Two-Sided Learning and Moral Hazard

Gonzalo Cisternas*
MIT Sloan

February 2014

Abstract

I study games of ex-ante symmetric uncertainty about a payoff-relevant state variable. In these games, a long-run player and a population of small players learn about a hidden state from a public signal that is subject to Brownian shocks. The long-run player can influence the small players' beliefs by affecting the publicly observable signal or by affecting the hidden state itself, in both cases in an additively separable way. The impact of the small players' beliefs on the long-run player's payoffs is nonlinear. I derive necessary conditions for Markov Perfect Equilibria in public beliefs that take the form of ordinary differential equations (ODEs). These ODEs capture how the long-run player's equilibrium actions optimally balance the size of marginal flow payoffs, cost-smoothing motives and ratchet forces, across different levels of public beliefs. I obtain verifiable sufficient conditions that ensure that a solution to this ODE is an equilibrium, and use them to show the existence of equilibria in environments where the underlying fundamental is a Gaussian process. Finally, I develop applications to evaluate monetary policy under partial information, to analyze government's incentives to manipulate official statistics, and to study reputation dynamics in labor markets.

Keywords: Continuous-time games, incentives, learning, asymmetric information, Brownian motion.

JEL codes: C73, D81, D82, D83.

---

*Email: gcistern@mit.edu. I would like to thank Yuliy Sannikov for his invaluable advice, Dilip Abreu and Stephen Morris for helpful discussions, and Alessandro Bonatti, Markus Brunnermeier, Alberto Cavallo, Hector Chade, Eduardo Faingold, Leandro Gorno, Diogo Guillén, Jay Lu, Tomasz Sadzik, Ramon van Handel and all seminar participants at Harvard-MIT, MIT Sloan, NYU Stern, Stanford GSB, Toulouse School of Economics, UCLA, UCSD and the University of Minnesota for their constructive comments.

1 Introduction

This paper develops a new class of continuous-time games for analyzing incentives in settings where agents learn about the relevant economic environment. In these games, a long-run player and a population of small individuals share a common prior regarding the initial value of a hidden state variable, and subsequently learn about its future evolution from a public signal that is subject to Brownian shocks. The long-lived player can nevertheless control the small players' beliefs about the unobserved state by taking costly hidden actions that affect the evolution of the public signal. Using continuous-time techniques, I study the actions that the long-run player takes in any equilibrium in which the small players perfectly anticipate his behavior, and thus beliefs remain aligned.

Strategic behavior driven by both imperfect observability of an agent's actions and incomplete information about the environment is pervasive in economics. In labor markets, when wages are based on perceived ability, workers exert effort in an attempt to manipulate their employers' beliefs about their unobserved skills (Holmstrom (1999)). Similarly, when the value a product is given by the public perception of its unobserved quality, a firm's investment policy will depend on the nature of the learning process of its costumers (Board and Meyer-ter-Vehn (2013)). In monetary policy, imperfect control over money growth allows a central bank to surprise an economy with high inflation when the population learns about the policymaker's preferences for stimulating the economy (Cukierman and Meltzer (1986)).

Unlike most of the existing literature on dynamic games of incomplete information, this paper is concerned with environments in which all agents are uncertain about the economic environment.[1] Two-sided learning and moral hazard appears when both a policymaker and a population of individuals learn about unobserved components of inflation, and monetary policy has imperfect control. It also arises when a government secretly manipulates inflation statistics in an attempt to anchor expectations about a hidden inflation trend. Similarly when a contractor and a government learn about the contractor's efficiency to deliver goods, and the contractor takes unobserved actions to affect performance. The methods developed in this paper are crucial for our understanding of dynamic incentives in settings where (i) there is ex-ante symmetric uncertainty about a payoff-relevant variable, and where (ii) imperfect monitoring gives rise to incentives that are driven by the possibility of *affecting the perceptions of others*.

Quantifying the incentives that arise in these environments imposes a host of technical challenges. In the class of games I analyze, the long-run player's payoffs are determined by the small players' actions, which in turn depend on their beliefs about the hidden state. Starting from a common prior, if the small players anticipate the long-run player's actions, public and private beliefs will remain aligned. However, the long-run player's incentives *on the equilibrium path* are determined by the benefits from hypothetical deviations *off the equilibrium path*. Evaluating the long-run player's off-equilibrium payoffs is difficult for two reasons. First, after a deviation takes place the long-run player acquires private information about the evolution of the unobserved state.[2] Second, and most importantly, off the equilibrium path the small players' beliefs become biased. The long-lived agent

[1] The most notable exception is Holmstrom (1999).

[2] This private information is persistent and changes stochastically over time, as it is driven by a learning process.

can thus condition his actions on the values that both private and public beliefs take, and the small players will construct beliefs using a wrong conjecture about equilibrium play.

I focus on Markov Perfect Equilibria (MPE) in pure strategies, with key state variables being the long-run player's private belief about the hidden state, and the *belief-asymmetry process*, which is a measure of the degree of discrepancy between private and public beliefs. In such a context, I develop a first-order approach for quantifying the value of a local deviation off the equilibrium path, which ultimately determines the long-run player's incentives on the margin. More specifically, I show that in any MPE in which beliefs remain aligned, the value that the long-run player attaches to inducing a small degree of asymmetry between private and public beliefs is the solution to an ordinary differential equation (ODE), which I refer to as the *incentives equation.* The incentives equation captures how the incentives to distort public beliefs vary across different levels of public opinion. Such variations occur when learning, the long-run player's preferences or the small players' actions are *nonlinear*.

