Optimal Long-term Contracting with Learning*

Zhiguo He[†] Bin Wei[‡] Jianfeng Yu[§] Feng Gao[¶]

October 2014

Abstract

We introduce uncertainty into the Holmstrom and Milgrom (1987) model to study optimal long-term contracting with learning. In the dynamic relationship, the agent's shirking not only reduces the current output, but also increases the agent's information rent due to the persistent belief manipulation effect. We characterize the optimal contract by solving a dynamic programming problem in which information rent is the unique state variable. We find that in the optimal contract, the optimal effort decreases stochastically over time, exhibiting a front-loaded pattern. Furthermore, the optimal contract exhibits an option-like feature in that incentives increase after good performance.

Keywords: Executive Compensation, Moral Hazard, Bayesian Learning, Hidden Information, Belief Manipulation, Private Savings, Continuous Time, Stock Options.

*We thank Philip Dybvig, Johannes Horner, Navin Kartik, Mark Loewenstein, Andrey Malenko, Hui Ou-Yang, Julien Prat, Yuliy Sannikov, Noah Williams, and participants at Federal Reserve Board, Cheung Kong Graduate School of Business, 2014 AFA Annual Meeting at Philadelphia, 2013 Econometric Society Winter Meeting at San Diego, 2012 Society of Economic Dynamics Meeting, and 2012 China International Conference in Finance for helpful comments. 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.

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

[‡]Federal Reserve Bank of Atlanta, 1000 Peachtree St. N.E., Atlanta, GA 30309. Phone: (1)404-498-8913. Email: bin.wei@atl.frb.org.

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

[¶]Tsinghua University, School of Economics and Management, Beijing, China. Email: gaof@sem.tsinghua.edu.cn.

1 Introduction

Many long-term contractual relationships feature learning, as uncertainty arises if either the project’s quality or the agent’s ability is unknown when a long-term contract is signed. Dynamic learning is most relevant for venture capital firms investing in companies with new technologies, or for firms hiring fresh graduates. Unfortunately, with learning, it is generally challenging to study long-term contracting (for reasons stated below), and the current literature offers little about what the optimal incentive dynamics looks like in this setting.

We introduce uncertainty and learning into the classic Holmstrom and Milgrom (1987) model with a Constant-Absolute-Risk-Aversion (CARA) agent. We choose Holmstrom and Milgrom (1987) for two reasons. First, in the Holmstrom and Milgrom (1987) model the optimal contract is linear. Against this benchmark, we show that uncertainty and learning make the optimal compensation contract option-like, i.e., working incentives rise following good performance. Second, Holmstrom and Milgrom (1987) has a tractable dynamic CARA-Normal framework which accommodates learning nicely. We consider an infinite-horizon variation of Holmstrom and Milgrom (1987) with stationary learning to maintain tractability.

In our model the principal signs a long-term contract with the agent. The observable output each period is the sum of the agent’s unobservable effort, the project’s unknown profitability (or the agent’s unknown ability), and some transitory noise. To focus on learning only (rather than adverse selection), we assume that both the principal and agent share a common prior on the project’s profitability when signing the long-term contract initially.

Different from Holmstrom and Milgrom (1987), incentive provisions across different times become intertemporally linked due to learning. The intertemporal linkage of incentive provisions is rooted in the “*hidden information*” problem.¹ Along the equilibrium path the principal knows as much as the agent knows, as both start with the common prior. However, along off-equilibrium paths the agent knows strictly more, because only the agent knows his actual effort which may deviate from the recommended level. Specifically, imagine that the agent has followed the recommended effort policy in the past, thus both parties share the same correct belief about the project’s profitability. If the agent shirks today, i.e., exerting some effort below the recommended level, then the lower effort decreases today’s output on average. Furthermore, with Bayesian learn-

¹This is in contrast to the standard “*hidden action*” dynamic agency models where the agent’s unobservable shirking has only a *short-lived* effect. For recent development of dynamic contracting in finance, see DeMarzo and Fishman (2007), Biais, Mariotti, Plantin, and Rochet (2007), DeMarzo and Sannikov (2006), He (2009), Piskorski and Tchisty (2010), DeMarzo, Fishman, He, and Wang (2012), Malenko (2013), etc.

ing, the principal who anticipates a higher effort today would mistakenly attribute today’s weak performance to lower profitability. Thus, by shirking today the agent can distort downward the principal’s inference about profitability from today onwards, which is long-lasting (i.e., persistent hidden information). This *belief manipulation* effect is beneficial to the agent, as the principal will mistakenly reward the agent later whenever the future performance beats the principal’s downward distorted expectations. We refer to this potential benefit due to off-equilibrium private information as the agent’s *information rent*.

In solving the optimal contract with learning, we need the information rent as another state variable in addition to the agent’s continuation value. The information rent captures the marginal benefit of the agent’s shirking due to the belief manipulation effect, and hence enters the agent’s incentive compatibility constraint. The higher the future incentives, the greater the information rent, and the lower the agent’s current working motivations. In fact, we show that the information rent can be conveniently expressed as the sum of properly discounted future incentives, and the agent’s optimal effort is simply the instantaneous incentive (i.e., pay-performance sensitivity) minus the information rent due to the belief manipulation effect. Finally, thanks to the CARA preference without the wealth effect,² the agent’s continuation value separates from the problem, and the optimal contract can be fully characterized by an ordinary differential equation (ODE) with the information rent as the only state variable.

We find two interesting features of the optimal contract. First, in our model the optimal effort policy, which is always distorted downward relative to the first-best benchmark, has a negative drift, thus exhibiting a front-loaded or time-decreasing pattern.³ In Section 5.1 we solve in closed form the optimal deterministic contract (i.e., the optimal one among the contracts in the subspace that implements time-varying deterministic incentives only), and show analytically that the optimal deterministic effort policy decreases over time. This pattern holds in the optimal stochastic contract, and the intuition comes from the belief manipulation effect. As mentioned, later incentives increase the agent’s current information rent for shirking. This implies that future pay-performance sensitivities hurt the agent’s motive for working in earlier periods, but not the other way around. Given that later incentives are more costly, the optimal contract implements lower effort in later

²The CARA preference features no wealth effect. Although the agent is promised with different levels of continuation values after a certain history of performance shocks, these promised continuation values themselves can be viewed as different promised wealth levels to the agent and thus do not affect the contracting problem looking forward.

³This is somewhat surprising. We have explained that under a given contract the information rent makes the agent want to work less in earlier periods, and casual readers might conclude that in the optimal contract the agent should work *less* in earlier periods. However, the opposite holds in the optimal contract: the principal will endogenously give higher incentives in earlier periods so that the agent work *harder* in earlier periods in equilibrium.

periods.⁴

Second, we find that the optimal effort policy is stochastic with higher incentives after good performance, exhibiting an *option-like* feature.⁵ The intuition comes from reducing the agent's belief manipulation in a long-term relationship. For a risk-averse agent, the amount of information rent not only depends on the benefits from belief manipulation that are proportional to future pay-performance sensitivities, but also the agent's marginal utility at these future states when receiving those benefits. Raising incentives after good performance introduces a negative correlation between pay-for-performance and marginal utility. That is, greater future benefits from belief manipulation are associated with the states when the agent cares less. Hence, the option-like compensation contract lowers the agent's information rent standing today.

This result of effort policy being history-dependent is more surprising given our setting. With a standard CARA-Normal setting and learning, as the posterior variance only changes over time deterministically (in our stationary setting, it is a constant), the resulting equilibrium effort profile is usually deterministic (e.g., Gibbons and Murphy, 1992; Holmstrom, 1999). In contrast, in our model with learning, the optimal long-term contract has an option-like feature in that pay-for-performance rises following good performance.

We emphasize that it is the combination of long-term contracting and learning that drives the front-loaded and option-like incentives. For long-term contracting but no learning, the model is a simple extension of Holmstrom and Milgrom (1987) and a constant effort policy is optimal (Section 4.3). Now with learning, the belief manipulation effect induces the agent to work less today given higher future incentives. The absence of commitment in the short-term relationships, however, implies that principals at different times will not take this belief manipulation effect into account. Section 5.2.3 shows that the absence of commitment again implies a constant effort process in equilibrium, thanks to the Gaussian setting with stationary Bayesian learning (as in Holmstrom, 1999).

⁴Interestingly, the pattern of time-decreasing effort policy in our paper with *post-contracting* information asymmetry is opposite to the dynamic contracting setting with *pre-contracting* asymmetric information in Garrett and Pavan (2012): in that paper, the agent privately observes his productivity when signing the contract; and under the assumption that the effect of the initial productivity on his future productivity is declining over time, the optimal effort policy is time-increasing. Intuitively, in Garrett and Pavan (2012), downward distortion required for rent extraction is more severe in earlier periods when the major friction is pre-contracting private information. It is intriguing that pre-contracting private information and post-contracting information have opposite predictions for the time-series pattern of effort distortion, but the difference also lies on the agent being risk-neutral without wealth constraint in Garrett and Pavan (2012). Relatedly, Sannikov (2014) allows for the possibility of the agent's current effort to affect future fundamentals, and finds that the optimal effort policy is increasing over time as well.

⁵Dittmann and Maug (2007) show that the optimal contract implied by standard principal-agent models almost never contains any options, a prediction contrary to practice. From this perspective, the endogenous *option-like* result is particularly interesting.

We rely on the specific setting (e.g., CARA preferences, Gaussian processes, etc) to fully characterize the optimal long-term contract with learning. For instance, we use the first-order approach to solve for the optimal contract, and the specific CARA preference allows us to verify the validity of the first-order approach by identifying an upper-bound of the agent’s deviation value (Section 3.4). Nevertheless, we emphasize that the above two contracting features are likely to hold qualitatively in a more general setting. Robustness exists because the economic forces driving these results do not depend on CARA preferences or Gaussian processes. In any long-term contracting environment with learning, the agent’s information rent due to belief manipulation—that is, the agent would like to shirk to distort the principal’s future belief downward—is general. The time-decreasing effort policy comes from the fact that later incentives enter the agent’s forward-looking information rent in earlier periods (but not the other way around). Furthermore, the option-like feature only relies on the concavity of the agent’s utility function, so that the marginal value of earning future (potential) belief manipulation benefit is lower for the agent with good performance hence higher compensation.

There is no doubt that in practice the use of options is pervasive. Hall and Liebman (1998), for example, find that there has been a large increase in the use of stock options in CEO compensation for incentive provisions. Interestingly, traditional static models typically do not predict option grants.⁶ For example, Dittmann and Maug (2007) calibrate a standard static structural model and find that most CEOs should hold more straight equity, hold no stock options, and receive lower salaries. Our model not only gives one mechanism for why option-like payoffs are desirable in managerial compensation, but also predicts that industries with higher uncertainty should grant more stock options to their managers. The latter cross-sectional predication is consistent with the evidence in Ittner, Lambert, and Larcker (2003) and Murphy (2003) who document more extensive use of stock options in new-economy firms (e.g., computer-related firms).

Our paper is closest to DeMarzo and Sannikov (2014) and Prat and Jovanovic (2014). Both papers deal with long-lasting belief manipulation effect in dynamic agency settings with learning, but restrict attention to the optimal contract that implements a constant first-best level of effort. Prat and Jovanovic (2014) focus on the role of intertemporal commitment in optimal contracting. DeMarzo and Sannikov (2014) impose limited liability constraint on the agent, and study the optimal payout and termination policies. In contrast, we solve for the optimal effort policy jointly with

⁶There are a few exceptions in a dynamic framework. For instance, Edmans and Gabaix (2011) show that the convexity of the contract depends on the marginal cost of effort. In Ju and Wan (2012), stock options become optimal when the agent has to be paid above certain subsistence level.

the optimal long-term compensation contract, and emphasize the general economic mechanisms that shape the optimal effort policy in long-term optimal contracting. Relatedly, the long-lasting belief manipulation also exists in Bergemann and Hege (1998, 2005) and Horner and Samuelson (2012).⁷

The topic of optimal contracting with endogenous learning also relates to the recent literature studying optimal long-term contracts with adverse selection and moral hazard (e.g., Baron and Besanko, 1984; Sung, 2005; Sannikov, 2007; Garrett and Pavan, 2012; Gershkov and Perry, 2012; Halac, Kartik, and Liu, 2012; and Cvitanic, Wan, and Yang, 2013). In general, when the agent has pre-contracting private information that is persistent, a mechanism design approach naturally arises; see Pavan, Segal and Toikka (2012) and Golosov, Troshkin, and Tsyvinski (2012), who use the first-order-approach to solve the agent’s problem.⁸ However, because our paper focuses on the problem without pre-contracting private information, we do not need to solve for the optimal menu for the agent’s truthful reporting when signing the contract.

The rest of the paper is organized as follows. Section 2 lays out the model, and Section 3 solves the agent’s problem. Section 4 solves the optimal contract by reformulating the principal’s problem which allows for the use of dynamic programming techniques. Section 5 discusses the implications of our model, and Section 6 concludes. All proofs are in the appendix.

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

$$dY_t = (\mu_t + \theta_t) dt + \sigma dB_t, \quad (1)$$

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 the constant $\sigma > 0$ is the volatility of cash flows. Moral

⁷There 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 between the principal and the agent about the agent’s ability, where the agent is dogmatic about his belief (i.e., the agent never updates his posterior belief about profitability from past performance), which eliminates the belief manipulation effect. In that paper, 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 manipulation effect. More recently, 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.

⁸This is the same approach used in Williams (2009, 2011) and Zhang (2009), who study persistent information in a continuous-time principal-agent setting. We also use the first-order-approach to solve the agent’s problem, and verify the validity of the first-order approach in Section 3.4.

hazard arises from the agent's unobservable effort choice. The risk-neutral principal (*she* later on) offers the Constant-Absolute-Risk-Averse (CARA) 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. Both parties can commit to the long-term relationship at $t = 0$, at which point the agent has no personal wealth and has an exogenous reservation utility of v_0 . Without loss of generality, we assume the principal has all the bargaining power.

Relative to Holmstrom and Milgrom (1987), we introduce the project's *unknown* profitability θ_t into the output process in equation (1). Equivalently one can interpret θ_t as the agent's unknown ability. The profitability θ_t is unobservable and thus calls for learning. If θ_t is perfectly observable (or absent), there is no need for learning; in the beginning of Section 4.3 we show that this case reduces to Holmstrom and Milgrom (1987).

We assume that profitability $\{\theta_t\}$ follows a martingale process so that

$$d\theta_t = \phi \sigma dB_t^\theta,$$

where the Brownian motion $\{B^\theta\}$ is independent to $\{B\}$ and $\phi > 0$ is a constant. At time 0, the principal and the agent share the common normal prior: $\theta_0 \sim \mathcal{N}(m_0, \Sigma_0^\theta)$. We mainly focus on stationary learning; in Appendix B we briefly discuss non-stationary learning for robustness check. For learning to be stationary, the prior uncertainty is assumed to satisfy $\Sigma_0^\theta = \sigma^2 \phi$, so that the posterior variance $\Sigma_t^\theta = \Sigma_0^\theta$ for all t and Bayesian updating is time-independent. When $\phi = 0$, our model features no uncertainty, and thus is reduced to the benchmark model of Holmstrom and Milgrom (1987).

We further assume that the agent can privately save (i.e., saving is unobservable) so as to smooth his consumption intertemporally. It is a well-known result that CARA preferences do not have a wealth effect, and the issue of private savings can be easily dealt with (e.g., Fudenberg, Holmstrom, and Milgrom, 1990; Williams, 2009; He, 2011). In fact, later we show that private savings allow us to separate the agent's continuation value from another state variable in deriving the optimal contract, which renders the tractability of our problem.

Private savings imply that the agent's actual consumption can differ from wage c_t . Denote the agent's actual consumption by $\hat{c}_t$ and actual effort by $\hat{\mu}_t$. Following Holmstrom and Milgrom

(1987), we assume that the agent has an exponential or CARA preference:⁹

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

where $a > 0$ is the agent's absolute risk aversion coefficient, and $g(\hat{\mu}_t) \equiv \frac{1}{2}\hat{\mu}_t^2$ is the instantaneous quadratic monetary cost of exerting effort $\hat{\mu}_t$. The quadratic form of $g(\cdot)$ simplifies our results, but our analysis goes through as long as $g(\cdot)$ is strictly increasing and strictly convex. For ease of technical proofs, we will assume that the feasible effort policies are bounded, i.e., there exists a sufficiently large constant L_μ so that $|\hat{\mu}_t| < L_\mu$.

2.1 Bayesian Learning and Effort

Recall that at time 0, the principal and the agent share the common normal prior $\theta_0 \sim \mathcal{N}(m_0, \Sigma_0^\theta)$. From now on we normalize $m_0 = 0$. 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 is $\mathcal{F}\{Y_s, \mu_s : 0 \leq s \leq t\}$, as the principal knows the recommended effort policy $\mu \equiv \{\mu_t\}$. However, the agent's information set also includes his actual effort policy $\hat{\mu} \equiv \{\hat{\mu}_t\}$, i.e., $\mathcal{F}\{Y_s, \mu_s, \hat{\mu}_s : 0 \leq s \leq t\}$. Intuitively, relative to the principal, the agent knows (weakly) more because he knows his actual past effort choices $\hat{\mu}$, which may deviate from the recommended policy μ . This distinction is important for our analysis.

