV_{pp}}{V_{ppp}}=\lim_{p\uparrow\widehat{p}}2V_p\cdot\frac{V_{pp}}{V_{ppp}}=-q,$$

Hence, we have

$$\lim_{p\uparrow\widehat{p}}\frac{V_{pp}}{V_{ppp}}=-\infty,\ \textit{and}\ \lim_{p\uparrow\widehat{p}}\frac{V_{ppp}}{V_{pp}}=0^-$$

From equation (87) and $q\neq 0$, we have

$$\lim_{p\uparrow\widehat{p}}\frac{V_p^2}{V_{pp}}\frac{V_{ppp}}{V_{pp}}=0,$$

and hence $q=\frac{1}{ar}$. As a result, $q=0$ or $q=\frac{1}{ar}$. ■

Lemma 8 *It is impossible to have* $q = \frac{1}{ar}$.

Proof. *Suppose* $q = \frac{1}{ar}$. *By plugging* $V_p(\widehat{p}) = V_{pp}(\widehat{p}) = 0$ *and* $q = \frac{1}{ar}$ *into the key ODE equation (68), we obtain*

$$V(\widehat{p}) = \frac{1}{2r} > V^{HM} = \frac{1}{2r}\frac{1}{1+ar\sigma^2},$$

which contradicts with the fact that our value function should be below the one derived from Holmstrom and Milgrom (1987). ■

Lemma 9 *It is impossible to have* $q = 0$ *either.*

Proof. *Suppose that* $q = 0$. *From equation (86), we have*

$$-1 = \frac{1}{2}\frac{(1+\hat{p})}{1+ar\sigma^2}\left(ar\lim_{n\to\infty}\frac{Q(\hat{p}) - Q(p_n)}{\hat{p} - p_n}\right),$$

which implies that $\lim_{n\to\infty}\frac{Q(\hat{p})-Q(p_n)}{\hat{p}-p_n} < 0$. *Using this property, we can obtain contradiction. To see this,* $Q(\hat{p}) = \lim_{n\to\infty}\frac{V_p^2(p_n)}{V_{pp}(p_n)} = -q = 0$ *at* $\widehat{p}$. *However, because* $\widehat{p}$ *is the first point so that* $V_{pp}(\widehat{p}) = 0$, *we know that* $V_{pp}(\widehat{p}-) < 0$ *which implies that*

$$Q(\widehat{p}-) = \frac{V_p^2(\widehat{p}-)}{V_{pp}(\widehat{p}-)} < 0.$$

As a result, $\lim_{n\to\infty}\frac{Q(\hat{p})-Q(p_n)}{\hat{p}-p_n} > 0$, *which is a contradiction.* ■

Combining the above three lemmas, we obtain that Case 3 is impossible. Thus, we complete the proof for Step 2.b. Since $V(0) = 0$, $V_p(0) = \frac{1}{\phi} > 0$, and $V(M_p) < 0$, and $V(p) < 0$ for $p \geq 0$, it follows that there exists $\overline{p} > 0$ such that $V'(\overline{p}) = 0$. The analysis above also implies that $V_{pp}(\overline{p}) < 0$ strictly. Therefore, the function is strictly concave and with positive denominator for $p \in [0, \overline{p})$. Before we proceed to Step 2.c, we show the following lemma on the property of $\overline{p}$, which is useful in the subsequent proof.

Lemma 10 $\overline{p} \leq \overline{p}^d$ *where* $\overline{p}^d \equiv \overline{p}_0^d = B^d/A^d$.

Proof. *From the key ODE equation (68), we know that at* $\overline{p}$,

$$V(\overline{p}) = H_2(\overline{p}) = \frac{1}{r}\left[\frac{1}{2}\frac{(1+\overline{p})^2}{1+ar\sigma^2} - \overline{p} - \frac{1}{2}\overline{p}^2\right].$$

Similarly, under the deterministic policy,

$$V^d\left(\overline{p}^d\right) = H_2\left(\overline{p}^d\right) = \frac{1}{r}\left[\frac{1}{2}\frac{\left(1+\overline{p}^d\right)^2}{1+ar\sigma^2} - \overline{p}^d - \frac{1}{2}\left(\overline{p}^d\right)^2\right].$$

It is straightforward that $H_2'(p) = \frac{1}{r}\left[\frac{1+p}{1+ar\sigma^2} - 1 - p\right] < 0$, *and hence to show* $\overline{p} \leq \overline{p}^d$, *it is enough to prove that* $V(p) > V^d(p)$. *Let's define function*

$$H(p) \equiv V(p) - V^d(p),$$

then the ODE equation (68) implies

$$rH(p) - p(\phi + r)H_p(p) = \frac{1}{2}\left[\frac{(1+p-\phi V_p(p))^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2(p)}{V_{pp}(p)}} - \frac{\left(1+p-\phi V_p^d(p)\right)^2}{1+ar\sigma^2}\right]. \tag{88}$$

It follows from $\hat{x} = \frac{V_{pp}(0)-V_{pp}^d(0)}{1/\phi+A^d} > 0$ *that* $V_p(0+) > V_p^d(0+)$, *and hence initially* $V(p)$ *is above* $V^d(p)$, *i.e.,* $H(0+) > 0$ *and* $H(0) = 0$.