This equation is of considerable importance. From an economic perspective, it delivers in a strikingly clean way all the forces that drive belief manipulation motives. From a computational perspective, it reduces the problem of computing Markov perfect equilibria from solving a complex partial differential equation (PDE), to computing solutions of a nonlinear ODE. From a technical perspective, it transforms the problem of checking global incentive compatibility (and thus, the problem of existence of MPE), to the problem of finding solutions to an ODE that must also satisfy some additional properties. Furthermore, this equation is obtained under very weak assumptions, with no major restrictions imposed on the functional form that payoffs can take. Two assumptions are nonetheless important. First, I restrict to learning processes that admit posterior distributions summarized by a one-dimensional state variable. Second, the long-run player's actions enter in an additively separable way into the corresponding dynamics.[3]

The generality of the methods developed in this paper – mostly driven by the advantage of continuous-time methods over traditional ones – makes them applicable to a wide range of economic environments. Furthermore, they enable us to address applied questions for which we did not have tools in the past. In Section 3, I study models of commitment in settings where agents have partial information about the relevant environment. For instance, if agents learn about unobserved components of inflation, what are the inefficiencies generated by a government that manipulates inflation statistics in an attempt to control the market's expectations about future inflation?[4] Or how does the inflationary bias introduced a policymaker who lacks commitment depend on the characteristics of the market's learning process? The difficulty in analyzing these environments is lies on the fact that a government's or policymaker's payoffs are best modeled as nonlinear functions of the belief about the unobserved state. Instead, the existing literature allowing for two-sided learning has analyzed value-creation in fairly linear environments. In the linear and additive model of career concerns developed by Holmstrom (1999), workers with the same tenure exert identical effort levels, despite their differences in perceived

[3]The first assumption is purely for tractability reasons. Relaxing the second one either introduces experimentation motives (something beyond the scope of this paper), or it conflicts the first assumption, as the learning present in the model requires keeping track of distributions.

[4]Manipulation of an official inflation statistic was documented by Cavallo (2012) for the case of Argentina.

abilities or wages. In the reputation model of Board and Meyer-ter-Vehn (2013), a firm's investments in quality are independent of its private information about current quality. As we move away from these tractable frameworks, incentives will depend on the value of becoming (hypothetically) privately informed, and such value will vary across different levels of public beliefs.

The incentives equation, as a necessary condition for equilibrium behavior, does not ensure that the long-run player does not benefit from inducing a large degree of belief asymmetry. In order to verify incentive compatibility globally, it is necessary to study the long-run player's payoffs for deviations of any size. Intuitively, the long-run player will engage in "double deviations" when he can exploit the benefit of being privately informed about the unobserved state. For the case of Gaussian learning I derive a *verification theorem* (Theorem 4.4) that states conditions under which a solution to the incentives equation is effectively a MPE, with the main condition precisely pertaining to a bound on the information rent that the long-run player can obtain if a deviation takes place. In Section 5 I verify analytically that this condition holds in the case in which the long-run player's marginal flow payoff is bounded and the underlying fundamental is a Brownian martingale, provided there is some discounting. Beyond this case, the corresponding condition can be verified ex-post on any numerical solution to the incentives equation.

For settings beyond the one just described, one can directly investigate the long-run player's value function. Interestingly, this function satisfies a new type of PDE characterized by its particular *nonlocal* structure: the local evolution of the long-run agent's utility *off* the equilibrium path depends on the marginal value of belief asymmetry *along* the equilibrium path. This is because the signal structure satisfies the full support assumption, so the small players always construct beliefs as if the long-run player had never deviated. This new class of PDEs correspond to a standard Hamilton-Jacobi-Bellman (HJB) equations that also satisfy the requirement that the small players must anticipate the long-run player's actions on the equilibrium path. Since the latter requirement corresponds an additional constraint on HJB equation, classic verification theorems apply.

I use this insight to show the existence of a linear MPE for a class of games with linear-quadratic structure (linear learning and quadratic payoffs) in which the mentioned PDEs admit analytic solutions. I exploit the tractability of this linear-quadratic framework in two applications. First, I show that once people recognize the possibility that a government can secretly manipulate inflation statistics to control people's expectations, the government of a high-inflation economy can become trapped into this type of manipulation, and even during recessions. Second, in an exercise that analyzes the role of monetary policy in affecting unobserved components of inflation, I determine how the size of the inflationary bias created by a central bank depend on key parameters of the market's learning process. In particular, when both the policymaker's control worsens and prices become less rigid, the incentives to surprise the economy with inflation decrease.

The paper is organized as follows. In Section 2 I study the class of games in general form, and derive necessary conditions for equilibrium behavior. Section 3 discusses applications and extensions of the methods presented in Section 2. Section 4 presents two verification theorems (sufficient conditions). Section 5 studies games of Gaussian learning and uniformly bounded marginal flow payoffs. Section 6 is devoted to studying games with a linear-quadratic structure. Section 7 concludes. All proofs are relegated to the Appendix.

1.1 Literature

Models of reputation in labor markets are a natural application of the class of games studied here. Holmstrom (1999) develops a linear and additive model of career concerns in which equilibrium effort depends only on tenure, and hence are independent of the worker's history (i.e. reputation). In the static model of Dewatripont et al. (1999), introducing complementarities between effort and skills in the output technology generates strategic complementarity between realized effort and conjectured effort. More recently, Bonatti and Hörner (2013) show how current and future effort can become strategic substitutes. More generally, the incentives equation states that equilibrium incentives will depend on a worker's past performance, on how reputation is expected to evolve over time (thus connecting effort decisions across time) and on the possibility of affecting the market's contemporaneous conjecture of equilibrium play. Finally, both Kovrijnykh (2007) and Martinez (2009) study discrete-time finite-horizon models of career uncertainty in which the necessary conditions for equilibrium incentives they derive depend on the worker's reputation. However, unlike in my paper, the question of global incentive-compatibility (and consequently, the validity of the first-order approach) is not addressed.

In the context of investment games with learning, Board and Meyer-ter-Vehn (2013) study a firm's reputation dynamics when the quality of a product is unobserved by its customers. In their model, the investment policy depends on the type of learning of its customers (good news or bad news), yet is independent of the firm's private information about quality. Board and Meyer-ter-Vehn (2010) obtain necessary conditions for a firm's investment policy in a model that extends the previous one by allowing for exit. In the case in which the firm is also learning about its product, investment depends on the value of inducing belief asymmetry, which now takes an integral form as valuations can jump.

