Theoretical Economics 17 (2022), 1589–1618 1555-7561/20221589

Dynamic delegation with a persistent state

Yi Chen
Johnson College of Busines, Cornell University

In this paper, I study the dynamic delegation problem in a principal–agent model wherein an agent privately observes a persistently evolving state, and the principal commits to actions based on the agent's reported state. There are no transfers. While the agent has state-independent preferences, the principal wants to match a state-dependent target. I solve the optimal delegation in closed form, which sometimes prescribes actions that move in the opposite direction of the target. I provide a simple necessary and sufficient condition for that to occur. Generically, the principal fares strictly better in the optimal delegation than in the babbling outcome. Over time, the principal is worse off in expectation, but the agent is better or worse off depending on the shape of the principal's state-dependent target.

Keywords. Communication, dynamic delegation, contrarian, quota mechanism, Brownian motion.

JEL classification. D82, D83, D86.

1. Introduction

In many organizations, decisions are not made by the person who actually holds and understands the most relevant information. Moreover, the transmission of this information is often hindered by conflicts of interest. The informed party may have a self-interested motive to mislead the uninformed party, who then takes action based on information that may have been miscommunicated.

A firm's headquarters, for example, will allocate resources to a division manager over time. The headquarters metes out resources in order to hit a target amount of allocation that depends on some state of the project, say, profitability, consumer taste, or technical

Yi Chen: yi.chen@cornell.edu
I am indebted to Dirk Bergemann, Johannes Hörner, and Larry Samuelson for their constant academic support. I am grateful to Marco Battaglini, Yeon-Koo Che, Chen Cheng, Eduardo Faingold, Simone Galperti, Yuan Gao, Marina Halac, Ryota Iijima, Justin Johnson, Navin Kartik, Nicolas Lambert, Fei Li, Elliot Lipnowski, Chiara Margaria, Dmitry Orlov, Gregory Pavlov, Ennio Stacchetti, Philipp Strack, Juuso Välimäki, Zhe Wang, Yiqing Xing, Xiye Yang, and John Zhu for extended discussions. I also thank the conference and seminar participants at Penn State University, Cornell University, Johns Hopkins University, University of Southern California, University of Arizona, Western University, Peking University HSBC Business School, Southern University of Science and Technology, the 2018 International Conference on Game Theory at Stony Brook, the 2018 Southern Economic Association Annual Conference, the 2019 North American Summer Meeting of the Econometric Society, the 2019 Osaka Workshop on Economics of Institutions and Organizations, and the 2020 North American Winter Meeting of the Econometric Society for helpful comments and suggestions. All errors are mine.

 https://doi.org/10.3982/TE4710

parameters, which develops slowly over time. Only the division manager directly observes this state, but he wants to receive more resources *regardless* of the state. Worse still, the headquarters receives profits only after a long lag, making it difficult to detect misrepresentations in the manager's reports. These severe conflicts of interest beg the question, "Does the headquarters benefit at all from the manager's information with a dynamic contract?" If so, how should it optimally act on the manager's reports to utilize information?

To answer these questions, I use a dynamic principal–agent model to investigate how the agent's persistent but changing private information can be best elicited and put to use. The agent privately observes a state, which evolves as a Brownian motion with a drift. The agent continuously reports the state to the principal, but has the ability to inflate or shade the report at any time. The principal observes nothing but the agent's reports and so commits to a dynamic contract specifying actions over time based on the reported history. The principal cares about the state because she incurs a flow cost that is quadratic in the gap between the action and her *target*, i.e., her state-dependent favorite action. The agent, by contrast, has a transparent motive independent of the state: the higher the action, the better.

Communication is ineffective in a *static* contracting environment where the agent has severely misaligned, state-independent preferences. Indeed, to neutralize the misreporting motive of the agent, the contract would have to assign the same expected action for all reported states. When it comes to *long-term* relationships with an evolving state, then the prospect of communication is better. In order to elicit truthful reports, the principal only needs to ensure that the agent's continuation payoff is independent of his current report. In other words, the principal commits to a fixed *quota* (Jackson and Sonnenschein (2007)), which is the discounted sum of actions. This leaves the principal with the optimization problem of intertemporally reallocating the quota to make the best use of the agent's information.

Using a recursive method, I solve the optimal contract in closed form. While tractability is usually difficult to obtain in a dynamic setting with persistent information, I address this obstacle by reducing the dimensionality of the recursive problem and transforming the nonlinear partial differential equation into two ordinary differential equations. The closed-form solution allows for comprehensive analysis of the optimal contract.

First, the optimal contract prescribes how the action responds to information at any time. If a contract stipulates an action moving in the same direction as the target, it is said to exhibit a *conformist* pattern; if instead the action moves in the opposite direction from the target, it exhibits a *contrarian* pattern. I find a simple necessary and sufficient condition for either pattern to occur, which involves the third derivative of the target function. To understand the intuition, suppose the state process has no drift and the target function is increasing. If a positive shock to the state boosts the current target, the principal will be inclined to raise the *current* level of action. Meanwhile, due to the persistence of the state, future targets are also expected to increase, tempting the principal to take higher actions in the *future*. Unfortunately, the agent's incentive constraints cannot allow both. Any increase in the current action necessitates lower future actions or vice versa. Whether a conformist or contrarian pattern emerges hinges on the trade-off between hitting the current target on the one hand and hitting future targets on the other. If the former goal dominates, a conformist pattern emerges; otherwise, we should observe a contrarian pattern. When the target function has a positive third derivative, the expected future target is more sensitive to state shocks than the current target is. Then, in the intertemporal trade-off, the principal optimally sacrifices the goal of hitting the current target in exchange for a better chance of hitting future targets, leading to a contrarian pattern.

Second, I show that communication generically improves upon babbling in the optimal dynamic contract. Communication is effective as long as the action is responsive to the reported state, whether in a conformist or a contrarian pattern. When the drift of the state is zero, the knife-edge exceptions occur when the target function is linear or quadratic. In these cases, a shock to the current state justifies an increase of current action as much as it demands a rise in the expected future actions; therefore, the principal's optimal choice is to stay put despite her degree of freedom in responding to information. The result implies that the curvature of the target function is not sufficient to guarantee effective communication; instead, a nonzero third derivative is required.

Third, the model delivers predictions on the welfare of the two parties as the state unfolds. The principal expects an ever-increasing cost over time, indicating an inevitable worsening of the match between the target and the action. This cost–backloading result holds even for patient players: the incentive constraints cause distortions to accumulate without bound, which weighs heavily on the principal precisely because of her patience. The agent, on the other hand, is not necessarily immiserated. The trend of his continuation payoff depends on the shape of the target function as well as on the state process.

I also explore three extensions of the main model. The first enriches the state process by allowing for mean reversion and, accordingly, weaker persistence of the state. Mean reversion undermines the responsiveness of the expected future target to the current state, because any shock to the current state decays over time. Consequently, a contrarian pattern is less likely to emerge as the principal finds it less appealing to sacrifice hitting the current target for the sake of hitting future ones. The second extension examines the optimal contract for a finite horizon. Even adding calendar time as an additional state variable, the problem is still solvable. While the basic insights remain, the finite horizon introduces a deadline effect. In the beginning, when the deadline is far in the future, the contract behaves similarly to the main model. As time goes on, the action becomes less and less responsive to information, and the agent's influence gradually vanishes. The third and final extension explores the implication of having a less patient agent relative to the principal. The difference in discount rates generally results in payoff front-loading for the agent, which is more likely to cause immiseration. Moreover, hitting future targets becomes less costly for the principal; thus, the intertemporal trade-off is more inclined to favor future targets and the contract appears more contrarian.

1.1 *A two-period example*

To illustrate the key trade-offs in a dynamic contracting problem, I present a two-period example. A state $\theta_t \in \mathbb{R}$ follows a random walk. In period $t = 1$, θ_1 is drawn from $\mathcal{N}(0, 1)$. In period $t = 2$, θ_2 evolves from θ_1 with noise,

$$\theta_2 = \theta_1 + \varepsilon,$$

where $\varepsilon \sim \mathcal{N}(0, 1)$ is independent of θ_1. In each period t, the agent privately learns θ_t and reports $\hat{\theta}_t$ to the principal, who takes action $x_t \in \mathbb{R}$. Monetary transfers are not available. The principal's total cost is $(x_1 - f(\theta_1))^2 + (x_2 - f(\theta_2))^2$, where $f(\cdot)$ is a time-invariant *target function*. The agent's total payoff is $x_1 + x_2$. A contract is a pair $(x_1(\hat{\theta}_1), x_2(\hat{\theta}_1, \hat{\theta}_2))$, mapping report histories into actions. Based on the revelation principle, I focus on truthful contracts. The principal solves

$$\min_{x_1(\cdot), x_2(\cdot,\cdot)} \quad \mathbb{E}\big[\big(x_1 - f(\theta_1)\big)^2 + \big(x_2 - f(\theta_2)\big)^2\big]$$

$$\text{subject to} \quad x_1(\theta_1) + \mathbb{E}\big[x_2(\theta_1, \theta_2)|\theta_1\big] \geq x_1(\hat{\theta}_1) + x_2(\hat{\theta}_1, \hat{\theta}_2) \quad \forall \theta_1, \hat{\theta}_1, \hat{\theta}_2 \tag{1}$$

$$x_2(\theta_1, \theta_2) \geq x_2(\theta_1, \hat{\theta}_2) \quad \forall \theta_1, \theta_2, \hat{\theta}_2. \tag{2}$$

Condition (2) requires that truth-telling in period 2 is optimal for the agent after a truthful report in period 1. Condition (1) governs the truth-telling incentive in period 1 for the agent, who weighs all possible reporting strategies.

Since the agent's payoff is state-independent, condition (2) implies that $x_2(\theta_1, \theta_2)$ does not depend on θ_2. Writing $x_2(\theta_1, \theta_2) = x_2(\theta_1)$ for short, condition (1) is simplified to $x_1(\theta_1) + x_2(\theta_1) \geq x_1(\hat{\theta}_1) + x_2(\hat{\theta}_1)$ for all θ_1 and $\hat{\theta}_1$. To satisfy this condition, we must have

$$x_1(\theta_1) + x_2(\theta_1) \equiv W, \tag{3}$$

where W is a constant, interpreted as the *quota* (total payoff) promised to the agent. This quota, as well as how x_1 and x_2 jointly respond to θ_1, are optimally chosen by the principal to minimize cost.

With the simplified incentive constraints, the optimal two-period contract is obtainable for general $f(\cdot)$ (see Appendix A.1). Specifically, for a linear target $f(\theta) = \theta$, we have $x_1(\theta_1) = x_2(\theta_1) = 0$, i.e., the outcome is "babbling" as the actions do not reflect information about the state. A quadratic target $f(\theta) = \theta^2$ does not lead to better utilization of information, as $x_1(\theta_1) = 1$ and $x_2(\theta_1) = 2$ for all θ_1. An interesting case arises when $f(\theta) = e^\theta$, for which $x_1(\theta_1) = \frac{1}{2}(e + \sqrt{e}) - \frac{1}{2}(\sqrt{e} - 1)e^{\theta_1}$ and $x_2(\theta_1) = \frac{1}{2}(e + \sqrt{e}) + \frac{1}{2}(\sqrt{e} - 1)e^{\theta_1}$. As the first-period target e^{θ_1} increases, the corresponding action x_1 decreases, in order for x_2 to increase in the next period.

Why does this pattern emerge? The answer lies in the shape of the target function. At an arbitrary state θ_1, suppose the optimal contract specifies actions x_1 and x_2 over the two periods. Given the principal's quadratic cost function, a marginal increase in

x_1 brings a marginal benefit of $2(f(\theta_1) - x_1)$ to the principal. Meanwhile, the quota mechanism forces x_2 to decrease, which imposes a marginal cost of $2(\mathbb{E}[f(\theta_2)|\theta_1] - x_2)$. Optimality requires them to cancel out. Now, a higher θ_1 raises the marginal benefit by $2f'(\theta_1)$, but also raises the marginal cost by $2\mathbb{E}[f'(\theta_2)|\theta_1]$.[1] When the slope of the target function is convex (e.g., $f(\theta) = e^{\theta}$), the latter effect dominates because $\mathbb{E}[f'(\theta_2)|\theta_1] > f'(\mathbb{E}[\theta_2|\theta_1]) = f'(\theta_1)$. To restore optimality, the principal should shift some quota from x_1 to x_2 despite the increased current target $f(\theta_1)$.

