

# Signaling with Private Monitoring\*

Gonzalo Cisternas and Aaron Kolb

November 13, 2023

## Abstract

We study dynamic signaling when the sender does not see the signals that her actions generate. The sender then uses her past play to forecast what a receiver believes, in turn forcing the receiver to forecast the previous forecast, and so forth. We identify a class of linear-quadratic-Gaussian games where this endogenous higher-order uncertainty can be handled. The *sender's second-order belief* is key: it is a private state that she controls, and it creates a new channel for information transmission. We examine the role of higher-order uncertainty and this new signaling channel in applications to macroeconomics, reputation, and trading: inflationary biases under discretion can be larger; career-concerned agents may benefit from not knowing their reputations; and informed trades can carry more price impact. We also introduce an existence method for boundary value problems that can be used in other dynamic games.

## 1 Introduction

The study of information transmission through actions has proven crucial for understanding phenomena as diverse as central bank ambiguity (Cukierman and Meltzer, 1986), reputation effects in industries (Milgrom and Roberts, 1982), and price discovery in financial markets (Kyle, 1985). Such *signaling games* have naturally progressed from settings where the actions of an informed “sender” are perfectly observed, to richer environments in which behavior is observed imperfectly through public signals with noise. Thus, a pervasive implicit assumption in this area of work is that the belief of the traditional “receiver” is commonly known: knowing what this player has seen, the sender is always certain of the inferences made.

---

\*Cisternas: Federal Reserve Bank of New York, [gonzalo.cisternas@ny.frb.org](mailto:gonzalo.cisternas@ny.frb.org). Kolb: Indiana University Kelley School of Business, 1309 E. Tenth St., Bloomington, IN 47405 [kolba@indiana.edu](mailto:kolba@indiana.edu). We thank Alessandro Bonatti, Marco del Negro, Wouter Dessein, Mehmet Ekmekci, Eduardo Faingold, Robert Gibbons, Nathan Kaplan, Marina Halac, Antoine Martin, Stephen Morris, Alessandro Pavan, Asani Sarkar, Andy Skrzypacz, Bruno Strulovici, Juuso Toikka, Vish Viswanathan, and numerous seminar audiences for useful conversations and comments. The views expressed in this paper are those of the authors and do not necessarily represent the position of the Federal Reserve Bank of New York or the Federal Reserve System.

This common knowledge breaks when there are private signals of behavior—and these signals can be important. In macroeconomics, private signals of aggregate variables can help explain the observed real effects of monetary shocks (Woodford, 2002). In the e-commerce industry, data brokers routinely collect imperfect signals of consumer behavior to secretly quantify unobserved consumer characteristics (Bonatti and Cisternas, 2020). In financial markets, certain traders can have a natural advantage in picking up signals of others’ trades (Yang and Zhu, 2020). Two fundamental questions emerge. How do informed individuals—monetary authorities, savvy consumers, sophisticated traders—behave when their actions generate private information available to others? Does it really matter in terms of economic outcomes that such senders do not know exactly what their relevant receivers believe?<sup>1</sup>

This paper offers the first framework for examining signaling games featuring private signals of actions. Before describing the model, let us use an example to motivate the core issue underlying settings with this feature. Consider a monetary authority (she) who has private information about the optimal level of inflation for an economy. The authority takes actions that transmit this information, but the signals observed by the private sector are noisy. As we argued, if those signals are *public*, the private sector’s belief is always known. But this means that the monetary authority disregards her past actions when forecasting such a belief: fixing any public history, counterfactually higher (lower) past actions only indicate that the shocks in the signal must have been smaller (larger) to generate the same realizations—which is what the private sector ultimately sees and uses to form its estimates.

Our point is not to argue that this type of phenomenon is implausible. Rather, that it seems less plausible than a situation in which the monetary authority reflects on her past behavior and, say, with the aid of additional data, tries to gauge what the private sector believes. This is exactly what would happen if the monetary authority ceases to observe the signals seen by the private sector: her history of play would regain statistical value, as higher past actions become indicative of higher beliefs by the private sector, and vice versa. This explicit use of past behavior to form *imperfect forecasts* of what others know, as natural as it seems, is a largely unexplored area within signaling games. Also, it is what distinguishes private monitoring from other simpler signal structures under which common knowledge may also break, such as when a receiver sees exogenous private signals of a sender’s type.

As the private sector receives signals of the sender’s actions, it forms a belief of the state of the economy. If the monetary authority cares about these estimates—say, because they shape inflation expectations—the monetary authority has to form a second-order belief: a belief about the private sector’s belief. As the authority relies on her past play to construct this forecast, and her actions carry her type, the private sector may now need to forecast

---

<sup>1</sup>We discuss the differences from these papers in the literature review section.

this belief. But the resulting third-order belief will be based on private signals again, and hence the forecasting problem gets restarted. In other words, attempting to enable signaling games to deliver imperfect estimates that non-trivially depend on past play leads to the classic problem of “forecasting the forecasts of others” (e.g., [Townsend, 1983](#)): but in our case, with players that have the power to individually affect the signals observed by others.

In the class of games that we study, a forward-looking sender (she) interacts with a receiver (he) over a finite horizon. Both players feature quadratic preferences, and they take actions continuously over time. Further, the sender has a time-invariant normally distributed type. All this is standard. The innovation, however, is that the receiver privately observes a signal that is linear in the sender’s action and that is distorted by additive Brownian noise. We close our baseline model with two extra ingredients. First, the receiver is myopic, an assumption meant to isolate how the sender’s behavior changes solely due to the emergence of higher-order beliefs—it can be relaxed, as we discuss in detail. Second, there is a public signal of the receiver’s action, also distorted by Brownian noise; this commonly observed signal serves as external data that the sender uses to learn about the receiver’s inferences.

We construct linear Markov equilibria (LMEs) in which, after all private histories, the following states are used linearly: the sender uses her type and the mean of her second-order belief, while the receiver uses the mean of her first-order belief; and both players also rely on a “public belief,” the mean of the belief about the sender’s type for someone who only observes the public signal (and assumes that the players follow their equilibrium strategies). The sender’s second-order belief always aggregates her past actions linearly, so it is her private information. Also, it is the only state of hers that she influences directly, paralleling the traditional control of a receiver’s commonly known belief in public settings. Thus, any equilibrium analysis requires establishing optimality with respect to this novel state.

The first contribution of the paper is to establish that no additional state variables are needed—the “beliefs about beliefs” problem is manageable. We do so via a *representation* of the sender’s second-order belief, along the path of play of linear Markov strategies, as a convex combination of the sender’s type and our public state—precisely reflecting estimates that combine internal (i.e., past play) and external (i.e., public) data. As the full-support monitoring makes deviations hidden, the receiver always believes that the representation holds, and the linear aggregation of histories in this linear-quadratic-Gaussian (LQG) world kicks in: via the representation, the receiver’s third-order belief combines his first-order belief and the public belief (hence the players’ use of the latter state), and iterating this logic up the hierarchy of beliefs reveals that our states are sufficient statistics (even after deviations, because these are hidden). Allowing for an *asymmetric*—namely, a private-public—monitoring structure is then key for making progress in this area of dynamic games.

Equipped with this finding, we can examine how the presence of higher-order uncertainty affects outcomes. A key fact is that, if we hold fixed the public signal's realizations, different types will now disagree about the receiver's belief: as different monetary authority types have behaved differently in the past, they will expect different inflation expectations by the market, even after seeing the *same data* about these expectations. This is captured by the second-order belief's dependence on the type in the representation; and it implies that there is separation—hence, information transmission—through the second-order belief channel.

In our leading example, a monetary authority evaluates how to set inflation to best balance her private information about the economy with her desire to stabilize output around a target. As is standard, the private sector is trying to forecast the authority's choice of inflation at all times, which now requires forecasting the authority's type. Meanwhile, the monetary authority must forecast the private sector's belief to determine how much inflation is needed to surprise the economy and affect employment. Thus, a distinctive feature relative to existing work is that here both parties are learning from each other as actions unfold.

Our tools enable us to compute the *inflationary bias* in the presence of higher-order uncertainty. This measure of wasteful inflation is lower than if the authority were myopic, but in general larger than if the private sector's beliefs were perfectly known (and the authority still forward-looking). In essence, our LME uncovers an attempt to dampen inflation expectations today relative to the myopic solution, and surprise the economy tomorrow at a lower cost; but the benefit of this dampening depends on how responsive inflation expectations are. The direct effect of higher-order uncertainty is to make these expectations more sluggish in responding to the authority's actions, in which case all types create more wasteful inflation.

Ex post, however, different types do influence output due to their private information. As higher types have chosen higher inflation, they expect higher expectations by the market. To stabilize output then, such higher types end up creating even more inflation; through the second-order belief channel, the separation of types is amplified. An interesting strategic effect arises due to this enhanced separation. A noisier public signal—which yields an environment with more higher-order uncertainty—induces a stronger reliance on past play in the authority's learning. As types separate more radically, there is more information transmission, which can eventually lead to *more responsive* market expectations and lower inflation: the notion that higher degrees of higher-order uncertainty always imply less responsive beliefs can break, because behavior is affected. The two-sided learning aspect of our model, coupled with the presence of players with the power to affect market variables, is key.

We explore our novel signaling channel in more depth in two other applications: reputation and trading. Succinctly, our reputation game shows that if a sender is concerned about maintaining a neutral reputation, she may prefer to be uncertain about how she is being per-

ceived, despite this impeding her ability to take the best actions to manage her reputation. The reason is that the new signaling channel can reduce the separation of types relative to the public case, and so less information may get transmitted in the first place. Our trading game instead shows that informed trades become costlier. Having traded more aggressively in the past, higher types think that their receivers have higher beliefs, and hence that they will buy more in the near future; this is superior information that can be exploited today. Since aggressive trading today reinforces aggressive trading tomorrow—thereby amplifying separation, and hence price impact—our trader slows down her purchases.

We conclude with the second technical contribution of the paper, which is to introduce a method for showing the existence of LMEs. The LQG structure naturally gives prominence to the means of posterior beliefs: these states aggregate the signals observed, and these signals are affected by actions. But posterior variances—capturing the players’ learning—are of great importance too, because they determine the sensitivity of posterior means to signal realizations, and hence they matter for the choice of coefficients attached to the belief states in the strategies. A complex feedback loop arises through this variance channel: the players’ conjectured strategy coefficients affect posterior variances, which in turn shape the evolution of the belief states, and thus ultimately affect the choice of coefficients themselves.

In this LQG world, establishing the existence of an LME equates to solving a key *boundary value problem* (BVP). This BVP consists of ordinary differential equations (ODEs) for the coefficients in the sender’s strategy, which are traced backward from the endgame by backward induction. But due to the feedback loop at play, these ODEs are coupled with two “learning” counterparts that are traced forward from initial values: one ODE for receiver’s posterior variance and another for the weight on the type in the representation, encoding the sender’s learning. The challenge here is the presence of multiple ODEs in both directions: existing work has dealt with settings in which only one learning ODE arises due to the players signaling and learning at the same rate. If the environment is asymmetric, and the players signal at different rates, the traditional methods used so far do not apply.

The feedback present hints at a fixed-point approach for proving that our BVP has a solution. What is perhaps less clear is how exactly to construct a fixed-point problem that is tractable and informative. We argue that an infinite-dimensional approach *over candidate solutions to the learning ODEs* is the best avenue. We discuss this method extensively: how it is a major step forward; how it effectively exploits the economics of the problem; and how it can be implemented in other (potentially asymmetric and not necessarily LQG) settings featuring a feedback among ODEs. Via this approach, we prove the existence of LMEs for horizon lengths up to threshold times with two key properties: they are inversely proportional to the prior variance about the sender’s type, but *independent of the discount*

*rate*. Thus, these times are conservative bounds and can be improved on a case-by-case basis.

We review the related literature next. Section 2 presents our leading application and the general model, and Section 3 covers the representation result and the sender’s best-response problem. Section 4 examines applications, and Section 5 studies the existence of LMEs. Section 6 discusses our existence method and assumptions. All proofs are in the Appendix.

**Related literature** Relative to two-player signaling games, our dynamic model features evolving beliefs that are not common knowledge and can explicitly depend on past play. These features are absent if signals of actions are noisy and public, as in Heinsalu (2018) and Ekmekci et al. (2022). While an explicit use of past play arises when actions are perfectly observed, sender’s estimates are usually perfect: in Kaya (2009), the sender’s actions are the only signal available, while in Kremer and Skrzypacz (2007), Daley and Green (2012) and Kolb (2015, 2019) observable actions are coupled with exogenous public signals of the type.

Private beliefs can nevertheless arise if there is multisided signaling—LQG models have proven useful in this area, provided the environment has sufficient public information or symmetry. For instance, Foster and Viswanathan (1996), Back et al. (2000), and Bonatti et al. (2017) examine settings with multiple informed agents, all of whom learn from a single imperfect public signal of behavior: past play is used to construct a “residual” signal about others’ behavior, and hence no higher-order beliefs are needed as states; further, since learning is symmetric, a multidimensional BVP never arises. Recently, Bonatti and Cisternas (2020) examine two-sided signaling when firms observe a private signal of a consumer’s history of past behavior; the prices firms set, however, fully reveal their beliefs.<sup>2</sup>

With private monitoring, players need to compute distributions over rivals’ histories to determine their actions. Not only do these histories grow over time, but the resulting distributions vary with a player’s own past behavior (the game’s structure changes after deviations; see Kandori, 2002). Constructing belief-dependent equilibria is then difficult, so past work—most of which features multi-sided private monitoring but no incomplete information—has taken different approaches, even looking for mixed-strategy equilibria where beliefs about histories are irrelevant (e.g., Ely and Välimäki, 2002). Belief-dependent equilibria do arise in Mailath and Morris (2002), where strategies take a *finite*-automaton form; players then form beliefs about those states, but beliefs depend on the observed private histories. Building on this, Phelan and Skrzypacz (2012) find equilibria by only looking at extreme beliefs of such states. We also “reduce” the inference problem—to a finite set of real-valued, evolving states—but we pin down the sender’s incentives at all values of her second-order belief. This latter state varies with the sender’s own past play, and the fact that it is spanned by the rest

---

<sup>2</sup>Private beliefs can also arise with *exogenous* private signals of a sender’s type (Feltovich et al., 2002; Cetemen and Margaria, 2020; Kolb et al., 2021), or if types exhibit correlation (e.g. Cetemen et al., 2023).

of the states only along the path of play reflects that the game changes after deviations.

Regarding our applications, the preferences in our monetary policy game coincide with those in the benchmark case in [Athey et al. \(2005\)](#), who use mechanism design to determine the optimal degree of discretion for a monetary authority; instead, we study discretionary policies, where private information is revealed gradually. More substantially, a distinctive feature of our model is the emergence of higher-order uncertainty from a monetary authority and a market learning about each other’s private beliefs: instead, in the games of [Cukierman and Meltzer \(1986\)](#) and [Faust and Svensson \(2001\)](#) only a market learns about the authority’s preferences, while in the decision problem of [Svensson and Woodford \(2004\)](#) it is the market that is fully informed. Other models featuring two-sided learning are [Svensson and Woodford \(2003\)](#) and [Cisternas \(2018\)](#), but the source of uncertainty are the economy’s fundamentals, which are unknown to everyone. Crucially, all these papers are LQG, and the signals seen by the relevant receivers are also public. In turn, private signals of behavior in macroeconomic models usually involve aggregate variables linked to infinitesimal agents: in [Woodford \(2002\)](#) firms observe private signals of nominal output, while in [Amador and Weill \(2012\)](#) agents see private signals of aggregate behavior when the information about an economy is dispersed.

Finally, on reputation, [Bouvard and Lévy \(2019\)](#) study a model with quadratic payoffs and symmetric Gaussian uncertainty in which beliefs are public in the linear equilibrium studied. And on trading, [Yang and Zhu \(2020\)](#) find that mixed-strategy equilibria can arise if there is leakage of an informed trader’s behavior; with only two rounds of trading, the problem of how a player’s own histories are aggregated to forecast a rival’s belief is absent.

## 2 Model and Leading Example

This section lays out the class of signaling games we study. To facilitate the exposition, we introduce a *monetary policy game* in the spirit of [Kydlan and Prescott \(1977\)](#) and [Barro and Gordon \(1983\)](#) as the leading application. We refer to this example to build intuition throughout our analysis, and we explore this and other applications in [Section 4](#).

### 2.1 Leading Example

**A canonical setting** We revisit the classic topic of output-inflation trade offs under discretion. A monetary authority interacts with a private sector (or market). We assume that the monetary authority experiences (ex post) flow payoffs that are proportional to

$$-(k + \hat{a}_t - a_t)^2 - (a_t - \theta)^2. \tag{1}$$

The variable  $\theta \in \mathbb{R}$  represents a time-invariant state of the economy that is privately observed by the monetary authority: e.g., a newly realized shock about whose magnitude the authority is superiorly informed. In turn, the variable  $a_t \in \mathbb{R}$  denotes the monetary authority's action at time  $t$ , which we interpret as the level of inflation at that instant. The last term in (1) then reflects the costs that different inflationary stimuli  $a_t$  can have given the new economic conditions perceived by the authority. On the other hand, the first term in (1) captures a desire to stabilize output around a target (normalized to zero) through a Phillips curve logic. That is, we assume that  $\hat{a}_t \in \mathbb{R}$  denotes the private sector's estimate of inflation  $a_t$ , while  $k$  is the natural rate of unemployment. Thus, the unemployment rate,  $k + \hat{a}_t - a_t$ , departs from its natural level in the opposite direction of unanticipated inflation.

Throughout this paper, the common prior is that  $\theta$  is normally distributed with mean  $\mu \in \mathbb{R}$  and variance  $\gamma^o > 0$ . Also, the monetary authority is forward-looking. And since inflation expectations summarize how market choices affect the authority, the private sector can be seen as a myopic player who, at all times, maximizes the expectation of

$$-(a_t - \hat{a}_t)^2. \tag{2}$$

These elements define an otherwise standard two-player game of asymmetric information. What is missing is a complete formulation of the players' information structure. What signals do the players see—in particular, the private sector? Before addressing this issue, it is useful to discuss the type of interdependence that arises in a game like this.

**Beliefs about beliefs** From (1)–(2) it is clear that the monetary authority will condition inflation on  $\theta$ . Because of this *type sensitivity*, the market will need to form a belief about  $\theta$  to forecast inflation. But the latter *first-order belief sensitivity* now forces the authority to form a belief about the private sector's belief to be able to move employment in the right direction (and magnitude) to stabilize output—a *second-order belief sensitivity* emerges.

Of course, this issue has no bite in a one-shot (e.g., simultaneous move) interaction. Concretely, first-order beliefs suffice to construct a linear equilibrium of the form

$$a = \frac{\theta + \mu}{2} + k \quad \text{and} \quad \hat{a} = \mu + k. \tag{3}$$

Key here is that the policymaker *knows what the market knows* when she acts—but what if this is not the case? For instance, agents in the economy may have imperfect knowledge of how monetary policy transmits to the economy. Moreover, such knowledge is likely idiosyncratic in nature. Inflation expectations then resemble a form of information held by the market, not readily available to the monetary authority, that she is trying to learn.



**A mixed public-private monitoring structure** To capture a situation like this, we let the private sector have access to a signal of inflation that is not available to the authority:

$$dY_t = a_t dt + \sigma_Y dZ_t^Y. \quad (4)$$

Here,  $Z^Y$  is a Brownian motion, while  $\sigma_Y$  is a volatility parameter. The latter is strictly positive so that  $Y$  indeed constitutes private information to the market: otherwise, the authority would know with certainty what the private sector has seen, and hence expects.<sup>3</sup>

In practice, monetary authorities periodically gather information about market expectations from a variety of sources. To capture this essential feature, we assume that there is a public signal of inflation expectations by the private sector, for example survey data:

$$dX_t = \hat{a}_t dt + \sigma_X dZ_t^X, \quad (5)$$

where  $Z^X \perp Z^Y$  is also a Brownian motion. The parameter  $\sigma_X$  is strictly positive again, which prevents the authority from perfectly learning what the market knows in real time.<sup>4</sup>

As the authority conditions her actions on  $\theta$ , she signals the new state of the economy to the market. At the same time, since she is uncertain about how her actions are interpreted, she ceases to know how the private sector’s belief has departed from its prior mean  $\mu$ . But estimating this belief is key to preventing output from becoming too destabilized, either because the private sector has misinterpreted the authority’s actions, or because of inadequate inflation surprises.<sup>5</sup> While the authority relies on (5) to learn about the market’s expectations, the problem is that her second-order belief will remain her private information—despite  $X$  being public. This means that the private sector will have to forecast this non-trivial belief; but because only the private sector observes  $Y$ , such estimates will be private, and the forecasting problem continues. The public nature of  $X$  allows us to successfully handle this “beliefs about beliefs” problem, in which the interdependence of the players’ actions is key.

**Economic implications** A credibility problem is the central tension in these policy games. Indeed, from the static *discretionary* equilibrium (3), unemployment takes value  $k$  after averaging across all possible types—i.e., it is unaffected from an ex ante perspective—while average inflation is  $\mu + k$ . Yet the *commitment* solution,  $a^c = \frac{\theta + \mu}{2}$ , achieves the same rate of average unemployment with an inflation level of just  $\mu$ . That is, the authority may want to

---

<sup>3</sup>Private signals of aggregate variables like (4) also arise in Woodford (2002), who stresses the difference between public information and information that individuals are actually aware of.

<sup>4</sup>The public signal could instead be about (true) employment,  $dX_t = (a_t - \hat{a}_t - k)dt + \sigma_X dZ_t$ , featuring both players’ actions. Section S.4 in the Supplementary Appendix extends our baseline model to this case.

<sup>5</sup>Central bank’s actions, or intentions of actions, may lead to unintended reactions by markets, even if the goal is not to generate a surprise. See Fisher (2017) on the well-known “taper tantrum” in 2013.

commit to a lower level of inflation, but this is ex post suboptimal.<sup>6</sup> How does the monetary authority signal its superior information while managing the market’s expectations as both parties embark on learning from each other? How do the incentives to create surprise inflation play out, and how do they relate to the classic credibility problem? In terms of inflation, does it matter that the authority does not know exactly what the market has seen?

## 2.2 A Class of Games

Our general class of games maintains the previous signal structure, but allows for general quadratic versions of (1)–(2).<sup>7</sup> A “sender” (she; e.g., a monetary authority) interacts with a “receiver” (he; e.g., a private sector) continuously over a time interval  $[0, T]$ ,  $T < \infty$ . The sender has a fixed private type  $\theta$  whose distribution  $\mathcal{N}(\mu, \gamma^\theta)$  is common knowledge. The players’ actions— $a_t$  for the sender and  $\hat{a}_t$  for the receiver—take values in  $\mathbb{R}$ . The sender is forward-looking: given action paths  $(a_t)_{t \in [0, T]}$  and  $(\hat{a}_t)_{t \in [0, T]}$ , her ex post payoff is given by

$$\int_0^T e^{-rt} u(a_t, \hat{a}_t, \theta) dt + e^{-rT} \psi(\hat{a}_T), \quad (6)$$

where  $r \geq 0$  is a discount rate and  $u : \mathbb{R}^3 \rightarrow \mathbb{R}$  and  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  are quadratic functions. For our leading application,  $u(a, \hat{a}, \theta) \propto -(k + \hat{a} - a)^2 - (a - \theta)^2$ , while  $\psi \equiv 0$ . Later, in Section 4.2, we examine a game with a non-trivial lump-sum terminal payoff  $\psi(\cdot)$ .

The receiver is assumed to be myopic, and thus concerned only about maximizing his flow utility at all instants of time. This player’s ex post time- $t$  payoff is denoted

$$\hat{u}(a_t, \hat{a}_t, \theta), \quad (7)$$

with  $\hat{u} : \mathbb{R}^3 \rightarrow \mathbb{R}$  quadratic. In our leading example,  $\hat{u}(a, \hat{a}, \theta) \propto -(a - \hat{a})^2$ , and so the private sector solves a *prediction problem* at all times. We discuss this myopia assumption and its connection with prediction problems at the end of this subsection (Remark 1).

We now turn to the specific conditions on  $(u, \hat{u}, \psi)$ ; subscripts on these functions denote partial derivatives. We begin with two concavity requirements. First, we assume  $u_{aa} = \hat{u}_{\hat{a}\hat{a}} = -1$ . The negative value ensures that the players’ best responses are well defined given the unbounded action space. That the value is  $-1$  simply amounts to a normalization of the players’ payoffs, which simplifies our calculations (in the monetary example, for instance, (1) is scaled by  $1/4$ , and (2) by  $1/2$ ). Second, we assume that  $\psi_{\hat{a}\hat{a}} \leq 0$ : this restriction is driven by the applications that we study, which include the case  $\psi \equiv 0$  as in our leading example.

<sup>6</sup>Hence, (3) is “discretionary.” See Supplementary Appendix Section S.1.6 for the commitment solution.

<sup>7</sup>This generalization is not exhaustive: see Section 4.3 for a game with a more general public signal.

The next conditions create sufficient interdependence in the players' actions, and hence a need for higher-order beliefs; they generalize our "sensitivity" notions from Section 2.1:

**Assumption 1.** (i)  $u_{a\theta} \neq 0$  (type sensitivity); (ii)  $|\hat{u}_{a\theta}| + |\hat{u}_{a\hat{a}}| \neq 0$  (first-order belief sensitivity); and (iii)  $|u_{a\hat{a}}| + |u_{\hat{a}\hat{a}}| + |\psi_{\hat{a}\hat{a}}| \neq 0$  (second-order belief sensitivity).

Part (i) is needed for there to be any signaling, and (ii) for the receiver's action to be sensitive to his private belief. This happens when the receiver cares about the type directly, encoded in  $\hat{u}_{a\theta} \neq 0$ , or when he does so indirectly through the sender's action,  $\hat{u}_{a\hat{a}} \neq 0$  (as in our leading example). Part (iii) then guarantees that the sender's behavior will explicitly depend on her forecast of the receiver's belief. This happens when her utility exhibits a non-trivial strategic interaction,  $u_{a\hat{a}} \neq 0$  (as in our leading application, through unemployment). But it also happens when  $u$  or  $\psi$  is nonlinear in the receiver's action ( $|u_{\hat{a}\hat{a}}| + |\psi_{\hat{a}\hat{a}}| \neq 0$ ).

Two observations about (iii) are in order. First, if this assumption is not imposed, the players' equilibrium strategies will depend only on first-order beliefs, and the economic insights are rendered standard (see Proposition 1 in the next Section).<sup>8</sup> Second, related to this point, in Section 5 we complement (iii) with mild technical conditions that ensure that our equilibrium always exhibits information transmission through the second-order belief.

The receiver privately observes  $Y$  carrying the sender's action as in (4). There is also a public signal  $X$  of the receiver's action, given by (5). Altogether, the linear dynamics (4)–(5), quadratic preferences (6)–(7), and Gaussian randomness  $(\theta, Z^Y, Z^X)$  constitute our class of LQG games. In what follows, we let  $\mathbb{E}_t[\cdot]$  and  $\hat{\mathbb{E}}_t[\cdot]$  denote the players' conditional expectations for  $t \in [0, T]$ , retaining  $\mathbb{E}[\cdot]$  for computing averages from an ex ante perspective.

**Remark 1.** *The receiver's myopia is conceptually useful for isolating the impact of a private signal on the sender's behavior. While it may seem special, it is not a major limitation for three reasons. First, we set up the sender's best-response problem assuming a general receiver. Second, if the receiver solves a prediction problem, the same equilibrium that we find also arises if he is forward-looking (Proposition 8). Third, in Appendix D, we show how our method for existence can be adapted to this case without major conceptual adjustments.*

**Strategies and Equilibrium.** Towards our solution concept, note that the noise terms  $(Z^X, Z^Y)$  have full support, so the players' actions are hidden from their counterparties. This means that the only off-path histories for any player are those in which that same individual has deviated, and hence that imposing sequential rationality does not refine the

---

<sup>8</sup>While the sender still needs to take an expectation over the receiver's belief, actions—on or off the path of play—do not depend on it in the equilibrium studied. This is because the model becomes largely linear.

set of equilibrium outcomes.<sup>9</sup> At this stage of our analysis then, it is without loss to use the Nash equilibrium concept, and thus leave behavior after deviations unspecified for now.