If the agent follows the recommended effort policy μ , the principal's posterior belief about θ_t is correct and fully summarized by the first two moments:

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

A standard filtering argument (e.g., Theorem 12.2 in Liptser and Shiriyayev, 1977) implies that $\Sigma_t^{\theta, \mu} = \sigma^2 \phi$ for all t , and

$$dm_t^\mu = \Sigma_t^{\theta, \mu} \frac{dY_t - (\mu_t + m_t^\mu) dt}{\sigma^2} = \sigma \phi dB_t^\mu, \text{ with } m_0 = 0, \quad (2)$$

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

$$dB_t^\mu = \frac{dY_t - (\mu_t + m_t^\mu) dt}{\sigma}. \quad (3)$$

⁹In the tradition of Holmstrom and Milgrom (1987), the CARA preference allows for negative consumption, i.e., both c_t and $\hat{c}_t$ can take negative values. This is different from DeMarzo and Sannikov (2014) in which the agent is protected by limited liability and hence the endogenous contract termination arises. It is unclear how the limited liability restriction affects the qualitative results of our paper.

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

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

The superscript $\hat{\mu}$ emphasizes the dependence on the agent's actual effort policy $\hat{\mu}$ (which the principal does not know). Similarly, $\Sigma_t^{\theta, \hat{\mu}} = \sigma^2 \phi$ for all t , and

$$dm_t^{\hat{\mu}} = \Sigma_t^{\theta, \hat{\mu}} \frac{dY_t - (\hat{\mu}_t + m_t^{\hat{\mu}}) dt}{\sigma^2} = \sigma \phi dB_t^{\hat{\mu}}, \text{ with } m_0 = 0, \quad (4)$$

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

$$dB_t^{\hat{\mu}} \equiv \frac{1}{\sigma} \left(dY_t - (\hat{\mu}_t + m_t^{\hat{\mu}}) dt \right). \quad (5)$$

The difference between the agent's belief $m_t^{\hat{\mu}}$ and the principal's belief m_t^{μ} will be important in our later analysis.

2.2 Formulating the Optimal Contracting Problem

We first 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 $\{c_t, \mu_t\}$ the agent's problem is

$$\begin{aligned} & \max_{\{\hat{c}, \hat{\mu}\}} \mathbb{E}_0^{\hat{\mu}} \left[\int_0^\infty e^{-rt} u(\hat{c}_t, \hat{\mu}_t) dt \right] \\ \text{s.t. } & dY_t = (\hat{\mu}_t + m_t^{\hat{\mu}}) dt + \sigma dB_t^{\hat{\mu}}, \\ & dS_t = rS_t dt + c_t dt - \hat{c}_t dt \text{ with } S_0 = 0, \end{aligned} \quad (6)$$

with the transversality condition, say S_t has to be bounded.¹⁰ Here, $\mathbb{E}^{\hat{\mu}}[\cdot]$ denotes the probability measure induced by the agent's effort policy $\{\hat{\mu}_t\}$, and $\{\hat{c}_t\}$ is the agent's actual consumption policy. Denote the optimal solution to equation (6) by $\{c_t^*, \mu_t^*\}$.

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 in equation (6) is $c_t^* = c_t$ and $\mu_t^* = \mu_t$, 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 a standard result in the literature, the following lemma shows that there is no loss of generality by restricting attention to incentive-compatible and no-savings contracts. The idea is similar to the revelation principle. The principal who can fully commit to the contract can save for the agent, and, once she knows the agent's actual effort policy, she will perform correct Bayesian updating based on that policy.

¹⁰We explicitly impose the assumption of private savings being bounded in Assumption 1 in the proof of Proposition 1 in Appendix A.3.

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

The optimal contract solves the principal's problem:

$$\max_{\{c_t, \mu_t\} \text{ is incentive-compatible and no-savings}} \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt} (dY_t - c_t dt) \right] \quad (7)$$

$$\text{s.t.} \quad dY_t = (\mu_t + m_t^\mu) dt + \sigma dB_t^\mu; \quad (8)$$

$$\mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt} u(c_t, \mu_t) dt \right] = v_0. \quad (9)$$

In equation (8), the principal's Bayesian updating of m_t^μ and dB_t^μ is based on the correct optimal effort policy $\{\mu_t\}$ taken by the agent in the incentive-compatible contract. Equation (9) is the agent's participation constraint at $t = 0$ for the agent with a reservation value v_0 . Since negative transfers are allowed, this participation constraint at $t = 0$ must bind.

3 The Agent's Problem

In this section we analyze the agent's problem. We illustrate heuristically the necessary conditions for a contract $\{c_t, \mu_t\}$ to be incentive-compatible and to induce no private savings. Formal proofs are delegated to Appendix A.3.

3.1 Continuation Value and Incentives

Given the incentive-compatible and no-savings contract $\{c_t, \mu_t\}$, the agent's continuation value, which is his expected payoff from the continuation contract, is defined as:

$$v_t \equiv \mathbb{E}_t^\mu \left[\int_t^\infty e^{-r(s-t)} u(c_s, \mu_s) ds \right]. \quad (10)$$

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

$$\begin{aligned} dv_t &= rv_t dt - u(c_t, \mu_t) dt + \beta_t (-arv_t) (dY_t - \mu_t dt - m_t^\mu dt) \\ &= rv_t dt - u(c_t, \mu_t) dt + \beta_t (-arv_t) \sigma dB_t^\mu. \end{aligned} \quad (11)$$

We can interpret β_t as the dollar incentives on the agent's unexpected performance. To see this, $\beta_t (-arv_t)$ can be interpreted as the incentive loading (in utilities) on the agent's unexpected performance $dY_t - m_t^\mu dt$. We show shortly that $-arv_t > 0$ is the agent's marginal utility from consumption at time t , i.e., $u_c(c_t, \mu_t)$. As a result, dividing utility incentives $\beta_t (-arv_t)$ by the marginal utility yields dollar incentives received by the agent.

Later we refer to pay-performance sensitivities $\{\beta_t\}$ simply as incentives. Throughout the paper, we impose a further technical condition for ease of our analysis. Essentially, we restrict the feasible incentive slopes $\{\beta_t\}$ to be bounded, i.e., there exists some sufficiently large constant M such that $\beta_t \in [-M, M]$. Since in the optimal contract endogenously the optimal incentives are bounded, this assumption is far from restrictive. As explained later, this assumption is in the same spirit as imposing the transversality condition.

3.2 No Savings

We first show that, under CARA preferences, no-savings conditions imply that

$$rv_t = u(c_t, \mu_t) = -\frac{1}{a} \exp[-a(c_t - g(\mu_t))]. \quad (12)$$

The argument is similar to He (2011). To see this, we first present the following lemma, where we use Π to denote any compensation contract.

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

$$v_t(S; \Pi) = v_t(0; \Pi) \cdot e^{-arS} = v_t \cdot e^{-arS}, \quad (13)$$

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 equation (10).

The driving force behind this result is simple. Due to CARA preferences, the agent's problem is translation-invariant with respect to his underlying wealth level, as evidenced by $u(c_s + rS, \mu_s) = e^{-arS} u(c_s, \mu_s)$. Thus, for a CARA agent, given the extra savings S , his new optimal policy is to take the optimal consumption-effort-learning policy without savings, and consume an extra rS more for all future dates. As suggested by equation (13), the optimal deviation value with savings S is just the original value without savings, multiplied by the adjusting factor e^{-arS} for extra consumption going forward.

By the optimality of the agent's consumption-savings policy in equation (6), his marginal utility from consumption must equal his marginal value of wealth, and equation (13) implies that:

$$u_c(c_t, \mu_t) = \left. \frac{\partial v_t(S; \Pi)}{\partial S} \right|_{S=0} = -arv_t. \quad (14)$$

Thus, equation (12) follows immediately from equation (14) by using the fact that under CARA preferences, the agent's utility level is linear in his marginal utility:

$$au(c_t, \mu_t) = -u_c(c_t, \mu_t). \quad (15)$$

Plugging equation (12) into equation (11), we find that for no-saving contracts, v_t follows an (exponential) martingale:¹¹

$$\begin{aligned} dv_t &= \beta_t (-arv_t) \sigma dB_t^\mu \\ \Leftrightarrow v_s &= v_t \exp \left(- \int_t^s ar\beta_u \sigma dB_u^\mu - \frac{1}{2} \int_t^s a^2 r^2 \beta_u^2 \sigma^2 du \right) \text{ for } s > t. \end{aligned} \tag{16}$$

We can also understand this result by combining two observations. First, the agent can smooth out his consumption intertemporally, and hence his marginal utility has to follow a martingale. Second, his continuation value v_t is linear in his marginal utility u_c because of equations (14) and (15). The fact that v_s/v_t only depends on $\{\beta_u\}_{s \leq u \leq t}$ in equation (16) is important for later analysis.

3.3 Effort and Belief Distortion

The 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 the principal's future beliefs about project profitability downward.

Consider the following thought experiment. Suppose that at time t the agent chooses an effort level $\hat{\mu}_t$ below the recommended effort μ_t , and thus output is lower than what is expected by the principal. Crucially, however, the principal thinks the agent is exerting an effort of μ_t —thus she (through learning) mistakenly attributes lower output to a lower value of profitability θ_t . In contrast, the agent updates profitability θ_t based on his true effort level $\hat{\mu}_t$, leading to a positive wedge $m_t^{\hat{\mu}} - m_t^\mu$ between the beliefs of the agent and principal. In other words, by shirking, the agent makes the principal (mistakenly) underestimate the project profitability. This belief manipulation is beneficial to the agent in a dynamic setting—when future outputs turn 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 formalize this effect. When the agent deviates from the recommended effort path μ by choosing effort policy $\hat{\mu}$, the principal's posterior mean estimate about θ_t is distorted downward.

¹¹Because $|\beta| < M$ is bounded, the local martingale $\{v_t\}$ is a martingale.

This distortion, denoted by Δ_t , has the following intuitive expression:¹²

$$\Delta_t \equiv m_t^{\hat{\mu}} - m_t^{\mu} = \phi \int_0^t e^{\phi(s-t)} (\mu_s - \hat{\mu}_s) ds. \quad (17)$$

Intuitively, the current belief distortion is a properly-weighted cumulative effort deviation in the past. When $\phi = 0$, the zero prior uncertainty $\Sigma_0^0 = \sigma^2 \phi = 0$ eliminates any belief divergence and the issue of belief manipulation is absent.

As suggested by equation (11), the contract relies on the agent's "unexpected" performance along the equilibrium path $dY_t - (\mu_t + m_t^{\mu}) dt$. This equals σdB_t^{μ} under the equilibrium measure and has a mean of zero. For the agent who deviates by exerting $\hat{\mu} \neq \mu$, under his information set the above "unexpected" performance no longer has zero mean. Imagine that the agent has deviated before t so that $\hat{\mu}_s \neq \mu_s$ where $s < t$. Even if the agent exerts the same effort at time t so that $\mu_t = \hat{\mu}_t$, equation (5) implies that

$$dY_t - (\hat{\mu}_t + m_t^{\hat{\mu}}) dt = dY_t - (\mu_t + m_t^{\mu}) dt \quad (18)$$

has zero mean under the agent's information set. Hence, the "unexpected" performance $dY_t - (\mu_t + m_t^{\mu}) dt$ displays a positive drift under the agent's information set:

$$dY_t - (\mu_t + m_t^{\mu}) dt = \underbrace{\left[dY_t - (\mu_t + m_t^{\mu}) dt \right]}_{\text{zero mean under agent's info. set}} + \Delta_t dt,$$

Intuitively, if the agent shirked a bit in the past $\hat{\mu}_s < \mu_s$ with $s < t$, then later on the principal would mistakenly think the project is worse than it actually is (under the agent's correct measure), i.e., $\Delta_t > 0$. As a result, the agent can easily beat the principal's expectation, which explains the positive drift $\Delta_t > 0$. As we emphasize later, this translates to a positive information rent to the agent, as long as the principal is providing incentives to reward the agent for his "unexpected" performance.

3.4 Incentive-Compatibility Constraint and Intuition

Proposition 1 characterizes the agent's incentive compatibility constraint, along with the equilibrium consumption and continuation value heuristically derived above. We provide a rigorous proof for Proposition 1 in Appendix A.3. There, we also show that the policy (c, μ) satisfying the (local) necessary conditions is also globally optimal, based on an argument used in Sannikov (2014).

¹²According to equations (2) and (4), the increment of Δ_t is given by:

$$d\Delta_t = dm_t^{\hat{\mu}} - dm_t^{\mu} = \phi \left(dY_t - (\hat{\mu}_t + m_t^{\hat{\mu}}) dt \right) - \phi \left(dY_t - (\mu_t + m_t^{\mu}) dt \right) = \phi (\mu_t - \hat{\mu}_t - \Delta_t) dt,$$

which leads to the expression of Δ_t in (17). Here, we have used the fact that both the principal and the agent share the same belief when signing the contract at $t = 0$, i.e., $\Delta_0 = 0$ as $m_0^{\mu} = m_0^{\hat{\mu}} = m_0$.

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

$$\mu_t = \underbrace{\beta_t}_{\text{instantaneous incentive}} - \underbrace{\mathbb{E}_t^\mu \left[\int_t^\infty \phi e^{-(\phi+r)(s-t)} \frac{\beta_s v_s}{v_t} ds \right]}_{\text{future information rent}} \quad (19)$$

$$= \beta_t - \mathbb{E}_t^\mu \left[\int_t^\infty \phi e^{-(\phi+r)(s-t)} \beta_s \exp \left(- \int_t^s ar \beta_u \sigma dB_u^\mu - \frac{1}{2} \int_t^s a^2 r^2 \beta_u^2 \sigma^2 du \right) ds \right] \quad (20)$$

In addition, equation (12) implies that consumption (or wage) follows

$$c_t = g(\mu_t) - \frac{\ln(-arv_t)}{a}, \quad (21)$$

and the continuation payoff from the contract is

$$v_t = v_0 \exp \left(- \int_0^t ar \beta_s \sigma dB_s^\mu - \frac{1}{2} \int_0^t a^2 r^2 \beta_s^2 \sigma^2 ds \right). \quad (22)$$

Under the transversality condition assumption imposed in the proof, the above first-order conditions are sufficient to ensure global optimality.

In a standard dynamic agency problem without profitability uncertainty (e.g., $\phi = 0$), the agent's effort μ_t at time t should only depend on the time- t incentive β_t offered by the contract (i.e., $\mu_t = \beta_t$; recall the quadratic effort cost $g(\mu_t) = \mu_t^2/2$). 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 equation (19). The forward-looking downward adjustment term represents the *information rent* to the agent. Intuitively, this term captures the marginal benefit of manipulating the principal's future belief downward.¹³ Also, the expression in (20) implies that the agent's continuation payoffs $\{v\}$ drop out, which allows us to write the agent's incentive-compatibility constraint independent of $\{v\}$. As shown later, this convenient property is crucial for the tractability of our problem.

3.4.1 Intuition of Incentive-Condition Constraint

The rest of this subsection is devoted to understanding the key incentive compatibility constraint (19). Consider the agent who reduces his effort to slightly below the recommended effort level μ_t ,

¹³This information rent 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, as in any typical moral hazard model, the marginal deviation benefit (marginal rent) is important in characterizing the agent's incentive-compatibility condition.

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

$$\mu^\epsilon = \begin{cases} \mu_s & \text{for } s \notin [t, t + dt]; \\ \mu_s - \epsilon & \text{otherwise.} \end{cases} \quad (23)$$

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

Standing at time t , the agent's total payoffs include his instantaneous utility $u(c_t, \mu_t)$ and his continuation payoff. In Appendix A.4 we show that one can write the agent's total payoff from time t onwards as

$$u(c_t, \mu_t - \epsilon) dt + v_t + \mathbb{E}_t^{\mu^\epsilon} \left[\int_t^\infty e^{-r(s-t)} dv_s \right], \quad (24)$$

where $\mathbb{E}_t^{\mu^\epsilon}$ emphasizes that the agent forms his expectation based on his information set induced by μ^ϵ . Using the result in equation (16), we can rewrite (24) heuristically as:

$$\begin{aligned} & u(c_t, \mu_t - \epsilon) dt + v_t + \mathbb{E}_t^{\mu^\epsilon} \left\{ \begin{aligned} & \beta_t (-arv_t) (dY_t (\mu_t - \epsilon) - \mu_t dt - m_t^\mu dt) + \\ & \int_{t+dt}^\infty e^{-r(s-t)} \beta_s (-arv_s) [dY_s - (\mu_s + m_s^\mu) ds] \end{aligned} \right\} \\ = & \underbrace{u(c_t, \mu_t - \epsilon) dt}_{\text{saving effort cost instantaneously}} + v_t + \\ & \mathbb{E}_t^{\mu^\epsilon} \left\{ \underbrace{\int_{t+dt}^\infty e^{-r(s-t)} \beta_s (-arv_s) \left[\underbrace{\left(dY_s - (\mu_s^\epsilon + m_s^{\mu^\epsilon}) dt \right)}_{\text{martingale under info set generated by } \mu} + \underbrace{\Delta_s ds}_{\text{future belief divergences}} \right]}_{\text{creating belief divergence persistently}} \right\} \end{aligned} \quad (25)$$

There should be another correction term in $(\mu_s^\epsilon - \mu_s) ds$ in the second equality, but it is zero because of (23), i.e., we consider a one-shot deviation at time t from the equilibrium effort policy.

There are two channels through which shirking at time t affects future changes in the agent's continuation value. The first channel captures the instantaneous performance effect, i.e., the agent's effort affects instantaneous performance dY_t and, thus, his continuation value. To see this, write performance $dY_t(\mu_t)$ over $[t, t + dt]$ as a function of time- t effort μ_t . Exerting effort $\mu_t - \epsilon$ hurts the short-term performance over $[t, t + dt]$ because

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

Modulated by incentives, this leads to a drop in the agent's continuation value by $\beta_t(-arv_t) \cdot \epsilon dt$, via the channel of "hurting performance instantaneously." The second channel is the persistent effect due to belief manipulation. As discussed in Section 3.3, the agent's shirking at time t shifts the belief divergence path $\{\Delta_s\}$ away from the equilibrium path $\{\Delta_s = 0\}$ for $s > t$, according to equation (17).