This paper is also related to the literature on optimal contracts in environments where actions have persistence, in the sense that deviations generate a wedge between the agent's and the principal's perceived distribution of all future payoff-relevant variables. In DeMarzo and Sannikov (2011), Jovanovic and Prat (2013) and He et al. (2014) output carries noisy information of both the firm's unobserved fundamentals and the agent's hidden actions, and optimal contracts under specific functional forms are derived. Finally, the necessary and sufficient conditions for equilibrium incentives that I derive are similar to the ones derived both by Williams (2011) in a context of persistent private information, and by Sannikov (2014) in a setting where actions have long-term impact on performance, in the characterizations of their optimal contracts. In all these settings time is continuous and additive separability is a crucial assumption.

To conclude, this paper contributes to the study of dynamic incentives using continuous-time techniques. In particular, I exploit the connection of stochastic control with the theory of differential equations to considerably expand the class of environments with incomplete information that can be analyzed. Two related papers studying continuous-time games between a large player and a continuum of small players are Faingold and Sannikov (2011) and Bohren (2014). The first paper studies reputation dynamics in settings with imperfectly observable actions and one-sided learning about the large player's fixed type. Bohren (2014) analyzes a general class of investment games with imperfectly observable actions but without learning. In both papers the equilibrium analysis is simplified by the fact that there is only one payoff-relevant variable.

2 Signal Manipulation Games: General Case

2.1 Set-up

Consider an economy in which one long-run player and a population of small players simultaneously learn about a hidden state variable. The unobserved process is denoted by $\theta := (\theta_t)_{t\geq 0}$ and it takes values in $\Theta \subseteq \mathbb{R}$. Hereinafter I refer to the hidden state variable as the *fundamentals*.

In a signal-jamming game fundamentals are exogenous, and the long-run player affects the signal from which the small players extract information about the unobserved state. More specifically, there is public signal $\xi := (\xi_t)_{t\geq 0}$ that takes the form

$$d\xi_t = (a_t + \theta_t)dt + \sigma_\xi dZ_t,\ t \geq 0, \tag{1}$$

where $Z := (Z_t)_{t\geq 0}$ is a Brownian motion independent of θ, and $\sigma_\xi > 0$ is a volatility parameter. The term a_t represents the degree of signal manipulation exerted by the long-lived player at time $t \geq 0$. The long-run player's manipulation choices are not observed by the rest of the economy and they take values in a set $A \subset \mathbb{R}$. Observe that the signal structure (1) satisfies the full-support assumption with respect the the long-run player's actions.

Notice that the long-run player effectively observes the component of the public signal that is not explained by signal manipulation, $Y_t := \xi_t - \int_0^t a_s ds,\ t \geq 0$. By definition

$$dY_t = \theta_t dt + \sigma_\xi dZ_t,\ t \geq 0, \tag{2}$$

from where we can see that Y is an exogenous process that is privately observed by the long-run agent. Equations (1) and (2) yield that signal manipulation has an *additive* structure in this paper. In the sequel, $\mathcal{F}_t^Y$ denotes the information generated by the process Y up to time t, and $\mathbb{F}^Y := (\mathcal{F}_t^Y)$ the (completed) filtration associated to Y. The corresponding analogous notation is used to denote the filtration generated by ξ.

The long-run player uses the information conveyed by Y to construct private beliefs about θ. The *private beliefs process* is denoted by $\rho := (\rho_t)_{t\geq 0}$ where

$$\rho_t(x) := \mathbb{P}(\theta_t \leq x | \mathcal{F}_t^Y),\ x \in \Theta,\ t \geq 0. \tag{3}$$

In this definition, $\mathbb{P}(\cdot|\mathcal{F}_t^Y)$ corresponds to the long-run player's posterior belief about θ_t given the observations $(Y_s : s \in [0,t])$, constructed via Bayes rule, $t \geq 0$.

The small players use the information conveyed by the public signal ξ to learn about the fundamentals. The *public belief* process is denoted by $\rho^* := (\rho_t^*)_{t\geq 0}$ and defined by

$$\rho_t^*(x) := \mathbb{P}^*(\theta_t \leq x | \mathcal{F}_t^\xi),\ x \in \Theta,\ t \geq 0, \tag{4}$$

where $\mathbb{P}^*(\cdot|\mathcal{F}_t^\xi)$ denotes the small players posterior belief about θ_t given the partial observations $(\xi_s : s \in [0,t])$ (constructed using Bayes rule) given their *subjective* belief about the distribution of ξ, as the latter depends on their conjecture of equilibrium play. Private and public beliefs coincide if, for instance, the small players have perfectly anticipated the long-run players past actions. However, moral hazard can give rise to potential divergence between private and public beliefs. I assume that $\rho_0(\cdot) = \rho_0^*(\cdot)$, so there is ex-ante symmetric uncertainty about the initial state of the underlying fundamental.

The small players act myopically because their individual actions are anonymous and they do not affect the population average observed by the long-run agent. Thus, at any instant of time the small players maximize their ex-ante flow payoffs. Ex-ante flow payoffs in turn depend on the small player's beliefs about the current value of the underlying fundamental θ_t, and on the action that they conjecture the long-run player is currently following, a^*_t, $t \geq 0$. I summarize the small players' actions in the best-response function

$$\mathtt{b}(\rho^*_t, a^*_t),\ t \geq 0.$$

The small player's actions affect the long-run player's flow payoff. Hence, the long-run player cares about the fundamentals to the extent that they influence the small player's learning process. Given any strategy $b := (b_t)_{t\geq 0}$ of the small players and any manipulation strategy $a := (a_t)_{t\geq 0}$, the long-run player's payoffs at time t take the form

$$\int_t^\infty e^{-r(s-t)}(u(b_s) - g(a_s))ds,\ t \geq 0. \tag{5}$$

The function $u : \mathbb{R} \to \mathbb{R}$ represents the component of the long-run player's flow utility that is affected by the small player's actions. The cost of manipulation is given by a differentiable function $g : A \to \mathbb{R}$ which is convex and that satisfies $g(0) = 0$.[5]

As it is traditional in the literature of stochastic control, I allow for both strategies and conjectured strategies to satisfy mild integrability conditions:

Definition 2.1. *A manipulation strategy* $a := (a_t)_{t\geq 0}$ *is said to be feasible if it corresponds to a progressively measurable process with respect to the information generated by* (ξ, Y)*, also satisfying the integrability condition*

$$\mathbb{E}\left[\int_0^t a_s^2 ds\right] < \infty,\ \forall t \geq 0. \tag{6}$$

A conjecture $a^* := (a^*_t)_{t\geq 0}$ *is said to be feasible if it corresponds to a progressively measurable process with respect to the information generated by* ξ *that also satisfies (6). In both cases,* $\mathbb{E}[\cdot]$ *corresponds to the expectation operator under the long-run player's probability measure of events.*

With this in hand, we can define the equilibrium concept:

Definition 2.2. *An equilibrium consists of (i) a feasible manipulation strategy of the long-run player* $a_t(\xi_s, s \in [0,t], \rho_t)$, $t \geq 0$*; (ii) a public strategy of the small players* $b_t(\xi_s, s \in [0,t])$, $t \geq 0$*; (iii) a private belief process* $\rho_t(\cdot|Y_s, s \in [0,t])$, $t \geq 0$*; and (iv) a public belief process* $\rho^*_t(\cdot|\xi_s : s \in [0,t])$, $t \geq 0$*, such that:*

(a) *The strategy* $a := (a_t)_{t\geq 0}$ *maximizes the long-run player's continuation payoff*

$$\mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}(u(b_s) - g(a_s))dt\right], \tag{7}$$

given the small player's public strategy $(b_t)_{t\geq 0}$ *and the long-run player's belief process* $\rho := (\rho_t)_{t\geq 0}$*, after any history* $(\xi_s : s \in [0,t], a_s, s \in [0,t])$, $t \geq 0$*;*

[5] Provided that ex-ante payoffs can be written as some function of the private and public beliefs, the analysis that follows can be extended to situations in which the long-run player's flow payoff $u(\cdot)$ also depends on the current state of θ.

(b) $b_t = \mathtt{b}(\rho_t^*, a_t^*)$ *for some measurable function* $\mathtt{b}$*, which is optimal given* ρ_t^* *and the action* $a_t^*(\xi_s, s \in [0,t]) = a_t(\xi_s, s \in [0,t], \rho_t^*)$*,* $t \geq 0$*;*

(c) *The long-run player's beliefs* $\rho_t = \rho_t(\cdot|Y_s, s \in [0,t])$ *are determined by Bayes' rule;*

(d) *The small player's beliefs* $\rho_t^* = \rho_t^*(\cdot|\xi_s, s \in [0,t])$ *are determined by Bayes' rule under the assumption that the long-run player has been following the strategy* $a_t^*(\xi_s, s \in [0,t]) = a_t(\xi_s, s \in [0,t], \rho_t^*)$*.*

Observe that the histories of the form $(\xi_s : s \in [0,t], \rho_t)$, $t \geq 0$, summarize all the payoff relevant information for the long-run player, so we can restrict the set of feasible manipulation strategies to feasible strategies as in (i). Part (ii) instead requires that the small player's actions depend on the information generated by ξ only.

Concerning the equilibrium conditions, (a) states that the long run player's strategy specifies actions both on and off the equilibrium path, and has to maximize future payoffs on and off the equilibrium path. In contrast, the optimality of the small player's actions is checked only on the equilibrium path (condition (b)), as the full support assumption makes any partial observation $(\xi_s : s \in [0,t])$, $t \geq 0$, consistent with equilibrium play. Condition (c) and (d) correspond to the consistency requirements that both belief processes ρ and ρ^* must be constructed via Bayes' rule using the strategies specified by equilibrium play. In particular, (d) states that the small player's beliefs are always constructed as if the long-run player is effectively following the strategy prescribed on the equilibrium path.

2.2 Learning and Belief Manipulation

In this section I present a unified approach for studying belief manipulation dynamics in settings where the posterior distributions ρ and ρ^* can be summarized by one-dimensional diffusions. Such settings correspond to fundamentals in the form of Gaussian diffusions, or in the form of two-state Markov-switching processes.[6] This general approach reveals that it is the additively separable technology – rather than the particular nature of fundamentals– the key assumption behind the characterization of incentives in the form of ODEs.[7]

2.2.1 Unidimensional Learning Processes

Definition 2.3. (Linear and Nonlinear Unidimensional Learning Processes).

(i) A private learning process is said to be linear (or Gaussian) if fundamentals evolve according to an Ornstein-Uhlenbeck process of the form

$$d\theta_t = -\kappa(\theta_t - \eta)dt + \sigma_\theta dZ_t^\theta, \; t \geq 0, \tag{8}$$

the signal process is given by (2), and the initial prior $\theta_0|\mathcal{F}_0$ *is normally distributed.*

[6] Going beyond these two classes the analysis becomes intractable, or the set of one-dimensional sufficient statistics cease to have economically meaningful interpretations.

[7] For instance, if posterior distributions are summarized by a finite set of one-dimensional state variables, incentives will be characterized by a multidimensional equation.

(ii) A private learning process is said to be nonlinear if fundamentals θ evolve as a two-state Markov chain and the signal structure is given by (2). That is, there exist $\ell < h \in \mathbb{R}$ such that θ takes values in $\Theta = \{h, \ell\}$, and there is a generator matrix

$$\Lambda := \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_0 & -\lambda_0 \end{bmatrix}, \tag{9}$$

where λ_i corresponds to the transition rate from state i to state j, $i, j \in \{\ell, h\}$.

Public learning processes are defined analogously replacing Y by ξ defined in (1).