From this perspective, a pure strategy for the sender specifies, at any time  $t$ , an action  $a_t \in \mathbb{R}$  as a function of the history  $(\theta, (X_s)_{s < t})$ . For the receiver,  $\hat{a}_t \in \mathbb{R}$  in turn conditions on the history  $((X_s)_{s < t}, (Y_s)_{s < t})$ ,  $t \in [0, T]$ .<sup>10</sup> Due to the Brownian information, some extra regularity is required: a strategy profile  $(a_t, \hat{a}_t)_{t \in [0, T]}$  is *admissible* if each component is square integrable (so payoffs are well-defined given  $(X, Y)$ ) and (4)–(5) admits a unique solution (so the realizations of  $(X, Y)$  are well-defined). We only consider such (pure) strategy profiles.<sup>11</sup>

**Definition 1** (Nash equilibrium). *A strategy profile  $(a_t, \hat{a}_t)_{t \geq 0}$  is a Nash equilibrium if: (i)  $(a_t)_{t \in [0, T]}$  maximizes  $\mathbb{E}_0 \left[ \int_0^T e^{-rt} u(a_t, \hat{a}_t, \theta) dt + e^{-rT} \psi(\hat{a}_T) \right]$ ; and (ii) for each  $t \in [0, T]$ ,  $\hat{a}_t$  maximizes  $\hat{\mathbb{E}}_t[\hat{u}(a_t, \hat{a}_t, \theta)]$  when  $(\hat{a}_s)_{s < t}$  has been followed.*

Given the LQG structure, it is natural to look for Nash equilibria in strategies that are *linear* functions of the signals observed by each player. This task, however, is not as simple as it seems. The difficulty is that the sender will need to evaluate deviations, and at those off-path histories she will condition on more information than  $(\theta, X)$ : for example, a monetary authority of any given type will behave differently depending on how aggressively she has created inflation, even if the data about inflation expectations observed (the history of the public signal  $X$ ) is the same. This issue of behavior non-trivially depending on past play is salient in our games in that allowing for such an endogenous history-dependence is essential for being able to find equilibria. As we show, however, this issue has a parsimonious resolution when the players think that their counterparties are using belief states linearly.

## 3 Equilibrium Analysis: Linear Markov Equilibria

### 3.1 Belief States: Overview and Intuition

**Linear Markov Strategies** Let us offer a brief overview of our construction. With incomplete information, we can look for equilibria where the signals observed by the players are aggregated into beliefs linked to the sender’s type. Further, due to the quadratic preferences,

<sup>9</sup>See Mailath and Samuelson (2006), pp. 395–396. With hidden actions, a Nash equilibrium fails to be sequentially rational only if it dictates suboptimal behavior for a player after her own deviation. Since such off-path histories are not reached, the same outcome arises if optimal behavior is specified after the deviation.

<sup>10</sup>Formally, pure strategies are real-valued stochastic processes that are *progressively measurable* with respect to the filtrations of the stochastic processes observed. See Chapter 1 in Karatzas and Shreve, 1991.

<sup>11</sup>Square integrability refers to  $\int_0^T a_t^2 dt$  and  $\int_0^T \hat{a}_t^2 dt$  being finite in expectation (and thus also the sender’s payoff). Our equilibrium will satisfy that (4)–(5) admits a unique solution in a strong (pathwise) sense, which ensures a unique solution in a weak (probability law) sense: there is a unique distribution on  $C([0, T])^2$ —the set of paths of  $(X, Y)$ —equipped with the Borelians, that is consistent with the players’ strategies.

the means of such posterior beliefs—henceforth, *beliefs*—will act as state variables that are used linearly. In addition to the sender’s type  $\theta$ , therefore, we will use the following states:

$$\hat{M}_t := \hat{\mathbb{E}}_t[\theta], \quad M_t := \mathbb{E}_t[\hat{M}_t], \quad \text{and} \quad L_t := \mathbb{E}[\theta | \mathcal{F}_t^X]. \quad (8)$$

Here,  $\hat{M}_t$  is the receiver’s belief about  $\theta$ , which uses the private observations of  $Y$  up to  $t$ . In turn,  $M_t$  is the sender’s second-order belief, i.e., her belief about the receiver’s first-order belief. Finally,  $L_t$  is the belief about  $\theta$  for someone who only observes the public signal  $X$ .

Recall our leading example,  $u(a, \hat{a}, \theta) \propto -(k + \hat{a} - a)^2 - (a - \theta)^2$  and  $\hat{u}(a, \hat{a}, \theta) \propto -(a - \hat{a})^2$ . The use of  $\theta$  by the authority, and of  $\hat{M}$  by the private sector, are obvious. The authority will then condition her actions on  $M$  due to output having an ideal target. For example, if the market expects higher types to generate more inflation ( $\hat{a}$  increases with  $\hat{M}$ ), an authority who perceives high values of  $\hat{M}$  will fear a drop in employment. This perception, encoded in a high value of  $M$ , will lead to inflationary stimuli to stabilize output. The presence of the “public state”  $L$  relates to the market having to forecast  $M$ , discussed shortly.

These beliefs obviously depend on the strategies used by the players. Our focus is on equilibria in which, on and off the path of play, the players use *linear Markov strategies*:

$$a_t = \beta_{0t} + \beta_{1t}M_t + \beta_{2t}L_t + \beta_{3t}\theta \quad (9)$$

$$\hat{a}_t = \delta_{0t} + \delta_{1t}\hat{M}_t + \delta_{2t}L_t. \quad (10)$$

The coefficients  $\beta_{it}$ ,  $i = 0, 1, 2, 3$ , and  $\delta_{jt}$ ,  $j = 0, 1, 2$ , will be differentiable functions of time, a dependence that must be allowed to capture end-game and learning effects. For example, returning to the credibility problem, the monetary authority may want to take actions to be perceived as a low type today, and then surprise the market with inflation tomorrow; but the profitability of this depends on how much time is left to enjoy higher employment.

**The sender’s second-order belief** Before entering the technical analysis, it is useful to elaborate on the state  $M$ . The starting point is that it will not only depend on the history of the public signal  $X$  (which carries the receiver’s action). Due to the private monitoring, it will also depend explicitly on the sender’s past *past behavior*; higher past actions indicate higher realizations of  $Y$ , so  $M$  is higher for any fixed public history of  $X$ . Intuitively, this captures a monetary authority who actively reflects on her past choices and, combined with data about inflation expectations, decides how to act. By contrast, if  $Y$  were public, the common knowledge of the receiver’s observations implies that  $M \equiv \hat{M}$ , i.e., beliefs are fully determined by the histories of  $Y$ —the authority can disregard her past actions to (perfectly) forecast inflation expectations, because she knows what the market has seen. This distinction

is important, as it has signaling implications. Concretely, since the sender’s actions depend on her type, and  $M$  non-trivially aggregates past actions if  $Y$  is not public, a link between  $M$  and  $\theta$  emerges in equilibrium. That is, there is signaling through the second-order belief.

To analyze these games, therefore, we need two different “expressions” for  $M$ . First, we need to know how  $M$  looks along the path of play of (9)–(10). This is because the receiver must anticipate the total informativeness of the sender’s actions in equilibrium ( $M$  included) to be able to correctly interpret  $Y$  and form his belief. Section 3.2 establishes such a *representation* of  $M$ , which is a backward-looking expression for this state. Second, to find an optimal strategy for the sender (and hence, the equilibrium coefficients in (9)–(10)), we need a forward-looking expression for  $M$  under deviations from (9): a law of motion for how the sender perceives her actions will affect  $\hat{M}$  for arbitrary continuation strategies. This step is needed to be able to set up the sender’s best-response problem in Section 3.3.

We note that the emergence of a second-order belief non-trivially affecting outcomes only happens if there is sufficient strategic interdependence, as captured by Assumption 1(iii).

**Proposition 1.** *Suppose that  $u_{a\hat{a}} = u_{\hat{a}a} = \psi_{\hat{a}a} = 0$ . In any equilibrium of the form (9)–(10),  $\beta_1 \equiv \beta_2 \equiv \delta_2 \equiv 0$ : at all histories, the sender’s action is affine in  $\theta$  and the receiver’s affine in  $\hat{M}$ . The strategies are also equilibria if  $Y$  is public (hence, they are independent of  $\sigma_X$ ).*

As an example, suppose that the authority’s payoff is linear in employment, i.e.,  $u(a, \hat{a}, \theta) = -(k + \hat{a} - a) - (a - \theta)^2$  (i.e., there are no losses from overheating the economy). While the authority has to construct a second-order belief to forecast  $\hat{a}$ , her flow payoff is linear in that state. Inflation surprises then trigger an impulse response of employment that is independent of the level that  $M$  takes: the states  $M$  and  $L$  are never used, so information transmission is as usual; and changing  $\sigma_X$ , or even making  $\hat{M}$  public, is irrelevant for outcomes. Away from this case, the receiver will have to non-trivially forecast  $M$ —the “beliefs about beliefs problem” is then at play, and the need for the public state  $L$  will arise, as we explain next.

## 3.2 Representation of the Second-Order Belief

**Representation** Suppose that the players follow the linear Markov strategies (9)–(10). Heuristically, given the LQG structure, it is natural to conjecture the representation

$$M_t = \chi_t \theta + (1 - \chi_t) L_t, \tag{11}$$

where  $L_t := \mathbb{E}[\theta | \mathcal{F}_t^X]$  and  $(\chi_t)_{t \in [0, T]}$  is deterministic. Intuitively, to forecast the receiver’s belief, the sender takes its public estimate  $\mathbb{E}[\hat{M}_t | \mathcal{F}_t^X]$ —which coincides with  $\mathbb{E}[\theta | \mathcal{F}_t^X]$  by the law of iterated expectations—and adjusts it based on her own private information stemming

from her past actions, which carry  $\theta$  under (9). Further, with Gaussian learning, one expects this adjustment to be linear and deterministic, encoded in the weight  $\chi$ .

The coefficient  $\chi$  is a measure of the sender's learning about the receiver's belief (relative to someone who only observes the public signal). Because  $\chi$  is linked to past play, it is naturally connected to the sender's past signaling. To quantify total information transmission, we insert (11) into (9), which yields the sender's actions along the path of (9)–(10),

$$a_t = \alpha_{0t} + \alpha_{2t}L_t + \alpha_{3t}\theta, \quad (12)$$

$$\text{where } \alpha_{0t} := \beta_{0t}, \quad \alpha_{2t} := \beta_{2t} + \beta_{1t}(1 - \chi_t), \quad \text{and} \quad \alpha_{3t} := \beta_{3t} + \beta_{1t}\chi_t. \quad (13)$$

The total weight on the type,  $\alpha_3$ , is the sender's *signaling coefficient*, which is modified by  $\beta_1\chi$  due to the higher-order uncertainty. Naturally, as more signaling takes place and the sender expects the receiver to have learned more about  $\theta$ ,  $\chi$  must increase in weight. But how exactly does  $\chi$  depend on past values of  $\alpha_3$ ? And is  $\alpha_3$  the only input to  $\chi$ ?

Our first contribution is to characterize the dependence of  $\chi$  on the coefficients in the strategies (9)–(10) via a *system of ODEs*. To this end, we introduce the posterior variance

$$\gamma_t := \hat{\mathbb{E}}_t[(\theta - \hat{M}_t)^2], \quad t \in [0, T],$$

which is a measure of the receiver's learning. (We have omitted the hat symbol for notational convenience. An expression for  $\hat{M}$  is presented in (A.1) in the appendix.) Our approach is constructive in that we assume that  $L$  in (9)–(11) is a generic state that depends only on the histories of  $X$  and time and confirm that it corresponds to  $\mathbb{E}[\theta|\mathcal{F}_t^X]$  under (9)–(10).<sup>12</sup>

**Lemma 1.** *Suppose that  $(X, Y)$  is driven by (9)–(10) and that the receiver conjectures (11), where  $(L_t)_{t \in [0, T]}$  throughout is a process that depends only on the public information. Then  $M$  satisfies (11) if and only if:  $L_t \equiv \mathbb{E}[\theta|\mathcal{F}_t^X]$  under (9)–(10);  $\chi_t = \mathbb{E}_t[(\hat{M}_t - M_t)^2]/\gamma_t$ ; and*

$$\dot{\gamma}_t = -\frac{\gamma_t^2(\beta_{3t} + \beta_{1t}\chi_t)^2}{\sigma_Y^2}, \quad \gamma_0 = \gamma^\circ, \quad (14)$$

$$\dot{\chi}_t = \frac{\gamma_t(\beta_{3t} + \beta_{1t}\chi_t)^2(1 - \chi_t)}{\sigma_Y^2} - \frac{\gamma_t\chi_t^2\delta_{1t}^2}{\sigma_X^2}, \quad \chi_0 = 0. \quad (15)$$

Also,  $0 < \gamma_t \leq \gamma^\circ$  and  $0 \leq \chi_t < 1$ ,  $t \in [0, T]$ , with strict inequalities over  $(0, T]$  if  $\beta_{3,0} \neq 0$ .<sup>13</sup>

<sup>12</sup>Briefly, the approach is as follows. Given (12) (with  $L$  general), the problem of learning  $\theta$  from  $Y$  is (conditionally) Gaussian (Liptser and Shiryaev, 1977, Theorems 12.6 and 12.7); a stochastic mean  $(\hat{M}_t)_{t \in [0, T]}$  and a deterministic variance  $(\gamma_t)_{t \in [0, T]}$  then emerge. The sender's problem of filtering  $\hat{M}$  using  $X$  under (10) is (conditionally) Gaussian again; here, the mean  $M$  will shift with the sender's history of play for any given history of  $X$ . Imposing that  $M$  coincides with (11) when (9) is followed allows us to pin down  $(\chi, L)$ .

<sup>13</sup>A law of motion for  $(L_t)_{t \in [0, T]}$  is presented in the next section. See also (A.10) in the Appendix.

By the lemma, the weight  $\chi$  corresponds to the ratio of the players’ posterior variances. This is not surprising, as the players’ learning is necessarily connected: if the sender has signaled her type more aggressively, she will expect the receiver to be more certain about it, so lower values of  $\gamma$  are associated with higher values of  $\chi$ . Mathematically, the system of ODEs (14)–(15) is fully coupled ( $\chi$  affects  $\dot{\gamma}$  while  $\gamma$  affects  $\dot{\chi}$ ). Also, higher values of the signaling coefficient  $\alpha_3 = \beta_3 + \beta_1\chi$  prompt a faster decay of  $\gamma$  and a faster growth of  $\chi$ .<sup>14</sup>

The system is just a local representation of how the *learning coefficients*  $\gamma$  and  $\chi$  are affected by past values of the coefficients guiding information transmission—in the particular case of the sender, the weights  $\beta_3$  and  $\beta_1$  on  $\theta$  and  $M$ , respectively, via  $\alpha_3$ . At  $t = 0$ , for instance, there is no higher-order uncertainty:  $M = \hat{M} = L = \mu$ , and so  $\chi_0 = 0$ . But if  $\beta_{3,0} = \alpha_{3,0} \neq 0$  at that instant—i.e., some signaling takes place then—we have that  $\gamma_t < \gamma^o$  and  $0 < \chi_t$  at future times, reflecting that some learning has occurred. Conversely,  $0 < \gamma_t$  and  $\chi_t < 1$  capture that, with finite signaling coefficients, learning is never complete.

Finally, the last term in (15) shows that  $\delta_1$ , the receiver’s signaling coefficient, is another input to  $\chi$ . Indeed, higher values of  $\delta_1$  improve the informativeness of the public signal, so the sender favors the public state  $L$  over her past history of play in her forecasting exercise: in the ODE (15), a higher signal-to-noise ratio  $\delta_{1t}^2/\sigma_X^2$  puts more downward pressure on  $\chi$ .

**Sufficient statistics** Our states are sufficient statistics if the players think that the counterparty is using them linearly. The key is that deviations are hidden. In particular, the receiver always believes that the sender is on path. By construction, the receiver assumes that the representation (11) holds when he has not deviated. But even if he deviates, he only affects the realizations of  $L$ , which is a function of the commonly observed signal  $X$  exclusively; the receiver then still believes that  $M$  is as in (11), given the current value of  $L$ . As the receiver uses  $L$ , via (11), to forecast  $M$ , the former public state is payoff relevant.

Thus, the sender thinks that the representation always holds from the *receiver’s perspective*—even if the sender has deviated because, again, deviations are hidden. This means that there is common knowledge that the receiver’s third-order belief combines the current values of  $\hat{M}$  and  $L$  linearly, on and off the path of play. The sender’s fourth-order belief will always be a linear combination between  $M$  and  $L$ , and we can continue inductively: given any history of play,  $(t, \theta, \hat{M}, M, L)$  summarize all the higher-order beliefs for our players.

All told, the representation is key for guaranteeing that the state space is “closed” when each player tries to forecast the other’s private histories. To find the coefficients associated with those states, however, our players have to optimize at those histories. This is the goal of the next section, where we examine  $M$  under arbitrary strategies.

---

<sup>14</sup>For instance, if  $\sigma_X = +\infty$  (the public signal is uninformative), the solution to (14)–(15) satisfies  $\chi = 1 - \frac{\gamma_t}{\gamma^o}$ , so  $\gamma$  and  $\chi$  are inversely related—see Lemma B.1 in the appendix, and Section 6 for a generalization.



### 3.3 The Sender's Best-Response Problem

Assessing the optimality of (9) requires the sender to evaluate deviations. Unlike with the receiver, the representation need not hold from the *sender's perspective* any more, because  $M$  explicitly depends on her past actions (Remark 2 below). At those off-path histories, her perception of the continuation game changes: a deviation resulting in a different  $M$  than in the representation for the same  $L$  leads her to behave differently, so the sender keeps track of  $M$  and  $L$  separately. Below are dynamics for  $(M, L)$  under arbitrary sender strategies.

**Lemma 2.** *From the sender's perspective, if she follows  $(a'_t)_{t \in [0, T]}$ ,*

$$dM_t = \frac{\gamma_t \alpha_{3t}}{\sigma_Y^2} (a'_t - [\alpha_{0t} + \alpha_{2t} L_t + \alpha_{3t} M_t]) dt + \frac{\gamma_t \chi_t \delta_{1t}}{\sigma_X} dZ_t \quad (16)$$

$$dL_t = \frac{\gamma_t^X \chi_t \delta_{1t}}{\sigma_X^2} [\delta_{1t} (M_t - L_t) dt + \sigma_X dZ_t], \quad (17)$$

where  $Z_t := \frac{1}{\sigma_X} [X_t - \int_0^t (\delta_{0s} + \delta_{1s} M_s + \delta_{2s} L_s) ds]$  is a Brownian motion, and  $\gamma_t^X := \frac{\gamma_t}{(1-\chi_t)} = \mathbb{E}[(\theta - L_t)^2 | \mathcal{F}_t^X]$ . Also,  $\mathbb{E}_t[(\hat{M}_t - M_t)^2] = \gamma_t \chi_t$  for any such  $(a'_t)_{t \in [0, T]}$ .

The importance of the second-order belief state is clear from (16)–(17): the sender's actions only enter  $M$ , so finding an equilibrium is inevitably linked to imposing optimality with respect to this state. This justifies the extended strategy (9) involving both  $M$  and  $L$ .

The law of motion of  $M$ , (16), encapsulates how the sender expects the receiver's private belief  $\hat{M}$  to evolve in response to different continuation strategies by the sender. On the other hand, changes in  $L$  matter for the sender's incentives because the receiver uses this state to predict  $M$  in his third-order belief exercise. To understand why  $M$  feeds into  $L$  in (17), suppose that the sender has taken “high” actions in the past: expecting high values of the receiver's belief, the sender predicts a gradual rise in  $L$  through the channel of the receiver's actions influencing  $X$ . Our applications will discuss this channel in more depth.

Finally, note that (16)–(17) depend on the learning coefficients  $(\gamma, \chi)$ . This originates from the learning of the receiver who, as argued, uses the representation to form his mean-variance pair  $(\hat{M}, \gamma)$ . Two observations are in order. First, the appearance of  $\chi$  makes it clear that having upfront knowledge of a representation is necessary for setting up a best-response problem. Second, while the importance of posterior means is obvious, posterior variances are critical too. The reason is that much of the complexity behind the fixed point at play in these games (Section 5) is due to this variance channel: while  $(\gamma, \chi)$  depends on the signaling coefficients in (9)–(10), this pair also shapes the responsiveness of  $(M, L)$  to the sender's actions, and so  $(\gamma, \chi)$  affects the choice of coefficients in the strategies themselves.<sup>15</sup>

<sup>15</sup>The responsiveness of  $(M, L)$  also depends on the signaling coefficients:  $\alpha_3$  enters  $M$ , while  $\delta_1$  affects  $L$ .

**Remark 2** ( $M$  and  $L$  at all histories). *Inserting the definition of  $Z_t$  in the lemma into (16) yields a dynamic for  $M$  that is linear in the same variable; its solution  $M_t$  is linear in  $(a_s, L_s, X_s)_{s < t}$ . But inserting the same expression for  $Z_t$  in (17) then yields that  $L_s$  is a linear function of  $(X_\tau)_{\tau < s}$  exclusively. Hence,  $M$  is linear in the latter history and past play.*

Towards the sender's best-response problem, recall that the receiver's action  $\hat{a} = \delta_0 + \delta_1 \hat{M} + \delta_2 L$  enters  $(u, \psi)$ , while  $\hat{M}$  is hidden from the sender. However, we can replace  $\hat{M}$  by  $M$  and optimize: the key is that the sender's posterior variance,  $\mathbb{E}_t[(\hat{M}_t - M_t)^2]$ , is invariant to her deviations in this LQG world (last part of Lemma 2), i.e., changes in the sender's actions only shift the receiver's belief.<sup>16</sup> Up to an additive constant, the sender maximizes

$$\mathbb{E}_0 \left[ \int_0^T e^{-rt} u(a_t, \delta_{0t} + \delta_{1t} M_t + \delta_{2t} L_t, \theta) dt + e^{-rT} \psi(\delta_{0T} + \delta_{1T} M_T + \delta_{2T} L_T) \right] \quad (18)$$

subject to the dynamics (16)–(17) of  $(M, L)$ , and where  $(\gamma, \chi)$  follow the ODEs (14)–(15).<sup>17</sup> Notably, the challenge of endogenous history dependence discussed in Section 2—that performing equilibrium analysis requires allowing for deviations, but those deviations non-trivially affect future play (through the inferences a player makes)—has a parsimonious resolution here: one needs to look at the sender's behavior at arbitrary values of  $(M, L)$ .<sup>18</sup>

The coefficients  $(\delta_0, \delta_1, \delta_2)$  in the receiver's strategy (10) are fully general thus far, so the same steps apply irrespective of the receiver's time preference. In what follows, we focus on a myopic receiver—the forward-looking case is discussed in Section 6. From this perspective, a tuple  $\vec{\beta} := (\beta_0, \beta_1, \beta_2, \beta_3)$  induces a *linear Markov equilibrium* (LME) if  $\beta_{0t} + \beta_{1t} M + \beta_{2t} L + \beta_{3t} \theta$  is an optimal policy for the sender when  $(\delta_0, \delta_1, \delta_2)$  satisfy the myopic-best reply condition

$$\hat{a}_t := \delta_{0t} + \delta_{1t} \hat{M}_t + \delta_{2t} L_t = \arg \max_{\hat{a}' \in \mathbb{R}} \mathbb{E}_t[\hat{u}(\alpha_{0t} + \alpha_{2t} L_t + \alpha_{3t} \theta, \hat{a}', \theta)] \quad (19)$$

for the receiver. Note that our notion of LME is *perfect* in that it specifies optimal behavior after deviations. Further, along the path of the policy,  $a_t = \alpha_{0t} + \alpha_{2t} L_t + \alpha_{3t} \theta$  by construction, as required by a (linear) Nash equilibrium.<sup>19</sup> We discuss the LMEs that arise in our applications next; the question of existence of LMEs is relegated to Section 5.

---

$M$  also responds to the sender's action in proportion to  $\gamma$  (inherited from  $\hat{M}$ ), and it is revised upward when  $a'_t > \mathbb{E}_t[\alpha_{0t} + \alpha_{2t} L_t + \alpha_{3t} \hat{M}_t]$ , i.e., when  $a_t > \mathbb{E}_t[\mathbb{E}_t[a_t]]$ . As for  $L$ , it is the variance of  $\theta$  given the public information,  $\gamma^X$ , that matters via  $\text{Cov}(\theta, dX_t | \mathcal{F}_t^X) = \delta_1 \gamma_t^X$ . By Lemma 1,  $\chi < 1$ , so  $L$  is always well-defined.

<sup>16</sup>Formally,  $\mathbb{E}_t[u(a_t, \hat{a}_t, \theta)] = \mathbb{E}_t[u(a_t, \delta_{0t} + \delta_{1t} M_t + \delta_{2t} L_t, \theta)] + \frac{1}{2} u_{\hat{a}\hat{a}} \delta_{1t}^2 \gamma_t \chi_t$ , and likewise for  $\psi$  at  $t = T$ .

<sup>17</sup>See the proof of Lemma 2 for the connection with the so-called *separation principle* in *decision problems* with unobserved controlled states (here,  $\hat{M}$ ): namely, the ability to separate estimation from optimization.

<sup>18</sup>The admissible strategies for this problem are the  $\mathbb{R}$ -valued, square-integrable processes  $(a_t)_{t \in [0, T]}$  that are  $(\theta, M, L)$ -progressively measurable; see Pham (2009) and the proof of Lemma 2.

<sup>19</sup>It is easy to see that sequential rationality is trivially satisfied for the receiver, even if forward-looking.

## 4 Applications

We examine three applications: our monetary policy game from Section 2.1; a reputation game motivated by career concerns; and a trading game using an extension of our methods.<sup>20</sup>

A central issue is how key equilibrium outcomes change due to the higher-order uncertainty present. To this end, we sometimes compare the LMEs found to those that arise when  $Y$  is public, where  $\hat{M}$  ( $\equiv M$ ) is the only evolving belief state used by the players. LMEs in these public cases can be computed within our model simply by setting  $\sigma_X = 0$  in the public signal, as observing the receiver’s action can allow the sender to recover  $\hat{M}$ . Further, the signaling coefficient  $\alpha_3$  simplifies to  $\beta_3$ : since all sender types agree on the receiver’s belief at all histories of  $Y$ , there is no variation in behavior across types through the belief channel.<sup>21</sup>

### 4.1 Monetary Policy Game Revisited

From Section 2, the relevant payoffs for our players—up to positive factors—are

$$\underbrace{\int_0^T e^{-rt} \{-(k + \hat{a}_t - a_t)^2 - (a_t - \theta)^2\} dt}_{\text{monetary authority}} \quad \text{and} \quad \underbrace{\hat{u}(a_t, \hat{a}_t, \theta) = -(\hat{a}_t - a_t)^2}_{\text{private sector}},$$

where  $a_t$  is inflation,  $\hat{a}_t$  the market’s forecast of it, and  $k + \hat{a}_t - a_t$  the unemployment rate.

It is immediate that the monetary authority cannot affect output on average (i.e.,  $\mathbb{E}[k + a_t - \hat{a}_t] \equiv k$ ); this is due to  $\hat{a}_t = \hat{\mathbb{E}}_t[a_t]$  and the law of iterated expectations. But since the authority’s equilibrium strategy is sequentially rational—i.e., the policy is *discretionary*—wasteful inflation is created on average in equilibrium. Figure 1 below plots time paths for the average *inflationary bias*, i.e., the time profile of ex ante equilibrium inflation in excess of  $\mu$ , the average inflation for an authority who is only concerned about inflation targeting. This is done for different values of  $\sigma_X$ : larger values of it imply noisier feedback to our sender and capture environments with larger higher-order uncertainty.

Three interesting phenomena arise in the figure. First, inflationary biases are lower than in the static benchmark solution, the horizontal line (except at  $t = T$  where myopic behavior is optimal). Second, and related to the role of higher-order uncertainty, the bias tends to be higher than in the public case: the dashed lines ( $\sigma_X > 0$ ) are higher than the solid line ( $\sigma_X = 0$ ). Third, inflationary biases for different values of  $\sigma_X$  can *cross*: high values of  $\sigma_X$  are associated with large (small) inflationary biases earlier (later) in the game.

<sup>20</sup>Existence results for the first two games fall under the umbrella of our general existence theorem of Section 5, while an existence result for our trading game is in Section S.4 in the Supplementary Appendix.

<sup>21</sup>If  $Y$  is private, directly setting  $\sigma_X = 0$  in our model leads to  $M = L = \hat{M}$ , and hence our LMEs coincide with those that arise when  $Y$  is public (where  $X$  is also ignored because it does not have statistical value).

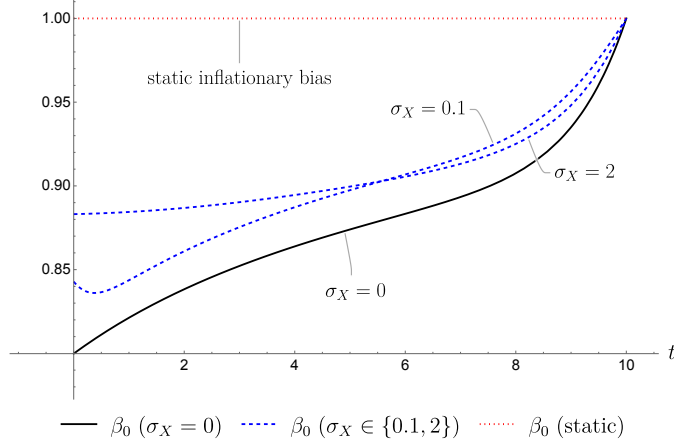


Figure 1: Inflationary bias for various values of  $\sigma_X$ . Parameters:  $(k, \gamma^o, r, \sigma_Y) = (1, 1, 1, 1.5)$ .