We show that the incentive-compatibility constraint in equation (19) is implied by equation (25). By “reducing effort cost instantaneously” in equation (25), the agent’s marginal gain from shirking at t is $-u_\mu(c_t, \mu_t) \cdot \epsilon dt$. Since $u_\mu(c_t, \mu_t) = -u_c(c_t, \mu_t) \mu_t = arv_t \mu_t$, this marginal gain is $(-arv_t) \mu_t \cdot \epsilon dt$. On the other hand, shirking “hurts performance instantaneously” in equation (25), which gives rise to a marginal cost of $\beta_t(-arv_t) \cdot \epsilon dt$. In standard models without belief manipulation, these two forces fully determine the agent’s trade-off in choosing his optimal effort at time t .

Now we analyze the novel term “creating belief divergence persistently” in equation (25). There, because $dY_s - (\mu_s^\epsilon + m_s^{\mu^\epsilon}) dt$ has zero mean, this term equals

$$\mathbb{E}_t^{\mu^\epsilon} \left[\int_t^\infty e^{-r(s-t)} \beta_s(-arv_s) \Delta_s ds \right]. \quad (26)$$

Recall equation (23) which says that the agent is shirking by ϵ only at $[t, t + dt]$ but never before or after. Because of the exponential decay indicated by equation (17), the belief divergence in any future time $s > t$ is

$$\Delta_s = \phi e^{-\phi(s-t)} \cdot \epsilon dt. \quad (27)$$

Intuitively, as new information flows in, this belief divergence persists but decays over time exponentially at the rate of ϕ . Plugging equation (27) into equation (26), the marginal impact of shirking via the channel of belief manipulation is (where the second term comes from the difference between the expectation terms $\mathbb{E}_t^{\mu^\epsilon}[\cdot]$ and $\mathbb{E}_t^\mu[\cdot]$ is in the order of ϵdt):¹⁴

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

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. The higher-powered the future incentives $\{\beta_s\}$, the greater the information rent. Finally, for a risk averse agent, the information rent depends on the agent’s future marginal utility $(-arv_s)$ when receiving those manipulation benefits, a result that is useful later to understand the option-like feature in optimal contracting.

Combining three pieces together (canceling ϵdt and ignoring higher order terms), and dividing both sides by time- t marginal utility $(-arv_t)$, the agent’s incentive-compatibility constraint is

$$\mu_t - \beta_t + \frac{1}{-arv_t} \mathbb{E}_t^\mu \left[\int_t^\infty \phi e^{-(\phi+r)(s-t)} \beta_s(-arv_s) ds \right] = 0,$$

which is just (19); equation (20) follows because of equation (22). We call the last term in equation (19) information rent, which is new to standard dynamic agency models.

¹⁴For the purpose of illustrating the intuition, this statement is meant to be heuristic. For rigorous argument, see the Appendix for the proof of Proposition 1.

3.4.2 Validity of the first-order approach

We will rely on the agent’s incentive-compatibility constraint (19), which is the agent’s first-order condition in his effort decision, to derive the optimal contract in Section 4.3. This is the so-called first-order approach, and in the dynamic agency literature it is challenging to show that the necessary local first-order condition for the agent’s problem is indeed sufficient for the agent’s global optimality. Readers can skip this subsection which is only for this technical (though important) issue, as it is not related to the derivation of optimal contracting studied later in Section 4.3.

In this paper we are able to guarantee the validity of the first-order approach. In the second part of the proof for Proposition 1, we show that the first-order conditions in Proposition 1 are sufficient to ensure the agent’s global optimality, by following an upper-bound approach employed in Sannikov (2014).

To illustrate the basic idea, suppose that the agent facing the employment contract has deviated in the past, by having saved a bit and/or shirked a bit. For private saving, it is clear that this agent’s deviation state is the saving balance S_t ; recall Lemma 2 and equation (13). For shirking which distorts the principal’s current and future beliefs, the relevant deviation state is less clear, but we identify one deviation state variable X_t that suffices to characterize the “upper bound” for the agent’s deviation value. The state variable X_t essentially accumulates past effort deviations properly, just as savings accumulate past consumption deviations; see equation (53) in Appendix A.3 for details. Given these two deviation states, we define a function

$$W(S_t, X_t) \equiv v_t \cdot \exp(-ar(S_t + X_t)),$$

where v_t is the agent’s equilibrium continuation payoff achieved by the equilibrium strategy satisfying the first-order conditions. Utilizing the convenient properties for CARA preferences, we prove that $W(S_t, X_t)$ is the *upper bound* of the agent’s deviation value, given his current deviation state-pair (S_t, X_t) .¹⁵ Importantly, this upper bound satisfies the property of $W(S_t = 0, X_t = 0) = v_t$. As a result, the strategy satisfying first-order conditions, which achieves this upper bound, is indeed optimal for the agent.

¹⁵It is worth noting that $W(S_t, X_t)$ is not the *exact* deviation value of the agent; it just provides an upper bound for the agent’s deviation value. We prove this result by showing that the auxiliary gain process $\int_0^t e^{-rs} u(\hat{c}_s, \hat{\mu}_s) ds + e^{-rt} W(S_t, X_t)$ follows a supermartingale for any feasible policy $\{\hat{c}_t, \hat{\mu}_t\}$. For technical reasons, we also assume that the space of deviation strategies satisfies the transversality condition, which is guaranteed if these two deviation states (S_t, X_t) are bounded. For more details, see the proof for Proposition 1.

4 The Principal's Problem and Optimal Contracting

From now on we focus on incentive-compatible contracts such that both parties will have the same information set. For ease of notation we use dB_t to denote dB_t^μ , and write $\mathbb{E}^\mu$ as $\mathbb{E}$. In Section 4.1 we first separate out the agent's continuation payoff v_t from the principal's problem. In Section 4.2 we then rewrite the principal's problem in a standard recursive form with only one state variable. This recursive form allows us to apply the dynamic programming technique to solve for the optimal contract in Section 4.3.

4.1 Rewriting the Principal's Problem

In light of Proposition 1, we first rewrite the principal's problem in equation (7). Proposition 1 establishes an important link between recommended effort $\{\mu_t\}$ and incentives $\{\beta_t\}$ in any incentive-compatible contracts. Moreover, the principal can choose the optimal $\{\beta_t^*\}$ to maximize her value, and the corresponding optimal consumption process $\{c_t^*\}$ and the optimal effort policy $\{\mu_t^*\}$ are determined by equation (21) and equation (19), respectively. Therefore, we can rewrite the principal's problem in equation (7) as

$$\max_{\{\beta_t\}} \mathbb{E} \left[\int_0^\infty e^{-rt} (dY_t - c_t dt) \right] \quad (28)$$

$$s.t. \quad dY_t = (\mu_t + m_t)dt + \sigma dB_t \text{ and } dm_t = \phi \sigma dB_t, \quad (29)$$

$$c_t = g(\mu_t) - \frac{\ln(-arv_t)}{a}, \text{ where } g(\mu_t) = \frac{1}{2}\mu_t^2, \quad (30)$$

$$dv_t = \beta_t(-arv_t)\sigma dB_t, \text{ given } v_0; \quad (31)$$

$$\mu_t = \beta_t - \mathbb{E}_t \left[\int_t^\infty \phi \beta_s e^{-(\phi+r)(s-t)} \exp \left(- \int_t^s ar \beta_u \sigma dB_u - \frac{1}{2} \int_t^s a^2 r^2 \beta_u^2 \sigma^2 du \right) ds \right]. \quad (32)$$

Here, equation (29) describes the dynamics of output and posterior belief; equations (30)-(32) are derived from equations (19)-(22) in Proposition 1.

Thanks to the CARA preference, the agent's continuation value v_t separates from the problem and the optimal contracting problem can be rewritten without v_t . Start from the principal's objective in equation (28). In Appendix A.5 we show that

$$\begin{aligned} & \mathbb{E} \left[\int_0^\infty e^{-rt} (dY_t - c_t dt) \right] \\ &= \underbrace{\mathbb{E} \left[\int_0^\infty e^{-rt} \mu_t dt \right]}_{\text{expected output}} - \underbrace{\left(-\frac{\ln(-arv_0)}{ar} \right)}_{\text{C.E. of outside option } v_0} - \mathbb{E} \left[\int_0^\infty e^{-rt} \left(\underbrace{g(\mu_t)}_{\text{effort cost}} + \underbrace{ar\sigma^2\beta_t^2/2}_{\text{risk comp.}} \right) dt \right], \quad (33) \end{aligned}$$

The discounted expected output comes from the agent's effort (recall the project's initial profitability $m_0 = 0$). The total compensation cost is the Certainty Equivalent (CE) $-\frac{\ln(-arv_0)}{ar}$ of delivering

the agent's outside option v_0 , plus the monetary effort cost $g(\mu_t) = \mu_t^2/2$ and the 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 lack of the wealth effect of CARA preferences.

Combining (33) with (32), the principal's problem is simplified to

$$\begin{aligned} \max_{\{\beta_t\}} \mathbb{E} \left[\int_0^\infty e^{-rt} \left(\mu_t - \frac{1}{2} \mu_t^2 - \frac{1}{2} ar \sigma^2 \beta_t^2 \right) dt \right] \\ \text{s.t. } \mu_t = \beta_t - \mathbb{E}_t \left[\int_t^\infty \phi \beta_s e^{-(\phi+r)(s-t)} \exp \left(- \int_t^s ar \beta_u \sigma dB_u - \frac{1}{2} \int_t^s a^2 r^2 \beta_u^2 \sigma^2 du \right) ds \right]. \end{aligned} \quad (34)$$

Importantly, only incentives $\{\beta\}$ enter the problem, but not the agent's continuation payoffs $\{v\}$.

4.2 Recursive Formulation

We now recursively formulate the principal's problem in (34) and solve it using the approach of dynamic programming. Define the information rent p_t as

$$p_t \equiv \mathbb{E}_t \left[\int_t^\infty \phi \beta_s e^{-(\phi+r)(s-t)} \exp \left(- \int_t^s ar \beta_u \sigma dB_u - \int_t^s \frac{1}{2} (ar \beta_u \sigma)^2 du \right) ds \right], \quad (35)$$

which, together with the Martingale Representation Theorem, implies that there exists some progressively measurable process $\{\sigma_t^p\}$ so that the dynamics of p_t follows (see Appendix A.6)

$$dp_t = [(\phi + r) p_t + \beta_t (ar \sigma \sigma_t^p - \phi)] dt + \sigma_t^p dB_t. \quad (36)$$

From now on, we interpret $\{\sigma_t^p, \beta_t\}$ as our control because the pair determines the drift and diffusion of p_t in equation (36). As we will derive σ_t^p and β_t as a function of the auxiliary state p_t , the control pair $\{\sigma_t^p, \beta_t\}$ gives the full history of $\{\beta_t : t \geq 0\}$ that we are after.

The information rent p_t serves as the only state variable for the principal when designing the optimal contract. Denote the principal's value function by $V(p)$. The Hamilton-Jacobi-Bellman (HJB) equation of the principal problem is straightforward:

$$rV(p) = \max_{\beta, \sigma^P} (\beta - p) - \frac{1}{2} (\beta - p)^2 - \frac{ar\sigma^2}{2} \beta^2 + V_p [(\phi + r)p + \beta (ar\sigma\sigma^P - \phi)] + \frac{1}{2} V_{pp} (\sigma^P)^2. \quad (37)$$

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

$$\beta = \frac{1 + p - \phi V_p}{1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{(V_p)^2}{V_{pp}}} \text{ and } \sigma^P = -ar\sigma\beta \frac{V_p}{V_{pp}}. \quad (38)$$

Plugging them back into the HJB equation (37), we have

$$rV = \frac{1}{2} \frac{(1 + p - \phi V_p)^2}{1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{V_p^2}{V_{pp}}} - p - \frac{1}{2} p^2 + V_p (\phi + r) p. \quad (39)$$

We solve the problem in equation (37) by analyzing the above Ordinary Differential Equation (ODE) in equation (39).

Recall that we restrict the feasible incentive slopes $\{\beta_t\}$ to be bounded, i.e., there exists some sufficiently large constant M such that $\beta_t \in [-M, M]$. This is in the same spirit as imposing the transversality condition, because given bounded incentives $\{\beta_t\}$ the promised information rent p —as expected future discounted incentives—is also bounded.¹⁶ The proof in the Appendix relies on this boundedness assumption, and there we also show that the derived optimal policy is independent of M when M is sufficiently large. For details, see Appendix A.6.

4.3 Optimal Contracting

We present the main result of this paper in this section. Before we start analyzing the optimal contract, we first consider the trivial benchmark case. Suppose that the profitability θ_t is observable. This is essentially the classic 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

$$\mu_t^{HM} = \beta_t^{HM} = \frac{1}{1 + ar\sigma^2}, \quad (40)$$

and the principal’s value is $V^{HM} = 1 / (2r (1 + ar\sigma^2))$. The optimal contract can be implemented by a constant equity share $1 / (1 + ar\sigma^2)$ (with proper benchmarking). Because the principal can ignore the direct information about θ_t , the value V^{HM} serves as an upper bound for our value function $V(p)$ when profitability is unobservable:

$$V(p) \leq V^{HM} = \frac{1}{2r (1 + ar\sigma^2)}. \quad (41)$$

To solve the optimal contract, we analyze the ODE in equation (39) with the boundary condition in equation (62) using the technique of dynamic programming. In order to ensure the concavity of $V(p)$, we impose the following parametric condition throughout the paper, which restricts ϕ to be

¹⁶Careful readers may wonder the boundary conditions that we impose in equation (39). The bounded information rent p naturally gives us the boundary conditions for equation (39). In Appendix we show that $\beta_t \in [-M, M]$ implies information rent $\{p_t\}$ to be bounded, and $\{p_t\}$ is absorbing on its boundaries. This allows us to calculate the principal’s value $V(p)$ when p_t hits its boundaries. In analyzing the HJB equation (39) that characterizes the optimal contract, we impose these boundary conditions.

relatively small:

$$\frac{\phi}{r} > a \left[2 \left(\frac{\phi}{r} + 1 \right)^3 - \phi \sigma^2 \right]. \quad (42)$$

The following proposition is our main result which characterizes the properties of the value function, and hence the optimal policy $\{\beta^*, \sigma^{p,*}\}$ as in equation (38).

Proposition 2 *Suppose that (42) holds, and suppose that M is sufficiently large (recall we restrict $\beta_t \in [-M, M]$). We have the following properties for $V(p) \in \mathbb{C}^2$ which characterize the optimal contract.*

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

In the optimal contract, the principal sets the initial information rent p_0^* to be $\bar{p}$. Afterwards the information rent p_t^* as the state variable evolves according to equation (36). The optimal control is characterized by equation (38). Interestingly, property 3 in Proposition 2 states that the information rent p_t^* will never wander out of an endogenous interval $[0, \bar{p}]$. Recall that the information rent p_t^* is the (properly) discounted promised future incentives. Property 3 suggests that 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, in which the optimal incentives $\left\{ \beta_t^{HM} = \frac{1}{1+a r \sigma^2} \right\}$ are isolated across periods and remain constant over time. In our model with learning, incentive provisions are linked across periods, and the optimal incentives $\{\beta_t^*\}$ become stochastic. Nevertheless, because the model primitives are stationary (CARA-Normal setting, additive technology in equation (29), and stationary learning), the information rent $\{p_t^*\}$ remains endogenously bounded in the optimal contract.

5 Model Implications

For better understanding our results, we first analyze the case in which we restrict the incentives $\{\beta_t\}$ to be deterministic. We then compare the optimal policies with stochastic incentives to

both the Holmstrom and Milgrom (1987) benchmark with observable θ_t , and the contract with optimal deterministic incentives. The discussion focuses on the two qualitative features of optimal contracting: front-loaded incentives and option-like incentives.

5.1 Contract with Deterministic Incentives

We will show that the optimal incentives are front-loaded in dynamic contracting with learning. This result is best illustrated when we constrain the incentives $\{\beta\}$ to be deterministic (but can vary over time). This case also provides an important benchmark for the fully stochastic optimal contract.

When we search for the optimal contract in the contracting space with deterministic incentives, we can derive the time-decreasing incentives analytically. Not surprisingly, as shown shortly, the pattern of time-decreasing incentives is shared by the fully stochastic optimal contract solved in Section 4.3. The reason that $\{\beta\}$ being deterministic helps is that we can move the conditional expectation in equation (35) inside the integral, so that:¹⁷

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

Here, information rent is the sum of properly discounted future deterministic incentives, and it is deterministic as well.

Denote by $V^d(p)$ the value function with deterministic policies. Since p_t is deterministic, its volatility σ^p is zero. Plugging $\sigma^p = 0$ into (37), we have $\beta^d(p) = (1 + p - \phi V_p^d) / (1 + ar\sigma^2)$, with the HJB equation as:

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

The following proposition solves the above ODE in closed-form.