□

The next result states that for the learning structures defined above, there is a unique one-dimensional diffusion $p := (p_t)_{t\geq 0}$ taking values in $\mathbb{R}$ that contains all the statistical information conveyed by ρ:

Lemma 2.4. (Law of motion of private beliefs). *Suppose that the long-run player's learning structure is either linear or nonlinear. Then, in each case, there exist a function $\mu : \mathbb{R} \to \mathbb{R}$, a time-dependent function $(\sigma_t)_{t\geq 0}$ and a $\mathbb{F}^Y$-Brownian motion $Z^Y := (Z_t^Y)_{t\geq 0}$ such that the diffusion*

$$dp_t = \mu(p_t)dt + \sigma_t dZ_t^Y, \; p_0 = p^o \tag{10}$$

takes values over the entire real line, and is a sufficient statistic for ρ. More specifically:

- (L) *When learning is linear, $p_t := \mathbb{E}[\theta_t|\mathcal{F}_t^Y]$, $t \geq 0$, $\mu(p) = -\kappa(p - \eta)$, $p \in \mathbb{R}$, $\sigma_t = \frac{\gamma_t}{\sigma_\xi}$, where $\gamma_t := \mathbb{E}[(\theta_t - p_t)^2|\mathcal{F}_t^Y]$, $t \geq 0$, and $Z_t^Y := \frac{1}{\sigma_\xi}\left(Y_s - \int_0^t p_s ds\right)$, $t \geq 0$. Moreover, the posterior variance process $\gamma := (\gamma_t)_{t\geq 0}$ solves the ODE $\dot{\gamma}_t = -2\kappa\gamma_t + \sigma_\theta^2 - \left(\frac{\gamma_t}{\sigma_\xi}\right)^2$, $t > 0$, $\gamma_0 = \gamma^0$. The pair (p^o, γ^o) is such that $\theta_0|\mathcal{F}_0 \sim \mathcal{N}(p^o, \gamma^o)$;*
- (NL) *When learning is nonlinear, we can take $p_t := \log\left(\frac{\mathbb{P}(\theta_t = h|\mathcal{F}_t^Y)}{1-\mathbb{P}(\theta_t = h|\mathcal{F}_t^Y)}\right)$, $t \geq 0$, $\sigma_t \equiv \frac{h-\ell}{\sigma_\xi}$ $\forall\, t \geq 0$, $\mu(p) = \lambda_1 \frac{e^p+1}{e^p} - \lambda_0(1 + e^p) - \frac{(h-\ell)^2}{2\sigma_\xi^2}\left(1 - \frac{2e^p}{1+e^p}\right)$, $p \in \mathbb{R}$, and $Z_t^Y := \frac{1}{\sigma_\xi}\left(Y_t - \int_0^t \frac{e^{p_s}}{1+e^{p_s}} ds\right)$, $t \geq 0$. The initial value p^o is such that $p^o = \log\left(\frac{\mathbb{P}(\theta_0 = h|\mathcal{F}_0^Y)}{1-\mathbb{P}(\theta_0 = h|\mathcal{F}_0^Y)}\right)$.*

In either case, the process Z^Y is called an innovation process.

Proof: See the Appendix.

□

Consider first the case of linear learning. Since the system is Gaussian, we have to keep track of the evolution of the posterior mean and the posterior variance only. They evolve according to

$$dp_t = -\kappa(p_t - \eta)dt + \frac{\gamma_t}{\sigma_\xi}\frac{dY_t - p_t dt}{\sigma_\xi}, \text{ and} \tag{11}$$

$$\dot{\gamma}_t = -2\kappa\gamma_t + \sigma_\theta^2 - \left(\frac{\gamma_t}{\sigma_\xi}\right)^2, \; t > 0, \tag{12}$$

respectively. Learning is said to be linear because p_t admits a solution that is linear in the partial observations $(Y_s : 0 \leq s \leq t)$, $t \geq 0$.

Observe that the posterior mean reverts toward η at the same rate $\kappa \geq 0$. Also, the response of the posterior mean to signal surprises (captured by the innovation process Z^Y) increases with the size of the mean-square error and decreases with the signal's volatility (σ_ξ). This implies that beliefs react more strongly in settings where either less information has been accumulated, or in settings where signals are more accurate. The posterior variance in turn evolves in a deterministic way, so its entire trajectory is perfectly anticipated at time zero. When $\gamma^o = \gamma^* := \sigma_\xi^2(\sqrt{\kappa^2 + \sigma_\theta^2/\sigma_\xi^2} - \kappa)$, the unique strictly positive stationary solution of (12), $\gamma_t \equiv \gamma^*$ for all $t \geq 0$, so the posterior mean follows a time-homogeneous diffusion. In what follows, I assume that this is the case, so $\sigma_t \equiv \sigma$ is a constant independent of time.[8]

Finally, when learning is nonlinear $\pi_t := \mathbb{P}(\theta_t = h|\mathcal{F}_t^Y)$, $t \geq 0$, follows the time-homogeneous diffusion

$$d\pi_t = (\lambda_1(1-\pi_t) - \lambda_0\pi_t)dt + \frac{(h-\ell)\pi_t(1-\pi_t)}{\sigma_\xi}\left(\frac{dY_t - \pi_t dt}{\sigma_\xi}\right), \; t > 0.$$

However, working with the log-likelihood ratio process $p := (p_t)_{t\geq 0}$, allows us to carry a single generic diffusion $dp_t = \mu(p_t)dt + \sigma_t dZ_t^Y$ regardless of the nature of the learning process.

2.2.2 The Belief-Asymmetry Process

The small players can only use the information conveyed by ξ to construct estimates about the current value of fundamentals. However, in order to form correct beliefs about the hidden state, they must undo the bias that the long-run player adds to the public signal. More specifically, a straightforward adaptation of Lemma 2.4 to the case of public beliefs yields that, for any feasible conjecture $a^* := (a_t^*)_{t\geq 0}$, there exists a process $Z^* := (Z_t^*)_{t\geq 0}$ that is a $\mathbb{F}^\xi-$Brownian motion from the small player's perspective such that the process $p^* := (p_t^*)_{t\geq 0}$ given by

$$dp_t^* = \mu(p_t^*)dt + \sigma dZ_t^*, \; p_0 = p^o. \tag{13}$$

is a sufficient statistic for the small players' learning process. Because the innovation process Z^* must capture unexpected realizations of the public signal given the small players limited information, Z^* takes the form

$$Z_t^* = \frac{1}{\sigma_\xi}\left(\xi_t - \int_0^t (a_s^* + \mathbb{E}^{a^*}[\theta_s|\mathcal{F}_s^\xi])ds\right), \; t \geq 0, \tag{14}$$

where $\mathbb{E}^{a^*}$ is the small players' subjective expectation operator under the assumption the the long-run player follows a^*. Intuitively, since at any point in time a fraction of the observed signal is attributable to signal manipulation, only the changes in the public signal that are unexplained by the manipulation strategy (namely, $d\xi_t - a_t^* dt$) convey relevant information about the underlying fundamental, $t \geq 0$.