We show that $\overline{p} < \overline{p}^d$. *Suppose it is not true so that* $\overline{p} > \overline{p}^d$. *Since* V *is concave over* $[0, \overline{p}]$,

$$V_p\left(\overline{p}^d\right) > V_p(\overline{p}) = 0 = V_p^d\left(\overline{p}^d\right),$$

and hence,

$$H_p\left(\bar{p}^d\right) = V_p\left(\bar{p}^d\right) - V_p^d\left(\bar{p}^d\right) > 0.$$

In addition, if $\overline{p} > \bar{p}^d$, *then* $H\left(\bar{p}^d\right) = V\left(\bar{p}^d\right) - V^d\left(\bar{p}^d\right) < V\left(\bar{p}\right) - V^d\left(\bar{p}^d\right) = H_2\left(\bar{p}\right) - H_2\left(\bar{p}^d\right) < 0$. *Because* $H\left(0+\right) > 0$, *it must be that* V *crosses* V^d *at some* $\widehat{p}$ *before* $\bar{p}^d$, *so that*

$$H\left(\widehat{p}\right) = 0, \text{ and } H_p\left(\widehat{p}\right) < 0.$$

As a result, there exists another point $p_1 \in \left[\widehat{p},\ \bar{p}^d\right]$ *so that*

$$H\left(p_1\right) < 0,\ H_p\left(p_1\right) = 0 \text{ which implies that } V_p\left(p_1\right) = V_p^d\left(p_1\right).$$

However, it follows from equation (88) that at $p = p_1$,

$$\begin{aligned} rH\left(p_1\right) &= \frac{1}{2}\left[\frac{\left(1+p_1-\phi V_p\left(p_1\right)\right)^2}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2(p_1)}{V_{pp}(p_1)}} - \frac{\left(1+p_1-\phi V_p^d\left(p_1\right)\right)^2}{1+ar\sigma^2}\right] \\ &= \frac{\left(1+p_1-\phi V_p\left(p_1\right)\right)^2}{2}\left[\frac{1}{1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2(p_1)}{V_{pp}(p_1)}} - \frac{1}{1+ar\sigma^2}\right] > 0, \end{aligned}$$

which contradicts with $H\left(p_1\right) < 0$. ■

Step 2.c: Concavity & Positivity of the Denominator at $p < 0$: Here we want to show that, for any $p < 0$,

$$1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2\left(p\right)}{V_{pp}\left(p\right)} > 0 \text{ and } V_{pp}\left(p\right) < 0.$$

Again, we prove these inequalities by contradiction, with similar argument as in Step 2.b. Suppose that the above two inequalities fail to hold at some point. Denote the largest point $\widehat{p} < 0$ at which at least one of the two inequalities does not hold. That is,

$$1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2\left(p\right)}{V_{pp}\left(p\right)} > 0 \text{ and } V_{pp}\left(p\right) < 0 \text{ for } p \in (\widehat{p}, 0]$$

Because $V_{pp}\left(p\right) < 0$ for $p \in (\widehat{p}, 0]$ and $V_p\left(0\right) = \frac{1}{\phi} > 0$, we have

$$\begin{aligned} V_p\left(p\right) &> V_p\left(0\right) > 0 \text{ for } p \in (\widehat{p}, 0] \\ V_p\left(\widehat{p}\right) &\geq V_p\left(0\right) > 0. \end{aligned}$$

Similar to the argument in Step 2.b, one of the three cases below holds at $\widehat{p}$:

$$\text{(Case 1) } 1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2\left(\widehat{p}\right)}{V_{pp}\left(\widehat{p}\right)} = 0 \text{ and } V_{pp}\left(\widehat{p}\right) < 0$$

or

$$\text{(Case 2) } 1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2\left(\widehat{p}\right)}{V_{pp}\left(\widehat{p}\right)} = 0 \text{ and } V_{pp}\left(\widehat{p}\right) = 0$$

or

$$\text{(Case 3) } 1+ar\sigma^2+a^2r^2\sigma^2\frac{V_p^2\left(\widehat{p}\right)}{V_{pp}\left(\widehat{p}\right)} > 0 \text{ and } V_{pp}\left(\widehat{p}\right) = 0.$$

We show that all the three cases are impossible below.