In Section 2, I lay out the formal model in continuous time with an infinite horizon. Continuous time allows for the gradual arrival of information and thus closed-form analysis. An infinite horizon avoids the deadline effect, keeps the stationarity of the problem, and enables the study of the asymptotics.

1.2 *Related literature*

This paper contributes to a closely related literature on allocation problems without monetary transfer. In a static setting, Jackson and Sonnenschein (2007) study the decision rule facing many replicas of the same allocation problem and propose a "quota mechanism" that links all allocations together, where efficiency is asymptotically achieved as the number of replicas increases. In a dynamic setting, repeated allocation games (e.g., Renault, Solan, and Vieille (2013), Margaria and Smolin (2018), Lipnowski and Ramos (2020)) feature an informed sender and an uninformed receiver, where the sender observes independent and identically distributed (i.i.d.) or persistent information and reports to the receiver for decision-making. When the receiver is assumed to have intertemporal commitment power, Frankel (2016) and Guo and Hörner (2020), among others, analyze dynamic allocation problems. In these models, dynamic versions of the quota mechanism arise in equilibrium strategy or optimal contract, where the quota is cashed out in a conformist pattern. My model also exhibits a dynamic quota for the agent, but the combination of a persistent state process and a general target function allows for intertemporal trade-offs that lead to potentially contrarian patterns. Such a counterintuitive pattern cannot arise in a static, multidimensional setting or in a dynamic model with finite state Markov chain and linear preferences.

This paper also builds on the literature on communication. Since Crawford and Sobel (1982) and Green and Stokey (2007) pioneered the field of sender–receiver games, a large body of scholarship has been produced (see Sobel (2013) for a comprehensive summary). Meanwhile, communication with a committed receiver has inspired the literature on delegation (see Holmstrom (1977), Melumad and Shibano (1991), Alonso and Matouschek (2008), Amador and Bagwell (2013)). This paper features a committed receiver and a privately informed sender, with evolving information and without transfers. It therefore expands on the models of dynamic delegation.

More broadly, other related works tackle various allocation problems with similar, but different, settings.[2] Bird and Frug (2019) study a dynamic contracting problem without transfer and implement the unique optimal contract by a deadline to earn rewards.

[1] It holds that $\mathrm{d}\mathbb{E}[f(\theta_2)|\theta_1]/\mathrm{d}\theta_1 = \mathbb{E}[f'(\theta_2)|\theta_1]$ due to the random walk assumption.

[2] Models of multiple competing agents (e.g., Ben-Porath, Dekel, and Lipman (2014) and de Clippel, Eliaz, Fershtman, and Rozen (2021)) find "strategic favoritism" as the optimal mechanism. There is also a larger

Boleslavsky and Lewis (2016) and Malenko (2019) study dynamic mechanisms of influence with either costly verification or noisy observation of the state.

The remainder of the paper is organized as follows. Section 2 lays out the setting for the continuous-time model. Section 3 reduces the agent's incentive constraint to a necessary condition and a stronger, sufficient condition. Section 4 solves and analyzes the optimal contract. Section 5 discusses three extensions of the main model, and Section 6 concludes.

2. The model

There is a principal (she) and an agent (he). Time $t \geq 0$ is continuous. A state θ evolves over time, but is observable only to the agent. The agent continuously makes potentially manipulated report $\hat{\theta}$ of the true state to the principal, who commits to action $x \in \mathbb{R}$ at all times based on the history of reports.

The state process θ starts at zero and evolves according to

$$\theta_t = \mu t + Z_t,$$

where Z is the standard Brownian motion on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The constant μ is the drift of the process. The volatility is constant and normalized to 1. The law of motion of θ is common knowledge. The state process is highly persistent as it features independent increments. In Section 5.1, I introduce mean reversion into the process as a less persistent counterpart.

Interests are misaligned. While the principal's favorite action is state-dependent, the agent only wishes to induce actions as high as possible. Specifically, given a state–action pair (θ, x), the principal suffers a quadratic flow cost $(x - f(\theta))^2$ from the gap between the action x and a state-dependent *target* $f(\theta)$. For ease of analysis, the function f is assumed to be piecewise $\mathcal{C}^2$. The agent's flow payoff is simply x, independent of the state. Intertemporally, the players share the same discount rate $r > 0$.

A strategy m of the agent is a θ-measurable process, such that his reported process $\hat{\theta}$ follows

$$\mathrm{d}\hat{\theta}_t = m_t \,\mathrm{d}t + \mathrm{d}\theta_t,$$

where $m_t \in \mathbb{R}$ represents the "intensity of misreporting" at instant t. The space of feasible strategies is set to

$$\mathcal{M} \equiv \left\{ m : \mathbb{E}\big[e^{2\overline{\alpha} \int_0^t m_s \,\mathrm{d}s}\big] < \infty \; \forall t, \text{ and } \lim_{t \to \infty} e^{-rt} \mathbb{E}\big[e^{2\overline{\alpha} \int_0^t m_s \,\mathrm{d}s}\big] = 0 \right\}$$

to exclude Ponzi-type strategies, where

$$\overline{\alpha} \equiv \frac{1}{2}\big(\sqrt{2r + \mu^2} - |\mu|\big) \tag{4}$$

is a positive constant. I show in Appendix A.11 that this restriction is not essential.

literature on dynamic contract with transfers (i.e., Fernandes and Phelan (2000), Battaglini (2005), Sannikov (2008), Williams (2011), Kapička (2013), DeMarzo and Sannikov (2016)), but the lack of transfers in my setting aggravates the agency problem because the continuation payoff of the agent can only be promised by a sequence of future allocations.

In the beginning, the principal commits to a contract x. It is a process adapted to the information generated by $\hat{\theta}$, specifying, at any time t, an action $x_t \in \mathbb{R}$ as a function of the history of reports $\hat{\theta}^t$ up to time t. There are no monetary transfers.

Given a contract–strategy pair (x, m), the total expected cost and payoff are, respectively,

$$U_P(x, m) = \mathbb{E}^m\left[\int_0^\infty re^{-rt}\big(x_t - f(\theta_t)\big)^2 \,\mathrm{d}t\right]$$
$$U_A(x, m) = \mathbb{E}^m\left[\int_0^\infty re^{-rt} x_t \,\mathrm{d}t\right],$$

where $\mathbb{E}^m$ denotes the expectation induced by strategy m. Hereafter, "payoff" and "cost" refer to the agent's total expected payoff and the principal's total expected cost, unless otherwise noted. The following regularity condition ensures finiteness of the above cost and payoff.

Assumption 1 (Regularity). *There exists* $\alpha_0 > 0$ *and* $\alpha_1 \in [0, \overline{\alpha})$ *such that* $|f(\theta)| \leq \alpha_0 e^{\alpha_1 |\theta|}$.

Intuitively, the condition prevents the target function from growing too exponentially in both directions. The constant $\overline{\alpha}$ is defined in (4). This condition is not too restrictive, as it allows for all piecewise polynomials and piecewise continuous bounded functions, among others.

The agent chooses a strategy m to maximize his payoff from a given contract x. The principal designs a contract x to minimize her cost given the agent's strategy in reaction to the contract. With the usual revelation principle argument (see Appendix A.2 for a formal proof), it is without loss of generality to focus on *truthful* contracts, i.e., those that make truth-telling ($m \equiv 0$) optimal for the agent among all strategies. Moreover, I focus on deterministic mechanisms, but, as is verified later, such a restriction is without loss of generality. Therefore, the principal solves

$$\min_{(x_t(\cdot))_{t\geq 0}} \quad \mathbb{E}\left[\int_0^\infty re^{-rt}\big(x_t(\theta^t) - f(\theta_t)\big)^2 \,\mathrm{d}t\right] \tag{5}$$

$$\text{subject to} \quad \mathbb{E}\left[\int_0^\infty re^{-rt} x_t(\theta^t) \,\mathrm{d}t\right] \geq \mathbb{E}\left[\int_0^\infty re^{-rt} x_t(\hat{\theta}^t) \,\mathrm{d}t\right],$$
$$\text{where } \hat{\theta}_t \equiv \theta_t + \int_0^t m_s \,\mathrm{d}s \quad \forall m \in \mathcal{M}. \tag{6}$$

The incentive constraint (6) stipulates that truth-telling leads to the highest payoff. While the constraint is expressed as of time zero, it also implies incentive compatibility at all later times, since the agent faces a decision problem with time-consistent preferences. This constraint implicitly assumes that the payoff of the agent is well defined on and off equilibrium, but such an assumption is innocuous because if a contract generates non-integrable payoffs for the agent, it must bring infinite cost to the principal, which is clearly suboptimal.

I will end this section with a few comments on the model. First, the quadratic-cost assumption is not essential for the qualitative results, but it greatly simplifies analysis because the cost-minimizing action facing an uncertain target is the mean of the target. The separability result (Lemma 2) directly benefits from this assumption. Second, one can alternatively define $f_t \equiv f(\theta_t)$ as the state. Here I choose to keep the process simple, while summarizing everything in f. Third, the insatiable preferences of the agent can be interpreted in Crawford and Sobel (1982) as taking the bias to infinity and, therefore, represent severely misaligned interests.

3. Incentives of the agent

This section reduces the incentive compatibility of the agent into a tractable form, in preparation for deriving the optimal contract. Specifically, a necessary condition for incentive compatibility is obtained in Section 3.1 from the first-order approach. It is then augmented to a sufficient condition in Section 3.2, which is later invoked to verify the optimality of the candidate solution in Section 4.

3.1 *Incentive compatibility: Necessary condition*

The dynamic first-order approach (Williams (2011), Kapička (2013), Pavan, Segal, and Toikka (2014), DeMarzo and Sannikov (2016)) derives a local version of the incentive constraints. To apply this method, I define a process $W = (W_t)_{t\geq 0}$ for any contract x,

$$W_t \equiv \mathbb{E}_t\left[\int_t^\infty re^{-r(s-t)} x_s \,\mathrm{d}s\right],$$

as the agent's on-path expected continuation payoff. The expectation $\mathbb{E}_t$ is conditional on the information generated by the state process up to time t.

To use the recursive method, I need a few state variables to summarize the history. The current state θ_t naturally serves as a state variable. The continuation payoff W_t is commonly used as another state variable in the literature (see Abreu, Pearce, and Stacchetti (1986) and Thomas and Worrall (1990), among others). Furthermore, due to the persistent private information, a third state variable, called the continuation *marginal* payoff, is usually required (Fernandes and Phelan (2000), Williams (2011), Kapička (2013), Guo and Hörner (2020)). However, in this paper I can drop the third state variable despite the persistence of information. This is because the agent's payoff is independent of the state, and, hence, his flow and continuation payoffs, both on and off path, are common knowledge. Even if the agent used strategy $m \neq 0$ and his private belief about the state diverged from the principal's, the continuation payoff would evolve as if $\hat{\theta}_t = \theta_t + \int_0^t m_s \,\mathrm{d}s$ was the realization of the true state and the agent had reported truthfully.

Given any contract x, the implied evolution of W can be written as a diffusion process according to Lemma 1 below.

LEMMA 1 (Martingale Representation Theorem). *For any contract x, there exists a process $\beta = (\beta_t)_{t\geq 0}$ adapted with respect to the information generated by θ^t such that*

$$dW_t = r(W_t - x_t)\,dt + r\beta_t \underbrace{(d\theta_t - \mu\,dt)}_{=dZ_t}. \tag{7}$$

The first term on the right-hand side of (7) represents the drift of W_t: it is replenished at rate r and at the same time drained as action x_t is taken to fulfill the promise. The second term, which is the diffusion, governs the incentives. On the equilibrium path, it holds that $d\hat{\theta}_t - \mu\,dt = d\theta_t - \mu\,dt = dZ_t$, which has zero mean. The multiplier $r\beta_t$ is interpreted as the instantaneous slope of the continuation payoff with respect to reported states or "strength of incentives." Suppose the instantaneous slope $r\beta_t$ is positive. By adding a drift $m > 0$ to the report for a short moment dt, the agent tricks the principal into believing that the state is $m\,dt$ higher than its true value, which boosts the agent's continuation payoff by $r\beta_t m\,dt$.[3] Similarly, if $r\beta_t < 0$, the agent can profit by shading the report. To deter a local deviation from truth-telling, it is necessary to keep $r\beta_t$ identically at zero. In other words, β_t always being zero is the only way to keep the discounted quota fixed. Proposition 1 formalizes the above reasoning.