[8] When learning is Gaussian and non-stationary incentives are characterized by a PDE. See Section 3 for an example based on Holmstrom (1999) in which this PDE admits a simple solution.

The long-run player can affect the small players' beliefs by controlling the evolution of the public signal. From his standpoint, the latter evolves according to

$$d\xi = (a_t + \mathbb{E}[\theta_t|\mathcal{F}_t^Y])dt + \sigma_\xi dZ_t^Y = (a_t + f(p_t))dt + \sigma_\xi dZ_t^Y,\ t \geq 0, \tag{15}$$

where Z^Y is his innovation process (Lemma 2.4), and $f : \mathbb{R} \to \mathbb{R}$ depends on the specific type of learning process.[9] Plugging (15) into (14), and then the resulting process into (13) yields

$$dp_t^* = [\mu(p_t^*) + \beta(a_t - a_t^*) + \beta(f(p_t) - f(p_t^*))]dt + \sigma dZ_t^Y,\ t \geq 0, \tag{16}$$

from the long-run player's standpoint, where $\beta = \sigma/\sigma_\xi$.[10]

Expression (16) shows how the long-run player can control the small players' belief. In particular, after a deviation takes place, the long-run player acquires private information about the evolution of the public signal ($f(p)$ term), which in turn gives him private information about the future evolution of public beliefs, and hence about his future payoffs.

Because we are interested in settings in which everyone shares the same uncertainty regarding the initial value of the underlying fundamental, it is natural to introduce a state variable that measures the wedge between private and public beliefs:

Proposition 2.5. *Suppose that the learning structure is unidimensional, and let (a, a^*) denote any feasible pair. Then, from the long-run player's perspective, the small players' belief process can be written as $p_t^* = p_t + \Delta_t$, $t \geq 0$, where the process $\Delta := (\Delta_t)_{t\geq 0}$ takes the form*

$$d\Delta_t = [-\phi(p_t, \Delta_t) + \beta(a_t - a_t^*)]dt, t > 0,\ \Delta_0 = \Delta^o, \tag{17}$$

with $\phi : \mathbb{R}^2 \to \mathbb{R}$ a function satisfying $\phi(p, 0) \equiv 0$ for all $p \in \mathbb{R}$. More specifically:

(L) *When learning is linear, $\phi(p,\Delta) = (\beta + \kappa)\Delta$ and $\beta = \sigma/\sigma_\xi = \gamma^*/\sigma_\xi^2$;*

(NL) *When learning is nonlinear, $\phi(p,\Delta) = -\lambda_1 \left[\frac{e^{p+\Delta}+1}{e^{p+\Delta}} - \frac{e^p+1}{e^p}\right] + \lambda_0[e^{p+\Delta} - e^p]$, and $\beta = \sigma/\sigma_\xi = (h - \ell)/\sigma_\xi^2$.*

Proof: Subtract (10) from (16) using the expressions for μ in Lemma 2.4, and the fact that $f(p) = p$ in the linear learning case and $f(p) = \frac{e^p}{e^p+1}$ when learning is nonlinear.

□

Suppose that both parties have the same prior about the initial value of the fundamental ($\Delta^o = 0$) and that the long-run player follows a^* up to time t. Because $\phi(p, 0) = 0$, beliefs will remain aligned up to instant. If at time t the long-run player decides to manipulate the signal above (below) the small players' expectations, he will become more (less) pessimistic than the small players about the current value of fundamentals.

[9] $f(p) = p$ in the linear learning case and $f(p) = \frac{e^p}{e^p+1}$ when learning is nonlinear.

[10] In the linear case, and as long as a_t^* takes the form $a^*(p_t^*)$ for some globally Lipschitz function $a^*(\cdot)$, the integrability condition in Definition 2.1 is enough to ensure the existence and uniqueness of a strong solution to (16), for any feasible strategy $a := (a_t)_{t\geq 0}$ (Theorem 1.3.15 in Pham (2009)). Hence, (16) is a well-defined object. See Sections 4 and 5 for more details.

Two final observations. First, because as time evolves past information becomes less of a good prediction of current fundamentals, belief discrepancies have an inherent tendency to disappear. More specifically, any stock of belief asymmetry resulting from a one-shot deviation from a^* at time t vanishes locally at a speed equal to $|\phi(p_t, \Delta_t)|$, $t \geq 0$. Second, observe that the small players' conjecture of equilibrium play acts as a threshold that the long-run player has to exceed in order to induce more belief asymmetry. Thus, in any equilibrium in which beliefs remain aligned, the small players construct beliefs using a strategy a^* such that, if the long-run player over-manipulates the public signal, the benefit from both boosting perceived fundamentals and having private information about it does not compensate the additional cost of effort; and if the long-run player under-manipulates, the loss in his payoffs outweighs both the savings from the under-manipulation, and the benefit from having private information about the hidden state.

2.3 Markov Perfect Equilibrium: Necessary Conditions

2.3.1 Markov Perfect Equilibrium

From the previous section, there exist a *private beliefs process* $p := (p_t)_{t\geq 0}$ and a *belief asymmetry process* $\Delta := (\Delta_t)_{t\geq 0}$, both taking values over the entire real line, and that fully summarize the long-run player's and the small players' beliefs $\rho := (\rho_t)_{t\geq 0}$ and $\rho^* := (\rho^*_t)_{t\geq 0}$. Their laws of motion take the form

$$\begin{aligned} dp_t &= \mu(p_t)dt + \sigma dZ_t^Y,\ p_0 = p^o, \text{ and} && (18)\\ d\Delta_t &= [-\phi(p_t, \Delta_t) + \beta(a_t - a^*_t)]dt,\ t > 0,\ \Delta_0 = \Delta^o, && (19)\end{aligned}$$

where μ, σ_p, ϕ and β are given in Lemma 2.4 and Propositions 2.5. While the long-run player's private belief evolves exogenously, he effectively controls the degree of belief asymmetry. In the sequel, I assume that the small players' best-response is of the form $\mathtt{b}(p^*, a^*)$ and observe that $\mathtt{b}(p^*_t, a^*_t) = \mathtt{b}(p_t + \Delta_t, a^*_t)$, $t \geq 0$.

