Signaling with Private Monitoring*

Gonzalo Cisternas and Aaron Kolb

March 8, 2024

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 observable, to noisy environments in which signals of behavior are imperfect, yet still commonly observed. In this paper, we depart from these canonical sender-receiver games, involving public signals exclusively, by allowing the receiver to privately observe a noisy signal of the sender's actions. This departure, and the methods that we develop, offer a path for addressing a whole new set of questions in dynamic signaling.

*Cisternas: Federal Reserve Bank of New York, `gonzalo.cisternas@ny.frb.org`. Kolb: Indiana University Kelley School of Business, `kolba@indiana.edu`. We thank the Editor and three anonymous referees for many helpful suggestions, as well as 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 audiences for useful comments. The views expressed in this paper are ours and do not necessarily represent those of the Federal Reserve Bank of New York or the Federal Reserve System.

The real-world relevance of private signals of others' behavior has been recognized since the study of oligopolies by Stigler (1964). But signals of this nature may also play a key role in a variety of other contexts: 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); and in financial markets, certain traders can have a natural advantage in picking up signals of others' trades (Yang and Zhu, 2020). How do informed individuals—monetary authorities, reputation-concerned 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?[1]

This paper offers the first framework for examining signaling games featuring private signals of actions. In the class of games studied, a forward-looking sender (she) interacts with a receiver (he) over a finite horizon. Both players have 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, which allows us to isolate how the sender's behavior changes solely due to the uncertainty that a private signal creates—this 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.

To illustrate the core issue underlying our class of games, consider the following example. Suppose that a monetary authority (the sender) 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, or market (the receiver), are noisy. If those signals are *public*, there is common knowledge of what the private sector has seen. But this means that the monetary authority disregards her past actions when forecasting the market's 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 signals—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 what would happen if the authority ceases to observe all the signals of

[1]We discuss the differences from these papers in the literature review section.

inflation seen by the market: her history of play retains statistical value—because higher past actions become indicative of higher beliefs by the private sector, and vice versa—while still allowing for the use of additional data to refine the estimates constructed with past play.

The challenge, of course, is that beliefs cease to be common knowledge in such a context. Consider our example. As the private sector receives signals of the monetary authority's actions, it forms a belief about the economy. If the authority cares about these estimates—because they shape inflation expectations—she 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 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 now with players that have the power to individually affect the signals observed by others. The rest of our applications—career concerns and insider trading, within the umbrella of models of reputation and financial markets discussed earlier—also feature this form of higher-order uncertainty, which is endogenous and dynamic, as it emerges and changes as play unfolds.[2]

Relative to games with private monitoring, we are able to construct equilibria in which the players non-trivially condition their actions on private states in the form of beliefs. Specifically, we focus on 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 his 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 aggregates her past actions linearly, so it is *always* 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

[2] The combination of (i) lack of common knowledge and (ii) a non-trivial use of past play to form beliefs is what distinguishes private monitoring from other signal structures beyond the imperfect public case. For instance, with perfect monitoring of the sender's actions, the sender uses her past play to forecast what the receiver knows, but the receiver's belief is public. Instead, if the receiver sees exogenous private signals of a sender's type, common knowledge breaks but the sender does not need to condition on past play.

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 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 therefore key for making progress in this area of dynamic games.

Relative to traditional (i.e., public) noisy signaling games, we uncover a new channel for information transmission: there is separation through the second-order belief. Indeed, given any fixed history of the public signal, 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. In equilibrium, this separation is captured by the second-order belief's dependence on the type in the representation. We examine the impact of this new signaling channel, and more generally of the higher-order uncertainty at play, in applications.

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 is 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.

The rest of the applications explore our novel signaling channel in more depth. Succinctly, our reputation game shows that if a career-concerned agent wants to be perceived as neutral, she may prefer to not know exactly what her reputation is, despite this impeding her ability to take the best actions to manage her reputation if she is perceived as biased. The reason is that the new signaling channel can reduce the separation of types relative to the public case, and so less information about the bias 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 the 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 loop 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 of 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 model and applications, while section 3 the representation result and the sender's best-response problem. Section 4 analyzes our applications. Section 5 is devoted to the existence of LMEs. Section 6 discusses our assumptions and further applications of our methods. The Appendix and Supplementary Appendix contain all the proofs.

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 with multisided private information—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 muldimensional 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.[3]

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;

[3]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).

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

Regarding our applications, a distinctive feature of our monetary policy game is the presence of higher-order uncertainty linked to a monetary authority and a market who learn about each other's private information. Instead, Cukierman and Meltzer (1986) and Faust and Svensson (2001) study discretionary (i.e., sequentially rational) equilibria in which only a market is gradually learning about the authority's preferences; Athey et al. (2005) study full commitment by an authority where revelation mechanisms lead to changes in her private information being revealed immediately; and in the decision problem of Svensson and Woodford (2004) it is an authority that learns from a fully informed market. Other models featuring two-sided learning are Svensson and Woodford (2003) and Cisternas (2018), but the source of uncertainty there is common to everyone. Most of these papers are LQG and have public signals of actions.[4] 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.

[4]The preferences in our monetary policy game coincide with those in the "benchmark case" in Athey et al. (2005). We thank two referees for pointing to this area of applicability and several of these references.

2 Model

We develop a *linear-quadratic-Gaussian* (LQG) framework for analyzing two-player dynamic signaling games featuring (i) an ex ante informed player whose actions are privately monitored by a second player and (ii) a public signal channel from this second player—who endogenously develops private information in the form of a belief—to the former. Thus, in these signaling games higher-order beliefs are at play, and private information flows in both directions. The general model is introduced next; Section 2.2 presents three applications.

2.1 An LQG Class of Games

There are two players, which we label as sender and receiver. They take actions $a_t \in \mathbb{R}$ for the sender and $\hat{a}_t \in \mathbb{R}$ for the receiver, continuously over a time interval $[0, T]$, with $T < \infty$. The sender possesses payoff-relevant private information, which we denote by $\theta \in \mathbb{R}$ and assume to be normally distributed with mean $\mu \in \mathbb{R}$ and variance $\gamma^o > 0$. At the outset, the receiver only knows this distribution, and this is common knowledge.

The sender is forward-looking, with her total ex post payoff taking the form

$$\int_0^T e^{-rt} u(a_t, \hat{a}_t, \theta)\, dt + e^{-rT} \psi(\hat{a}_T), \tag{1}$$

where $r \geq 0$ is a discount rate, while $u : \mathbb{R}^3 \to \mathbb{R}$ and $\psi : \mathbb{R} \to \mathbb{R}$ are quadratic functions, the latter to accommodate the potential presence of lump-sum terminal payoffs. On the other hand, the receiver is assumed to be myopic, and thus concerned only about maximizing his (expected) flow utility in every instant: the receiver's ex post time-t payoff is denoted

$$\hat{u}(a_t, \hat{a}_t, \theta), \tag{2}$$

with $\hat{u} : \mathbb{R}^3 \to \mathbb{R}$ also quadratic. (Remark 1 below discusses the myopia assumption.)

Turning to the information structure, we assume that, as time progresses, the receiver has access to a *private* signal Y of the sender's actions that evolves as

$$dY_t = a_t dt + \sigma_Y dZ_t^Y. \tag{3}$$

Here, Z^Y is a one-dimensional Brownian motion, while $\sigma_Y > 0$ is a volatility parameter; the strict positivity of the latter scalar ensures that Y indeed constitutes private information to the receiver (otherwise, the sender knows what the receiver has seen; namely, her actions).

Finally, we assume that there is a *public* noisy signal X of the receiver's actions given by

$$dX_t = \hat{a}_t dt + \sigma_X dZ_t^X, \tag{4}$$

where $Z^X \perp Z^Y$ is also a (one-dimensional) Brownian motion. Unless otherwise stated, we assume $\sigma_X > 0$, which prevents the sender from perfectly inferring the receiver's belief in real time. Altogether, the quadratic preferences (1)–(2), linear dynamics (3)–(4), and Gaussian randomness (θ, Z^Y, Z^X) define a class of LQG games.

To make these games non-trivial, we need to impose some conditions on $(u, \hat{u}, \psi)$; let subscripts on these functions denote partial derivatives. We begin with two concavity requirements. First, we assume $u_{aa} = \hat{u}_{\hat{a}\hat{a}} = -1$. That is, the players' utilities are strictly concave with respect to their own actions, thus ensuring that best responses are well-defined despite the unbounded action space. That the value is -1 simply amounts to a normalization of the players' payoffs, which simplifies our calculations. 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 $\psi \equiv 0$.

The next conditions create sufficient interdependence in the players' utilities so that computing higher-order beliefs is needed to determine best responses.

Assumption 1. *(i) $u_{a\theta} \neq 0$ (type sensitivity); (ii) $|\hat{u}_{\hat{a}\theta}| + |\hat{u}_{a\hat{a}}| \neq 0$ (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, while (ii) is needed for the receiver's action to be sensitive to his private belief. The latter happens when the receiver cares about the type either directly ($\hat{u}_{\hat{a}\theta} \neq 0$) or indirectly through the sender's action ($\hat{u}_{a\hat{a}} \neq 0$). 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 strategic interaction term ($u_{a\hat{a}} \neq 0$). But it also happens when u or ψ is nonlinear in the receiver's action ($|u_{\hat{a}\hat{a}}| + |\psi_{\hat{a}\hat{a}}| \neq 0$).

Condition (iii), or second-order belief sensitivity, is fundamental for being able to depart from the traditional models studied so far. Indeed, 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).[5] Second, related to this point, in Section 5 we complement (iii) with mild technical conditions that ensure that there always is information transmission in equilibrium (second-order belief effect included).

Remark 1. *The receiver's myopia is conceptually useful for isolating the impact of a private signal on the sender's behavior from a confounding strategic response by the receiver. While*

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

it may seem restrictive, 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*, as in our monetary policy and reputation applications, 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 adjustments.*

In what follows, we let $\mathbb{E}_t[\cdot]$ and $\hat{\mathbb{E}}_t[\cdot]$ denote the players' conditional expectations, and we use $\mathbb{E}[\cdot|\mathcal{F}_t^X]$ for conditional estimates that use X but not (θ, Y), $t \in [0, T]$. We retain $\mathbb{E}[\cdot]$ (without any subindex or conditioning) for computing averages from an ex ante perspective.