PROPOSITION 1 (Necessity). *The incentive constraint (6) implies $\beta_t = 0$ for all $t \geq 0$ a.s. $\mathbb{P}$.*

According to the proposition, truth-telling necessitates the entire disentanglement of the agent's continuation payoff from his current report, a property already found in the two-period example. As a direct implication of the proposition, the agent faces a deterministic discounted quota.[4] To see this, let $\beta_t = 0$ in (7). Then the evolution of W_t is completely pinned down by the path of x_t, and $W_0 = \int_0^\infty re^{-rt}x_t\,dt$ as long as $e^{-rt}W_t \to 0$. In other words, the agent is guaranteed a fixed discounted quota W_0 up front in any incentive compatible contract. With Proposition 1 simplifying the incentive constraints, the principal's remaining task is to optimize her response to information, as is analyzed in Section 4.

3.2 *Incentive compatibility: Sufficient condition*

The necessary condition in Proposition 1 is almost sufficient, except that the contract also has to prevent Ponzi-style global deviations with an infinite horizon. The following proposition claims that if a transversality condition holds for all strategies of the agent, then no global deviation is profitable and the necessary condition $\beta_t = 0$ becomes sufficient.

PROPOSITION 2 (Sufficiency). *Let W_t^m denote the process of continuation payoff when the agent employs strategy m. If $\beta_t = 0$ for all $t \geq 0$ a.s. $\mathbb{P}$ and $\lim_{t\to\infty} e^{-rt}\mathbb{E}W_t^m = 0$ for all $m \in \mathcal{M}$, then the incentive constraint (6) is satisfied.*

[3]The flow payoff also changes, but in continuous time, the change is of order $o(dt)$.
[4]This property is typical in the literature on dynamic quota mechanisms, e.g., Frankel (2016).

4. Optimal contract

This section presents the optimal contract and its properties. I first derive a candidate solution using the necessary condition from Proposition 1 in lieu of the original incentive constraints, and then verify it by checking the sufficient condition from Proposition 2.

4.1 *Solution to the recursive problem*

I use dynamic programming to solve a relaxed recursive problem where the incentive constraints are replaced by the weaker version in Proposition 1. As is conjectured in Section 3.1, the optimal contract is governed by two state variables: the state θ and the continuation payoff W. In Theorem 1, I formally verify that the candidate contract is indeed the solution to the original problem (5)–(6).

I define $C(\theta, W)$ as the cost function of the principal. From (7) and Proposition 1, the W process evolves according to $\mathrm{d}W_t = r(W_t - x_t)\,\mathrm{d}t$, without volatility. Hence, the cost function must satisfy the Hamilton–Jacobi–Bellman (HJB) equation

$$rC(\theta, W) = \min_x r\big(x - f(\theta)\big)^2 + r(W - x)C_W(\theta, W) + \mu C_\theta(\theta, W) + \frac{1}{2}C_{\theta\theta}(\theta, W). \qquad (8)$$

The right-hand side consists of four terms. The first term is the normalized flow cost. The second and third terms are expected changes of cost caused by the drift in W and θ, respectively. The last term is the Itô term, generated by the volatility of θ. In order for the recursive form to correctly represent the original problem, the cost and the payoff must also satisfy the transversality condition

$$\lim_{t\to\infty} e^{-rt}C(\theta_t, W_t) = 0, \qquad \lim_{t\to\infty} e^{-rt}W_t = 0 \quad \text{a.s. } \mathbb{P}. \qquad (9)$$

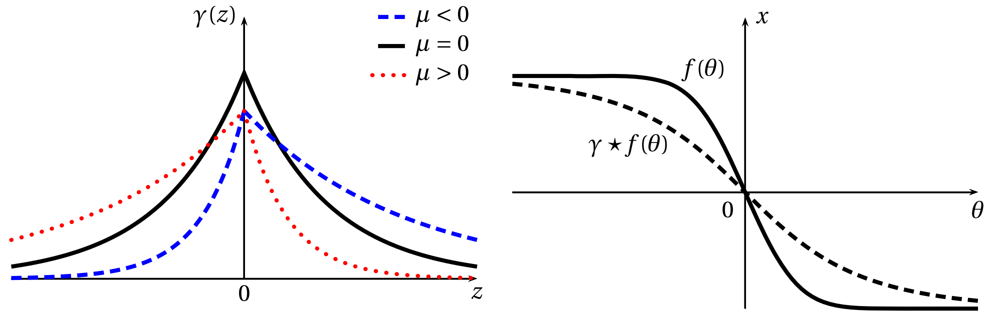
Conditions (8) and (9) define a problem that is intractable in general, but the quadratic cost structure admits the following separability result that simplifies the problem.

Lemma 2 (Separability). *If contract x minimizes $\mathbb{E}[\int_0^\infty re^{-rt}(x_t - f(\theta_t))^2\,\mathrm{d}t]$ subject to $\int_0^\infty re^{-rt}x_t = W_0$ a.s. $\mathbb{P}$, then for any constant $a \in \mathbb{R}$, contract $x + a$ minimizes $\mathbb{E}[\int_0^\infty re^{-rt}(x_t - f(\theta_t))^2\,\mathrm{d}t]$ subject to $\int_0^\infty re^{-rt}x_t = W_0 + a$ a.s. $\mathbb{P}$.*

The lemma suggests that the policy function satisfies $\partial x(\theta, W)/\partial W = 1$. The first-order condition of (8) also requires $x = f(\theta) + \frac{1}{2}C_W(\theta, W)$. Under these two conditions, the cost function must be quadratic in W, i.e., $C(\theta, W) = (W - g(\theta))^2 + h(\theta)$, where g and h are functions to be determined. Then (8) and (9) are transformed into a pair of second-order differential equations with constant coefficients:

$$g''(\theta) + 2\mu g'(\theta) - 2rg(\theta) = -2rf(\theta), \qquad \lim_{\theta\to\pm\infty} e^{\mu\theta - \sqrt{\mu^2+2r}|\theta|}g(\theta) = 0 \qquad (10)$$

$$h''(\theta) + 2\mu h'(\theta) - 2rh(\theta) = -2g'(\theta)^2, \qquad \lim_{\theta\to\pm\infty} e^{\mu\theta - \sqrt{\mu^2+2r}|\theta|}h(\theta) = 0. \qquad (11)$$



(a) Asymmetric Laplace Distribution (b) Convolution $\gamma \star f$, when $\mu = 0$

FIGURE 1. The asymmetric Laplace distribution and its convolution with a target function.

Plugging in the solutions $g(\theta) = \gamma \star f(\theta)$ and $h(\theta) = \frac{1}{r}\gamma \star (\gamma \star f)'^2(\theta)$, I arrive at the unique candidate solution to (8) and (9),

$$C(\theta, W) \equiv \big(W - \gamma \star f(\theta)\big)^2 + \frac{1}{r}\gamma \star (\gamma \star f)'^2(\theta) \tag{12}$$

$$x(\theta, W) \equiv f(\theta) - \gamma \star f(\theta) + W, \tag{13}$$

where γ is an asymmetric Laplace distribution

$$\gamma(z) \equiv \frac{r}{\sqrt{\mu^2 + 2r}} e^{\mu z - \sqrt{\mu^2+2r}|z|}$$

and $\gamma \star f$ is the convolution between the two functions.

The function γ decays exponentially on both sides of zero, but at potentially different rates. Figure 1(a) shows the graph of γ for different values of the drift μ. It is symmetric when $\mu = 0$, but is otherwise skewed by the drift.

The convolution $\gamma \star f(\theta)$ is interpreted as the *expected discounted future target*, conditional on the current state θ. Specifically, Fubini's theorem shows

$$\gamma \star f(\theta) = \mathbb{E}\left[\int_0^\infty re^{-rt} f(\theta_t)\,\mathrm{d}t \,\middle|\, \theta_0 = \theta\right], \qquad \mathrm{d}\theta_t = \mu\,\mathrm{d}t + \mathrm{d}Z_t.$$

Therefore, the convolution $\gamma \star f$ serves to smooth the target function f with kernel γ, giving more probability weight to states in the neighborhood of the current state θ and less weight to distant states. Nonnegativity of a function is preserved under this convolution. Moreover, since $(\gamma \star f)' = \gamma \star f'$ for differentiable f, the nonnegativity of the n-order derivative of a function is also preserved for any $n \geq 1$. Figure 1(b) illustrates a typical target function f and its convoluted version $\gamma \star f$, when $\mu = 0$.

Formally, Theorem 1 verifies that the candidate solution indeed solves the original problem (5)–(6). Since the auxiliary state variable W appears in the policy function (13), I replace it with the primitives when stating the optimal contract below.

Theorem 1 (Optimal Contract). *The principal's cost function and policy function are described by (12) and (13). Therefore, the principal's minimum cost is* $\gamma \star (\gamma \star f)'^2(0)/r$, *and the unique optimal contract reads*

$$x_t(\theta^t) \equiv x(\theta_t, W_t) = f(\theta_t) - \gamma \star f(\theta_t) + \underbrace{\gamma \star f(0)}_{W_0} + \underbrace{r \int_0^t \big(\gamma \star f(\theta_s) - f(\theta_s)\big)\, \mathrm{d}s}_{W_t - W_0}. \tag{14}$$

The theorem has several implications. First, rearranging terms in the policy function yields the Euler equation

$$\underbrace{x(\theta, W) - f(\theta)}_{\text{current distortion}} = \underbrace{W - \gamma \star f(\theta)}_{\text{future distortion}}.$$

The left-hand side is the gap between the current action x and the current target $f(\theta)$ or, simply, the *current distortion*. The right-hand side is the gap between the expected discounted future actions W and the expected discounted future targets $\gamma \star f(\theta)$, i.e., the *future distortion*. Due to the principal's convex flow cost, these distortions must be balanced intertemporally at optimum. As a result, the agent's continuation payoff W is smoothly cashed out over time.

Second, the cost function $C(\theta, W)$ is convex in W, implying that a random contract cannot further reduce cost for the principal. Indeed, if a random contract assigns a nondegenerate distribution to a set of deterministic contracts, the principal can alternatively choose the mean action according to this set of contracts. From this, the continuation payoff W also takes the mean value, which saves costs due to convexity. This argument, similar to that of Strausz (2006), verifies that it is without loss of generality to focus on deterministic contracts.

Third, by minimizing $C(0, W_0)$ with respect to W_0, I obtain the agent's ex ante payoff at $W_0 = \gamma \star f(0)$, i.e., the expected discounted future targets evaluated at the initial state. In other words, the actions should optimally match the targets *on average*.

Fourth, the minimized cost, $\gamma \star (\gamma \star f)'^2(0)/r$, is interpreted as the *agency cost*. This is because in a complete information problem where the agent receives W in expectation, the minimized cost could have been lowered to $(W - \gamma \star f(0))^2$, such that the extra term in (12) is indeed the agency cost. As the players become patient ($r \to 0$), the agency cost $\gamma \star (\gamma \star f)'^2(0)/r$ does not necessarily approach zero, because a Brownian motion does not have a stationary distribution. Indeed, the distortion $W_t - \gamma \star f(\theta_t)$ accumulates through the evolution of the contract, and the cost from the distortion grows over time. The ever-growing future costs weigh heavily at time zero due to lack of discounting. Below are two simple cases in point.

Example 1 (Linear). Consider a linear target function $f(\theta) = -\theta$ and let $\mu = 0$. Then $\gamma \star f(\theta) = -\theta$ and the agency cost reduces to $\gamma \star (\gamma \star f)'^2(0)/r = 1/r$. As $r \to 0$, the agency cost diverges to infinity. ◊

Example 2 (Binary). Consider a binary target function $f(\theta) = \mathbb{1}\{\theta \geq \overline{\theta}\}$ and let $\mu = 0$. Then $\gamma \star f(\theta)$ becomes $1 - \frac{1}{2}e^{-\sqrt{2r}|\theta-\overline{\theta}|}$ for $\theta \geq \overline{\theta}$ and $\frac{1}{2}e^{-\sqrt{2r}|\theta-\overline{\theta}|}$ for $\theta < \overline{\theta}$. The agency cost simplifies to $\frac{1}{3}e^{-\sqrt{2r}|\overline{\theta}|} - \frac{1}{6}e^{-2\sqrt{2r}|\overline{\theta}|}$. As $r \to 0$, it converges to $\frac{1}{6}$ regardless of $\overline{\theta}$. Therefore, even with a bounded target function, the agency cost does not vanish for patient players. ◊

4.2 *Response to information*

The optimal contract is the key to answering the following questions. First, how should the principal optimally respond to the agent's report? Second, does the principal always benefit from the agent's information? I start by defining the following terms to capture the response to information in the optimal contract.

