

Dynamic Information Provision: Rewarding the Past and Guiding the Future

Ian Ball*

23 January 2023

Abstract

I study the optimal provision of information in a long-term relationship between a sender and a receiver. The sender observes a persistent, evolving state and commits to send signals over time to the receiver, who sequentially chooses public actions that affect the welfare of both players. I solve for the sender's optimal policy in closed form: the sender reports the value of the state with a delay that shrinks over time and eventually vanishes. Even when the receiver knows the current state, the sender retains leverage by threatening to conceal the future evolution of the state.

Keywords: dynamic information design; delayed reporting.

JEL Classification: D82, D83, D86.

*Department of Economics, MIT, ianball@mit.edu. I thank Dirk Bergemann for thoughtful and friendly guidance throughout the project. An earlier version of this paper appeared as the second chapter of my Ph.D. dissertation at Yale. For helpful discussions, I thank Alexandre Belloni, Ben Brooks, Eduardo Faingold, Mira Frick, Drew Fudenberg, Simone Galperti, Marina Halac, Johannes Hörner, Ryota Ijima, Navin Kartik, Emir Kamenica, Deniz Kattwinkel, Nicolas Lambert, Chiara Margaria, Stephen Morris, Giuseppe Moscarini, Larry Samuelson, Andy Skrzypacz, Philipp Strack, and Alex Wolitzky. I thank audiences at Yale, Allerton, Stony Brook, and the Stanford Summer Institute in Theoretical Economics (SITE) for useful feedback. Finally, I thank Yifan Dai for excellent research assistance.

1 Introduction

How can information be used in place of money as a reward to motivate behavior? In many relationships, controlling information provides powerful leverage. A leaker sharing protected information with a media outlet can use the promise of additional leaks to demand slanted coverage. In many organizations, from large firms to the US military, turf wars impede efficient information exchange (Herrera et al., 2017). In such organizations, where preferences are misaligned and transfers are often infeasible, refusing to share information can be an important bargaining chip.

In these relationships, between individuals or organizational units, one party controls information that is necessary to guide efficient decisions. When preferences are misaligned, this information can be withheld as a punishment for past behavior. In this paper, I illustrate how the dual role of information shapes the dynamics of information provision within a long-term relationship.

I consider a stylized model in continuous time with a sender (she) and a receiver (he). There is a payoff-relevant state, which follows a diffusion process. The sender observes the evolution of this state, and she sends signals over time to the receiver, who sequentially chooses public actions that affect the welfare of both players. The sender and receiver have partially aligned quadratic preferences. The receiver wants to match his action with the state, but the sender wants the receiver to shift his action above the state by a fixed bias.

The only instrument available to the sender is the control of information. The sender commits to a dynamic information policy, which assigns a signal distribution to each private history of past states, signals, and actions. Over time, as the sender observes the evolving state and the receiver's actions, she sends the signals prescribed by the information policy. The receiver observes the sender's signals and chooses actions sequentially. Since the receiver is forward-looking, he considers the effect of his actions both on his flow payoff and on the informativeness of future signals. The sender, anticipating the receiver's best response, chooses an information policy to maximize her expected

utility.

The fundamental tradeoff for the sender is between *precision* and *bias*. Providing the receiver with information about the current state has two effects. On the one hand, the receiver can more precisely tailor his current action to the state, making both players better off. On the other hand, since the state process is persistent, the receiver’s uncertainty about future states is reduced, making future information less valuable to him. Hence, the receiver is less willing to bias his actions.

The sender faces a complex, nested optimization problem. I reduce the sender’s problem in two steps.

First, I change the domain of optimization from information policies to *decision rules*, i.e., state-dependent distributions of actions over time. A decision rule is a best response to some information policy if and only if it is a best response to a canonical information policy—the associated direct, grim-trigger policy that makes direct action recommendations and cuts off all future information if any recommendation is violated. Decision rules that can be induced in this way are called *obedient*.

Second, I observe that the players’ payoffs and the obedience constraint can be expressed in terms of two statistics of a decision rule—the bias and variance. At a given time, the bias¹ is the difference between the action recommended to the receiver and the receiver’s bliss point (his expectation of the state). The variance is the posterior variance of the current state, given the history of action recommendations. The bias and variance are stochastic processes, but I show that deterministic bias and variance processes—termed paths—trace out the Pareto frontier.

Bayesian updating restricts which variance paths are feasible. For a variance path to be induced by some information policy, a necessary condition called Bayes plausibility is that the variance never increases by more than it would in the absence of new information. I show that this condition is also sufficient. Theorem 1 says that every Bayes-plausible posterior variance

¹This (action) bias is a statistic computed from a decision rule. The (preference) bias is a parameter of the sender’s utility function.

path can be induced by *delayed reporting*. At each time t , the sender's action recommendation reveals the state realization at some earlier time $\varphi(t)$.

The sender's problem is therefore reduced to optimizing over obedient, Bayes-plausible bias–variance paths. I use Lagrangian relaxation and dynamic programming to find the solution in closed form. The optimal bias and variance functions are stated in Theorem 2. There are two cases.

If the sender's bias is sufficiently small relative to the volatility of the state process, then the sender can induce her first-best decision rule. She keeps the receiver perfectly informed of the state but threatens to cut off all future information if the receiver ever deviates from the *sender's* optimal action.

If the sender's bias is sufficiently large relative to the volatility of the state process, then the sender cannot induce her first-best decision rule. Hence, the bias–precision tradeoff is in force. The optimal policy has two phases. First, in the *transition phase*, the sender gradually reduces the receiver's uncertainty about the current state while gradually narrowing the gap between the recommended action and the receiver's bliss point. At some finite time, the current state is fully revealed, and the *stationary phase* begins. Thereafter, the sender keeps the receiver perfectly informed of the current state. The sender demands that the receiver maintain a fixed bias between his action and the state. If the receiver deviates at any time, the sender cuts off all future information.

The rest of the paper is organized as follows. Section 1.1 reviews related literature. Section 2 presents the model. Section 3 solves a two-period example. The main analysis begins in Section 4, where I introduce delayed reporting and simplify the sender's problem. Section 5 describes the optimal information policy and analyzes comparative statics. Section 6 solves for the optimal policy with multidimensional states and actions. The components of the state are revealed sequentially, in order of increasing persistence. Section 7 is the conclusion. Measure-theoretic definitions are in Appendix A. Proofs are in Appendix B.

1.1 Related literature

My paper joins a growing literature on dynamic information design. The first papers (Renault et al., 2017; Ely, 2017) study the optimal provision of information (about an evolving state) to a myopic receiver who acts repeatedly.² Since the receiver is myopic, his action choices depend only on his beliefs about the current state, not the promise of future information.

The closest dynamic information design papers (Smolin, 2021; Ely and Szydlowski, 2020; Orlov et al., 2020) study the optimal provision of information over time to a forward-looking receiver who chooses when to stop.³ To stop optimally, the receiver considers his current belief about the state and also the future information he will receive if he waits. The sender can delay information transmission in order to entice the receiver to wait longer. In my paper, by contrast, the receiver chooses a rich action at each time. Therefore, the sender has a countervailing motive to reveal information so that the receiver can adjust his action to the current state.⁴ This precision motive drives gradual information revelation.⁵

The dynamics in my model are broadly similar to the backloading of rewards in dynamic principal-agent models with cash constraints (Lazear, 1981; Harris and Holmstrom, 1982; Thomas and Worrall, 1994).⁶ Withholding information, unlike money, directly entails inefficiency. Closest to my paper is the apprenticeship model in the independent work of Fudenberg and Rayo (2019, hereafter FR), which builds upon the framework of Garicano and Rayo (2017).

²The working paper version Ely (2015) finds that the solution is unchanged if the receiver is forward-looking. This is a special feature of the binary-action setting.

³Other dynamic information design papers explore different issues: costly communication (Honryo, 2018); private information held by the receiver (Au, 2015); and the effect of a decision deadline (Bizotto et al., 2021).

⁴In subsequent work by Kaya (2022), the receiver chooses effort each period. But there is no precision motive because the sender prefers more effort, no matter the state.

⁵Orlov et al. (2020) find an equilibrium with gradual information revelation in a model in which the sender has only within-period commitment. If the sender’s bias against stopping is sufficiently strong, then promising delayed information is not credible. In my model, gradualism arises with or without intertemporal commitment; see the discussion of commitment in Section 2.2.

⁶In a quite general (complete information) principal–agent setting, Ray (2002) shows that in all efficient “self-enforcing agreements,” the continuation value of the agent is backloaded.

In FR, the principal is endowed with a perfectly divisible unit of knowledge that can be costlessly transmitted to the agent, who is cash-constrained. The principal selects a contract specifying paths of knowledge, effort, and wages for the agent. At any time, the agent can walk away with his current knowledge stock. While immediate knowledge transmission is efficient, the principal's optimal contract transmits knowledge gradually in order to suppress the value of the agent's outside option.

In my model, the receiver's posterior variance and action bias, respectively, play similar roles to the principal's stock of untransmitted knowledge and the agent's excess effort in FR. But there are important differences. First, in my model the information provided by the sender grows stale over time (because the state evolves). So even if the receiver knows the current state, it is still *feasible* for the sender to demand bias from the receiver. Second, the players' payoffs in my model are not transferable. The curvature in the Pareto frontier makes it *optimal* for the sender to demand bias after fully revealing the state. The stationary phase, featuring perfect information transmission and biased actions, is new to my model.

2 Model

Time is continuous and the horizon is infinite. There are two players: a sender (she) and a receiver (he). At each time $t \in [0, \infty)$, the state $\theta_t \in \mathbf{R}$ is realized and the receiver chooses an action $a_t \in \mathbf{R}$. Flow utilities for the sender and receiver are given by

$$u_S(a_t, \theta_t) = -(a_t - \theta_t - \beta)^2, \quad u_R(a_t, \theta_t) = -(a_t - \theta_t)^2.$$

The receiver wants to match his action with the state, but the sender wants the receiver to shift his action above the state by a bias β . Without loss, assume $\beta > 0$. Both players discount future flow utilities at exponential rate r .

The initial state θ_0 is normally distributed with mean μ_0 and variance σ_0^2 .

The state then follows the stochastic differential equation

$$d\theta_t = \kappa\theta_t dt + \sigma dZ_t, \quad (1)$$

where $\{Z_t\}_{t \geq 0}$ is a standard Brownian motion, independent of the initial state θ_0 .⁷ The volatility parameter σ is strictly positive. Assume $2\kappa < r$. Hence, a constant action yields finite expected utility for both players. The process can be explosive ($\kappa > 0$), mean-reverting ($\kappa < 0$), or a Brownian motion ($\kappa = 0$).

The state distribution is common knowledge, but the state realizations are observed only by the sender. The sender also observes the receiver's actions. Initially, the sender commits to a dynamic *information policy* S , which consists of a signal realization space \mathbf{S} together with a signal distribution for each history of past states, signals, and actions. Given this policy, the receiver faces a dynamic decision problem. At each time t , having observed the signals sent up to time t , the receiver chooses an action.⁸ The receiver learns about the state only from the sender's signals. In particular, the receiver does not observe past flow payoffs.

2.1 Sender's problem

The sender's information policy and the receiver's action choices jointly determine a stochastic process $A = \{A_t\}_{t \geq 0}$, which I called a *decision rule* as in [Bergemann and Morris \(2016\)](#). A decision rule A yields utilities

$$u_i(A) = \mathbf{E} \left[\int_0^\infty r e^{-rt} u_i(A_t, \theta_t) \right], \quad i = S, R.$$

A decision rule A is *compatible* with an information policy S if at each time t , the sender's signals up to time t (which depend on the receiver's previous actions A_s for $s < t$) provide sufficient information for the receiver to select

⁷The process is normalized to have zero drift if $\theta_t = 0$. The results do not change if (1) is replaced with $d\theta_t = (\mu_t + \kappa\theta_t) dt + \sigma dZ_t$, for any deterministic time-dependent drift μ_t .

⁸Restricting to pure strategies is without loss. Any randomization by the receiver can be incorporated into the sender's signal.

action A_t .⁹ A decision rule A is a best response to an information policy S if A maximizes the receiver’s utility over all decision rules compatible with S . The sender maximizes her utility $u_S(A)$ over all pairs (S, A) with the property that A is a best response to S .