2.2 Applications

We examine three applications: monetary policy, reputation, and trading. The first two games are under the umbrella of Section 2, while the last is based on an extension. We rely on the monetary policy example to explain the general model and intuitions when needed.

Monetary policy The following monetary policy game is in the spirit of Kydland and Prescott (1977) and Barro and Gordon (1983). There is a monetary authority (the sender) and a private sector, or market (the receiver). Terminal payoffs for the authority are absent ($\Psi \equiv 0$) while flow payoffs read as follows (the multiplicative factors deliver $u_{aa} = \hat{u}_{\hat{a}\hat{a}} = -1$):

$$\underbrace{u(a, \hat{a}, \theta) = \frac{1}{4}\left[-(k + \hat{a} - a)^2 - (a - \theta)^2\right]}_{\text{monetary authority}} \quad \text{and} \quad \underbrace{\hat{u}(a, \hat{a}, \theta) = -\frac{1}{2}(a_t - \hat{a}_t)^2}_{\text{private sector}}. \tag{5}$$

We interpret the authority's type θ as private information regarding a newly realized shock to the economy, while a_t corresponds to the authority's choice of inflation at t. In turn, $\hat{u}$ reveals that $\hat{a}_t$ is always $\hat{\mathbb{E}}_t[a_t]$, i.e., the private sector's estimate of current inflation.

A myopic private sector is reduced form for agents who cannot affect aggregate variables; estimates of inflation matter for the private sector because they are used to set nominal wages. On the other hand, u is reduced form for an authority who faces an output-inflation trade off: while it is costly to set inflation away from θ—the second term—doing so can bring output closer to a target (normalized to zero). This is the first term in u, displaying a Phillips curve: the unemployment rate, $k + \hat{a}_t - a_t$, deviates from its natural level k in the opposite direction of unanticipated inflation, $a - \hat{a}$. Finally, Y as in (3) reflects a private sector that has access to a signal of inflation not available to the authority. For instance, this can capture agents in the economy who have imperfect knowledge of how monetary policy transmits to the economy, or who misinterpret central banks' actions; such knowledge and interpretations are inherently private.[6] In turn, our public signal X as in (4) resembles survey data that is periodically gathered by central banks to learn about market expectations.[7]

Assumption 1 clearly holds. In particular, second-order belief sensitivity (Assumption 1(iii)), is satisfied because u exhibits a bliss point with respect to unemployment, leading to $|u_{a\hat{a}}| + |u_{\hat{a}\hat{a}}| \neq 0$. This means that the authority will need to forecast $\hat{a}$—and hence the market's estimate of θ, depending on Y—to set inflation. While the public signal X will not eliminate the higher-order inferences at play, it will make the problem manageable.

A credibility problem is the central tension in monetary policy games like this one. To illustrate, consider the linear equilibrium of a one-shot (e.g., simultaneous move) interaction:

$$a = \frac{\theta + \mu}{2} + k \;\text{ and }\; \hat{a} = \mu + k. \tag{6}$$

Averaging across all types, $\mathbb{E}[k + \hat{a} - a] \equiv k$—i.e., unemployment 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 average rate of unemployment with inflation of just μ. That is, the authority may want to lower inflation by k—the static *inflationary bias*—but creating surprise inflation is ex post optimal.[8] How does the monetary authority manage the market's expectations as both parties embark on learning from each other? How do the incentives to create surprise inflation evolve, and how do they relate to this credibility problem? In terms of inflation, does it matter that the authority does not know what the market has seen?

Reputation and Trading Briefly, we study two additional examples in Sections 4.2 and 4.3. The first is a reputation game: θ is the intensity of a bias on a relevant issue, while $\mu = 0$ (a normalization) is interpreted as the *unbiased type*. The players' payoffs are

$$\text{sender: } \frac{1}{2}\left[-\int_0^T e^{-rt}(a_t - \theta)^2 dt - e^{-rT}\psi \hat{a}_T^2\right]; \quad \text{receiver: } -\frac{1}{2}(\hat{a}_t - \theta)^2,$$

where $\psi > 0$. Note that the sender has a long-term concern to appear as unbiased: this is because the receiver tries to predict the bias at all times, and the sender's terminal payoff is maximized when $\hat{a}_T = \mu = 0$. Assumption 1(iii) holds because $\psi > 0$. Second, to leverage

[6] Woodford (2002) stresses the difference between public information and information that individuals are actually aware of. He models the latter with private signals of aggregate variables like (3). On the other hand, central banks' 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.

[7] 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. Also, θ could be interpreted as a preference parameter of the authority, and k as the authority's output target (in this case, with the natural rate of unemployment normalized to zero).

[8] Hence, (6) is "discretionary." See Supplementary Appendix Section S.1.6 for the commitment solution.

our methods beyond the baseline model, we consider the following trading game:

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

where $\mathbb{E}[\theta|\mathcal{F}_t^X]$ is the price of an asset of true value θ, based on an "order flow" $dX_t = (a_t + \hat{a}_t)dt + \sigma_X dZ_t^X$. That is, the sender now affects the public signal, and there is a "third action": prices set by market makers using the public information only. The interaction term $\mathbb{E}[\theta|\mathcal{F}_t^X]a_t$ in the sender's payoff plays the role of $u_{a\hat{a}} \neq 0$ in Assumption 1(iii). The sender cares that her actions create information to the receiver not available to price setters.

2.3 Strategies and Equilibrium

Towards a solution concept for our general model, note that the noise terms (Z^X, Z^Y) in (3)-(4) 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 set of equilibrium outcomes.[9] 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 $(X_s, Y_s)_{s<t}$, $t \in [0, T]$.[10] 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 (3)–(4) admits a unique solution (so the realizations of (X, Y) are well-defined). We only consider such (pure) strategy profiles.[11]

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

[9]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.

[10]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.

[11]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 (3)–(4) 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.

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 (θ, 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 at the core of our games in that allowing for such an endogenous history-dependence is essential for being able to find equilibria. As we show, this complex 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, the means of such posterior beliefs—henceforth, *beliefs*—will act as state variables that are used linearly. In addition to the sender's type θ, therefore, we will use the following states:

$$\hat{M}_t := \hat{\mathbb{E}}_t[\theta], \quad M_t := \mathbb{E}_t[\hat{M}_t], \quad \text{and} \quad L_t := \mathbb{E}[\theta|\mathcal{F}_t^X]. \tag{7}$$

Here, $\hat{M}_t$ is the receiver's belief about θ, 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 θ for someone who only observes the public signal X.

To illustrate, consider 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 θ by the authority, and hence of $\hat{M}$ by the private sector, are obvious. As the authority sets inflation then, she ceases to know how $\hat{M}$ has departed from its prior mean μ. But estimating this belief is key to preventing output from becoming too destabilized from its target level—either because the private sector has misinterpreted the authority's actions, or because the inflation surprises have been inadequate. 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 if these expectations are not met (recall that unemployment reads $k + \hat{a} - a$). This perception, encoded in a high value of M, will lead to inflationary stimuli partially aimed at stabilizing 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*:

$$a_t = \beta_{0t} + \beta_{1t} M_t + \beta_{2t} L_t + \beta_{3t}\theta \tag{8}$$

$$\hat{a}_t = \delta_{0t} + \delta_{1t}\hat{M}_t + \delta_{2t} L_t. \tag{9}$$

The coefficients β_{it}, $i = 0, 1, 2, 3$, and δ_{jt}, $j = 0, 1, 2$, will be differentiable functions of time, a dependence that must be allowed in order to capture end-game and learning effects. For example, returning to the credibility problem, the authority may want to take actions to induce low inflation expectations by the market, and then surprise the latter 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 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 $M \equiv \hat{M}$, so 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 θ emerges in equilibrium. That is, there is *signaling through the second-order belief.*

To analyze these games then, we need two different expressions for M. First, we need to know how M looks along the path of play of (8)–(9); in particular, its exact dependence on θ. This is because the receiver must anticipate the total informativeness of the sender's actions in equilibrium—i.e., M included—to be able to correctly interpret Y and form his belief. Lemma 1 in 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, we need a forward-looking expression for M under deviations from (8): a general law of motion for M, capturing how the sender perceives $\hat{M}$ will respond to arbitrary continuation strategies. We present this law of motion in Lemma 2 in Section 3.3. Importantly, this dynamic will depend on the representation derived in Lemma 1, because the sensitivity of M to the sender's actions—inherited from $\hat{M}$—will depend on the sender's total signaling as perceived by the receiver, which includes the dependence of M on θ that arises in equilibrium. In summary, these two lemmata are crucial for setting up the sender's best-response problem at the end of Section 3.3, which we then use to determine the equilibrium coefficients in (8)–(9).

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 $u_{a\hat{a}} = u_{\hat{a}\hat{a}} = \psi_{\hat{a}\hat{a}} = 0$. In any equilibrium of the form (8)–(9), $\beta_1 \equiv \beta_2 \equiv \delta_2 \equiv 0$: at all histories, the sender's action is affine in θ and the receiver's affine in $\hat{M}$. The strategies are also equilibria if Y is public (hence, they are independent of σ_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 σ_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 (8)–(9). Heuristically, given the LQG structure, it is natural to conjecture the representation

$$M_t = \chi_t \theta + (1 - \chi_t) L_t, \tag{10}$$

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 θ under (8). Further, with Gaussian learning, one expects this adjustment to be linear and deterministic, encoded in the weight χ.

The coefficient χ is a measure of the sender's learning about the receiver's belief (relative to someone who only observes the public signal). Because χ is linked to past play, it is naturally connected to the sender's past signaling. To quantify total information transmission, we insert (10) into (8), which yields the sender's actions along the path of (8)–(9),

$$a_t = \alpha_{0t} + \alpha_{2t} L_t + \alpha_{3t} \theta, \tag{11}$$

$$\text{where} \quad \alpha_{0t} := \beta_{0t}, \quad \alpha_{2t} := \beta_{2t} + \beta_{1t}(1 - \chi_t), \quad \text{and} \quad \alpha_{3t} := \beta_{3t} + \beta_{1t}\chi_t. \tag{12}$$

The total weight on the type, α_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 θ, χ must increase in weight. But how exactly does χ depend on past values of α_3? And is α_3 the only input to χ?