Proposition 3 *Within the class of deterministic policies, the value function for the optimal deterministic contract is quadratic*

$$V^d(p) = -\frac{1}{2} A^d p^2 + B^d p. \quad (43)$$

The evolution of information rent, incentive, and effort are given by:

$$\begin{aligned} p_t^d &= \frac{B^d}{A^d} e^{-\lambda t}, \beta_t^d = \frac{1 + A^d \phi}{1 + ar\sigma^2} p_t^d, \\ \mu_t^d &= \beta_t^d - p_t^d = \frac{A^d \phi - ar\sigma^2}{1 + ar\sigma^2} p_t^d, \end{aligned} \quad (44)$$

¹⁷This is because of the property of exponential martingale (recall that $\{\beta\}$ are bounded): $\mathbb{E}_t \left[\exp \left(- \int_t^s ar\beta_u \sigma dB_u - \frac{1}{2} \int_t^s a^2 r^2 \beta_u^2 \sigma^2 du \right) \right] = 1$.

where $\lambda \equiv -\phi - r + \frac{1+A^d\phi}{1+ar\sigma^2}\phi > 0$, $B^d \equiv 1/\phi$ and

$$A^d \equiv \frac{(2\phi + r)ar\sigma^2 + r + \sqrt{(2\phi + r)^2 a^2 r^2 \sigma^4 + 2ar\sigma^2 \left[(\phi + r)^2 + \phi^2\right] + r^2}}{2\phi^2}. \quad (45)$$

The above proposition shows that in the optimal deterministic contract, the information rent p_t^d , the incentive β_t^d , and the optimal effort μ_t^d all follow certain exponentially decaying paths (toward zero), with the same rate of $-\lambda$. Moreover, at $t = 0$, from equation (44) we have

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

Thus, the initial optimal effort level is below the observable profitability benchmark, and the optimal effort path is decreasing over time.

The optimality of the front-loaded effort policies comes from the forward looking nature of information rent. From the agent's incentive-compatibility condition in equation (19), the belief manipulation effect implies that giving incentives later tends to make the agent 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. Clearly, this result relies on the commitment ability in long-term contracting. In fact, Section 5.2.3 will show that when relationships are short-term, the resulting incentives and effort policies are constant over time.

Both Prat and Jovanovic (2014) and our model find front-loaded incentives to be optimal. Because Prat and Jovanovic (2014) implements a constant effort,¹⁸ the forward looking nature of information rent implies that the compensation contract has to offer front-loaded incentives. Our model allows the optimal contract to adjust on the effort margin (not just incentives), and cheaper incentive provisions in earlier periods naturally push the optimal contract to implement a front-loaded effort profile.

The front-loaded effort policies also arise in models with career concerns (e.g., Gibbons and Murphy, 1992, Holmstrom, 1999), but through a distinct mechanism. There, agents in their early careers face *higher uncertainty* in their abilities, and, thus, work harder to impress the market (but the market will not be fooled in equilibrium, a standard signal jamming problem). This force is not

¹⁸Prat and Jovanovic (2014) assume that the effort cost is linear over the feasible interval $[0, 1]$ and focus on implementing the highest effort level 1 (which is also assumed in DeMarzo and Sannikov (2014)). In addition, Prat and Jovanovic (2014) study the non-stationary case where the underlying profitability θ (as a parameter) never changes, and as time passes both parties get to learn the true profitability eventually. In Appendix B we show that the pattern of time-decreasing effort pattern is robust to this assumption.

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 $r = 0.5, a = 0.6, \sigma = 1.5, \phi = 0.022$.

present in our stationary model, as the uncertainty of the profitability/ability (i.e., the posterior variance of θ_t) stays constant over time.

Although both Gibbons and Murphy (1992) and our paper feature a front-loaded effort policy, the predictions regarding optimal incentive profiles are different. Due to career concerns, in Gibbons and Murphy (1992) the agent works hard even without high-powered in-job incentives. In contrast, all incentives in our model are from the long-term contract, and the front-loaded effort profile requires a front-loaded incentive contract. This is a common feature in dynamic contracting model with full commitment and learning, such as in Prat and Jovanovic (2014) and DeMarzo and Sannikov (2014).

5.2 Value Function and Optimal Policies

Now we return to our optimal contract among the space of fully stochastic incentives, and illustrate two qualitative properties of our optimal contract. First, as shown in Section 5.1 with the special case of deterministic contracts, the fully stochastic optimal contract features front-loaded incentives. Second, the optimal management of the agent's information rent leads to an option-like feature in

the optimal contract, i.e., incentives rise after good performance.

From now on we always refer to the optimal policies, and without risk of confusion we omit the superscript “*”. Figure 1 plots the value function $V(p)$, the optimal control $\{\beta(p), \sigma^p(p)\}$, and the associated optimal policy $\mu_t(p) = \beta_t(p) - p$ 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 Panel A in Figure 1. Panel B 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 the information rent rises after good performance shocks, an interesting property that will be discussed shortly.

What drives the stochastic contract to be superior to the deterministic one? It is because the stochastic contract implements a more efficient effort policy, closer to the higher Holmstrom and Milgrom (1987) effort benchmark. In Panel C, we show that the incentive $\beta(p)$ sits above the deterministic counterpart for almost the entire range (and, thus, gets closer to the Holmstrom and Milgrom (1987) benchmark level), except for low p ’s (i.e., close to zero). A similar pattern holds for the implemented effort $\mu(p) = \beta(p) - p$ in Panel D of Figure 1.

Interestingly, though not evident in Figure 1, when p is close to zero both the incentive slope $\beta(p)$ and the effort $\mu(p)$ drop below their deterministic counterparts. This result is a robust feature of the model. Indeed, with the aid of asymptotic analysis in equation (97) in Appendix C, one can analytically verify that the difference 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.

5.2.1 Option-like Incentives

In our model it is optimal to implement a history dependent effort policy. This is surprising: as the posterior variance only changes over time deterministically with a standard CARA-Normal setting and learning (in our stationary setting, it is a constant in particular), usually the resulting equilibrium effort profile is a deterministic process as well (e.g., Holmstrom, 1999).

To understand the economic mechanism that drives this result, we study how the history-dependent effort policies improve over 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, $(\beta'(p_t) - 1) \sigma^p dB_t$).

As shown in the top panels in Figure 2, these diffusion terms are positive. Thus, in contrast to the 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, pay-for-performance rises following good performance, suggesting that 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 rent in a long-term relation. As explained in Section 3.4, 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 rent:

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

The information rent captures the agent’s additional future rewards when the principal mistakenly attributes the higher-profitability-driven good performance to the agent’s effort, and this is why future incentives $\{\beta_s\}$ matter. Equally important, for a risk-averse agent, the amount of information rent also depends on his marginal utilities $u_c(c_s, \mu_s)$ when receiving manipulation benefits at future states.

Because future incentives β_s and future marginal utilities $u_c(c_s, \mu_s)$ enters the information rent p_t multiplicatively, a negative correlation between β_s and $u_c(c_s, \mu_s)$ lowers p_t . Intuitively, information rent can be reduced if the contract allocates greater belief manipulation benefits in states where the agent cares less. Interestingly, the option-like feature achieves this negative correlation. To see this, following a positive output shock, the agent becomes wealthier, hence a lower marginal utility $u_c(c_s, \mu_s) = -arv_s$.¹⁹ By making the optimal contract option-like, the principal raises incentives after good performance and, thus, imposes a negative correlation between incentives and marginal utility.

Now let us get back to 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 a high marginal utility). Because the information rent p is positively correlated with performance as indicated by Panel B in Figure 2, the stochastic optimal contract imposes lower incentives in the states where p gets close to zero.

¹⁹Formally, 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.

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 $r = 0.5$, $a = 0.6$, $\sigma = 1.5$, $\phi = 0.022$.

5.2.2 Time-decreasing effort policies

For the optimal contract with general stochastic incentives, we plot the drift of incentives β and effort μ in the two bottom panels in Figure 2. We have shown analytically in Section 5.1 that, because of the forward-looking information rent, the effort policy is decreasing over time in the optimal deterministic contract. Not surprisingly, this pattern persists in the fully stochastic optimal contract. Graphically, the front-loaded effort policy is reflected by the negative drifts in the bottom panels in Figure 2.

5.2.3 What if contractual relationship is short-term?

We want to emphasize that the above two features of our optimal contract is due to the interaction between long-term contracting and learning. The case of observable θ_t shuts down learning, and the Holmstrom and Milgrom (1987) result implies the optimal effort and incentives are constant over time. What if there is learning, but the contractual relationship is short-term?

Imagine the following setting with short-term contractual relationship, where a long-lived agent with unknown ability θ_t is working for a continuum of principals. At any time $t > 0$ there is one

principal who signs a short-term incentive contract with the agent. The relationship, however, only lasts for the interval $[t, t + dt]$. The short-term contract consists of a fixed wage α_t , an incentive β_t , and the recommended effort μ_t , so that at the beginning of $t + dt$ the agent receives a compensation flow of

$$\alpha_t dt + \beta_t (dY_t - \mu_t dt - m_t dt),$$

given date t belief, $\mathbb{E}_t[\theta_t] = m_t$, and output performance dY_t over $[t, t + dt]$. At the end of period $t + dt$, the relationship breaks, and the agent signs another contract $\{\alpha_{t+dt}, \beta_{t+dt}\}$ with another principal indexed by $t + dt$. Importantly, short-term relationships rule out inter-period commitment, implying each principal takes other principals' equilibrium offers as given.

For simplicity, to determine the history of fixed wages $\{\alpha_t\}$, we assign all the bargaining power to principals. We have the following proposition.

Proposition 4 *Suppose that contractual relationships are short-term and principals have all the bargaining power. Then in equilibrium incentives β_t^{ST} are constant over time:*

$$\beta_t^{ST} = \frac{\phi + r}{r + ar\sigma^2(\phi + r)} \text{ for all } t,$$

and the equilibrium effort μ_t^{ST} is constant over time as well

$$\mu_t^{ST} = \frac{r}{\phi + r} \beta_t^{ST} = \frac{r}{r + ar\sigma^2(\phi + r)} \text{ for all } t.$$

When the principals have all the bargaining power, Proposition 1 still applies to the agent's problem.²⁰ Thus, given today's incentive β_t^{ST} and future incentives $\{\beta_{t+s}^{ST} : s > 0\}$, the agent exerts $\mu_t^{ST} = \beta_t^{ST} - p_t^{ST}$, where p_t^{ST} is the discounted future incentives $\{\beta_{t+s}^{ST}\}$ defined as in equation (35). The time- t principal takes p_t^{ST} as given and maximizes the expected output $\beta_t^{ST} - p_t^{ST} + m_t$, minus the total compensation which is the sum of the effort cost $(\beta_t^{ST} - p_t^{ST})^2/2$ and the risk compensation $\frac{1}{2}ar\sigma^2(\beta_t^{ST})^2$. Ignoring the given project quality m_t , the time- t principal maximizes the flow payoff in equation (34) only:

$$\max_{\beta_t^{ST}} (\beta_t^{ST} - p_t^{ST}) - \frac{1}{2} (\beta_t^{ST} - p_t^{ST})^2 - \frac{ar\sigma^2}{2} (\beta_t^{ST})^2 \Rightarrow \beta_t^{ST} = \frac{1 + p_t^{ST}}{1 + ar\sigma^2} \quad (46)$$

Thus, with short-term contracting, each principal ignores the long-term consequence of the short-term incentives offered by herself. Stationarity implies both β_t^{ST} and p_t^{ST} are constants.

²⁰When the agent does not have any bargaining power, the proof in Proposition 4 shows that for the agent's problem, the short-term incentives $\{\beta_t\}$ here play the same role as the incentives $\{\beta_t\}$ in long-term contracts analyzed in Proposition 1.

Intuitively, without commitment, in short-term contracting each principal at different points of time solves his individual myopic (and stationary) problem in equation (46). In contrast, with long-term contracting, a single principal not only maximizes the flow payoff in equation (46), but also takes into account the effect of β_{t+s} on the forward-looking information rent p_t .²¹ This forward looking force in the full commitment environment, combined with learning, makes the optimal effort policy time-decreasing and stochastic.

5.3 Empirical Predictions

Our model has a few key implications which can be tested in the data. First, our model suggests that it is more efficient to assign higher incentives after good performance due to the lower marginal utility of the agent at that time. That is, the optimal contract derived in our dynamic setting with learning features an option-like payoff. On the other hand, traditional static models typically do not predict option grants to managers.²² For example, Dittmann and Maug (2007) calibrate a standard static structural model and find that most CEOs should hold more straight equity, hold no stock options, and receive lower salaries. In the data, there is strong evidence that the use of options is pervasive (e.g., Hall and Liebman (1998)).

Second, the option-like feature that incentives increase after good performance can be delivered in other ways than rewarding the manager stock options *ex ante*. Indeed, Core and Guay (1999) find that the annual grant of options and stocks to CEO is increasing in past stock returns. More recently, Bergman and Jenter (2006) document that option and stock grants per employee are increasing in past stock returns, and He, Li, Wei, and Yu (2014) report that incentives increase with past firm-level profitability. In addition, our model also implies that managerial incentives should be procyclical at the aggregate level. The idea is simple: even though aggregate economic conditions should be indexed out in the optimal contract, the fact that the agent's marginal utility in good time is low implies that it is relatively cheap to assign incentive there. This prediction is consistent with the empirical finding in Eisfeldt and Rampini (2008), who show that the Hodrick-Prescott filtered executive compensation is remarkably procyclical.

²¹This result is in contrast to Fudenberg, Holmstrom, and Milgrom (1990), in which learning is not present. They show that, with dynamic moral hazard only, the optimal long-term contract can be implemented by short-term ones under CARA preferences. In a way, their result suggests that commitment itself is not that important. In contrast, our model shows that the commitment in long-term contracting is important because of the long-lasting belief manipulation effect with endogenous learning.

²²There are a few exceptions. For instance, Dittmann, Maug, and Spalt (2010) show that a loss-averse utility function can produce convex contracts, and Dittmann and Yu (2010) show that convex contracts can be obtained if the agent also affects firm risk. Moreover, in a dynamic setting, Edmans and Gabaix (2011) show that the convexity of the contract depends on the marginal cost of effort. In Ju and Wan (2012), stock options become optimal when the agent has to be paid above a certain subsistence level.

Third, our model predicts that the industry or the firm with higher uncertainty should have more option-based contracts for managerial compensation. Indeed, both Ittner, Lambert, and Larcker (2003) and Murphy (2003) find that new-economy firms (such as, computer, software, the Internet, or telecommunication firms) tend to grant more stock options to managers than old-economy firms. This evidence is consistent with our model prediction since the new economy tends to be the industry with higher uncertainty.

There is one caveat in linking our optimal contracting results to compensation contracts in practice. As emphasized, we focus on long-term contracting in which full commitment is possible, which is theoretically appealing as it gives the upper bound of other long-term relations with partial commitment. In practice without full commitment, career concerns (Gibbons and Murphy, 1992, Holmstrom, 1999) are another important force, especially when the labor market is mobile and agents/workers can easily move. Therefore, our model applies more to the situation where human capital is more firm-specific and thus the long-term job security is a primary concern.

6 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, and thus, a “hidden information” problem. Utilizing the convenient property of CARA preference, we show that the optimal contracting can be reformulated to a dynamic programming problem with only one state variable, and characterize the optimal contract by the solution to an ODE. We show that the optimal effort decreases with tenure on average, and the optimal contract exhibits an option-like feature in the sense that incentives/effort rise after positive performance shocks. These two properties rely on the combination of learning and long-term contracting, as we show the resulting effort profile is constant over time if we have either short-term contracting with learning, or long-term contracting without learning (Holmstrom and Milgrom, 1987).

Although we are only able to give a full characterization of the optimal long-term contract with learning under a specific setting (e.g., CARA preferences, Gaussian processes, etc), the above two key qualitative results, i.e., front-loaded effort policy and option-like feature, will likely be robust to more general settings. The main reason we think the results extend is that the economic force behind these results do not depend on CARA preferences or Gaussian processes. The agent’s information

rent due to belief manipulation, i.e., the agent’s inclination to shirk to distort the principal’s future belief downward, is general in any long-term contracting environment with learning. The higher costs of providing incentives later leads to the time-decreasing effort policy, simply because later incentives enter the agent’s forward-looking information rent in earlier periods (but not the other way around). Additionally, the option-like feature comes from the fact that the agent is risk-averse so that the marginal value of in earning belief manipulation benefits is lower after good performance.

References

- Adrian, Tobias, and Mark M. Westerfield, 2009, Disagreement and learning in a dynamic contracting model, *Review of Financial Studies* 22, 3873–3906.
- Baron, D. P., and D. Besanko, 1984, Regulation and information in a continuing relationship, *Information Economics and Policy* 1, 267–302.
- Bergemann, D. and U. Hege, 1998, Venture capital financing, moral hazard, and learning, *Journal of Banking & Finance* 22, 703–735.
- Bergemann, D. and U. Hege, 2005, The financing of innovation: Learning and stopping, *RAND Journal of Economics* 36, 719–752.
- Bergman, Nittai and Dirk Jenter, 2007, Employee sentiment and stock option compensation, *Journal of Financial Economics* 84, 667–712.
- Biais, B., T. Mariotti, G. Plantin, and J. Rochet, 2007, Dynamic security design: Convergence to continuous time and asset pricing implications, *Review of Economic Studies* 74, 345–90.
- Core, John and Wayne Guay, 1999, The use of equity grants to manage optimal equity incentive levels, *Journal of Accounting and Economics* 28, 151–184.
- Core, J., Guay, W., Larcker, D., 2003. Executive equity compensation and incentives: a survey. *Economic Policy Review* 9, 27–50.
- 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.
- Cvitanic, Jakša, Xuhu Wan, and Huali Yang, 2013, Dynamics of contract design with screening, *Management Science* 59, 1229–1244.
- DeMarzo, P., M. Fishman, Z. He, and N. Wang, 2012, Dynamic agency and the q-theory of investment, *Journal of Finance* 67, 2295–2340.
- DeMarzo, P., and Y. Sannikov, 2006, Optimal security design and dynamic capital structure in a continuous-time agency model, *Journal of Finance* 61, 2681–2724.
- DeMarzo, Peter M., and Yuliy Sannikov, 2014, Learning, termination and payout policy in dynamic incentive contracts, working paper, Stanford University and Princeton University.
- Dittmann, Ingold and Ernst Maug, 2007, Lower salaries and no options? On the optimal structure of executive pay, *Journal of Finance* 62, 303–343.