2.2 Discussion of assumptions

Monitoring The sender perfectly observes the receiver’s actions. Therefore, the sender can condition future information on past actions. Without monitoring, the receiver would act myopically, always matching his action with his expectation of the state. Therefore, the sender would perfectly reveal the state at all times.

Commitment The sender is assumed to have full dynamic commitment power. That is, the sender can commit to action-contingent signals within each “period” and across “periods.”¹⁰ One-period commitment is crucial, as in static Bayesian persuasion (Kamenica and Gentzkow, 2011). Multi-period commitment is not necessary in the main case of interest.¹¹ Under the optimal full-commitment policy, the receiver is never promised more than his reservation utility. Therefore, the sender can achieve the same decision rule with an alternative policy that uses only one-period punishments. If the receiver deviates, the sender provides no information next period. Thereafter, the sender reverts to the policy that is optimal given the receiver’s current beliefs. This reversion does not increase the receiver’s continuation payoff, so the receiver’s payoff from deviating is unchanged.

Unobserved payoffs The receiver does not observe his own flow payoffs. With quadratic utility, flow payoffs perfectly reveal the trajectory of the state, so this assumption is necessary to preserve the receiver’s uncertainty. In the

⁹This circularity makes the formal definition delicate; see Appendix A.

¹⁰This distinction can be formalized in a discrete-time approximation of the continuous-time model.

¹¹That is, if $\beta \geq \sigma/\sqrt{r - 2\kappa}$. Otherwise, the receiver is promised more than his reservation utility, and multi-period commitment is needed.

motivating applications, it seems reasonable that the receiver could experience the cost of his own state uncertainty without learning the state realization. In any case, the forces in the model should apply as long as the sender retains some informational advantage over the receiver.

3 Two-period example

This section illustrates the sender’s bias–precision tradeoff in an example with two periods, $t = 1, 2$. Flow payoffs are as in the main model. Each player i maximizes the discounted sum of flow payoffs $u_{i,1} + \delta u_{i,2}$. The state process is given by

$$\theta_1 \sim N(0, \sigma_1^2), \quad \theta_2 = \rho\theta_1 + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2),$$

where ε is independent of θ_1 .

This example can be analyzed backwards. In the second period, the receiver acts myopically since this is the last period of the game. Full disclosure maximizes flow payoffs for both players, so it is optimal for the sender to fully disclose the state on-path, i.e., if the receiver obeys the sender’s first-period recommendation. If the receiver disobeys the first-period recommendation, it is optimal to impose the maximal punishment by providing no information.

Now consider the first period. If the sender’s signal induces posterior variance v_1 , the sender can demand that the receiver bias his action away from his posterior mean by b_1 , provided that

$$b_1^2 \leq \delta(\rho^2 v_1 + \sigma^2). \tag{2}$$

This inequality captures the bias–precision tradeoff. The right side—the discounted residual variance of θ_2 without additional information—is the receiver’s discounted loss from not learning the state in the second period. When this constraint binds, the sender must pay a price of $1/(\delta\rho^2)$ in higher variance per unit of squared bias.

The sender chooses b_1 and v_1 to maximize her total (on-path) discounted payoff $-(b_1 - \beta)^2 - v_1 - \delta\beta^2$, subject to (2). There are three cases.

- (a) $\beta \leq \sqrt{\delta}\sigma$. Here, (2) is not binding, so the optimum is $\hat{b}_1 = \beta$ and $\hat{v}_1 = 0$.
- (b) $\rho = 0$. Here, v_1 does not appear in (2). Revealing θ_1 does not affect the receiver's uncertainty about θ_2 , so there is no tradeoff between bias and precision. The optimum is $\hat{v}_1 = 0$ and $\hat{b}_1 = \min\{\beta, \sqrt{\delta}\sigma\}$.
- (c) $\beta > \sqrt{\delta}\sigma$ and $\rho \neq 0$. This is the main case of interest. Here, (2) must hold with equality; otherwise, the sender would strictly prefer to reveal more in the first period. The first-order condition for the optimal bias equates the direct marginal benefit $2(\beta - b_1)$ of increasing b_1 with the marginal cost $2b_1/(\delta\rho^2)$ from the required increase in variance.¹²

I now describe in detail the comparative statics for the optimal bias \hat{b}_1 and variance \hat{v}_1 since the same results hold for the continuous-time solution during the transition phase. As the preference bias β increases, the sender withholds more information (higher \hat{v}_1) in order to induce greater action bias (higher \hat{b}_1). As the volatility σ increases, the bias \hat{b}_1 is unchanged (since σ does not enter the bias first-order condition), but the variance \hat{v}_1 decreases since less variance is needed to induce the same level of bias. As δ and ρ^2 increase, the receiver's loss from not learning θ_2 is more sensitive to the first-period variance v_1 . That is, the price $1/(\delta\rho^2)$ of bias (in units of variance) decreases. Hence, \hat{b}_1 increases. The effect on total "spending" \hat{v}_1 is ambiguous.

4 Reducing the dimension of the problem

In this section, I change the domain of the sender's optimization problem from the space of all information policies to the simpler space of obedient decision

¹²The solution is given by

$$\hat{b}_1 = \beta \frac{\delta\rho^2}{1 + \delta\rho^2}, \quad \hat{v}_1 = \beta^2 \frac{\delta\rho^2}{(1 + \delta\rho^2)^2} - \frac{\sigma^2}{\rho^2}.$$

As a function of v_1 , the sender's payoff is concave, so it is optimal to induce a constant variance. Here I assume that $\sigma_1^2 \geq \hat{v}_1$, so there exists a signal structure inducing this variance \hat{v}_1 .

rules with deterministic bias and variance.

4.1 Obedient decision rules

For a given information policy, the sender’s utility depends only on the decision rule induced by the receiver’s best response. Therefore, it suffices to optimize over every decision rule that can be induced as a best response to some information policy. I now characterize such decision rules.

Given a decision rule A , let \mathcal{F}_t^A denote the σ -algebra generated by A_s for $s \leq t$. This σ -algebra represents the minimal information that the receiver must have by time t in order to follow A . The receiver gets exactly this minimal information if the sender makes direct action recommendations.

Decision rule A is *obedient* if for each $t \geq 0$,¹³

$$-\mathbf{E} \left[\int_t^\infty r e^{-r(s-t)} (A_s - \theta_s)^2 ds \middle| \mathcal{F}_t^A \right] \geq -\frac{\sigma^2 + r \operatorname{var}(\theta_t | \mathcal{F}_t^A)}{r - 2\kappa}. \quad (3)$$

Inequality (3) is the *time- t obedience constraint*. Conditional on the history of actions up to time t , the receiver’s expected continuation value from following the decision rule is at least as large as his reservation utility—his expected continuation value from deviating at time t , forfeiting all future information, and choosing $A_s = \mathbf{E}[\theta_s | \mathcal{F}_t^A] = e^{\kappa(s-t)} \mathbf{E}[\theta_t | \mathcal{F}_t^A]$ for $s \geq t$.

Proposition 1 (Obedience)

Decision rule A is a best response to some information policy if and only if A is obedient.

The necessity of obedience is clear. If obedience is violated, then at some time t the receiver can profit by deviating to myopic play on a nontrivial subset of action histories. For sufficiency, I show that any obedient decision rule A is a best response to a canonical information policy—the direct, grim-trigger information policy that sends signals $S_t = A_t$, provided that the receiver has followed all past recommendations. If the receiver ever deviates, this policy

¹³All inequalities involving conditional expectations are interpreted almost surely.

sends uninformative signals forever after. The time- t obedience constraint ensures that it is not profitable for the receiver to start deviating at time t .

The sender's problem is now reduced to maximizing over all obedient decision rules. I next show that it suffices to maximize over simpler summary statistics of these decisions rules.

4.2 Bias and variance

For any decision rule A , define the induced bias and variance processes by

$$B_t = A_t - \mathbf{E}[\theta_t | \mathcal{F}_t^A] \quad \text{and} \quad V_t = \text{var}(\theta_t | \mathcal{F}_t^A).$$

The *bias* B_t is the gap at time t between the action and the receiver's expectation of the state, given the action recommendations up until time t . The *variance* V_t is the posterior variance of the state θ_t , given the action recommendations up until time t . The players' payoffs and the obedience condition (3) can be expressed in terms of the bias and variance processes only.¹⁴ In general, B_t and V_t are random, but I show that the Pareto frontier is traced out by decision rules with deterministic bias and variance.

Proposition 2 (Deterministic bias and variance)

For each obedient decision rule A , there exists an obedient decision rule A' with deterministic bias and variance such that both players weakly prefer A' to A .

Here is a sketch of the proof. Consider an obedient decision rule A with bias and variance processes $\{B_t\}_{t \geq 0}$ and $\{V_t\}_{t \geq 0}$. Let $b(t) = \mathbf{E}[B_t]$ and $v(t) = \mathbf{E}[V_t]$ for each t . I will construct a new decision rule A' with deterministic bias b and deterministic variance v . Both players weakly prefer A' to A because their flow payoffs are linear in variance and strictly concave in bias. The receiver's reservation utility is linear in variance, so A' is obedient. To complete the proof of Proposition 2, I need to construct a signal structure that induces the

¹⁴In particular, the expected state $\mathbf{E}[\theta_t | \mathcal{F}_t^A]$ does not appear by itself. This holds because the sender's bliss point is $\theta_t + \beta$, not $\alpha\theta_t + \beta$ for some $\alpha \neq 1$.

desired variance function v . The next section introduces a convenient general construction called delayed reporting.

4.3 Bayes plausibility and delayed reporting

First I introduce a necessary condition for a variance process to be consistent with Bayesian updating. Suppose that at time t , the posterior variance of θ_t is V_t . If no additional information is provided, then at time $t + h$, the posterior variance of θ_{t+h} is $\eta(V_t, h)$, where

$$\eta(v, h) = \begin{cases} e^{2\kappa h}v + (e^{2\kappa h} - 1)\frac{\sigma^2}{2\kappa} & \text{if } \kappa \neq 0, \\ v + \sigma^2 h & \text{if } \kappa = 0. \end{cases} \quad (4)$$

A function $v: [0, \infty) \rightarrow [0, \infty)$ is *Bayes-plausible* if

- (i) $v(0) \leq \sigma_0^2$;
- (ii) $v(s) \leq \eta(v(t), s - t)$ for all $s > t \geq 0$.

By the law of total variance, a necessary condition for a variance process $\{V_t\}_{t \geq 0}$ to be induced by some decision rule is that the function $v(t) = \mathbf{E} V_t$ is Bayes-plausible. Part **i** is the *initial variance constraint*: the sender's initial disclosure cannot increase the expected variance. Part **ii** is the *no-disclosure upper bound*: the receiver's expected variance never increases faster than it would if the sender provided no information. For deterministic variance processes, which I call *paths*, Bayes plausibility is also a sufficient condition, as I now show.

To prove sufficiency, I introduce *delayed reporting*. At each time t , the sender reports the exact realization of the state at a previous time. Initially, there is no previous state to report, so I define a fictitious history before time 0 to be used as a randomization device. Let $Y = \{Y_t\}_{t \geq 0}$ be an independent standard Brownian motion. For $t < 0$, let $\theta_t = \theta_0 + Y_{-t}$.

Delayed reporting is parameterized by a *reporting function*, defined to be a weakly increasing function $\varphi: [0, \infty) \rightarrow [-\infty, \infty)$ satisfying $\varphi(t) \leq t$ for each

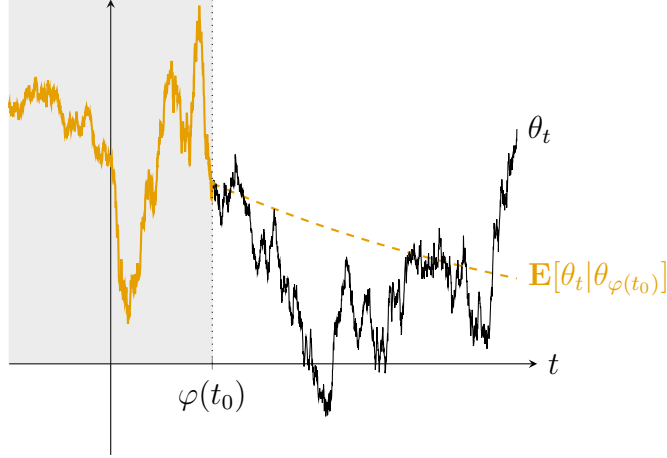


Figure 1. Delayed report at time t_0 , and the receiver's updated expectations

t . At each time t , the sender reports $\theta_{\varphi(t)}$. Let $\varphi(t) = -\infty$ if the sender has provided no information by time t . Figure 1 shows a snapshot of the receiver's information under delayed reporting. At time t_0 , the receiver's beliefs about θ_t for $t \geq \varphi(t_0)$ depend only on the value of $\theta_{\varphi(t_0)}$. The dashed curve shows the conditional expectation function $\mathbf{E}[\theta_t | \theta_{\varphi(t_0)}] = e^{\kappa(t-t_0)}\theta_{\varphi(t_0)}$.