Our first contribution is to characterize the dependence of χ on the coefficients in the strategies (8)–(9) via a *system of ODEs*. To this end, we introduce the posterior variance

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

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 (8)–(10) 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 (8)–(9).[12]

Lemma 1. *Suppose that (X, Y) is driven by (8)–(9) and that the receiver conjectures (10), where $(L_t)_{t\in[0,T]}$ throughout is a process that depends only on the public information. Then M satisfies (10) if and only if: $L_t \equiv \mathbb{E}[\theta|\mathcal{F}_t^X]$ under (8)–(9); $\chi_t = \mathbb{E}_t[(\hat{M}_t - M_t)^2]/\gamma_t$; and*

$$\dot{\gamma}_t = -\frac{\gamma_t^2(\beta_{3t}+\beta_{1t}\chi_t)^2}{\sigma_Y^2}, \qquad \gamma_0 = \gamma^o, \tag{13}$$

$$\dot{\chi}_t = \frac{\gamma_t(\beta_{3t}+\beta_{1t}\chi_t)^2(1-\chi_t)}{\sigma_Y^2} - \frac{\gamma_t\chi_t^2\delta_{1t}^2}{\sigma_X^2}, \qquad \chi_0 = 0. \tag{14}$$

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

By the lemma, the weight χ 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 γ are associated with higher values of χ. Mathematically, the system of ODEs (13)–(14) is fully coupled (χ affects $\dot{\gamma}$ while γ affects $\dot{\chi}$), and higher values of the signaling coefficient $\alpha_3 = \beta_3 + \beta_1\chi$ prompt a faster decay of γ and a faster growth of χ.[13]

The system is a local representation of how the *learning coefficients* γ and χ are affected by past values of the coefficients guiding information transmission—in the particular case of the sender, the weights β_3 and β_1 on θ and M, respectively, via α_3. At $t = 0$, for instance, there is no higher-order uncertainty: $M = \hat{M} = L = \mu$, and so $\chi_0 = 0$. But if $\beta_{3,0} = \alpha_{3,0} \neq 0$

[12] Briefly, the approach is as follows. Given (11) (with L general), the problem of learning θ 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 (9) 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 (10) when (8) is followed allows us to pin down (χ, L).

[13] For instance, if $\sigma_X = +\infty$ (the public signal is uninformative), the solution to (13)–(14) satisfies $\chi = 1 - \frac{\gamma_t}{\gamma^o}$, so γ and χ are inversely related—see Lemma B.1 in the appendix, and Section 6 for a generalization.

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 (14) shows that δ_1, the receiver's signaling coefficient, is another input to χ. Indeed, higher values of δ_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 (14), a higher signal-to-noise ratio δ_{1t}^2/σ_X^2 puts more downward pressure on χ.[14]

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 (10) 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 (10), given the current value of L. As the receiver uses L, via (10), 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.

3.3 The Sender's Best-Response Problem

Assessing the optimality of (8) 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 leads her to behave differently, even for the same L, so the sender keeps track of M and L separately—below are their dynamics under arbitrary sender strategies.

[14] A law of motion for $L_t \equiv \mathbb{E}[\theta|\mathcal{F}_t]$ is presented in the next section. See also (A.10) in the Appendix.

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

$$dM_t = \frac{\gamma_t\alpha_{3t}}{\sigma_Y^2}(a'_t - [\alpha_{0t} + \alpha_{2t}L_t + \alpha_{3t}M_t])dt + \frac{\gamma_t\chi_t\delta_{1t}}{\sigma_X}dZ_t \tag{15}$$

$$dL_t = \frac{\gamma_t^X\chi_t\delta_{1t}}{\sigma_X^2}[\delta_{1t}(M_t - L_t)dt + \sigma_X dZ_t], \tag{16}$$

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 (15)–(16): since the sender's actions only enter M, the task of finding an equilibrium is inevitably linked to imposing optimality with respect to this state. This justifies our use of an extended strategy such as (8) involving both M and L as separate states.

The law of motion of M, (15), 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 (16), 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 (15)–(16) depend on the learning coefficients (γ, χ) of Lemma 1: for instance, M responds to the sender's action in proportion to $\gamma_t\alpha_{3t} = \gamma_t[\beta_{3t} + \beta_{1t}\chi_t]$. This dependence originates from the receiver's learning who uses the representation to form his mean-variance pair $(\hat{M}, \gamma)$: more uncertainty (higher γ) or stronger signaling (larger α_3) imply a more responsive $\hat{M}$, a property that M inherits. Two observations are in order. First, the appearance of χ reveals that an equilibrium representation for M is necessary to set up a best-response problem. Second, while the importance of posterior means is obvious, posterior variances are critical too. Indeed, much of the complexity behind the fixed point at play in these games (which we study in Section 5) is due to this variance channel: while (γ, χ) depends on the signaling coefficients in (8)–(9), this pair also shapes the responsiveness of (M, L) to the sender's actions, and so (γ, χ) affects the choice of strategy coefficients too.[15]

Remark 2 (M and L at all histories)**.** *Inserting the definition of* Z_t *in the lemma into* (15) *yields a dynamic for* M *that is linear in the same variable; its solution* M_t *is linear*

[15] Similarly, the public state L responds to news X in proportion to the receiver's signaling coefficient δ_1 and the variance of θ given the public information, γ^X, via $\text{Cov}(\theta, dX_t|\mathcal{F}_t^X) = \delta_{1t}\chi_t\gamma_t^X dt$. By Lemma 1, $\chi < 1$, so L is well-defined. Also, note that M adjusts upward when $a'_t > \mathbb{E}_t[\hat{\mathbb{E}}_t[a_t]] = \mathbb{E}_t[\alpha_{0t} + \alpha_{2t}L_t + \alpha_{3t}\hat{M}_t]$.

in $(a_s, L_s, X_s)_{s<t}$*. But inserting the same expression for* Z_t *in* (16) *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.*

Equipped with these two lemmata, we can set up the sender's best-response problem. Recall that the receiver's action $\hat{a} = \delta_0 + \delta_1 \hat{M} + \delta_2 L$ enters the pair (u, ψ), where the state $\hat{M}$ is hidden from the sender. In this LQG world, however, we can simply 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 (last part of Lemma 2), i.e., changes in the sender's actions only shift the receiver's belief.[16] Up to an additive constant, the sender's problem is to maximize

$$\mathbb{E}_0\left[\int_0^T e^{-rt} u(a_t, \delta_{0t} + \delta_{1t} M_t + \delta_{2t} L_t, \theta) dt + e^{-rT} \psi(\delta_{0T} + \delta_{1T} M_T + \delta_{2T} L_T)\right] \tag{17}$$

subject to the dynamics (15)–(16) of (M, L) from Lemma 2, and where (γ, χ) follow the ODEs (13)–(14) from Lemma 1.[17] Notably, the history-dependence challenge described in Section 2—that performing equilibrium analysis requires allowing for deviations, but those deviations endogenously affecting future play in a non-trivial manner through the inferences a player makes—has a parsimonious resolution here: one just needs to look at the sender's behavior at arbitrary values of (M, L).[18]

The coefficients $(\delta_0, \delta_1, \delta_2)$ in the receiver's strategy (9) 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

$$\hat{a}_t := \delta_{0t} + \delta_{1t}\hat{M}_t + \delta_{2t} L_t = \arg\max_{\hat{a}' \in \mathbb{R}} \hat{\mathbb{E}}_t[\hat{u}(\alpha_{0t} + \alpha_{2t} L_t + \alpha_{3t}\theta, \hat{a}', \theta)] \tag{18}$$

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. (Unlike with the sender, all the payoff-relevant histories for the receiver are reachable on path, so the sequential rationality requirement is covered by the Nash equilibrium concept, as in Definition 1.) We discuss the LMEs that arise in our applications next; the question of existence of LMEs is relegated to Section 5.

[16] 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 ψ at $t = T$.

[17] 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.

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

4 Applications

In this section, we examine the three applications introduced in Section 2.2.[19] A central issue is how key equilibrium outcomes change due to the presence of higher-order uncertainty. 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 α_3 simplifies to β_3 in that case: 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.[20]

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 output cannot be affected from an ex ante perspective, i.e., averaging across all possible monetary authority types: $\mathbb{E}[k+a_t-\hat{a}_t]\equiv k$ holds 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, for each type, 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*, ex ante equilibrium inflation in excess of μ, the average inflation for an authority who is only concerned about inflation targeting. Our focus will be on this measure of wasteful inflation, which we depict for different values of σ_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 σ_X can *cross*: high values of σ_X are associated with large (small) inflationary biases earlier (later) in the game.

[19] For general existence conditions for the first two, see Section 5. For the third game, see Remark 3.

[20] 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).

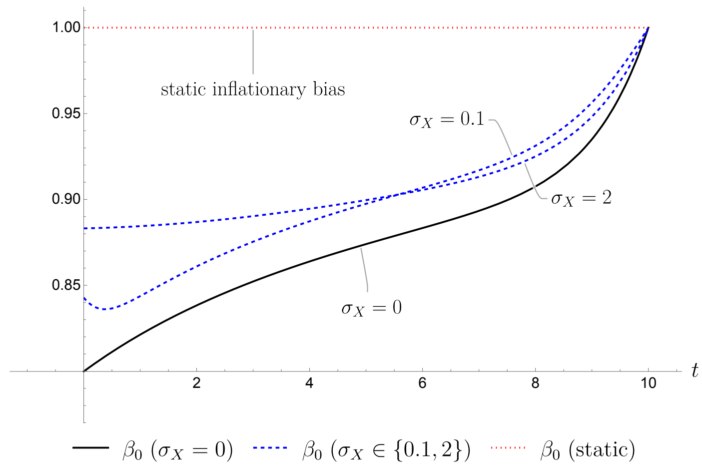


Figure 1: Inflationary bias for various values of σ_X. Parameters: $(k, \gamma^o, r, \sigma_Y) = (1, 1, 1, 1.5)$.