Definition 1 (Response to Information). Suppose $f'(\theta) \neq 0$ exists at state θ. Contract x is called

(a) *conformist* at state θ if $\partial x(\theta, W)/\partial\theta$ has the same sign of $f'(\theta)$

(b) *contrarian* at state θ if $\partial x(\theta, W)/\partial\theta$ has the opposite sign of $f'(\theta)$

(c) *unresponsive* at state θ if $\partial x(\theta, W)/\partial\theta = 0$.

Contract x is called *babbling* if it is unresponsive at all states.

The conformist, contrarian, and unresponsive properties are locally defined at any given state. Moreover, the optimal contract (14) implies that these properties are attached to the current state, not the history. The babbling property, on the other hand, is a global characterization of the entire target function.

Theorem 2 below proposes two equivalent criteria to check the direction in which the contract should respond to information. The first criterion directly compares f' and $(\gamma \star f)'$. The second involves higher order derivatives.

Theorem 2 (Conformist or Contrarian). *At any state θ such that $f'(\theta) \neq 0$, the following statements are equivalent:*

(i) The optimal contract is conformist (resp. contrarian, unresponsive) at state θ.

(ii) It holds that $(\gamma \star f)'/f' < 1$ (resp. > 1, $= 1$) at state θ.

(iii) It holds that $(2\mu(\gamma \star f)'' + (\gamma \star f)''')f' < 0$ (resp. > 0, $= 0$) at state θ.

According to the theorem, it is fairly common for the optimal contract to exhibit contrarian actions, at least for some states. But why should the contrarian response ever be optimal? After all, it is tempting to match a higher target with an increased action. Here is the intuition. The slope f' of the target function can be loosely interpreted as the *information sensitivity* of the principal. A steeper target function means that the principal's bliss point is more sensitive to the state. For simplicity, consider $f' > 0$ in the following argument. If the state experiences a positive shock $d\theta > 0$, then the current target increases by $f'(\theta)\,d\theta$, and the principal is tempted to raise the current action. At the same time, due to the persistence of the state, the expected discounted future target increases by $(\gamma \star f)'(\theta)\,d\theta$, motivating the principal to elevate future actions as well. However, she cannot achieve both due to incentive constraints: higher current action inevitably leads to lower future actions and vice versa. Therefore, when $f'(\theta)$ is larger than $(\gamma \star f)'(\theta)$, the *current* information sensitivity is higher than its *future* counterpart, and the intertemporal trade-off favors the current quality of match. In this way, the action moves along with the target at the cost of the future quality of match, causing the contract to be conformist at this state. If, instead, the future information sensitivity $(\gamma \star f)'(\theta)$ is larger, then the principal optimally sacrifices the current quality of match in order to reduce the future distortion, leading to a contrarian response from the contract.

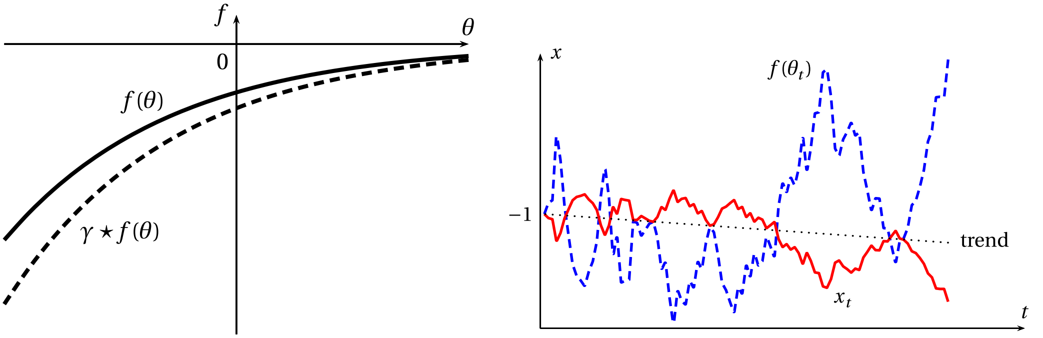
The potentially contrarian response to information constitutes a major departure from the literature on allocation problems (Jackson and Sonnenschein (2007), Renault, Solan, and Vieille (2013), Guo and Hörner (2020)). While an incentive compatible contract must feature a quota, the use of a quota in the literature has a "conformist" bent: spend the quota when the state demands higher actions and save it otherwise. In this model, however, the use of a quota can be either conformist or contrarian depending on the target function and the state process.

When is the contract more likely to be contrarian? Assuming $f'(\theta) > 0$, part (iii) of Theorem 2 boils down to two additive terms:

$$\underbrace{2\mu(\gamma \star f)''(\theta)}_{\text{drift effect}} + \underbrace{(\gamma \star f)'''(\theta)}_{\text{volatility effect}}.$$

The first term captures the *drift effect*. If the expected future target $\gamma \star f$ is convex, i.e., if the information sensitivity is increasing, then a positive drift μ boosts the relative importance of future quality of match, contributing to a more contrarian contract. The second term captures the *volatility effect*. If $\gamma \star f$ has a positive third derivative, i.e., if the information sensitivity is convex, then by Jensen's inequality, the information sensitivity tends to increase in the future. This also favors a more contrarian contract. The intuition for the case $f'(\theta) < 0$ is similar. The following three examples provide sample paths of optimal contracts for different target functions. For simplicity, $\mu = 0$ in all three examples.

Example 3 (Exponential). Consider an exponential target function $f(\theta) = b_0 e^{b_1\theta}$, where $|b_1| < \frac{1}{2}\sqrt{2r}$ such that Assumption 1 is satisfied. Then the expected future target is $\gamma \star f(\theta) = f(\theta) \cdot 2r/(2r - b_1^2)$. Figure 2(a) shows both f and $\gamma \star f$ for $b_0 = -1$ and $b_1 = -0.7$. Since $\gamma \star f'/f' = 2r/(2r - b_1^2) > 1$, Theorem 2(ii) predicts a contrarian pattern at all states. Figure 2(b) displays simulated paths of the target $f(\theta_t)$ and the action x_t. The auxiliary dotted curve represents the deterministic part of the action path, where all Brownian shocks are removed. Since $f(\theta_t)$ and x_t respond to new information in opposite directions, they never lie on the same side of the dotted curve. ◊

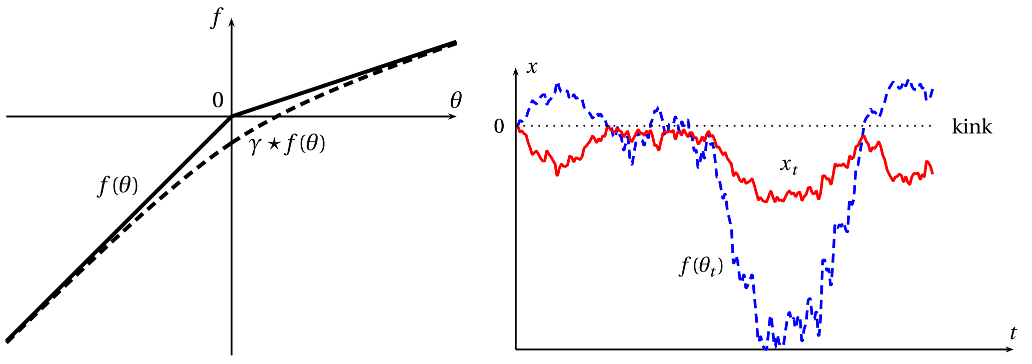


(a) Target (solid) and future target (dashed) (b) Target path (dashed) and action path (solid)

FIGURE 2. Exponential target: $r = 1$, $\mu = 0$, and $f(\theta) = -e^{-0.7\theta}$.

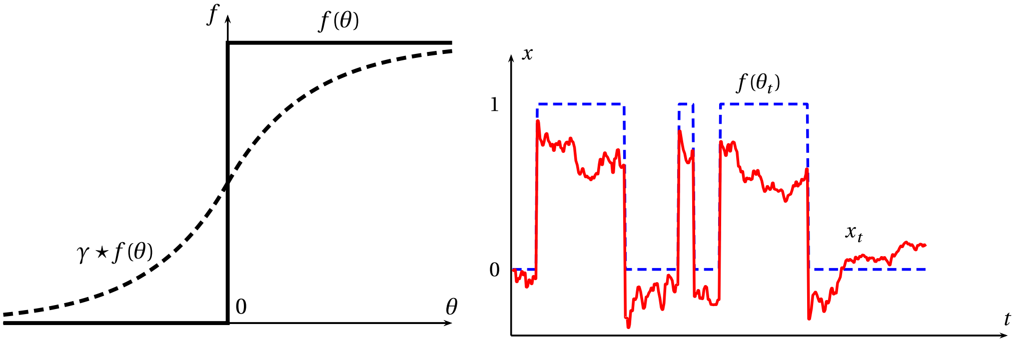
The key property of an exponential target f is that f''' and f' always have the same sign, so that $\gamma \star f$ is an amplified version of f. Whenever the current target increases, its future counterpart increases by even more.

EXAMPLE 4 (Kinked). Consider a kinked target function $f(\theta) = b_0\theta$ if $\theta < 0$ and $f(\theta) = b_1\theta$ if $\theta \geq 0$, where $b_0, b_1 > 0$ and $b_0 \neq b_1$. Then the expected future target is $\gamma \star f(\theta) = f(\theta) + e^{-\sqrt{2r}|\theta|}(b_1 - b_0)/(2\sqrt{2r})$. These functions are shown in Figure 3(a) for $b_1 = \frac{1}{3} < 1 = b_0$. Since $(\gamma\star f)' - f' = \frac{1}{2}(b_1 - b_0)e^{-\sqrt{2r}|\theta|}$ for $\theta > 0$ and $(\gamma\star f)' - f' = \frac{1}{2}(b_0 - b_1)e^{-\sqrt{2r}|\theta|}$ for $\theta < 0$, according to Theorem 2(ii), the action is contrarian at all states where f' is the smaller between b_0 and b_1, and is conformist otherwise. Figure 3(b) plots the simulated paths for the target $f(\theta_t)$ and the action x_t, where the two paths co-move whenever the target $f(\theta)$ is below its kink (i.e., on its steeper segment) and move out of phase otherwise. ♢



(a) Target (solid) and future target (dashed) (b) Target path (dashed) and action path (solid)

FIGURE 3. Kinked target: $r = 1$, $\mu = 0$, and $f(\theta) = \begin{cases} \theta, & \text{if } \theta < 0 \\ \frac{1}{3}\theta, & \text{if } \theta \geq 0. \end{cases}$



(a) Target (solid) and future target (dashed) (b) Target path (dashed) and action path (solid)

FIGURE 4. Binary target: $r = 1$, $\mu = 0$, and $f(\theta) = \mathbb{1}\{\theta \geq 0.4\}$.

Intuitively, the target's slope f' takes only two values, and, therefore, the expected future target $\gamma \star f$ must have a slope in between. On the flatter (steeper) segment of the target function, the current target is less (more) responsive to shocks than its future counterpart. This explains the coexisting patterns in the same contract. Notably, both a kinked target and an exponential target can be increasing and concave, but the optimal contracts are qualitatively different. Therefore, concavity or convexity of the target function alone is insufficient to determine the pattern of the contract.

EXAMPLE 5 (Binary). We revisit Example 2 for the binary target. The expected future target can be rewritten as $\gamma \star f(\theta) = f(\theta) - \frac{1}{2}e^{-\sqrt{2r}|\theta-\overline{\theta}|}\operatorname{sgn}(\theta - \overline{\theta})$. For $\overline{\theta} = 0.4$, these functions are plotted in Figure 4(a). Since f' is either zero or undefined, Definition 1 does not apply. For $\theta \neq \overline{\theta}$, $(\gamma \star f)' - f' = \sqrt{r/2}e^{-\sqrt{2r}|\theta-\overline{\theta}|} > 0$, so that the action always moves in the opposite direction of the state. At $\theta = \overline{\theta}$, (13) implies that the action jumps along with the target, which is "conformist" in a broader sense. Figure 4(b) simulates the time paths for the target and the action. Every time the target jumps, the optimal action follows suit. At other times, the action still responds to the state despite the constant target. $\Diamond$

This result obtains from the nature of the expected future target. The expected future target $\gamma \star f$ is everywhere increasing because it factors in the upward jump at $\theta = \overline{\theta}$. A marginal increase in the state, although not affecting the target, increases the expected future target and hence demands a shift of resources from present to future.