- Dittmann, I., E. Maug, and O. Spalt, 2010, Sticks or carrots? Optimal CEO compensation when managers are loss-averse, *Journal of Finance* 65, 2015–2050.
- Dittmann, I., and K. C. Yu, 2010, How Important are risk-taking incentives for executive compensation? *Working Paper*, Erasmus University–Rotterdam.
- Edmans, Alex, and Xavier Gabaix, 2011, Tractability in incentive contracting, *Review of Financial Studies* 24, 2865–2894.
- Eisfeldt, Andrea and Adriano Rampini, 2008, Managerial incentives, capital reallocation, and the business cycle, *Journal of Financial Economics* 87, 177–199.
- 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.
- Garrett, Daniel F., and Alessandro Pavan, 2012, Managerial turnover in a changing world, *Journal of Political Economy* 120, 879–925.
- Gershkov, A. and M. Perry, 2012, Dynamic contracts with moral hazard and adverse selection, *Review of Economic Studies* 79, 268–306.
- Gibbons, Robert and Kevin J. Murphy, 1992, Optimal incentive contracts in the presence of career concerns: theory and evidence.” *Journal of Political Economy* 100, 468–505.
- Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski, 2012, Optimal dynamic taxes, working paper, Princeton, Yale, and Minnesota.
- Halac, Marina, and Navin Kartik, Qingmin Liu, 2012, Optimal contracts for experimentation, working paper, Columbia University.
- Hall, Brian J., and Kevin J. Murphy, 2003, The trouble with stock options, *Journal of Economic Perspectives* 17, 49–70.
- Hall, Brian and Jeffrey Liebman, 1998, Are CEOs really paid like bureaucrats, *Quarterly Journal of Economics* 113, 653–691.
- He, Zhiguo, 2009, Optimal executive compensation when firm size follows geometric brownian motion, *Review of Financial Studies* 22, 859–892.
- He, Zhiguo, 2011, A model of dynamic compensation and capital structure, *Journal of Financial Economics* 100, 351–366.
- He, Zhiguo, Si Li, Bin Wei, and Jianfeng Yu, 2014, Uncertainty, risk, and incentives: Theory and evidence, *Management Science* 60, 206–226.
- 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, 303–328.
- Horner, J. and L. Samuleson, 2012, Incentives for experimenting agents, working paper, Yale Univesrity.

Ittner, Christopher, Richard Lambert, and David Larcker, 2003, The structure and performance consequences of equity grants to employees of new economy firms, *Journal of Accounting and Economics* 34, 89–127.

Ju, Nengjiu, and Xuhu Wan, Optimal compensation and pay-performance sensitivity in a continuous-time principal-agent model, *Management Science*, forthcoming.

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

Malenko, Andrey, 2013, Optimal Dynamic Capital Budgeting, working paper, MIT.

Murphy, Kevin, 2003, Stock-based pay in new economy firms, *Journal of Accounting and Economics* 34, 129–147.

Pavan Alessandro, Ilya Segal, and Juuso Toikka, 2012, Dynamic mechanism design, working paper, Northwestern, Stanford, and MIT.

Prat, Julien, and Boyan Jovanovic, 2014, Dynamic Contracts when Agent’s Quality is Unknown, forthcoming, *Theoretical Economics*.

Piskorski, T., and A. Tchistyi, 2010, Optimal Mortgage Design, *Review of Financial Studies* 23, 3098–3140.

Revuz, Daniel, Yor, Marc, 1999, *Continuous Martingale and Brownian Motion*, 3rd, Springer.

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

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

Sannikov, Yuliy, 2014, Moral hazard and long-run incentives, working paper, Princeton University.

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

Sung, Jaeyoung, 2005, Optimal contracts under adverse selection and moral hazard: A continuous-time approach, *Review of Financial Studies* 18, 1021–1073.

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

Williams, Noah, 2011, Persistent private information, *Econometrica* 79, 1233–1274.

Zhang, Yuzhe, 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 He (2011). Consider any contract $\Pi = \{c_t, \mu_t\}$ which induces an optimal policy $\{c_t^*, \mu_t^*\}$ from the agent with a value v_0^* , so that

$$\begin{aligned} v_0^* &= \mathbb{E}^{\mu^*} \left[\int_0^\infty e^{-rt} u(c_t^*, \mu_t^*) dt \right] \\ \text{s.t. } dY_t &= \left(\mu_t^* + m_t^{\mu^*} \right) dt + \sigma dB_t^{\mu^*}, \\ dS_t &= rS_t dt + c_t dt - c_t^* dt \text{ with } S_0 = 0, \end{aligned}$$

The principal knows the resulting optimal effort policy $\{\mu_t^*\}$ and she updates her belief according to $\{\mu_t^*\}$ rather than the recommended effort policy $\{\mu_t\}$. From the agent's budget equation, we have

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

which gives the agent's optimal savings path. Note that if S_t is bounded then 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 $\Pi^* = \{c_t^*, \mu_t^*\}$. Clearly, taking consumption-effort policy $\{c_t^*, \mu_t^*\}$ is feasible for the agent with no private-savings. Now we show that $\{c_t^*, \mu_t^*\}$ is optimal for the agent given this contract.

Suppose, counter-factually, that given the contract Π^* , the agent finds that $\{c'_t, \mu'_t\}$ yields a strictly higher payoff $v'_0 > v_0^*$ in her problem, with associated savings path

$$S'_t = \int_0^t e^{r(t-s)} (c'_s - c_s^*) ds,$$

which satisfies the transversality condition. Formally, we have

$$\begin{aligned} v'_0 &= \mathbb{E}^{\mu'} \left[\int_0^\infty e^{-rt} u(c'_t, \mu'_t) dt \right] > v_0^* \\ \text{s.t. } dY'_t &= (\mu'_t + m_t^{\mu'}) dt + \sigma dB_t^{\mu'}, \\ dS'_t &= rS'_t dt + c'_t dt - c_t^* dt \text{ with } S_0 = 0. \end{aligned}$$

Now we construct an contradiction to the preassumption that “given $\Pi = \{c_t, \mu_t\}$ the agent's optimal policy is $\{c_t^*, \mu_t^*\}$ with a value of v_0^* .” Suppose that given $\Pi = \{c_t, \mu_t\}$, the agent takes the policy $\{c'_t, \mu'_t\}$ instead of $\{c_t^*, \mu_t^*\}$ which is claimed to be optimal. Because $v'_0 > v_0^*$ this alternative policy strictly dominates $\{c_t^*, \mu_t^*\}$; the only thing left is to verify whether the consumption plan is feasible given some saving policy. But, the saving policy $S''_t = S_t + S'_t = \int_0^t e^{r(t-s)} (c_s - c'_s) ds$ achieves $\{c'_t\}$ given the income process $\{c_t\}$, because

$$\begin{aligned} dS''_t &= dS_t + dS'_t = rS_t dt + c_t dt - c_t^* dt + [rS'_t dt + c'_t dt - c_t^* dt] \\ &= r(S_t + S'_t) dt + c_t dt - c'_t dt \\ &= rS''_t dt + c_t dt - c'_t dt, \end{aligned}$$

which also satisfies the transversality condition $\lim_{T \rightarrow \infty} \mathbb{E} [e^{-rT} S''_T] = 0$ if both S_t and S'_t satisfy the transversality condition. Thus, given the original contract Π , the saving rule $\{S''_t\}$ supports $\{c'_t, \mu'_t\}$ but delivers a strictly higher payoff v'_0 . This contradicts with the optimality of $\{c_t^*, \mu_t^*\}$ under the contract Π .

Finally, because the principal knows that $\{c_t^*, \mu_t^*\}$ is optimal for the agent, the principal still correctly knows the agent's actual optimal effort policy $\{\mu_t^*\}$ and thus perform the correct Bayes updating, and her payoff is the same as that under the contract $\Pi = \{c_t, \mu_t\}$. Hence it is without loss of generality to focus on contracts that are incentive-compatible and no-savings.

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

$$\begin{aligned} \max_{\{\hat{c}_s\}, \{\hat{\mu}_s\}} \quad & \mathbb{E}^{\hat{\mu}} \left[\int_t^\infty -\frac{1}{a} e^{-a(\hat{c}_s - \frac{1}{2}\hat{\mu}_s^2) - r(s-t)} ds \right] \\ \text{s.t.} \quad & dS_s = rS_s ds + c_s ds - \hat{c}_s ds, \quad S_t = S, \quad s > t \\ & dY_s = (\hat{\mu}_t + m_t^{\hat{\mu}}) dt + \sigma dB_s^{\hat{\mu}}, \end{aligned} \tag{47}$$

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

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

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

We claim that the solution to this problem is $\{c_s^* - rS, \mu_s^*\}$, and therefore the value is $v_t(0; \Pi) = e^{arS} v_t(S; \Pi)$. There are two steps to show this. First, this solution is feasible. Second, suppose that there exists another policy $\{\hat{c}'_s, \hat{\mu}'_s\}$ that is superior to $\{c_s^* - rS, \mu_s^*\}$, so that the associated value $v'_t(0; \Pi) > e^{-arS} v_t(S; \Pi)$. Consider $\{\hat{c}'_s + rS, \hat{\mu}'_s\}$, which is feasible to the problem in equation (47). Under this plan, however, the agent's objective is

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

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

A.3 Proof for Proposition 1

A.3.1 Necessity

Denote the $\{c, \mu\}$ to be the agent's (proposed) optimal consumption-effort policy given the compensation contract that satisfies the first-order condition stated in the proposition. The agent's continuation payoff v_t follows $dv_t = (-arv_t) \beta_t \sigma dB_t^\mu$ where $\{\beta\}$ are incentives specified by the contract. We will use the following property of $\{v\}$ later:

$$v_t = v_0 \exp \left(\int_0^t ar\beta_u \sigma dB_u^\mu - \int_0^t 0.5a^2 r^2 \beta_u^2 \sigma^2 du \right) v_0 = v_0 - \int_0^t arv_s \beta_s \sigma dB_s^\mu. \quad (48)$$

It is to show that when $|\beta| < M$ is bounded, v_t follows a martingale (Revuz and Yor, 1999, P139). This also verifies that v_t is the agent's equilibrium continuation payoff following the equilibrium consumption-effort policy.

We now establish the necessary conditions stated in the proposition by considering deviation strategies on effort and consumption policies respectively. First consider the deviation policy in effort, i.e. $\{\hat{c}_t, \hat{\mu}_t\} = \{c_t, \mu_t + \varepsilon \delta_t\}$, where the deviation policy $\{\delta_t \neq 0\}$ is arbitrary. Due to CARA preference we have

$$u(\hat{c}_t, \hat{\mu}_t) = u(c_t, \mu_t) e^{a\mu_t \varepsilon \delta_t + 0.5a\varepsilon^2 \delta_t^2}.$$

The agent's value under the deviation policy indexed by ε is simply

$$\hat{v}_0(\varepsilon) \equiv \mathbb{E}_0^{\hat{\mu}} \left[\int_0^\infty e^{-rt} r v_t e^{a\mu_t \varepsilon \delta_t + 0.5a\varepsilon^2 \delta_t^2} dt \right];$$

note that the expectation is under the measure induced by deviating effort profile $\hat{\mu}$. Equations (3), (5) and (17) imply that the effort profile $\hat{\mu}$, which depends on ε , induces a change of measure relative to μ by

$$\begin{aligned} \sigma dB_t^\mu - \sigma dB_t^{\hat{\mu}} &= (\hat{\mu}_t + m_t^{\hat{\mu}} - \mu_t - m_t^\mu) dt = (\varepsilon \delta_t + m_t^{\hat{\mu}} - m_t^\mu) dt \\ &= \left(\varepsilon \delta_t - \phi \int_0^t e^{-\phi(t-\tau)} (\varepsilon \delta_\tau) d\tau \right) = (\varepsilon \delta_t - \varepsilon \Delta_t) dt, \end{aligned} \quad (49)$$

where, as in (17), we denote

$$\Delta_s \equiv \phi \int_0^s e^{-\phi(s-u)} \delta_u du. \quad (50)$$

Hence we introduce the exponential martingale N_t , indexed by ε :

$$N_t(\varepsilon) \equiv \exp \left(\int_0^t \frac{\varepsilon \delta_s - \varepsilon \Delta_s}{\sigma} dB_s^\mu - \int_0^t \frac{(\varepsilon \delta_s - \varepsilon \Delta_s)^2}{2\sigma^2} ds \right), \quad \text{with } N_0(\varepsilon) = 1,$$

so that according to Girsanov theorem, we have

$$\widehat{v}_0(\varepsilon) = \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt} r v_t e^{a\mu_t \varepsilon \delta_t + 0.5a\varepsilon^2 \delta_t^2} dt \right] = \mathbb{E}_0^\mu \left[\int_0^\infty N_t(\varepsilon) e^{-rt} r v_t e^{a\mu_t \varepsilon \delta_t + 0.5a\varepsilon^2 \delta_t^2} dt \right].$$

Now we take derivative of $\widehat{v}_0(\varepsilon)$ with respect to ε , and evaluate it at $\varepsilon = 0$. Because

$$\left. \frac{dN_t(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = N_t(\varepsilon) \cdot \left[\int_0^t \frac{\delta_s - \Delta_s}{\sigma} dB_s^\mu - \int_0^t \frac{\varepsilon (\delta_s - \Delta_s)^2}{\sigma^2} ds \right] \Big|_{\varepsilon=0} = \int_0^t \frac{\delta_s - \phi \Delta_s}{\sigma} dB_s^\mu,$$

we have

$$\begin{aligned} \left. \frac{d\widehat{v}_0(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} &= \left\{ \mathbb{E}_0^\mu \left[\int_0^\infty \frac{dN_t(\varepsilon)}{d\varepsilon} \cdot e^{-rt} r v_t e^{a\mu_t \varepsilon \delta_t + 0.5a\varepsilon^2 \delta_t^2} dt \right] + \mathbb{E}_0^\mu \left[\int_0^\infty N_t(\varepsilon) e^{-rt} r v_t \frac{d(e^{a\mu_t \varepsilon \delta_t + 0.5a\varepsilon^2 \delta_t^2})}{d\varepsilon} dt \right] \right\} \Big|_{\varepsilon=0} \\ &= \underbrace{\mathbb{E}_0^\mu \left[\int_0^\infty \left(\int_0^t \frac{\delta_s - \Delta_s}{\sigma} dB_s^\mu \right) \cdot e^{-rt} r v_t dt \right]}_{AA} + \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt} a r v_t \mu_t \delta_t dt \right]. \end{aligned} \quad (51)$$

The first term AA equals to (using (48)):

$$\begin{aligned} AA &= \mathbb{E}_0^\mu \left[\int_0^\infty r e^{-rt} \left(\int_0^t \frac{\delta_s - \Delta_s}{\sigma} dB_s^\mu \right) \cdot \left(v_0 - \int_0^t a r v_s \beta_s \sigma dB_s^\mu \right) dt \right] \\ &= \mathbb{E}_0^\mu \left[\int_0^\infty r e^{-rt} \left(\int_0^t \frac{\delta_s - \Delta_s}{\sigma} dB_s^\mu \right) \left(- \int_0^t a r v_s \beta_s \sigma dB_s^\mu \right) dt \right] = -\mathbb{E}_0^\mu \left[\int_0^\infty r e^{-rt} \int_0^t a r \beta_s v_s (\delta_s - \Delta_s) ds \right] \\ &= -\mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt} a r \beta_t v_t (\delta_t - \Delta_t) dt \right], \end{aligned}$$

where the last line uses change of order of integration. Plugging this result back into (51), and using (48), we have

$$\begin{aligned} \left. \frac{d\widehat{v}_0(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} &= a r v_0 \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt - \int_0^t a r \beta_u \sigma dB_u^\mu - \int_0^t 0.5a^2 r^2 \beta_u^2 \sigma^2 du} [(\mu_t - \beta_t) \delta_t + \beta_t \Delta_t] dt \right] \\ &= a r v_0 \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt - \int_0^t a r \beta_u \sigma dB_u^\mu - \int_0^t 0.5a^2 r^2 \beta_u^2 \sigma^2 du} [(\mu_t - \beta_t) \delta_t] dt \right] \\ &\quad + \underbrace{a r v_0 \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt - \int_0^t a r \beta_u \sigma dB_u^\mu - \int_0^t 0.5a^2 r^2 \beta_u^2 \sigma^2 du} \beta_t \Delta_t dt \right]}_{BB}. \end{aligned} \quad (52)$$

Let us simplify the term BB in (52) further. Denote $Z_t \equiv - \int_0^t a r \beta_u \sigma dB_u^\mu - \int_0^t 0.5a^2 r^2 \beta_u^2 \sigma^2 du$ so that $\mathbb{E}_0^\mu [\exp(Z_t)] = \exp(Z_s)$ for $s < t$ under the condition of $|\beta| < M$ being bounded (see, e.g., Revuz and Yor, 1999, P139). Then by changing the order of integration, we can get (recall 50)

$$\begin{aligned} BB &= \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt + Z_t} \beta_t \Delta_t dt \right] = \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt + Z_t} \beta_t \phi \left(\int_0^t e^{-\phi(t-s)} \delta_s ds \right) dt \right] \\ &= \mathbb{E}_0^\mu \left[\int_0^\infty \delta_s \phi \mathbb{E}_s^\mu \left[\int_s^\infty e^{-rt + Z_t} e^{-\phi(t-s)} \beta_t dt \right] ds \right] \\ &= \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rs + Z_s} \delta_s \phi \mathbb{E}_s^\mu \left[\int_s^\infty e^{Z_t - Z_s} e^{-(r+\phi)(t-s)} \beta_t dt \right] ds \right]. \end{aligned}$$

As a result, we have

$$\left. \frac{d\widehat{v}_0(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = a r v_0 \mathbb{E}_0^\mu \left[\int_0^\infty e^{-rt + Z_t} \delta_t \left(\mu_t - \beta_t + \phi \mathbb{E}_t^\mu \left[\int_t^\infty \beta_s e^{-(r+\phi)(s-t)} e^{Z_s - Z_t} ds \right] \right) dt \right]$$

This implies that any incentive-compatible and no-saving policy must satisfy

$$\mu_t = \beta_t - \phi \mathbb{E}_t^\mu \left[\int_t^\infty \beta_s e^{-(r+\phi)(s-t)} e^{Z_s - Z_t} ds \right], a.s.$$

Otherwise, we can choose negative δ_t when $\mu_t > \beta_t - \phi \mathbb{E}_t^\mu \left[\int_t^\infty \beta_s e^{-(r+\phi)(s-t)} e^{Z_s - Z_t} ds \right]$, and positive δ_t when $\mu_t < \beta_t - \phi \mathbb{E}_t^\mu \left[\int_t^\infty \beta_s e^{-(r+\phi)(s-t)} e^{Z_s - Z_t} ds \right]$ (note that arv_0 is negative). Then a deviation strategy $\{c_t, \mu_t + \varepsilon \delta_t\}$ for sufficiently small ε will be profitable, leading to a contradiction.