**LME coefficients and dynamics** The LME supporting these outcomes is as follows.

**Proposition 2.** *Suppose that  $r \geq 0$  and  $\sigma_X \in (0, \infty)$ . In any LME, the coefficients satisfy  $\beta_{1t} + \beta_{2t} + \beta_{3t} = 1$  and  $\alpha_{3t} := \beta_{3t} + \beta_{1t}\chi_t > 0$ . On the equilibrium path, therefore,*

$$a_t = \beta_{0t} + \alpha_{3t}\theta + (1 - \alpha_{3t})L_t \quad \text{and} \quad \hat{a}_t = \beta_{0t} + \alpha_{3t}\hat{M}_t + (1 - \alpha_{3t})L_t. \quad (20)$$

Further, there is  $T^\dagger > 0$  such that, for all  $T < T^\dagger$  the following can be shown analytically:  $\beta_{0t} < k$  for  $t < T$ ;  $\beta_{1t}, \beta_{2t} \in (0, 1/2)$  while  $\beta_{3t} \in [1/2, 1)$ ; and  $\alpha_{3t} \in (1/2, 1)$ .<sup>22</sup>

From (20), equilibrium inflation  $a_t$  preserves the structure of the static equilibrium from Section 2,  $a^{\text{static}} = k + \frac{1}{2}\theta + \frac{1}{2}\mu$ , albeit with time-varying weights: it is a convex combination between the type and the public belief about it, plus an intercept. In particular, note that the intercept  $(\beta_{0t})_{t \in [0, T]}$  is the average *inflationary bias* from Figure 1, as  $\mathbb{E}[a_t] = \beta_{0t} + \mu$ .

Towards understanding Figure 1, let us examine the equilibrium coefficients in the LME from the proposition. To this end, we write the authority's payoff (up to a constant) as:

$$\underbrace{2k(a_t - \hat{a}_t)}_{\text{stimulus}} \quad \underbrace{-(a_t - \hat{a}_t)^2}_{\text{output stabilization}} \quad \underbrace{-(a_t - \theta)^2}_{\text{inflation targeting}}. \quad (21)$$

The coefficient  $\beta_0$  is linked to the stimulus part in (21). In particular, a forward-looking authority understands that an inflationary bias as in the static case is dynamically costly: large inflation surprises *anchor* the private sector's belief at higher levels, making it more costly to boost employment in the future. All types deviate downward from the static bias—i.e.,  $\beta_0 < k$ —to dampen  $M$  today (a proxy for the market's current belief), and create more

<sup>22</sup> $T^\dagger$  depends on parameters; see discussion after Theorem 1. Numerically, the findings hold beyond  $T^\dagger$ .

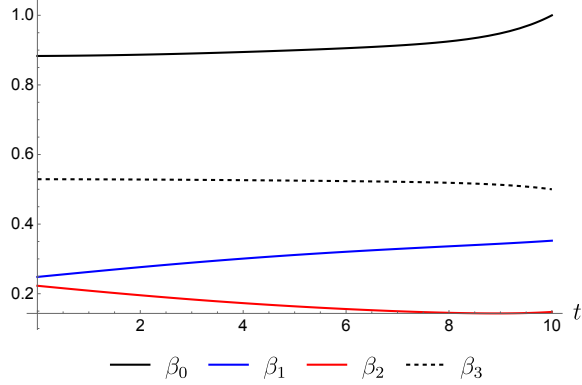


Figure 2: The sender’s strategy coefficients, for  $(k, r, \gamma^o, \sigma_X, \sigma_Y) = (1, 1, 1, 2, 1.5)$ .

inflation, at a lower cost, at a future date—i.e.,  $\beta_0$  is predominantly increasing. From an ex ante standpoint, the authority exhibits an apparent greater commitment to low inflation.<sup>23</sup>

The authority correctly expects lower future inflation costs in this LME because, ex post, different types do affect outcome variables, such as market expectations. This brings us to the rest of the coefficients,  $(\beta_1, \beta_2, \beta_3)$ , and how they connect with the remaining terms in (21). Consider Figure 2. The starting point is that authorities with higher types will choose higher inflation, so  $\beta_3 > 0$ . In response, higher private sector “types”  $\hat{M}$  will set higher inflation expectations  $\hat{a}$ . But if these expectations are not met, unemployment can grow (e.g., through excessive wage demands) and higher values of  $M$  must then be accompanied with higher inflation, so  $\beta_1 > 0$ . By the same logic, both parties will attach a positive weight to  $L$ : high values of  $L$  indicate high values of  $M$  again, ultimately leading to  $\beta_2 > 0$ .<sup>24</sup>

Altogether, higher market beliefs about  $\theta$  bring high inflation expectations, which in equilibrium are self-fulfilling. Inflation then becomes costly to the authority, because it is set far from  $\theta$ . A forward-looking authority would like to *guide* the market in the direction of  $\theta$  and surprise it with relatively high inflation tomorrow (via  $\beta_0$ ), when it is cheaper to do so. Note that this guiding motive is captured by the last two terms in (21), as setting  $a = \hat{a} = \theta$  in them eliminates any potential losses for the monetary authority.

Thus, in equilibrium, the weight attached to  $\theta$  in the extended strategy is larger than in the static case,  $\beta_3 > 1/2$ : the authority signals her type more aggressively through this channel to gradually guide  $\hat{M}$  toward it. And as more time remains in the game, it is more profitable to stick to inflation targeting—i.e.,  $\beta_3$  is decreasing—because there is more time to enjoy the future benefits. In the process, however, output dynamics arise due to

<sup>23</sup>This type of downward deviation is reminiscent of the *ratchet effect*. See Cisternas (2018) for an application to monetary policy symmetric incomplete information where an analogous effect is at play.

<sup>24</sup>Because output and inflation have equal weights in (21), overweighing the importance of the type comes at the cost of the coefficients in the rest of the states, so  $\beta_1 + \beta_2 + \beta_3 \equiv 1$ .

$a - \hat{a} = \alpha_3(\theta - \hat{M}) \neq 0$ . The fact that the total signaling coefficient  $\alpha_3$  is less than 1 reflects the authority’s concern about surprising the market too much in this guiding process, as the signals observed by the private sector are imperfect, and hence can destabilize output.

The classic credibility problem has a neat dynamic manifestation here. The authority tries to build a reputation for *transparency* early on, understood as carefully guiding the market’s belief towards its true type, with little wasteful inflation early on. But as she succeeds, her temptation to boost the economy increases, and her reputation becomes compromised.

**Higher-order uncertainty** We now turn to examining how the presence of higher-order uncertainty affects the inflationary bias  $\beta_0$ . To do so, it is useful to consider the case in which  $Y$  is public, where the LME is  $a_t = \beta_{0t}^{\text{pub}} + \beta_{3t}^{\text{pub}}\theta + (1 - \beta_{3t}^{\text{pub}})\hat{M}_t$  and  $\hat{a}_t = \beta_{0t}^{\text{pub}} + \hat{M}_t$ .<sup>25</sup> From the authority’s perspective, the contrast is between inflation expectations of the form

$$\underbrace{\beta_{0t} + \alpha_{3t}M_t + (1 - \alpha_{3t})L_t}_{= \mathbb{E}_t[\hat{a}_t], \text{ when } \sigma_X > 0} \quad \text{versus} \quad \underbrace{\beta_{0t}^{\text{pub}} + \hat{M}_t}_{= \hat{a}_t, \text{ when } \sigma_X = 0} .$$

The problem of forecasting others’ forecasts has the “direct” effect of making market expectations less sensitive to the authority’s actions: the weight  $\delta_1^{\text{pub}} \equiv 1$  attached to  $\hat{M}$  in the public case (right term) is now split into a weight  $\delta_1 = \alpha_3$  on  $M$ —the state that is directly controlled—and a weight of  $\delta_2 = 1 - \alpha_3$  attached to  $L$ —a state that updates more slowly because it is only indirectly controlled via inducing changes in the receiver’s action. (Note the importance of the receiver’s signaling—encoded in  $\delta_1$ —for this finding.)

Thus, all else equal, market expectations tend to become more sluggish. As a result, creating inflation surprises become less costly from a dynamic perspective, and the incentive to deviate downward from the static solution is weakened. The tangible consequence of the beliefs about beliefs problem is that a higher inflationary bias is likely to arise.

**Proposition 3.** *Fix  $r > 0$  and  $\sigma_X > 0$ . (i) There is  $T^\dagger > 0$  (depending on parameters), such that for all  $T < T^\dagger$ ,  $\beta_0$  is uniformly higher than its counterpart  $\beta_0^{\text{pub}}$  over  $[0, T)$ . (ii) if  $r = 0$  and  $\sigma_X = +\infty$ ,  $\beta_0$  is higher than the public counterpart for all horizons  $T > 0$ .*

Higher inflationary biases are guaranteed if there is sufficient immediacy and/or higher-order uncertainty: in both cases, the authority manages market expectations mainly through influencing  $M$ , the channel that has a more immediate—albeit reduced—response. For instance, given any degree of non-trivial discounting, we can always find meaningful horizons such that the public channel  $L$  is not able to compensate the reduced sensitivity of  $\hat{a}$  to  $M$ . On the other hand, the extent of higher-order uncertainty faced by the authority is maximized

---

<sup>25</sup>We establish this in Proposition S.1 in the Supplementary Appendix.

if  $\sigma_X = +\infty$ : the public state  $L$  naturally becomes  $\mu$ , and hence cannot be affected. In this case, the inflationary bias is higher for all horizons, even if the authority is patient.<sup>26</sup>

But there are indirect effects too: varying the degree of higher-order uncertainty faced by the authority affects her signaling. Due to this second channel, the time paths for inflationary bias can cross as  $\sigma_X$  varies. To this end, consider Figure 3 below, depicting the total *signaling coefficient*  $\alpha_3 := \beta_3 + \beta_1\chi$  for various values of  $\sigma_X$ :

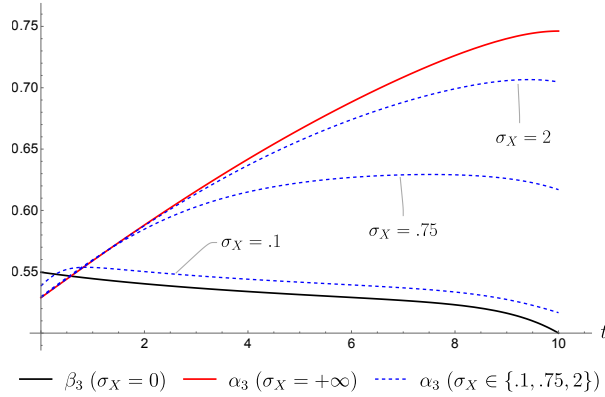


Figure 3: Sender’s signaling coefficient as  $\sigma_X$  varies. Other parameters:  $(k, \gamma^o, r, \sigma_Y) = (1, 1, 1, 1.5)$

From the plot, the signaling coefficient is decreasing for the most part when  $\sigma_X$  is low (e.g.,  $\sigma_X = 0.1$ ), but the opposite occurs when  $\sigma_X$  is large (e.g.,  $\sigma_X = 2$ ). In the former case, there is mild higher-order uncertainty, so  $\beta_3$  dominates in  $\alpha_3$ ; but as argued, the “guiding motive” weakens over time. At the other end, the extent of signaling through the second-order belief channel,  $\beta_1\chi$ , dominates. Indeed, as higher types create more inflation ( $\beta_3 > 0$ ), their past behavior will lead them to expect higher beliefs by the market, i.e., to develop high values of  $M$ . To stabilize output, these types will now create even more inflation ( $\beta_1 > 0$ ), effectively amplifying the separation of types. This effect builds up over time as  $M$  increasingly reflects the type in the representation (11) and the guiding motive decays (both  $\chi$  and  $\beta_1$  grow).

As  $\sigma_X$  grows,  $L$  becomes more sluggish, and so  $M$  becomes the main channel to guide the market. However, with a reduced sensitivity of  $\mathbb{E}_t[\hat{a}_t]$  to  $M_t$  (the direct effect) this task becomes harder for all types: early in the game,  $\beta_3$  falls more and more below the public case (i.e., closer to the myopic solution)—see Figure 3 close to  $t = 0$ , where  $\alpha_3 \approx \beta_3$ . Large degrees of higher-order uncertainty are then associated with weaker signaling, hence with less sensitive expectations, and ultimately a larger inflationary bias. As time progresses, things reverse: our new form of signaling builds up, expanding the wedge in information transmission at high versus low values of  $\sigma_X$ , as Figure 3 shows. Inflation expectations  $\hat{a}$

<sup>26</sup>If  $\sigma_X = +\infty$ , the reliance on past play is maximized. The representation (11) holds with  $L \equiv \mu$ , while the ODEs (14)–(15) for  $(\gamma, \chi)$  in Lemma 1 simplify dramatically. See Lemma B.1 in Appendix B.

gain more sensitivity when  $\sigma_X$  is large, and hence lower (in relative terms) inflationary biases arise later in the game, as Figure 1 shows after  $t = 6$ . We can formalize this crossing of the signaling coefficients by comparing two settings with minimal and maximal higher-order uncertainty:  $\sigma_X \in \{0, +\infty\}$ . We exploit the analytical solutions when  $r = 0$  for comparisons.

**Proposition 4.** *Suppose that  $\sigma_X \in \{0, +\infty\}$ . Then, an LME exists for all  $T > 0$  and  $r \geq 0$ . Further, if  $r = 0$ , it can be shown analytically that the signaling coefficient in the public case is higher (lower) at  $t = 0$  ( $t = T$ ) than its counterpart for  $\sigma_X = \infty$ .*

The crossing of signaling coefficients for interior, extreme opposite, levels of  $\sigma_X$  then follows by continuity—that this crossing occurs before the  $\beta_0$  coefficients cross is because, when  $\sigma_X$  is large, the contribution of  $\alpha_3$  to the sensitivity of inflation expectations needs to build up to offset that  $M$  itself has become less responsive to due to the receiver’s learning.<sup>27</sup>

In conclusion, the logic that higher-order uncertainty leads to more sluggish beliefs can then be challenged if the signals used endogenously become more informative. The two-sided learning aspect of our game and the presence of agents who can affect signals are key.<sup>28</sup>

## 4.2 Application 2: Reputation for Neutrality

Consider now the following payoffs (up to positive factors so that  $u_{aa} = \hat{u}_{\hat{a}\hat{a}} = 1$ ):

$$\text{sender: } - \int_0^T e^{-rt} (a_t - \theta)^2 dt - e^{-rT} \psi \hat{a}_T^2, \quad \text{receiver: } - (\hat{a}_T - \theta)^2,$$

where  $\psi > 0$ . The interpretation is that  $\theta$  represents a privately known *bias* on a relevant issue, with the prior mean—which we now set to  $\mu = 0$  for notational convenience—capturing an unbiased type. The sender (e.g., a politician or expert), however, faces a conflict: she finds it costly to take actions away from her type ( $-(a_t - \theta)^2$  term) but she benefits from appearing as unbiased at a terminal time  $T$ ; this is because the receiver wants to predict the sender’s bias (i.e.,  $a_t = \hat{M}_t$ ,  $t \in [0, T]$ ) and being perceived as neutral at  $T$  (i.e.,  $\hat{a}_T = \hat{M}_T = \mu = 0$ ) eliminates the terminal loss. The receiver could be a news outlet that gets private signals  $Y$

<sup>27</sup>By Proposition 3, the crossing of the inflationary biases for  $\sigma_X = 0$  and  $+\infty$  becomes degenerate, in that it is backloaded to  $t = T$ . Also, the non-monotonicity of  $\alpha_3$  for  $\sigma_X \in (0, +\infty)$  is the net effect of the guiding motive and the signaling through  $M$  moving in opposite direction: the proof of Proposition 4 shows that the signaling coefficient is decreasing when  $\sigma_X = 0$  and increasing when  $\sigma_X = \infty$ . The non-monotone pattern of  $\beta_0$  in some plots is a consequence of this non-monotonicity of  $\alpha_3$ , via the same sensitivity logic.

<sup>28</sup>Higher-order uncertainty instead leading to lower (higher) inflation earlier (later) on can happen for low  $r$  and  $\sigma_X$ . Indeed, with a precise public signal, the players rely heavily on  $L$ . A patient monetary authority can then manage inflation expectations through this state early on, which is more sensitive as  $\delta_1 = \alpha_3$  grows; but as time progresses,  $M$  gains more immediate prominence, and we fall into the domain of Proposition 3. The logic, however, applies for settings close to public, and hence the differences in inflation are small.



of the sender’s behavior and that reports its perception of the bias;<sup>29</sup> the reporting process  $X$  is imperfect, but fair on average (the shocks have zero mean)—and naturally public.

Does having access to better information, as measured by a more precise signal  $X$ , benefit this career-concerned agent? Clearly, one advantage of more precise information is that actions can be better tailored to one’s reputation. Let us first look at the LME of this game.

**Proposition 5.** *Suppose that  $r \geq 0$  and  $\sigma_X \in (0, \infty)$ . In any LME, the sender’s strategy satisfies  $\beta_{0t} = 0$  and  $\beta_{1t}, \beta_{2t} \leq 0 < \beta_{3t} \leq 1$  for all  $t \in [0, T]$ , with all inequalities strict over  $[0, T)$ , while  $\hat{a}_t = \hat{M}_t$  (i.e.,  $\delta_1 = 1$  and  $\delta_0 = \delta_2 \equiv 0$ ). Moreover,  $\alpha_{3t} := \beta_{3t} + \beta_{1t}\chi_t \in (0, 1)$ .*

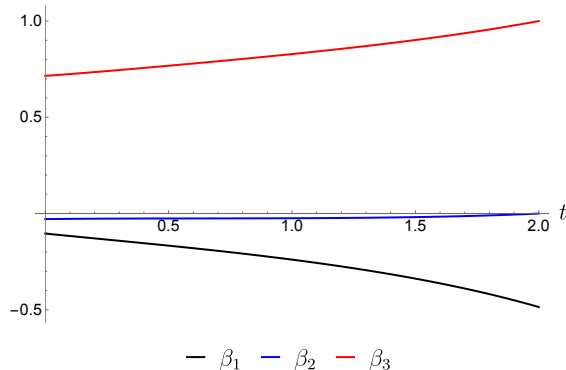


Figure 4: Sender’s strategy coefficients in the reputation game:  $(\psi, \gamma^o, r, \sigma_X, \sigma_Y) = (1, 2, .1, 1, 1)$ .

To understand these coefficients, consider Figure 4. First, if the sender were myopic, she would attach a weight of 1 to  $\theta$  at all times, precisely the terminal value of  $\beta_3$ . The sender then deviates from this value at earlier times in an effort to manage her reputation. Specifically, from a time- $t$  perspective, her reputational concerns are captured by

$$-e^{-r(T-t)}\psi\mathbb{E}_t[\hat{M}_T^2] = -e^{-r(T-t)}\psi(\mathbb{E}_t[M_T^2] + \chi_T\gamma_T), \quad (22)$$

where  $\chi_T\gamma_T$  is the variance of the sender’s second-order belief (Lemma 2). Two conclusions immediately follow. First, since higher types take higher actions ( $\alpha_3 > 0$ ) due to their higher biases, these types will anticipate greater upward drift in their reputation  $M$  all else equal. To preempt a large terminal loss, the sender moderates her actions, resulting in  $\beta_{3t} < 1$  until time  $T$ ; this deviation is stronger earlier in the game, as more time is left to reap the benefits of it. Second, senders with (from their perspective) biased reputations  $M_t$  expect to be perceived as biased at the end, so they will attempt corrective actions early on: the weight  $\beta_1$  on  $M$  is negative so as to prevent this state from growing. And since  $M_t$  becomes a better predictor of  $M_T$  as time progresses, such corrections becomes stronger:  $\beta_{1t}$  is decreasing.

<sup>29</sup>Actions such as voting, contributions, favors, statements to groups of influence, etc. often have a private nature, and hence are likely to be leaked with error, justifying the noise in  $Y$ .

It is noteworthy that  $L$  is used in the strategies despite *never* appearing in the players' payoffs. Recall that, via the representation, high values of  $L$  indicate to the receiver that  $M$  is high, and hence that a strong correction (i.e., a low  $dY$ ) should be observed. If the sender does not meet the receiver's expectations, the sender predicts that her reputation will deteriorate upwards. The case  $\theta = 0$ , at an off-path history where  $M_t = 0$ , illustrates this issue: despite being truly unbiased, and also believing she is perceived as such, this type chooses  $a_t = \beta_2 L_t$  (which is negative if  $L_t > 0$ ) because she expects to be perceived as biased tomorrow otherwise. As this predictability ceases to matter at  $T$ ,  $\beta_{2T} = 0$  in Figure 4.

**Second-order belief and concealment** As more extreme types take more extreme actions in equilibrium, they will also develop more extreme second-order beliefs  $M$ ; hence, they will correct their reputations more aggressively than less extreme types. Our new signaling channel now goes *against* information transmission:  $\beta_1 \chi$  and  $\beta_3$  have opposite signs. But this creates scope for less separation, and hence a better chance to conceal the bias. A subtle trade-off emerges: with higher-order uncertainty, the sender loses her ability to take the best actions to manage her reputation, but she may transmit less information in the first place. Conversely, in a public setting, the sender can perfectly tailor her actions to her reputation, but types do not separate through their beliefs. To make our point, we examine the cases  $\sigma_X = 0$  and  $\sigma_X = +\infty$  when  $r = 0$ , which are analytically simple.

**Proposition 6.** *Suppose that  $r = 0$  and  $\psi < \sigma_Y^2/\gamma^\circ$ . For all  $T > 0$ , there is a unique LME for  $\sigma_X \in \{0, +\infty\}$ . The sender's ex ante payoff is higher when  $\sigma_X = +\infty$  than when  $\sigma_X = 0$ .*

Our career-concerned sender can benefit from being uncertain about how she is perceived when  $\psi < \sigma_Y^2/\gamma^\circ$ . Otherwise, changes in beliefs are costly (large  $\psi$ , due to an acute concavity), or beliefs themselves change frequently (due to a large initial uncertainty  $\gamma^\circ$  or more informative  $Y$  signal), which increases the “better tailoring” benefit of more precise feedback.

### 4.3 Application 3: Trading and Leakage

Consider a public signal of the form  $dX_t = (a_t + \hat{a}_t)dt + \sigma_X dZ_t^X$ , and payoffs

$$\text{sender: } \int_0^T \left[ (\theta - \mathbb{E}[\theta | \mathcal{F}_t^X]) a_t - \frac{a_t^2}{2} \right] dt; \quad \text{receiver: } (\theta - \mathbb{E}[\theta | \mathcal{F}_t^X]) \hat{a}_t - \frac{\hat{a}_t^2}{2}.$$

The sender is an informed trader who knows the fundamental value  $\theta$  of an asset, while the receiver is an ex ante uninformed investor who sees a leakage  $Y$  of the sender's trades.<sup>30</sup> The

---

<sup>30</sup>Yang and Zhu (2020) argue that, by handling retail order flow (proxy for noise trading), proprietary trading firms can construct private signals of institutional investors' (proxy for informed traders) behavior.

term  $\mathbb{E}[\theta|\mathcal{F}_t^X]$  corresponds to the asset’s price at time  $t$ , based on the public *total order flow*  $X$ , as in Kyle (1985). Thus,  $(\theta - \mathbb{E}[\theta|\mathcal{F}_t^X])dt$  captures the trading gains for each unit bought over  $[t, t + dt)$ . In turn,  $a_t^2/2$  denotes other types of transaction costs.<sup>31</sup>

This game departs from our baseline model because (i) the public signal carries the sender’s action and (ii) there is a “third action”—the price—based exclusively on the public information. Our analysis from Sections 2–3 extends to these cases nonetheless. The key is that the identity  $L_t = \mathbb{E}[\theta|\mathcal{F}_t^X]$  still holds in the analogous representation result for  $M$ , which enables us to explore environments where  $\mathbb{E}[\theta|\mathcal{F}_t^X]$  is directly payoff relevant.

In models of this kind, a key question is how the sender exploits the mispricing  $\theta - L$  accounting for how her trades affect future prices. Here, such trades also generate private information for others. The following proposition characterizes the structure of any LME—Section S.4 in the Supplementary Appendix provides an existence result.

**Proposition 7.** *Suppose  $\sigma_X \in (0, \infty)$ . In any LME,  $\beta_{0t} = 0$  and  $\beta_{1t} + \beta_{2t} + \beta_{3t} = 0$ . Thus, along the path of play,  $a_t = \alpha_{3t}(\theta - L_t)$  with  $\alpha_{3t} > 0$ , while the receiver follows  $\hat{a}_t = \hat{M}_t - L_t$ . Finally, the equilibrium price satisfies  $dL_t = \Lambda_t dX_t$  where  $\Lambda_t := \frac{\gamma_t^X(\alpha_{3t} + \chi_t)}{\sigma_X^2}$ .*

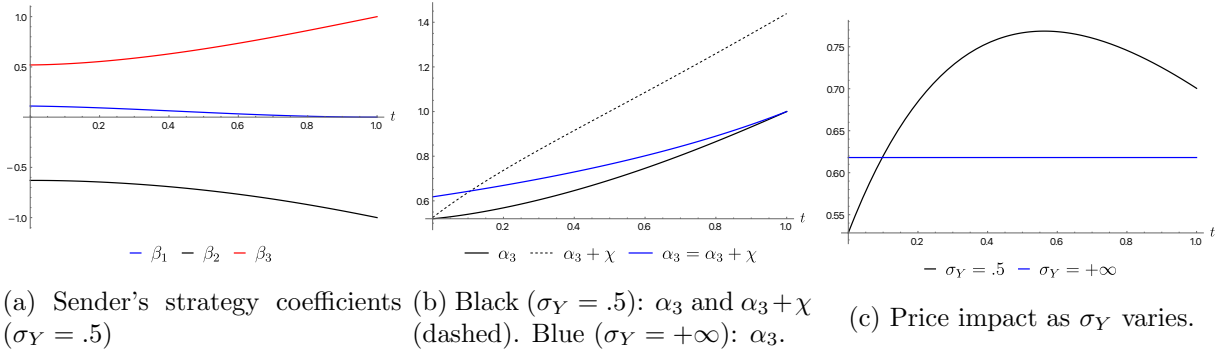


Figure 5: Trading game:  $(\gamma^\circ, r, \sigma_X) = (1, 0, 1)$ .

As higher sender types expect a larger profit per unit traded, the equilibrium weight on the type  $\beta_3$  is positive. At the same time, all types scale back their purchases as the price increases, and so  $\beta_2 < 0$ . Further, as the endgame approaches and the concern about future prices fades,  $\beta_3$  and  $\beta_2$  move toward the myopic solution monotonically—see Figure 5a.<sup>32</sup>

As is usual, the sender’s equilibrium trades are based on the size of the current mispricing via  $\alpha_3(\theta - L_t)$ , with  $\alpha_3 > 0$  also shaping the responsiveness of prices via  $\Lambda$ ; in other words, endogenous price impact makes it costly for the sender to place large trades. What is novel is

<sup>31</sup>E.g., taxes from trades (Subrahmanyam, 1998). Additional costs from large “long” positions also arise from limited resources within a fund; and on the “short” side, due to the use of brokers for borrowing shares.

<sup>32</sup>As usual, we can sign coefficients for horizon lengths that depend on parameters:  $\beta_{1t} > 0$ ,  $\beta_{3t} \in (0, 1)$ , and  $\beta_{2t} < 0$ ; while  $\frac{d\alpha_{3t}}{dt} > 0$ ,  $\frac{d\beta_{3t}}{dt} > 0$ , and  $\frac{d\beta_{1t}}{dt} < 0$ . See Supplementary Appendix Section S.4.4.

how the presence of private monitoring affects the form that price impact takes, which brings us to the important role that  $\beta_1$  plays. To illustrate, suppose that the sender has deviated by trading more aggressively, resulting in a higher  $M$  than implied by the representation (11). In practice, this means that the sender thinks that the receiver is optimistic about the asset, and that this optimism will eventually get incorporated into prices: the persistent state  $M$  enters the law of motion for the price  $L$  (just as in the law of motion (17) for  $L$  in our baseline model). That is, an extra layer of price impact emerges.

This extra layer is a form of price predictability that can be exploited by the sender. Indeed, using  $\beta_{1t} + \beta_{2t} + \beta_{3t} = 0$ , her extended strategy reads  $a_t = \beta_{3t}(\theta - L_t) + \beta_{1t}(M_t - L_t)$ , thus capturing that the sender exploits both forms of superior information relative to market makers (who believe  $\mathbb{E}[\hat{M}_t | \mathcal{F}_t^X] = L_t$ );  $\beta_1 > 0$  then reflects an incentive to buy more aggressively in anticipation of higher future prices, an effect that decays over time ( $\beta_{1T} = 0$ ) (also in Figure 5a). A sequentially rational trader recognizes that any individual trade now carries more price impact:  $\Lambda$  in the proposition features  $\beta_3$  augmented by  $\beta_1\chi + \delta_1\chi = \beta_1\chi + \chi$ , the sender’s signaling through  $M$  plus the receiver’s own trades.<sup>33</sup> The trader then slows down her purchases for fear of high future prices, and  $\alpha_3$  falls below the “no-leak” benchmark case  $\sigma_Y = +\infty$  (in which the receiver does not trade at all)—see Figure 5b. In turn, in the same figure, the order flow’s informativeness (sender and receiver combined),  $\alpha_3 + \chi$ , is depicted in dashed: it is low early on due to the sender’s reduced signaling, but it builds up as the contribution of  $\chi$  grows.