Case 1. We show this case is impossible. Suppose this is true. Then it must be true that

$$\frac{V_p^2\left(\widehat{p}\right)}{V_{pp}\left(\widehat{p}\right)} = -\frac{1+ar\sigma^2}{a^2r^2\sigma^2}.$$

Moreover, for the value function to be bounded, it must be true that

$$1+\widehat{p}-\phi V_p\left(\widehat{p}\right) = 0.$$

Since $V_p(\widehat{p}) \geq V_p(0)$ the above equation however implies

$$\widehat{p} = \phi V_p(\widehat{p}) - 1 \geq 0,$$

which contradicts with $\widehat{p} < 0$.

Case 2. This case is impossible, because $V_p(\widehat{p}) > 0$ and $V_{pp}(\widehat{p}) = 0$ imply $\lim_{p \downarrow \widehat{p}} \frac{V_p^2(\widehat{p})}{V_{pp}(\widehat{p})} = -\infty$, which contradicts with $1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{V_p^2(\widehat{p})}{V_{pp}(\widehat{p})} = 0$.

Case 3. Note that Case 3 occurs only in the situation that

$$V_p(p) \to 0^+ \text{ and } V_{pp}(p) \to 0^- \text{ when } p \downarrow \widehat{p}.$$

Otherwise, a non-zero $V_p(\widehat{p})$ implies $\frac{V_p^2(\widehat{p})}{V_{pp}(\widehat{p})} = -\infty$, a contradiction to the fact that $1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{V_p^2(p)}{V_{pp}(p)} > 0$ for all $\widehat{p} < p < 0$. On the other hand, as shown above,

$$V_p(\widehat{p}) \geq V_p(0) > 0,$$

which contradicts with the fact that $V_p(p) \to 0^+$ when $p \downarrow \widehat{p}$.

Step 2.d: Absorbing Lower Boundary 0 and Entrance-no-exit Upper Boundary $\overline{p}$:

Lemma 11 *The lower bound $p = 0$ is absorbing. The upper boundary $\overline{p}$ is entry-no-exit, i.e., $\sigma_p = 0$ and $\mu_p < 0$. Therefore, $p_t \in [0, \overline{p}]$ always.*

Proof. We have shown that there exists an upper boundary $\overline{p}$ such that

$$V_p(\overline{p}) = 0.$$

This implies that $\beta(\overline{p}) = \frac{1+\overline{p}}{1+ar\sigma^2}$ and $\sigma^P(\overline{p}) = 0$, and

$$\begin{aligned} \left.\frac{dp}{dt}\right|_{p=\overline{p}} &= (\phi + r) p_t + \beta_t (ar\sigma\sigma^p - \phi) \\ &= (\phi + r) p_t - \beta_t \phi = \phi(\overline{p} - \beta(\overline{p})) + r p_t \\ &= \left(1 - r\overline{p}\frac{1}{\phi}\right) \frac{\phi}{(V_{pp}\phi - 1)} M. \end{aligned}$$

Since $V_{pp} < 0$, the above equation is negative if and only if $1 - r\overline{p}\frac{1}{\phi} > 0$. Hence, it follows from $A^d > r\left(\frac{1}{\phi}\right)^2$, (where $A^d > 0$ is the coefficient on the quadratic term for the deterministic value function), we have

$$1 - r\bar{p}^d \frac{1}{\phi} = 1 - r\left(\frac{1}{\phi}\right)^2 \frac{1}{A^d} > 0.$$

Therefore, it follows from Lemma 10 that $\bar{p} \leq \bar{p}^d$, and thus $1 - r\bar{p}\frac{1}{\phi} \geq 1 - r\bar{p}^d\frac{1}{\phi} > 0$.

By a similar argument, we can prove that $p = 0$ is absorbing, since

$$\left.\frac{dp}{dt}\right|_{p=\underline{p}} = \beta_t (ar\sigma\sigma^p - \phi) = 0.$$

■

This lemma is important because it implies that V is determined by the policy within the region $[0, \overline{p}]$. Because V is in C^2 and $V_{pp} < 0$ strictly, we know that the policy function

$$\beta(p) = \frac{1 + p - \phi V_p}{1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{V_p^2}{V_{pp}}}$$

is bounded. Therefore, we can pick some arbitrary M to bound it.