A reporting function φ induces a variance path v if, for all $t \geq 0$,

$$\text{var}(\theta_t | \theta_{\varphi(t)}) = v(t),$$

with the convention that $\text{var}(\theta_t | \theta_{-\infty}) = \text{var}(\theta_t)$.

Theorem 1 (Delayed reporting)

Each Bayes-plausible variance path is induced by some reporting function.

If a variance path v is induced by a reporting function φ , then for any bias path b , the following decision rule has deterministic bias b and variance v :

$$A_t = \mathbf{E}[\theta_t | \theta_{\varphi(t)}] + b(t). \quad (5)$$

Here, A_t is random because the conditional expectation is random.

Delayed reporting simplifies the receiver's belief-updating process. Instead of aggregating the state information contained in the entire signal history, the

receiver forms his belief at each time t from the time- t signal realization alone. Moreover, once φ crosses 0, the sender’s signal is a deterministic function of the state history—the sender does not need to commit to randomization. One example of delayed reporting, with a different state process, is the optimal email notification policy in [Ely \(2017\)](#).¹⁵

5 Optimal information policy

I now solve for the sender’s optimal bias and variance functions. The associated decision rule with delayed reporting, given in (5), is a best response to the direct, grim-trigger information policy.

The sender’s problem is to choose bias and variance functions b and v to solve

$$\begin{aligned} \text{maximize} \quad & - \int_0^\infty r e^{-rt} [(b(t) - \beta)^2 + v(t)] dt \\ \text{subject to} \quad & - \int_t^\infty r e^{-r(s-t)} [b^2(s) + v(s)] ds \geq - \frac{\sigma^2 + rv(t)}{r - 2\kappa}, \quad t \geq 0 \\ & 0 \leq v(s) \leq \eta(v(t), s - t), \quad s > t \geq 0 \\ & 0 \leq v(0) \leq \sigma_0^2. \end{aligned}$$

The first constraint is obedience. The last two constraints impose Bayes plausibility.

Remark (Pareto frontier). Solving the sender’s problem for arbitrary bias β immediately yields the entire Pareto frontier. Fix π in $[0, 1)$. Maximizing the social welfare function $\pi u_S + (1 - \pi)u_R$ is equivalent to maximizing the utility of a different sender with bias $\pi\beta$. This follows from the decomposition

$$\pi[(b - \beta)^2 + v] + (1 - \pi)[b^2 + v] = (b - \pi\beta)^2 + v + \pi(1 - \pi)\beta^2.$$

¹⁵In [Ely \(2017\)](#), the state is binary, indicating whether an unread email is waiting. The optimal policy “beeps” after an email arrives, but with a delay of length Δ . Thus, the reporting function is $\varphi(t) = \max\{t - \Delta, 0\}$. In this setting, the initial state is known (no emails are waiting), so revealing θ_0 provides no information. Thus, there is no need for a fictitious history.

5.1 Obedience is binding

Under the optimal policy, the obedience constraint must be active whenever the variance is strictly positive. Otherwise, the sender could strictly improve her payoff by reducing the variance v over a small time interval where the constraint is slack. This perturbation relaxes the earlier obedience constraints and leaves the later obedience constraints unchanged.¹⁶

The variance function, when positive (and differentiable), satisfies the differential equation

$$(r - 2\kappa)b^2(t) = 2\kappa v(t) + \sigma^2 - v'(t). \quad (6)$$

This is the continuous-time analogue of the binding obedience constraint (2) in the two-period example. On the right side, the first two terms capture the evolution of the receiver's posterior variance if he receives no additional information. The receiver is willing to bias his action by $b(t)$ only if the sender provides information that reduces his posterior variance (relative to its exogenous evolution) at rate $(r - 2\kappa)b^2(t)$. Hence, $r - 2\kappa$ is the price (in terms of variance) of inducing squared bias. This is the differential analogue of the price $\delta^{-1}\rho^{-2}$ in the two-period example, with $\delta = e^{-r}$ and $\rho = e^\kappa$. The higher the price, the more information the sender must initially withhold to induce a desired bias path. If the receiver is more impatient (higher r) or the process is less persistent (lower κ), the receiver demands a greater reduction in variance to select a given level of bias.

If the variance first hits zero at some time t , then by continuity, the time- t obedience constraint must be active. Starting at time t , the relaxed continuation problem, with subsequent obedience constraints dropped, is to choose $b(s)$ and $v(s)$ for $s > t$ to maximize the sender's continuation payoff subject to

$$-\int_t^\infty r e^{-r(s-t)} [b^2(s) + v(s)] ds = -\frac{\sigma^2}{r - 2\kappa}.$$

¹⁶In the formal proof, the optimal policy is derived directly, without first showing that obedience is active. The informal arguments in this section are intended to build intuition.

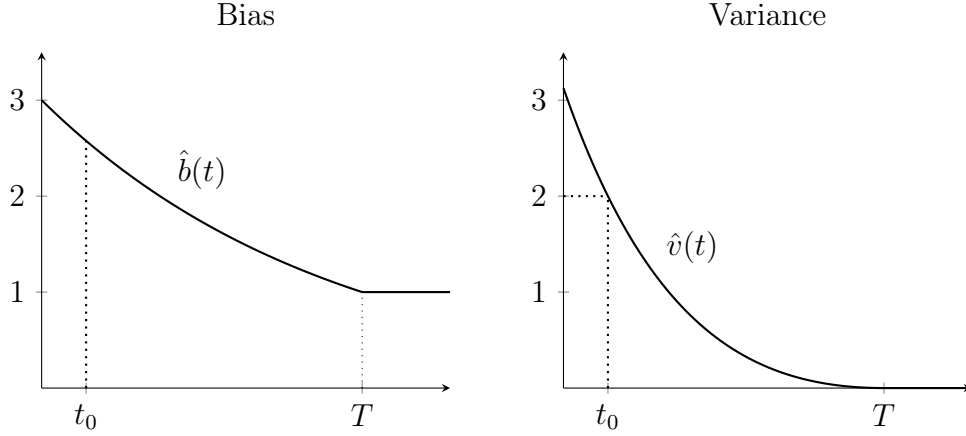


Figure 2. Optimal bias and variance functions

The sender's flow payoff is decreasing in v and concave in b , so her continuation payoff is maximized by the stationary policy with $v(s) = 0$ and $b(s) = \min\{\beta, \sigma/\sqrt{r - 2\kappa}\}$, for $s > t$. Since $v(t) = 0$, this policy satisfies the obedience constraint at time t and hence at all subsequent times. Therefore, this policy is optimal in the original problem. It will form the stationary phase of the solution.

5.2 Optimal policy

The main result characterizes the optimal bias and variance paths.

Theorem 2 (Optimal bias and variance)

The optimal bias \hat{b} and variance \hat{v} are unique and given as follows.¹⁷

- I. If $\beta \leq \sigma/\sqrt{r - 2\kappa}$, then $\hat{b}(t) = \beta$ and $\hat{v}(t) = 0$.

¹⁷Here x_+ denotes the positive part $\max\{x, 0\}$ of a real number x . In some expressions, the persistence parameter κ appears in the denominator. The results still hold for $\kappa = 0$ if each expression is replaced with its limit as κ tends to 0.

II. If $\beta > \sigma/\sqrt{r-2\kappa}$, then

$$\begin{aligned}\hat{b}(t) &= \frac{\sigma}{\sqrt{r-2\kappa}} e^{(r-2\kappa)(T-t)_+}, \\ \hat{v}(t) &= -\frac{\sigma^2}{2\kappa} + \frac{\sigma^2(r-2\kappa)}{2(r-\kappa)} \left(\kappa^{-1} e^{-2\kappa(T-t)_+} + (r-2\kappa)^{-1} e^{2(r-2\kappa)(T-t)_+} \right),\end{aligned}$$

where the full-disclosure time T takes the unique value for which the inequalities $\hat{b}(0) \leq \beta$ and $\hat{v}(0) \leq \sigma_0^2$ both hold, at least one with equality.

There are two cases. If $\beta \leq \sigma/\sqrt{r-2\kappa}$, then the sender can induce her first-best decision rule $A_t = \theta_t + \beta$. At each time, the receiver's continuation value $-\beta^2$ from this rule is weakly greater than his reservation utility $-\sigma^2/(r-2\kappa)$ from forfeiting all information and acting myopically.

Hereafter, I focus on the interesting case in which $\beta > \sigma/\sqrt{r-2\kappa}$. The sender's first-best decision rule is not obedient, so the bias-precision tradeoff is in force. Figure 2 plots the optimal policy for a fixed set of parameters.¹⁸ The optimal policy has two phases—a *transition phase* until the full-disclosure time T , and a *stationary phase* after time T . Figure 3 plots the reporting function that implements this optimal policy. As time t approaches T , the delay $t - \varphi(t)$ tends to 0. Then $\varphi(t) = t$ for $t \geq T$. There is a kink when φ crosses 0 because the fictitious history has a different distribution than the true state process.

To derive the optimal policy, I first drop the initial variance constraint. I solve this relaxed problem by attaching a suitable multiplier to each time- t obedience constraint. Then I integrate over these constraints to form the Lagrangian. In the solution of the relaxed problem, the optimal bias takes the form

$$\max \left\{ \beta e^{-(r-2\kappa)t}, \frac{\sigma}{\sqrt{r-2\kappa}} \right\}. \quad (7)$$

In (7), the bias equals β at time 0. At each time t during the transition phase, the flow benefit of increasing the bias equals the shadow cost of reducing the variance. This shadow cost is initially zero. Over time, as the variance decreases, this shadow cost increases, and the bias moves away from β . The

¹⁸Here, $\beta = 3$, $r = 3$, $\kappa = -0.5$, $\sigma = 2$, and σ_0^2 is large enough that the initial variance constraint does not bind.

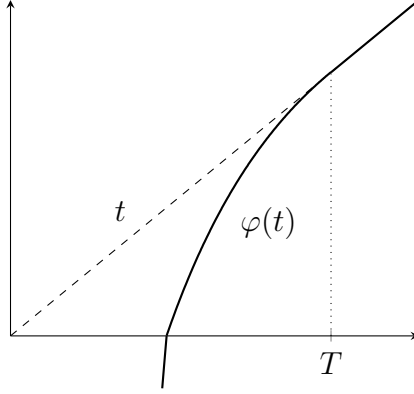


Figure 3. Optimal reporting function

stationary phase begins at time T when the variance hits zero and the bias hits $\sigma/\sqrt{r-2\kappa}$.

If σ_0^2 is larger than the initial variance in the relaxed solution, then the relaxed solution solves the original problem. In this case, the sender makes an initial disclosure to reduce the receiver's variance below σ_0^2 . Conversely, if σ_0^2 is smaller than the initial variance in the relaxed solution, then the relaxed solution is not feasible. Under the relaxed solution, the variance drifts downward and hits σ_0^2 at some time t_0 . The solution of the original problem is the continuation policy from time t_0 onwards. In this case, there is no initial disclosure of information, and the sender never induces her preferred action bias β . In Figure 2, the time t_0 is indicated for initial variance $\sigma_0^2 = 2$.

Corollary 1 (Optimum with deterministic state)

If $\sigma = 0$, the optimal bias \hat{b} and variance \hat{v} are unique and given by

$$\hat{b}(t) = b_0 e^{-(r-2\kappa)t}, \quad \hat{v}(t) = \frac{r-2\kappa}{2(r-\kappa)} b_0^2 e^{-2(r-2\kappa)t},$$

where b_0 is the minimum of β and $\sqrt{2\sigma_0^2(r-\kappa)/(r-2\kappa)}$.