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,*

$$a_t = \beta_{0t} + \alpha_{3t}\theta + (1 - \alpha_{3t})L_t \quad \text{and} \quad \hat{a}_t = \beta_{0t} + \alpha_{3t}\hat{M}_t + (1 - \alpha_{3t})L_t. \tag{19}$$

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 $\alpha_{3t} \in (1/2, 1)$.[21]

From (19), 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. It is useful to write the authority's payoff (up to a constant) as:

$$\underbrace{2k(a_t - \hat{a}_t)}_{\text{stimulus}} \quad \underbrace{-(a_t - \hat{a}_t)^2}_{\text{output stabilization}} \quad \underbrace{-(a_t - \theta)^2}_{\text{inflation targeting}} \;. \tag{20}$$

The coefficient β_0 is linked to the stimulus part in (20). 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

[21] $T^\dagger$ depends on parameters; see discussion after Theorem 1. Numerically, the findings hold beyond $T^\dagger$.

inflation, at a lower cost, at a future date—i.e., β_0 is predominantly increasing. From an ex ante standpoint, the authority exhibits an apparent greater commitment to low inflation.[22]

The authority correctly expects lower future inflation costs in this LME because, ex post, different types do affect outcome variables due to their superior information. This brings us to the rest of the coefficients, $(\beta_1, \beta_2, \beta_3)$, and how they connect with the remaining terms in (20). 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 $\beta_2 > 0$.[23]

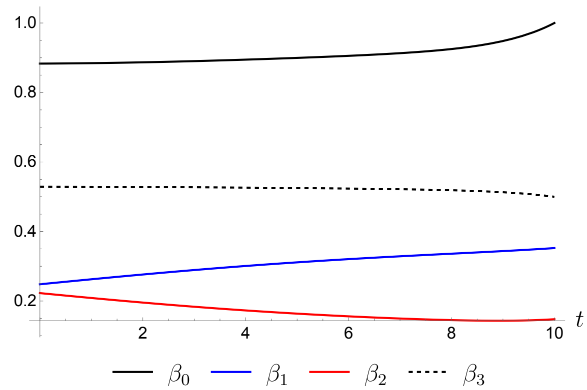


Figure 2: The sender's strategy coefficients, for $(k, r, \gamma^o, \sigma_X, \sigma_Y) = (1, 1, 1, 2, 1.5)$.

Altogether, high market beliefs about θ bring high inflation expectations, which in equilibrium are self-fulfilling. Inflation then becomes costly to the authority in this case, because it must be set far from θ. A forward-looking authority would like to *guide* the market in the direction of θ and, as this happens, surprise it with higher inflation tomorrow (via β_0), when it is cheaper to do so. Note that this guiding motive is captured by the last two terms in (20), as setting $a = \hat{a} = \theta$ in them eliminates any potential losses for the monetary authority.

In equilibrium, this guiding motive has a clean representation: the weight attached to θ in the extended strategy is larger than in the static case, $\beta_3 > 1/2$, capturing that the authority tries to convey her type more aggressively through this channel in order to steer $\hat{M}$ toward it. Further, as more time remains in the game, it is more profitable to stick to inflation targeting—i.e., β_3 is decreasing—because there is more time to enjoy the lower future costs

[22]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.

[23]Because output and inflation have equal weights in (20), overweighing the importance of the type comes at the cost of the coefficients in the rest of the states, so $\beta_1 + \beta_2 + \beta_3 \equiv 1$.

of inflation. In the process, however, output dynamics arise due to $a - \hat{a} = \alpha_3(\theta - \hat{M}) \neq 0$, but output volatility is costly (output stabilization term in (20)). The fact that the total signaling coefficient α_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 demonstrates both transparency and restraint early on, understood as carefully guiding the market's belief towards its true type and generating little wasteful inflation, respectively. But as she succeeds in transmitting her private information, her ability to exercise restraint becomes compromised.

Higher-order uncertainty We now turn to examining how the presence of higher-order uncertainty affects the inflationary bias β_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 $\hat{a}_t = \beta_{0t}^{\text{pub}} + \hat{M}_t$.[24] 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 δ_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 $T < T^\dagger$, β_0 is uniformly higher than its counterpart β_0^{pub} over $[0, T)$. (ii) If $r = 0$ and $\sigma_X = +\infty$, β_0 is higher than the public counterpart* for all horizons $T > 0$.

From the proposition, 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—but

[24] We establish this in Proposition S.1 in the Supplementary Appendix. See also Appendix B.2 here.

now 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 if $\sigma_X = +\infty$: the public state L naturally becomes μ, and hence cannot be affected. In this case, the inflationary bias is higher for all horizons, even if the authority is patient.[25]

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 σ_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 σ_X:

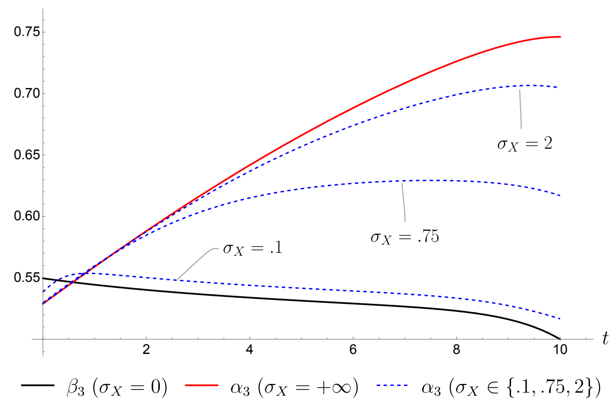


Figure 3: Sender's signaling coefficient as σ_X varies. Other parameters: $(k, \gamma^o, r, \sigma_Y) = (1, 1, 1, 1.5)$

From the plot, the signaling coefficient is decreasing for the most part when σ_X is low (e.g., $\sigma_X = 0.1$), while the opposite occurs when σ_X is large (e.g., $\sigma_X = 2$). In the former case, there is mild higher-order uncertainty, so β_3 dominates in α_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 market beliefs, 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 (10) and the guiding motive decays (both χ and β_1 grow over time).

As σ_X increases, 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, β_3 falls more and more below the public

[25] If $\sigma_X = +\infty$, the reliance on past play is maximized. The representation (10) holds with $L \equiv \mu$, while the ODEs (13)–(14) for (γ, χ) in Lemma 1 simplify dramatically. See Lemma B.1 in Appendix B.

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 σ_X, as Figure 3 shows. Inflation expectations $\hat{a}$ gain more sensitivity when σ_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 σ_X then follows by continuity—that this crossing occurs before the β_0 coefficients cross is because, when σ_X is large, the contribution of α_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.[26]

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.[27]

4.2 Application 2: Reputation for Neutrality

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

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

where $\psi > 0$. The sender (e.g., a politician or expert) has a *bias* θ on a relevant issue; the prior mean—set to $\mu = 0$ for notational convenience—captures an unbiased type. The

[26] 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 α_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 β_0 in some plots is a consequence of this non-monotonicity of α_3, via the same sensitivity logic.

[27] Higher-order uncertainty instead leading to lower (higher) inflation earlier (later) on can happen for low r and σ_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.

sender 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 is predicting the sender's bias (i.e., $a_t = \hat{M}_t$ at all times) and when $\hat{a}_T = \hat{M}_T = \mu = 0$ there is no terminal loss. The receiver could be a news outlet that gets private signals Y of the sender's behavior and that reports its perception of the bias;[28] 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., $\delta_1 \equiv 1$ and $\delta_0 \equiv \delta_2 \equiv 0$). Moreover, $\alpha_{3t} := \beta_{3t} + \beta_{1t}\chi_t \in (0,1)$.*

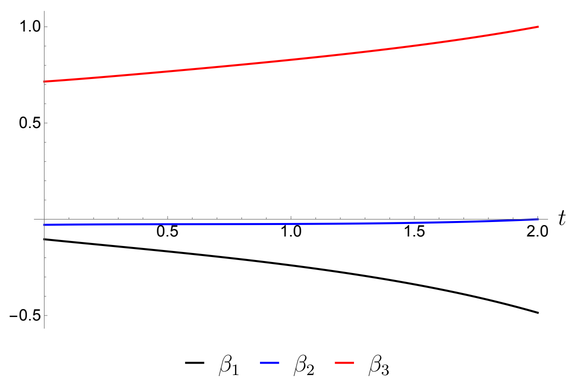


Figure 4: Sender's strategy coefficients in the reputation game: $(\psi, \gamma^o, r, \sigma_X, \sigma_Y) = (1, 2, .1, 1, 1)$.

To understand these coefficients, consider Figure 4. First, if the sender were myopic, she would attach a weight of 1 to θ at all times, precisely the terminal value of β_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

$$-e^{-r(T-t)}\psi\mathbb{E}_t[\hat{M}_T^2] = -e^{-r(T-t)}\psi(\mathbb{E}_t[M_T^2] + \chi_T\gamma_T), \tag{21}$$

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 $\beta_{3t} < 1$

[28] 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.

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 β_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: β_{1t} is decreasing.

It is noteworthy that L is used in the strategies despite *never* appearing in the players' payoffs. To see why, recall that, via the representation, high values of L indicate to the receiver that M is high, and hence that a strong correction (or 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 β_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 $\psi < \sigma_Y^2/\gamma^o$. 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 $\psi < \sigma_Y^2/\gamma^o$. Otherwise, changes in beliefs are costly (large ψ, due to an acute concavity), or beliefs themselves change frequently (due to a large initial uncertainty γ^o 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

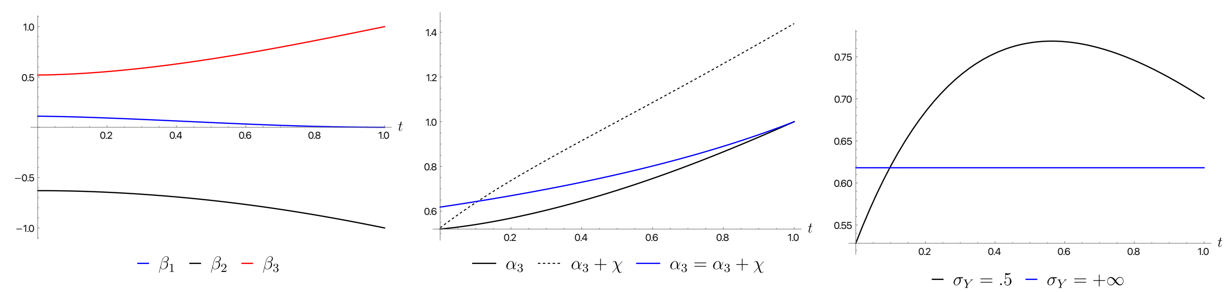
$$\text{sender: } \int_0^T \left[(\theta - \mathbb{E}[\theta|\mathcal{F}_t^X])a_t - \frac{a_t^2}{2}\right] dt; \quad \text{receiver: } (\theta - \mathbb{E}[\theta|\mathcal{F}_t^X])\hat{a}_t - \frac{\hat{a}_t^2}{2}.$$

The sender is an informed trader who knows the fundamental value θ of an asset, and who submits a market order for a_t shares at t. The the receiver is an ex ante uninformed investor who sees a leakage Y of the sender's trades, and submits orders labeled by $\hat{a}_t$, $t \in [0, T]$.[29] The 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 $[t, t+dt)$, while $a_t^2/2$ and $\hat{a}_t^2/2$ encode other types of transaction costs.[30]

Remark 3. *Note that the public signal also carries the sender's action. Our analysis from Sections 2–3 can be extended to this more general case because a representation for M analogous to Lemma 1 continues to hold. Thus, we can examine games featuring a "third action" coming from a player who only sees X (here, market makers who set a price $\mathbb{E}[\theta|\mathcal{F}_t^X]$ for the asset); note that, because of the representation, this is also true in the baseline model. See Supplementary Appendix Section S.4, where we: extend our model to both players affecting X; present the proof for the proposition below; and establish an existence result for this game.*

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.