A.7.3 Step 3: Verification

Define the auxiliary gain process, as a function of the contract Π, to be

$$G_t(\Pi) = \int_0^t e^{-rs}\left((\beta_s - p_s) - \frac{1}{2}(\beta_s - p_s)^2 - \frac{1}{2}a^2r^2\sigma^2\beta_s^2\right)ds + e^{-rt}V(p).$$

Define τ as the hitting time of p reaches $\pm M_p$, which could be infinite. Obviously, $G_\tau(\Pi)$ is the actual payoff from the contract Π. For given t, it is easy to show that

$$\mathbb{E}_t\left[e^{rt}dG_t\right] = \left[\begin{array}{c} -rV(p) + (\beta_t - p_t) - \frac{1}{2}(\beta_t - p_t)^2 - \frac{ar\sigma^2}{2}\beta_t^2 \\ +V_p\left[(\phi + r)p_t + \beta_t\left(ar\sigma\sigma^P - \phi\right)\right] + \frac{1}{2}V_{pp}\left(\sigma_t^P\right)^2 \end{array}\right]dt + V_p\sigma_t^P dB_t.$$

Therefore,

$$dG_t = \mu_G(p)\,dt + e^{-rt}V_p\sigma_t^P dB_t.$$

Due to construction of ODE in the HJB equation, under the optimal policy Π^* we have $\mu_G(p) = 0$ while for other policies we have $\mu_G(p) \leq 0$. Also, we restrict the policy $\{\sigma_t^P\}$ is well-behaved (square integrable in the usual sense) so that $\int_0^\infty e^{-rt}V_p\sigma_t^P dB_t$ is a martingale. Therefore under the optimal contract $\mathbb{E}[G_\tau(\Pi^*)] = G_0(\Pi^*) = V(p_0)$.

Given any $T > 0$, we have

$$\begin{aligned}\mathbb{E}[G_\tau(\Pi)] &= \mathbb{E}\left[G_{T\wedge\tau}(\Pi) + \mathbf{1}_{T\leq\tau}\left(\int_T^\tau e^{-rs}\left((\beta_s - p_s) - \frac{1}{2}(\beta_s - p_s)^2 - \frac{1}{2}a^2r^2\sigma^2\beta_s^2\right)ds + e^{-r\tau}V(p_\tau)\right)\right] \\ &\leq G_0 + e^{-rT}\mathbb{E}\left[\int_T^\tau e^{-r(s-T)}\frac{1}{2}ds\right]\end{aligned}$$

where $\mathbb{E}\left[\int_T^\tau e^{-r(s-t)}\frac{1}{2}ds\right]$ is the first-best project value. Therefore let $T \to \infty$, we have $\mathbb{E}[G_\tau(\Pi)] \leq G_0 = V(p)$. This implies that the proposed conract solves the relaxed problem.

A.7.4 Step 4: $\bar{p}$ is independent of M

Now we show that $\bar{p}$ is independent of M. Take some sufficiently large M_1, and consider the solution obtained with the upper entry-no-exit boundary $\bar{p}_1$. Note that $\bar{p}_1 < M_1$ strictly, because $V(\bar{p}_1; M_1) > 0$ while at $M_{p,1} = \left(1 + \frac{r}{\phi}\right)M_1$ the value is strictly negative. And, for $p \in [0, \bar{p}_1]$, we have

$$V(p; M_1) = \mathbb{E}\left[\int_0^\infty e^{-rt}\left((\beta_t - p_t) - \frac{(\beta_t - p_t)^2}{2} - \frac{1}{2}a^2r^2\sigma^2\beta_t^2\right)dt\middle| p_0 = p\right].$$

It is clear that given $\bar{p}_1$, this function is independent of M_1, because under the optimal policy $p \in [0, \bar{p}_1]$, and M does not affect the flow payoff per se (note that β_t is assumed to be unconstrained in the relaxed problem).

Now consider $M_2 \in (\bar{p}_1, M_1)$. The next lemma follows.

Lemma 12 *$V(p; M_1) \geq V(p; M_2)$ for $p \in [0, \bar{p}_1]$.*

Proof. *Denote the corresponding M_p's as $M_{p,i}$. Since $M_1 > M_2$ and $M_{p,1} > M_{p,2}$, the policy space for semi-constrained problem with M_1 is strictly larger than the policy space in problem with M_2. To see this, note that for the problem 1 (with M_1), the principal can choose $\beta_s = \left(1 + \frac{r}{\phi}p\right)$ for $s > t$ once $p_t \in [M_{p,2}, M_{p,1}]$, which is exactly the constraint for the policy space of problem 2 with M_2. As a result, $V(p; M_2) \leq V(p; M_1)$ for $p \in [0, \bar{p}_1]$.* ■

However, given M_2, consider the exact same policy under M_1 with endogenous upper entry-no-exit boundary $\bar{p}_1$, which generates the same value as $V(p; M_1)$. As a result, the policy under M_1 also solves the problem with M_2. Therefore we must have the same solution for both M_i's, and $\bar{p}_1 = \bar{p}_2$. QED.