Below, Theorem 3 captures the knife-edge cases where the contract becomes babbling, in which the principal optimally stays unresponsive to the agent's reports. Such communication failures arise only for a nongeneric set of target functions. Otherwise, the principal always finds a direction to intertemporally reallocate actions to reduce cost.

THEOREM 3 (Impossibility). *(i) For $\mu = 0$, the contract is babbling if and only if the target is almost everywhere identical to $c_0 + c_1\theta + c_2\theta^2$ for some constants c_0, c_1, c_2.*

(ii) *For* $\mu \neq 0$, *the contract is babbling if and only if the target is almost everywhere identical to* $c_0 + c_1\theta + c_2 e^{-2\mu\theta}$ *for some constants* c_0, c_1, c_2.

According to the theorem, the curvature of the target function is not sufficient to guarantee the gains from information transmission. For example, when $\mu = 0$, effective information transmission requires a nonzero curvature of the information sensitivity or, equivalently, a nonzero third derivative of the target function. This is why quadratic target functions lead to babbling in part (i) of the theorem.

Although linear and quadratic target functions do not admit effective communication when the state is a Brownian motion without drift, this is specific to the state process. When the state follows some other process, say the Ornstein–Uhlenbeck process (see Section 5.1), information is utilized even with these target functions.

4.3 *Evolution of the contract*

Next, I discuss the stochastic evolution of the optimal contract on path, characterizing the dynamics of cost and payoff as the contract is executed over time.

Proposition 3 (Cost and Payoff Dynamics). *(i) The principal's continuation cost is a submartingale, i.e.,* $\mathbb{E}_t[\mathrm{d}C(\theta_t, W_t)]/\mathrm{d}t \geq 0$.

(ii) The agent's continuation payoff monotonically increases (resp. decreases) over time if $\gamma \star f(\theta) - f(\theta) > 0$ *(resp.* < 0*) for all* θ.

Part (i) of Proposition 3 claims that the principal faces statistically growing continuation costs $C_t \equiv C(\theta_t, W_t)$ as the contract is executed over time. Intuitively, this is because the current state is realized while future states can only be predicted. At the beginning, the minimized cost is $C(0, \gamma \star f(0))$. If, at a later time $t > 0$, the state becomes zero again, then the continuation payoff W_t governed by the incentive constraints will have almost surely wandered away from $\gamma \star f(0)$, and the continuation cost will have increased. The back-loading of the principal's cost is typical in dynamic mechanism design without transferable utilities (e.g., Guo and Hörner (2020)).

Part (ii) implies that the agent does not necessarily end up immiserated; instead, the trajectory of his payoffs depends on the shape of the target function. When the target function is convex, the expected future target $\gamma \star f$ is always larger than the current target f. Therefore, the agent's continuation payoff increases over time because the principal wants the action path to accommodate such an overall trend. When the target function is concave, the opposite is true and the agent is immiserated. In sum, the front-loading or back-loading of the agent's payoff is driven by the expected evolution of the target function.

5. Extensions

This section extends the main model in three directions. First, I consider a less persistent state process by introducing mean reversion. Second, I consider a finite time horizon to explore the nonstationary behavior of the contract. Finally, I study the effect of having an agent who is less patient than the principal.

5.1 *Mean reverting state process*

The persistence of the state is demonstrably important to the intertemporal trade-off: future information sensitivity can sometimes outweigh its current counterpart because a shock to the state will echo in the distant future. In the main model, the persistence is high in the sense that any shock is permanent without decay. In this section, I weaken the persistence and allow for mean reversion. This change in the state process has two implications: that the state has a stationary distribution and that the increment of the state is negatively correlated with the state. The mean reversion setting brings the model closer to the common wisdom found in the existing literature on dynamic allocation and explains why contrarian patterns rarely arise there.

The persistence is weaker when the state exhibits mean reversion. In this subsection, I consider an Ornstein–Uhlenbeck process of the form

$$\mathrm{d}\theta_t = -\phi\theta_t\,\mathrm{d}t + \mathrm{d}Z_t,$$

where $\phi \geq 0$ is a constant representing the strength of mean reversion. When $\phi = 0$, the process reduces to a special case of the main model. It can be shown that Proposition 1 still holds as a necessary condition. With a procedure similar to that used in the main model, the cost and policy functions are obtained as

$$C(\theta, W) = \big(W - \gamma_\phi \circ f(\theta)\big)^2 + \frac{1}{r}\gamma_\phi \circ (\gamma_\phi \circ f)'^2(\theta), \qquad x(\theta, W) = W + f(\theta) - \gamma_\phi \circ f(\theta),$$

where the operation $\gamma_\phi \circ f$ produces the unique solution $g = \gamma_\phi \circ f$ to the second-order differential equation

$$g''(\theta) - 2\phi\theta g'(\theta) - 2rg(\theta) = -2rf(\theta), \qquad \lim_{\theta\to\pm\infty} e^{-\sqrt{\frac{r}{2}}|\theta|}g(\theta) = 0. \tag{15}$$

While an explicit solution to (15) is not obtainable in general, one can derive an alternative expression for $\gamma_\phi \circ f$ by means of a forward stochastic differential equation

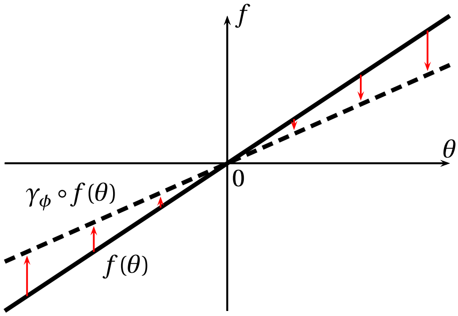
$$\gamma_\phi \circ f = \mathbb{E}\left[\int_0^\infty re^{-rt}f(\theta_t)\,\mathrm{d}t \,\middle|\, \theta_0 = \theta\right], \qquad \mathrm{d}\theta_t = -\phi\theta_t\,\mathrm{d}t + \mathrm{d}Z_t.$$

Even with mean reversion, the term $\gamma_\phi \circ f(\theta)$ is again interpreted as the *expected discounted future target*, similar to that in the main model. Whether the contract is conformist or contrarian depends now on the comparison between f' and $(\gamma_\phi \circ f)'$.

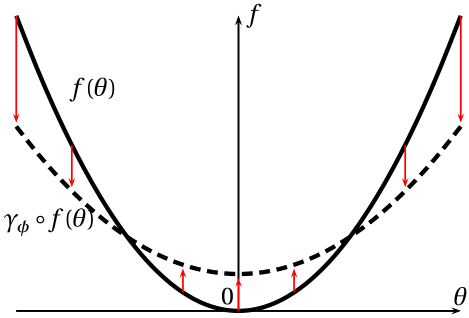
It is not easy to directly compare optimal contracts with different parameters ϕ. That being said, comparison is possible in the special case where f is a polynomial. When $f(\theta) = \sum_{k=0}^n b_k\theta^k$, the future target function $\gamma_\phi \circ f(\theta)$ is a polynomial of the same order, but the coefficient on the highest order n is dampened toward zero and becomes $b_n \cdot r/(r + n\phi)$. Therefore,

$$\lim_{\theta\to\pm\infty} \frac{(\gamma_\phi \circ f)'(\theta)}{f'(\theta)} = \frac{r}{r + n\phi} < 1,$$

meaning that the contract is conformist whenever the state is sufficiently far from zero. Intuitively, states far away from zero have a strong tendency to drift back and, hence,



(a) A linear target function



(b) A quadratic target function

FIGURE 5. The effect of mean reversion ($r = 1$, $\phi = 0.5$). (a) The target $f(\theta) = \theta$ (solid) and the future target $\gamma_\phi \circ f(\theta)$ (dashed). (b) The target $f(\theta) = \theta^2$ (solid) and the future target $\gamma_\phi \circ f(\theta)$ (dashed).

weigh less in the expected future target than they do in the case of zero mean reversion. As a result, the future information sensitivity $(\gamma_\phi \circ f)'(\theta)$ gives disproportionally large probability weight to states near zero, attenuating itself below $f'(\theta)$ when $|\theta|$ is large.

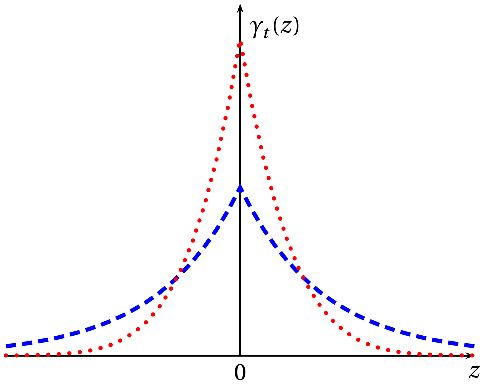
Figure 5 shows two simple examples where f is a polynomial. Figure 5(a) features a linear target where $\gamma_\phi \circ f$ is flatter than f and, hence, the contract is conformist everywhere. According to Theorem 3, without mean reversion, we would have ended up with a babbling outcome. Figure 5(b) plots the case of a quadratic target. The future target $\gamma_\phi \circ f$ has a dampened slope compared to f, and the contract is conformist at all states except zero. Again, the contract does not lead to babbling, although it would have if $\phi = 0$.

5.2 *Finite horizon*

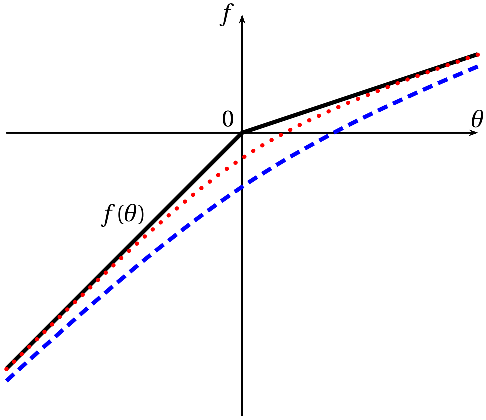
In some economic applications, the time horizon for a contract is relatively short, because the principal is either unable or unwilling to commit for a long period. The contracting horizon can also be short because the state or cost becomes visible to the principal after some period of time. The finite horizon introduces nonstationarity to the contracting environment and creates a deadline effect on top of the patterns found in the main model.

For simplicity, we assume in this subsection that $\mu = 0$. The contracting horizon is $T > 0$, and both players discount at rate $r > 0$. Let $W_t \equiv \mathbb{E}_t[\int_t^T e^{-r(s-t)} x_s \, ds]/(1 - e^{-r(T-t)})$ be the normalized expected continuation payoff of the agent at time t, and write $C(\theta, W, t)$ as the principal's normalized cost function when the state is θ, the agent's continuation payoff is W, and the calendar time is t. The HJB equation is modified to

$$rC(\theta, W, t) = \min_x r\bigl(x - f(\theta)\bigr)^2 + r(W - x)C_W(\theta, W, t) + \bigl(1 - e^{-r(T-t)}\bigr)C_t(\theta, W, t) + \frac{1}{2}\bigl(1 - e^{-r(T-t)}\bigr)C_{\theta\theta}(\theta, W, t).$$



(a) The kernel γ_t at various t



(b) A target function f (solid) with $\gamma_t \star f$ at various t

FIGURE 6. The change of γ_t over time, showing the deadline effect. Parameters: $r = 1$, $\mu = 0$, $T = 4$, $t = 0$ (dashed), and $t = 3.6$ (dotted).