The implication is that price impact  $\Lambda$  begins below the no-leak counterpart, but eventually surpasses it, as shown in Figure 5c. This time pattern highlights the role of the endogenous correlation of the player’s private information, via the receiver’s endogenous type. Alternatively, when such correlation is exogenous (because the players’ private information itself is, as in Foster and Viswanathan, 1996; Back et al., 2000) price impact is initially high if correlation is positive, due to intense competition early on.<sup>34</sup>

## 5 Existence of Linear Markov Equilibria

In this section we show that the problem of finding LMEs is effectively one of solving a system of ODEs with a mix of initial and terminal conditions—a “boundary value problem”

<sup>33</sup>The appearance of  $\chi$  multiplying  $\delta_1 \equiv 1$  stems from  $\text{Cov}(\hat{M}_t, \theta | \mathcal{F}_t^X) = \chi_t \gamma_t^X$  from the perspective of market makers, as they must also use the representation to construct a (second-order) belief about  $\hat{M}$ .

<sup>34</sup>It is possible to show that there is a non-zero measure of times at which  $\Lambda_t > \Lambda_t^{no\ leak}$ ; see Section S.4.4 in the Supplementary Appendix. Also in that section, numerical calculations show that the hump-shaped pattern for  $\Lambda$  remains such if the receiver is forward-looking—the fall in  $\Lambda$  is due to the increased market maker’s learning (i.e.,  $\gamma^X$  falling) eventually dominating when quadratic costs bound trading intensity.

(or ‘BVP’). We provide time horizons for which such a problem admits a solution.

**Setting up a BVP** We postulate a quadratic value function for the sender of the form

$$V(\theta, m, \ell, t) = v_{0t} + v_{1t}\theta + v_{2t}m + v_{3t}\ell + v_{4t}\theta^2 + v_{5t}m^2 + v_{6t}\ell^2 + v_{7t}\theta m + v_{8t}\theta\ell + v_{9t}m\ell,$$

where  $v_i$ ,  $i = 0, \dots, 9$  are differentiable functions of time. We can then write the Hamilton-Jacobi-Bellman (HJB) equation for the sender’s problem: for all  $t < T$ ,

$$rV = \sup_{a'} \left\{ \tilde{u}_t(a', \mathbb{E}_t[\hat{a}_t], \theta) + V_t + \mu_M(a')V_m + \mu_L V_\ell + \frac{\sigma_M^2}{2} V_{mm} + \sigma_M \sigma_L V_{m\ell} + \frac{\sigma_L^2}{2} V_{\ell\ell} \right\}, \quad (23)$$

where  $\tilde{u}_t(\cdot) := u(\cdot) + \frac{1}{2}u_{\hat{a}\hat{a}}\delta_{1t}^2\gamma_t\chi_t$  (due to  $\mathbb{E}_t[\hat{M}_t^2] = \mathbb{E}_t[M_t^2] + \gamma_t\chi_t$  by Lemma 2) and where  $\mu_M(a')$  and  $\mu_L$  (respectively,  $\sigma_M$  and  $\sigma_L$ ) denote the drifts (respectively, volatilities) in (16) and (17). Note that, via  $\tilde{u}$  and  $(\mu_M, \mu_L, \sigma_M, \sigma_L)$ , (23) implicitly depends on a tuple  $\vec{\beta} := (\beta_0, \beta_1, \beta_2, \beta_3)$  used by the receiver to form beliefs about the sender and construct his best response  $\hat{a}_t$  (as in (19)). Also,  $(\mu_M, \mu_L, \sigma_M, \sigma_L)$  depend on  $(\gamma, \chi)$  (see (16)–(17)).

Let  $a(\theta, m, \ell, t)$  denote the maximizer in (23). The first-order condition (FOC) reads

$$\frac{\partial u}{\partial a}(a(\theta, m, \ell, t), \delta_{0t} + \delta_{1t}m + \delta_{2t}\ell, \theta) + \underbrace{\frac{\gamma_t \alpha_{3t}}{\sigma_Y^2}}_{dM_t/da_t} \underbrace{[v_{2t} + 2v_{5t}m + v_{7t}\theta + v_{9t}\ell]}_{V_m(\theta, m, \ell, t)} = 0, \quad (24)$$

which is linear in  $a(\theta, m, \ell, t)$  and  $(\theta, m, \ell)$ . Solving for  $a(\theta, m, \ell, t)$  in (24) and imposing the equilibrium condition  $a(\theta, m, \ell, t) = \beta_{0t} + \beta_{1t}m + \beta_{2t}\ell + \beta_{3t}\theta$  yields a linear equation in  $(\theta, m, \ell)$ . We then impose that the terms attached to these variables on each side, and the constants, coincide. Having linked  $\vec{\beta}$  with  $(v_2, v_5, v_7, v_9)$ , one can always obtain a system of ODEs for  $v_i$ ,  $i = 0, \dots, 9$  using the HJB equation. But this problem permits a reduction:

1. We can instead solve for  $(v_2, v_5, v_7, v_9)$  as a function of  $\vec{\beta}$  and  $(\gamma, \chi)$  (see (C.1)–(C.4) in the Appendix); the associated mapping is well-defined provided that  $\alpha_3$  and  $\gamma$  never vanish, which will be the case in the equilibrium we construct (more in this shortly);
2. We can then insert the resulting expressions into (23), along with  $a(\theta, m, \ell, t) = \beta_{0t} + \beta_{1t}m + \beta_{2t}\ell + \beta_{3t}\theta$ , to obtain a system of ODEs for  $(\beta_0, \beta_1, \beta_2, \beta_3)$  and the remaining coefficients in  $V$ . These ODEs are coupled with those of  $(\gamma, \chi)$  because this pair shapes the law of motion of  $(M, L)$ . The resulting system of ODEs can be further reduced by eliminating  $(v_0, v_1, v_3, v_4, \beta_0)$  which are “downstream” of the remaining variables.<sup>35</sup>

---

<sup>35</sup>Note that  $(v_0, v_1, v_4)$  are the coefficients of the constant,  $\theta$ - and  $\theta^2$ -terms in the sender’s value function, none of which the sender controls, so they have no impact on the rest of the system. Meanwhile, the equations for  $(\beta_0, v_3)$  are coupled as these encode the *deterministic* component of the sender’s incentive to manipulate

This procedure yields a system of ODEs for  $(\beta_1, \beta_2, \beta_3, v_6, v_8, \gamma, \chi)$ ; see Appendix C.<sup>36</sup> To this system we add *boundary conditions*. First,  $\gamma$  and  $\chi$  satisfy exogenous initial conditions  $\gamma_0 = \gamma^\circ > 0$  and  $\chi_0 = 0$  capturing the players' uncertainty at  $t = 0$ . Second, the remaining variables satisfy *endogenous* terminal conditions determined by the (Bayes) Nash equilibrium of the static game of two-sided incomplete information that takes place at time  $T$ . If there are no terminal payoffs (i.e.,  $\psi \equiv 0$ ), these conditions read as

$$\beta_{0T} = \frac{u_{a0} + u_{a\hat{a}}\hat{u}_{\hat{a}0}}{1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}}, \quad \beta_{1T} = \frac{u_{a\hat{a}}[u_{a\theta}\hat{u}_{\hat{a}a} + \hat{u}_{\hat{a}\theta}]}{1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T}, \quad \beta_{2T} = \frac{u_{a\hat{a}}^2\hat{u}_{\hat{a}a}[u_{a\theta}\hat{u}_{\hat{a}a} + \hat{u}_{\hat{a}\theta}](1 - \chi_T)}{(1 - u_{a\hat{a}}\hat{u}_{\hat{a}a})(1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T)}, \quad (25)$$

$$\beta_{3T} = u_{a\theta}, \quad v_{6T} = v_{8T} = 0, \quad \text{where } u_{x0} := u_x(0, 0, 0) \text{ and } \hat{u}_{x0} := \hat{u}_x(0, 0, 0), \quad x \in \{a, \hat{a}\}.$$

This fully specifies a BVP that the coefficients  $(\beta_1, \beta_2, \beta_3, v_6, v_8)$ , along with  $(\gamma, \chi)$ , must satisfy in a LME. (See Section S.3.2 in the Supplementary Appendix for the case  $\psi \neq 0$ .)

Towards establishing the existence of LMEs, we state technical conditions on primitives.

**Technical conditions** We require that the aforementioned static terminal game always admits an equilibrium: in (25), all the denominators must be different from zero *after every possible history* of the game, encoded in the possible values that  $\chi_T$  can take. Thus, we need  $u_{a\hat{a}}\hat{u}_{\hat{a}a} < 1$ . The idea is that the sender's *best-response* is linear in  $\hat{a}$  with slope  $u_{a\hat{a}}$ , while the receiver's counterpart has slope  $\hat{u}_{\hat{a}a}$  on  $a$  (due to  $u_{aa} = \hat{u}_{\hat{a}\hat{a}} = 1$ ); with complete information then, the players' best responses would (generically) never intersect if  $u_{a\hat{a}}\hat{u}_{\hat{a}a} = 1$ . Due to the higher-order inferences, however, it is  $1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}$  and  $1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T$  that can never vanish: since  $\chi_T \in [0, 1)$  by Lemma 1, we deduce that the best-response functions always intersect when  $1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T$  never changes sign, and so we require that  $u_{a\hat{a}}\hat{u}_{\hat{a}a} < 1$ .

Second, we want the sender to always signal her type, so that there is always information transmission via the second-order belief. Indeed, recalling that the sender's actions read  $\alpha_0 + \alpha_2 L + \alpha_3 \theta$  along the path of play, a coefficient  $\alpha_3$  that never vanishes guarantees that types' actions are always ranked in the same way. A sufficient condition is that, for all  $\chi_T$ ,

$$\alpha_{3T} := \beta_{1T}\chi_T + \beta_{3T} \underbrace{=}_{\text{using (25)}} \frac{u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}\chi_T}{1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T} \neq 0,$$

i.e., that the terminal value of  $\alpha_3$  never vanishes. Since  $\chi_T \in [0, 1)$ , the numerator never vanishes if  $u_{a\theta}$  and  $u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}$  have the same sign.<sup>37</sup> Our technical conditions read:

---

beliefs, which by definition is independent of the values that the beliefs take.

<sup>36</sup>We present the system after a change of variables that simplifies the ODEs. The fact that  $v_6$  and  $v_8$  cannot be eliminated is a consequence of the sender indirectly affecting  $L$  via changes in  $M$  (see (17)).

<sup>37</sup>If  $\alpha_3 < 0$  in equilibrium, higher types take lower actions, yet still develop higher second-order beliefs, due to the weights that  $M$  attaches to past actions being negative (the receiver responds negatively to high

**Assumption 2.** *Flow payoffs satisfy (i)  $u_{a\hat{a}}\hat{u}_{a\hat{a}} < 1$  and (ii)  $u_{a\theta}(u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\theta}) > 0$ .*

**Existence of LME** Establishing the existence of a solution to the BVP is nontrivial not only because solutions to the ODEs must exist over the whole time horizon, but also because they must land at potentially endogenous values. This issue is challenging when there are multiple ODEs in both directions: the “behavior” ODEs for  $(\vec{\beta}, v_6, v_8)$  are traced backward from their terminal values by backward induction, while the “learning” ODEs for  $(\gamma, \chi)$  are traced forward from their initial values. In BVPs with only one ODE going in one of the directions—say, forward—a traditional one-dimensional shooting argument applies: introduce a guess variable for the candidate (unknown) terminal value of that forward ODE, then trace all variables backward in time using that guess variable as the initial condition, and argue via the intermediate value theorem that some guess hits the target (the exogenous initial condition). With multiple ODEs in both directions, this method does not apply.<sup>38</sup>

Our core problem of existence of LME, however, is a fixed-point one, which has a natural manifestation here: the learning coefficients  $(\gamma, \chi)$  depend on the signaling that takes place during the game, but the signaling coefficients depend on the learning counterparts because they affect  $(M, L)$  in the sender’s problem. Our approach exploits this logic by constructing a fixed-point problem over candidate coefficients as functions over  $[0, T]$ . We first explain how this *infinite-dimensional* approach works, and later elaborate on its convenience.

Begin with an arbitrary pair  $\lambda := (\gamma, \chi)$  that proxies for solutions to the learning ODEs. We require this pair to live a closed-convex domain  $\Lambda$  that nests all functions  $(\gamma, \chi)$  that can be obtained as solutions to their coupled ODEs (14)–(15) for continuous  $(\beta_1, \beta_3)$  satisfying a particular uniform bound. Equipped with  $\lambda$ , we “shoot back”: we pose an *initial-value problem* in time-reversed form consisting of the ODEs for  $(\vec{\beta}, v_6, v_8)$  taking  $\lambda$  as an input, and where the initial conditions for the ODEs are given by the static time- $T$  conditions of the game (which may depend on  $\lambda_T$ ). We then derive a sufficient condition on the time horizon such that: (i) this initial-value problem has a unique solution for all  $\lambda \in \Lambda$ ; (ii) the solution satisfies the uniform bound referred to above (we expand on this after the theorem); and (iii) the solution is continuous in  $\lambda$ . We then “shoot forward”: we feed the resulting  $(\beta_1, \beta_3)$  pair into the learning ODEs for  $(\gamma, \chi)$  to get a solution for this system that we denote  $\boldsymbol{\lambda}$ . As we prove, this two-step procedure delivers a mapping  $\lambda \in \Lambda \mapsto \boldsymbol{\lambda} \in \Lambda$  that is continuous, and Schauder’s infinite dimensional fixed-point theorem applies. By construction, the fixed-point coefficients found induce an LME. Figure 6 illustrates one iteration of this procedure.

---

observations of  $Y$ ). Note that the receiver can suspend information transmission, when  $\delta_1 = \hat{u}_{a\theta} + \hat{u}_{a\hat{a}}[\beta_{3t} + \beta_{1t}\chi_t]$  vanishes: the sender simply ignores  $X$ , while  $M$  updates via her play. Such a suspension is only temporary due to  $(\beta_1, \beta_3)$  changing over time—and it can never happen if exactly one of  $(\hat{u}_{a\theta}, \hat{u}_{a\hat{a}})$  is zero.

<sup>38</sup>Special cases for which the one-dimensional shooting is applicable are discussed in Section 6.

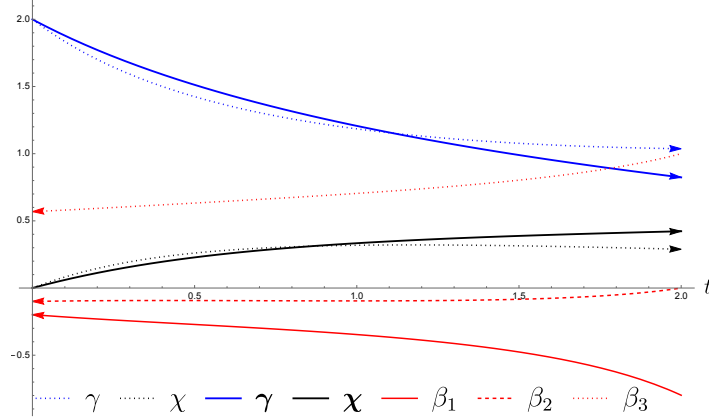


Figure 6: One iteration of our method (reputation game). Start with the dotted curves that point right, capturing learning coefficients. With these, generate candidate equilibrium coefficients, in red pointing left. Use the latter to generate solutions to the learning ODEs in solid pointing right.

We can now state our main theorem, which guarantees existence of LME for time horizons that are robust to the discount rate, for the entire class of games. Recall that  $\psi(\cdot)$  is the sender’s terminal payoff, which depends on the receiver’s terminal action.

**Theorem 1.** *Suppose Assumptions 1 and 2 hold. If  $\psi$  is linear (including  $\psi \equiv 0$ ) or  $\psi$  is not too concave, there exists a scalar  $C > 0$  independent of  $(r, \gamma^o)$  such that, if  $T < C/\gamma^o$ , there exists an LME for all  $r \geq 0$ . In this equilibrium,  $\alpha_3$  never vanishes.*

To understand why the horizons found are proportional to  $1/\gamma^o$ , recall that the receiver’s belief is less responsive to new information as  $\gamma$  falls and learning progresses. With less responsive beliefs, the sender’s incentives to deviate from myopic behavior decrease. The ODEs for the behavior coefficients  $(\vec{\beta}, v_6, v_8)$  are thus proportional to  $\gamma$ , thereby limiting the growth of any solution. Hence, as  $\gamma^o$  falls, we can find longer time horizons over which such solutions can be uniformly bounded, which guarantees their existence. (This also applies to the times  $T^\dagger$  in Propositions 2 and 3, i.e., they expand as  $\gamma^o$  falls.) As for the concavity condition, a non-trivial terminal payoff can make the static Nash equilibrium at  $T$  more complex due to “last minute” incentives. A technical lower bound on the second derivative of  $\psi$  then allows us to extract a sufficiently regular selection of static equilibria for all possible  $(\chi_T, \gamma_T)$  over  $[0, 1] \times [0, \gamma^o]$ , which we need for continuity reasons. This lower bound depends on parameters, and sometimes it never binds (our reputation game being an example).

While Theorem 1 is obviously very general, it is natural to ask why we use an infinite-dimensional approach. Interestingly, the reason relates to the economics of the problem: behavior is determined via backward induction, while Bayesian updating naturally moves going forward. In the BVP, these properties materialize in two ways. First, it is only when the behavior ODEs are traced backward that greater discounting limits their growth; we



exploit this to find times for existence that apply for all  $r \geq 0$ . Second, the learning ODEs always admit solutions if traced forward, but not necessarily backwards from generic values. To leverage this structure then, only the forward or backward component of the system of ODEs must be used in each “shooting” step, which means that candidate solutions for the remaining ODEs are needed as inputs. With *functions* needed as inputs, the approach must be infinite dimensional, beginning with either the forward or backward component.

Finally, all the steps in this existence technique can be refined: we can include more general terminal payoffs; obtain tighter uniform bounds (we only use the degree of the polynomials involved); and potentially find horizons of existence that increase with the discount rate (because behavior must be closer to myopic as  $r$  increases, and an LME for myopic players exists for all  $T$  by Assumption 2).<sup>39</sup> In the next section, we discuss this method in light of the existing literature and identify areas for future applicability.

## 6 Discussion

**Forward-looking receiver** If the receiver is forward-looking, the states  $(t, \hat{M}, L)$  are still sufficient for him. However, since he can actively control  $L$  via his actions influencing  $X$ , his incentives may change relative to the LME that we have found.

Importantly, this does not occur in our first two applications. To see why, consider the monetary policy game, and suppose that the private sector deviates from its myopic best response  $\hat{E}_t[a_t]$  over  $[t, t + dt)$ . By doing so, it incurs a loss over that instant. To see that there is no future gain, observe that since the deviation is hidden, the market continues thinking that the authority takes actions according to  $\alpha_3\theta + (1 - \alpha_3)L$ ; in this expression, only  $L$  is affected by the deviation because  $\alpha_3$  is deterministic. In other words, the private sector has not improved its ability to predict  $a_t$  at future times, which is what it cares about: the informativeness of  $Y$ , via  $\alpha_3$ , is unaffected, while the state  $L$  is always observed anyways. The same logic applies to the reputation game, and more generally to *prediction* problems.

**Proposition 8.** *Suppose that  $\hat{u}(a, \hat{a}, \theta) = -\frac{1}{2}(c_0 + c_1\theta + c_2a - \hat{a})^2$ , with  $c_0, c_1, c_2 \in \mathbb{R}$ , and that an LME in our baseline model exists. Then, for all  $\hat{r} \geq 0$ , the same LME arises when the receiver has the payoff  $\int_0^T e^{-\hat{r}t} \hat{u}(a_t, \hat{a}_t, \theta) dt + e^{-\hat{r}T} \hat{u}(a_T, \hat{a}_T, \theta)$ .*

Beyond these settings, non-trivial dynamic incentives can arise. The most relevant pertain to the receiver’s incentives to manipulate the sender’s belief via “jamming”  $L$ ; but since

---

<sup>39</sup>To find horizons that apply for all  $r \geq 0$ , we perform two modifications to the BVP before constructing a fixed point. For expositional ease, we defer a detailed explanation of those modifications and the underlying motivation to the proof in Appendix C (see ‘**Centering**’ and ‘**Auxiliary variable**’ steps).

the latter state is a publicly observed belief, these incentives are well-understood (Holmström, 1999). Crucially, our methods can be adapted to study the case of a general forward-looking receiver, which we discuss in detail in Appendix D. In a nutshell, while no additional learning ODEs would arise, the “backward part” of the BVP would have to be augmented with ODEs for the (now) dynamic coefficients  $(\delta_0, \delta_1, \delta_2)$  in the receiver’s strategy. Our two-step shooting method can then be applied to the new enlarged system.<sup>40</sup>

**Private-value environments and one-dimensional shooting** Our general analysis requires two learning dynamics,  $\gamma$  and  $\chi$ , because our players may signal at very different rates. We define a *private value* environment as one where  $\hat{u}_{\hat{a}\theta} = 0$ , i.e., the receiver’s best reply does not *directly* depend on  $\theta$ . There, the players signal at proportional rates (i.e.,  $\delta_1 = \hat{u}_{\hat{a}a}\alpha_3$ ), and the environment gains strategic symmetry:

**Proposition 9.** *If  $\hat{u}_{\hat{a}\theta} = 0$ ,  $\chi_t = \frac{c_1 c_2 (1 - [\gamma_t / \gamma^o]^d)}{c_1 + c_2 [\gamma_t / \gamma^o]^d}$  for some positive scalars  $c_1, c_2$  and  $d$ . Thus,  $\chi_t \in [0, c_2)$  when  $\gamma_t \in (0, \gamma^o]$ , where  $c_2$  is increasing in  $\sigma_X$  with  $\lim_{\sigma_X \rightarrow 0} c_2 = 0$  and  $\lim_{\sigma_X \rightarrow +\infty} c_2 = 1$ .*

When the players signal at similar rates (as defined above) there is a one-to-one mapping between  $\gamma$  and  $\chi$ . Thus, there is effectively just one ODE going forward, and traditional one-dimensional shooting methods apply (see Bonatti et al., 2017). Two observations are in order. First, the horizons for which we can guarantee the existence of LME in Theorem 1 are no worse than in the one-dimensional case. The reason is that, in both approaches, the horizons found are pinned down by ensuring that the *behavior* ODEs admit solutions (where the learning coefficients are used as inputs, and are bounded by Lemma 1). Thus, our infinite-dimensional method is the “right” extension of the one-dimensional shooting case.

Second, Proposition 9 is a contribution in itself: analogous one-to-one mappings are easy to guess when types come from symmetric distributions (e.g., Foster and Viswanathan, 1996), but in our case the sender’s type is exogenously fixed, while the receiver’s type changes and its distribution is determined in equilibrium. Alternatively, by extending to a second-order belief and involving a system of ODEs, our representation is a novel result in the literature. The upper bound  $c_2$  for  $\chi$  confirms that less weight is given to the type as  $\sigma_X$  falls and the public signal improves, as we anticipated in Section 3.2.

**Other applications of the existence method** Our fixed-point technique can be readily applied in two other subareas of LQG games: games of one-sided noisy signaling involving multidimensional types, and games with multi-sided private information and noisy public signals, where types have different prior variances. In both cases, one or many receivers have

---

<sup>40</sup>In the general case, it is more convenient to skip the reduction steps 1 and 2 after (24) and state the system solely in terms of the value function coefficients  $\vec{v}$ . See Appendix D for details.

to construct a deterministic variance-covariance *matrix* when learning about multiple types from linear strategies. Thus, multiple “learning” ODEs naturally arise.<sup>41</sup>

But our method is also applicable to other settings where equilibrium variables are encoded in ODEs. Consider search models for over-the-counter markets, where investors and dealers trade assets bilaterally (e.g., [Duffie et al., 2005](#)). “Behavior ODEs” capturing willingness to pay arise from Bellman equations, and these ODEs depend on the masses of agents looking to trade through contact rates. Meanwhile, these masses obey deterministic dynamics. With stationary solutions, the ODEs become simple algebraic equations. But the forward-backward nature matters out of steady state, such as when an entry/exit mechanism operates: the participation decision depends on future market profitability, so utilities enter the ODEs for the masses of various market participants—a feedback loop emerges.<sup>42</sup> An infinite horizon economy can then be approximated by finite-horizon versions.<sup>43</sup>

**Two-sided private monitoring** If the receiver’s actions were privately monitored too, the analysis would be more complicated. Note that the receiver would also have to rely on his past play to forecast the sender’s “ $M$ ” in his third-order belief exercise. The resulting linear aggregate of past actions need not coincide with his contemporaneous first-order belief: past actions carry the receiver’s past beliefs, and beliefs change over time. This means that the sender now has to non-trivially forecast a historical average of past values of the receiver’s first-order belief in this fourth-order belief step. Whether there are representation results that make this problem manageable—and, equally important, whether further moving up the belief hierarchy matters for economic outcomes—is an open question.

**Linear-quadratic-Gaussian structure** The tractability of static LQG models is demonstrated by the vast number of applications in which they have been used. Less obvious is what to expect from LQG structures in dynamic environments featuring complex information structures like ours. This paper demonstrates that, despite the substantial gap in difficulty when transitioning to the latter world, it is still possible to obtain new answers and insights and to develop methodological tools that can be implemented in other domains. It is our belief that the perhaps stylized nature of these games, rather than being a limitation, is an asset that helps uncover forces that are robust to other, more nonlinear settings.

---

<sup>41</sup>See [Cetemen \(2020\)](#), who uses a finite-dimensional fixed-point method from an earlier version of our paper, suited for undiscounted games. Multiple learning variables also arise in [Foster and Viswanathan \(1994\)](#), where types can be multi-dimensional and asymmetric; their fixed-point problem is confronted numerically.

<sup>42</sup>Our fixed point admits a “temporal” formulation: learning depends on past behavior, which in turn depends on future learning/behavior via backward induction. A similar circularity also arises purely from learning in LQG decision problems with forward-looking variables. See [Svensson and Woodford \(2003\)](#).

<sup>43</sup>[Bonatti et al. \(2017\)](#) use this approach to show the existence of an LME in an infinite-horizon version of their model of dynamic oligopoly with incomplete information.

## Appendix A: Proofs for Section 3

**Preliminary results.** We state standard results on ODEs (Teschl, 2012) which we use in the proofs that follow. Let  $f(t, x)$  be continuous from  $[0, T] \times \mathbb{R}^n$  to  $\mathbb{R}^n$ , where  $T > 0$ .

- Peano's Theorem (Theorem 2.19, p. 56): There exists  $T' \in (0, T)$ , such that there is at least one solution to the IVP  $\dot{x} = f(t, x)$ ,  $x(0) = x_0$  over  $t \in [0, T']$ .

If, moreover,  $f$  is locally Lipschitz continuous in  $x$ , uniformly in  $t$ , then:

- Picard-Lindelöf Theorem (Theorem 2.2, p. 38): For  $(t_0, x_0) \in [0, T] \times \mathbb{R}^n$ , there is an open interval  $I$  over which the IVP  $\dot{x} = f(t, x)$ ,  $x(t_0) = x_0$  admits a unique solution.

- Comparison theorem (Theorem 1.3, p. 27): If  $x(\cdot), y(\cdot)$  are differentiable,  $x(t_0) \leq y(t_0)$  for some  $t_0 \in [0, T]$ , and  $\dot{x}_t - f(t, x(t)) \leq \dot{y}_t - f(t, y(t)) \forall t \in [t_0, T]$ , then  $x(t) \leq y(t) \forall t \in [t_0, T]$ . If, moreover,  $x(t) < y(t)$  for some  $t \in [t_0, T]$ , then  $x(s) < y(s) \forall s \in [t, T]$ .

**Proof of Lemma 1.** Let  $L$  in (11) denote a square-integrable process that is  $X$ -progressively measurable. Inserting (11) into (9) yields  $a_t = \alpha_{0t} + \alpha_{2t}L_t + \alpha_{3t}\theta$  which the receiver thinks drives  $Y$ , where  $\alpha_{0t} = \beta_{0t}$ ,  $\alpha_{2t} = \beta_{2t} + \beta_{1t}(1 - \chi_t)$ , and  $\alpha_{3t} = \beta_{3t} + \beta_{1t}\chi_t$ .

The receiver's filtering problem is then conditionally Gaussian. Specifically, define

$$d\hat{Y}_t := dY_t - [\alpha_{0t} + \alpha_{2t}L_t]dt = \alpha_{3t}\theta dt + \sigma_Y dZ_t^Y,$$

which are in the receiver's information set, and where the last equalities hold from his perspective. By Theorems 12.6 and 12.7 in Liptser and Shiryaev (1977), his posterior belief is Gaussian with mean  $\hat{M}_t$  and variance  $\gamma_{1t}$  (simply  $\gamma_t$  in the main body) that evolve as

$$d\hat{M}_t = \frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y^2} [d\hat{Y}_t - \alpha_{3t}\hat{M}_t dt] \quad \text{and} \quad \dot{\gamma}_{1t} = -\frac{\gamma_{1t}^2 \alpha_{3t}^2}{\sigma_Y^2}. \quad (\text{A.1})$$

(These expressions still hold after deviations, which go undetected.)