This solution is the limit of the main solution as σ tends to 0. With $\sigma = 0$, the state evolves deterministically. If the receiver learns the current state, he can perfectly predict its future trajectory, so he will take his first-best decisions forever after. Since the receiver's marginal utility from action bias vanishes

when the action bias is zero, it is optimal for the sender to sacrifice some precision in order to induce a slight bias.¹⁹

5.3 Comparative statics

Say that a function *increases* in response to a parameter change if it strictly increases at some point and does not decrease at any point.

Proposition 3 (Comparative statics)

Suppose that $\beta > \sigma/\sqrt{r-2\kappa}$ and the initial variance constraint is not active.

1. The optimal bias function \hat{b} is increasing in β and σ and decreasing in $r - 2\kappa$.
2. The optimal variance function \hat{v} is increasing in β and decreasing in σ .

In the stationary phase, $\hat{b}(t) = \sigma/\sqrt{r-2\kappa}$ and $\hat{v}(t) = 0$, so the bias is increasing in σ and decreasing in $r - 2\kappa$. In the transition phase, the comparative statics for the bias and variance paths are the same as those for the first-period bias and variance in the two-period example (with discount factor $\delta = e^{-r}$ and discrete-time persistence $\rho = e^\kappa$). As the preference bias β increases, the sender withholds more information to induce greater action bias. As the volatility σ increases, the *sensitivity* of the receiver’s reservation utility to the variance does not change, but the *level* of the receiver’s reservation utility decreases, relaxing the obedience constraint. Thus, the optimal bias \hat{b} over the transition phase does not change, but the variance \hat{v} decreases.²⁰ Recall from (6) that $r - 2\kappa$ is the price (in variance) of inducing bias. As this price increases, the bias decreases. The effect of r and κ on the variance, however,

¹⁹This contrasts with Fudenberg and Rayo’s (2019) transferable utility model. With $\sigma = 0$ and $\kappa = 0$, a variance path is Bayes-plausible if and only if it is weakly decreasing. This is identical to Fudenberg and Rayo’s (2019) constraint on the path of untransmitted knowledge. Under their optimal policy, however, this untransmitted knowledge hits zero in finite time. Thereafter, the agent keeps the whole surplus and chooses effort efficiently.

²⁰The shadow cost of reducing variance remains the same because of the linearity assumptions: payoffs are linear in the variance, and the state follows a linear stochastic differential equation. If the initial variance constraint binds, then \hat{b} does increase in σ over the transition phase.

is ambiguous. As r increases and κ decreases, more variance is required to induce a given level of bias, but the optimal bias path is lower. Which effect dominates depends on other parameter values and can change over time.

6 Multidimensional states and actions

The main model studies *how much* information the sender provides over time. With a multidimensional state, I now study *which* information the sender provides at each time. Returning to one of the motivating examples, suppose that the main unit in an organization has private information about two different evolving situations that are relevant to another unit. In order to reward the other unit for making concessions, the solution below suggests that it is optimal to hold out information about the less mean-reverting situation for longer, revealing it only after fully disclosing the more mean-reverting situation.

Suppose that the state is n -dimensional, denoted $\theta_t = (\theta_{t,1}, \dots, \theta_{t,n}) \in \mathbf{R}^n$. The components of the initial state θ_0 are independent. Each component $\theta_{0,i}$ has a normal distribution $N(\mu_i, \sigma_{0,i}^2)$. Thereafter, the components evolve independently. Each component $\theta_{t,i}$ follows the linear stochastic differential equation

$$d\theta_{t,i} = \kappa_i \theta_{t,i} dt + \sigma_i dZ_{t,i},$$

where $Z_i = \{Z_{t,i}\}_{t \geq 0}$ is a standard Brownian motion, and Z_1, \dots, Z_n and θ_0 are mutually independent. For each i , assume $\sigma_i > 0$ and $2\kappa_i < r$. Order the components by increasing persistence, so $\kappa_1 \leq \dots \leq \kappa_n$.

At each time t , the receiver chooses an action $a_t = (a_{t,1}, \dots, a_{t,n}) \in \mathbf{R}^n$. The flow utilities for the sender and receiver are given by

$$u_S(a_t, \theta_t) = - \sum_{i=1}^n (a_{t,i} - \theta_{t,i} - \beta_i)^2, \quad u_R(a_t, \theta_t) = - \sum_{i=1}^n (a_{t,i} - \theta_{t,i})^2.$$

The sender's preference bias is a vector $\beta = (\beta_1, \dots, \beta_n) \in \mathbf{R}^n$.

With minor modifications to the argument in the main model, it can be shown that the optimal policy induces deterministic bias and variance func-

tions, which are now vector-valued. For each i , let

$$b_i(t) = A_{t,i} - \mathbf{E}[\theta_{t,i} | \mathcal{F}_t^A] \quad \text{and} \quad v_i(t) = \text{var}(\theta_{t,i} | \mathcal{F}_t^A).$$

Write $b(t) = (b_1(t), \dots, b_n(t))$ and $v(t) = (v_1(t), \dots, v_n(t))$. Define a separate variance-updating function η_i as in (4) for each component i .

The sender chooses vector-valued bias and variance functions b and v to solve

$$\begin{aligned} \text{maximize} \quad & - \int_0^\infty r e^{-rt} \sum_{i=1}^n [(b_i(t) - \beta_i)^2 + v_i(t)] dt \\ \text{subject to} \quad & - \int_t^\infty r e^{-r(s-t)} \sum_{i=1}^n [b_i^2(s) + v_i(s)] ds \geq - \sum_{i=1}^n \frac{\sigma_i^2 + r v_i(t)}{r - 2\kappa_i}, \quad t \geq 0 \\ & 0 \leq v_i(s) \leq \eta_i(v_i(t), s - t), \quad i = 1, \dots, n, \quad s > t \geq 0 \\ & 0 \leq v_i(0) \leq \sigma_{0,i}^2, \quad i = 1, \dots, n. \end{aligned}$$

The time- t obedience constraint is the sum of the single-dimensional time- t obedience constraints for each component. It is feasible to separately choose the single-dimensional optimal policy for each component, but this is generally suboptimal.

As in the main model, the obedience constraint must be active at each time t . The variance vector (when differentiable) satisfies the differential equation

$$\|b(t)\|^2 = \sum_{i=1}^n \frac{2\kappa_i v_i(t) + \sigma_i^2 - v_i'(t)}{r - 2\kappa_i}. \quad (8)$$

This is the multidimensional analogue of (6). Only the magnitude of the bias vector appears in the obedience constraint, so it is optimal to always choose $b(t)$ parallel to β . With a single-dimensional state, bias has price $r - 2\kappa$ in units of variance, which is the only currency. With n state components, there are n currencies. The price of bias is $r - 2\kappa_i$ in units of v_i . The volatility parameters do not enter the price because they do not affect the *sensitivity* of the reservation utility to the current variance. Since the variances v_1, \dots, v_n

have the same effect on the sender's flow payoff, the sender prefers to pay in the currency with the lowest price, i.e., the highest κ_i . If the initial variance $\sigma_{0,n}^2$ is sufficiently large, the sender compensates the agent with information about only the most persistent component n . All other components are revealed immediately. If $\sigma_{0,n}^2$ is too small, then the sender must withhold information about other components as well. The components are revealed sequentially, in order of increasing persistence, so that the largest bill is paid at the lowest price.

To state the theorem, let $\hat{\sigma}_j^2 = \sum_{i=1}^j \sigma_i^2 / (r - 2\kappa_i)$ for $j = 1, \dots, n$. Set $\hat{\sigma} = \hat{\sigma}_n$. Observe from (8) that $\hat{\sigma}$ is the magnitude of bias that the sender can induce while keeping the receiver perfectly informed.

Theorem 3 (Optimum with multidimensional state)

If the state is n -dimensional, then the following vector-valued bias and variance paths, \hat{b} and \hat{v} , are optimal. The optimum is unique if $\kappa_1 < \dots < \kappa_n$.

I. *If $\|\beta\| \leq \hat{\sigma}$, then $\hat{b}(t) = \beta$ and $\hat{v}(t) = 0$.*

II. *If $\|\beta\| > \hat{\sigma}$, then for some uniquely determined critical component i_0 and full-disclosure times $0 = t_1 = \dots = t_{i_0-1} < t_{i_0} \dots < t_n$,*²¹

$$\hat{b}(t) = \hat{\sigma} \exp\left(\sum_{i=i_0}^n (r - 2\kappa_i)(t_i - t \vee t_{i-1})_+\right) \frac{\beta}{\|\beta\|}.$$

For $i < i_0$, we have $\hat{v}_i(t) = 0$. For $i \geq i_0$, the variance \hat{v}_i is defined

²¹If $\sigma_{0,i}^2 = 0$ for all i , then $(t_1, \dots, t_n) = 0$, so technically $i_0 = n + 1$. Outside of this edge case, we have $i_0 \leq n$ and the full-disclosure times are pinned down by the following conditions:

- (i) $\|\hat{b}(0)\| \leq \|\beta\|$, $\hat{v}_{i_0}(0) \leq \sigma_{0,i_0}^2$, and $\hat{v}_i(0) = \sigma_{0,i}^2$ for $i > i_0$;
- (ii) $\|\hat{b}(0)\| = \|\beta\|$ or $(i_0, \hat{v}_1(0)) = (1, \sigma_{0,1}^2)$.

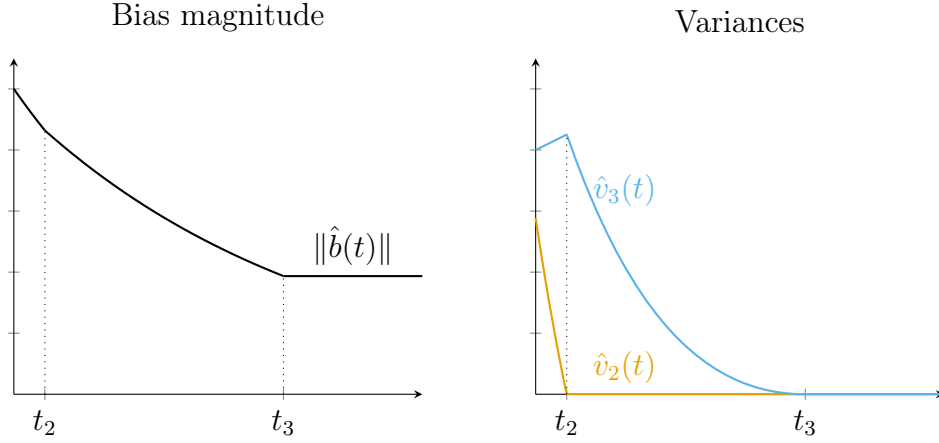


Figure 4. Optimal information policy with three-dimensional state

piecewise. For $t \geq t_{i-1}$,

$$\begin{aligned} \hat{v}_i(t) = & \frac{\hat{\sigma}_i^2(r - 2\kappa_i)}{2\kappa_i} \left(e^{-2\kappa_i(t_i-t)_+} - 1 \right) \\ & + \frac{\hat{\sigma}^2(r - 2\kappa_i)e^{2\tau_{i+1}}}{2(r - \kappa_i)} \left(e^{2(r-2\kappa_i)(t_i-t)_+} - e^{-2\kappa_i(t_i-t)_+} \right), \quad (9) \end{aligned}$$

where $\tau_j = \sum_{i=j}^n (r - 2\kappa_i)(t_i - t_{i-1})$ for $j = 1 \dots, n$. For $t < t_{i-1}$,

$$\hat{v}_i(t) = -\frac{\sigma_i^2}{2\kappa_i} + \left(\hat{v}_i(t_{i-1}) + \frac{\sigma_i^2}{2\kappa_i} \right) e^{2\kappa_i(t-t_{i-1})}.$$

I focus on the interesting case in which $\|\beta\| > \hat{\sigma}$. Figure 4 plots the optimal policy in an example with three components.²² (The variance $v_1(t)$ is always zero, so it is not plotted.) The critical component is $i_0 = 2$. At time $t_1 = 0$, the sender fully discloses component 1 and makes a partial disclosure about component 2. Then the sender gradually provides information about component 2, while keeping the receiver fully informed of component 1 and providing no information about component 3. At time t_2 , component 2 is fully revealed. Then the sender gradually reveals component 3, while keeping the receiver fully informed of components 1 and 2. At time t_3 , the entire state is

²²Here, $\|\beta\| = 5$, $r = 3$, $(\kappa_1, \kappa_2, \kappa_3) = (-0.75, -0.25, 0.25)$, and $\sigma_1 = \sigma_2 = \sigma_3 = 2$.

fully revealed, and the stationary phase begins: $\|\hat{b}(t)\| = \hat{\sigma}$ and $\hat{v}(t) = 0$ for $t \geq t_3$. Between the threshold times t_{i-1} and t_i , the bias magnitude decays exponentially at rate $r - 2\kappa_i$, the optimal rate of decay in the single-dimensional model with component i as the state.