Even with three state variables, the above system is still solvable thanks to the quadratic flow cost. Following the same procedure as in the main model, I find the cost and policy functions taking similar forms

$$C(\theta, W, t) = \big(W - \gamma_t \star f(\theta)\big)^2 + \frac{1 - e^{-r(T-t)}}{r} \gamma_t \star (\gamma_t \star f)'^2(\theta)$$
$$x(\theta, W, t) = W + f(\theta) - \gamma_t \star f(\theta),$$

where

$$\gamma_t(z) \equiv \frac{\sqrt{r}}{2\sqrt{2}\big(1 - e^{-r(T-t)}\big)} \cdot \left(e^{-\sqrt{2r}|z|} \operatorname{Erfc}\left(\frac{\sqrt{2}|z| - 2\sqrt{r}(T-t)}{2\sqrt{T-t}}\right) - e^{\sqrt{2r}|z|} \operatorname{Erfc}\left(\frac{\sqrt{2}|z| + 2\sqrt{r}(T-t)}{2\sqrt{T-t}}\right)\right)$$

is a time-dependent kernel. Figure 6(a) plots the kernel at different calendar times. As t increases, the kernel is gradually concentrated around zero. In the limit as $t \to T$, the kernel collapses to a Dirac delta function. Holding t fixed while extending T to infinity, the kernel converges to a Laplace distribution as in the main model. Once again, $\gamma_t \star f(\theta) = \mathbb{E}[\int_t^T e^{-r(s-t)} f(\theta_s)\, \mathrm{d}s | \theta_t = \theta]$ is the expected discounted future targets. Figure 6(b) shows this $\gamma_t \star f$ evaluated at different times. As time passes, the "future" is shorter, and, therefore, less probability weight is given to states far from θ in the above expectation. In the limit, we have $\lim_{t\to T} \gamma_t \star f(\theta) = f(\theta)$ for all θ at which f is continuous.

How does the response to information change over time as the contract approaches the end of the horizon? This is determined by the comparison between f' and $(\gamma_t \star f)'$, according to the policy function. With a finite horizon, the expected future information sensitivity $(\gamma_t \star f)'$ depends not only on the state, but also on calendar time t. Due to the required smoothness of function f, $\lim_{t\to T}(\gamma_t \star f)'(\theta) = f'(\theta)$ almost everywhere. In other words, the gap $f' - (\gamma_t \star f)'$ tends to zero and the responsiveness to information vanishes as the deadline approaches. As a result, the agent loses his influence on the action over time, and the information transmission gradually reduces to babbling. This is consistent with the two-period example, wherein the second-period information is disregarded.

5.3 *Less patient agent*

In some agency problems, the principal has a longer horizon and is thus more patient than the agent. Krasikov, Lamba, and Mettral (2020) study a dynamic contracting model with unequal discounting, where the patience gap generates a front-loading motive for the agent's payoff that is interacting with the initial back-loading force. In this extension, I study how the patience gap affects the optimal response to information.

Suppose the principal discounts at rate $r > 0$, while the agent has a higher rate $\rho > r$. To keep notation simple, I set $\mu = 0$. Since the incentive constraint is evaluated on behalf of the agent, the first-order condition (FOC) now requires $\mathrm{d}W_t = \rho(W_t - x_t)\,\mathrm{d}t$. In the HJB equation for the cost function, however, it is the principal's discount rate r that takes place:

$$rC(\theta, W) = \min_x r\big(x - f(\theta)\big)^2 + \rho(W - x)C_W(\theta, W) + \frac{1}{2}C_{\theta\theta}(\theta, W).$$

The policy function can be similarly obtained, leading to the Euler equation

$$x(\theta, W) - f(\theta) = \frac{2\rho - r}{\rho}\big(W - \gamma_\rho \star f(\theta)\big),$$

where γ_ρ has the same expression as γ in the main model except that $\mu = 0$ and r is replaced by ρ.

Two immediate changes arise from the policy function. First, the current action no longer cashes out the promised continuation payoff W proportionally. Instead, since $(2\rho - r)/\rho > 1$, the principal front-loads payoffs to the agent to exploit the difference in patience.

Second, the response to information is now determined by $\partial x/\partial\theta = f'(\theta) - \gamma_\rho \star f'(\theta) \cdot (2\rho - r)/\rho$. Directly, the future information sensitivity is amplified by a factor $(2\rho - r)/\rho > 1$, making it easier to overtake the current information sensitivity. Indirectly, the distribution γ_ρ is more concentrated around zero due to the higher ρ, pulling the future information sensitivity toward its current counterpart. While the overall effect is ambiguous, the direct effect dominates in many cases. As an example, let $f(\theta) = \theta$. The contract in the main model is babbling, but now since $\rho > r$, the principal can benefit from the information to some extent. Since $(f'(\theta) - \frac{2\rho - r}{\rho}\gamma_\rho \star f'(\theta))/f'(\theta) = 1 - (2\rho - r)/\rho < 0$, the contract is always contrarian.

6. Concluding remarks

The fact that the optimal contract can, as demonstrated in this paper, behave in a contrarian manner is a noteworthy departure from the existing literature that explores similar problems. The contrarian pattern is unique to an intrinsically dynamic situation, wherein information unfolds over time. For contrast, consider a static setting wherein the agent observes the realization of the entire state path before reporting the path to the principal at once. In that case, the principal always responds to information in a conformist pattern, because with all states available, the current state no longer plays the dual role of justifying current action and predicting future actions.

The contrarian pattern of the contract can be viewed as a new implication for agency problems; it never arises if there are no conflicts of interest. The more aligned the preferences are, the less likely it is that the optimal contract will be contrarian. The principal's action moving in the opposite direction of the agent's report should not be interpreted as distrust or punishment; instead, it can be understood as the principal's efficient way of utilizing information in the presence of conflicting interests.

Appendix

A.1 *Solving the two-period contract*

The IC's are simplified to one equation: $x_1(\theta_1) + x_2(\theta_1) = W$. Plug this back into the objective to obtain the unconstrained problem

$$\min_{x_1(\cdot), W} \mathbb{E}\big[\big(x_1(\theta_1) - f(\theta_1)\big)^2 + \mathbb{E}\big[\big(W - x_1(\theta_1) - f(\theta_2)\big)^2 | \theta_1\big]\big].$$

For every θ_1, the FOC with respect to $x_1(\theta_1)$ gives $x_1(\theta_1) = \frac{1}{2}W + \frac{1}{2}(f(\theta_1) - \mathbb{E}[f(\theta_2)|\theta_1])$. Plugging this into the objective and taking FOC with respect to W, we have $W = \mathbb{E}f(\theta_1) + \mathbb{E}f(\theta_2)$. Replacing W in the expression of $x_1(\theta_1)$ with the above, we have the solution for x_1: $x_1^*(\theta_1) = \frac{1}{2}(f(\theta_1) - \mathbb{E}[f(\theta_2)|\theta_1]) + \frac{1}{2}(\mathbb{E}f(\theta_1) + \mathbb{E}f(\theta_2))$. We then find x_2 from the incentive compatibility (IC) condition: $x_2^*(\theta_1) = -\frac{1}{2}(f(\theta_1) - \mathbb{E}[f(\theta_2)|\theta_1]) + \frac{1}{2}(\mathbb{E}f(\theta_1) + \mathbb{E}f(\theta_2))$.

A.2 *Revelation principle*

Lemma 3 (Revelation Principle). *Given any contract x that implements a mapping from state paths into action paths, there exists a truthful contract $x^\dagger$ that implements the same mapping.*

Proof. Suppose the given contract x induces a (not necessarily truthful) strategy $m \in \mathcal{M}$, which generates a mapping from state paths into action paths. Let $M_t \equiv \int_0^t m_s \, \mathrm{d}s$ be the accumulated misreporting. Consider a new contract $x^\dagger$ such that $x_t^\dagger(\hat{\theta}^t) \equiv x_t((\hat{\theta} + M)^t)$. If the truth-telling strategy $m^\dagger$ is not optimal for the agent under this contract, there exists a strategy $m' \in \mathcal{M}$ along with $M_t' \equiv \int_0^t m_s' \, \mathrm{d}s$ such that $\mathbb{E}[\int_0^\infty re^{-rt} x_t^\dagger((\theta +$

$M')^t)\,\mathrm{d}t] > \mathbb{E}[\int_0^\infty re^{-rt}x_t^\dagger(\theta^t)\,\mathrm{d}t]$. Contradiction arises as $m+m' \in \mathcal{M}$ outperforms m in the original contract:

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty re^{-rt}x_t\big((\theta+M+M')^t\big)\,\mathrm{d}t\right] &= \mathbb{E}\left[\int_0^\infty re^{-rt}x_t^\dagger\big((\theta+M')^t\big)\,\mathrm{d}t\right] \\ &> \mathbb{E}\left[\int_0^\infty re^{-rt}x_t^\dagger\big(\theta^t\big)\,\mathrm{d}t\right] \\ &= \mathbb{E}\left[\int_0^\infty re^{-rt}x_t\big((\theta+M)^t\big)\,\mathrm{d}t\right].\end{aligned}$$

The new contract $x^\dagger$ implements the original mapping from θ^t to x^t by construction, $\forall t$. □

A.3 *Proof of Lemma 1*

Given a contract x, define the process of the agent's total payoff evaluated at time 0 but with information at time t,

$$\hat{W}_t^0 \equiv \int_0^t re^{-rs}x_s\,\mathrm{d}s + e^{-rt}W_t,$$

which is a martingale because for any $0 \le t' \le t$,

$$\mathbb{E}_{t'}\hat{W}_t^0 = \int_0^{t'} re^{-rs}x_s\,\mathrm{d}s + \mathbb{E}_{t'}\left[\int_{t'}^t re^{-rs}x_s\,\mathrm{d}s\right] + e^{-rt}\mathbb{E}_{t'}\left[\int_t^\infty re^{-r(s-t)}x_s\,\mathrm{d}s\right] = \hat{W}_{t'}^0.$$

By Theorem 1.3.13 in Karatzas and Shreve (1991), the martingale $\hat{W}_t^0$ has a right-continuous-with-left-limit (RCLL) modification. Therefore, by Theorem 3.4.15 in the same book, the martingale has a representation:

$$\hat{W}_t^0 = \hat{W}_0^0 + \int_0^t re^{-rs}\beta_s\,\mathrm{d}Z_s \quad \forall t \ge 0.$$

Subtracting the two expressions for $\hat{W}_t^0$ and then differentiating with respect to t, we have

$$\mathrm{d}W_t = r(W_t - x_t)\,\mathrm{d}t + r\beta_t\,\mathrm{d}Z_t = r(W_t - x_t)\,\mathrm{d}t + r\beta_t(\mathrm{d}\hat{\theta}_t - \mu\,\mathrm{d}t),$$

which has an equivalent integral form $W_t = W_0 + \int_0^t r(W_s - x_s)\,\mathrm{d}s + \int_0^t r\beta_s\,\mathrm{d}Z_s$.

A.4 *Proof of Proposition 1*

For any strategy $m \in \mathcal{M}$, Novikov's condition is satisfied. By the Girsanov theorem, there exists a martingale Y with $Y_t \equiv e^{\int_0^t m_s\,\mathrm{d}Z_s - \frac{1}{2}\int_0^t m_s^2\,\mathrm{d}s}$, serving as the Radon–Nikodym derivative between the measure induced by m and the measure under truth-telling. It evolves according to $\mathrm{d}Y_t = Y_t m_t\,\mathrm{d}Z_t$ with $Y_0 = 1$. Besides Y_t, the cumulative misreporting $M_t = \int_0^t m_s\,\mathrm{d}s$ is also a state variable, with evolution $\mathrm{d}M_t = m_t\,\mathrm{d}t$. Then the agent's payoff from a strategy $m \in \mathcal{M}$ is $\mathbb{E}[\int_0^\infty re^{-rt}Y_t x_t\,\mathrm{d}t]$.

Let p^Y be the costate variable for the drift of Y and let q^Y be the costate for the volatility of Y. Let p^M and q^M be the counterparts for M. The agent's current value Hamiltonian is $rYx + q^Y Ym + p^M m$.

The first-order condition for $m = 0$ to be optimal, evaluated at $m = 0$, $Y = 1$, is

$$q^Y + p^M = 0. \tag{16}$$

The Euler equations for Y and M, evaluated at $m = 0$, $Y = 1$, are

$$\begin{aligned} \mathrm{d}p^Y &= r\big(p^Y - x\big)\,\mathrm{d}t + q^Y\,\mathrm{d}Z_t \\ \mathrm{d}p^M &= rp^M\,\mathrm{d}t + q^Y\,\mathrm{d}Z_t, \end{aligned} \tag{17}$$

with transversality conditions $\lim_{t\to\infty} p_t^Y e^{-rt} = 0$ and $\lim_{t\to\infty} p_t^M e^{-rt} = 0$. The solution to the above backward stochastic differential equations are $p_t^Y = \mathbb{E}_t[\int_t^\infty re^{-r(s-t)} x_s\,\mathrm{d}s] = W_t$ and $p_t^M = 0$, where W_t is the agent's continuation payoff defined in Section 3.1. Hence, by comparing (17) and (7), we have $q^Y = r\beta$. Plugging this back into (16) and using the fact that $p^M = 0$, we have the necessary condition $\beta = 0$.