The sender can affect  $\hat{M}_t$  via her choice of actions. Indeed, using that  $d\hat{Y}_t = (a_t - \alpha_{0t} - \alpha_{2t}L_t)dt + \sigma_Y dZ_t^Y$  from her standpoint,

$$d\hat{M}_t = (\kappa_{0t} + \kappa_{1t}a_t + \kappa_{2t}\hat{M}_t)dt + B_t^Y dZ_t^Y, \quad \text{where} \quad (\text{A.2})$$

$$\kappa_{1t} = \alpha_{3t}\gamma_{1t}/\sigma_Y^2, \quad \kappa_{0t} = -\kappa_{1t}[\alpha_{0t} + \alpha_{2t}L_t], \quad \kappa_{2t} = -\alpha_{3t}\kappa_{1t}, \quad B_t^Y = \alpha_{3t}\gamma_{1t}/\sigma_Y. \quad (\text{A.3})$$

On the other hand, since the sender always thinks that the receiver is on path, the public signal evolves, from her perspective, as  $dX_t = (\delta_{0t} + \delta_{1t}\hat{M}_t + \delta_{2t}L_t)dt + \sigma_X dZ_t^X$ . Because the dynamics of  $\hat{M}$  and  $X$  have drifts that are affine in  $\hat{M}$ —with intercepts and slopes that are in the sender's information set—and deterministic volatilities, the pair  $(\hat{M}, X)$  is conditionally

Gaussian. Thus, by the filtering equations in Theorem 12.7 in [Liptser and Shiryaev \(1977\)](#),  $M_t := \mathbb{E}_t[\hat{M}_t]$  and  $\gamma_{2t} := \mathbb{E}_t[(M_t - \hat{M}_t)^2]$  satisfy

$$dM_t = \underbrace{(\kappa_{0t} + \kappa_{1t}a_t + \kappa_{2t}M_t)dt}_{=\mathbb{E}_t[(\kappa_{0t} + \kappa_{1t}a_t + \kappa_{2t}\hat{M}_t)dt]} + \frac{\gamma_{2t}\delta_{1t}}{\sigma_X^2}[dX_t - (\delta_{0t} + \delta_{1t}M_t + \delta_{2t}L_t)dt] \quad (\text{A.4})$$

$$\dot{\gamma}_{2t} = 2\kappa_{2t}\gamma_{2t} + (B_t^Y)^2 - (\gamma_{2t}\delta_{1t}/\sigma_X)^2, \quad (\text{A.5})$$

where  $dZ_t := [dX_t - (\delta_{0t} + \delta_{1t}M_t + \delta_{2t}L_t)dt]/\sigma_X$  is a Brownian motion from the sender's standpoint.<sup>44</sup> Observe that since (A.4) is linear, one can solve for  $M_t$  as an *explicit* function of past actions  $(a_s)_{s<t}$  and past realizations of the public history  $(X_s)_{s<t}$ .

Inserting  $a_t = \beta_{0t} + \beta_{1t}M_t + \beta_{2t}L_t + \beta_{3t}\theta$  in (A.4) and collecting terms yields  $dM_t = [\hat{\kappa}_{0t} + \hat{\kappa}_{1t}M_t + \hat{\kappa}_{2t}L_t + \hat{\kappa}_{3t}\theta]dt + \hat{B}_t dX_t$ , where,

$$\begin{aligned} \hat{\kappa}_{0t} &= \left(\frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y^2}\right)(\beta_{0t} - \alpha_{0t}) - \delta_{0t}\frac{\gamma_{2t}\delta_{1t}}{\sigma_X^2}, & \hat{\kappa}_{1t} &= \left(\frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y^2}\right)(\beta_{1t} - \alpha_{3t}) - \delta_{1t}\frac{\gamma_{2t}\delta_{1t}}{\sigma_X^2}, \\ \hat{\kappa}_{2t} &= \left(\frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y^2}\right)(\beta_{2t} - \alpha_{2t}) - \delta_{2t}\frac{\gamma_{2t}\delta_{1t}}{\sigma_X^2}, & \hat{\kappa}_{3t} &= \left(\frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y^2}\right)\beta_{3t}, \quad \text{and} \quad \hat{B}_t = \frac{\gamma_{2t}\delta_{1t}}{\sigma_X^2}. \end{aligned}$$

Now let  $R(t, s) = \exp(\int_s^t \hat{\kappa}_{1u} du)$ . Since  $M_0 = \mu$ , we have

$$M_t = R(t, 0)\mu + \theta \int_0^t R(t, s)\hat{\kappa}_{3s} ds + \int_0^t R(t, s)[\hat{\kappa}_{0s} + \hat{\kappa}_{2s}L_s] ds + \int_0^t R(t, s)\hat{B}_s dX_s.$$

Imposing (11) yields the equations

$$\chi_t = \int_0^t R(t, s)\hat{\kappa}_{3s} ds \quad \text{and} \quad L_t = \frac{R(t, 0)\mu + \int_0^t R(t, s)[\hat{\kappa}_{0s} + \hat{\kappa}_{2s}L_s] ds + \int_0^t R(t, s)\hat{B}_s dX_s}{1 - \chi_t}.$$

The validity of the construction boils down to finding a solution to the previously stated equation for  $\chi$  that takes values in  $[0, 1)$ . Indeed, when this is the case, it is easy to see that

$$dL_t = \frac{L_t[\hat{\kappa}_{1t} + \hat{\kappa}_{2t} + \hat{\kappa}_{3t}]dt + \hat{\kappa}_{0t}dt + \hat{B}_t dX_t}{1 - \chi_t}, \quad (\text{A.6})$$

from which it is easy to conclude that  $L$  is a (linear) function of  $X$  as conjectured.

We will find a solution to the  $\chi$ -equation that is  $C^1$  with values in  $[0, 1)$ . Differentiating

---

<sup>44</sup>Theorem 12.7 in [Liptser and Shiryaev \(1977\)](#) is stated for actions that depend on  $(\theta, X)$  exclusively, but it also applies to those that condition on past play (i.e., on  $M$ ). Indeed, from (A.2),  $\hat{M}_t = \hat{M}_t^\dagger + A_t$  where  $\hat{M}_t^\dagger = \hat{M}_t^\dagger[Z_t^Y; s < t]$  and  $A_t = \int_0^t e^{\int_0^s \kappa_{2u} du} \kappa_{1s} a_s ds$ . Applying the theorem to  $(\hat{M}_t^\dagger, X_t)_{t \in [0, T]}$ , yields a posterior mean  $M_t^\dagger$  and variance  $\gamma_{2t}^\dagger$  for  $\hat{M}_t^\dagger$  such that  $M_t^\dagger + A_t = M_t$  as in (A.4) and  $\gamma_{2t} = \gamma_{2t}^\dagger$ .

$\chi_t = \int_0^t R(t, s) \hat{\kappa}_{3s} ds$  then yields an ODE for  $\chi$  as below that is coupled with  $\gamma_1$  and  $\gamma_2$ :

$$\dot{\gamma}_{1t} = -\gamma_{1t}^2 (\beta_{3t} + \beta_{1t} \chi_t)^2 / \sigma_Y^2 \quad (\text{A.7})$$

$$\dot{\gamma}_{2t} = -2\gamma_{2t} \gamma_{1t} (\beta_{3t} + \beta_{1t} \chi_t)^2 / \sigma_Y^2 + \gamma_{1t}^2 (\beta_{3t} + \beta_{1t} \chi_t)^2 / \sigma_Y^2 - (\gamma_{2t} \delta_{1t})^2 / \sigma_X^2 \quad (\text{A.8})$$

$$\dot{\chi}_t = \gamma_{1t} (\beta_{3t} + \beta_{1t} \chi_t)^2 (1 - \chi_t) / \sigma_Y^2 - (\delta_{1t} \chi_t) (\gamma_{2t} \delta_{1t}) / \sigma_X^2. \quad (\text{A.9})$$

In the proof of Lemma A.1 (presented next), we take the system above as a primitive and establish that  $\chi = \gamma_2 / \gamma_1$ . Equipped with this, we set  $\gamma_2 = \chi \gamma_1$  in the third ODE, and after writing  $\gamma$  for  $\gamma_1$ , the first and third ODEs become (14)–(15). The same Lemma A.1 establishes that  $0 < \gamma_t \leq \gamma^\circ$  and  $0 \leq \chi_t < 1$ , with strict inequalities for all  $t > 0$  if  $\beta_{3,0} \neq 0$ .

Finally, after plugging in the expressions that define  $(\vec{\kappa}, \hat{B})$ , (A.6) becomes

$$dL_t = (\ell_{0t} + \ell_{1t} L_t) dt + B_t dX_t, \quad \text{where} \quad (\text{A.10})$$

$$\ell_{0t} = -\frac{\gamma_t \chi_t \delta_{0t} \delta_{1t}}{\sigma_X^2 (1 - \chi_t)}, \quad \ell_{1t} = -\frac{\gamma_t \chi_t \delta_{1t} (\delta_{1t} + \delta_{2t})}{\sigma_X^2 (1 - \chi_t)}, \quad B_t = \frac{\gamma_t \chi_t \delta_{1t}}{\sigma_X^2 (1 - \chi_t)}. \quad (\text{A.11})$$

That  $L_t$  coincides with  $\mathbb{E}[\theta | \mathcal{F}_t^X]$  is proved in the Supplementary Appendix.  $\square$

The next lemma shows the existence and uniqueness of a solution to the ODE-system (14)–(15) for  $\gamma$  and  $\chi$ , a property that we exploit in our existence technique. The lemma also establishes the remaining steps from the proof of Lemma 1:

**Lemma A.1** (Learning ODEs). *Suppose that  $(\beta_1, \beta_3, \delta_1)$  are continuous. Then, there is a unique solution to (14)–(15), which satisfies  $0 < \gamma_t \leq \gamma^\circ$  and  $0 \leq \chi_t < 1$  for all  $t \in [0, T]$ , with strict inequalities over  $(0, T]$  if  $\beta_{3,0} \neq 0$ . The same conclusions hold if  $\delta_{1t} = \hat{u}_{a\theta} + \hat{u}_{a\alpha} \alpha_{3t}$ .*

*Proof.* Consider the system in  $(\gamma_1, \gamma_2, \chi)$  from the proof of Lemma 1. By Peano's Theorem, a solution exists in some interval  $[0, T')$  where  $T' > 0$ . And since the system is locally Lipschitz continuous in  $(\gamma_1, \gamma_2, \chi)$  uniformly in  $t \in [0, T]$ , the solution is unique over any interval of existence by the Picard-Lindelöf Theorem. By applying the comparison theorem to  $\gamma_1$  and the zero function, we obtain  $\gamma_1 > 0$ ; and clearly,  $\dot{\gamma}_1 \leq 0$  so  $\gamma_1 \leq \gamma^\circ$ . Hence,  $\gamma_2 / \gamma_1$  is well-defined, and it is easy to verify that it satisfies the  $\chi$ -ODE. Since the solution is unique whenever it exists, we conclude that  $\chi = \gamma_2 / \gamma_1$ , as promised in Lemma 1; in other words,  $\chi_t = \mathbb{E}_t[(M - \hat{M})^2] / \hat{\mathbb{E}}_t[(\theta - \hat{M})^2]$ . We can therefore substitute  $\gamma_2 = \chi \gamma_1$  into (A.7) and (A.9) and abbreviate  $\gamma_1$  to  $\gamma$  to obtain (14)–(15). Next, we apply the comparison theorem to the pairs  $(0, \chi)$  and  $(\chi, 1)$  to obtain  $0 \leq \chi < 1$ .

Using these bounds, we argue that the solution to (14)-(15) exists over  $[0, T]$ . Suppose by way of contradiction that the maximum interval of existence is  $[0, \tilde{T})$ . Then since  $(\gamma, \chi)$  and their derivatives are bounded, the solution can be extended to  $\tilde{T}$ . If  $\tilde{T} = T$ , we are done, and if  $\tilde{T} < T$ , by Peano's Theorem the solution can be further extended to  $\tilde{T} + \epsilon$  for some  $\epsilon > 0$ , contradicting that  $[0, \tilde{T})$  is the maximum interval of existence. We conclude that the solution exists over the whole horizon  $[0, T]$ .

If, moreover,  $\beta_{3,0} \neq 0$ , then  $\dot{\gamma}_{1,0} < 0$  and  $\dot{\chi}_0 > 0$ . Hence, by continuity of  $\dot{\gamma}_1$  and  $\dot{\chi}$ , there exists  $\epsilon > 0$  such that  $\gamma_{1t} < \gamma^o$  and  $\chi_t > 0$  for all  $t \in (0, \epsilon)$ , and by the comparison theorem, these strict inequalities hold up to time  $T$ .

Lastly, suppose that  $\delta_{1t} = \hat{u}_{a\theta} + \hat{u}_{a\hat{a}}\alpha_{3t} = \hat{u}_{a\theta} + \hat{u}_{a\hat{a}}(\beta_{1t}\chi_t + \beta_{3t})$ , where  $(\beta_1, \beta_3)$  are differentiable. Then the system (14)-(15) changes in that the functional form of the operator is altered (since  $\chi$  enters  $\delta_1$ ), but importantly, it still satisfies the conditions for the Peano and Picard-Lindelöf theorems, and the arguments above go through.  $\square$

*Proof of Lemma 2.* Using (A.3), the dynamic (A.4) for  $M$  becomes  $dM_t = \frac{\gamma_t \alpha_{3t}}{\sigma_Y^2} (a_t - [\alpha_{0t} + \alpha_{2t}L_t + \alpha_{3t}M_t])dt + \frac{\chi_t \gamma_t \delta_{1t}}{\sigma_X} dZ_t$ , where  $dZ_t := [dX_t - (\delta_{0t} + \delta_{1t}M_t + \delta_{2t}L_t)dt]/\sigma_X$  a Brownian motion from the sender's standpoint. As for the law of motion of  $L$ , this follows from (A.10) using (A.11) and that  $dX_t = (\delta_{0t} + \delta_{2t}L_t + \delta_{1t}M_t)dt + \sigma_X dZ_t$  from the sender's perspective.

Regarding variances, that  $\mathbb{E}_t[(M_t - \hat{M}_t)^2] = \chi_t \gamma_t$  is independent of deviations by the sender follows from the proof of Lemma 1 (which shows that  $\gamma_2 := \mathbb{E}_t[(M_t - \hat{M}_t)^2]$  depends only on conjectured (linear) strategies) and that of A.1 (which establishes that  $\gamma_2 = \chi\gamma$ ). Finally, that  $\gamma_t^X := \mathbb{E}[(\theta - L_t)^2 | \mathcal{F}_t^X] = \frac{\gamma_t}{1 - \chi_t}$  is established in Lemma S.8 in Section S.4.1. in the Supplementary appendix, for a more general signal structure.

We conclude with three observations about the best-response problem that follow thanks to this lemma. First, from (A.2) and (A.4),  $\hat{M}_t - M_t$  depends only on the conjectured linear Markov strategy by the sender, not on her actual choice of strategy. This means that  $Z_t$  is also unresponsive to deviations due to  $\sigma_X dZ_t = \delta_{1t}(\hat{M}_t - M_t)dt + \sigma_X dZ_t^X$  under the true data-generating process. This 'strategic independence' enables us to fix an exogenous Brownian motion  $Z$  in the laws of motion of  $M$  and  $L$ , and solve the sender's best-response problem with those dynamics to find her best response. This is what happens in the traditional *separation principle* for decision problems in which an unobserved state (here,  $\hat{M}$ ) is controlled—see, for instance, Liptser and Shiryaev (1977), Chapter 16, or Bensoussan (1992), Chapters 2 and 7. An important distinction relative to this literature is that, due to the strategic nature of our environment, estimation and optimization are not completely separated: while in both worlds the optimal control depends on posterior estimates, our estimates non-trivially depend on the optimal control through the players' (correct) conjectures of equilibrium coefficients, which enter the posterior variances. This phenomenon generates a type of feedback loop

that arises only in equilibrium, but not after deviations (because actions are unobserved to the counterparty). We can then split estimation and optimization in sequence as in those decision problems, but the two steps remain non-trivially connected in equilibrium.

Second, from (18), (A.4)–(A.5) and the proof of Lemma A.1, no additional state variables are needed, since  $\gamma_{2t} := \mathbb{E}_t[(M_t - \hat{M}_t)^2] = \chi_t \gamma_t$  is deterministic. Third, as argued in footnote 18, the set of admissible strategies for the best-response problem consists of all square-integrable processes that are progressively measurable with respect to  $(\theta, M, L)$ . This set is the traditional one in stochastic control (Chapters 1.3 and 3.2 in Pham, 2009): it obviously does not restrict to a linear use of the states, and the dynamics of  $M$  and  $L$  admit (strong) solutions in this space, so the strategy profile consisting of an admissible strategy for the sender and a linear Markov one for the receiver are also admissible in the sense of Section 2. Note that this set is richer than that in Definition 1, due to the conditioning on  $M$ .  $\square$

*Proof of Proposition 1.* The proof uses the BVP from Section 5 to find LMEs in the general case. The idea is as follows. Suppose that  $u_{a\hat{a}} = \hat{u}_{a\hat{a}} = \psi_{\hat{a}\hat{a}} = 0$ . First, the players' myopic behavior at time  $T$  implies the terminal values  $\beta_{1T} = \beta_{2T} = \delta_{2T} = v_{6T} = v_{8T} = 0$  in any LME (where  $v_6$  and  $v_8$  are the coefficients on  $L^2$  and  $ML$  in the sender's candidate quadratic value function from Section 5); this can be seen from the myopic equilibrium (25), where only  $\beta_{0T}$  changes if  $\psi_{\hat{a}}(0) \neq 0$  (see Section S.3.2 in the Supplementary Appendix). Now, consider the ODEs in Appendix C that form our BVP: by simple inspection, the constant  $(0, 0, 0, 0)$  satisfies the ODEs and terminal values for  $(\beta_1, \beta_2, v_6, v_8)$ . By the Picard-Lindelöf theorem, these are the unique solutions, so  $a = \beta_{0t} + \beta_{3t}\theta$  and  $\hat{a} = \delta_{0t} + \delta_{1t}\hat{M}$  in any LME. Using the same ODEs, it is easy to see that  $(\beta_3, \gamma)$  are independent of  $(\sigma_X, (\chi_t)_{t \in [0, T]})$ . Regarding  $\beta_0$ , the same is true due to  $v_3 \equiv 0$ , the coefficient on  $L$  in the value function, reflecting that the public signal has no use in this case; see Section S.3.4 in the Supplementary Appendix.

Next, suppose  $Y$  is public and that players are conjectured to follow the strategies above. The effect of making  $Y$  public is that  $M \equiv L \equiv \hat{M}$ , with  $\hat{M}$  being the same function of the history of  $Y$  as in the non-public case. Since  $\delta_2 \equiv 0$ , the sender's best response problem is unchanged, and since  $\beta_1 \equiv \beta_2 \equiv 0$ , the receiver's problem is also unchanged. Thus, the strategy profile continues to be an equilibrium (again, irrespective of  $(\chi_t)_{t \in [0, T]}$  and  $\sigma_X$ ).  $\square$

## Appendix B: Proofs for Section 4

### B.1: Proof of Proposition 2

This proof relies on the proof of Theorem 1 in Section C. Consider the time-reversed ODEs for the strategy coefficients, with  $(\gamma, \chi)$  as an input, and initial conditions given by static



equilibrium that arises at  $t = T$ . One can check that  $v_6 = \sigma_Y^2[-1 + 2\beta_1(1 - \chi) + \alpha_3]/(4\alpha_3\gamma) - v_8/2$  and  $\beta_2 = 1 - \beta_1 - \beta_3$  satisfy the ODEs and initial conditions for  $(v_6, \beta_2)$ ; by the Picard-Lindelöf Theorem, these are the unique solutions, and hence  $\beta_1 + \beta_2 + \beta_3 \equiv 1$ . As for  $\alpha_3$ , note that its terminal value is  $\beta_{1T}\chi_T + \beta_{3T} = \frac{1}{2-\chi_T} > 0$ , and its (backward) ODE is

$$\dot{\alpha}_{3t} = -r\alpha_{3t}[\alpha_{3t}(2 - \chi_t) - 1] + \frac{2\alpha_{3t}^3\gamma_t\chi_t}{\sigma_X^2\sigma_Y^2(1 - \chi_t)} \left\{ \sigma_Y^2\chi_t[1 - \alpha_{3t} - \beta_{1t}(1 - \chi_t)] + \alpha_{3t}\gamma_tv_{8t} \right\}.$$

Applying the comparison theorem to  $\alpha_3$  (going backward in time) establishes  $\alpha_3 > 0$ . Now on the equilibrium path,  $a_t = \beta_{0t} + \alpha_{3t}\theta + \alpha_{2t}L_t = \beta_{0t} + \alpha_{3t}\theta + (1 - \alpha_{3t})L_t$ , where  $\alpha_2 \equiv 1 - \alpha_3$  follows from  $\beta_1 + \beta_2 + \beta_3 = 1$ . The receiver thus plays  $\hat{a}_t = \hat{\mathbb{E}}_t[a_t] = \beta_{0t} + \alpha_{3t}\hat{M}_t + (1 - \alpha_{3t})L_t$ .

To prove the remaining inequalities, we again use the proof of Theorem 1 and (within it) the proof of Theorem C.1. The broad idea is that we can always find non-trivial horizons (depending on parameters), such that the solutions to our ODEs cannot have grown enough to violate the inequalities of interest. Concretely, define  $\rho > 0$  as in Step 2 of the latter proof. For all  $K \in (0, 1)$ , there exists  $T(\gamma^\circ; K) \in \Omega(1/\gamma^\circ)$  such that for all  $T < T(\gamma^\circ; K)$ , there exists a solution to the BVP in  $(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$  (where  $\tilde{\beta}_2$ ,  $\tilde{v}_6$ , and  $\tilde{v}_8$  are defined as in the proof of Theorem 1) with the following properties:  $\chi \in [0, 1]$ ; the myopic coefficients  $(\beta_{1t}^m, \tilde{\beta}_{2t}^m, \beta_{3t}^m) = \left(\frac{1}{2(2-\chi_t)}, \frac{1}{2(2-\chi_t)}, \frac{1}{2}\right)$  are bounded in magnitude by  $\rho$ ;  $(\beta_1, \tilde{\beta}_2, \beta_3)$  differ from their myopic counterparts by at most  $K$  and similarly  $|\tilde{v}_6|, |\tilde{v}_8| \leq K$ ; and  $(\gamma, \chi)$  are Lipschitz continuous with uniform Lipschitz constants that depend on  $\rho$  and  $K$  but not  $T$ , and that are in  $O(\gamma^\circ)$ . Hence, given any constant  $K$ , there exists  $T^\dagger(\gamma^\circ; K) \in \Omega(1/\gamma^\circ)$  such that in addition to the earlier bounds, we also have  $|\gamma_t - \gamma^\circ| \leq K$  and  $|\chi_t| \leq K$ . Define  $g(K) = \frac{K}{4(2-K)}$ , which is an upper bound on  $\|(\beta_{1t}^m, \tilde{\beta}_{2t}^m, \beta_{3t}^m) - (1/4, 1/4, 1/2)\|_\infty$ , where  $\|\cdot\|_\infty$  denotes the sup norm, when  $\chi_t$  takes values in  $[0, K]$ ; observe that  $g(K) \rightarrow 0$  as  $K \rightarrow 0$ . Now for arbitrary  $K > 0$ , for all  $T < T^\dagger(\gamma^\circ; K)$ , by the triangle inequality we have  $\|(\beta_{1t}, \tilde{\beta}_{2t}, \beta_{3t}) - (1/4, 1/4, 1/2)\|_\infty \leq K + g(K)$ . Taking  $K$  sufficiently small ensures that  $\beta_{3t} \in (0, 1)$ ,  $\alpha_{3t} \in (0, 1)$ , and  $\beta_{1t}, \tilde{\beta}_{2t} \in (0, 1/2)$ , and thus  $\beta_2 = (1 - \chi_t)\tilde{\beta}_{2t} \in (0, 1/2)$ .

Next, we prove  $\beta_{3t} \geq 1/2$  and  $\alpha_{3t} > 1/2$  for  $t < T$  for appropriately chosen  $T$ . We show that  $\dot{\beta}_3 < 0$ . Write the (forward) ODE for  $\beta_3$  as  $\dot{\beta}_{3t} = f^{\beta_3}(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$ , where  $f^{\beta_3}$  is of class  $C^1$ . It is easy to check that  $f^{\beta_3}(\mathbf{z}) = -\frac{\gamma^\circ}{16\sigma_Y^2} < 0$ , where  $\mathbf{z} := (\gamma^\circ, 0, 1/4, 1/4, 1/2, 0, 0)$  is the vector of coefficients in the equilibrium of the trivial ( $T = 0$ ) game. For any  $K > 0$  and  $T < T^\dagger(\gamma^\circ; K)$ , by construction,  $\|(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8) - \mathbf{z}\|_\infty \leq K + g(K)$ . Thus by continuity of  $f^{\beta_3}$ , for sufficiently small  $K$ , if  $T < T^\dagger(\gamma^\circ; K)$ ,  $\dot{\beta}_{3t} < 0$  for all  $t \in [0, T]$ . Given  $\beta_{3T} = 1/2$ , this implies  $\beta_{3t} \geq 1/2$  for all  $t$  and  $\beta_{3t} > 1/2$  for  $t < T$ . In turn, for all  $t > 0$  we have  $\beta_{1t}\chi_t > 0$  and thus  $\alpha_{3t} > \beta_{3t} \geq 1/2$ . Since  $\alpha_{3,0} = \beta_{3,0} > 1/2$ ,  $\alpha_{3t} > 1/2$  for all  $t$ .

To show  $\beta_{0t} < k$  for  $t < T$  ( $T$  depending on parameters), it suffices to show  $\dot{\beta}_{0t} > 0$

(because  $\beta_{0T} = k$ ). Let  $\alpha_{3t}^m = \frac{1}{2-\chi_t}$  and  $\tilde{v}_3 = \frac{\gamma v_3}{1-\chi}$ . The (forward) ODEs for  $(\beta_0, \tilde{v}_3)$  are

$$\begin{aligned}\dot{\beta}_{0t} &= r \frac{\alpha_{3t}}{\alpha_{3t}^m} (\beta_{0t} - k) + \frac{\alpha_{3t}^2 \gamma_t}{\sigma_Y^2 \sigma_X^2} \left[ k \sigma_X^2 - 2 \tilde{v}_{3t} \alpha_{3t} \chi_t - 2 (\beta_{0t} - k) \chi_t (\tilde{v}_{8t} \alpha_{3t} + \sigma_Y^2 \tilde{\beta}_{2t} \chi_t) \right] \\ \dot{\tilde{v}}_{3t} &= \frac{\beta_{0t} \gamma_t (1 - \alpha_{3t})}{2(1 - \chi_t)} + \tilde{v}_{3t} \left( r + \frac{\alpha_{3t}^2 \gamma_t \chi_t}{\sigma_X^2} \right).\end{aligned}$$

(See `spm.nb` on our websites.) Write the  $\beta_0$ -ODE by  $\dot{\beta}_0 = f^{\beta_0}(\beta_0, \tilde{v}_3, \gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$ , where  $f^{\beta_0}$  is of class  $C^1$ . It is easy to check that  $f^{\beta_0}(k, 0, \mathbf{z}) = \frac{k\gamma^\circ}{4\sigma_Y^2} > 0$  for  $\mathbf{z}$  as above. Extending the bounding arguments above to encompass  $(\beta_0, \tilde{v}_3)$ , for any  $K$  there exists  $T^\dagger(\gamma^\circ; K) \in \Omega(1/\gamma^\circ)$  such that for  $T < T^\dagger(\gamma^\circ; K)$ ,  $\|(\beta_0, \tilde{v}_3, \gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8) - (k, 0, \mathbf{z})\|_\infty \leq K + g(K)$ ; thus, by continuity of  $f^{\beta_0}$ , for  $K$  sufficiently small,  $\dot{\beta}_{0t} > 0$  for all  $t$ .

## B.2: Proofs of Propositions 3 and 4

The full details are in Section S.1 in the Supplementary appendix. We describe them briefly.

**(I)** The proof of Proposition 3(i)—finding horizons such that  $\beta_0$  when  $\sigma_X > 0$  is always larger than its counterpart when  $Y$  is public (or  $\sigma_X = 0$ ),  $\beta_0^{pub}$ —is similar to the proof that establishes the bounds for the strategy coefficients in Proposition 2. The idea is that  $\beta_{0T} = \beta_{0T}^{pub} = k$  while  $\dot{\beta}_{0T} < \dot{\beta}_{0T}^{pub}$ . Thus,  $\beta_{0t} > \beta_{0t}^{pub}$  in a neighborhood  $(T - \epsilon, T]$ . Hence, for  $T$  depending on parameters, one can ensure that the solutions do not violate this ranking.