A.8 Proof for Theorem 1

We now show that the relaxed problem solves the original problem. As explained before, our original problem in (33) has more stringent constraints than the relaxed problem: in the original problem we require $\beta_t \leq M$ always, while for the relaxed problem we only require that $\beta_t \leq M$ whenever p_t hits $\pm M_p$. As a result, we have $V(p) \geq V^C(p)$ always. Here, $V^C(p)$ denotes the value for the principal's original problem.

To show our theorem, it suffices to show that in the region $[0, \overline{p}]$, we have $V(p) = V^C(p)$, i.e., the relaxed problem and the original problem achieve the same value for sufficiently high M. Take the solution $V(p)$ and its corresponding incentive policy $\beta^M(\cdot)$; and define

$$B(M) \equiv \max_{0 \leq p \leq \overline{p}} \left| \beta^M(p) \right|,$$

where $\beta^M(p)$ emphasizes the possibility of the dependence of the optimal policy on the parameter value M. If we can show that we can choose sufficiently high M so that $B(M) \leq M$ holds, then the additional constraints are never binding in the original problem, and then both problems share the same solution obtained in Proposition 4.

We show that $B(M)$ is independent of M for sufficiently high M, which immediately implies our result. To show this, it is sufficient to show that both the relaxed value function $V(p)$ and $\overline{p}$ are independent of M. We have shown that $\overline{p}$ is independent of M when M is sufficiently high. Moreover, since the endogenous state p never go outside the region $[0, \overline{p}]$, we have

$$V(p) = \mathbb{E}_0 \left[\int_0^\infty e^{-rt} \left((\beta_t - p_t) - \frac{(\beta_t - p_t)^2}{2} - \frac{1}{2} a^2 r^2 \sigma^2 \beta_t^2 \right) dt \middle| p_0 = p \right]$$

to be independent of M. As a result, $\max_{0 \leq p \leq \overline{p}} \left| \beta^M(p) \right|$ is independent of M, and our result follows.

B Appendix B: Non-stationary Learning

B.1 Formulating the problem

Suppose that the unknown parameter θ is a constant. Define $\phi_t \equiv \Sigma_t^\theta / \sigma^2$, hence

$$dm_t = \Sigma_t^\theta \frac{dY_t - (\mu_t + m_t)\, dt}{\sigma^2} = \phi_t \sigma \frac{dY_t - (\mu_t + m_t)\, dt}{\sigma} \equiv \phi_t \sigma dB_t^\mu, \text{ and } \phi_t = \frac{\phi_0}{1 + \phi_0 t}, \tag{89}$$

where we still use ϕ_t without risk of confusion. According to Prat and Jovanovic (2010), the agent's incentive compatibility constraint is

$$\begin{aligned} \mu_t &= \beta_t - \mathbb{E}_t \left[\int_t^T \phi_s \beta_s e^{-r(s-t)} e^{\left(- \int_t^s a r \beta_u \sigma dB_u - \frac{1}{2} \int_t^s a^2 r^2 \beta_u^2 \sigma^2 du \right)} ds \right] \\ &= \beta_t - \int_t^T \phi_s \beta_s e^{-r(s-t)} ds. \end{aligned}$$

where the second equation we invoke the restriction that $\{\beta\}$ being deterministic. Thus, the principal's problem can be written as (where T can take a value of infinity)

$$\begin{aligned} & \max_{\{\beta_t\}} \int_0^T e^{-rt} \left(\mu_t - \frac{1}{2} \mu_t^2 - \frac{1}{2} a r \sigma^2 \beta_t^2 \right) dt \\ s.t. \qquad & \mu_t = \beta_t - \int_t^T \phi_s \beta_s e^{-r(s-t)} ds. \end{aligned}$$

Define $p_t \equiv \int_t^T \phi_s \beta_s e^{-r(s-t)} ds = e^{rt} \int_t^T \phi_s \beta_s e^{-rs} ds$ so that

$$\mu_t = \beta_t - p_t,\ p_t' = -\phi_t \beta_t + r p_t \Rightarrow \beta_t = -\frac{1}{\phi_t} \left(p_t' - r p_t \right).$$

Under this transformation, the objective becomes $\max_{\{p_t\}} \int_0^T L(t, p_t, p_t')\, dt$, so that

$$L(t, p_t, p_t') \equiv e^{-rt} \left(-\frac{1}{\phi_t} \left(p_t' - r p_t \right) - p_t - \frac{1}{2} \left(\frac{1}{\phi_t} \left(p_t' - r p_t \right) + p_t \right)^2 - \frac{a r \sigma^2}{2} \left(\frac{1}{\phi_t} \left(p_t' - r p_t \right) \right)^2 \right).$$

with the constraint that $p_T = 0$.