The necessary conditions for the equilibrium consumption plan are much more standard. Fixing μ , it is easy to show that the necessary first-order condition for the agent's consumption-saving problem is that his marginal utility from consumption, i.e., $u_c(c_t, \mu_t)$ follows a martingale. Because $u_c(c_t, \mu_t) = -au(c_t, \mu_t)$ for exponential utility, and $v_t = \mathbb{E}_t^\mu \left[\int_t^\infty e^{-r(s-t)} u(c_t, \mu_t) dt \right]$, the result follows easily. QED.

A.3.2 Sufficiency

Consider any alternative policy $\{\hat{c}, \hat{\mu}\}$ deviating from the original policy $\{c, \mu\}$, with an expected payoff $\mathbb{E}_0^{\hat{\mu}} \left[\int_0^\infty e^{-rs} u(\hat{c}_s, \hat{\mu}_s) ds \right]$. To prove that this deviation payoff cannot exceed the equilibrium payoff v , we are following the idea in Sannikov (2014) by constructing an upper bound for the deviation policies.

To construct the upper bound for deviation payoffs, we will show that it is sufficient to keep track of two deviation state variables that matter to the agent's potential deviation value. The first deviating state variable captures the agent's private saving

$$S_t \equiv \int_0^t e^{r(t-s)} (c_s - \hat{c}_s) ds.$$

The second deviating state variable, denoted by X_t , captures the persistent belief manipulation effect:

$$\begin{aligned} X_t &\equiv \int_0^t e^{r(t-s)} \left[(\beta_s - \mu_s) (\mu_s - \hat{\mu}_s) + \beta_s (m_s^\mu - m_s^{\hat{\mu}}) \right] ds \\ &= \int_0^t e^{r(t-s)} \left[(\beta_s - \mu_s) (\mu_s - \hat{\mu}_s) + \beta_s \left(\phi \int_0^s e^{\phi(u-s)} (\hat{\mu}_u - \mu_u) du \right) \right] ds, \end{aligned} \quad (53)$$

where we have used $m_t^\mu - m_t^{\hat{\mu}} = \phi \int_0^t e^{\phi(u-t)} (\hat{\mu}_u - \mu_u) du$ as in (17). Denote by $\delta_s \equiv \mu_s - \hat{\mu}_s$ the effort deviation at s , we can rewrite X_t as

$$\begin{aligned} X_t &= \int_0^t e^{r(t-s)} (\beta_s - \mu_s) \delta_s ds - \int_0^t e^{r(t-s)} \beta_s \left(\phi \int_0^s e^{\phi(u-s)} \delta_u du \right) ds \\ &= \int_0^t e^{r(t-s)} (\beta_s - \mu_s) \delta_s ds - \int_0^t e^{r(t-s)} \delta_s \phi \left(\int_s^t e^{-(r+\phi)(u-s)} \beta_u du \right) ds \\ &= \int_0^t e^{r(t-s)} \delta_s \left[\beta_s - \mu_s - \phi \left(\int_s^t e^{-(r+\phi)(u-s)} \beta_u du \right) \right] ds, \end{aligned}$$

which gives an explicit expression as the history of deviation $\{\delta_s\}$.

Given these two variables, we construct a candidate of an upper bound for the agent's deviation value, which is defined as

$$W(S_t, X_t) \equiv v_t \exp(-ar(S_t + X_t)), \quad (54)$$

with the evolution of (note dS_t and dX_t do not have diffusion terms)

$$dW_t = e^{-ar(S_t + X_t)} dv_t - arv_t e^{-ar(S_t + X_t)} (dS_t + dX_t). \quad (55)$$

If $W(S_t, X_t)$ is indeed the upper bound of deviation value, then for an agent who has not deviated yet with $S_t = X_t = 0$, the upper bound of his deviation value is just v_t .

There are two points worth making:

1. There is some nonlinearity involved between the manipulated belief divergence and the private savings for the agent's upper bound deviation value, because X_t and S_t enter the upper bound in a multiplicative way, i.e.

$$W(S_t, X_t) = v_t e^{-arS_t} e^{-arX_t}.$$

Hence, $W(S_t, X_t)$ captures the potential inter-dependence between deviating incentives of consumption and effort.

2. There might be other state variable, e.g. p_t , that affects the contract faced by the agent, which matters for the agent's *exact* deviation value. However, we are not trying characterize the agent's *exact* deviation value (which p_t should matter); rather, we are identifying an upper bound for the deviation value, in which case p_t not necessarily matters.²³ As we show below, this particular upper bound for deviation depends only on the equilibrium payoff v_t as in (54).

We further require the following assumptions on the agent's deviation strategies for the usual transversality conditions, which are standard in infinite-horizon consumption/saving problems.

Assumption 1 (Transversality conditions). There exist (however large) positive constants L_s and L_x , so that the agent's deviation strategies satisfy

$$|S_t| < L_s, \text{ and } |X_t| < L_x.$$

We have the following key lemma which shows that $W(S_t, X_t)$ is the upper bound of the agent's deviation value.

Lemma 3 *Facing the contract $\{c_t, \mu_t\}$, suppose that the agent's deviation history leads to a pair of deviation states as (S_t, X_t) at time t . Then the agent's deviation value from time t onwards is bounded above by $W(S_t, X_t)$, if the assumption of transversality conditions holds.*

Proof. We first give the outline of the argument. To prove $W(S_t, X_t)$ is an upper bound for the agent's deviation value, define the auxiliary gain process G_t associated with any feasible policies $\{\hat{c}, \hat{\mu}\}$ to be

$$G_t(\{\hat{c}_s, \hat{\mu}_s\}_{s=0}^\infty) \equiv \int_0^t e^{-rs} u(\hat{c}_s, \hat{\mu}_s) ds + e^{-rt} W(S_t, X_t).$$

Clearly, $\mathbb{E}_0^\mu[G_\infty] = \mathbb{E}_0^\mu[\int_0^\infty e^{-rs} u(\hat{c}_s, \hat{\mu}_s) ds]$ is the expected payoff under the feasible policy, given the transversality condition $\lim_{s \rightarrow \infty} \mathbb{E}_0^\mu[e^{-rt} W_t] = 0$ (which is implied by Assumption 1 for transversality conditions). On the other hand, $G_0 = W(S_0, X_0)$ is the proposed upper bound of agent's deviation value given the current relevant deviation states (S_0, X_0) . Obviously, one sufficient condition for $\mathbb{E}_0^\mu[G_\infty] \leq G_0 = W(S_0, X_0)$, i.e., the upper bound (54) is valid, is that the auxiliary gain process G_t is a supermartingale for any deviation policy from the agent's perspective.

Now we start proof. For ease of notation, denote

$$\hat{u}_t \equiv u(\hat{c}_t, \hat{\mu}_t), \quad u_t \equiv u(c_t, \mu_t), \quad \text{with } u_t = rv_t.$$

Differentiating G_t with respect to t , and using (55) we find that

$$e^{rt} dG_t = (\hat{u}_t - rW_t) dt + dW_t = (\hat{u}_t - rW_t) dt + e^{-ar(S_t+X_t)} dv_t - arv_t e^{-ar(S_t+X_t)} (dS_t + dX_t).$$

Using $\sigma dB_t^\mu = \sigma d\hat{B}_t^\mu + (\hat{\mu}_t + m_t^\mu - \mu_t - m_t^\mu) dt$ and (48) we have

$$\begin{aligned} e^{rt} dG_t &= [\hat{u}_t - rW_t] dt - ar\beta_t v_t e^{-ar(S_t+X_t)} \left[\sigma d\hat{B}_t^\mu + (\hat{\mu}_t + m_t^\mu - \mu_t - m_t^\mu) dt \right] - arv_t e^{-ar(S_t+X_t)} (dS_t + dX_t) \\ &= -ar\beta_t v_t e^{-ar(S_t+X_t)} \sigma d\hat{B}_t^\mu + [\hat{u}_t - u_t e^{-ar(S_t+X_t)}] dt \\ &\quad - arv_t e^{-ar(S_t+X_t)} \left(\beta_t [\hat{\mu}_t - \mu_t - (m_t^\mu - m_t^\mu)] dt + dS_t + dX_t \right) \\ &= [\dots] d\hat{B}_t^\mu + u_t e^{-ar(S_t+X_t)} \left\{ e^{-a(\hat{c}_t - \frac{1}{2}\hat{\mu}_t^2) + a(c_t - \frac{1}{2}\mu_t^2) + ar(S_t+X_t)} - 1 \right\} dt \\ &\quad + u_t e^{-ar(S_t+X_t)} \left\{ -a\beta_t [\hat{\mu}_t - \mu_t - (m_t^\mu - m_t^\mu)] dt - a(dS_t + dX_t) \right\}, \end{aligned}$$

where [...] stands for $-ar\beta_t v_t e^{-ar(S_t+X_t)} \sigma$. Using the fact that

$$dS_t = (rS_t + c_t - \hat{c}_t)dt, \text{ and } dX_t = rX_t dt + [(\beta_t - \mu_t)(\mu_t - \hat{\mu}_t) + \beta_t(m_t^\mu - m_t^\mu)] dt,$$

²³One simple example will be that $V(a, p) \equiv a - p^2 \leq a \equiv W(a)$ which is independent of p .

we know that the drift of $e^{rt}dG_t$ equals

$$u_t e^{-ar(S_t+X_t)} \left\{ e^{-a(\hat{c}_t - \frac{1}{2}\hat{\mu}_t^2) + a(c_t - \frac{1}{2}\mu_t^2) + ar(S_t+X_t)} - 1 \right\} + u_t e^{-ar(S_t+X_t)} \{ -a\mu_t [\hat{\mu}_t - \mu_t] - a((rS_t + c_t - \hat{c}_t) + rX_t) \}.$$

But because $e^x \geq 1 + x$ and $u_t < 0$, the drift is bounded above by

$$\begin{aligned} & u_t e^{-ar(S_t+X_t)} \left\{ -a \left(\hat{c}_t - \frac{1}{2}\hat{\mu}_t^2 \right) + a \left(c_t - \frac{1}{2}\mu_t^2 \right) + arS_t + arX_t - a\mu_t [\hat{\mu}_t - \mu_t] - a((rS_t + c_t - \hat{c}_t) + rX_t) \right\} \\ = & u_t e^{-ar(S_t+X_t)} \left(\frac{a}{2} (\hat{\mu}_t^2 - \mu_t^2) + a\mu_t (\mu_t - \hat{\mu}_t) \right) = u_t e^{-ar(S_t+X_t)} \left(\frac{a}{2} (\hat{\mu}_t - \mu_t)^2 \right) \leq 0, \end{aligned}$$

which implies a negative drift for dG_t .

To summarize, we have shown that

$$dG_t = \text{negative_drift} - e^{-rt} ar\beta_t v_t e^{-ar(S_t+X_t)} \sigma dB_t^{\hat{\mu}}.$$

To ensure that G_t is a supermartingale, we need to ensure that

$$\mathbb{E}_0^{\hat{\mu}} \left[\int_0^T e^{-rt} ar\beta_t v_t e^{-ar(S_t+X_t)} \sigma dB_t^{\hat{\mu}} \right] = 0 \text{ for all } T.$$

Since $|S_t|$ and $|X_t|$ are bounded, we only need to ensure the square integrability condition (Revuz and Yor, 1999, P139):

$$\mathbb{E}_0^{\hat{\mu}} \left[\int_0^T (e^{-rt} v_t)^2 dt \right] < \infty \text{ for all } T.$$

Under $\hat{\mu}$, using (49) we have

$$\frac{dv_t}{v_t} = -ar\beta_t \sigma dB_t^{\hat{\mu}} = -ar\beta_t \sigma dB_t^{\hat{\mu}} - ar\beta_t (\delta_t - \Delta_t) dt,$$

which implies that

$$v_t = v_0 \exp \left[\int_0^t -ar\beta_s \sigma dB_s^{\hat{\mu}} - \int_0^t 0.5a^2 r^2 \beta_s^2 \sigma^2 ds - \int_0^t ar\beta_s [\delta_s - \Delta_s] ds \right].$$

Denote the bounds for effort and incentives by L_μ and M , so that $|\mu_t| < L_\mu$ and $|\beta_t| < M$. Then we have

$$\begin{aligned} \left| \int_0^t ar\beta_s [\delta_s - \Delta_s] ds \right| &= ar \left| \int_0^t \delta_s \left[\beta_s - \phi \int_s^t e^{\phi(s-u)} \beta_u du \right] ds \right| < ar \int_0^t |\delta_s| \left| \beta_s - \phi \int_s^t e^{\phi(s-u)} \beta_u du \right| ds \\ &< ar \int_0^t 2L_\mu \max \left(\beta_s, \phi \int_s^t e^{\phi(s-u)} \beta_u du \right) ds < 2arL_\mu Mt, \\ \left| \int_0^t 0.5a^2 r^2 \beta_s^2 \sigma^2 ds \right| &< 0.5a^2 r^2 \sigma^2 M^2 t. \end{aligned}$$

Hence,

$$\begin{aligned} \int_0^T (e^{-rt} v_t)^2 dt &= \int_0^T (e^{-rt} v_0)^2 \exp \left[\int_0^t -2ar\beta_s \sigma dB_s^{\hat{\mu}} - \int_0^t a^2 r^2 \beta_s^2 \sigma^2 ds - \int_0^t 2ar\beta_s [\delta_s - \Delta_s] ds \right] dt \\ &< \int_0^T (v_0)^2 e^{4arL_\mu Mt - 2rt} \exp \left[\int_0^t -2ar\beta_s \sigma dB_s^{\hat{\mu}} - \int_0^t 2a^2 r^2 \beta_s^2 \sigma^2 ds \right] \exp \left[\int_0^t a^2 r^2 \beta_s^2 \sigma^2 ds \right] dt \\ &< \int_0^T (v_0)^2 e^{4arL_\mu Mt + a^2 r^2 \sigma^2 M^2 t - 2rt} \exp \left[\int_0^t -2ar\beta_s \sigma dB_s^{\hat{\mu}} - \int_0^t 2a^2 r^2 \beta_s^2 \sigma^2 ds \right] dt. \end{aligned}$$

Because $\int_0^t (2ar\beta_s \sigma)^2 ds < (2arM\sigma)^2 t$ for all t , $\exp \left[\int_0^t -2ar\beta_s \sigma dB_s^{\hat{\mu}} - \int_0^t 2a^2 r^2 \beta_s^2 \sigma^2 ds \right]$ is an exponential martingale under the measure induced by $\hat{\mu}$. Therefore, for all T , we have

$$\begin{aligned} \mathbb{E}_0^{\hat{\mu}} \left[\int_0^T (e^{-rt} v_t)^2 dt \right] &< \int_0^T e^{4arL_\mu Mt + a^2 r^2 \sigma^2 M^2 t - 2rt} (v_0)^2 \mathbb{E}_0^{\hat{\mu}} \left[\exp \left(\int_0^t -2ar\beta_s \sigma dB_s^{\hat{\mu}} - \int_0^t 2a^2 r^2 \beta_s^2 \sigma^2 ds \right) \right] dt \\ &= (v_0)^2 \int_0^T e^{4arL_\mu Mt + a^2 r^2 \sigma^2 M^2 t - 2rt} dt < \infty. \end{aligned}$$

Now given the fact that G_t is a supermartingale, we have

$$W(S_0, X_0) = G_0 \geq \mathbb{E}_0^{\hat{\mu}} \lim_{t \rightarrow \infty} [G_t] = \mathbb{E}_0^{\hat{\mu}} \left[\int_0^\infty e^{-rs} u(\hat{c}_s, \hat{\mu}_s) ds + \lim_{t \rightarrow \infty} [e^{-rt} W_t] \right] = \mathbb{E}_0^{\hat{\mu}} \left[\int_0^\infty e^{-rs} u(\hat{c}_s, \hat{\mu}_s) ds \right],$$

which is the agent's deviation payoff. Here, the last equality requires the transversality condition which is ensured by the assumption of bounded $|S_t|$ and $|X_t|$. This implies that the bound (54) is indeed the upper bound for the agent's deviation value. ■

We have shown that $W(S_t, X_t)$ is an upper bound for the agent's potential deviation value given the deviated states (S_t, X_t) . Then, for an agent who has not deviated yet with $S_t = X_t = 0$, the upper bound of his deviation value is just v_t . Because the equilibrium strategy achieves this upper bound v_t , the equilibrium strategy is indeed globally optimal. As a result, we have shown that the equilibrium strategy which achieves v is indeed optimal. Q.E.D.