**(II)** The remaining statements pertain to  $\sigma_X \in \{0, +\infty\}$ . LMEs can be found as follows.

**Case  $\sigma_X = 0$ .** We characterize an LME where  $a_t = \beta_{0t} + \beta_{1t} \hat{M}_t + \beta_{3t} \theta$  and  $\hat{a}_t = \hat{\mathbb{E}}_t[a_t] = \beta_{0t} + (\beta_{1t} + \beta_{3t}) \hat{M}_t$ . Further, we prove that in any such an equilibrium must satisfy  $\beta_1 + \beta_3 \equiv 1$ , so the sender effectively recovers  $\hat{M}$ . The laws of motion for the receiver's belief are

$$d\hat{M}_t = \frac{\beta_{3t} \gamma_t}{\sigma_Y^2} \left\{ dY_t - \underbrace{[(\beta_{0t} + (\beta_{1t} + \beta_{3t}) \hat{M}_t) dt]}_{=\hat{\mathbb{E}}_t[a_t]} \right\} \quad \text{and} \quad \dot{\gamma}_t = - \left( \frac{\beta_{3t} \gamma_t}{\sigma_Y} \right)^2, \quad (\text{B.1})$$

with  $\hat{M}_0 = \mu$  and  $\gamma_0 = \gamma^\circ$ . Clearly,  $(\theta, \hat{M}_t, t)$  are the relevant states for the sender's problem, as  $\hat{M}_t$  is public. Let  $V$  denote the sender's value function. The HJB equation is

$$rV = \sup_{a \in \mathbb{R}} \left\{ \frac{1}{4} [-(a - \theta)^2 - (a - \hat{a}_t - k)^2] + \frac{\beta_{3t} \gamma_t}{\sigma_Y^2} [a - \beta_{0t} - (\beta_{1t} + \beta_{3t}) m] V_m + \frac{\beta_{3t}^2 \gamma_t^2}{2\sigma_Y^2} V_{mm} + V_t \right\}.$$

We guess  $V(\theta, m, t) = v_{0t} + v_{1t} \theta + v_{2t} m + v_{3t} \theta^2 + v_{4t} m^2 + v_{5t} \theta m$  to derive a system of ODEs for  $(\beta_0, \beta_1, \beta_3)$  subject to terminal conditions  $(\beta_{0T}, \beta_{1T}, \beta_{3T}) = (k, 1/2, 1/2)$ ; these ODEs depend on  $\gamma$ , which evolves according to (B.1) with initial condition  $\gamma_0 = \gamma^\circ$ . This defines a

BVP in  $(\beta_1, \beta_3, \gamma)$ . Equipped with a solution, it is easy to recover  $\beta_0$  and the value function coefficients; showing that  $\beta_1 + \beta_3 \equiv 1$  is done using the same approach as in Proposition 2.

Regarding existence, with only one ODE going forward we can use a shooting method. Specifically, we construct an IVP consisting of ODEs for  $(\beta_1, \beta_3, \gamma)$  in reversed time, and with an auxiliary variable  $\gamma^F$  for the (reversed)  $\gamma$  ODE. Via an intermediate-value theorem argument as in Bonatti et al. (2017), there must be a  $\gamma^F > 0$  such that  $\gamma_T = \gamma^o$  while all the other ODEs in the IVP admit solutions; see the Supplementary Appendix for the details.

**Case  $\sigma_X = +\infty$**  Since  $L \equiv \mu$  in this case, we look for an equilibrium in which the sender plays  $a_t = \beta_{0t} + \beta_{1t}M_t + \beta_{2t}\mu + \beta_{3t}\theta$ . The following representation result holds:

**Lemma B.1** (Belief Representation). *Assume  $\sigma_X = +\infty$ . Suppose the receiver expects  $a_t = \alpha_{0t} + \alpha_{2t}\mu + \alpha_{3t}\theta$ , where  $\alpha_0 = \beta_0$ ,  $\alpha_2 = \beta_2 + \beta_1(1 - \chi)$ ,  $\alpha_3 = \beta_3 + \beta_1\chi$ ,  $\chi = 1 - \gamma/\gamma^o$ , and  $\dot{\gamma}_t := \hat{\mathbb{E}}_t[(\theta - \hat{M}_t)^2]$ . Then  $\dot{\gamma}_t = -\left(\frac{\gamma_t\alpha_{3t}}{\sigma_Y^2}\right)^2$ . Moreover, if the sender follows  $a_t = \alpha_{0t} + \alpha_{2t}\mu + \alpha_{3t}\theta$ , then  $M_t = \chi_t\theta + (1 - \chi_t)\mu$  holds at all times.*

That the states  $(\theta, M_t, t)$  are sufficient on and off path for the sender is as in the paper—refer to the Supplementary Appendix for the details and the proof of the previous lemma.

The sender then controls  $M$ , which evolves as  $dM_t = \frac{\alpha_{3t}\dot{\gamma}_t}{\sigma_Y^2}(a - \alpha_{0t} - \alpha_{2t}\mu - \alpha_{3t}M_t)dt$ , where  $M_0 = \mu$ ,  $\dot{\gamma}_t = -\left(\frac{\alpha_{3t}\dot{\gamma}_t}{\sigma_Y^2}\right)^2$ ,  $\alpha_3 = \beta_3 + \beta_1\chi$ , and  $\chi = 1 - \gamma/\gamma^o$ . The HJB equation is

$$rV(\theta, m, \mu, t) = \sup_{a \in \mathbb{R}} \left\{ \frac{1}{4}[-(a - \theta)^2 - (a - k - \alpha_{0t} - \alpha_{2t}\mu - \alpha_{3t}m)^2 - \alpha_{3t}^2\gamma_t\chi_t] + V_t + \frac{\alpha_{3t}\dot{\gamma}_t}{\sigma_Y^2}(a - \alpha_{0t} - \alpha_{2t}\mu - \alpha_{3t}m)V_m \right\}.$$

We guess  $V(\theta, m, \mu, t) = v_{0t} + v_{1t}\theta + v_{2t}m + v_{3t}\mu + v_{4t}\theta^2 + v_{5t}m^2 + v_{6t}\mu^2 + v_{7t}\theta m + v_{8t}\theta\mu + v_{9t}m\mu$  and take analogous steps to those in the  $\sigma_X = 0$  case, leading to a BVP for  $(\beta_0, \beta_1, \beta_2, \beta_3, \gamma)$  (because of the one-to-one mapping between  $\gamma$  and  $\chi$ ). With only one ODE going forward, the argument for existence is analogous  $\sigma_X = 0$  case. Finally, we can show that  $\beta_1 + \beta_2 + \beta_3 \equiv 1$  and  $a_t = \beta_{0t} + (1 - \alpha_{3t})\mu + \alpha_{3t}\theta$ , with  $\alpha_3 \in (1/2, 1)$ .

(III) The comparisons of signaling coefficients and inflationary biases when  $\sigma_X \in \{0, +\infty\}$  exploit the use of analytic solutions when  $r = 0$ . See Sections S.1.4 and S.1.5.

## Appendix C: Proofs for Section 5

**Overview of approach** Our overall proof strategy consists of reducing the HJB equation (23) subject to the equilibrium condition (24) to a suitable boundary value problem that we

then solve using a fixed-point argument. The BVP will contain ODEs linked to behavior—hence, involving terminal conditions—and also the learning ODEs for  $(\gamma, \chi)$  that have initial conditions. The fixed point will be over pairs of functions  $(\gamma, \chi)$ : a pair  $(\gamma^*, \chi^*)$  that generates mutual best responses that in turn induce learning ODEs whose solution is  $(\gamma^*, \chi^*)$ .

This overarching goal requires several intermediate steps, which we label *core subsystem*, *centering*, *auxiliary variable*, *fixed point* and *verification*; we provide brief explanations of these when they arise. Throughout the proof, we refer to the *myopic equilibrium coefficients*

$$(\beta_{0t}^m, \beta_{1t}^m, \beta_{2t}^m, \beta_{3t}^m) = \left( \frac{u_{a0} + u_{a\hat{a}}\hat{u}_{\hat{a}0}}{1 - u_{a\hat{a}}\hat{u}_{a\hat{a}}}, \frac{u_{a\hat{a}}(u_{a\theta}\hat{u}_{a\hat{a}} + \hat{u}_{\hat{a}\theta})}{1 - u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t}, \frac{u_{a\hat{a}}^2\hat{u}_{a\hat{a}}(u_{a\theta}\hat{u}_{a\hat{a}} + \hat{u}_{\hat{a}\theta})(1 - \chi_t)}{(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t)}, u_{a\theta} \right),$$

which correspond to the sender's strategy coefficients in the unique linear Bayes Nash equilibrium involving states  $(\theta, M, \hat{M}, L)$  of the static game with flow utilities  $(u, \hat{u})$  if the receiver believes  $M_t = \chi_t\theta_t + (1 - \chi_t)L_t$ . By Assumption 2,  $(\beta_{0t}^m, \beta_{1t}^m, \beta_{2t}^m, \beta_{3t}^m)$  is well defined and  $\alpha_t^m := \beta_{1t}^m\chi_t + \beta_{3t}^m \neq 0$  for all  $\chi_t \in [0, 1]$ . Henceforth, given  $\chi_t$ , we write  $\beta_{it}^m$  and  $\alpha_t^m$  to refer to these functions of  $\chi_t$ , suppressing the dependence on  $\chi_t$ , and we abbreviate  $\alpha_3$  to  $\alpha$ .

**Core subsystem:** *We show that the problem of existence of LME reduces to a core subsystem in  $(\gamma, \chi, \vec{\beta}, v_6, v_8)$ , where  $\vec{\beta} := (\beta_1, \beta_2, \beta_3)$ , and perform a change of variables for  $(\beta_2, v_6, v_8)$ ; we denote the new system by  $(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$ .*

The first thing to note is that  $\alpha_t := \beta_{1t}\chi_t + \beta_{3t} \neq 0$  for all  $t \in [0, T]$  in any LME. Indeed, if  $\alpha_t = 0$ , it is then easy to verify from the HJB equation that  $\beta_{it} = \beta_{it}^m$  for  $i \in \{0, 1, 2, 3\}$ : since the sender's actions transmit no information, both players must be using myopic best responses. But this implies that  $\alpha_t = \alpha_t^m \neq 0$  in such an LME, a contradiction. Second, since the coefficients  $(\beta_0, \beta_1, \beta_2, \beta_3)$  and  $\chi$  will be continuous, it follows that  $\gamma_t > 0$  at all times by Lemma A.1. From the HJB equation, it is easy to see that

$$v_{2t} = -\sigma_Y^2 [u_{a0} + u_{a\hat{a}}\hat{u}_{\hat{a}0} - (1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})\beta_{0t}] / (\alpha_t\gamma_t) \quad (\text{C.1})$$

$$v_{5t} = -\sigma_Y^2 [u_{a\hat{a}}\hat{u}_{\hat{a}\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}\alpha_t - \beta_{1t}] / (2\alpha_t\gamma_t) \quad (\text{C.2})$$

$$v_{7t} = -\sigma_Y^2 [u_{a\theta} - \beta_{3t}] / (\alpha_t\gamma_t) \quad (\text{C.3})$$

$$v_{9t} = -\sigma_Y^2 [u_{a\hat{a}}\hat{u}_{a\hat{a}}\beta_{1t}(1 - \chi_t) - \beta_{2t}(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})] / (\alpha_t\gamma_t). \quad (\text{C.4})$$

Expressions (C.1)-(C.4) allow us to eliminate  $v_i$  and  $\dot{v}_i$ ,  $i \in \{2, 5, 7, 9\}$ , in the HJB equation to get a system of ODEs for  $(\gamma, \chi, \beta_0, \vec{\beta}, v_0, v_1, v_3, v_4, v_6, v_8)$ —as a last step we verify that our  $(\alpha, \gamma)$  satisfy  $|\alpha_t||\gamma_t| > 0$  all  $t \in [0, T]$ , recovering the value function through (C.1)-(C.4).

The full system of ODEs can be found in the Mathematica file `spm.nb` on our websites— we omit them in favor of stating the core subsystem with which we will be working below. The omitted system has three properties easily verified by inspection in the same file:

- (i) the ODEs for  $(\vec{\beta}, v_6, v_8)$  do not contain  $(v_0, v_1, v_3, v_4, \beta_0)$ ;
- (ii) given  $(\vec{\beta}, v_6, v_8)$ ,  $(v_0, v_1, v_3, v_4, \beta_0)$  form a non-homogeneous linear ODE system; and
- (iii)  $(\vec{\beta}, v_6, v_8)$  carries  $(1 - \chi)$  in the denominator.

Parts (i) and (ii) imply that we can focus on the sub-system  $(\vec{\beta}, v_6, v_8)$ , as any linear system with continuous coefficients admits a unique solution for all times (Teschl, 2012, Corollary 2.6).<sup>45</sup> Part (iii) reflects that the dynamic for  $L$  carries a denominator of that form; by Lemma A.1, however, we know that  $\chi \in [0, 1)$  if the coefficients are continuous.

It is then convenient to use the change of variables  $(\tilde{\beta}_2, \tilde{v}_6, \tilde{v}_8) = (\beta_2/(1 - \chi), v_6\gamma/(1 - \chi)^2, v_8\gamma/(1 - \chi))$  that eliminates this denominator in the resulting system for the functions  $(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$ —because  $(\chi, \gamma)$  only depend on  $(\beta_1, \beta_3)$  directly, it follows that  $\chi \in [0, 1)$  and  $\gamma > 0$  in any solution to this system, and we trivially recover  $(\beta_2, v_6, v_8)$ .<sup>46</sup>

We can now state the core subsystem of ODEs for  $(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$  with which we will be working. Recall that  $\delta_{1t} = \hat{u}_{a\theta} + \hat{u}_{a\hat{a}}(\beta_{1t}\chi_t + \beta_{3t})$ .

$$\begin{aligned}
\dot{\tilde{v}}_{6t} &= \tilde{v}_{6t}[r + \alpha_t^2\gamma_t/\sigma_Y^2 + 2\delta_{1t}^2\gamma_t\chi_t/\sigma_X^2] - (\gamma_t/2) \left\{ \beta_{1t}^2\hat{u}_{a\hat{a}}[2u_{a\hat{a}} + u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}}] \right. \\
&\quad \left. + \tilde{\beta}_{2t}(2\beta_{1t} + \tilde{\beta}_{2t})[-1 + 2u_{a\hat{a}}\hat{u}_{a\hat{a}} + u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}}^2] \right\} \\
\dot{\tilde{v}}_{8t} &= \tilde{v}_{8t}[r + \delta_{1t}^2\gamma_t\chi_t/\sigma_X^2] - \gamma_t \left\{ (\tilde{\beta}_2 + \beta_{1t})[u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}] - \beta_{1t}\beta_{3t} \right\} \\
\dot{\beta}_{1t} &= r\frac{\alpha_t}{\alpha_t^m}[\beta_{1t} - \beta_{1t}^m] - \gamma_t[\sigma_X^2\sigma_Y^2(u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t)]^{-1} \times \\
&\quad \left\{ \tilde{\beta}_{2t}2(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})\sigma_Y^2\delta_{1t}^2\chi_t(u_{a\theta} + \beta_{1t}\chi_t) + \beta_{1t}^2[\sigma_X^2\alpha_t(u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t) + (1 - 2u_{a\hat{a}}\hat{u}_{a\hat{a}})\sigma_Y^2\delta_{1t}^2\chi_t^2] \right. \\
&\quad + \beta_{1t}\sigma_X^2\alpha_t[\hat{u}_{a\hat{a}}(u_{a\hat{a}} + u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}})\alpha_t^2\chi_t + \hat{u}_{a\theta}(u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t) + \alpha_t(-u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}} + 2u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}}\hat{u}_{a\theta}\chi_t)] \\
&\quad - \beta_{1t}\sigma_Y^2u_{a\hat{a}}\delta_{1t}^2\chi_t(2u_{a\theta}\hat{u}_{a\hat{a}} + \hat{u}_{a\theta}\chi_t) + \delta_{1t}^2\tilde{v}_{8t}\alpha_t\chi_t(\beta_{1t} - u_{a\hat{a}}\hat{u}_{a\hat{a}}) \\
&\quad \left. - \sigma_X^2\delta_{1t}\alpha_t[u_{a\hat{a}}(u_{a\theta}\hat{u}_{a\hat{a}} - u_{a\theta}\alpha_t) - u_{\hat{a}\hat{a}}u_{a\theta}\delta_{1t}] \right\}. \\
\dot{\tilde{\beta}}_{2t} &= r\frac{\alpha_t}{\alpha_t^m}[\tilde{\beta}_{2t} - \tilde{\beta}_{2t}^m] - \gamma_t[\sigma_X^2\sigma_Y^2(u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t)(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})]^{-1} \times \\
&\quad \left\{ \delta_{1t}^2\alpha_t\chi_t[2\tilde{v}_{6t}(u_{a\theta} + u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t) - u_{a\hat{a}}^2\hat{u}_{a\hat{a}}\hat{u}_{a\theta}\tilde{v}_{8t}] \right. \\
&\quad + \tilde{\beta}_{2t}\sigma_X^2\alpha_t[\hat{u}_{a\hat{a}}(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})(u_{a\hat{a}} + u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}})\alpha_t^2\chi_t + \hat{u}_{a\theta}(u_{a\theta} + u_{a\hat{a}}u_{a\theta} - u_{a\hat{a}}u_{a\theta}\hat{u}_{a\hat{a}} + u_{\hat{a}\hat{a}}u_{a\theta}\hat{u}_{a\hat{a}} \\
&\quad \left. + [u_{a\hat{a}}^2 + u_{\hat{a}\hat{a}}]\hat{u}_{a\theta}\chi_t) + \alpha_t(u_{a\theta}\hat{u}_{a\hat{a}}[1 - u_{a\hat{a}}\hat{u}_{a\hat{a}}] + u_{a\theta}[-1 + 2u_{a\hat{a}}\hat{u}_{a\hat{a}} + u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}}^2]) \right\}
\end{aligned}$$

<sup>45</sup>Intuitively,  $(v_0, v_1, v_4)$  are the coefficients of the constant,  $\theta$ - and  $\theta^2$ -terms in the sender's value function, none of which the sender controls, so they do not affect the rest of the system. The equations for  $(\beta_0, v_3)$  are coupled and encode the *deterministic* component of the sender's incentive to manipulate beliefs; they do not enter the sub-system for  $(\vec{\beta}, v_6, v_8)$  but depend on the latter through the signal-to-noise ratio in  $Y$ .

<sup>46</sup>Our method for finding intervals of existence of LME relies on bounding solutions to ODEs uniformly, and this denominator would unnecessarily complicate that task since there is no upper bound on  $1/(1 - \chi)$  that applies to all environments. This change of variables is akin to working with  $\tilde{L} = (1 - \chi)L$  instead of  $L$ .

$$\begin{aligned}
& + \hat{u}_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t [u_{a\hat{a}}^2 + 2u_{\hat{a}\hat{a}} - u_{a\hat{a}} u_{\hat{a}\hat{a}} \hat{u}_{a\hat{a}}]) + \delta_{1t} [\sigma_X^2 u_{a\hat{a}} \hat{u}_{a\hat{a}} \alpha_t (u_{a\theta} u_{\hat{a}\hat{a}} \delta_{1t} + u_{a\hat{a}} [u_{a\theta} \alpha_t - u_{\hat{a}\theta} \hat{u}_{a\hat{a}}])] \\
& - \tilde{\beta}_{2t} \sigma_Y^2 (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}}) \delta_{1t}^2 \chi_t [u_{a\theta} (1 - 2u_{a\hat{a}} \hat{u}_{a\hat{a}}) + \chi_t (u_{a\hat{a}} \hat{u}_{\hat{a}\theta} - \beta_{1t} [1 - 2u_{a\hat{a}} \hat{u}_{a\hat{a}}])] \\
& + \alpha_t (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}}) [\tilde{\beta}_{2t} \tilde{v}_{8t} \delta_{1t}^2 \chi_t - \sigma_X^2 \beta_{1t}^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)] + 2\sigma_Y^2 \delta_{1t}^2 \tilde{\beta}_{2t}^2 \chi_t^2 (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}})^2 \\
& + \beta_{1t} \delta_{1t} [\sigma_X^2 \alpha_t (u_{a\hat{a}} + u_{\hat{a}\hat{a}} \hat{u}_{a\hat{a}}) (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t) + \sigma_Y^2 \delta_{1t} \chi_t u_{a\hat{a}} u_{a\theta} \hat{u}_{a\hat{a}} (1 - 2u_{a\hat{a}} \hat{u}_{a\hat{a}})] \Big\} \\
\dot{\beta}_{3t} = & r \frac{\alpha_t}{\alpha_t^m} [\beta_{3t} - \beta_{3t}^m] - \gamma_t [\sigma_X^2 \sigma_Y^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)]^{-1} \times \\
& \Big\{ \tilde{\beta}_{2t} 2(1 - u_{a\hat{a}} \hat{u}_{a\hat{a}}) \sigma_Y^2 \delta_{1t}^2 \chi_t^2 (\beta_{3t} - u_{a\theta}) - \beta_{1t}^2 \chi_t [\sigma_X^2 \alpha_t (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t) + \sigma_Y^2 \delta_{1t}^2 \chi_t^2 (1 - 2u_{a\hat{a}} \hat{u}_{a\hat{a}})] \\
& - \beta_{1t} \alpha_t \sigma_X^2 [\hat{u}_{a\hat{a}} (u_{a\hat{a}} + u_{\hat{a}\hat{a}} \hat{u}_{a\hat{a}}) \alpha_t^2 \chi_t^2 + \hat{u}_{\hat{a}\theta} \chi_t (u_{\hat{a}\theta} + u_{\hat{a}\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)] \\
& + \alpha_t ([u_{\hat{a}\theta} \hat{u}_{a\hat{a}} - u_{a\hat{a}} \hat{u}_{\hat{a}\theta}] \chi_t - u_{a\theta} + [u_{a\hat{a}} + 2u_{\hat{a}\hat{a}} \hat{u}_{a\hat{a}}] \hat{u}_{\hat{a}\theta} \chi_t^2) - \beta_{1t} \delta_{1t}^2 \chi_t^2 \sigma_Y^2 (1 - 2u_{a\hat{a}} \hat{u}_{a\hat{a}}) (u_{a\theta} - \alpha_t) \\
& + \delta_{1t}^2 \alpha_t \chi_t \tilde{v}_{8t} (\beta_{3t} + \chi_t u_{a\hat{a}} \hat{u}_{\hat{a}\theta}) + \sigma_X^2 \delta_{1t} \alpha_t [(u_{a\hat{a}} u_{\hat{a}\theta} - u_{\hat{a}\hat{a}} u_{a\theta}) \hat{u}_{\hat{a}\theta} \chi_t + (u_{a\hat{a}} + u_{\hat{a}\hat{a}} \hat{u}_{a\hat{a}}) \alpha_t^2 \chi_t \\
& + \alpha_t (u_{\hat{a}\theta} - \chi_t [u_{a\theta} (u_{a\hat{a}} + u_{\hat{a}\hat{a}} \hat{u}_{a\hat{a}}) - u_{\hat{a}\hat{a}} \hat{u}_{\hat{a}\theta}]) \Big\} \\
\dot{\gamma}_t = & -(\beta_{1t} \chi_t + \beta_{3t})^2 \gamma_t^2 / \sigma_Y^2, \quad \dot{\chi}_t = \gamma_t [(\beta_{1t} \chi_t + \beta_{3t})^2 (1 - \chi_t) / \sigma_Y^2 - \delta_{1t}^2 \chi_t^2 / \sigma_X^2].
\end{aligned}$$

This system has two initial conditions  $(\gamma_0, \chi_0) = (\gamma^\circ, 0)$ . It also has terminal conditions for  $(\beta_{1T}, \tilde{\beta}_{2T}, \beta_{3T}, \tilde{v}_{6T}, \tilde{v}_{8T})$  that depend on whether there are terminal payoffs. In what follows, we focus on the case without terminal payoffs—i.e., where the terminal conditions are  $(\beta_{1T}^m, \tilde{\beta}_{2T}^m, \beta_{3T}^m, 0, 0)$ —postponing the discussion of terminal payoffs to the end of the analysis. We note that the remaining denominators never vanish thanks to Assumption 2, and that all the ODEs carry  $r$ -independent terms that scale linearly in  $\gamma$ ; this latter property will allow us to find horizons for existence that are inversely proportional to  $\gamma^\circ$ .

**Centering:** *To exploit discounting, we focus on the centered system  $(\gamma, \chi, \beta_1^c, \tilde{\beta}_2^c, \beta_3^c, \tilde{v}_6, \tilde{v}_8)$ , where  $(\beta_1^c, \tilde{\beta}_2^c, \beta_3^c)$  denotes  $(\beta_1, \tilde{\beta}_2, \beta_3)$  net of the myopic counterpart.* The tuple  $(\beta_1, \tilde{\beta}_2, \beta_3)$  is constructed going backward in time from its terminal value as with backward induction in discrete time. One would expect higher discount rates to pull these coefficients towards the myopic values more strongly, thereby facilitating the existence of LME. Indeed, the term  $-r \frac{\alpha}{\alpha^m} (\beta_i - \beta_i^m)$  in the time-reversed version of the  $\beta_i$ -ODE reflects this fact as long as  $\alpha := \beta_1 \chi + \beta_3$  does not change sign. To exploit the effect of discounting when finding intervals of existence, it is then useful to introduce the *centered* coefficients, i.e.,  $x_{it}^c := x_{it} - x_{it}^m$  for  $x \in \{\beta_1, \tilde{\beta}_2, \beta_3\}$ , and work with the ODEs of  $(\beta_1^c, \tilde{\beta}_2^c, \beta_3^c, \tilde{v}_6, \tilde{v}_8)$  in backward form.<sup>47</sup>

The next lemma states the key properties of this backward centered system, noting that (i) the RHS of the ODEs for  $(\beta_1, \tilde{\beta}_2, \beta_3)$  above are polynomials in  $(\beta_1, \tilde{\beta}_2, \beta_3) = (\beta_1^c + \beta_1^m, \tilde{\beta}_2^c + \tilde{\beta}_2^m, \beta_3^c + \beta_3^m)$ , (ii)  $(\beta_1^m, \tilde{\beta}_2^m, \beta_3^m)$  are functions of  $\chi$  and are independent of  $r$ , (iii)  $(\dot{\beta}_1^m, \dot{\tilde{\beta}}_2^m, \dot{\beta}_3^m)$

<sup>47</sup>Note that a backward first-order ODE of a function  $f$  is obtained by differentiating  $\tilde{f} = f(T - t)$ , and hence only differs with the original one in the sign. We maintain the labels to avoid notational burden.

carry a factor of  $\gamma$  through  $\dot{\chi}$ , and (iv)  $\alpha_t^m = \frac{u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t}{1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t}$ . (The proof is straightforward and hence omitted.) Without fear of confusion, in the lemma and in what follows we denote the solution to the backward system by  $(\beta_1^c, \tilde{\beta}_2^c, \beta_3^c, \tilde{v}_6, \tilde{v}_8)$  (and unless otherwise stated, we always refer to the backward system when invoking this tuple). Also, let  $\vec{\beta}^c := (\beta_1^c, \tilde{\beta}_2^c, \beta_3^c)$ .

**Lemma C.1.** *For  $x \in \{\beta_1, \tilde{\beta}_2, \beta_3\}$  and  $y \in \{\tilde{v}_6, \tilde{v}_8\}$ , the (backward) ODEs that  $x^c$  and  $y$  satisfy have the form*

$$\begin{aligned} \dot{x}_t^c &= -r x_t^c \frac{\alpha_t}{\alpha_t^m} + \frac{\gamma_t h_x(\vec{\beta}^c, \tilde{v}_{6t}, \tilde{v}_{8t}, \chi_t)}{\sigma_X^2 \sigma_Y^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)^{n_{1,x}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t)^{n_{2,x}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}})^{n_{3,x}}} \\ \dot{y}_t &= -y_t [r + \gamma_t R_y(\vec{\beta}^c, \tilde{v}_{6t}, \tilde{v}_{8t}, \chi_t)] + \frac{\gamma_t h_y(\vec{\beta}^c, \chi_t)}{\sigma_X^2 \sigma_Y^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)^{n_{1,y}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t)^{n_{2,y}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}})^{n_{3,y}}}, \end{aligned}$$

where  $n_{i,x}, n_{i,y} \in \mathbb{N}$ ,  $i = 1, 2, 3$ , and  $h_x, h_y$ , and  $R_y \geq 0$  are polynomials.<sup>48</sup> The initial conditions are  $(\vec{\beta}_0^c, \tilde{v}_{60}, \tilde{v}_{80}) = (0, 0, 0, 0, 0)$ .

In particular, notice that (i) the terms not containing  $r$  continue scaling with  $\gamma$ , (ii) the denominators are bounded away from zero, and (iii) the discount rate term pushes any solution towards zero when  $\alpha$  does not change sign. We turn to this issue in the next step.