B.2 Euler equation and its sufficiency for optimality

We have a standard problem of calculus of variation, and the Euler equation for this problem is:

$$L_p\left(s, p_s, p_s'\right) = \frac{dL_{p'}\left(s, p_s, p_s'\right)}{ds}. \tag{90}$$

Because

$$\begin{aligned}
L_p\left(s, p_s, p_s'\right) &= e^{-rs}\left(\frac{r}{\phi_s} - 1 + \left(\frac{1}{\phi_s}\left(p_s' - rp\right) + p_s\right)\left(\frac{r}{\phi_s} - 1\right) + ar\sigma^2\left(\frac{1}{\phi_s}\left(p_t' - rp\right)\right)\frac{r}{\phi_s}\right), \\
L_{p'}\left(s, p_s, p_s'\right) &= e^{-rs}\left[-\frac{1}{\phi_s} - \frac{1}{\phi_s}\left(\frac{1}{\phi_s}\left(p_s' - rp\right) + p_s\right) - ar\sigma^2\left(\frac{1}{\phi_s}\right)^2\left(p_s' - rp_s\right)\right],
\end{aligned}$$

conducting algebraic simplifications on the Euler equation (90) yields

$$0 = -\frac{1}{\phi_s}\left(2p_s' - rp + \frac{1}{\phi_s}\left(p_s'' - rp_s'\right)\right) - 2ar\sigma^2\left(\frac{1}{\phi_s}\right)\left(p_s' - rp_s\right) - ar\sigma^2\left(\frac{1}{\phi_s}\right)^2\left(p_s'' - rp_s'\right).$$

Therefore, from the above equality, we obtain that the optimal policy must satisfy:

$$\frac{1}{\phi_s}p_s'' = -2p_s' + \frac{1}{\phi_s}rp_s' + rp\left(1 + \frac{ar\sigma^2}{1 + ar\sigma^2}\right) \tag{91}$$

Combining with $p_T = 0$ one can solve the optimal path using initial condition p_0. Then maximizing over p_0 one can find the solution to the original problem.

Before we proceed to prove the main result in the next sebsection, we further show that the first-order optimality of Euler equation is sufficient for global optimality. Notice $L_{pp} < 0$ and $L_{p'p'} < 0$. Furthermore, for any x and y, we aim to show that

$$L_{pp}y^2 + 2L_{pp'}xy + L_{p'p'}x^2 \leq 0, \tag{92}$$

where the equality holds only for $x = y = 0$. Notice that

$$\begin{aligned}
e^{rs}\left[L_{pp}y^2 + 2L_{pp'}xy + L_{p'p'}x^2\right] &= -\left(\left(\frac{r}{\phi_s} - 1\right)^2 + ar\sigma^2\left(\frac{r}{\phi_s}\right)^2\right)y^2 \\
&+2xy\left(\left(\frac{1}{\phi_s}\left(\frac{r}{\phi_s} - 1\right) + ar\sigma^2\left(\frac{r}{\phi_s}\right)^2\right)\right) - x^2\left(1 + ar\sigma^2\right)\left(\frac{1}{\phi_s}\right)^2.
\end{aligned}$$

Thus, it is sufficient to show the following:

$$\left[\left(\frac{r}{\phi_s} - 1\right) + ar\sigma^2\frac{r}{\phi_s}\right]^2 < \left[\left(\frac{r}{\phi_s} - 1\right)^2 + ar\sigma^2\left(\frac{r}{\phi_s}\right)^2\right]\left(1 + ar\sigma^2\right).$$

The left-hand-side is $\left(\frac{r}{\phi_s}\left(1 + ar\sigma^2\right) - 1\right)^2$, while the right-hand-side is

$$\begin{aligned}
&\left[\left(1 + ar\sigma^2\right)\left(\frac{r}{\phi_s}\right)^2 - \frac{2r}{\phi_s} + 1\right]\left(1 + ar\sigma^2\right) \\
= &\left(1 + ar\sigma^2\right)^2\left(\frac{r}{\phi_s}\right)^2 - \frac{2r}{\phi_s}\left(1 + ar\sigma^2\right) + 1 + ar\sigma^2 \\
= &\left(\frac{r}{\phi_s}\left(1 + ar\sigma^2\right) - 1\right)^2 + ar\sigma^2 > LHS.
\end{aligned}$$

Thus, equation (92) holds, and Euler equation is both necessary and sufficient for optimality (Theorem 2.3 in Chapter one of Fleming and Rishel (1975)).