A.4 Appendix for Section 1

To see the first line is the agent's total payoff from time t onwards given any effort policy μ and c , 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 $\mathbb{E}_t^\mu [e^{rt} G(\infty)] = e^{rt} G(t) + \mathbb{E}_t^\mu \left[\int_t^\infty e^{-r(s-t)} dv_s \right]$, which is $u(c_t, \mu_t) dt + v_t + \mathbb{E}_t^\mu \left[\int_t^\infty e^{-r(s-t)} dv_s \right]$ by ignoring utilities occurring before t . Under equilibrium effort, dv_s is martingale increment and thus $\mathbb{E}_t^\mu \left[\int_t^\infty e^{-r(s-t)} dv_s \right] = 0$.

A.5 Appendix for Section 4.1

First of all, as m_t follows a martingale with $m_0 = 0$, we have

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

And, for wage cost, we have,

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

where we have used equation (31) in the last equality and the fact that β_t 's are bounded (so $\mathbb{E} \left[\int_0^\infty e^{-rt} \frac{dv_t}{v_t} dt \right] = -\mathbb{E} \left[\int_0^\infty e^{-rt} ar \beta_t \sigma dB_t \right] = 0$).

A.6 Appendix for Section 4.2

Define $\pm M_p \equiv \pm \frac{\phi M}{\phi + r}$. The following lemma shows that the information rent p_t is bounded in $[-M_p, M_p]$ if incentives $\{\beta_t\}$ are bounded in $[-M, M]$. Further, the boundaries for p_t are absorbing.

Lemma 4 Suppose that $\beta_t \in [-M, M]$ where M is a given constant. Then the state variable p_t reaches $\pm M_p$ if and only if $\beta_s = \pm M, \forall s \geq t$, which implies that $\pm M_p$ are absorbing states for p . As a result, when $p = \pm M_p$, $V(\pm M_p)$ is quadratic in M as in (62).

Proof. Suppose that the control variable is constrained such that $\beta_t \in [-M, M]$, where $M > 0$ is an arbitrarily large, but fixed constant. Recall the definition of p_t , and we have

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

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:

$$\begin{aligned} V(\pm M_p) &= \int_0^\infty e^{-rt} \left(\pm \frac{rM}{\phi+r} - \frac{1}{2} \left(\frac{rM}{\phi+r} \right)^2 - \frac{1}{2} a^2 r^2 \sigma^2 M^2 \right) ds \\ &= \pm \frac{M}{\phi+r} - \frac{r}{2} \left[\frac{1}{(\phi+r)^2} + a^2 \sigma^2 \right] M^2. \end{aligned} \quad (56)$$

Now we derive the evolution of p_t when it lies inside $(-M_p, M_p)$. 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 the 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 $d\left(\frac{1}{v_t}\right) = \frac{1}{v_t} (ar \beta_t \sigma dB_t + (ar \beta_t \sigma)^2 dt)$. Hence, according to Ito's lemma,

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

Therefore, we have

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

which is equation (36).

A.7 Proof for Proposition 2

Now we write problem (34) in the standard dynamic programming language. Substituting $\mu_t = \beta_t - p_t$ in the principal's objective, we have

$$\begin{aligned} \max_{\{\beta_t, \sigma_t^P\}} \mathbb{E} \left\{ \int_0^\infty e^{-rt} \left[(\beta_t - p_t) - \frac{1}{2} (\beta_t - p_t)^2 - \frac{1}{2} ar \sigma^2 \beta_t^2 \right] dt \right\} \\ \text{s.t. } dp_t &= [(\phi+r) p_t + \beta_t (ar \sigma \sigma_t^P - \phi)] dt + \sigma_t^P dB_t \text{ for all } t > 0, \text{ and } p_0 = p \\ \beta_s &\in [-M, M], p_t \in [-M_p, +M_p], \text{ and } p_t = \pm M_p \text{ are absorbing.} \end{aligned} \quad (57)$$

We are after the optimal policy $\{\beta_t^*, \sigma_t^{P,*}\}$ as functions of the state variable p_t .

A.7.1 Step 0: Relaxed problem and parameter restrictions

For ease of argument, we first consider the principal's relaxed maximization problem given M (and $M_p = \frac{\phi M}{\phi+r}$):

$$V(p; M) \equiv \max_{\{\beta_t, \sigma_t^p\}} \mathbb{E} \left[\int_0^\infty e^{-rt} \left(\mu_t - \frac{1}{2} \mu_t^2 - \frac{1}{2} ar\sigma^2 \beta_t^2 \right) dt \right] \quad (58)$$

$$s.t. \ dp_t = [(\phi+r)p_t + \beta_t(ar\sigma\sigma_t^p - \phi)]dt + \sigma_t^p dB_t \text{ for all } t > 0, \text{ and } p_0 = p$$

$$p_t \text{ is absorbing at } \pm M_p,$$

$$\beta_t \text{ can exceed } M \text{ (but remains finite) when } p_t \in (-M_p, M_p). \quad (59)$$

Problem (58) is a relaxed version of the principal's original problem (34), due to equation (59). Essentially, given M , in the original problem (34) we require $\beta_t \in [-M, M]$ for any time t ; while in problem (58) we only require $\beta_t \in [-M, M]$ whenever p_t hits $\pm M_p$ in light of Lemma 4. Thus, when $p = \pm M_p$, the boundary conditions are the same between these two problems. However, this relaxation helps because the relaxed problem (58) allows us to use the interior first-order condition of β when $p \in (-M_p, M_p)$. We will show that for sufficiently large M the value achieved in the relaxed problem is the same as that in the original problem, which implies that the solution to the relaxed problem is also that to the original problem.

Let us define two useful functions, which prove to be useful in saving notations.

$$H_1(p) \equiv \frac{1}{\phi}p - \frac{1}{2} \frac{r}{\phi^2} p^2 - \frac{a\sigma^2(\phi+r)^2}{2\phi^2} p^2,$$

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

And, the constant $\bar{p}^d$ in Proposition 2 is defined as:

$$\bar{p}^d \equiv \frac{2\phi}{(2\phi+r)ar\sigma^2 + r + \sqrt{(2\phi+r)^2 a^2 r^2 \sigma^4 + 2ar\sigma^2[(\phi+r)^2 + \phi^2] + r^2}}.$$

We state assumptions required for the proof below.

Assumption 2: The parameters satisfy the following condition:

$$a \left[2 \left(\frac{\phi}{r} + 1 \right)^3 - \phi\sigma^2 \right] < \frac{\phi}{r}. \quad (60)$$

Assumption 3: 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, with the boundary condition (62) given in equation (56):

$$rV(p) = \frac{1}{2} \frac{(1+p-\phi V_p(p))^2}{1+ar\sigma^2+a^2r^2\sigma^2 \frac{(V_p(p))^2}{V_{pp}}} - p - \frac{1}{2} p^2 + V_p(p)(\phi+r)p, \quad (61)$$

$$s.t. V(\pm M_p) = H_1(\pm M_p). \quad (62)$$

We proceed our proof with the following four steps.

Step 1: We first focus on the ODE in equation (61). 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 equation (61) 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 in equation (61) 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.7.2 Step 1: $V(0) = 0$ and $V_p(0) = \frac{1}{\phi}$.

We show that $V(0) = 0$ and $V_p(0) = \frac{1}{\phi}$ by following the ensuing three steps.

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

$$1 + \underline{p} - \phi V_p(\underline{p}) = 0. \tag{63}$$

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 (63) and (61) imply $V(0) = 0$ and $V_p(0) = \frac{1}{\phi}$. Differentiating the HJB in equation (61) 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 (63), imply that

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

Therefore, either $\underline{p} = 0$ or $V_{pp}(\underline{p}) = 0$. We first rule out the case of $V_{pp}(\underline{p}) = 0$. Note that this case implies the denominator of the first term in the right hand side of equation (61) is infinite, so that V must satisfy

$$rV(\underline{p}) = -\underline{p} - \frac{1}{2}\underline{p}^2 + (1 + \underline{p})\left(1 + \frac{r}{\phi}\right)\underline{p}, \tag{64}$$

Further, it follows from Taylor expansion that

$$\begin{aligned} V(\underline{p} + \epsilon) &= V(\underline{p}) + \frac{1}{\phi}(1 + \underline{p})\epsilon + o(\epsilon^2) \\ V_p(\underline{p} + \epsilon) &= \frac{1}{\phi}(1 + \underline{p}) + o(\epsilon). \end{aligned} \tag{65}$$

Thus, evaluating the HJB equation (61) at $\underline{p} + \epsilon$, we have

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

where the second equality uses the fact $\frac{V_p^2(\underline{p}+\epsilon)}{V_{pp}(\underline{p}+\epsilon)}$ goes to infinity as ϵ goes to zero because the continuity of $V_{pp}(\underline{p})$ implies that $V_{pp}(\underline{p}+\epsilon)$ is at the order of $o(1)$, and the last equality follows from equation (64). But this contradicts with equation (65) since they do not match at the second order ϵ^2 . As a result, $\underline{p} = 0$ and thus $V_p(0) = \frac{1}{\phi}$.

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

$$V_{pp}(\underline{p}) = -\frac{a^2r^2\sigma^2 V_p^2(\underline{p})}{1 + ar\sigma^2} = -\frac{a^2r^2\sigma^2}{1 + ar\sigma^2} \frac{1}{\phi^2} (1 + \underline{p})^2, \quad (66)$$

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

$$\frac{(1 + p - \phi V_p)^2}{1 + ar\sigma^2 + a^2r^2\sigma^2 \frac{V_p^2}{V_{pp}}} = \frac{2(1 + p - \phi V_p)(1 - \phi V_{pp})}{a^2r^2\sigma^2 \frac{2V_p V_{pp}^2 - V_p^2 V_{ppp}}{V_{pp}^2}}. \quad (67)$$

Differentiating the HJB equation (61), we have,

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

Plugging in equation (67) into the above equation, we find that the two terms in the first line cancel each other, and

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

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

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

or we have $V_{pp}(\underline{p}) = 0$ and $\underline{p} = -1$ due to equation (66). Furthermore, in the second case, we have

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

In Lemma 7 below,²⁴ we will show that for any p , if

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

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

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 (61) implies that

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

Thus, it follows from equation (38) that the policy function $\beta(p)$ at $p = 0$ has value

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

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

²⁴It 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.

A.7.3 Step 2: Prove the concavity and positivity

We will use the results of optimal deterministic contract (i.e., the optimal contract within the space in which $\{\beta_t\}$ are deterministically time-varying) with value function $V^d(p)$ studied in Proposition 3 in Section 5.1.

Step 2.a: Concavity & Positivity of the Denominator at $p = 0$: We want to show that under assumption 2-3, 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. Define the following constants

$$A^d \equiv \frac{(2\phi + r)ar\sigma^2 + r + \sqrt{(2\phi + r)^2 a^2 r^2 \sigma^4 + 2ar\sigma^2[(\phi + r)^2 + \phi^2] + r^2}}{2\phi^2}, \quad A^o \equiv r/\phi^2, \text{ and } B^d \equiv 1/\phi.$$

Lemma 5 *Denote*

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

Under the condition in equation (42), for $a > 0$, we have $\chi^o < \chi$ and

$$A^d - A^o > \frac{ar\sigma^2}{\phi} \left(1 + \frac{r}{\phi}\right) > \frac{2a^2\sigma^2\phi}{r(1 + ar\sigma^2)} \left(1 + \frac{r}{\phi}\right)^4. \quad (68)$$

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

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

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

$$A^d - A^o > \frac{1}{\phi} \left(\frac{r}{\phi} (1 + ar\sigma^2) + 2ar\sigma^2 - \frac{r}{\phi} \right) > \frac{ar\sigma^2}{\phi} \left(1 + \frac{r}{\phi} \right). \quad (69)$$

Furthermore, Assumption 2 (i.e., equation (42)) implies that

$$\begin{aligned} \frac{\phi}{r} + a\phi\sigma^2 &> 2a \left(\frac{\phi}{r} + 1 \right)^3 \Leftrightarrow \left(\frac{r}{\phi} \right)^3 \left(\frac{\phi}{r} + a\sigma^2\phi \right) > 2a \left(1 + \frac{r}{\phi} \right)^3 \\ \Leftrightarrow \left(\frac{r}{\phi} \right)^2 (1 + ar\sigma^2) &> 2a \left(1 + \frac{r}{\phi} \right)^3 \\ \Leftrightarrow \frac{ar\sigma^2}{\phi} \left(1 + \frac{r}{\phi} \right) &> \frac{2a^2\sigma^2\phi}{r(1 + ar\sigma^2)} \left(1 + \frac{r}{\phi} \right)^4. \end{aligned} \quad (70)$$

Thus, combining equation (69) and equation (70), we complete the proof for equation (68).

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

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

Combining this with the equality in equation (68), we have

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

Furthermore, notice that

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

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

Lemma 6 *Under condition equation (42) and for $a > 0$, there exists a unique root $x_0 \in (0, \chi^o)$ to the following cubic polynomial:*

$$\begin{aligned}G(x) \equiv & \left(A^d - x \left(\frac{1}{\phi} + A^d\right)\right) (1-x)^2 + \frac{2x}{1 + \phi A^d} \left(\left(A^d - x \left(\frac{1}{\phi} + A^d\right)\right) (1 + ar\sigma^2) - a^2 r^2 \frac{\sigma^2}{\phi^2}\right) \\ & - \left(A^d - x \left(\frac{1}{\phi} + A^d\right)\right) + \frac{a^2 r^2 \sigma^2}{1 + ar\sigma^2} \frac{\sigma^2}{\phi^2}.\end{aligned}\quad (71)$$

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

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

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

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

where the last inequality is due to the following fact

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

Third, 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:

$$G''(x) = 4\left(\frac{1}{\phi} + A^d\right) \left(1 - x - \frac{1 + ar\sigma^2}{1 + \phi A^d}\right) + 2\left(A^d - x \left(\frac{1}{\phi} + A^d\right)\right). \quad (72)$$

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

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

where the last inequality follows equation (68). Thus, both terms in equation (72) are positive, 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

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

Together with equation (68), we obtain

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

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 in equation (61) 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 (analyzed in Section 5.1 Proposition 3)

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

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

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

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

$$\lim_{p \downarrow 0} \frac{V - V^d}{p} = \lim_{p \downarrow 0} \frac{1}{2rp} \left[\frac{(1 + p - \phi V_p)^2}{1 + ar\sigma^2 + a^2 r^2 \sigma^2 \frac{V_p^2}{V_{pp}}} - \frac{(1 + p - \phi V_p^d)^2}{1 + ar\sigma^2} \right] + \left(\frac{\phi}{r} + 1 \right) \lim_{p \downarrow 0} (V_p - V_p^d). \quad (73)$$

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

$$\lim_{p \downarrow 0} \frac{V - V^d}{p} = \lim_{p \downarrow 0} \frac{V - 0}{p} - \lim_{p \downarrow 0} \frac{V^d - 0}{p} = V_p(0) - V_p^d(0) = \lim_{p \downarrow 0} (V_p - V_p^d). \quad (74)$$

Thus, combining equation (73) and equation (74) yields

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

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

$$-\frac{\phi}{r} \lim_{p \downarrow 0} \frac{V_p - V_p^d}{p} = \lim_{p \downarrow 0} \frac{1}{2rp^2} \left[\frac{(1 + p - \phi V_p)^2}{1 + ar\sigma^2 + a^2r^2\sigma^2 \frac{V_p^2}{V_{pp}}} - \frac{(1 + p - \phi V_p^d)^2}{1 + ar\sigma^2} \right]. \quad (75)$$

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

$$\lim_{p \downarrow 0} \frac{V_p - V_p^d}{p} = \lim_{p \downarrow 0} \frac{V_p - (1/\phi)}{p} - \lim_{p \downarrow 0} \frac{V_p^d - (1/\phi)}{p} = V_{pp}(0) - V_{pp}^d(0), \quad (76)$$

and

$$\lim_{p \downarrow 0} \frac{(1 + p - \phi V_p)^2}{p^2} = \lim_{p \downarrow 0} \frac{2(1 + p - \phi V_p)(1 - \phi V_{pp})}{2p} = (1 - \phi V_{pp}(0))^2. \quad (77)$$

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

$$1 + p - \phi V_p^d = 1 + p - \phi(-A^d p + B^d) = (1 + \phi A^d)p. \quad (78)$$

Substituting the above equations (76), (77), and (78) back into equation (75), we have

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

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 (71) 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

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

Therefore,

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

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

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

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:

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

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

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

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

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

or

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

or

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

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

$$T(p) \equiv 1 + p - \phi V_p \tag{79}$$

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

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

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

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

which contradicts 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 also impossible.