A.5 *Proof of Proposition 2*

Suppose the agent uses an arbitrary strategy $m \in \mathcal{M}$, so that the reported process is $\hat{\theta} = \theta + M$, where $M_t = \int_0^t m_t\,\mathrm{d}t$. The resulting action and continuation payoff processes are denoted as x^m and W^m. Because $\hat{\theta}$ is in the support of of θ, these two processes evolve as if $\hat{\theta}$ was the true state and the agent reported truthfully. Therefore, plugging in $\beta_t \equiv 0$, we have $\mathrm{d}W_t^m = r(W_t^m - x_t^m)\,\mathrm{d}t$ and, thus, $x_t^m\,\mathrm{d}t = -(e^{rt}/r)\,\mathrm{d}(e^{-rt}W_t^m)$. The agent's payoff from strategy m is

$$\lim_{t\to\infty} \mathbb{E}\int_0^t re^{-rs}x_s^m\,\mathrm{d}s = W_0 - \lim_{t\to\infty} e^{-rt}\mathbb{E}W_t^m,$$

which means that as long as $\lim_{t\to\infty} e^{-rt}\mathbb{E}W_t^m = 0$ for all $m \in \mathcal{M}$, the agent's payoff is always W_0 regardless of his strategy.

A.6 *Proof of Lemma 2*

Assume toward a contradiction that another contract $\hat{x} + a \neq x + a$ achieves $\mathbb{E}[\int_0^\infty re^{-rt}(\hat{x}_t + a - f(\theta_t))^2] < \mathbb{E}[\int_0^\infty re^{-rt}(x_t + a - f(\theta_t))^2]$ and $\int_0^\infty re^{-rt}(\hat{x}_t + a) = W_0 + a$. This means $\int_0^\infty re^{-rt}\hat{x}_t = W_0$ and

$$\begin{aligned} &\mathbb{E}\bigg[\int_0^\infty re^{-rt}\big(\hat{x}_t - f(\theta_t)\big)^2\bigg] \\ &\quad = \mathbb{E}\bigg[\int_0^\infty re^{-rt}\big(\hat{x}_t + a - f(\theta_t)\big)^2\bigg] - a^2 - 2a\mathbb{E}\bigg[\int_0^\infty re^{-rt}\big(\hat{x}_t - f(\theta_t)\big)\bigg] \\ &\quad < \mathbb{E}\bigg[\int_0^\infty re^{-rt}\big(x_t + a - f(\theta_t)\big)^2\bigg] - a^2 - 2a\mathbb{E}\bigg[\int_0^\infty re^{-rt}\big(x_t - f(\theta_t)\big)\bigg] \\ &\quad = \mathbb{E}\bigg[\int_0^\infty re^{-rt}\big(x_t - f(\theta_t)\big)^2\bigg], \end{aligned}$$

which contradicts the fact that x is the minimizer of $\mathbb{E}[\int_0^\infty re^{-rt}(x_t - f(\theta_t))^2]$ subject to $\int_0^\infty re^{-rt}x_t = W_0$.

A.7 *Proof of Theorem 1*

The proof takes two steps. In Step 1, I show that the candidate solution, (12) and (13), indeed achieves the lowest cost in the relaxed problem. In Step 2, I show that the candidate solution also satisfies the global IC conditions.

Step 1. For any contract $\hat{x}$ satisfying the IC necessary condition and the transversality condition of the agent, define $\hat{W}$ as the resulting continuation payoff process where $\hat{W}_0 = W_0$, and define

$$\hat{C}_t^0 \equiv \int_0^t re^{-rs}\big(\hat{x}_s - f(\theta_s)\big)^2 \,\mathrm{d}s + e^{-rt}C(\theta_t, \hat{W}_t) \tag{18}$$

as the total cost process evaluated at time t. In this process, the policy follows the arbitrary contract $\hat{x}$ until time t and then the candidate cost function takes place as continuation, promising $\hat{W}_t$ as continuation payoff. The goal is to show that $\hat{C}_t^0$ is a martingale if $\hat{x}$ coincides with the optimal policy (13) and a submartingale if not. The total differential for $\hat{C}_t^0$ is

$$\begin{aligned} e^{rt}\,\mathrm{d}\hat{C}_t^0 &= r\big(\hat{x}_t - f(\theta_t)\big)^2 \,\mathrm{d}t - rC(\theta_t, \hat{W}_t)\,\mathrm{d}t + r(\hat{W}_t - \hat{x}_t)C_W(\theta_t, \hat{W}_t)\,\mathrm{d}t \\ &\quad + C_\theta(\theta_t, \hat{W}_t)\,\mathrm{d}Z_t + \mu C_\theta(\theta_t, \hat{W}_t)\,\mathrm{d}t + \frac{1}{2}C_{\theta\theta}(\theta_t, \hat{W}_t)\,\mathrm{d}t \\ &= C_\theta(\theta_t, \hat{W}_t)\,\mathrm{d}Z_t + r(\hat{x}_t - x_t)\big(\hat{x}_t + x_t - 2f(\theta_t) - C_W(\theta_t, \hat{W}_t)\big)\,\mathrm{d}t \\ &= C_\theta(\theta_t, \hat{W}_t)\,\mathrm{d}Z_t + r(\hat{x}_t - x_t)^2\,\mathrm{d}t, \end{aligned}$$

where $x_t = x(\theta_t, \hat{W}_t)$ is the candidate policy, the second equality follows from the HJB equation (8), and the third equality utilizes the policy function (13). It is clear that

$$e^{rt}\frac{\mathbb{E}_t\,\mathrm{d}\hat{C}_t^0}{\mathrm{d}t} = r(\hat{x}_t - x_t)^2 \geq 0,$$

with equality if and only if $\hat{x}_t = x_t$.

In the following discussion, I show that the arbitrary contract $\hat{x}$ does not achieve a lower cost, given the agent's truthful report. For any initial value W_0,

$$C(0, W_0) = \hat{C}_0^0 \leq \mathbb{E}\hat{C}_\infty^0 = \mathbb{E}\int_0^\infty re^{-rs}\big(\hat{x}_s - f(\theta_s)\big)^2\,\mathrm{d}s + \lim_{t\to\infty}\mathbb{E}e^{-rt}C(\theta_t, \hat{W}_t).$$

If $\mathbb{E}\int_0^\infty re^{-rs}(\hat{x}_s - f(\theta_s))^2\,\mathrm{d}s = \infty$, then this contract $\hat{x}$ results in infinite cost. Now suppose $\mathbb{E}\int_0^\infty re^{-rs}(\hat{x}_s - f(\theta_s))^2\,\mathrm{d}s < \infty$, i.e., $\mathbb{E}\int_0^\infty(\hat{x}_s - f(\theta_s))^2\,\mathrm{d}(-e^{-rs}) < \infty$. This means, with respect to the finite product measure, $\hat{x} - f(\theta) \in \mathcal{L}^2$. At the same time, it is straightforward to verify that $f(\theta) \in \mathcal{L}^2$ under Assumption 1. By the closure to addition of $\mathcal{L}^2$, one arrives at the conclusion that $\mathbb{E}\int_0^\infty re^{-rs}\hat{x}_s^2\,\mathrm{d}s < \infty$. For any $\hat{W}$ satisfying the agent's transversality condition, we must have $\hat{W}_t = \mathbb{E}_t \int_0^\infty \hat{x}_{t+s}\,\mathrm{d}(-e^{-rs})$. Therefore,

$$\begin{aligned} e^{rt}\mathbb{E}\hat{W}_t^2 &= e^{rt}\mathbb{E}\left(\int_0^\infty \hat{x}_{t+s}\,\mathrm{d}\big(-e^{-rs}\big)\right)^2 \\ &\le e^{rt}\mathbb{E}\int_0^\infty \hat{x}_{t+s}^2\,\mathrm{d}\big(-e^{-rs}\big) = \mathbb{E}\int_t^\infty re^{-rs}\hat{x}_s^2\,\mathrm{d}s < \infty, \end{aligned}$$

following Hölder's inequality. Letting $t \to \infty$, we have $\lim_{t\to\infty} \mathbb{E}[e^{rt}\hat{W}_t^2] = 0$. Moreover, it is straightforward to verify that $\mathbb{E}[e^{rt}(\gamma \star f(\theta_t))^2]$ vanishes. Hence, $\lim_{t\to\infty} \mathbb{E}[e^{-rt}C(\theta_t, \hat{W}_t)] = 0$ by noticing that $(\hat{W}_t - \gamma \star f(\theta_t))^2 \le 2(\hat{W}_t^2 + (\gamma \star f(\theta_t))^2)$. This means $C(0, W_0) \le \mathbb{E}\int_0^\infty re^{-rs}(\hat{x}_s - f(\theta_s))^2\,\mathrm{d}s$.

Step 2. It remains to verify that the sufficient condition (the condition in Proposition 2) for IC is satisfied. Suppose the agent adopts an arbitrary misreporting strategy $m \in \mathcal{M}$, so that the reported process is $\hat{\theta} = \theta + M$, where $M_t = \int_0^t m_t\,\mathrm{d}t$. The resulting action and continuation payoff processes are denoted as x^m and W^m. There exists $T > 0$ such that $\mathbb{E}e^{\sqrt{2r+\mu^2}M_t} < e^{rt}$ for all $t > T$. For any such $t > T$,

$$\begin{aligned} \int_0^t \mathbb{E}e^{\alpha_1(\theta_s+M_s)}\,\mathrm{d}s &\le \int_0^t \sqrt{\mathbb{E}e^{2\alpha_1\theta_s}}\sqrt{\mathbb{E}e^{2\alpha_1 M_s}}\,\mathrm{d}s \\ &\le \int_0^t e^{\alpha_1(\mu+\alpha_1)s}\big(\mathbb{E}e^{2\overline{\alpha}M_s}\big)^{\frac{\alpha_1}{2\overline{\alpha}}}\,\mathrm{d}s \\ &= \int_0^T e^{\alpha_1(\mu+\alpha_1)s}\big(\mathbb{E}e^{2\overline{\alpha}M_s}\big)^{\frac{\alpha_1}{2\overline{\alpha}}}\,\mathrm{d}s + \int_T^t e^{\alpha_1(\mu+\alpha_1)s}\big(\mathbb{E}e^{2\overline{\alpha}M_s}\big)^{\frac{\alpha_1}{2\overline{\alpha}}}\,\mathrm{d}s \\ &\le \int_0^T e^{\alpha_1(\mu+\alpha_1)s}\big(\mathbb{E}e^{2\overline{\alpha}M_s}\big)^{\frac{\alpha_1}{2\overline{\alpha}}}\,\mathrm{d}s + \frac{e^{\alpha_2 t} - e^{\alpha_2 T}}{\alpha_2}, \end{aligned}$$

where $\alpha_2 \equiv \alpha_1(\mu + \alpha_1) + r\alpha_1/(2\overline{\alpha}) < r$. Hence, the first term in the last line is finite while the second term grows slower than e^{rt}. Similarly, $\int_0^t \mathbb{E}e^{-\alpha_1(\theta_s+M_s)}\,\mathrm{d}s$ grows slower than e^{rt} too. With the candidate policy function,

$$\begin{aligned} \left|\frac{\mathrm{d}W_t^m}{\mathrm{d}t}\right| &= r\big|\gamma \star f(\theta_t + M_t) - f(\theta_t + M_t)\big| \\ &\le r\alpha_0\frac{2\alpha_1\mu + 4r - \alpha_1^2}{2\alpha_1\mu + 2r - \alpha_1^2}e^{-\alpha_1(\theta_t+M_t)} + r\alpha_0\frac{-2\alpha_1\mu + 4r - \alpha_1^2}{-2\alpha_1\mu + 2r - \alpha_1^2}e^{\alpha_1(\theta_t+M_t)}; \end{aligned}$$

therefore, with the above analysis, $\lim_{t\to\infty} e^{-rt}\mathbb{E}W_t^m = \lim_{t\to\infty} e^{-rt}\mathbb{E}|W_t^m| = 0$.

Finally, to obtain the lowest cost for the principal as well as the explicit optimal contract, I set W_0 optimally at $\gamma \star f(0)$. Also, with the policy function, the continuation payoff evolves as

$$W_t = W_0 + \int_0^t \mathrm{d}W_s = \gamma \star f(0) + r\int_0^t \big(\gamma \star f(\theta_s) - f(\theta_s)\big)\,\mathrm{d}s.$$

Hence, plug W_t into (13) to obtain (14).