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 $\Lambda_t := \frac{\gamma_t^X(\alpha_{3t}+\chi_t)}{\sigma_X^2}$.*



(a) Sender's strategy coefficients ($\sigma_Y = .5$)

(b) Black ($\sigma_Y = .5$): α_3 and $\alpha_3 + \chi$ (dashed). Blue ($\sigma_Y = +\infty$): α_3.

(c) Price impact as σ_Y varies.

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

As higher sender types expect a larger profit per unit traded, the equilibrium weight on the type β_3 is positive. At the same time, all types scale back their purchases as the price

[29] 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.

[30] 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.

increases, and so $\beta_2 < 0$. Further, as the endgame approaches and the concern about future prices fades, β_3 and β_2 move toward the myopic solution $(\beta_{3T}, \beta_{2T}) = (1, -1)$ monotonically—see Figure 5a.[31]

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 Λ; in other words, endogenous price impact makes it costly for the sender to place large trades. What is novel is how the presence of private monitoring affects the form that price impact takes, which brings us to the important role that β_1 plays. To illustrate, suppose that the sender has deviated by trading more aggressively, resulting in a higher M than implied by the representation (10). 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 (16) for L in our baseline model). That is, an extra layer of price impact emerges.

This additional 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 $\mathbb{E}[\hat{M}_t|\mathcal{F}_t^X] = L_t$); that $\beta_1 > 0$ 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: Λ in the proposition features β_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.[32] The trader then slows down her purchases for fear of high future prices, and α_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 χ grows over time.

The implication is that price impact Λ 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.[33]

[31] 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.

[32] The appearance of χ 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 $\hat{M}$.

[33] 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

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" (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 $v_{i\cdot}$, $i=0,...,9$ are differentiable functions of time. We can then write the Hamilton-Jacobi-Bellman (HJB) equation for the sender's problem: for all $t<T$,

$$rV=\sup_{a'}\left\{\tilde{u}_t(a',\mathbb{E}_t[\hat{a}_t],\theta)+V_t+\mu_M(a')V_m+\mu_L V_\ell+\frac{\sigma_M^2}{2}V_{mm}+\sigma_M\sigma_L V_{m\ell}+\frac{\sigma_L^2}{2}V_{\ell\ell}\right\}, \quad (22)$$

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 μ_L (respectively, σ_M and σ_L) denote the drifts (respectively, volatilities) in (15) and (16). Note that, via $\tilde{u}$ and $(\mu_M,\mu_L,\sigma_M,\sigma_L)$, (22) 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 (18)). Also, $(\mu_M,\mu_L,\sigma_M,\sigma_L)$ depend on (γ,χ) (see (15)–(16)).

Let $a(\theta,m,\ell,t)$ denote the maximizer in (22). The first-order condition (FOC) reads

$$\frac{\partial u}{\partial a}(a(\theta,m,\ell,t),\delta_{0t}+\delta_{1t}m+\delta_{2t}\ell,\theta)+\underbrace{\frac{\gamma_t\alpha_{3t}}{\sigma_Y^2}}_{dM_t/da_t}\underbrace{[v_{2t}+2v_{5t}m+v_{7t}\theta+v_{9t}\ell]}_{V_m(\theta,m,\ell,t)}=0, \quad (23)$$

which is linear in $a(\theta,m,\ell,t)$ and (θ,m,ℓ). Solving for $a(\theta,m,\ell,t)$ in (23) 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 (θ,m,ℓ). We then impose that the terms attached to these variables on each side, and the constants, coincide, allowing us to link $\vec{\beta}$ with (v_2,v_5,v_7,v_9). At this point, one can always obtain a system of ODEs for v_i, $i=0,...,9$ using the HJB equation. But there is a reduction:

1. We can instead solve for (v_2,v_5,v_7,v_9) as a function of $\vec{\beta}$ and (γ,χ) (see (C.1)-(C.4) in the Appendix); the associated mapping is well-defined provided that α_3 and γ 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 (22), 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

in the Supplementary Appendix. Also in that section, numerical calculations show that the hump-shaped pattern for Λ remains such if the receiver is forward-looking—the fall in Λ is due to the increased market maker's learning (i.e., γ^X falling) eventually dominating when quadratic costs bound the intensity of trades.

coefficients in V. These ODEs are coupled with those of (γ, χ) 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.[34]

This procedure yields a system of ODEs for $(\beta_1, \beta_2, \beta_3, v_6, v_8, \gamma, \chi)$; see Appendix C.[35] To this system we add *boundary conditions*. First, γ and χ satisfy exogenous initial conditions $\gamma_0 = \gamma^o > 0$ 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

$$\beta_{0T} = \frac{u_{a0} + u_{a\hat{a}}\hat{u}_{\hat{a}0}}{1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}},\ \beta_{1T} = \frac{u_{a\hat{a}}[u_{a\theta}\hat{u}_{\hat{a}a} + \hat{u}_{\hat{a}\theta}]}{1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T},\ \beta_{2T} = \frac{u_{a\hat{a}}^2\hat{u}_{\hat{a}a}[u_{a\theta}\hat{u}_{\hat{a}a} + \hat{u}_{\hat{a}\theta}](1 - \chi_T)}{(1 - u_{a\hat{a}}\hat{u}_{\hat{a}a})(1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T)},$$
$$\beta_{3T} = u_{a\theta}, \quad v_{6T} = v_{8T} = 0, \text{ where } u_{x0} := u_x(0,0,0) \text{ and } \hat{u}_{x0} := \hat{u}_x(0,0,0),\ x \in \{a, \hat{a}\}. \tag{24}$$

This fully specifies a BVP that the coefficients $(\beta_1, \beta_2, \beta_3, v_6, v_8)$, along with (γ, χ), 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 (24), all the denominators must be different from zero *after every possible history* of the game, encoded in the possible values that χ_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 will look for equilibria in which the sender always signals her type in equilibrium. As it turns out, a sufficient condition for the total signaling coefficient α_3 to never vanish is that its terminal value, α_{3T}, is non-zero after all possible histories of the game.

[34] Note that (v_0, v_1, v_4) are the coefficients of the constant, θ- and θ^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 (β_0, v_3) are coupled as these encode the *deterministic* component of the sender's incentive to manipulate beliefs, which by definition is independent of the values that the beliefs take.

[35] 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 (16)).

With the aid of (24) then, we require that

$$\alpha_{3T} := \beta_{1T}\chi_T + \beta_{3T} = \frac{u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}\chi_T}{1 - u_{a\hat{a}}\hat{u}_{\hat{a}a}\chi_T}$$

must never vanish.[36] Since $\chi_T \in [0, 1)$ for all histories of play, the numerator will never be zero if $u_{a\theta}$ and $u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}$ have the same sign. Our technical conditions are:

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

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 (γ, χ) 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.[37]

Our core problem of existence of LME, however, is a fixed-point one, which we already anticipated: the learning coefficients (γ, χ) 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 Λ that nests all functions (γ, χ) that can be obtained as solutions to their coupled ODEs (13)–(14) for continuous (β_1, β_3) satisfying a particular uniform bound. Equipped with λ, 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 λ as an input, and

[36]That $\alpha_{3T} \neq 0$ implies that α_3 does not change sign is established in the proof of our main theorem. Since sender's actions read $\alpha_0 + \alpha_2 L + \alpha_3\theta$ along the path of play, it is easy to conclude that an α_3 that never changes sign implies that higher types do hold higher second-order beliefs in equilibrium. (While higher types do take lower actions if $\alpha_3 < 0$, the weights that M attaches to past actions are then negative, and the same conclusion holds.) Note that the receiver can suspend information transmission, when $\delta_1 = \hat{u}_{\hat{a}\theta} + \hat{u}_{\hat{a}a}[\beta_{3t} + \beta_{1t}\chi_t]$ vanishes: the sender simply ignores X, while M updates via her play. Such a suspension is only temporary due to (β_1, β_3) changing over time—and it can never happen if exactly one of $(\hat{u}_{\hat{a}\theta}, \hat{u}_{\hat{a}a})$ is zero.

[37]Special cases for which the one-dimensional shooting is applicable are discussed in Section 6.

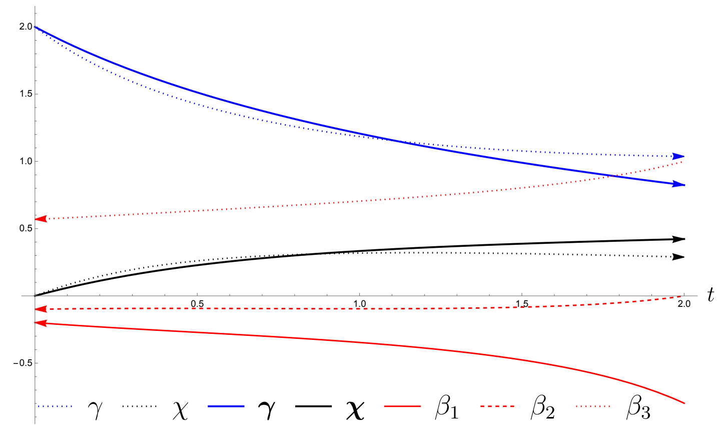


Figure 6: One iteration of our method (reputation game). Start with the dotted curves that point right, capturing learning coefficients. With these, generate candidate equilibrium coefficients, in red pointing left. Use the latter to generate solutions to the learning ODEs in solid pointing right.

where the initial conditions for the ODEs are given by the static time-T conditions of the game (which may depend on λ_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 λ. We then "shoot forward": we feed the resulting (β_1, β_3) pair into the learning ODEs for (γ, χ) 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.