B.3 Time-decreasing effort policy

The following sequence of argument shows that the optimal initial information rent $p_0^* > 0$, and the associated effort policy μ_t^* is decreasing over time.

Step 1. From Euler equation we show that p never changes sign.

If $p(0) = p_0 > 0$ and $p(T) = 0$, but p turns negative somewhere inbetween. Then there there must exist some point $\widehat{t}$ so that

$$p\left(\widehat{t}\right) < 0,\ p'\left(\widehat{t}\right) = 0 \text{ but } p''\left(\widehat{t}\right) \geq 0$$

This contradicts with the Euler equation (91):

$$p''\left(\widehat{t}\right)\left(\frac{1}{\phi_{\widehat{t}}}\right) = rp\left(\widehat{t}\right)\left(1 + \frac{ar\sigma^2}{1+ar\sigma^2}\right) < 0.$$

Similarly we can show that if $p(0) < 0$ and $p(T) = 0$ then $p < 0$ always. Therefore p never changes sign.

Step 2. Now we prove that if $p_0 > 0$ then the optimal effort policy goes down with time. From equation (91), and $\mu_t = \beta_t - p_t = -\frac{1}{\phi_t}\left(p_t' - rp_t\right) - p_t$, we have

$$\mu_t' = -\frac{1}{\phi_t}p_t'' + \frac{1}{\phi_t}rp_t' - 2p_t' + rp_t = -rp_t\frac{ar\sigma^2}{1+ar\sigma^2} < 0,$$

since $p_t > 0$ always. Here we also used the fact that $\left(\frac{1}{\phi_t}\right)' = 1$ from (89). Similarly, if $p_0 < 0$, the optimal effort policy goes up with time.

Step 3. We show the optimal $p_0 > 0$. First, note that a positive p_0 strictly improve the principal's value over $p_0 = 0$. When $p_0 = 0$, it follows that $p(t) = 0$, $\forall t \in [0,T]$ satisfies the Euler equation, and hence it is optimal with zero value. Now suppose that $\beta_t = \epsilon$ always where $\epsilon > 0$ is sufficiently small, so that

$$\mu_t = \beta_t - \int_t^T \phi_s \epsilon e^{-r(s-t)}ds,\ p_t = \int_t^T \phi_s \epsilon e^{-r(s-t)}ds \text{ with } p_0 = \int_0^T \phi_s \epsilon e^{-r(s-t)}ds.$$

Then the value from this policy must be strictly positive, as the benefit $\int_0^T e^{-rt}\mu_t dt$ is in the order of ϵ while the cost $\int_0^T e^{-rt}\left(-\frac{1}{2}\mu_t^2 - \frac{1}{2}ar\sigma^2\beta_t^2\right)dt$ is in second order of ϵ^2. Because the constant policy might not be optimal, the policy that satisfies the Euler equation should be better, i.e., $V(p_0) > 0$. Finally we rule out $p_0 < 0$ being optimal. To see this, we know that when $p_0 < 0$ then the optimal effort policy $\{\mu_t\}$ is increasing over time. However,

$$\mu_T = -\frac{1}{\phi_T}\left(p_T' - rp_T\right) - p_T = -\frac{1}{\phi_T}p_T' \leq 0.$$

The last inequality follows from the fact that $p_t \leq 0$ for $t < T$ while $p_T = 0$. As a result, the optimal effort policy $\{\mu_t\}$ is negative always. Thus, if $p_0 < 0$, then the objective $\int_0^T e^{-rt}\left(\mu_t - \frac{1}{2}\mu_t^2 - \frac{1}{2}ar\sigma^2\beta_t^2\right)dt$ is negative as well.

Combining the above three steps, we complete the proof that the optimal effort policy is decreasing over time.