**Auxiliary variable:** *To exploit discounting, we introduce an auxiliary variable  $\tilde{\alpha} \neq 0$  and work with an ODE-system for  $(\gamma, \chi, \beta_1^c, \tilde{\beta}_2^c, \beta_3^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$ . Observe that  $\alpha$  will indeed never vanish in any solution to the centered system. In fact, a tedious but straightforward exercise shows that in *backward* form,  $\alpha = \beta_1 \chi + \beta_3$  satisfies*

$$\begin{aligned} \dot{\alpha}_t &= \alpha_t \left\{ -r \left( \frac{\alpha_t}{\alpha_t^m} - 1 \right) + \gamma_t [\sigma_X^2 \sigma_Y^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)]^{-1} \times \right. \\ &\quad \left\{ \delta_{1t} [\beta_{1t} \chi_t + \beta_{3t}] \sigma_X^2 u_{\hat{a}\theta} + \delta_{1t} \chi_t [\delta_{1t} \chi_t \sigma_Y^2 (2\tilde{\beta}_{2t} [1 - u_{a\hat{a}} \hat{u}_{a\hat{a}}] + \beta_{1t} [1 - 2u_{a\hat{a}} \hat{u}_{a\hat{a}}]) \right. \\ &\quad \left. \left. + (\beta_{1t} \chi_t + \beta_{3t}) (\delta_{1t} \tilde{v}_{8t} + \sigma_X^2 [u_{\hat{a}\hat{a}} \delta_{1t} + u_{a\hat{a}} (\beta_{1t} \chi_t + \beta_{3t})]) \right] \right\}, \end{aligned} \quad (\text{C.5})$$

with initial condition  $\alpha_0 = \alpha_0^m = \frac{u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_0}{1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_0}$  (here, for consistency,  $\chi_0$  is the terminal value of  $\chi$  going forward in time). By Assumption 2,  $\alpha_0^m$  always has the same sign as  $u_{a\theta}$  because  $\chi_0 \in [0, 1]$ . Also, the right-hand side of (C.5) is proportional to  $\alpha$ , so it vanishes at  $\alpha \equiv 0$ . By the comparison theorem,  $\alpha$  is always nonzero, as the ODE is locally Lipschitz continuous in  $\alpha$  uniformly in time. Moreover, since  $\alpha^m$  never changes sign,  $\alpha/\alpha^m > 0$ .

However, our fixed point argument will input *general*  $(\gamma, \chi)$  pairs into the backward ODEs of Lemma C.1, pairs that need not solve the learning ODEs (or even be differentiable). Thus, we will not be able to use a comparison argument like that above to show that each induced

<sup>48</sup>More precisely, we have  $n_{1,x} = 1$ ,  $n_{1,y} = 0$ , and  $n_{3,\beta_1} = n_{3,\beta_3} = 0$ .

$\alpha := \beta_1\chi + \beta_3$  never changes sign for any  $(\gamma, \chi)$ , allowing us to exploit the discount rate.

To circumvent this difficulty, we augment the BVP with an auxiliary variable  $\tilde{\alpha}$  to serve as a proxy for  $\alpha$  in the  $r$  term in the centered system; by construction, it will share the sign of  $\alpha^m$  and, in any solution to the BVP, will coincide with  $\alpha$ . Specifically, observe that using the decomposition  $x = x^c + x^m$  for  $x \in \{\beta_1, \tilde{\beta}_2, \beta_3\}$  yields that the  $r$ -independent term inside the outer brace of (C.5) is of the form  $\frac{\gamma_t h_\alpha(\vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \chi_t)}{\sigma_X^2 \sigma_Y^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)^{n_{1,\alpha}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t)^{n_{2,\alpha}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}})^{n_{3,\alpha}}}$ , where  $h_\alpha$  is a polynomial and  $n_{j,\alpha} \in \mathbb{N}$ ,  $j = 1, 2, 3$ . We introduce the (backward) linear ODE

$$\dot{\tilde{\alpha}}_t = \tilde{\alpha}_t \left\{ -r \left( \frac{\alpha_t}{\alpha_t^m} - 1 \right) + \frac{\sigma_X^{-2} \sigma_Y^{-2} \gamma_t h_\alpha(\vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \chi_t)}{(u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)^{n_{1,\alpha}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t)^{n_{2,\alpha}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}})^{n_{3,\alpha}}} \right\} \quad (\text{C.6})$$

with initial condition  $\tilde{\alpha}_0 = \alpha_0^m$ . That is, the right-hand side of (C.6) is exactly as the one in (C.5) except for  $\tilde{\alpha}$  now multiplying the bracket. The exact same application of the comparison argument between  $\alpha$  and 0 shows that  $\tilde{\alpha}$  never vanishes over its interval of existence *for any pair*  $(\gamma, \chi)$  Lipschitz taking values in  $[0, \gamma^o] \times [0, 1]$ , and  $\tilde{\alpha}/\alpha^m > 0$ .

Our augmented BVP then consists of the ODEs of  $x^c = \beta_1^c, \tilde{\beta}_2^c, \beta_3^c$  in Lemma C.1 with a modified  $r$ -term of the form  $-r x_t^c \frac{\tilde{\alpha}_t}{\alpha_t^m}$ , i.e., with  $\tilde{\alpha}$  replacing  $\alpha$  in the numerator of the fraction accompanying  $r$ . It also includes: the ODEs of  $y = \tilde{v}_6, \tilde{v}_8$ ; the learning ODEs (14)-(15); and the ODE (C.6) of  $\tilde{\alpha}$ .<sup>49</sup> The resulting system of ODEs—denote it  $\dot{\mathbf{z}}_t = F(\mathbf{z}_t)$ , where  $\mathbf{z} := (\gamma, \chi, \vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$ —is such that each component of  $F(\mathbf{z})$  is a polynomial divided by a product of powers of  $1 - u_{a\hat{a}} \hat{u}_{a\hat{a}}$ ,  $1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t$ , and  $u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t$ . Since the latter are bounded away from zero,  $F$  is of class  $C^1$ . We verify at the end of the proof that any solution to this augmented BVP satisfies that  $\alpha := \beta_1\chi + \beta_3$  coincides with  $\tilde{\alpha}$  by construction.<sup>50</sup>

**Fixed point:** Use a fixed-point argument to show that there are horizon lengths of order  $1/\gamma^o$  such that the augmented BVP admits a solution. We will prove the following result:

**Theorem C.1.** *Under Assumptions 1 and 2, there is a strictly positive function  $T(\gamma^o) \in \Omega(1/\gamma^o)$  such that if  $T < T(\gamma^o)$ , there exists a solution to the BVP in  $\mathbf{z} = (\gamma, \chi, \vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$ .*

*Proof.* The proof consists of converting the BVP into a fixed point problem over pairs  $\lambda := (\gamma, \chi)$  in a suitable set. Specifically, for a given  $\lambda$  we can first solve the backward initial value problem (IVP) in the variables  $(\vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$  that takes  $\lambda$  as an input. Second, we can solve the forward IVP for the two learning coefficients that takes as an input the solution from the previous step. This procedure generates a continuous mapping from candidate  $\lambda$  paths in a suitable set to itself, to which we apply Schauder's fixed point theorem.

<sup>49</sup>For consistency, the  $\alpha_t$  in the  $r$ -term in (C.6) and in (14)-(15) must be written as  $(\beta_{1t}^c + \beta_{1t}^m)\chi_t + \beta_{3t}^c + \beta_{3t}^m$ .

<sup>50</sup>In a slight abuse of notation,  $\dot{\mathbf{z}}_t = F(\mathbf{z}_t)$  assumes that the ODEs have been stated in only one direction.



**Step 1:** Define the domain for our fixed point equation. Let  $\mathcal{C}$  denote the Banach space of continuous functions from  $[0, T]$  to  $\mathbb{R}$ , equipped with the sup norm  $\|\cdot\|_\infty$  defined by  $\|x\|_\infty := \sup\{|x_t| : t \in [0, T]\}$ . (To economize on notation, we use  $\|\cdot\|_\infty$  to denote the supremum norm for objects of all other dimensions too.) By the Arzela-Ascoli theorem (Ok, 2007, p. 198), the space of uniformly bounded functions with a common Lipschitz constant is a compact subspace of  $\mathcal{C}$ . In particular, for all  $\rho, K > 0$ , define  $\Gamma(\rho + K) \subset \mathcal{C}$  as the space of uniformly Lipschitz continuous functions  $\gamma : [0, T] \rightarrow [0, \gamma^\circ]$  with uniform Lipschitz constant  $(\gamma^\circ)^2(2[\rho + K])^2/\sigma_Y^2$  that satisfy  $\gamma_0 = \gamma^\circ$ . Likewise, let  $X(\rho + K) \subset \mathcal{C}$  denote the space of Lipschitz continuous functions  $\chi : [0, T] \rightarrow [0, 1]$  with uniform Lipschitz constant  $\gamma^\circ [(2[\rho + K])^2/\sigma_Y^2 + (|\hat{u}_{a\theta}| + |\hat{u}_{a\tilde{a}}|2[\rho + K])^2/\sigma_X^2]$  that satisfy  $\chi_0 = 0$ . Thus, the product  $\Lambda(\rho + K) := \Gamma(\rho + K) \times X(\rho + K)$  is a compact subspace of  $\mathcal{C}^2$ .

We note that these Lipschitz constants are motivated by a bounding exercise of the  $\gamma$  and  $\chi$  ODEs that uses  $|\beta_i^c| < K$  and  $|\beta_i^m| < \rho$ , implying that  $|\beta_i| < \rho + K$ ,  $i = 1, 3$ . Below, we shall construct horizons over which any solution satisfies this property.

**Step 2:** Given  $(\gamma, \chi) \in \Lambda(\rho + K)$ , define a backward initial value problem (IVP) for  $(\tilde{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$ , and establish sufficient conditions for this IVP to have a unique solution. For any function  $x$ , let us use  $\hat{x}_{(\cdot)} := x_{T-(\cdot)}$  to emphasize the time-reversed version of  $x$  whenever convenient (not to be confused with the hat notation used in the main body). Given any  $\lambda \in \Lambda(\rho + K)$ , where  $(\rho, K) \in \mathbb{R}_{++}^2$ , we can define a (backward) IVP consisting of the ODEs for  $(\tilde{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$  previously stated, but where  $\hat{\lambda}$  is used in place of the solutions of the learning ODEs. We write this problem as

$$\dot{\mathbf{b}}_t = \mathbf{f}^{\hat{\lambda}}(\mathbf{b}_t, t) \quad \text{s.t.} \quad \mathbf{b}_0 = (0, 0, 0, 0, 0, \alpha^m(\hat{\lambda}_0)), \quad (\text{IVP}^{\text{bwd}}(\hat{\lambda}))$$

where the use of boldface distinguishes solutions to this IVP from those of our original BVP. We write  $\mathbf{b}(\cdot; \lambda)$  for the solution as a functional of the input  $\lambda$ . The extra dependence on time in the right hand side of  $(\text{IVP}^{\text{bwd}}(\hat{\lambda}))$  is due to the role of  $\lambda$  in the system.

For all  $\lambda_t \in [0, \gamma^\circ] \times [0, 1]$ , let  $\mathbf{B}(\lambda_t) := (\beta_1^m(\lambda_t), \tilde{\beta}_2^m(\lambda_t), \beta_3^m(\lambda_t), 0, 0, \alpha^m(\lambda_t))$ . From here, we define  $\rho := \sup_{\lambda_t \in [0, \gamma^\circ] \times [0, 1]} \|\mathbf{B}_{-6}(\lambda_t)\|_\infty > 0$ , with  $\mathbf{B}_{-i}$  denoting as usual the vector  $\mathbf{B}$  excluding  $B_i$ .<sup>51</sup> For arbitrary  $K > 0$ , we now establish sufficient conditions for  $(\text{IVP}^{\text{bwd}}(\hat{\lambda}))$  to have a unique solution for each  $\lambda \in \Lambda(\rho + K)$ .

**Lemma C.2.** Fix  $\gamma^\circ, K > 0$ . There exists a threshold  $T(\gamma^\circ; K) > 0$  such that if  $T < T(\gamma^\circ; K)$ , then for all  $\lambda \in \Lambda(\rho + K)$ , a unique solution  $\mathbf{b}(\cdot; \lambda)$  to  $(\text{IVP}^{\text{bwd}}(\hat{\lambda}))$  exists over  $[0, T]$  and satisfies  $\|\mathbf{b}_i(\cdot; \lambda)\|_\infty < K$  for all  $i \in \{1, \dots, 5\}$ . Moreover,  $T(\gamma^\circ; K) \in \Omega(1/\gamma^\circ)$ .

<sup>51</sup>We exclude  $\tilde{\alpha}$  from the definition of  $\rho$  because it does not enter the ODEs for the learning coefficients explicitly, and hence it does not affect the definition of  $\Lambda(\rho + K)$ .

*Proof.* Fix any  $\lambda \in \Lambda(\rho + K)$ . Since  $\lambda$  is continuous in  $t$  and  $f^\lambda$  is of class  $C^1$  with respect to  $\mathbf{b}_t$ ,  $f^\lambda$  is locally Lipschitz continuous in  $\mathbf{b}_t$ , uniformly in  $t$ . By Peano's theorem, a local solution exists; and by the Picard-Lindelöf theorem, solutions are unique given existence. Given  $K > 0$ , we now construct  $T(\gamma^\circ; K)$  such that a solution exists over  $[0, T]$  and satisfies  $\|\mathbf{b}_i(\cdot; \lambda)\|_\infty < K$  for  $i \in \{1, \dots, 5\}$ .

We state two facts that hold over any interval of existence. First, using the ODEs adapted from Lemma C.1 (using  $\tilde{\alpha}$  instead of  $\alpha$  in the  $r$  terms), we have for  $i \in \{1, 2, 3\}$  and  $j \in \{4, 5\}$

$$\mathbf{b}_{it} = \int_0^t e^{-r \int_s^t \frac{\tilde{\alpha}_u}{\alpha_u^m} du} \hat{\gamma}_s h_i(\mathbf{b}_s, \hat{\chi}_s) ds \quad \text{and} \quad \mathbf{b}_{jt} = \int_0^t e^{-\int_s^t (r + \hat{\gamma}_u R_j(\mathbf{b}_u, \hat{\chi}_u)) du} \hat{\gamma}_s h_j(\mathbf{b}_s, \hat{\chi}_s) ds.$$

Here,  $h_i$  and  $h_j$  include the denominators that were factored out of  $h_x$  and  $h_y$  in Lemma C.1, and do not contain  $\tilde{\alpha}$ ;  $R_j$  is only a relabeling of  $R_y$  from the same lemma. Second, as long as the conjectured bounds  $|\mathbf{b}_{it}| < K$  for  $i \in \{1, 2, \dots, 5\}$  hold, a direct bounding exercise on  $h_i$  that uses  $\chi_t \in [0, 1]$  yields the existence of a scalar  $h_i(K)$  such that  $|\hat{\gamma}_s h_i(\mathbf{b}_s, \hat{\chi}_s)| \leq \gamma^\circ h_i(K)$ ,  $i \in \{1, 2, \dots, 5\}$ , where we have used that  $\gamma_t \in [0, \gamma^\circ]$  at all times.

Equipped with the equations above for  $\mathbf{b}_i$  and with  $h_i(K)$ ,  $i \in \{1, \dots, 5\}$ , notice that the bound  $|\mathbf{b}_{it}| < K$  clearly holds for small  $t$ . And as long as it holds,  $\tilde{\alpha}$  is finite because  $\mathbf{b}_{6t}$  has the form  $\alpha_0^m e^{\int_0^t G_s ds}$  with  $|G_s| < +\infty$  as the latter depends only on  $(\mathbf{b}_{-6}, \hat{\chi})$  at time  $s \in [0, t]$ . Moreover,  $\tilde{\alpha}/\alpha_t^m > 0$  (see ‘**Auxiliary variable**’). Thus, for  $i \in \{1, 2, 3\}$  and  $j \in \{4, 5\}$ ,

$$\begin{aligned} |\mathbf{b}_{it}| &\leq \int_0^t e^{-r \int_s^t \frac{\tilde{\alpha}_u}{\alpha_u^m} du} \gamma^\circ h_i(K) ds \leq \int_0^t \gamma^\circ h_i(K) ds = t \gamma^\circ h_i(K) \\ |\mathbf{b}_{jt}| &\leq \int_0^t e^{-\int_s^t (r + \hat{\gamma}_u R_j(\mathbf{b}_u, \hat{\chi}_u)) du} \gamma^\circ h_j(K) ds \leq \int_0^t \gamma^\circ h_j(K) ds = t \gamma^\circ h_j(K), \end{aligned}$$

where we have used that the exponential term is less than 1. Imposing that the right-hand sides above are themselves smaller than  $K$  leads us to  $T(\gamma^\circ; K) := \min_{i \in \{1, \dots, 5\}} \frac{K}{\gamma^\circ h_i(K)} > 0$  such that (IVP<sup>bwd</sup>( $\hat{\lambda}$ )) with  $T < T(\gamma^\circ; K)$  by construction admits a unique solution satisfying  $|\mathbf{b}_{-6}| < K$  for all  $\lambda \in \Lambda(\rho + K)$ . Moreover, since  $T(\gamma^\circ; K)$  is independent of  $r$ , the statement holds for all  $r \geq 0$ ; also  $T(\gamma^\circ; K) \in \Omega(1/\gamma^\circ)$ .<sup>52</sup>  $\square$

In what follows, assume  $T < T(\gamma^\circ; K)$ . Lemma C.2 implies that  $\lambda \in \Lambda(\rho + K) \mapsto \mathbf{b}(\cdot; \lambda)$  is a well-defined function linking  $\lambda$  paths to corresponding solutions to the backward IVP. We can then define the functional

$$q(\lambda) := (\hat{\mathbf{b}}_1(\cdot; \lambda), \hat{\mathbf{b}}_3(\cdot; \lambda)) + (B_1(\lambda_{(\cdot)}), B_3(\lambda_{(\cdot)}))$$

<sup>52</sup>It is clear from the argument that  $\tilde{\alpha}$  is also uniformly bounded for all  $\lambda \in \Lambda(\rho + K)$ . Also, the linearity of the  $\tilde{\alpha}$ -ODE (C.6) implies that the interval of existence is constrained only by the ODEs for  $\mathbf{b}_i$ ,  $i \in \{1, \dots, 5\}$ .

that for each  $\lambda$  delivers the induced “total” ‘ $\beta_1$ ’ and ‘ $\beta_3$ ’ forward-looking coefficients—the centered components delivered by the previous IVP plus the myopic counterparts—that we will use as an input in the learning ODEs below. (Clearly, each  $q(\lambda)$  function is a continuous function of time.) The continuity of this functional is key for our fixed-point argument.

**Step 3:** *The operator  $\lambda \mapsto q(\lambda)$  is continuous and  $\|q(\lambda)\|_\infty < \rho + K$  for all  $\lambda \in \Lambda(\rho + K)$ .* Let us show, more generally, that  $\lambda \mapsto \hat{\mathbf{b}}(\cdot; \lambda)$  is continuous; since  $\lambda \mapsto \mathbf{B}_i(\lambda_{(\cdot)})$  is clearly continuous due to  $\beta_i^m = \beta_i^m(\chi_{(\cdot)})$  being of class  $C^1$ ,  $i \in \{1, 3\}$ , the result will follow. To this end, we make use of the following lemma, proved in the Supplementary Appendix.

**Lemma C.3.** *Let  $X \subseteq \mathbb{R}^n$ ,  $Y \subseteq \mathbb{R}^m$  and  $U \subseteq \mathbb{R}^n$  be compact sets. Consider  $F : X \times Y \rightarrow U$  of class  $C^1$  and  $\omega : Y \rightarrow X$ . Suppose  $\mathcal{Y} \subset C([0, T]; Y)$  is a collection of functions such that for all  $y \in \mathcal{Y}$ , the initial value problem IVP( $y$ ) defined by  $\dot{x}_t = F(x_t, y_t)$  and  $x_0 = \omega(y_0)$  admits a solution defined over  $[0, T]$ . Then there exist constants  $k_1$  and  $k_2$  (depending on  $T$ ) such that for all  $y^1, y^2 \in \mathcal{Y}$ , the corresponding solutions  $x^i$  to IVP( $y^i$ ) satisfy*

$$\|x_t^1 - x_t^2\|_\infty \leq k_1 \|\omega(y_0^1) - \omega(y_0^2)\|_\infty + k_2 \sup_{s \in [0, T]} \|y_s^1 - y_s^2\|_\infty, \quad \text{for all } t \in [0, T].$$

Now consider any  $\lambda^1, \lambda^2 \in \Lambda(\rho + K)$ . We apply Lemma C.3 to:  $x = \mathbf{b}$ ;  $y^i = \hat{\lambda}^i$ ,  $i = 1, 2$ ;  $\omega(\cdot) = (0, 0, 0, 0, 0, \alpha^m(\cdot))$ ;  $F(x_t, y_t) := f^\lambda(\mathbf{b}_t, t)$ ; and  $X$  and  $Y$  the hypercubes defined by the uniform bounds on  $\mathbf{b}$  and  $\lambda$ , respectively. Using that  $\|x\|_\infty = \|\hat{x}\|_\infty$ , we obtain

$$\|\hat{\mathbf{b}}(\cdot; \lambda^1) - \hat{\mathbf{b}}(\cdot; \lambda^2)\|_\infty = \sup_{t \in [0, T]} \|\mathbf{b}_t(\lambda^1) - \mathbf{b}_t(\lambda^2)\|_\infty \leq k_1 |\alpha^m(\lambda_T^1) - \alpha^m(\lambda_T^2)| + k_2 \|\lambda^1 - \lambda^2\|_\infty,$$

for some constants  $k_1$  and  $k_2$ . Since  $\lambda_T \mapsto \alpha^m(\lambda_T)$  is continuous, it follows that  $\|\hat{\mathbf{b}}(\cdot; \lambda^1) - \hat{\mathbf{b}}(\cdot; \lambda^2)\|_\infty \rightarrow 0$  as  $\|\lambda^1 - \lambda^2\|_\infty \rightarrow 0$ , yielding the desired result.

Finally,  $\|q(\lambda)\|_\infty < \rho + K$  follows from  $\|\hat{\mathbf{b}}_i(\cdot; \lambda)\|_\infty < K$  and  $\|\mathbf{B}_i(\lambda_T)\|_\infty < \rho$ ,  $i = 1, 3$ .

**Step 4:** *Construct a continuous self-map on  $\Lambda(\rho + K)$  using the IVP for the learning ODEs.* Take  $\lambda \in \Lambda(\rho + K)$  and define the IVP for  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$

$$\dot{\boldsymbol{\lambda}}_t = f^{q(\lambda)}(\boldsymbol{\lambda}_t, t) \quad \text{s.t.} \quad \boldsymbol{\lambda}_0 = (\gamma^o, 0), \quad (\text{IVP}^{\text{fwd}}(q(\lambda)))$$

consisting of the two (forward) learning ODEs (14)-(15) that use as input  $q(\lambda) = (q_1(\lambda), q_2(\lambda))$  playing the role of  $(\beta_1, \beta_3)$ —here, the first (second) entry of the system corresponds to the  $\gamma$ -ODE ( $\chi$ -ODE), while the boldface convention aims at distinguishing between inputs  $\lambda$  via  $q$  and induced solutions  $\boldsymbol{\lambda}$  to this IVP. Importantly, because for all  $\lambda \in \Lambda(\rho + K)$  the function  $q(\lambda)$  is continuous in time, Lemma A.1 gives existence and uniqueness of a solution

to  $(\text{IVP}^{\text{fwd}}(q(\lambda)))$  defined over  $[0, T]$  that satisfies  $\boldsymbol{\lambda}_t \in (0, \gamma^o] \times [0, 1]$  for all such times.

Next, we argue that  $\boldsymbol{\lambda} \in \Lambda(\rho + K)$ . By construction,  $\boldsymbol{\lambda}_0 := (\boldsymbol{\lambda}_{1,0}, \boldsymbol{\lambda}_{2,0}) = (\gamma^o, 0)$ , and as noted above,  $\boldsymbol{\lambda}_t \in (0, \gamma^o] \times [0, 1]$  for all  $t \in [0, T]$ . Moreover, from the  $\gamma$ -ODE and  $\chi$ -ODE, we have that

$$\begin{aligned} |\dot{\boldsymbol{\lambda}}_{1t}| &= \left| -\frac{\boldsymbol{\lambda}_{1t}^2([q_2(\lambda)]_t + [q_1(\lambda)]_t \boldsymbol{\lambda}_{2t})^2}{\sigma_Y^2} \right| \leq (\gamma^o)^2 (2[\rho + K])^2 / \sigma_Y^2 \quad \text{and similarly} \\ |\dot{\boldsymbol{\lambda}}_{2t}| &\leq \gamma^o [(2[\rho + K])^2 / \sigma_Y^2 + (|\hat{u}_{a\theta}| + |\hat{u}_{a\hat{a}}|(2[\rho + K]))^2 / \sigma_X^2] \end{aligned}$$

for all  $t \in [0, T]$ . Since the Lipschitz bounds in the definition of  $\Lambda(\rho + K)$  are satisfied,  $\boldsymbol{\lambda} \in \Lambda(\rho + K)$ .

Finally, by Lemma C.3 applied to  $(\text{IVP}^{\text{fwd}}(q(\lambda)))$  by setting  $x = \boldsymbol{\lambda}$ ,  $y = q(\lambda)$ ,  $\omega(y_0) = (\gamma^o, 0)$ ,  $F(x_t, y_t) = f^{q(\lambda)}(\boldsymbol{\lambda}_t, t)$ ,  $X = [0, \gamma^o] \times [0, 1]$  and  $Y = [-\rho - K, \rho + K]^2$ , we conclude that  $q \mapsto \boldsymbol{\lambda}(q)$  is continuous. Since  $\lambda \mapsto q(\lambda)$  is continuous (Step 3), it follows that  $g(\lambda) := \boldsymbol{\lambda}(q(\lambda))$  is a continuous map from  $\Lambda(\rho + K)$  to itself.

**Step 5:** *Show that  $g$  has a fixed point.* By Step 1,  $\Lambda(\rho + K)$  is a nonempty, compact, convex Banach space, and by Step 4,  $g$  is a continuous map from  $\Lambda(\rho + K)$  to itself. By Schauder's Theorem (Zeidler, 1986, Corollary 2.13), there exists  $\lambda^* \in \Lambda(\rho + K)$  such that  $\lambda^* = g(\lambda^*)$ . It is clear, by construction, that  $(\lambda^*, \hat{\mathbf{b}}(\cdot; \lambda^*))$ , with  $\mathbf{b}(\cdot; \lambda^*)$  the solution to  $(\text{IVP}^{\text{bwd}}(\hat{\lambda}))$  under  $\lambda = \lambda^*$ , is a solution to the centered-augmented BVP under study. Finally, maximizing  $T(\gamma^o; K)$  over  $K > 0$  yields a  $T(\gamma^o) > 0$  that has the form  $C/\gamma^o$ .  $\square$

**Verification:** *Recover first a solution to the original BVP, and then to the full HJB equation.* We verify that the solution to the centered-augmented BVP induces a solution to the original BVP stated in the ‘**Core subsystem**’ section. To do this, we first note that any solution to the former BVP must satisfy the identity  $\tilde{\alpha} \equiv \alpha$ , where  $\alpha_t := \beta_{1t}\chi_t + \beta_{3t}$ ,  $\beta_{1t} := \beta_{1t}^c + \beta_{1t}^m$  and  $\beta_{3t} := \beta_{3t}^c + \beta_{3t}^m$ —consequently,  $(\gamma, \chi, \vec{\beta}^c, \tilde{v}_6, \tilde{v}_8)$  solves the centered system defined in the ‘**Centering**’ step. Indeed, using the definition of the myopic coefficients as well as the ODEs for  $\chi, \beta_{1t}^c$ , and  $\beta_{3t}^c$  yields that  $\alpha$  in backward form satisfies

$$\dot{\alpha}_t = -r\tilde{\alpha}_t(\alpha_t/\alpha_t^m - 1) + \alpha_t \frac{\gamma_t h_\alpha(\vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \chi_t)}{\sigma_X^2 \sigma_Y^2 (u_{a\theta} + u_{a\hat{a}} \hat{u}_{\hat{a}\theta} \chi_t)^{n_{1,\alpha}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}} \chi_t)^{n_{2,\alpha}} (1 - u_{a\hat{a}} \hat{u}_{a\hat{a}})^{n_{3,\alpha}}}.$$

Relative to (C.6), therefore, the  $r$ -term as well as the last fraction multiplying  $\alpha$  coincide. Call this last term  $H_t$ —a continuous function of time—and observe that  $p := \alpha - \tilde{\alpha}$  satisfies the ODE  $\dot{p}_t = p_t H_t$  with initial condition  $p_0 = 0$  due to  $\alpha_0 = \tilde{\alpha}_0 = \alpha_0^m$  (recall that time is being reversed). By uniqueness,  $p_t \equiv 0$  for all  $t \in [0, T]$ , confirming that  $\alpha \equiv \tilde{\alpha}$ .

Given this equivalence, it follows that  $(\gamma, \chi, \beta_1, \vec{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8) = (\lambda^*, \hat{\mathbf{b}}_{-6}(\cdot; \lambda^*) + \mathbf{B}_{-6}(\lambda^*))$

solves by construction the BVP stated in the ‘**Core subsystem**’ section. Moreover, as argued in Step 4 in the proof of Theorem C.1,  $\gamma > 0$  and  $\chi < 1$ , so we can invert the change of variables  $(\tilde{\beta}_2, \tilde{v}_6, \tilde{v}_8) = (\beta_2/(1 - \chi), v_6\gamma/(1 - \chi)^2, v_8\gamma/(1 - \chi))$  to obtain  $(\beta_2, v_6, v_8)$ . And since  $\alpha = \tilde{\alpha}$  never vanishes (see ‘**Auxiliary variable**’ section) and  $\gamma > 0$ , we can recover the rest of the coefficients as explained in the ‘**Core subsystem**’ section.