7 Conclusion

This paper studies the optimal provision of information in a long-term relationship. The sender cannot induce actions that are biased in her own favor and also precisely tailored to the state. She resolves this tradeoff with a non-stationary policy that reveals information gradually over time. Initially, actions are biased but imprecise. Over a transition phase, actions become more precise but less biased, eventually reaching a stationary phase of perfect precision and constant bias. Throughout the relationship, the receiver chooses these biased actions in order to continue receiving information from the sender. The dynamics of this informational relationship arise endogenously, not because the sender comes to “trust” the receiver, but as the solution to a dynamic incentive problem.

A Formal definition of strategies

Fix a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ that is sufficiently rich to carry all the random objects introduced below. The driving process $Z = \{Z_t\}_{t \geq 0}$ is a standard Brownian motion. The initial state θ_0 has a normal distribution $N(\mu_0, \sigma_0^2)$, independent of Z . Let

$$\theta_t = \theta_0 e^{\kappa t} + \sigma \int_0^t e^{\kappa(t-s)} dZ_s, \quad (10)$$

where the integral is an Itô integral. This process satisfies the stochastic differential equation in (1). Let $\{\mathcal{F}_t^\theta\}$ denote the filtration generated by $\{\theta_t\}$,²³ and set $\mathcal{F}_\infty^\theta = \sigma(\bigcup_{t \geq 0} \mathcal{F}_t^\theta)$.

As a randomization device, the sender is endowed with a standard Brownian motion $Y = \{Y_t\}_{t \geq 0}$ that is independent of $\mathcal{F}_\infty^\theta$. Let $\tilde{\mathcal{F}}_t^\theta = \sigma(\sigma(Y) \cup \mathcal{F}_t^\theta)$. The filtration $\{\tilde{\mathcal{F}}_t^\theta\}$ represents the exogenous information available to the sender.

To define the receiver's action paths, let $D[0, \infty)$ denote the space of all cadlag functions from $[0, \infty)$ to \mathbf{R} .²⁴ Equip $D[0, \infty)$ with the cylindrical filtration $\{\mathcal{C}_{t-}\}$, where \mathcal{C}_{t-} denotes the σ -algebra generated by the projection maps for times strictly before t .

An *information policy* consists of a measurable space $(\mathbf{S}, \mathcal{S})$ and a map

$$S: [0, \infty) \times \Omega \times D[0, \infty) \rightarrow \mathbf{S}$$

that is adapted to the filtration $\{\tilde{\mathcal{F}}_t^\theta \otimes \mathcal{C}_{t-}\}$ on $\Omega \times D[0, \infty)$. The interpretation is that the sender's signal at time t can depend on (i) the sender's exogenous information, through $\tilde{\mathcal{F}}_t^\theta$, and (ii) actions taken by the receiver strictly before time t , through \mathcal{C}_{t-} . Denote the entire information policy by S .

A *decision rule* is a real-valued $\{\tilde{\mathcal{F}}_t^\theta\}$ -adapted stochastic process on (Ω, \mathcal{F}) with cadlag sample paths. By Karatzas and Shreve (1998, Proposition 1.13,

²³That is, $\mathcal{F}_t^\theta = \sigma(\theta_s : s \leq t)$ for each t .

²⁴Cadlag functions are right-continuous and have left limits. The space $D[0, \infty)$ is called a Skorokhod space. See Billingsley (1968) for further details.

p. 5), a decision rule is progressively measurable with respect to $\{\tilde{\mathcal{F}}_t^\theta\}$.

Next, I define the compatibility of a decision rule with an information policy. Once this definition is in place, a decision rule A is a *best response* to an information policy S if (i) A is compatible with S , and (ii) $u_R(A) \geq u_R(A')$ for all decision rules A' compatible with S . Unfortunately, perfect monitoring in continuous time poses technical challenges (Simon and Stinchcombe, 1989; Bergin and MacLeod, 1993). In settings with continuous actions, I am not aware of a satisfactory restriction on strategies that avoids all of these technical problems.²⁵ Fortunately, the obedience characterization (Proposition 1) is robust to the exact definition of strategies. To demonstrate this, I take the following axiomatic approach.

The sender is restricted to some subcollection of *admissible* information policies. Each admissible information policy is associated with a collection of decision rules that are *compatible* with that policy. I assume that these notions of admissibility and compatibility satisfy the following two conditions.

- C1. If a decision rule A is compatible with an admissible information policy S , then so is any decision rule A' defined as follows. For some fixed time t_0 and some event G in $\mathcal{F}_{t_0}^A$, let $A'_t(\omega) = A_t(\omega)$ whenever $t < t_0$ or $\omega \notin G$; otherwise, define A' so that GA'_t is $\mathcal{F}_{t_0}^A$ -measurable for all $t \geq t_0$.²⁶

If the receiver can follow the decision rule A under information policy S , then starting at time t_0 , conditional on the event G , the receiver can select different actions, using information that was available at time t_0 .

²⁵The challenge is that the sender chooses how her signals depends on the receiver's actions in the *arbitrarily recent* past. If the dependence of the signals on actions is exogenous and noisy, e.g., through the drift of a Brownian motion (Sannikov, 2008), then decision rules can be defined with respect to an exogenous filtration. Without this exogenous structure, the intuitive measurability conditions become self-referential. Without further restrictions, these measurability conditions do not exclude pathological decision rules in which the receiver uses the sender's feedback rule to instantaneously transmit to himself exogenous information that he never receives directly.

Existing methods cover only settings in which the players choose when to switch between discrete actions. The grid method of Simon and Stinchcombe (1989) and the inertia strategies of Bergin and MacLeod (1993) are defined in deterministic environments. Kamada and Rao (2021) introduce a new approach for stochastic settings.

²⁶Here and below, I denote a set and its indicator function by the same symbol.

C2. The following trigger information policies are all admissible. Let X be a decision rule. Let $T: D[0, \infty) \times D[0, \infty) \rightarrow \mathbf{R}$ be a function such that, for each time t , the event $[T < t]$ is in $\mathcal{C}_{t-} \otimes \mathcal{C}_{t-}$. The (X, T) -trigger information policy S is defined by

$$S(t, \omega, a) = X_{t \wedge T(X(\omega), a)}(\omega).$$

Since X is $\{\tilde{\mathcal{F}}_t^\theta\}$ -progressively measurable, S is $\{\tilde{\mathcal{F}}_t^\theta \otimes \mathcal{C}_{t-}\}$ -adapted.

A decision rule A is compatible with the (X, T) -trigger policy S if and only if

$$[A_t \in B] \cap [T(X, A) \wedge t < s] \in \mathcal{F}_s^X,$$

for every Borel set B and all times s and t . This implies, in particular, that A is $\{\mathcal{F}_t^X\}$ -adapted.

The interpretation is that the receiver's decision cannot depend on the realizations of the process X after it is stopped by the sender.

It can be checked that the definition of compatibility (with trigger policies) in C2 satisfies C1. Thus, it is consistent with the axioms to restrict the sender to trigger information policies and to define compatibility by C2. Of course, there are many other classes of reasonable information policies. If we include additional policies in the admissible set, then the obedience characterization (Proposition 1) still holds, as long as compatibility with these new information policies is defined in a way that is consistent with C1.

B Proofs

In the proofs below, I use the following form of the law of total variance. For any square-integrable random variable X and any sub- σ -algebras \mathcal{G} and \mathcal{H} satisfying $\mathcal{G} \supset \mathcal{H}$,

$$\text{var}(X|\mathcal{H}) = \mathbf{E}[\text{var}(X|\mathcal{G})|\mathcal{H}] + \text{var}(\mathbf{E}[X|\mathcal{G}]|\mathcal{H}) \geq \mathbf{E}[\text{var}(X|\mathcal{G})|\mathcal{H}]. \quad (11)$$

Taking \mathcal{H} to be the trivial σ -algebra gives the usual law of total variance.

If X and Y are square integrable, and X is measurable with respect to \mathcal{G} , then

$$\mathbf{E}(X - Y)^2 \geq \text{var}(X - Y) \geq \mathbf{E}[\text{var}(X - Y|\mathcal{G})] = \mathbf{E}[\text{var}(Y|\mathcal{G})], \quad (12)$$

where the middle inequality uses the usual law of total variance and the last inequality uses the \mathcal{G} -measurability of X .

B.1 Proof of Proposition 1

First, I prove that obedience is necessary. I prove the contrapositive. Let A be a decision rule that is not obedient. That is, there exists some time t and some positive-measure set G in \mathcal{F}_t^A such that

$$-\mathbf{E} \left[G \int_t^\infty r e^{-r(s-t)} (A_s - \theta_s)^2 ds \right] < -\mathbf{E} \left[G \frac{\sigma^2 + r(\mathbf{E}[\theta_t|\mathcal{F}_t^A] - \theta_t)^2}{r - 2\kappa} \right]. \quad (13)$$

Define a new decision rule A' by setting $A'_s = e^{\kappa(s-t)} \mathbf{E}[\theta_t|\mathcal{F}_t^A]$ for $s \geq t$ on G , and setting A' equal to A otherwise. By (13), it follows that $u_R(A') > u_R(A)$. If A is compatible with some information policy, then so is A' (by C1), so A cannot be a best response to any policy.

In order to prove that obedience is sufficient, I first define the direct, grim-trigger information policy S associated to a fixed decision rule A as the trigger decision rule (from C2) with $X = A$ and

$$T(a, a') = \inf \left\{ t \geq 0 : \int_0^t (a_s - a'_s)^2 ds > 0 \right\},$$

where the infimum of the empty set equals ∞ . Using Fubini's theorem, it can be shown that the event $[T < t]$ is in $\{\mathcal{C}_{t-} \otimes \mathcal{C}_{t-}\}$, as required. The information policy S is defined by

$$S(t, \omega, a) = A_{t \wedge T(A(\omega), a)}(\omega).$$

Now I prove that obedience is sufficient. Let A be an obedient decision rule. Let S be the associated direct, grim-trigger information policy. I claim that A is a best response to S . Clearly A is compatible with S , since $T(A(\omega), A(\omega)) = \infty$ for all ω . Let $A' = \{A'_t\}$ be an arbitrary decision rule that is compatible with S . I claim that $u_R(A) \geq u_R(A')$. This inequality holds trivially if $u_R(A') = -\infty$, so assume that $u_R(A')$ is finite.

Define the random time T' by $T'(\omega) = T(A(\omega), A'(\omega))$. I approximate T' from above by simple functions. For each n , define the simple random times T_n by

$$T_n = \begin{cases} j/2^n & \text{if } (j-1)/2^n \leq T' < j/2^n, \quad j = 1, \dots, n2^n, \\ \infty & \text{if } T' \geq n. \end{cases}$$

By construction, $T_n > T'$ and $T_n \downarrow T'$. Let A^n be the decision rule that agree with A if $t < T_n$ and agrees with A' if $t \geq T_n$. Since $u_R(A')$ and $u_R(A)$ are both finite, Lebesgue's dominated convergence theorem implies that $u_R(A^n) \rightarrow u_R(A')$. Therefore, it suffices to check that $u_R(A^n) \leq u_R(A)$ for each n .