A.8 *Proof of Theorem 2*

It suffices to show that $\partial x/\partial\theta$ exists whenever f' does, and

$$\frac{\partial x}{\partial\theta} = f'(\theta) - (\gamma\star f)'(\theta) = -\frac{2\mu(\gamma\star f)''(\theta) + \sigma^2(\gamma\star f)'''(\theta)}{2r}.$$

In (14), taking the derivative of $x_t(\theta^t)$ with respect to θ_t, one has

$$\left.\frac{\partial x_t}{\partial\theta_t}\right|_{\theta_t=\theta} = f'(\theta) - (\gamma\star f)'(\theta).$$

Also, recall that $g = \gamma\star f$ is the solution to (10). Further differentiation with respect to θ gives

$$-\frac{1}{2r}(\gamma\star f)'''(\theta) - \frac{\mu}{r}(\gamma\star f)''(\theta) = f'(\theta) - (\gamma\star f)'(\theta).$$

Finally, by definition of γ, we can verify that $\gamma\star f$ is continuously differentiable. Therefore, $\partial x/\partial\theta = f'(\theta) - (\gamma\star f)'(\theta)$ exists whenever f' does.

A.9 *Proof of Theorem 3*

For part (i), the "if" direction can be verified by plugging in $f(\theta) = c_0 + c_1\theta + c_2\theta^2$. The resulting action path $x_t = f(0) + c_2 t$ is deterministic, achieving babbling. For the "only if" direction, define $\hat{C}_t^0$ the same way as in (18) for the babbling contract $\hat{x}$. The drift of $\hat{C}_t^0$ satisfies $e^{rt}\mathbb{E}_t[\mathrm{d}\hat{C}_t^0]/\mathrm{d}t = r(x_t - \hat{x}_t)^2$. In order to achieve the babbling cost, we need $x_t = \hat{x}_t$ almost surely, which means the optimal policy should be state-independent almost surely. Through (13), this requires $f - \gamma\star f$ to be a constant for almost all θ. When $\mu = 0$, this implies that $(\gamma\star f)'' = -2r(f - \gamma\star f)$ is a constant almost everywhere. From the smoothness of f, $\gamma\star f$ is twice differentiable, so that $(\gamma\star f)''$ is a constant, meaning $\gamma\star f(\theta) = \tilde{c}_0 + c_1\theta + c_2\theta^2$. This integral equation has the unique continuous solution $f(\theta) = (\tilde{c}_0 - c_2/r) + c_1\theta + c_2\theta^2$, where $\tilde{c}_0 - c_2/r$ can be rewritten as c_0. Modification of the above on a zero-measure set generates an equivalence class.

For part (ii), the "if" direction can be verified by plugging in $f(\theta) = c_0 + c_1\theta + c_2 e^{-2\mu\theta}$. The resulting action path $x_t = f(0) + c_1\mu t$ is deterministic. For the "only if" direction, repeat the same procedure in the proof of part (ii). When $\mu \neq 0$, this implies that $(\gamma\star f)'' + 2\mu(\gamma\star f)' = -2r(f - \gamma\star f)$ is a constant, meaning $\gamma\star f(\theta) = \tilde{c}_0 + c_1\theta + c_2 e^{-2\mu\theta}$. This integral equation has the unique continuous solution $f(\theta) = (\tilde{c}_0 - c_1\mu/r) + c_1\theta + c_2 e^{-2\mu\theta}$, where $\tilde{c}_0 - c_1\mu/r$ can be rewritten as c_0. Modification of the above on a zero-measure set generates an equivalence class.

A.10 *Proof of Proposition 3*

To show (i), note that

$$\begin{aligned}\frac{\mathbb{E}_t[\mathrm{d}C_t]}{\mathrm{d}t} &= \mu C_\theta(\theta_t, W_t) + r(W_t - x_t)C_W(\theta_t, W_t) + \frac{1}{2}C_{\theta\theta}(\theta_t, W_t)\\ &= r\bigl(C(\theta_t, W_t) - \bigl(x_t - f(\theta_t)\bigr)^2\bigr)\\ &= \gamma\star\bigl((\gamma\star f)'^2\bigr)(\theta_t) \geq 0,\end{aligned}$$

where the first equality follows from Ito's lemma, the second follows from the HJB equation (8), the third follows by plugging in the policy and cost functions, and the inequality comes from the fact that $(\gamma \star f)'^2 \geq 0$ and the convolution preserves the sign.

To show (ii), we start from the law of motion of W_t implied by the FOC,

$$\mathrm{d}W_t = r(W_t - x_t)\,\mathrm{d}t = r\big(\gamma \star f(\theta_t) - f(\theta_t)\big)\,\mathrm{d}t = \left(\mu(\gamma \star f)'(\theta_t) + \frac{1}{2}(\gamma \star f)''(\theta_t)\right)\mathrm{d}t,$$

where the second equality holds by the policy function, and the third equality comes from (10).

A.11 *Relaxing the strategy set* $\mathcal{M}$

The strategy set $\mathcal{M}$ puts an exponential limit on the speed that the agent can lie. It is assumed for technical simplicity. Now, I proceed to remove it. Without it, the global incentives can be problematic if the agent lies exponentially at a high rate. In so doing, the agent secures high flow payoffs while the continuation payoff explodes to $-\infty$. Nevertheless, I show that the cost function in the main model is still approachable by constructing a sequence of contracts that have costs approaching $C(0, W_0)$.

Consider the optimal contract truncated at time T. Before the deadline T, execute the original optimal contract. At time T, the action is frozen forever at $x_T = W_T$, so that the continuation payoff of the agent is promised even after the deadline. Obviously, with a finite horizon, the agent's Ponzi-style global deviation fails, as all past deviations factor into W_T at the deadline.

I claim that this contract yields a cost $C^T(0, W_0)$ that approaches $C(0, W_0)$ as $T \to \infty$. At time T, the difference between $C^T(0, W_0)$ and $C(0, W_0)$ is

$$\begin{aligned}
&C^T(\theta_T, W_T) - C(\theta_T, W_T) \\
&\quad = \gamma \star \big(W - f(\theta_T)\big)^2 - \big(W - \gamma \star f(\theta_T)\big)^2 - \frac{\sigma^2}{r}\gamma \star \big((\gamma \star f)'^2\big)(\theta_T) \\
&\quad = \gamma \star f^2(\theta_T) - \big(\gamma \star f(\theta_T)\big)^2 - \frac{\sigma^2}{r}\gamma \star \big((\gamma \star f)'^2\big)(\theta_T) \\
&\quad \leq \gamma \star f^2(\theta_T) \\
&\quad \leq 2\alpha_0^2 + \alpha_0^2 r\left(\frac{e^{-2\alpha_1\theta_T}}{r + 2\alpha_1\big(\mu - \alpha_1\sigma^2\big)} + \frac{e^{2\alpha_1\theta_T}}{r - 2\alpha_1\big(\mu + \alpha_1\sigma^2\big)}\right),
\end{aligned}$$

where the second inequality follows from Assumption 1. Hence,

$$\begin{aligned}
C^T(0, W_0) - C(0, W_0) &= e^{-rT}\mathbb{E}\big[C^T(\theta_T, W_T) - C(\theta_T, W_T)\big] \\
&\leq \alpha_0^2 e^{-rT}\left(\frac{e^{2\alpha_1(-\mu T + \alpha_1\sigma^2 T)}}{r + 2\alpha_1\big(\mu - \alpha_1\sigma^2\big)} + \frac{e^{2\alpha_1(\mu T + \alpha_1\sigma^2 T)}}{r - 2\alpha_1\big(\mu + \alpha_1\sigma^2\big)}\right),
\end{aligned}$$

where the last expression vanishes as $T \to \infty$, by Assumption 1.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1986), "Optimal cartel equilibria with imperfect monitoring." *Journal of Economic Theory*, 39, 251–269. [1596]

Alonso, Ricardo and Niko Matouschek (2008), "Optimal delegation." *The Review of Economic Studies*, 75, 259–293. [1593]

Amador, Manuel and Kyle Bagwell (2013), "The theory of optimal delegation with an application to tariff caps." *Econometrica*, 81, 1541–1599. [1593]

Battaglini, Marco (2005), "Long-term contracting with Markovian consumers." *American Economic Review*, 95, 637–658. [1594]

Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman (2014), "Optimal allocation with costly verification." *American Economic Review*, 104, 3779–3813. [1593]

Bird, Daniel and Alexander Frug (2019), "Dynamic non-monetary incentives." *American Economic Journal: Microeconomics*, 11, 111–150. [1593]

Boleslavsky, Raphael and Tracy R. Lewis (2016), "Evolving influence: Mitigating extreme conflicts of interest in advisory relationships." *Games and Economic Behavior*, 98, 110–134. [1594]

Crawford, Vincent P. and Joel Sobel (1982), "Strategic information transmission." *Econometrica*, 50, 1431–1451. [1593, 1596]

de Clippel, Geoffroy, Kfir Eliaz, Daniel Fershtman, and Kareen Rozen (2021), "On selecting the right agent." *Theoretical Economics*, 16, 381–402. [1593]

DeMarzo, Peter M. and Yuliy Sannikov (2016), "Learning, termination, and payout policy in dynamic incentive contracts." *The Review of Economic Studies*, 84, 182–236. [1594, 1596]

Fernandes, Ana and Christopher Phelan (2000), "A recursive formulation for repeated agency with history dependence." *Journal of Economic Theory*, 91, 223–247. [1594, 1596]

Frankel, Alexander (2016), "Discounted quotas." *Journal of Economic Theory*, 166, 396–444. [1593, 1597]

Green, Jerry R. and Nancy L. Stokey (2007), "A two-person game of information transmission." *Journal of Economic Theory*, 135, 90–104. [1593]

Guo, Yingni and Johannes Hörner (2020), "Dynamic allocation without money." Working paper. [1593, 1596, 1602, 1605]

Holmstrom, Bengt (1977), *On Incentives and Control in Organizations.* Ph.D. thesis, Stanford University. [1593]

Jackson, Matthew O. and Hugo F. Sonnenschein (2007), "Overcoming incentive constraints by linking decisions." *Econometrica*, 75, 241–257. [1590, 1593, 1602]

Kapička, Marek (2013), "Efficient allocations in dynamic private information economies with persistent shocks: A first-order approach." *Review of Economic Studies*, 80, 1027–1054. [1594, 1596]

Krasikov, Ilia, Rohit Lamba, and Thomas Mettral (2020), "Implications of unequal discounting in dynamic contracting." Available at SSRN 3142932. [1609]

Lipnowski, Elliot and Joao Ramos (2020), "Repeated delegation." *Journal of Economic Theory*, 188, 105040. [1593]

Malenko, Andrey (2019), "Optimal dynamic capital budgeting." *The Review of Economic Studies*, 86, 1747–1778. [1594]

Margaria, Chiara and Alex Smolin (2018), "Dynamic communication with biased senders." *Games and Economic Behavior*, 110, 330–339. [1593]

Melumad, Nahum D. and Toshiyuki Shibano (1991), "Communication in settings with no transfers." *The RAND Journal of Economics*, 22, 173–198. [1593]

Pavan, Alessandro, Ilya Segal, and Juuso Toikka (2014), "Dynamic mechanism design: A Myersonian approach." *Econometrica*, 82, 601–653. [1596]

Renault, Jérôme, Eilon Solan, and Nicolas Vieille (2013), "Dynamic sender–receiver games." *Journal of Economic Theory*, 148, 502–534. [1593, 1602]

Sannikov, Yuliy (2008), "A continuous-time version of the principal-agent problem." *The Review of Economic Studies*, 75, 957–984. [1594]

Sobel, Joel (2013), "Giving and receiving advice." *Advances in Economics and Econometrics*, 1, 305–341. [1593]

Strausz, Roland (2006), "Deterministic versus stochastic mechanisms in principal–agent models." *Journal of Economic Theory*, 128, 306–314. [1600]

Thomas, Jonathan and Tim Worrall (1990), "Income fluctuation and asymmetric information: An example of a repeated principal-agent problem." *Journal of Economic Theory*, 51, 367–390. [1596]

Williams, Noah (2011), "Persistent private information." *Econometrica*, 79, 1233–1275. [1594, 1596]

Co-editor Simon Board handled this manuscript.

Manuscript received 19 January, 2021; final version accepted 6 December, 2021; available online 22 December, 2021.