C Appendix C: Asymptotic Analysis

Expansion in the Deterministic Case

We first expand the value function in the deterministic case. We write the value function as ,

$$\begin{aligned} V^d(p) &= V^o(p) + a_r Q_1^d(p) + a_r^2 Q_2^d(p) + a_r^3 Q_3^d(p) + a_r^4 Q_4^d(p) + o\left(a_r^4\right), &(93)\\ \overline{p}^d &= \overline{p}^o - a_r q_1^d - a_r^2 q_2^d - a_r^3 q_3^d - a_r^4 q_4^d + o\left(a_r^4\right), &(94)\end{aligned}$$

and based on the HJB equation one can derive that

$$\begin{aligned} Q_1^d &= -\frac{\sigma^2}{2r}\left(1+\frac{r}{\phi}\right)^2 p^2, Q_2^d = \frac{\phi^2\sigma^4}{2r^3}\left(1+\frac{r}{\phi}\right)^2 p^2,\\ Q_3^d &= -\frac{\phi^4\sigma^6}{2r^5}\left(1+\frac{r}{\phi}\right)^2\left[1+\left(1+\frac{r}{\phi}\right)^2\right]p^2,\\ Q_4^d &= \frac{\phi^6\sigma^8}{2r^8}\left(1+\frac{r}{\phi}\right)^2\left[\frac{r^2}{\phi^2}\left(2+\frac{r}{\phi}\right)^2 + 5\left(1+\frac{r}{\phi}\right)^2\right]p^2. \end{aligned}$$

We can also show from the condition $V_p^d\left(\overline{p}^d\right)=0$ that

$$\begin{aligned} q_1^d &= \frac{\phi^3\sigma^2}{r^3}\left(1+\frac{r}{\phi}\right)^2,\ q_2^d=-\frac{\phi^5\sigma^4}{r^5}\left(1+\frac{r}{\phi}\right)^2\left[1+\left(1+\frac{r}{\phi}\right)^2\right] \\ q_3^d &= \frac{\phi^7\sigma^6}{r^7}\left(1+\frac{r}{\phi}\right)^2\left(\left(1+\frac{r}{\phi}\right)^2+\left[1+\left(1+\frac{r}{\phi}\right)^2\right]^2\right). \end{aligned}$$

The resulting optimal incentive $\beta^d(p)$ is given by

$$\beta^d=\left(1+\frac{r}{\phi}\right)p\left\{\begin{array}{c} 1+a_r\frac{\phi\sigma^2}{r}-a_r^2\frac{\phi^3\sigma^4}{r^3}\left[\frac{r^2}{\phi^2}+\left(1+\frac{r}{\phi}\right)\right] \\ +a_r^3\frac{\phi^5\sigma^6}{r^5}\left[\frac{r^4}{\phi^4}+\frac{r^2}{\phi^2}\left(1+\frac{r}{\phi}\right)+\left(1+\frac{r}{\phi}\right)\left(1+\left(1+\frac{r}{\phi}\right)^2\right)\right] \end{array}\right\}.$$

Expansion in the General Case

We consider the following approximations up to the fourth order of magnitude of a. That is,

$$\begin{aligned} V(p) &= V^o(p)+a_rQ_1(p)+a_r^2Q_2(p)+a_r^3Q_3(p)+a_r^4Q_4(p)+o\left(a_r^4\right), \\ \overline{p} &= \overline{p}^o-a_rq_1-a_r^2q_2-a_r^3q_3-a_r^4q_4+o\left(a_r^4\right). \end{aligned}$$

One can derive that

$$\begin{aligned} Q_1(p) &= Q_1^d(p)=-\frac{\sigma^2}{2r}\left(1+\frac{r}{\phi}\right)^2p^2 \\ Q_2(p) &= Q_2^d(p)+\frac{\sigma^2}{2\phi^2}\left(1+\frac{r}{\phi}\right)^2p^2\left(p-\overline{p}^o\right)^2 \\ Q_3(p) &= Q_3^d-\frac{\sigma^4}{r^2}\left(1+\frac{r}{\phi}\right)^2p^2\left[\left(p-\overline{p}^o\right)^2+\left(1+\frac{r}{\phi}\right)p\left(p-\overline{p}^o\right)-\frac{1}{2}\left(1+\frac{r}{\phi}\right)^2\left(p^2-\left(\overline{p}^o\right)^2\right)\right] \\ Q_4(p) &= Q_4^d(p)-\frac{\phi\sigma^6}{r^3}\left(1+\frac{r}{\phi}\right)^4p^4+\frac{\phi^4\sigma^6}{2r^6}\left(1+\frac{r}{\phi}\right)^6p^2+\left(p-\overline{p}^o\right)[...], \end{aligned}$$

and

$$\begin{aligned} q_1 &= q_1^d=\frac{\phi^3\sigma^2}{r^3}\left(1+\frac{r}{\phi}\right)^2 \\ q_2 &= q_2^d=-\frac{\phi^5\sigma^4}{r^5}\left(1+\frac{r}{\phi}\right)^2\left[1+\left(1+\frac{r}{\phi}\right)^2\right] \\ q_3 &= q_3^d+\frac{\phi^5\sigma^4}{r^6}\left(1+\frac{r}{\phi}\right)^3. \end{aligned}$$

The detailed derivation is available upon request.