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 ψ is linear (including $\psi \equiv 0$) or ψ is not too concave, there exists a scalar $C > 0$ independent of (r, γ^o) such that, if $T < C/\gamma^o$, there exists an LME for all $r \geq 0$. In this equilibrium, α_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 γ 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 γ, thereby limiting the growth of any solution. Hence, as γ^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 γ^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 ψ then allows us to extract a sufficiently regular selection of static equilibria for all possible (χ_T, γ_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 evolves 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).[38] 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 $\hat{\mathbb{E}}_t[a_t]$ 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 α_3 is deterministic. In other words, the private

[38]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).

sector has not improved its ability to predict a_t at future times, which is what it cares about: the informativeness of Y, via α_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* $\hat{u}(a, \hat{a}, \theta) = -\frac{1}{2}(c_0 + c_1\theta + c_2 a - \hat{a})^2$, *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 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.[39]

Private-value environments and one-dimensional shooting Our general analysis requires two learning dynamics, γ and χ, 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 θ. 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* σ_X *with* $\lim_{\sigma_X \to 0} c_2 = 0$ *and* $\lim_{\sigma_X \to +\infty} c_2 = 1$.

When the players signal at similar rates (as defined above) there is a one-to-one mapping between γ and χ. 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),

[39] In the general case, it is more convenient to skip the reduction steps 1 and 2 after (23) and state the system solely in terms of the value function coefficients $\vec{v}$. See Appendix D for details.

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 χ confirms that less weight is given to the type as σ_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 to construct a deterministic variance-covariance *matrix* when learning about multiple types from linear strategies. Thus, multiple "learning" ODEs naturally arise.[40]

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.[41] An infinite horizon economy can then be approximated by finite-horizon versions.[42]

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

[40]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.

[41]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).

[42]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.

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.

Data Availability Statement

The code underlying this research can be found at https://dx.doi.org/10.5281/zenodo.10607470. There are no data associated with this research.

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 (10) denote a square-integrable process that is X-progressively measurable. Inserting (10) into (8) 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 γ_{1t} (simply γ_t in the main body) that evolve as

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

(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,

$$\begin{aligned} d\hat{M}_t &= (\kappa_{0t} + \kappa_{1t}a_t + \kappa_{2t}\hat{M}_t)dt + B_t^Y dZ_t^Y, \quad \text{where} && \text{(A.2)} \\ \kappa_{1t} = \alpha_{3t}\gamma_{1t}/\sigma_Y^2, \quad \kappa_{0t} &= -\kappa_{1t}[\alpha_{0t} + \alpha_{2t}L_t],\ \kappa_{2t} = -\alpha_{3t}\kappa_{1t},\ B_t^Y = \alpha_{3t}\gamma_{1t}/\sigma_Y. && \text{(A.3)} \end{aligned}$$

On the other hand, since the sender always thinks that the receiver is on path, the public signal evolves, from her perspective, as $dX_t = (\delta_{0t} + \delta_{1t}\hat{M}_t dt + \delta_{2t}L_t)dt + \sigma_X dZ_t^X$. 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

$$\begin{aligned} dM_t &= \underbrace{(\kappa_{0t} + \kappa_{1t}a_t + \kappa_{2t}M_t)dt}_{=\mathbb{E}_t[(\kappa_{0t}+\kappa_{1t}a_t+\kappa_{2t}\hat{M}_t)dt]} + \frac{\gamma_{2t}\delta_{1t}}{\sigma_X^2}[dX_t - (\delta_{0t} + \delta_{1t}M_t + \delta_{2t}L_t)dt] && \text{(A.4)} \\ \dot{\gamma}_{2t} &= 2\kappa_{2t}\gamma_{2t} + (B_t^Y)^2 - (\gamma_{2t}\delta_{1t}/\sigma_X)^2, && \text{(A.5)} \end{aligned}$$

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.[43] 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,

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

[43] Theorem 12.7 in Liptser and Shiryaev (1977) is stated for actions that depend on (θ, 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 $\hat{M}^\dagger$ such that $M^\dagger + A_t = M_t$ as in (A.4) and $\gamma_{2t} = \gamma_{2t}^\dagger$.

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 (10) 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 χ that takes values in $[0,1)$. Indeed, when this is the case, it is easy to see that

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

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

We will find a solution to the χ-equation that is C^1 with values in $[0,1)$. Differentiating $\chi_t = \int_0^t R(t,s)\hat{\kappa}_{3s} ds$ then yields an ODE for χ as below that is coupled with γ_1 and γ_2:

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

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

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

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 γ for γ_1, the first and third ODEs become (13)–(14). The same Lemma A.1 establishes that $0 < \gamma_t \leq \gamma^o$ 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 $(\vec{\hat{\kappa}}, \hat{B})$, (A.6) becomes

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

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

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

The next lemma shows the existence and uniqueness of a solution to the ODE-system (13)–(14) for γ and χ, 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* (13)–(14)*, which satisfies* $0 < \gamma_t \leq \gamma^o$ *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* $\delta_{1t} = \hat{u}_{\hat{a}\theta} + \hat{u}_{a\hat{a}}\alpha_{3t}$*.*

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 γ_1 and the zero function, we obtain $\gamma_1 > 0$; and clearly, $\dot{\gamma}_1 \leq 0$ so $\gamma_1 \leq \gamma^o$. Hence, γ_2/γ_1 is well-defined, and it is easy to verify that it satisfies the χ-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, $\chi_t = \mathbb{E}_t[(M - \hat{M})^2]/\hat{\mathbb{E}}[(\theta - \hat{M})^2]$. We can therefore substitute $\gamma_2 = \chi\gamma_1$ into (A.7) and (A.9) and abbreviate γ_1 to γ to obtain (13)-(14). 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 (13)-(14) exists over $[0, T]$. Suppose by way of contradiction that the maximum interval of existence is $[0, \tilde{T})$. Then since (γ, χ) 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 $\delta_{1t} = \hat{u}_{\hat{a}\theta} + \hat{u}_{a\hat{a}}\alpha_{3t} = \hat{u}_{\hat{a}\theta} + \hat{u}_{a\hat{a}}(\beta_{1t}\chi_t + \beta_{3t})$, where (β_1, β_3) are differentiable. Then the system (13)-(14) changes in that the functional form of the operator is altered (since χ enters δ_1), but importantly, it still satisfies the conditions for the Peano and Picard-Lindelöf theorems, and the arguments above go through. □

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 (17), (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 (θ, 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 it is richer than that in Definition 1, due to the conditioning on M. □

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 (24), where only β_{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 (β_3, γ) are independent of $(\sigma_X, (\chi_t)_{t\in[0,T]})$. Regarding β_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 σ_X). □

Appendix B: Proofs for Section 4

In this section, we prove Proposition 2 and cover the main elements of the proofs of Propositions 3 and 4. The remaining details and the proofs for Sections 4.2 and 4.3 can be found in Sections S.1, S.2, and S.4, respectively, in the Supplementary Appendix.

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 (γ, χ) 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, β_2); by the Picard-Lindelöf Theorem, these are the unique solutions, and hence $\beta_1 + \beta_2 + \beta_3 \equiv 1$. As for α_3, note that its terminal value is $\beta_{1T}\chi_T + \beta_{3T} = \frac{1}{2-\chi_T} > 0$, and its (backward) ODE is

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

Applying the comparison theorem to α_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 $T(\gamma^o; K) \in \Omega(1/\gamma^o)$ such that for all $T < T(\gamma^o; K)$, 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 $(\beta^m_{1t}, \tilde{\beta}^m_{2t}, \beta^m_{3t}) = \left(\frac{1}{2(2-\chi_t)}, \frac{1}{2(2-\chi_t)}, \frac{1}{2}\right)$ are bounded in magnitude by ρ; $(\beta_1, \tilde{\beta}_2, \beta_3)$ differ from their myopic counterparts by at most K and similarly $|\tilde{v}_6|, |\tilde{v}_8| \leq K$; and (γ, χ) are Lipschitz continuous with uniform Lipschitz constants that depend on ρ and K but not T, and that are in $O(\gamma^o)$. Hence, given any constant K, there exists $T^\dagger(\gamma^o; K) \in \Omega(1/\gamma^o)$ such that in addition to the earlier bounds, we also have $|\gamma_t - \gamma^o| \leq K$ and $|\chi_t| \leq K$. Define $g(K) = \frac{K}{4(2-K)}$, which is an upper bound on $||(\beta^m_{1t}, \tilde{\beta}^m_{2t}, \beta^m_{3t}) - (1/4, 1/4, 1/2)||_\infty$, where $||\cdot||_\infty$ denotes the sup norm, when χ_t takes values in $[0, K]$; observe that $g(K) \to 0$ as $K \to 0$. Now for arbitrary $K > 0$, for all $T < T^\dagger(\gamma^o; K)$, 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 β_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^{β_3} is of class C^1. It is easy to check that $f^{\beta_3}(\mathbf{z}) = -\frac{\gamma^o}{16\sigma_Y^2} < 0$, where $\mathbf{z} := (\gamma^o, 0, 1/4, 1/4, 1/2, 0, 0)$ is the vector of coefficients in the equilibrium of the trivial ($T = 0$) game. For any $K > 0$ and $T < T^\dagger(\gamma^o; K)$, 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^{β_3}, for sufficiently small K, if $T < T^\dagger(\gamma^o; K)$, $\dot{\beta}_{3t} < 0$ 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 $\alpha^m_{3t} = \frac{1}{2-\chi_t}$ and $\tilde{v}_3 = \frac{\gamma v_3}{1-\chi}$. The (forward) ODEs for $(\beta_0, \tilde{v}_3)$ are

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

(See `spm.nb` on our websites.) Write the β_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^{β_0} is of class C^1. It is easy to check that $f^{\beta_0}(k, 0, \mathbf{z}) = \frac{k\gamma^o}{4\sigma_Y^2} > 0$ for $\mathbf{z}$ as above. Extending the bounding arguments above to encompass $(\beta_0, \tilde{v}_3)$, for any K there exists $T^\dagger(\gamma^o; K) \in \Omega(1/\gamma^o)$ such that for $T < T^\dagger(\gamma^o; K)$, $||(\beta_0, \tilde{v}_3, \gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8) - (k, 0, \mathbf{z})||_\infty \leq K + g(K)$; thus, by continuity of f^{β_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 β_0 when $\sigma_X > 0$ is always larger than its counterpart when Y is public (or $\sigma_X = 0$), β_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

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

with $\hat{M}_0 = \mu$ and $\gamma_0 = \gamma^o$. 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 γ, which evolves according to (B.1) with initial condition $\gamma_0 = \gamma^o$. This defines a BVP in $(\beta_1, \beta_3, \gamma)$. Equipped with a solution, it is easy to recover β_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 γ^F for the (reversed) γ 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 $\gamma_t := \hat{\mathbb{E}}_t[(\theta - \hat{M}_t)^2]$. 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 (θ, 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 $dM_t = \frac{\alpha_{3t}\gamma_t}{\sigma_Y^2}(a - \alpha_{0t} - \alpha_{2t}\mu - \alpha_{3t}M_t)\,dt$, where $M_0 = \mu$, $\dot{\gamma}_t = -\left(\frac{\alpha_{3t}\gamma_t}{\sigma_Y}\right)^2$, $\alpha_3 = \beta_3 + \beta_1\chi$, and $\chi = 1 - \gamma/\gamma^o$. The HJB equation is

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

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 γ and χ). 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 (22) subject to the equilibrium condition (23) 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 (γ, χ) that have initial conditions. The fixed point will be over pairs of functions (γ, χ): a pair (γ^*, χ^*) that generates mutual best responses that in turn induce learning ODEs whose solution is (γ^*, χ^*). For brevity, we often use α for the sender's signaling coefficient (i.e., we omit subscript 3).