That our resulting  $V$  is effectively the sender’s value function follows directly from Theorem 3.5.2 in Pham (2009): specifically, (i) since  $V$  is quadratic, it trivially satisfies the growth condition in the theorem; (ii)  $V$  satisfies the HJB equation by construction; and (iii) the controlled dynamics  $(M, L)$  are linear in equilibrium with coefficients that are of class  $C^1$ , and hence they admit a unique strong solution.

To verify the admissibility of the (induced) on-path strategy profile, note that, along the path of play, the laws of motion for  $\hat{M}$  and  $L$ , (A.1) and (A.6), are given by  $dX_t = [\alpha_{0t} + \alpha_{2t}L_t + \alpha_{3t}\theta]dt + \sigma_Y dZ_t^Y$  and  $dY_t = [\delta_{0t} + \delta_{1t}\hat{M}_t + \delta_{2t}L_t]dt + \sigma_Y dZ_t^Y$ . Thus, these on-path dynamics can be written as  $dX_t = A_X(t, W)dt + \sigma_X dZ_t^X$  and  $dY_t = A_Y(t, W)dt + \sigma_Y dZ_t^Y$ , where  $W := (\theta, X, Y)$ —in other words,  $W$  satisfies a *vector* stochastic differential equation (SDE) (where the  $\theta$  SDE trivially has a drift and volatility identically equal to zero). The nature of this vector SDE is *functional*, in that  $A_X(t, W)$  and  $A_Y(t, W)$  carry *functions* in the second argument: these operators are integrals that depend on values of  $X$  and/or  $Y$  for all times prior to  $t$  (i.e.,  $A_X(t, \cdot)$  and  $A_Y(t, \cdot)$  are *non-anticipative* linear functionals). The initial condition of this SDE is  $(\theta_0, X_0, Y_0) = (\theta, 0, 0)$  (recall that  $\theta_t = \theta$  for all  $t \geq 0$  in this SDE). By the extension of Theorem 4.6 in Liptser and Shiryaev (1977) to multidimensional processes (see the note on p. 143), this SDE admits a unique strong solution; call it  $W^* := (X^*, Y^*)$ . Letting  $(\Omega, \mathcal{F}, \mathbb{P})$  the probability space in which  $(Z^X, Z^Y)$  is a standard two-dimensional Brownian motion, the law of this process,  $\mathbb{P} \circ W^{-1}(\cdot)$ , where  $W^{-1}$  denotes the inverse image of  $W$ , is a distribution on the measurable space  $C([0, T]) \times C([0, T])$  equipped with Borel  $\sigma$ -algebra. Thus, there is a unique solution in a probability-law (weak) sense that is consistent with the players equilibrium strategies.

Finally, by the same result in Liptser and Shiryaev (1977), the fact that  $(\theta, 0, 0)$  has a finite second moment ensures not only that  $(X^*, Y^*)$  inherits that property, but also the associated second moments can be uniformly bounded over  $[0, T]$ . Since the coefficients in the strategies are of class  $C^1$ , we conclude that the players’ on-path strategies are square integrable. Progressive measurability then follows from  $(X^*, Y^*)$  having continuous paths and being adapted (due to the non-anticipative nature of the SDE). Thus, the (on-path) strategy profile is admissible in the sense of Section 2.

We extend our existence result to the case of terminal payoffs in the following corollary, proved in the Supplementary Appendix. The bound on curvature ensures that we can select

an equilibrium of the static terminal game with sufficient regularity for our method.

**Corollary C.2.** *There exist  $C_\psi \in \{-\infty\} \cup (-\infty, 0)$  and  $C_T > 0$ , both independent of  $(r, \gamma^\circ)$ , such that if  $\psi_{\hat{a}\hat{a}} \in (C_\psi/\gamma^\circ, 0]$  and  $T < C_T/\gamma^\circ$ , a linear Markov equilibrium exists for all  $r \geq 0$ . Moreover,  $\alpha_3$  never vanishes.*

## Appendix D: Proofs for Section 6

Before proving Proposition 8, we describe how our fixed point method can accommodate a receiver with general discount rate  $\hat{r} \geq 0$ . We first need laws of motion for  $(\hat{M}, L)$  for the receiver's best response problem, which we obtain from (A.1) and (A.10), using  $\hat{a}'_t$  in  $dX_t$ .

**Lemma D.1.** *From the receiver's perspective, if he follows  $(\hat{a}'_t)_{t \in [0, T]}$ ,*

$$d\hat{M}_t = \frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y} dZ_t \quad (\text{D.1})$$

$$dL_t = \frac{\gamma_t^X \chi_t \delta_{1t}}{\sigma_X^2} [\hat{a}'_t - (\delta_{0t} + [\delta_{1t} + \delta_{2t}]L_t) + \sigma_X dZ_t^X], \quad (\text{D.2})$$

where  $Z_t := \frac{1}{\sigma_Y} [Y_t - \int_0^t (\alpha_{0s} + \alpha_{2s}L_s + \alpha_{3s}\hat{M}_s) ds]$  is a Brownian motion.

**Generalizing the approach** The receiver's HJB equation then reads,

$$\begin{aligned} \hat{r}\hat{V} = \sup_{\hat{a}'} \left\{ \hat{u}(\alpha_{0t} + \alpha_{2t}\ell + \alpha_{3t}\hat{m}, \hat{a}', \hat{m}) + \gamma_t \left( \frac{\hat{u}_{aa}}{2} \alpha_{3t}^2 + \frac{\hat{u}_{\theta\theta}}{2} + \hat{u}_{a\theta} \alpha_{3t} \right) + \hat{V}_t \right. \\ \left. + \mu_{\hat{M}} \hat{V}_{\hat{m}} + \mu_L(\hat{a}') \hat{V}_\ell + \frac{\sigma_{\hat{M}}^2}{2} \hat{V}_{\hat{m}\hat{m}} + \frac{\sigma_L^2}{2} \hat{V}_{\ell\ell} \right\}, \quad t < T, \end{aligned} \quad (\text{D.3})$$

where  $\mu_{\hat{M}} (= 0)$ ,  $\sigma_{\hat{M}}$ ,  $\mu_L(\hat{a}')$ , and  $\sigma_L$  are the drift and noise in (D.1) and drift and noise in (D.2), respectively. (There is no  $\hat{V}_{\hat{m}\ell}$  term since the future innovations in  $\hat{M}$  and  $L$  are uncorrelated from the receiver's perspective.) Moreover, with an LQG structure we guess a solution of the form

$$\hat{V}(\hat{m}, \ell, t) = \hat{v}_{0t} + \hat{v}_{1t}\hat{m} + \hat{v}_{2t}\ell + \hat{v}_{3t}\hat{m}^2 + \hat{v}_{4t}\ell^2 + \hat{v}_{5t}\hat{m}\ell, \quad (\text{D.4})$$

where  $v_j, j = 0, \dots, 5$ , are differentiable functions of time. Imposing that the conjectured linear strategy for the receiver satisfies his first order condition, we obtain three equations relating  $(\delta_0, \delta_1, \delta_2)$  to  $(\vec{\beta}, \hat{v}_2, \hat{v}_4, \hat{v}_5)$ . Analogously for the sender, we get four equations relating  $(\beta_0, \beta_1, \beta_2, \beta_3)$  to  $(\vec{\delta}, v_2, v_5, v_7, v_9)$ . The combined system of seven equations can be solved to

obtain expressions for  $(\vec{\beta}, \vec{\delta})$  in terms of  $(v_2, v_5, v_7, v_9, \hat{v}_2, \hat{v}_4, \hat{v}_5)$  and  $(\gamma, \chi)$ . (Our bounding exercise will ensure that these expressions have denominators bounded away from zero.)

Substituting these equations into both players' HJB equations and the laws of motion for  $(\gamma, \chi)$ , we obtain a system of ODEs in  $(\vec{v}, \vec{\hat{v}}) := (v_0, \dots, v_9, \hat{v}_0, \dots, \hat{v}_5)$  and  $(\gamma, \chi)$ . The induced BVP has  $v_{iT} = \hat{v}_{jT} = 0$  for  $i = 0, \dots, 9$  and  $j = 0, \dots, 5$  as terminal conditions (absent lump sum payoffs at  $T$ ), while the initial conditions  $(\gamma_0, \chi_0) = (\gamma^\circ, 0)$ .

To conclude, this BVP can be transformed into a fixed point problem in  $(\gamma, \chi)$  using identical steps. Indeed, after an appropriate change of variables (see `spm.nb`) the (new) ODEs for  $x \in \{v_0, \dots, v_9, \hat{v}_0, \dots, \hat{v}_5\}$  can be written in backward form as

$$\dot{x}_t = -x_t r + \gamma_t R(\vec{v}_t, \vec{\hat{v}}_t, \gamma_t, \chi_t), \quad x_0 = 0, \quad (\text{D.5})$$

where  $R$  is a ratio of polynomials with denominator bounded away from zero when  $(\vec{v}, \vec{\hat{v}})$  is bounded by  $K$  sufficiently small. Hence, solutions exist for sufficiently small  $K$ . As in the proof of Theorem C.1, such a bound  $K$  also yields Lipschitz bounds on  $\gamma$  and  $\chi$  to use in defining the domain for our fixed point argument. The form of (D.5) implies that, starting with  $(\gamma, \chi)$  in this domain, there exists  $T(\gamma^\circ) \in \Omega(1/\gamma^\circ)$  independent of  $r$  such that solutions  $(\vec{v}, \vec{\hat{v}})$  traced backward are bounded by  $K$ . In turn,  $(\gamma, \chi)$  traced forward lie in the original domain, creating a self-map.

*Proof of Proposition 8.* It suffices to show that, given any LME when the receiver is myopic, each player's strategy is still a best response to the other player's strategy when the receiver is forward-looking. Since the best-response property already holds for the sender, we focus on the receiver. The receiver's HJB equation under a prediction problem is

$$\begin{aligned} \hat{r}\hat{V} = \sup_{\hat{a}'} \left\{ -\frac{1}{2}(c_0 + c_1\hat{m} + c_2[\alpha_{0t} + \alpha_{3t}\hat{m} + \alpha_{2t}\ell] - \hat{a}')^2 - \frac{1}{2}(c_1 + c_2\alpha_{3t})^2\gamma_t + \hat{V}_t \right. \\ \left. + \mu_{\hat{M}}\hat{V}_{\hat{m}} + \mu_L(\hat{a}')\hat{V}_\ell + \frac{\sigma_{\hat{M}}^2}{2}\hat{V}_{\hat{m}\hat{m}} + \frac{\sigma_L^2}{2}\hat{V}_{\ell\ell} \right\}, \quad t < T, \end{aligned} \quad (\text{D.6})$$

where we have used  $\hat{\mathbb{E}}_t[a_t] = \alpha_{0t} + \alpha_{3t}\hat{M} + \alpha_{2t}L$  and  $\hat{\mathbb{E}}_t[(\theta - \hat{M}_t)^2] = \gamma_t$ , while for  $t = T$ ,

$$\hat{V}(\hat{m}, \ell, T) = \sup_{\hat{a}'} \frac{1}{2} \left\{ -(c_0 + c_1\hat{m} + c_2[\alpha_{0T} + \alpha_{3T}\hat{m} + \alpha_{2T}\ell] - \hat{a}')^2 - (c_1 + c_2\alpha_{3T})^2\gamma_T \right\}. \quad (\text{D.7})$$

Now let  $(\delta_{0t}^m, \delta_{1t}^m, \delta_{2t}^m) = (c_0 + c_2\alpha_{0t}, c_1 + c_2\alpha_{3t}, c_2\alpha_{2t})$  denote the myopic strategy coefficients and  $\hat{a}_t^m = \delta_{0t}^m + \delta_{1t}^m\hat{M}_t + \delta_{2t}^m L_t$  the myopic policy. It is easy to see that  $\hat{a}_t^m$  attains the supremum

in (D.7) and the first quadratic term vanishes, so (D.7) yields the terminal condition

$$\hat{V}(\hat{m}, \ell, T) = -\frac{1}{2}(c_1 + c_2\alpha_{3T})^2\gamma_T. \quad (\text{D.8})$$

Note that this terminal payoff is independent of  $(\hat{m}, \ell)$ . In the same spirit, we conjecture a solution to the HJB where the value function depends only on time. In this case, (D.6) reduces to  $\hat{r}\hat{V} = \sup_{\hat{a}'} \left\{ -\frac{1}{2}(c_0 + c_1\hat{m} + c_2[\alpha_{0t} + \alpha_{3t}\hat{m} + \alpha_{2t}\ell] - \hat{a}')^2 - \frac{1}{2}(c_1 + c_2\alpha_{3t})^2\gamma_t + \hat{V}_t \right\}$ . It is easy to see that for all  $t < T$ , the right hand side is maximized at the myopic policy  $\hat{a}_t$ , at which point the first quadratic loss term vanishes, so the HJB equation further reduces to

$$\hat{r}\hat{V} = -\frac{1}{2}(c_1 + c_2\alpha_{3t})^2\gamma_t + \hat{V}_t. \quad (\text{D.9})$$

Simple integration using (D.8) and (D.9) yields the solution

$$\hat{V}(t) = -\frac{1}{2} \int_t^T e^{-\hat{r}(s-t)} (c_1 + c_2\alpha_{3s})^2 \gamma_s dt - \frac{1}{2} e^{-\hat{r}(T-t)} (c_1 + c_2\alpha_{3T})^2 \gamma_T, \quad (\text{D.10})$$

which is indeed a function only of time. We conclude that the myopic policy is optimal.  $\square$

*Proof of Proposition 9.* Suppose  $\delta_1 = \hat{u}_{a\hat{a}}\alpha_3$ , in which case the  $\chi$ -ODE boils down to

$$\dot{\chi}_t = \gamma_t \alpha_{3t}^2 \left( \frac{1 - \chi_t}{\sigma_Y^2} - \frac{(\hat{u}_{a\hat{a}}\chi_t)^2}{\sigma_X^2} \right) =: -\gamma_t \alpha_{3t}^2 Q(\chi_t).$$

Conjecture  $f(\chi_t) = \gamma_t$  for all  $t \geq 0$ , where  $f : [0, \bar{\chi}] \rightarrow [0, \gamma^o]$ , some  $\bar{\chi} \in (0, 1]$ , is differentiable. In this case,  $f'(\chi_t)\dot{\chi}_t = \dot{\gamma}_t$ . When  $\alpha_{3t} \neq 0$ ,  $\frac{f'(\chi_t)}{f(\chi_t)} = \frac{\Sigma}{Q(\chi_t)}$ . Hence, we solve the ODE  $\frac{f'(\chi)}{f(\chi)} = \frac{\Sigma}{Q(\chi)}$  for  $\chi \in (0, \bar{\chi})$  where  $f(0) = \gamma^o$ .

To this end, let  $c_2 := \frac{\sqrt{1/\sigma_Y^4 + 4(\hat{u}_{a\hat{a}})^2/[\sigma_X\sigma_Y]^2} - 1/\sigma_Y^2}{2(\hat{u}_{a\hat{a}}/\sigma_X)^2}$  and  $-c_1 := \frac{-\sqrt{1/\sigma_Y^4 + 4(\hat{u}_{a\hat{a}})^2/[\sigma_X\sigma_Y]^2} - 1/\sigma_Y^2}{2(\hat{u}_{a\hat{a}}/\sigma_X)^2}$  be the roots of the convex quadratic  $Q$  above. Note that these are well-defined since  $\hat{u}_{a\hat{a}}$  and Assumption 1 part (ii) imply that  $\hat{u}_{a\hat{a}} \neq 0$ . Clearly,  $-c_1 < 0 < c_2$ . Also,  $c_2 \leq 1$  as  $Q(1) \geq 0$ . Thus,  $\frac{\Sigma}{Q(\chi)} = -\frac{\sigma_X^2 \Sigma}{(\hat{u}_{a\hat{a}})^2(c_1 + c_2)} \left[ \frac{1}{\chi + c_1} - \frac{1}{\chi - c_2} \right]$  is well defined (and negative) over  $[0, c_2)$  with  $1/(\chi + c_1) > 0$  and  $-1/(\chi - c_2) > 0$  over the same domain. We can then set  $\bar{\chi} = c_2$  and solve  $\int_0^\chi \frac{f'(s)}{f(s)} ds = -\frac{\sigma_X^2 \Sigma}{(\hat{u}_{a\hat{a}})^2(c_1 + c_2)} \log \left( \frac{\chi + c_1}{c_2 - \chi} \frac{c_2}{c_1} \right)$ , which yields the decreasing function  $f(\chi) = f(0) \left( \frac{c_1}{c_2} \right)^{1/d} \left( \frac{c_2 - \chi}{\chi + c_1} \right)^{1/d}$ , where  $1/d = \sigma_X^2 \Sigma / [(\hat{u}_{a\hat{a}})^2(c_1 + c_2)] > 0$ . Imposing  $f(0) = \gamma^o$  and inverting yields  $\chi(\gamma) = f^{-1}(\gamma)$  as in the lemma. Note that  $\chi(\gamma^o) = 0$  and  $\chi(0) = c_2$ .

We now verify that  $\chi(\gamma)$  satisfies the  $\chi$ -ODE (even when  $\alpha_3 = 0$ ). We have

$$\frac{d(\chi(\gamma_t))}{dt} = \frac{\alpha_{3t}^2 \gamma_t}{\sigma_Y^2 [c_1 + c_2(\gamma/\gamma^o)^d]^2} c_1 c_2 d [c_1 + c_2] \left( \frac{\gamma_t}{\gamma^o} \right)^d.$$



By construction, moreover,  $c_1 c_2 = c_1 - c_2 = \frac{\sigma_X^2}{\sigma_Y^2 (\hat{u}_{aa})^2}$ , which follows from equating the first- and zero-order coefficients in  $Q(\chi) = \hat{u}_{aa}^2 \chi^2 / \sigma_X^2 + \chi / \sigma_Y^2 - 1 / \sigma_Y^2 = \hat{u}_{aa}^2 (\chi - c_2)(\chi + c_1) / \sigma_X^2$ . Thus,  $dc_1 c_2 = c_1 + c_2$ . On the other hand,

$$\frac{[\hat{u}_{aa} \chi(\gamma)]^2}{\sigma_X^2} = \frac{\hat{u}_{aa}^2}{\sigma_X^2} \left[ c_1 c_2 \frac{1 - (\gamma/\gamma^o)^d}{c_1 + c_2 (\gamma/\gamma^o)^d} \right]^2 = \frac{c_1^2 (1 - c_2)}{\sigma_Y^2} \left[ \frac{1 - (\gamma/\gamma^o)^d}{c_1 + c_2 (\gamma/\gamma^o)^d} \right]^2$$

where we used that  $c_1^2 c_2^2 / \sigma_X^2 = c_1^2 (1 - c_2) / \sigma_Y^2$  follows from  $\hat{u}_{aa}^2 c_2^2 / \sigma_X^2 = (1 - c_2) / \sigma_Y^2$  by definition of  $c_2$ . Thus, the right-hand side of the  $\chi$ -ODE evaluated at our candidate  $\chi(\gamma)$  satisfies

$$\gamma_1 \alpha_3^2 \left( \frac{1 - \chi}{\sigma_Y^2} - \frac{(\hat{u}_{aa} \chi)^2}{\sigma_X^2} \right) \Big|_{\chi=\chi(\gamma)} = \frac{\alpha_3^2 \gamma_1}{\sigma_Y^2} \left( 1 - \chi - c_1^2 (1 - c_2) \left[ \frac{1 - (\gamma/\gamma^o)^d}{c_1 + c_2 (\gamma/\gamma^o)^d} \right]^2 \right).$$

Thus, using that  $c_1 c_2 d = c_1 + c_2$  in our expression for  $d(\chi(\gamma_t))/dt$ , it suffices to show that  $[c_1 + c_2]^2 \left( \frac{\gamma_t}{\gamma^o} \right)^d = (1 - \chi)[c_1 + c_2 (\gamma/\gamma^o)^d]^2 - c_1^2 (1 - c_2)[1 - (\gamma/\gamma^o)^d]^2$ . Using that  $\chi[c_1 + c_2 (\gamma/\gamma^o)^d] = 1 - (\gamma/\gamma^o)$ , it is easy to conclude that this equality reduces to three equations  $0 = c_1^2 - c_1^2 c_2 - c_1^2 + c_1^2 c_2$ ,  $(c_1 + c_2)^2 = 2c_1 c_2 - c_1 c_2 (c_2 - c_1) + 2c_1^2 (1 - c_2)$  and  $0 = c_2^2 + c_1 c_2^2 - c_1^2 (1 - c_2)$ , capturing the conditions on the constant,  $(\gamma/\gamma^o)^d$  and  $(\gamma/\gamma^o)^{2d}$ , respectively. The first condition is trivially satisfied, and the third is easy to verify; by canceling common terms, the second condition is also a rearrangement of this identity. Thus,  $\chi(\gamma)$  as postulated satisfies the  $\chi$ -ODE; by uniqueness,  $\chi = \chi(\gamma)$ .

We now prove the final statement of the lemma. When  $\gamma_t \in (0, \gamma^o]$ , we have  $\chi_t = \frac{c_1 c_2 (1 - [\gamma_t/\gamma^o]^d)}{c_1 + c_2 [\gamma_t/\gamma^o]^d} < \frac{c_1 c_2}{c_1} = c_2$ . Now  $c_2$  simplifies to  $\frac{\sqrt{\sigma_X^4 + 4\sigma_Y^2 \sigma_X^2 \hat{u}_{aa}^2} - \sigma_X^2}{2\hat{u}_{aa}^2 \sigma_Y^2} = \frac{4\sigma_Y^2 \hat{u}_{aa}^2}{2\hat{u}_{aa}^2 \sigma_Y^2 (\sqrt{1 + 4\sigma_Y^2 \hat{u}_{aa}^2 / \sigma_X^2} + 1)}$ , which by inspection is increasing in  $\sigma_X$  and has limits  $\lim_{\sigma_X \rightarrow 0} c_2 = 0$  and  $\lim_{\sigma_X \rightarrow +\infty} c_2 = 1$   $\square$

## References

- AMADOR, M. AND P.-O. WEILL (2012): “Learning from private and public observations of others’ actions,” *Journal of Economic Theory*, 147, 910–940.
- ATHEY, S., A. ATKESON, AND P. J. KEHOE (2005): “The optimal degree of discretion in monetary policy,” *Econometrica*, 73, 1431–1475.
- BACK, K., C. H. CAO, AND G. A. WILLARD (2000): “Imperfect competition among informed traders,” *The Journal of Finance*, 55, 2117–2155.
- BARRO, R. J. AND D. B. GORDON (1983): “Rules, discretion and reputation in a model of monetary policy,” *Journal of monetary economics*, 12, 101–121.

- BENSOUSSAN, A. (1992): “Stochastic control of partially observable systems,” (*No Title*).
- BONATTI, A. AND G. CISTERNAS (2020): “Consumer scores and price discrimination,” *The Review of Economic Studies*, 87, 750–791.
- BONATTI, A., G. CISTERNAS, AND J. TOIKKA (2017): “Dynamic oligopoly with incomplete information,” *The Review of Economic Studies*, 84, 503–546.
- BOUVARD, M. AND R. LÉVY (2019): “Horizontal Reputation and Strategic Audience Management,” *Journal of the European Economic Association*.
- CETEMEN, D. (2020): “Efficiency in Repeated Partnerships,” *Available at SSRN 3724923*.
- CETEMEN, D., G. CISTERNAS, A. KOLB, AND S. VISWANATHAN (2023): “Activist Trading Dynamics,” *FRB of New York Staff Report*.
- CETEMEN, D. AND C. MARGARIA (2020): “Signaling What You Don’t Want,” *Available at SSRN 3732674*.
- CISTERNAS, G. (2018): “Two-sided learning and the ratchet principle,” *The Review of Economic Studies*, 85, 307–351.
- CUKIERMAN, A. AND A. H. MELTZER (1986): “A theory of ambiguity, credibility, and inflation under discretion and asymmetric information,” *Econometrica: journal of the econometric society*, 1099–1128.
- DALEY, B. AND B. GREEN (2012): “Waiting for News in the Market for Lemons,” *Econometrica*, 80, 1433–1504.
- DUFFIE, D., N. GÂRLEANU, AND L. H. PEDERSEN (2005): “Over-the-counter markets,” *Econometrica*, 73, 1815–1847.
- EKMEKCI, M., L. GORNO, L. MAESTRI, J. SUN, AND D. WEI (2022): “Learning from manipulable signals,” *American Economic Review*, 112, 3995–4040.
- ELY, J. C. AND J. VÄLIMÄKI (2002): “A robust folk theorem for the prisoner’s dilemma,” *Journal of Economic Theory*, 102, 84–105.
- FAUST, J. AND L. E. SVENSSON (2001): “Transparency and credibility: Monetary policy with unobservable goals,” *International Economic Review*, 42, 369–397.
- FELTOVICH, N., R. HARBAUGH, AND T. TO (2002): “Too Cool for School? Signalling and Countersignalling,” *The RAND Journal of Economics*, 33, 630–649.

- FISHER, S. (2017): “Monetary Policy Expectations and Surprises,” *Board of Governors of the Federal Reserve System*.
- FOSTER, F. D. AND S. VISWANATHAN (1994): “Strategic trading with asymmetrically informed traders and long-lived information,” *Journal of Financial and Quantitative Analysis*, 29, 499–518.
- (1996): “Strategic trading when agents forecast the forecasts of others,” *The Journal of Finance*, 51, 1437–1478.
- HEINSALU, S. (2018): “Dynamic Noisy Signaling,” *American Economic Journal: Microeconomics*, 10, 225–249.
- HOLMSTRÖM, B. (1999): “Managerial incentive problems: A dynamic perspective,” *The review of Economic studies*, 66, 169–182.
- KANDORI, M. (2002): “Introduction to repeated games with private monitoring,” *Journal of economic theory*, 102, 1–15.
- KARATZAS, I. AND S. E. SHREVE (1991): *Brownian motion and stochastic calculus*, vol. 113, Springer Science & Business Media.
- KAYA, A. (2009): “Repeated signaling games,” *Games and Economic Behavior*, 66, 841–854.
- KOLB, A., M. PEASE, D. W. SACKS, AND J. QUICK (2021): “Blind Disclosure,” *Available at SSRN 3765529*.
- KOLB, A. M. (2015): “Optimal entry timing,” *Journal of Economic Theory*, 157, 973–1000.
- (2019): “Strategic real options,” *Journal of Economic Theory*, 183, 344–383.
- KREMER, I. AND A. SKRZYPACZ (2007): “Dynamic signaling and market breakdown,” *Journal of Economic Theory*, 133, 58–82.
- KYDLAND, F. E. AND E. C. PRESCOTT (1977): “Rules rather than discretion: The inconsistency of optimal plans,” *Journal of political economy*, 85, 473–491.
- KYLE, A. S. (1985): “Continuous auctions and insider trading,” *Econometrica*, 1315–1335.
- LIPTSER, R. S. AND A. SHIRYAEV (1977): *Statistics of Random Processes 1, 2*, Springer-Verlag, New York.

- MAILATH, G. J. AND S. MORRIS (2002): “Repeated games with almost-public monitoring,” *Journal of Economic Theory*, 102, 189–228.
- MAILATH, G. J. AND L. SAMUELSON (2006): *Repeated Games and Reputations*, Oxford University Press.
- MILGROM, P. AND J. ROBERTS (1982): “Predation, reputation, and entry deterrence,” *Journal of economic theory*, 27, 280–312.
- OK, E. A. (2007): *Real analysis with economic applications*, vol. 10, Princeton University Press.
- PHAM, H. (2009): *Continuous-time stochastic control and optimization with financial applications*, vol. 61, Springer Science & Business Media.
- PHELAN, C. AND A. SKRZYPACZ (2012): “Beliefs and private monitoring,” *Review of Economic Studies*, 79, 1637–1660.
- SUBRAHMANYAM, A. (1998): “Transaction taxes and financial market equilibrium,” *The Journal of Business*, 71, 81–118.
- SVENSSON, L. E. AND M. WOODFORD (2003): “Indicator variables for optimal policy,” *Journal of monetary economics*, 50, 691–720.
- (2004): “Indicator variables for optimal policy under asymmetric information,” *Journal of Economic Dynamics and Control*, 28, 661–690.
- TESCHL, G. (2012): *Ordinary differential equations and dynamical systems*, vol. 140, American Mathematical Society.
- TOWNSEND, R. M. (1983): “Forecasting the forecasts of others,” *Journal of Political Economy*, 91, 546–588.
- WOODFORD, M. (2002): “Imperfect common knowledge and the effects of monetary policy,” *In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps.*
- YANG, L. AND H. ZHU (2020): “Back-running: Seeking and hiding fundamental information in order flows,” *The Review of Financial Studies*, 33, 1484–1533.
- ZEIDLER, E. (1986): *Nonlinear functional analysis vol. 1: Fixed-point theorems*, Springer-Verlag.