Fix n and let t_1, \dots, t_K denote the finite values that T_n takes with positive probability. For each k , let I_k be the indicator for the event that $T_n = t_k$. We have

$$u_R(A) - u_R(A^n) = \sum_{k=1}^K \mathbf{E} \left[I_k \int_{t_k}^{\infty} r e^{-rt} ((A'_t - \theta_t)^2 - (A_t - \theta_t)^2) dt \right]. \quad (14)$$

For $t \geq t_k$, the random variable $I_k A'_t$ is $\mathcal{F}_{t_k}^A$ -measurable by C2, so (12) gives

$$\mathbf{E}[I_k (A'_t - \theta_t)^2] \geq \mathbf{E}[I_k \text{var}(\theta_t | \mathcal{F}_{t_k})] = \mathbf{E}[I_k \eta(\text{var}(\theta_{t_k} | \mathcal{F}_{t_k}), t - t_k)], \quad (15)$$

where η is the function defined in (4). To see that each expectation in (14) is nonnegative, change the order of integration, substitute in (15), and then use the time- t_k obedience constraint for A . We conclude that $u_R(A) \geq u_R(A^n)$.

B.2 Proof of Proposition 2

Let A be an obedient decision rule. Define the bias and variance functions b and v by

$$b(t) = \mathbf{E}[A_t - \theta_t] \quad \text{and} \quad v(t) = \mathbf{E}[\text{var}(\theta_t | \mathcal{F}_t^A)].$$

First, I check that v is Bayes-plausible. By the law of total variance,

$$v(0) = \mathbf{E}[\text{var}(\theta_0 | \mathcal{F}_0^A)] \leq \text{var}(\theta_0) = \sigma_0^2.$$

For $s > t$, the stronger law of total variance in (11) gives

$$\mathbf{E}[\text{var}(\theta_s | \mathcal{F}_s^A) | \mathcal{F}_t^A] \leq \text{var}(\theta_s | \mathcal{F}_t^A) = \eta(\text{var}(\theta_t | \mathcal{F}_t^A), s - t), \quad (16)$$

where the equality can be derived from (10), using Itô's isometry and the $\{\tilde{\mathcal{F}}_t^\theta\}$ -adaptedness of A . Since η is linear in its first argument, taking expectations in (16) gives $v(s) \leq \eta(v(t), s - t)$.

Since v is Bayes-plausible, it follows from Theorem 1 (proven below, without appealing to Proposition 2) that there exists a reporting function φ that induces v . Define the decision rule A' by

$$A'_t = \mathbf{E}[\theta_t | \theta_{\varphi(t)}] + b(t).$$

Take expectations in the time- t obedience constraint for A to get

$$\mathbf{E} \left[\int_t^\infty r e^{-r(s-t)} (A'_s - \theta_s)^2 ds \right] \leq \frac{\sigma^2 + r \mathbf{E}[\text{var}(\theta_t | \theta_{\varphi(t)})]}{r - 2\kappa},$$

which is exactly the time- t obedience constraint for A' . Thus, A' is obedient.

It remains to check that both players weakly prefer A' to A . This holds because

$$\mathbf{E}(A_t - \theta_t)^2 = \mathbf{E}[B_t^2] + \mathbf{E}[V_t] \geq (\mathbf{E} B_t)^2 + \mathbf{E}[V_t] = \mathbf{E}(A'_t - \theta_t)^2,$$

and similarly, $\mathbf{E}(A_t - \theta_t - \beta)^2 \geq \mathbf{E}(A'_t - \theta_t - \beta)^2$.

B.3 Proof of Theorem 1

Let $v: [0, \infty) \rightarrow [0, \infty)$ be Bayes-plausible. Then v has the following *monotonicity* property. If $v(t) \leq \eta(w, t)$ for some fixed w and t , then for $s > t$ we have

$$v(s) \leq \eta(v(t), s - t) \leq \eta(\eta(w, t), s - t) = \eta(w, s),$$

where the second inequality holds because η is strictly increasing in its first argument. Moreover, the second inequality holds strictly if $v(t) < \eta(w, t)$.

Define the reporting function φ implicitly by the following piecewise system (which separates into cases according to the sign of $\varphi(t)$):

$$\begin{cases} \varphi(t) = -\infty & \text{if } v(t) = \eta(\sigma_0^2, t), \\ v(t) = \eta((1/\sigma_0^2 - 1/\varphi(t))^{-1}, t) & \text{if } \eta(0, t) < v(t) < \eta(\sigma_0^2, t), \\ v(t) = \eta(0, t - \varphi(t)) & \text{if } 0 \leq v(t) \leq \eta(0, t). \end{cases}$$

This construction ensures that φ induces v , provided that φ is a well-defined reporting function.

First I check that φ is well-defined. By the initial variance constraint and the monotonicity property, $v(t) \leq \eta(\sigma_0^2, t)$ for all $t \geq 0$, so the cases are exhaustive. In the second case, the solution is unique because η is strictly increasing in its first argument. In the third case, the solution is unique because $\eta(0, \cdot)$ is strictly increasing and $\eta(0, 0) = 0$.

Now I check that φ is a reporting function. For a fixed t , if the second case obtains, we get $\varphi(t) \leq 0$. If the third case obtains, we get $0 \leq \varphi(t) \leq t$. The monotonicity property ensures that as time passes, it is only possible to move from an earlier case to a later case (and not the reverse). Within each case, the monotonicity property ensures that v is weakly increasing in time (since η is strictly increasing in its first argument and the function $\eta(0, \cdot)$ is strictly increasing).

B.4 Proof of Theorem 2

Assume $\beta > \sigma/\sqrt{r - 2\kappa}$, for otherwise the result is clear. The full-disclosure time T in the theorem statement is well-defined because the expressions for $\hat{v}(0)$ and $\hat{b}(0)$, as functions of T , are strictly increasing.

Drop the no-disclosure bounds on the variance to obtain the relaxed problem²⁷

$$\begin{aligned}
& \text{minimize} && \int_0^\infty r e^{-rt} [(b(t) - \beta)^2 + v(t)] dt \\
& \text{subject to} && \int_t^\infty r e^{-r(s-t)} [\hat{b}(s)^2 + v(s)] ds \leq \frac{\sigma^2 + rv(t)}{r - 2\kappa}, \quad t \geq 0 \\
& && v(t) \geq 0, \quad t \geq 0 \\
& && v(0) \leq \sigma_0^2.
\end{aligned} \tag{17}$$

It is straightforward to check that \hat{v} satisfies the no-disclosure bounds. I will prove that (\hat{b}, \hat{v}) is the unique solution of (17). The proof is separated into two parts. The first part proves optimality. In the second part, the claimed uniqueness is stated precisely and then proved.

Optimality First, drop the initial variance constraint:

$$\begin{aligned}
& \text{minimize} && \int_{t_0}^\infty r e^{-rt} [(b(t) - \beta)^2 + v(t)] dt \\
& \text{subject to} && \int_t^\infty r e^{-r(s-t)} [b^2(s) + v(s)] ds \leq \frac{\sigma^2 + rv(t)}{r - 2\kappa}, \quad t \geq 0 \\
& && v(t) \geq 0, \quad t \geq 0.
\end{aligned} \tag{18}$$

Define auxiliary functions \tilde{b} and \tilde{v} by the corresponding expressions for b and v in the theorem statement, but with T defined by

$$\beta e^{-(r-2\kappa)T} = \frac{\sigma}{\sqrt{r - 2\kappa}}. \tag{19}$$

²⁷To simplify notation, I work with losses rather than utilities throughout the appendix.

Below, I will show that (\tilde{b}, \tilde{v}) solves (18). I claim that this implies that (\hat{b}, \hat{v}) solves (17). If $\sigma_0^2 \geq \tilde{v}(0)$, then (\tilde{b}, \tilde{v}) solves (17), and $(\tilde{b}, \tilde{v}) = (\hat{b}, \hat{v})$. If $\sigma_0^2 < \tilde{v}(0)$, then there is a unique time t_0 such that $\tilde{v}(t_0) = \sigma_0^2$. Since the time- t_0 obedience constraint is active, it follows from Bellman's principle of optimality that the map $t \mapsto (\tilde{b}(t_0 + t), \tilde{v}(t_0 + t))$, which equals (\hat{b}, \hat{v}) , solves (17). Otherwise, replacing the time- t_0 continuation policy in (\tilde{b}, \tilde{v}) with a solution of (17) would strictly increase the sender's payoff, while preserving all obedience constraints in (18).

To prove that (\tilde{b}, \tilde{v}) solves (18), attach nonnegative Lagrange multipliers $e^{-rt}\lambda(t)$ to each time- t obedience constraint and $re^{-rt}\mu(t)$ to each time- t non-negativity constraint. Integrate over these constraints to form the Lagrangian

$$\begin{aligned} L(b, v; \lambda, \mu) &= \int_0^\infty re^{-rt} [(b(t) - \beta)^2 + v(t)] dt \\ &\quad + \int_0^\infty e^{-rt}\lambda(t) \left\{ \int_t^\infty re^{-r(s-t)} [b^2(s) + v(s)] ds - \frac{\sigma^2 + rv(t)}{r - 2\kappa} \right\} dt \\ &\quad - \int_0^\infty re^{-rt}\mu(t)v(t) dt. \end{aligned}$$

After splitting the term in braces,²⁸ the double integral in the obedience constraint can be rearranged as

$$\int_0^\infty \int_t^\infty re^{-rs}\lambda(t) [b^2(s) + v(s)] ds dt = \int_0^\infty re^{-rs} \left(\int_0^s \lambda(t) dt \right) [b^2(s) + v(s)] ds,$$

where I have switched the order of integration by Tonelli's theorem. Next, swap the dummy variable names s and t in this integral, and define the accumulated multiplier

$$\Lambda(t) = \int_0^t \lambda(s) ds.$$

²⁸It suffices to define the Lagrangian for functions (b, v) that yield finite loss for the sender. For such functions, both integrals are finite as long as λ and μ are bounded, as they will be below.

After these simplifications, we have

$$L(b, v; \lambda, \mu) = \int_0^\infty r e^{-rt} \ell(b(t), v(t); \lambda(t), \mu(t)) dt - \frac{\sigma^2}{r - 2\kappa} \int_0^\infty e^{-rt} \lambda(t) dt,$$

where $\ell(b(t), v(t); \lambda(t), \mu(t))$ equals

$$(b(t) - \beta)^2 + \Lambda(t)b^2(t) + (1 + \Lambda(t) - \lambda(t)/(r - 2\kappa) - \mu(t))v(t). \quad (20)$$

Define the nonnegative multipliers by

$$(\hat{\lambda}(t), \hat{\mu}(t)) = \begin{cases} ((r - 2\kappa)e^{(r-2\kappa)t}, 0) & \text{if } t < T, \\ (0, e^{(r-2\kappa)T}) & \text{if } t \geq T. \end{cases}$$

With these multipliers, the coefficient on $v(t)$ vanishes, and the integrand in (20) becomes

$$(b(t) - \beta)^2 + (e^{(r-2\kappa)(t \wedge T)} - 1)b^2(t).$$

This expression is convex in $b(t)$ and the first-order condition gives

$$b(t) = \beta e^{-(r-2\kappa)(t \wedge T)} = \frac{\sigma}{\sqrt{r - 2\kappa}} e^{(r-2\kappa)(T-t)_+} = \tilde{b}(t),$$

where the middle equality uses the identity $t \wedge T = T - (T - t)_+$ and the definition of T in (19).

It follows that $(\tilde{b}(t), \tilde{v}(t))$ minimizes $\ell(\cdot, \cdot; \hat{\lambda}(t), \hat{\mu}(t))$ for each time t , hence (\tilde{b}, \tilde{v}) minimizes $L(\cdot, \cdot; \hat{\lambda}, \hat{\mu})$. It can be checked that (\tilde{b}, \tilde{v}) satisfies every obedience constraint with equality, so the Kuhn–Tucker conditions are satisfied.