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 χ_t, we write β_{it}^m and α_t^m to refer to these functions of χ_t, suppressing the dependence on χ_t.

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* (β_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 $\beta_{it} = \beta^m_{it}$ 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 $\alpha_t = \alpha^m_t \neq 0$ in such an LME, a contradiction. Second, since the coefficients $(\beta_0, \beta_1, \beta_2, \beta_3)$ and χ 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

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

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

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

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

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 (α, γ) 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).[44] 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

[44] Intuitively, (v_0, v_1, v_4) are the coefficients of the constant, θ- and θ^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 (β_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.

$(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8)$—because (χ, γ) only depend on (β_1, β_3) directly, it follows that $\chi \in [0, 1)$ and $\gamma > 0$ in any solution to this system, and we trivially recover (β_2, v_6, v_8).[45]

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 $\delta_{1t} = \hat{u}_{\hat{a}\theta} + \hat{u}_{a\hat{a}}(\beta_{1t}\chi_t + \beta_{3t})$.

$$
\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)\Big\{\beta_{1t}^2\hat{u}_{a\hat{a}}[2u_{a\hat{a}} + u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}}] \\
&\quad + \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]\Big\} \\
\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_{\hat{a}\theta}\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}_{\hat{a}\theta}\chi_t)]^{-1}\times \\
&\quad \Big\{\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}_{\hat{a}\theta}\chi_t) + (1 - 2u_{a\hat{a}}\hat{u}_{a\hat{a}})\sigma_Y^2\delta_{1t}^2\chi_t^2] \\
&\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}_{\hat{a}\theta}(u_{\hat{a}\theta} + u_{\hat{a}\hat{a}}\hat{u}_{\hat{a}\theta}\chi_t) + \alpha_t(-u_{a\theta} + u_{\hat{a}\theta}\hat{u}_{a\hat{a}} + 2u_{\hat{a}\hat{a}}\hat{u}_{a\hat{a}}\hat{u}_{\hat{a}\theta}\chi_t)] \\
&\quad - \beta_{1t}\sigma_Y^2 u_{a\hat{a}}\delta_{1t}^2\chi_t(2u_{a\theta}\hat{u}_{a\hat{a}} + \hat{u}_{\hat{a}\theta}\chi_t) + \delta_{1t}^2\tilde{v}_{8t}\alpha_t\chi_t(\beta_{1t} - u_{a\hat{a}}\hat{u}_{\hat{a}\theta}) \\
&\quad - \sigma_X^2\delta_{1t}\alpha_t[u_{a\hat{a}}(u_{\hat{a}\theta}\hat{u}_{\hat{a}\theta} - u_{a\theta}\alpha_t) - u_{\hat{a}\hat{a}}u_{a\theta}\delta_{1t}]\Big\}. \\
\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}_{\hat{a}\theta}\chi_t)(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})]^{-1}\times \\
&\quad \Big\{\delta_{1t}^2\alpha_t\chi_t[2\tilde{v}_{6t}(u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}\chi_t) - u_{a\hat{a}}^2\hat{u}_{a\hat{a}}\hat{u}_{\hat{a}\theta}\tilde{v}_{8t}] \\
&\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}_{\hat{a}\theta}(u_{\hat{a}\theta} + u_{a\hat{a}}u_{a\theta} - u_{a\hat{a}}u_{\hat{a}\theta}\hat{u}_{a\hat{a}} + u_{\hat{a}\hat{a}}u_{a\theta}\hat{u}_{a\hat{a}} \\
&\quad + [u_{a\hat{a}}^2 + u_{\hat{a}\hat{a}}]\hat{u}_{\hat{a}\theta}\chi_t) + \alpha_t(u_{\hat{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] \\
&\quad + \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}_{\hat{a}\theta}])] \\
&\quad - \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}}])] \\
&\quad + \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 \\
&\quad + \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 \\
&\quad \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}})] \\
&\quad - \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) \\
&\quad + \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) \\
&\quad + \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 \\
&\quad + \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\}
\end{aligned}
$$

[45] 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.

$$\dot{\gamma}_t = -(\beta_{1t}\chi_t + \beta_{3t})^2\gamma_t^2/\sigma_Y^2, \quad \dot{\chi}_t = \gamma_t\left[(\beta_{1t}\chi_t + \beta_{3t})^2(1-\chi_t)/\sigma_Y^2 - \delta_{1t}^2\chi_t^2/\sigma_X^2\right].$$

This system has two initial conditions $(\gamma_0, \chi_0) = (\gamma^o, 0)$. 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 γ; this latter property will allow us to find horizons for existence that are inversely proportional to γ^o.

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 β_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.[46]

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 χ and are independent of r, (iii) $(\dot{\beta}_1^m, \dot{\tilde{\beta}}_2^m, \dot{\beta}_3^m)$ carry a factor of γ 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*

$$\dot{x}_t^c = -rx_t^c\frac{\alpha_t}{\alpha_t^m} + \frac{\gamma_t h_x(\vec{\beta}^c, \tilde{v}_{6t}, \tilde{v}_{8t}, \chi_t)}{\sigma_X^2\sigma_Y^2(u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}\chi_t)^{n_{1,x}}(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_t)^{n_{2,x}}(1 - u_{a\hat{a}}\hat{u}_{a\hat{a}})^{n_{3,x}}}$$

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

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.*[47] *The initial*

[46] 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.

[47] More precisely, we have $n_{1,x} = 1$, $n_{1,y} = 0$, and $n_{3,\beta_1} = n_{3,\beta_3} = 0$.

conditions are $(\vec{\beta}^c_0, \tilde{v}_{60}, \tilde{v}_{80}) = (0,0,0,0,0)$.

In particular, notice that (i) the terms not containing r continue scaling with γ, (ii) the denominators are bounded away from zero, and (iii) the discount rate term pushes any solution towards zero when α 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* $(\gamma, \chi, \beta^c_1, \tilde{\beta}^c_2, \beta^c_3, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$. Observe that α 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

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

with initial condition $\alpha_0 = \alpha^m_0 = \frac{u_{a\theta} + u_{a\hat{a}}\hat{u}_{\hat{a}\theta}\chi_0}{1 - u_{a\hat{a}}\hat{u}_{a\hat{a}}\chi_0}$ (here, for consistency, χ_0 is the terminal value of χ going forward in time). By Assumption 2, α^m_0 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 α, so it vanishes at $\alpha \equiv 0$. By the comparison theorem, α is always nonzero, as the ODE is locally Lipschitz continuous in α uniformly in time. Moreover, since α^m never changes sign, $\alpha/\alpha^m > 0$.

However, our fixed point argument will input *general* (γ, χ) 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 $\alpha := \beta_1\chi + \beta_3$ never changes sign for any (γ, χ), 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 α in the r term in the centered system; by construction, it will share the sign of α^m and, in any solution to the BVP, will coincide with α. 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_α is a polynomial and $n_{j,\alpha} \in \mathbb{N}$, $j = 1,2,3$. We introduce the (backward) linear ODE

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

with initial condition $\tilde{\alpha}_0 = \alpha^m_0$. 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 α and 0 shows that $\tilde{\alpha}$ never vanishes over its interval of existence *for any pair* (γ, χ) *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 $-rx_t^c \frac{\tilde{\alpha}_t}{\alpha_t^m}$, i.e., with $\tilde{\alpha}$ replacing α in the numerator of the fraction accompanying r. It also includes: the ODEs of $y = \tilde{v}_6, \tilde{v}_8$; the learning ODEs (13)-(14); and the ODE (C.6) of $\tilde{\alpha}$.[48] 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.[49]

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 λ 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 λ 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 λ paths in a suitable set to itself, to which we apply Schauder's fixed point theorem.

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 $\gamma : [0, T] \to [0, \gamma^o]$ with uniform Lipschitz constant $(\gamma^o)^2(2[\rho + K])^2/\sigma_Y^2$ that satisfy $\gamma_0 = \gamma^o$. Likewise, let $X(\rho + K) \subset \mathcal{C}$ denote the space of Lipschitz continuous functions $\chi : [0, T] \to [0, 1]$ with uniform Lipschitz constant $\gamma^o\left[(2[\rho + K])^2/\sigma_Y^2 + (|\hat{u}_{\hat{a}\theta}| + |\hat{u}_{a\hat{a}}|2[\rho + K])^2/\sigma_X^2\right]$ 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 the these Lipschitz constant are motivated by a bounding exercise of the γ and χ 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.

[48]For consistency, the α_t in the r-term in (C.6) and in (13)-(14) must be written as $(\beta_{1t}^c + \beta_{1t}^m)\chi_t + \beta_{3t}^c + \beta_{3t}^m$.

[49]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 2: *Given $(\gamma,\chi) \in \Lambda(\rho+K)$, define a backward initial value problem (IVP) for $(\vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$, 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 $(\vec{\beta}^c, \tilde{v}_6, \tilde{v}_8, \tilde{\alpha})$ previously stated, but where $\hat{\lambda}$ is used in place of the solutions of the learning ODEs. We write this problem as

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

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 λ. The extra dependence on time in the right hand side of ($\mathrm{IVP}^{\mathrm{bwd}}(\hat{\lambda})$) is due to the role of λ in the system.