I restrict the analysis to *pure* strategies, which implies that the small players will perfectly anticipate the long-run players' realized actions when starting from a common prior. Since $p = p^*$ is the unique payoff-relevant state variable on the equilibrium path, it is natural to study Markov Perfect Equilibria (MPE) in p^*.

Definition 2.6. *A measurable function $a^* : \mathbb{R} \to A$ is a MPE if there exists a tuple (a, b, p, p^*) consisting of a manipulation strategy of the long-run player $a := (a_t)_{t\geq 0}$, a public action profile of the small players $b := (b_t)_{t\geq 0}$, a private belief process $p := (p_t)_{t\geq 0}$, and a public belief process $p^* := (p^*_t)_{t\geq 0}$ such that*

(i) *Given any feasible strategy $\breve{a} := (\breve{a}_s)_{s\geq 0}$ and any private history $(\xi_s : s \in [0,t], \breve{a}_s : s \in [0,t])$ that leads to $p_t = p^*_t$, the long-run player's optimal action at time t is of the form $a_t = a^*(p^*_t)$, $t \geq 0$;*

(ii) *After all public histories, the small player's best response function is of the form $b_t = \mathtt{b}(p^*_t, a^*(p^*_t))$, $t \geq 0$,*

(iii) *The pair $(a_t, a^*(p_t + \Delta_t), 0))_{t\geq 0}$ is feasible and it induces a unique solution Δ to (17) defined over the entire real line almost surely,* *and (a, b, p, p^*) satisfies (a)-(d) in Definition 2.2.*

Observe that in the previous definition there are no markovian restrictions on the long-run player's behavior off the equilibrium path. However, it is natural to conjecture that the long-lived player's optimal actions will depend only on the current value that (p, Δ) takes. This is because, once the small players construct beliefs using a markovian conjecture, the long-run player's optimization problem becomes one of stochastic control subject to the dynamics of p and Δ.

2.3.2 Necessary Conditions

Suppose that the small players conjecture that the long-run player's equilibrium actions take the form $(a^*(p_t^*))_{t\geq 0}$ for some measurable function $a^* : \mathbb{R} \to A$. Given this *markovian conjecture* the long-run player's problem consists of choosing a manipulation strategy $a := (a_t(\xi_s : s \in [0, t], p_t))_{t\geq 0}$, such that after all private histories the continuation strategy $(a_s)_{s\geq t}$ maximizes his expected discounted utility

$$\mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} (u(\mathtt{b}(p_s + \Delta_s, a^*(p_s + \Delta_s))) - g(a_s)) ds \right]$$

subject to the dynamics (18) and (19) with initial values $(p_t, \Delta_t) = (\overline{p}, \overline{\Delta})$, for all $(\overline{p}, \overline{\Delta}) \in \mathbb{R}^2$ and $t \geq 0$. I redefine $u(\mathtt{b}(\cdot, \cdot))$ to be $u(\cdot, \cdot)$ unless otherwise stated.

Let $V^{a^*}(\overline{p}, \overline{\Delta})$ denote the long-run player's value function associated with the previous problem, which takes a^* as an input. Then, if an optimal control exists and the value function is smooth enough, V^{a^*} satisfies the Hamilton-Jacobi-Bellman (HJB) equation[11]

$$\begin{aligned} rV^{a^*}(p, \Delta) &= \sup_{a\in A} \Big\{ u(p + \Delta, a^*(p + \Delta)) - g(a) + \mu(p) V_p^{a^*}(p, \Delta) + \frac{1}{2}\sigma^2 V_{pp}^{a^*}(p, \Delta) \\ &\quad + [-\phi(p, \Delta) + \beta(a - a^*(p + \Delta))] V_\Delta^{a^*}(p, \Delta) \Big\}, \ (p, \Delta) \in \mathbb{R}^2. \quad (20) \end{aligned}$$

Consequently, if an optimal control $\alpha := (\alpha_t)_{t\geq 0}$ exists, it has to satisfy

$$\alpha_t = \arg\max_{a\in A} \{a\beta V_\Delta^{a^*}(p_t, \Delta_t) - g(a)\}, \ t \geq 0. \quad (21)$$

Hence, if a MPE $a^* : \mathbb{R} \to A$ exists, the following *equilibrium condition* must hold:

$$a^*(p) = \arg\max_{a\in A} \{a\beta V_\Delta^{a^*}(p, 0) - g(a)\}, \ p \in \mathbb{R}. \quad (22)$$

The following results summarizes the discussion so far:

Proposition 2.7. (Global incentives). *Assume that a MPE $a^* : \mathbb{R} \to A$ exists and the associated value function V^{a^*} is of class $C^{2,1}(\mathbb{R}^2)$. Then, the long-run player's value function $V^{a^*}(p, \Delta)$ satisfies the partial differential equation (PDE)*

$$\begin{aligned} rV(p, \Delta) &= \sup_{a\in A} \Big\{ u(p + \Delta, a^*(p + \Delta)) - g(a) + \mu(p) V_p(p, \Delta) + \frac{1}{2}\sigma^2 V_{pp}(p, \Delta) \\ &\quad + [-\phi(p, \Delta) + \beta(a - a^*(p + \Delta))] V_\Delta(p, \Delta) \Big\}, \ (p, \Delta) \in \mathbb{R}^2 \quad (23) \end{aligned}$$

[11]Subscripts p and Δ denote partial derivatives with respect to p and Δ, respectively.

$$s.t. \qquad a^*(p) \in \arg\max_{a\in A} \{\beta V_\Delta(p,0)a - g(a)\}. \tag{24}$$

□

Observe that the PDE (23)-(24) is non-standard. In addition to having the fixed point condition (24), this PDE is *nonlocal*, as the local behavior of the long-run player's continuation value around a point (p,Δ) depends on the value attached to a marginal deviation off the equilibrium at the point $(p+\Delta,0)$. This is clearly seen in regions where on-path incentives are interior, so the small players use

$$a^*(p+\Delta) = (g')^{-1}(\beta V_\Delta(p+\Delta,0)).$$

as their conjecture about equilibrium play. This type of non-localness, and consequently, this type of PDE, seems to be new.