We then consider Case 3. Note that Case 3 occurs only in the situation that

$$V_p(p) \rightarrow 0^+ \text{ and } V_{pp}(p) \rightarrow 0^- \text{ when } p \uparrow \hat{p}.$$

Otherwise, a non-zero $V_p(\hat{p})$ implies $\frac{V_p^2(\hat{p})}{V_{pp}(\hat{p})} = -\infty$, 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 $p < \hat{p}$. Therefore, we can define a positive finite number $q \geq 0$ such that there exists some subsequence $\{p_n = \hat{p} - \epsilon_n\}$ approaching $\hat{p}$ from left so that

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

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

Taking the difference on the key ODE in equation (61) at $p = p_n$ and $p = \hat{p}$, for any sequence $\{p_n\} \rightarrow \hat{p}$ we have

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

Simplifying the above equation further, we obtain

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

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

$$arq - 1 = \frac{1}{2} \frac{1 + \hat{p}}{1 + ar\sigma^2 - a^2 r^2 \sigma^2 q} ar \left(\lim_{n \rightarrow \infty} \frac{Q(\hat{p}) - Q(p_n)}{\hat{p} - p_n} \right). \tag{80}$$

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

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

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

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 $\{p_n\} \rightarrow \hat{p}$ we have

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

Thus, $V_{ppp}(\hat{p})$ exists, and hence using $Q(p) = \frac{V_p^2(p_n)}{V_{pp}(p_n)}$ equation (80) implies that

$$1 - arq = \frac{1}{2} \frac{1 + \hat{p}}{1 + ar\sigma^2 - a^2 r^2 \sigma^2 q} \left[ar \frac{V_p^2(\hat{p})}{V_{pp}(\hat{p})} \frac{V_{ppp}(\hat{p})}{V_{pp}(\hat{p})} \right]. \quad (81)$$

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

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

which implies that

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

Hence, we have

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

From equation (81) and $q \neq 0$, we have

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

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 $V_p(\hat{p}) = V_{pp}(\hat{p}) = 0$ and $q = \frac{1}{ar}$ into the key ODE in equation (61), we obtain $V(\hat{p}) = \frac{1}{2r} > V^{HM} = \frac{1}{2r} \frac{1}{1 + ar\sigma^2}$, contradicting condition (41). ■

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

Proof. Suppose that $q = 0$. From equation (80), we have

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

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

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

As a result, $\lim_{n \rightarrow \infty} \frac{Q(\hat{p}) - Q(p_n)}{\hat{p} - p_n} > 0$, 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 $\bar{p} > 0$ such that $V'(\bar{p}) = 0$. The analysis above also implies that $V_{pp}(\bar{p}) < 0$ strictly. Therefore, the function is strictly concave with a positive denominator for $p \in [0, \bar{p}]$. Before we proceed to Step 2.c, we show the following lemma on the property of $\bar{p}$, which is useful in the subsequent proof.

Lemma 10 We have $\bar{p} \leq \bar{p}^d$ where $\bar{p}^d \equiv B^d/A^d = \frac{2\phi}{(2\phi+r)ar\sigma^2+r+\sqrt{(2\phi+r)^2a^2r^2\sigma^4+2ar\sigma^2[(\phi+r)^2+\phi^2]+r^2}}$.

Proof. From the key ODE equation (61), we know that at $\bar{p}$,

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

Similarly, under the deterministic policy,

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

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 $\bar{p} \leq \bar{p}^d$, 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 (61) implies

$$rH(p) - p(\phi+r)H_p(p) = \frac{1}{2} \left[\frac{(1+p-\phi V_p(p))^2}{1+ar\sigma^2+a^2r^2\sigma^2 \frac{V_p^2(p)}{V_{pp}(p)}} - \frac{(1+p-\phi V_p^d(p))^2}{1+ar\sigma^2} \right]. \quad (82)$$

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

$$V_p(\bar{p}^d) > V_p(\bar{p}) = 0 = V_p^d(\bar{p}^d),$$

and hence,

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

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

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

As a result, there exists another point $p_1 \in [\hat{p}, \bar{p}^d]$ such that

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

However, it follows from equation (82) that at $p = p_1$,

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

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

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

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

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

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

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

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

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

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

or

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

or

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

We show that all three cases are impossible below.

Case 1. Suppose this is true. Then it must be true that

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

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

$$1 + \hat{p} - \phi V_p(\hat{p}) = 0.$$

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

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

which contradicts $\hat{p} < 0$.

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

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

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

Otherwise, a non-zero $V_p(\hat{p})$ implies $\frac{V_p^2(\hat{p})}{V_{pp}(\hat{p})} = -\infty$, 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 $\hat{p} < p < 0$. On the other hand, as shown above $V_p(\hat{p}) \geq V_p(0) > 0$, contradicting the fact that $V_p(p) \rightarrow 0^+$ when $p \downarrow \hat{p}$.

Step 2.d: Absorbing lower boundary 0 & entrance-no-exit upper boundary $\bar{p}$:

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

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

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

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

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

Since $V_{pp} < 0$, the above equation is negative if and only if $1 - r\bar{p}\frac{1}{\phi} > 0$. 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 that

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

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

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

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

■

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

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

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

A.7.4 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 when 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

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

Therefore,

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

Due to construction of the 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, since V_p is bounded, and we restrict the policy $\{\sigma_t^P\}$ to be well-behaved (square integrable in the usual sense), $\int_0^t e^{-rt} V_p \sigma_s^P dB_s$ 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

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

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

A.7.5 Step 4: $\bar{p}$ is independent of M

We now 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

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

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 the semi-constrained problem with M_1 is strictly larger than the policy space in the 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 $\bar{p}_1 = \bar{p}_2$.

A.7.6 Step 5: Relaxed problem solves the original problem

We now show that the relaxed problem (58) solves the original problem (57). As explained before, our original problem in equation (57) has more stringent constraints than the relaxed problem (58): 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 $[0, \bar{p}]$, 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

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

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 both problems share the same solution obtained in Proposition 2.

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 $\bar{p}$ are independent of M . We have shown that $\bar{p}$ is independent of M when M is sufficiently high. Moreover, since the endogenous state p never go outside the region $[0, \bar{p}]$, 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, $\max_{0 \leq p \leq \bar{p}} |\beta^M(p)|$ is independent of M , and our result follows.

A.8 Proof for Proposition 3

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:

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

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

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

Rearranging the above equation, we have

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

which gives the solution for A^d :

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

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 equation (36), 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 equation (68) 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.9 Proof for Proposition 4

Suppose along the equilibrium path the agent's continuation payoff is v_t . Similar to equation (16), we want to show that $dv_t = -arv_t\beta_t(dY_t - \mu_t - m_tdt)$ where β_t is the short-term incentive slope offered along the equilibrium path. Because the agent's future rents are always zero (principals have all the bargaining power), it is easy to show that under the optimal saving policy the private saving balance follows

$$S_t = -\frac{1}{ar} \ln(-arv_t),$$

with consumption policy

$$c_t = g(\mu_t) - \frac{1}{a} \ln(-arv_t).$$

Because the principal has all the bargaining power, the fixed wage α_t satisfies

$$\alpha_t = g(\hat{\mu}_t) + \frac{1}{2}ar\sigma^2\beta_t^2. \quad (83)$$

Intuitively, the principal reimburses the agent's effort cost $\frac{\hat{\mu}_t^2}{2}$, and compensates the risk premium $\frac{1}{2}ar\sigma^2\beta_t^2$ borne by the agent; they are just enough to convince the agent to take the offer. Then, in equilibrium, $\mu_t = \hat{\mu}_t$ and the agent's budget constraint reads

$$\begin{aligned} dS_t &= rS_tdt - c_tdt + \alpha_tdt + \beta_t(dY_t - \mu_t - m_tdt) \\ &= \frac{1}{2}ar\sigma^2\beta_t^2dt + \beta_t(dY_t - \mu_t - m_tdt). \end{aligned}$$

Since $v_t = -\frac{1}{ar} \exp(-arS_t)$, using Ito's lemma we have

$$\begin{aligned} dv_t &= \exp(-arS_t) dS_t + \frac{ar}{2} \exp(-arS_t) dS_t^2 \\ &= -arv_t \left[\frac{1}{2}ar\sigma^2\beta_t^2dt + \beta_t(dY_t - \mu_t - m_tdt) \right] - \frac{ar}{2}arv_t\sigma^2\beta_t^2dt \\ &= -arv_t\beta_t(dY_t - \mu_t - m_tdt). \end{aligned}$$

Thus, the agent's continuation value process is identical to equation (16). This also verifies that α_t in equation (83) is the minimum fixed wage needed to attract the agent.

Proposition 1 implies that the agent's incentive compatibility constraint satisfies

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

Importantly, because the principal t takes future β_{t+s} as given, the principal t is taking p_t as given and choosing β_t to maximize

$$\begin{aligned} &\mathbb{E}_t[dY_t]/dt - \mathbb{E}_t[\alpha_tdt - \beta_t(dY_t - \mu_tdt - m_tdt)]/dt \\ &= \mu_t + m_t - \alpha_t = (\beta_t - p_t) + m_t - g(\beta_t - p_t) - \frac{1}{2}ar\sigma^2\beta_t^2. \end{aligned}$$

Hence, the first-order condition for the optimal incentive β_t is $1 - (\beta_t - p_t) - ar\sigma^2\beta_t = 0$, which implies that

$$\beta_t = \frac{1 + p_t}{1 + ar\sigma^2}.$$

We will see this optimality condition does not hold in the long-term contracting case.

Conjecture that $\beta_t = \beta^{ST}$ and $p_t = p^{ST}$ are constants (we will verify this property shortly.) Then, since $p^{ST} = \frac{\phi}{\phi+r}\beta^{ST}$,

$$\beta^{ST} = \frac{\phi + r}{r + ar\sigma^2(\phi + r)}, p^{ST} = \frac{\phi}{r + ar\sigma^2(\phi + r)},$$

and the equilibrium effort is

$$\mu^{ST} = \beta^{ST} - p^{ST} = \frac{r}{r + ar\sigma^2(\phi + r)}.$$

Now let us rule out the case of timing varying β . Recall that the allowable set of β_t is bounded by say $[-M, M]$. Define $\bar{\beta} \equiv \sup \{\beta_t\} \in [-M, M]$. The optimality of short-term incentive implies that there exists t so that $p_t = (1 + ar\sigma^2)\bar{\beta} - 1 - \varepsilon$ for some sufficiently small ε . On the other hand, similar to the argument in Lemma 4, $p_t \leq \frac{\phi}{\phi+r}\bar{\beta}$, which implies

$$(1 + ar\sigma^2)\bar{\beta} - 1 - \varepsilon \leq \frac{\phi}{\phi+r}\bar{\beta} \Rightarrow \left(\frac{r}{\phi+r} + ar\sigma^2\right)\bar{\beta} \leq 1 + \varepsilon. \quad (84)$$

Similarly, define $\underline{\beta} \equiv \inf \{\beta_t\} \in [-M, M]$ we will have

$$(1 + ar\sigma^2)\underline{\beta} - 1 + \varepsilon \geq \frac{\phi}{\phi+r}\underline{\beta} \Rightarrow -\left(\frac{r}{\phi+r} + ar\sigma^2\right)\underline{\beta} \leq -1 + \varepsilon. \quad (85)$$

Summing equation (84) and (85), we have

$$\bar{\beta} - \underline{\beta} \leq \frac{2\varepsilon}{\frac{r}{\phi+r} + ar\sigma^2}.$$

Since ε is arbitrarily small, it must be $\bar{\beta} = \underline{\beta}$, and β_t is constant. QED.

B Appendix B: Non-stationary Learning

For tractability reasons we assume stationary learning in our model as in DeMarzo and Sannikov (2014). A non-stationary learning setting is also widely adopted when economic agents face parameter uncertainty (e.g., Prat and Jovanovic, 2014): the underlying profitability θ (as a parameter) never changes, and as time passes both parties eventually get to learn the true profitability. Does the front-loaded feature of the optimal effort profile depend on whether or not learning is stationary?

When learning is non-stationary, the posterior updating rules are

$$dm_t = \Sigma_t^\theta \frac{dY_t - (\mu_t + m_t)dt}{\sigma^2} \equiv \frac{\Sigma_t^\theta}{\sigma} dB_t^\mu, \text{ and } \Sigma_t^\theta = \frac{\sigma^2 \Sigma_0^\theta}{\sigma^2 + \Sigma_0^\theta t}, \quad (86)$$

The most salient difference lies in the conditional variance Σ_t^θ . 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 and vanishes asymptotically.

If learning is non-stationary, on the equilibrium path eventually everyone learns θ perfectly, which implies that over time there are fewer and fewer noises in output dY_t . This is important, as optimal contracting is essentially an inference problem (Holmstrom, 1979). When time goes by, the signal-to-noise ratio of dY_t goes up, which implies that *information quality* goes up in designing the optimal contract. In other words, this information quality effect potentially pushes the optimal effort policy to be back-loaded (i.e., getting closer to the first-best level) to later periods with less uncertainty. This is an opposite force against the information rent effect which favors front-loaded effort policies.²⁵

Which force is stronger, the information rent effect or the information quality effect? The following proposition shows that, under deterministic policies, even with non-stationary learning the optimal effort policy still decreases with time, suggesting that the information rent effect dominates the information quality effect. We leave future research to investigate the general optimal contract with non-stationary learning.

Proposition 5 *Restrict attention to the deterministic policies. With non-stationary learning so that the posterior variance follows equation (86), the optimal effort policy is decreasing over time.*

²⁵ Although both parties learn profitability θ perfectly asymptotically so that we are in Holmstrom and Milgrom (1987), the optimal contract might not implement the level implied by Holmstrom and Milgrom (1987) in the distant future, simply because it could leave too much information rents to the agent in earlier periods.

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

$$dm_t = \Sigma_t^\theta \frac{dY_t - (\mu_t + m_t) dt}{\sigma^2} = \phi_t \sigma \frac{dY_t - (\mu_t + m_t) dt}{\sigma} \equiv \phi_t \sigma dB_t^\mu, \text{ and } \phi_t = \frac{\phi_0}{1 + \phi_0 t}, \quad (87)$$

where we still use ϕ_t without risk of confusion. According to Prat and Jovanovic (2014), the agent's incentive-compatibility constraint is

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

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

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

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

$$\mu_t = \beta_t - p_t, \quad p'_t = -\phi_t \beta_t + rp_t \Rightarrow \beta_t = -\frac{1}{\phi_t} (p'_t - rp_t).$$

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

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

with the constraint that $p_T = 0$.

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

$$L_p(s, p_s, p'_s) = \frac{dL_{p'}(s, p_s, p'_s)}{ds}. \quad (88)$$

Because

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

conducting algebraic simplifications on the Euler equation (88) yields

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

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

$$\frac{1}{\phi_s} p''_s = -2p'_s + \frac{1}{\phi_s} rp'_s + rp \left(1 + \frac{ar\sigma^2}{1 + ar\sigma^2} \right) \quad (89)$$

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 subsection, we further show that the first-order optimality of the 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

$$L_{pp}y^2 + 2L_{pp'}xy + L_{p'p'}x^2 \leq 0, \quad (90)$$

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

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

Thus, it is sufficient to show the following:

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

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

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

Thus, equation (90) holds, and the Euler equation is both necessary and sufficient for optimality (Theorem 2.3 in Chapter one of Fleming and Rishel (1975)).

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

Step 1. From the Euler equation we show that p never changes signs. If $p(0) = p_0 > 0$, $p(T) = 0$ but p turns negative somewhere in between, then there must exist some point $\hat{t}$ so that

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

This contradicts the Euler equation (89):

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

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

Step 2. We now prove that if $p_0 > 0$ then the optimal effort policy goes down with time. From equation (89), and $\mu_t = \beta_t - p_t = -\frac{1}{\phi_t} (p'_t - rp_t) - p_t$, 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 equation (87). Similarly, if $p_0 < 0$, the optimal effort policy goes up with time.

Step 3. We show that 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

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

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 $\{\mu_t\}$ increases over time. However,

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

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 always negative. Thus, if $p_0 < 0$, then the objective $\int_0^T e^{-rt} (\mu_t - \frac{1}{2}\mu_t^2 - \frac{1}{2}ar\sigma^2\beta_t^2) dt$ 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

C.1 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 a risk-neutral agent, i.e., $a = 0$. In this case, the optimal deterministic policy obtained in Section 5.1 achieves first-best (and hence is optimal among all possible contracts).

Lemma 13 *When $a = 0$, the deterministic policy in Proposition 3 is optimal with the first-best effort level $\mu^o = 1$. The corresponding value function, denoted by $V^o(p)$, is given by*

$$V^o(p) = -\frac{1}{2}A^o p^2 + B^o p, \quad (91)$$

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

$$\bar{p}_t^o = \bar{p}^o = \frac{B^o}{A^o} = \frac{\phi}{r}. \quad (92)$$

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:

$$V(p; a) = Q_0(p) + \sum_{i=1}^I a_r^i Q_i(p) + o(a_r^I), \quad (93)$$

$$\bar{p} = q_0^* - \sum_{j=1}^J a_r^j q_j^* + o(a_r^J), \quad (94)$$

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 equation (91), and $q_0^* = \bar{p}^o$ as given in equation (92). 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 equation (38).

The optimal deterministic contract studied in Proposition 3 is helpful in deriving Q_i 's in equation (93) and q_i^* 's in equation (94). Let $\bar{p}^d \equiv \frac{B^d}{A^d} > 0$ which satisfies $V_p^d(\bar{p}^d) = 0$. By expanding both $V^d(p; a)$ and $\bar{p}^d = B^d/A^d$ in the forms of equation (93), 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 $\bar{p}$ in the stochastic optimal contract starts to differ from the deterministic counterpart $\bar{p}^d$ in the third order a_r^3 . The next proposition summarizes our results.

Proposition 6 *When a is relatively small, we have the following approximations:*

$$\bar{p} = \bar{p}^d - a_r^3 \frac{\phi^5 \sigma^4}{r^6} \left(1 + \frac{r}{\phi}\right)^3 + o(a_r^3), \quad (95)$$

$$V(\bar{p}) = V^d(\bar{p}^d) + a_r^4 \left[\frac{\phi^6 \sigma^6}{r^8} \left(1 + \frac{r}{\phi}\right)^5 + \frac{\phi^5 \sigma^6}{r^7} \left(1 + \frac{r}{\phi}\right)^3 \right] + o(a_r^4). \quad (96)$$

Furthermore, we show that

$$\beta(p) = \beta^d(p) - a_r^2 \frac{\sigma^2}{\phi} \left(1 + \frac{r}{\phi}\right) p(p - \bar{p}^o) \left[(p - \bar{p}^o) + \left(1 + \frac{r}{\phi}\right) p \right] + o(a_r^2), \quad (97)$$

$$\sigma^p(p) = a_r \sigma \left(1 + \frac{r}{\phi}\right) \left(1 + a_r \frac{\phi \sigma^2}{r}\right) p(\bar{p} - p) + o(a_r^2). \quad (98)$$

Proof. Available upon request. ■