For all $\lambda_t \in [0,\gamma^o] \times [0,1]$, let $\mathrm{B}(\lambda_t) := (\beta_1^m(\lambda_t), \tilde{\beta}_2^m(\lambda_t), \beta_3^m(\lambda_t), 0, 0, \alpha^m(\lambda_t))$. From here, we define $\rho := \sup_{\lambda_t \in [0,\gamma^o]\times[0,1]} ||\mathrm{B}_{-6}(\lambda_t)||_\infty > 0$, with B_{-i} denoting as usual the vector B excluding B_i.[50] For arbitrary $K > 0$, we now establish sufficient conditions for ($\mathrm{IVP}^{\mathrm{bwd}}(\hat{\lambda})$) to have a unique solution for each $\lambda \in \Lambda(\rho+K)$.

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

Proof. Fix any $\lambda \in \Lambda(\rho+K)$. Since λ is continuous in t and $\mathrm{f}^{\hat{\lambda}}$ is of class C^1 with respect to $\mathbf{b}_t$, $\mathrm{f}^{\hat{\lambda}}$ 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 $T(\gamma^o;K)$ 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 α in the r terms), we have for $i \in \{1,2,3\}$ and $j \in \{4,5\}$

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

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

[50] We exclude $\tilde{\alpha}$ from the definition of ρ because it does not enter the ODEs for the learning coefficients explicitly, and hence it does not affect the definition of $\Lambda(\rho+K)$.

that uses $\chi_t \in [0,1]$ yields the existence of a scalar $h_i(K)$ such that $|\hat{\gamma}_s h_i(\mathbf{b}_s, \hat{\chi}_s)| \leq \gamma^o h_i(K)$, $i \in \{1, 2, \dots, 5\}$, where we have used that $\gamma_t \in [0, \gamma^o]$ 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 $j \in \{4, 5\}$,

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

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 $T(\gamma^o; K) := \min_{i \in \{1,\dots,5\}} \frac{K}{\gamma^o h_i(K)} > 0$ such that (IVP$^{\text{bwd}}(\hat{\lambda})$) with $T < T(\gamma^o; K)$ by construction admits a unique solution satisfying $|\mathbf{b}_{-6}| < K$ for all $\lambda \in \Lambda(\rho + K)$. Moreover, since $T(\gamma^o; K)$ is independent of r, the statement holds for all $r \geq 0$; also $T(\gamma^o; K) \in \Omega(1/\gamma^o)$.[51] □

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

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

that for each λ delivers the induced "total" 'β_1' and 'β_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 $\lambda \mapsto \mathrm{B}_i(\lambda_{(\cdot)})$ 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 $F : X \times Y \to U$ of class C^1 and $\omega : Y \to X$. 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 $x_0 = \omega(y_0)$*

[51] 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 $\mathbf{b}_i$, $i \in \{1, \dots, 5\}$.

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

$$||x^1_t - x^2_t||_\infty \leq k_1 ||\omega(y^1_0) - \omega(y^2_0)||_\infty + k_2 \sup_{s\in[0,T]} ||y^1_s - y^2_s||_\infty, \qquad \textit{for all } t \in [0,T].$$

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

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

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

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

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 $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$

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

consisting of the two (forward) learning ODEs (13)-(14) that use as input $q(\lambda) = (q_1(\lambda), q_2(\lambda))$ playing the role of (β_1, β_3)—here, the first (second) entry of the system corresponds to the γ-ODE (χ-ODE), while the boldface convention aims at distinguishing between inputs λ 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 (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 γ-ODE and χ-ODE, we have that

$$|\dot{\boldsymbol{\lambda}}_{1t}| = \left| -\frac{\boldsymbol{\lambda}^2_{1t}([q_2(\lambda)]_t + [q_1(\lambda)]_t \boldsymbol{\lambda}_{2t})^2}{\sigma^2_Y} \right| \leq (\gamma^o)^2 (2[\rho+K])^2/\sigma^2_Y \qquad \text{and similarly}$$
$$|\dot{\boldsymbol{\lambda}}_{2t}| \leq \gamma^o \left[(2[\rho+K])^2/\sigma^2_Y + (|\hat{u}_{\hat{a}\theta}| + |\hat{u}_{a\hat{a}}|(2[\rho+K]))^2/\sigma^2_X \right]$$

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 (IVP$^{\text{fwd}}$($q(\lambda)$)) by setting $x = \boldsymbol{\lambda}$, $y = q(\lambda)$, $\omega(y_0) = (\gamma^o, 0)$, $F(x_t, y_t) = \mathrm{f}^{q(\lambda)}(\boldsymbol{\lambda}_t, t)$, $X = [0,\gamma^o] \times [0,1]$ 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 (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/γ^o. □

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 $\beta_{3t} := \beta_{1t}^c + \beta_{1t}^m$—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 χ, β_{1t}^c, and β_{3t}^c yields that α 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 α 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 $(\gamma, \chi, \beta_1, \tilde{\beta}_2, \beta_3, \tilde{v}_6, \tilde{v}_8) = (\lambda^*, \hat{\mathbf{b}}_{-6}(\cdot;\lambda^*) + \mathrm{B}_{-6}(\lambda^*))$ 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 (β_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 θ 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 σ-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, γ^o), such that if $\psi_{\hat{a}\hat{a}} \in (C_\psi/\gamma^o, 0]$ and $T < C_T/\gamma^o$, a linear Markov equilibrium exists for all $r \geq 0$. Moreover, α_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 $(\hat{a}'_t)_{t\in[0,T]}$,*

$$d\hat{M}_t = \frac{\alpha_{3t}\gamma_{1t}}{\sigma_Y} dZ_t \tag{D.1}$$

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

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,

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

where $\mu_{\hat{M}}$ $(= 0)$, $\sigma_{\hat{M}}$, $\mu_L(\hat{a}')$, and σ_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.) Morevoer, with an LQG structure we guess a solution of the form

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

where $v_{j\cdot}, j = 0, \ldots, 5$, 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 (γ, χ). (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 (γ, χ), we obtain a system of ODEs in $(\vec{v}, \vec{\hat{v}}) := (v_0, \ldots, v_9, \hat{v}_0, \ldots, \hat{v}_5)$ and (γ, χ). The induced BVP has $v_{iT} = \hat{v}_{jT} = 0$ for $i = 0, \ldots, 9$ and $j = 0, \ldots, 5$ as terminal conditions (absent lump sum payoffs at T), while the initial conditions $(\gamma_0, \chi_0) = (\gamma^o, 0)$.

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

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

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 γ and χ to use in defining the domain for our fixed point argument. The form of (D.5) implies that, starting with (γ, χ) in this domain, there exists $T(\gamma^o) \in \Omega(1/\gamma^o)$ independent of r such that solutions $(\vec{v}, \vec{\hat{v}})$ traced backward are bounded by K. In turn, (γ, χ) 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

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

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 $t = T$,

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

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

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

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

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

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

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

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

Proof of Proposition 9. Suppose $\delta_1 = \hat{u}_{a\hat{a}}\alpha_3$, in which case the χ-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 $f : [0,\bar{\chi}) \to [0,\gamma^o]$, 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 $\hat{u}_{\hat{a}\theta}$ 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 χ-ODE (even when $\alpha_3 = 0$). We have

$$\frac{d(\chi(\gamma_t))}{dt} = \frac{\alpha_{3t}^2\gamma_t}{\sigma_Y^2[c_1 + c_2(\gamma/\gamma^o)^d]^2}c_1c_2d[c_1+c_2]\left(\frac{\gamma_t}{\gamma^o}\right)^d.$$

By construction, moreover, $c_1c_2 = c_1 - c_2 = \frac{\sigma_X^2}{\sigma_Y^2(\hat{u}_{a\hat{a}})^2}$, which follows from equating the first- and zero-order coefficients in $Q(\chi) = \hat{u}_{a\hat{a}}^2\chi^2/\sigma_X^2 + \chi/\sigma_Y^2 - 1/\sigma_Y^2 = \hat{u}_{a\hat{a}}^2(\chi-c_2)(\chi+c_1)/\sigma_X^2$. Thus, $dc_1c_2 = c_1 + c_2$. On the other hand,

$$\frac{[\hat{u}_{a\hat{a}}\chi(\gamma)]^2}{\sigma_X^2} = \frac{\hat{u}_{a\hat{a}}^2}{\sigma_X^2}\left[c_1c_2\frac{1-(\gamma/\gamma^o)^d}{c_1+c_2(\gamma/\gamma^o)^d}\right]^2 = \frac{c_1^2(1-c_2)}{\sigma_Y^2}\left[\frac{1-(\gamma/\gamma^o)^d}{c_1+c_2(\gamma/\gamma^o)^d}\right]^2$$

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

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

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 $0 = c_1^2 - c_1^2 c_2 - c_1^2 + c_1^2 c_2$, $(c_1+c_2)^2 = 2c_1c_2 - c_1c_2(c_2-c_1) + 2c_1^2(1-c_2)$ 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 χ-ODE; by uniqueness, $\chi = \chi(\gamma)$.

We now prove the final statement of the lemma. When $\gamma_t \in (0, \gamma^o]$, we have $\chi_t = \frac{c_1c_2(1-[\gamma_t/\gamma^o]^d)}{c_1+c_2[\gamma_t/\gamma^o]^d} < \frac{c_1c_2}{c_1} = c_2$. Now c_2 simplifies to $\frac{\sqrt{\sigma_X^4 + 4\sigma_Y^2\sigma_X^2\hat{u}_{a\hat{a}}^2} - \sigma_X^2}{2\hat{u}_{a\hat{a}}^2\sigma_Y^2} = \frac{4\sigma_Y^2\hat{u}_{a\hat{a}}^2}{2\hat{u}_{a\hat{a}}^2\sigma_Y^2\left(\sqrt{1+4\sigma_Y^2\hat{u}_{a\hat{a}}^2/\sigma_X^2}+1\right)}$, which by inspection is increasing in σ_X and has limits $\lim_{\sigma_X \to 0} c_2 = 0$ and $\lim_{\sigma_X \to +\infty} c_2 = 1$ $\square$

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.

Stigler, G. J. (1964): "A theory of oligopoly," *Journal of Political Economy*, 72, 44–61.

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.