Uniqueness I claim that if a function (b, v) solves (17), then $(b(t), v(t)) = (\hat{b}(t), \hat{v}(t))$ for almost every time t . To see this, suppose (b, v) solves (17). Then $b(t) = \hat{b}(t)$ for almost every t ; otherwise, $(b/2 + \hat{b}/2, v/2 + \hat{v}/2)$ is a feasible strict improvement. Hence, $v(t) = \hat{v}(t)$ for almost every t ; otherwise, $(b, v \wedge \hat{v})$

is a feasible strict improvement, where $v \wedge \hat{v}$ denotes the pointwise minimum.²⁹

B.5 Proof of Corollary 1

For uniqueness, follow the argument from the proof of Theorem 2 in Appendix B.4. For optimality, observe that the sender's objective, denoted $u_S(b, v)$, is independent of the volatility parameter σ . In the relaxed problem (without the initial variance constraint) from (18), the feasible set is increasing (with respect to set inclusion) in σ . For each $\sigma > 0$, let $(\tilde{b}_\sigma, \tilde{v}_\sigma)$ denote the solution of (18) when the volatility equals σ . Define $(\tilde{b}_0, \tilde{v}_0)$ by taking $b_0 = \beta$ in the expressions from the statement of Corollary 1. Fix a positive sequence $\{\sigma_n\}$ satisfying $\sigma_n \downarrow 0$. It suffices to check that $u_S(\hat{b}_{\sigma_n}, \hat{v}_{\sigma_n}) \rightarrow u_S(\hat{b}_0, \hat{v}_0)$. Observe that $(\tilde{b}_{\sigma_n}, \tilde{v}_{\sigma_n})$ converges pointwise to $(\tilde{b}_0, \tilde{v}_0)$. By Proposition 3, $(\tilde{b}_{\sigma_n}(t) - \beta)^2 + \tilde{v}_{\sigma_n}(t)$ is monotonically increasing in n for each t , so by Lebesgue's monotone convergence theorem, $u_S(\tilde{b}_{\sigma_n}, \tilde{v}_{\sigma_n}) \rightarrow u_S(\tilde{b}_0, \tilde{v}_0)$. Therefore, $(\tilde{b}_0, \tilde{v}_0)$ is optimal in (18), so (\hat{b}_0, \hat{v}_0) is optimal in (17) by Bellman's principle of optimality.

B.6 Proof of Proposition 3

Assume that $\beta > \sigma/\sqrt{r - 2\kappa}$ and the initial variance constraint is not active.

We have

$$\hat{b}(t) = \max \left\{ \beta e^{-(r-2\kappa)t}, \frac{\sigma}{\sqrt{r-2\kappa}} \right\},$$

so the comparative statics for \hat{b} are clear. For the variance, observe that the expression for \hat{v} in the theorem statement is strictly increasing as a function of T and it is otherwise independent of β . Since the full-disclosure time T is strictly increasing in β , it follows that \hat{v} is increasing in β .

Finally, I check that \hat{v} is decreasing in σ . Fix σ_1 and σ_2 with $0 < \sigma_1 < \sigma_2$. For each volatility σ_i , denote the optimal bias–variance pair by (\hat{b}_i, \hat{v}_i) and the

²⁹In fact, a slightly stronger result holds. Since (\hat{b}, \hat{v}) satisfies the obedience constraint with equality for every time t , it follows that $v(t) \geq \hat{v}(t)$ for every time t . If (b, v) also satisfies the no-disclosure upper bounds, then v must be lower semicontinuous, and hence $v(t) = \hat{v}(t)$ for every time t .

full-disclosure time by T_i . We have $T_1 > T_2$. For $t \geq T_1$, we have $\hat{v}_1(t) = \hat{v}_2(t) = 0$. For $T_2 \leq t < T_1$, we have $\hat{v}_1(t) > \hat{v}_2(t) = 0$. Finally, for $t < T_2$, observe that $\hat{b}_i(t)$ is independent of i and

$$(r - 2\kappa)\hat{b}_i^2(t) = 2\kappa\hat{v}_i(t) + \sigma_i^2 - \hat{v}_i'(t),$$

hence

$$\hat{v}_1'(t) - \hat{v}_2'(t) = 2\kappa(\hat{v}_1(t) - \hat{v}_2(t)) + \sigma_1^2 - \sigma_2^2 < 2\kappa(\hat{v}_1(t) - \hat{v}_2(t)).$$

Over the interval $(0, T_2)$, the function $f(s) = \hat{v}_2(T_2 - s) - \hat{v}_1(T_2 - s)$ satisfies $f'(s) \leq -2\kappa f(s)$. By Grönwall's inequality, $f(s) \leq e^{-2\kappa s} f(0)$ for all s in $[0, T_2]$, hence

$$\hat{v}_1(t) - \hat{v}_2(t) \geq e^{-2\kappa}(\hat{v}_1(T_2) - \hat{v}_2(T_2)) > 0,$$

for all t in $[0, T_2]$.

B.7 Proof of Theorem 3

Assume $\|\beta\| > \hat{\sigma}$, for otherwise the result is clear. It can be shown that the full-disclosure times are well-defined. The details are in the last subsection of the proof. Taking as given that these times are well-defined, I prove the result.

Drop the no-disclosure bounds except those starting at time 0 to obtain the relaxed problem

$$\begin{aligned} & \text{minimize} && \int_0^\infty r e^{-rt} \left[\|b(t) - \beta\|^2 + \sum_{i=1}^n v_i(t) \right] dt \\ & \text{subject to} && \int_t^\infty r e^{-r(s-t)} \left[\|b(s)\|^2 + \sum_{i=1}^n v_i(s) \right] ds \leq \sum_{i=1}^n \frac{\sigma_i^2 + r v_i(t)}{r - 2\kappa_i}, \quad t \geq 0 \\ & && 0 \leq v_i(t) \leq \eta_i(\sigma_{0,i}^2, t), \quad i = 1, \dots, n, \quad t \geq 0. \end{aligned} \tag{21}$$

It is straightforward to check that \hat{v} satisfies the dropped no-disclosure bounds. I prove that (\hat{b}, \hat{v}) solves (21), and then I check uniqueness.

Optimality To handle the initial variance constraint, consider an auxiliary problem. If $i_0 = 1$ and $\|\hat{b}(0)\| < \|\beta\|$, define t_0 so that $\hat{b}(t_0) = \beta$ (where $\hat{b}(t)$ is defined by the expression in the theorem statement for all real t). Otherwise, set $t_0 = 0$. The auxiliary problem is to choose functions b and v on $[t_0, \infty)$ to solve

$$\begin{aligned}
& \text{minimize} && \int_{t_0}^{\infty} r e^{-rt} \left[\|b(t) - \beta\|^2 + \sum_{i=1}^n v_i(t) \right] dt \\
& \text{subject to} && \int_t^{\infty} r e^{-r(s-t)} \left[\|b(s)\|^2 + \sum_{i=1}^n v_i(s) \right] ds \leq \sum_{i=1}^n \frac{\sigma_i^2 + r v_i(t)}{r - 2\kappa_i}, \quad t \geq t_0 \\
& && v_i(t) \geq 0, \quad i = 1, \dots, n, \quad t \geq t_0 \\
& && v_i(t) \leq \eta_i(v_i(t_0), t - t_0), \quad i = 1, \dots, n, \quad t \geq t_0.
\end{aligned} \tag{22}$$

Define auxiliary functions \tilde{b} and \tilde{v} on $[t_0, \infty)$ by the expressions for \hat{b} and \hat{v} (with the new definition of t_0). Below, I will show that (\tilde{b}, \tilde{v}) solves (22). I claim that this implies that (\hat{b}, \hat{v}) solves (21). If $t_0 = 0$, then (\tilde{b}, \tilde{v}) solves (21), and (\hat{b}, \hat{v}) equals (\tilde{b}, \tilde{v}) . If $t_0 < 0$, then $\tilde{v}_i(0) = \sigma_{0,i}^2$ for all $i = 1, \dots, n$. Since the time-0 obedience constraint is active, it follows from Bellman's principle of optimality that the restriction of (\tilde{b}, \tilde{v}) to $[0, \infty)$, which equals (\hat{b}, \hat{v}) , solves (21). Otherwise, replacing the time-0 continuation policy in (\tilde{b}, \tilde{v}) with a solution of (21) would strictly increase the sender's payoff, while preserving all obedience constraints in (22).

To prove that (\tilde{b}, \tilde{v}) solves (22), attach nonnegative Lagrange multipliers $e^{-rt}\lambda(t)$ to each time- t obedience constraint, $r e^{-rt}\mu_i(t)$ to the nonnegativity constraint on $v_i(t)$, and $r e^{-rt}\gamma_i(t)$ to the no-disclosure bound on $v_i(t)$. Integrate over these constraints to form the Lagrangian $L(b, v, \lambda, \mu, \gamma)$. Simplifying as in the single-dimensional case, we have

$$\begin{aligned}
L(b, v; \lambda, \mu, \gamma) &= \int_{t_0}^{\infty} r e^{-rt} \ell(b(t), v(t), \lambda(t), \mu(t), \gamma(t)) dt \\
&\quad - \sum_{i=1}^n \frac{\sigma_i^2}{r - 2\kappa_i} \int_{t_0}^{\infty} e^{-rt} \lambda(t) dt
\end{aligned}$$

$$- \sum_{i=1}^n \int_{t_0}^{\infty} r e^{-rt} \gamma_i(t) \eta_i(v_i(t_0), t - t_0) dt,$$

where the integrand $\ell(b(t), v(t), \lambda(t), \mu(t), \gamma(t))$ equals

$$\|b(t) - \beta\|^2 + \Lambda(t) \|b(t)\|^2 + \sum_{i=1}^n \left(1 + \Lambda(t) - \frac{\lambda(t)}{r - 2\kappa_i} - \mu_i(t) + \gamma_i(t) \right) v_i(t). \quad (23)$$

Next, define the multipliers as follows. First, define $j: [t_0, t_n) \rightarrow \{i_0, \dots, n\}$ by $j(t) = i$ if $t \in [t_{i-1}, t_i)$. Let

$$S(t) = \int_{t_0}^t (r - 2\kappa_{j(s)}) ds = \sum_{i=i_0}^n (r - 2\kappa_i) (t \wedge t_i - t_{i-1})_+.$$

Let

$$\hat{\lambda}(t) = \begin{cases} (r - 2\kappa_{j(t)}) e^{S(t)} & \text{if } t < t_n \\ 0 & \text{if } t \geq t_n. \end{cases}$$

It can be shown that that $\hat{\Lambda}(t) = e^{S(t)} - 1$. Next, for $i = 1, \dots, n$, let

$$(\hat{\mu}_i(t), \hat{\gamma}_i(t)) = \begin{cases} (0, 2(\kappa_i - \kappa_{j(t)})(r - 2\kappa_i)^{-1} e^{S(t)}) & \text{if } t < t_{i-1}, \\ (0, 0) & \text{if } t_{i-1} \leq t < t_i, \\ (2(\kappa_{j(t)} - \kappa_i)(r - 2\kappa_i)^{-1} e^{S(t)}, 0) & \text{if } t_i \leq t. \end{cases}$$

By construction, the coefficients on each $v_i(t)$ in the Lagrangian vanish. The integrand in (23) becomes

$$\|b(t) - \beta\|^2 + (e^{S(t)} - 1) \|b(t)\|^2.$$

This expression is convex in $b(t)$ and the first-order condition gives $b(t) = \beta \exp(-S(t))$. From the definition of t_0 , we have

$$\beta = \hat{b}(t_0) = \hat{\sigma} \exp(S(t_n)) \frac{\beta}{\|\beta\|}.$$

Combining these identities, it follows that

$$b(t) = \hat{\sigma} \exp(S(t_n) - S(t)) \frac{\beta}{\|\beta\|} = \tilde{b}(t).$$

It follows that $(\tilde{b}(t), \tilde{v}(t))$ minimizes $\ell(\cdot, \cdot; \hat{\lambda}(t), \hat{\mu}(t), \hat{\gamma}(t))$ for each time t , hence (\tilde{b}, \tilde{v}) minimizes $L(\cdot, \cdot; \hat{\lambda}, \hat{\mu})$. It can be checked that (\tilde{b}, \tilde{v}) satisfies every obedience constraint with equality, so the Kuhn–Tucker conditions are satisfied.

Uniqueness Here, the argument is different from the single-dimensional case. By Bellman’s principle optimality, it suffices to show that (\tilde{b}, \tilde{v}) is the (almost everywhere) unique solution of (22). If a function (b, v) solves (22), then $L(b, v; \hat{\lambda}, \hat{\mu}, \hat{\gamma}) = L(\tilde{b}, \tilde{v}; \hat{\lambda}, \hat{\mu}, \hat{\gamma})$ and (b, v) satisfies complementary slackness with the multipliers $(\hat{\lambda}, \hat{\mu}, \hat{\gamma})$. For each t ,

$$\operatorname{argmax}_{(x,y) \in \mathbf{R} \times \mathbf{R}_+} \ell(x, y; \hat{\lambda}(t), \hat{\mu}(t), \hat{\gamma}(t)) = \{\hat{b}(t)\} \times \mathbf{R}_+.$$

Therefore, $b(t) = \hat{b}(t)$ for almost all t . For each t , we have $\hat{\lambda}(t) > 0$. If $\kappa_1 < \dots < \kappa_n$, then for each i , we have $\hat{\gamma}_i(t) > 0$ for $t < t_{i-1}$ and $\hat{\mu}_i(t) > 0$ for $t \geq t_i$. Therefore, complementary slackness implies that $v(t) = \hat{v}(t)$ for almost all t .³⁰

Definition of full-disclosure times Suppose that $\sigma_{0,i}^2 > 0$ for some i , for otherwise the full-disclosure times are all 0.

In the theorem statement, the bias and variance functions are expressed in terms of the future disclosure times. To prove that these full-disclosure times are well-defined, I express the bias and variance functions in terms of their values at time 0. The formal procedure follows.

³⁰In fact, this equality must hold for every t , by the same argument as in the single-dimensional case. If the κ_i agree for i in some subinterval I of $\{1, \dots, n\}$, then $\hat{\mu}_i(t) = \hat{\gamma}_i(t) = 0$ for all t in $T = \cup_{i \in I} [t_{i-1}, t_i)$. Only the sum $\sum_{i \in I} v_i(t)$ is pinned down for t in T . It is optimal to select any Bayes-plausible choices of $v_i(t)$, for i in I and t in T , that induce the correct sum.

Let \mathcal{V} contain the zero n -vector together with all n -vectors of the form $(0_{i_0-1}, \nu_{i_0}, \sigma_{0,i_0+1}^2, \dots, \sigma_{0,n}^2)$, for some component $i_0 \in \{1, \dots, n\}$ and some ν_{i_0} in $(0, \sigma_{0,i_0}^2]$, where 0_{i_0-1} denotes a zero vector with $i_0 - 1$ components. The set \mathcal{V} is totally ordered by the usual componentwise order. Fix $\alpha \in [\hat{\sigma}, \infty)$ and $\nu \in \mathcal{V}$. Define functions b and v and disclosure times t_1, \dots, t_n as follows. Set $t_i = 0$ for $i \leq i_0 - 1$. Let $v(0) = \nu$ and $b(0) = \alpha\beta/\|\beta\|$. For each i , given that $b(t)$ and $v(t)$ are defined for $t \leq t_{i-1}$, define the time t_i and the values $(b(t), v(t))$ for t in $(t_{i-1}, t_i]$ as follows. For the bias,

$$b(t) = b(t_{i-1})e^{-(r-2\kappa_i)(t-t_{i-1})}.$$

For $j \leq i - 1$, we have $v_j(t) = 0$. For $j \geq i + 1$, we have $v_j(t) = \eta(v_j(t_{i-1}), t - t_{i-1})$. Finally, $v_i(t)$ can be expressed explicitly, but it is more convenient to observe that it is the unique solution of the differential equation

$$v_i'(t) = 2\kappa_i v_i(t) - (r - 2\kappa_i)(\|b(t)\|^2 - \hat{\sigma}_i^2),$$

with the given boundary value $v_i(t_{i-1})$. Let t_i be the smallest time t such that either $v_i(t) = 0$ or $\|b(t)\| = \hat{\sigma}$. If $\|b(t_i)\| = \hat{\sigma}$, set $T = t_i$ and terminate the procedure. If $\|b(t_i)\| > \hat{\sigma}$, either proceed to the next step if $i < n$ or else set $T = t_n$ if $i = n$. This procedure determines a function $f: [\hat{\sigma}, \infty) \times \mathcal{V} \rightarrow \mathbf{R}$ by setting

$$f(\alpha, \nu) = \|b(T)\| - \hat{\sigma} - (v_1(T) + \dots + v_n(T)).$$

This function f is continuous in the parameters of the problem. I claim that f satisfies the following monotonicity properties:

- (i) For each $\nu \in \mathcal{V}$, the function $f(\cdot, \nu)$ is strictly single-crossing from below.
- (ii) For each $\alpha \in [\hat{\sigma}, \infty)$, the function $f(\alpha, \cdot)$ is strictly single-crossing from above.

First I complete the proof, taking these properties as given. Fix $\nu \in \mathcal{V}$. We have $f(\hat{\sigma}, \nu) = -(\nu_1 + \dots + \nu_n) \leq 0$. It can be shown that for α sufficiently large, $f(\alpha, \nu) > 0$. By continuity and (i), there is a unique value $\alpha^*(\nu) \in [\hat{\sigma}, \infty)$ such

that $f(\alpha^*(\nu), \nu) = 0$. By (i) and (ii), the function $\alpha^*: \mathcal{V} \rightarrow [\hat{\sigma}, \infty)$ is strictly increasing. Observe that $\alpha^*(0) = \hat{\sigma}$. Let $\hat{\alpha} = \alpha^*(\sigma_{0,1}^2, \dots, \sigma_{0,n}^2)$. By continuity, the image $\alpha^*(\mathcal{V})$ is the interval $[\hat{\sigma}, \hat{\alpha}]$, so α^* has a right inverse $(\alpha^*)^{-1}: [\hat{\sigma}, \hat{\alpha}] \rightarrow \mathcal{V}$. Therefore, taking $\hat{b}(0) = (\|\beta\| \wedge \hat{\alpha})\beta/\|\beta\|$ and $\hat{v}(0) = (\alpha^*)^{-1}(\|\beta\| \wedge \hat{\alpha})$ and then applying the procedure above gives the full-disclosure times t_1, \dots, t_n .

To prove (i), fix some nonzero $\nu \in \mathcal{V}$. (The result is clear with $\nu = 0$ since $f(\alpha, 0) = \alpha - \hat{\sigma}$ for all α .) Let i_0 denote the index of the first nonzero component of ν . Fix $\alpha, \bar{\alpha} \geq \hat{\sigma}$ with $\bar{\alpha} > \alpha$. Apply the procedure above from the initial conditions (α, ν) and $(\bar{\alpha}, \nu)$ to obtain $(b, v, t_1, \dots, t_n, T)$ and $(\bar{b}, \bar{v}, \bar{t}_1, \dots, \bar{t}_n, \bar{T})$, respectively. Suppose $f(\alpha, \nu) \geq 0$. Hence, $v(T) = 0$. I show that $f(\bar{\alpha}, \nu) > 0$. Since $\kappa_1 \leq \dots \leq \kappa_n$, it can be shown using Grönwall's inequality that:

- (a) $\bar{t}_i < t_i$ for $i \geq i_0$;
- (b) $\|\bar{b}(t)\| > \|b(t)\|$ for $t \leq \min\{\bar{T}, T\}$;
- (c) $\bar{v}_i(t) \leq v_i(t)$ for $t \leq \min\{\bar{T}, T\}$.

In particular, we cannot have $\bar{T} > T$ for then $\bar{v}(T) \leq v(T) = 0$. So $\bar{T} \leq T$, and since \bar{b} is strictly decreasing, $\|\bar{b}(\bar{T})\| > \|b(T)\| \geq \hat{\sigma}$. Thus, $f(\bar{\alpha}, \nu) > 0$.

To prove (ii), fix $\alpha > \hat{\sigma}$. (The result is clear with $\alpha = \hat{\sigma}$ since $f(\hat{\sigma}, \nu) = -(\nu_1 + \dots + \nu_n)$ for all ν .) Fix $\nu, \bar{\nu} \in \mathcal{V}$ with $\bar{\nu} > \nu$. Apply the procedure above from the initial conditions (α, ν) and $(\alpha, \bar{\nu})$ to obtain $(b, v, t_1, \dots, t_n, T)$ and $(\bar{b}, \bar{v}, \bar{t}_1, \dots, \bar{t}_n, \bar{T})$, respectively. Suppose $f(\alpha, \nu) \leq 0$. Hence, $\|b(T)\| = \hat{\sigma}$, which implies that ν is nonzero. I show that $f(\alpha, \bar{\nu}) < 0$. Since $\kappa_1 \leq \dots \leq \kappa_n$, it can be shown using Grönwall's inequality that:

- (a) for $i \geq i_0$, if $\bar{v}_i(\bar{T}) = 0$, then $\bar{t}_i > t_i$;
- (b) $\|\bar{b}(t)\| \leq \|b(t)\|$ for $t \leq \min\{\bar{T}, T\}$;
- (c) $\bar{v}_i(t) \geq v_i(t)$ for $t \leq \min\{\bar{T}, T\}$.

In particular, we cannot have $\bar{T} > T$ for then $\bar{b}(T) \leq b(T) = \hat{\sigma}$. So $\bar{T} \leq T$, and hence $\bar{v}_n(\bar{T}) > 0$, for otherwise $t_n < \bar{T} \leq T$, which is a contradiction. Thus, $f(\alpha, \bar{\nu}) < 0$.

References

- AU, P. H. (2015): “Dynamic Information Disclosure,” *RAND Journal of Economics*, 46, 791–823. [5]
- BERGEMANN, D. AND S. MORRIS (2016): “Bayes Correlated Equilibrium and the Comparison of Information Structures in Games,” *Theoretical Economics*, 11, 487–522. [7]
- BERGIN, J. AND W. B. MACLEOD (1993): “Continuous Time Repeated Games,” *International Economic Review*, 34, 21–37. [27]
- BILLINGSLEY, P. (1968): *Convergence of Probability Measures*, Wiley. [26]
- BIZOTTO, J., J. RÜDIGER, AND A. VIGIER (2021): “Dynamic Persuasion with Outside Information,” *American Economic Journal: Microeconomics*, 13, 179–194. [5]
- ELY, J. C. (2015): “Beeps,” Working paper. [5]
- (2017): “Beeps,” *American Economic Review*, 107, 31–53. [5, 15]
- ELY, J. C. AND M. SZYDLOWSKI (2020): “Moving the Goalposts,” *Journal of Political Economy*, 128, 486–506. [5]
- FUDENBERG, D. AND L. RAYO (2019): “Training and Effort Dynamics in Apprenticeship,” *American Economic Review*, 109, 3780–3812. [5, 20]
- GARICANO, L. AND L. RAYO (2017): “Relational Knowledge Transfers,” *American Economic Review*, 107, 2695–2730. [5]
- HARRIS, M. AND B. HOLMSTROM (1982): “A Theory of Wage Dynamics,” *Review of Economic Studies*, 49, 315–333. [5]
- HERRERA, H., E. REUBEN, AND M. M. TING (2017): “Turf Wars,” *Journal of Public Economics*, 152, 143–153. [2]
- HONRYO, T. (2018): “Dynamic Persuasion,” *Journal of Economic Theory*, 178, 36–58. [5]
- KAMADA, Y. AND N. RAO (2021): “Strategies in Stochastic Continuous-Time Games,” Working paper. [27]
- KAMENICA, E. AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 2590–2615. [8]
- KARATZAS, I. AND S. SHREVE (1998): *Brownian Motion and Stochastic Cal-*

- culus*, vol. 113 of *Graduate Texts in Mathematics*, Springer, 2 ed. [26]
- KAYA, A. (2022): “Paying with Information,” Working paper. [5]
- LAZEAR, E. P. (1981): “Agency, Earnings Profiles, Productivity, and Hours Restrictions,” *American Economic Review*, 71, 606–620. [5]
- ORLOV, D., A. SKRZYPACZ, AND P. ZRYUMOV (2020): “Persuading the Principal to Wait,” *Journal of Political Economy*, 128, 2543–2578. [5]
- RAY, D. (2002): “The Time Structure of Self-Enforcing Agreements,” *Econometrica*, 70, 547–582. [5]
- RENAULT, J., E. SOLAN, AND N. VIEILLE (2017): “Optimal Dynamic Information Provision,” *Games and Economic Behavior*, 104, 329–349. [5]
- SANNIKOV, Y. (2008): “A Continuous-Time Version of the Principal–Agent Problem,” *Review of Economic Studies*, 75, 957–984. [27]
- SIMON, L. K. AND M. B. STINCHCOMBE (1989): “Extensive Form Games in Continuous Time: Pure Strategies,” *Econometrica*, 57, 1171–1214. [27]
- SMOLIN, A. (2021): “Dynamic Evaluation Design,” *American Economic Journal: Microeconomics*, 13, 300–331. [5]
- THOMAS, J. AND T. WORRALL (1994): “Foreign Direct Investment and the Risk of Expropriation,” *Review of Economic Studies*, 61, 81–108. [5]