The PDE (23)-(24) is typically hard to visualize. However, it is possible to extract properties of equilibrium behavior without the need of fully solving such equation. More specifically, observe that as long as the long-run player's actions are interior and beliefs are aligned, the long-run player's incentives are driven by $V_\Delta(p,0)$. The next result derives a necessary condition for the value attached to inducing a small belief discrepancy in any equilibrium in which beliefs remain aligned:

Theorem 2.8. (**On path behavior**) *Assume that a MPE $a^*:\mathbb{R}\to\mathbb{R}$ exists and that the associated value function V^{a^*} is of class $C^{2,1}(\mathbb{R}^2)$. Furthermore, suppose that at a level of public beliefs p*

(i) *on path incentives are interior and*

(ii) *on path incentives are locally twice continuously differentiable with respect to public beliefs, i.e., there exists a neighborhood $\mathcal{O}$ of p such that $a^*(\cdot)\in C^2(\mathcal{O})$.*

Then, $g'(a^*(\cdot)) = \beta V_\Delta^{a^*}(\cdot,0)$, where $V_\Delta^{a^*}(\cdot,0)$ satisfies the ODE in $p\mapsto V_\Delta(p,0)$

$$\begin{aligned}\tilde{r}(p,V_\Delta(p,0))V_\Delta(p,0) \;&=\; u_p(p,g'^{-1}(\beta V_\Delta(p,0))) + u_a(p,g'^{-1}(\beta V_\Delta(p,0)))\frac{d}{dp}g'^{-1}(\beta V_\Delta(p,0)) \\ &\quad + V_{\Delta p}(p,0)\mu(p) + \frac{1}{2}\sigma^2 V_{\Delta pp}(p,0),\; p\in\mathcal{O},\end{aligned} \tag{25}$$

with $\tilde{r}(p,V_\Delta(p,0)) := r + \phi_\Delta(p,0) + \beta\frac{d}{dp}g'^{-1}(\beta V_\Delta(p,0))$.

Proof: Differentiate the PDE (23)-(24) with respect to Δ and evaluate at $\Delta=0$.

□

Equation (25) is a recursive expression for how the payoff from a hypothetical *local* deviation off the equilibrium path varies across different levels of public beliefs, when the small players are perfectly anticipating the long-run player's actions. I refer to (25) as the *incentives equation*. Because when a MPE exists the corresponding mapping $p\mapsto V_\Delta(p,0)$ will satisfy this ODE, the incentives equation reduces the problem of finding MPE from solving a non-standard PDE, to studying the solutions of a nonlinear ODE.

The incentives equation optimally balances 3 forces – marginal flow payoffs, cost smoothing and ratcheting– in order to generate an optimal inter-temporal pattern of manipulation. More specifically, the long-run player must be indifferent between exerting the last unit of signal manipulation "today" (the right-hand side in (25)) and delaying it to "tomorrow". Exerting signal manipulation today is beneficial because it immediately affects the long-run player's flow payoffs (first two terms on the right hand side of (25)), and because it allows him to smooth out the costs of signal manipulation over time (term $p \mapsto \mu(p)V_{p\Delta}(p,0) + \frac{1}{2}\sigma_p^2 V_{pp\Delta}(p,0)$).[12] In particular, if the benefits from signal manipulation are expected to change at high rates in the near future, then, because manipulation costs are convex, it is optimal to start investing in belief distortion today.

Postponing manipulation to an instant later allows the long-run player to save the costs associated with his investments in belief asymmetry depreciating over time. These depreciation costs are captured in the *rate of return* on belief asymmetry

$$\tilde{r}(p, V_\Delta(p,0)) = r + \phi_\Delta(p,0) + \beta \frac{d}{dp} g'^{-1}\left(\beta V_\Delta(p,0)\right),\ p \in \mathbb{R},$$

in the left-hand side of (25). Given any fixed discount rate $r > 0$, the higher $\tilde{r}(p, V_\Delta(p,0))$ is, the lower the return from manipulating the public belief.

The rate of return $\tilde{r}$ is determined endogenously in equilibrium. Recall that given any markovian conjecture $\mathtt{a}(\cdot, 0)$, the belief-asymmetry process (19) evolves according to

$$d\Delta_t = [-\phi(p_t, \Delta_t) + \beta(a_t - a^*(p_t + \Delta_t))]dt,\ t \geq 0.$$

Thus, in any MPE in which beliefs remain aligned, a local deviation off the equilibrium path generates a flow dividend that depreciates (locally) at a rate equal to

$$\begin{aligned} \frac{\partial}{\partial \Delta}\left(\phi(p,\Delta) + \beta a^*(p+\Delta)\right)\Big|_{\Delta=0} &= \phi_\Delta(p,0) + \beta \frac{\partial}{\partial p} a^*(p) \\ &= \phi_\Delta(p,0) + \beta \frac{\partial}{\partial p} g'^{-1}(\beta V_\Delta(p,0)). \end{aligned}$$

The first term $\phi_\Delta(p,0)$ corresponds to the rate at which belief asymmetry inherently decays over time as information accumulates. The second term captures the returns from affecting the standard of manipulation the long-run agent will face in the near future (ratcheting). In fact, since meeting a high standard is costly for the long-run agent, if the second term is large, the total return for manipulating public beliefs will be low.[13]

---

[12] The expected rate of change of the marginal value of belief asymmetry satisfies $\lim_{h\to 0} \frac{\mathbb{E}_t[V_\Delta(p_{t+h},\Delta_{t+h})] - V_\Delta(p_t,\Delta_t)}{h}\Big|_{\Delta_t=0} = \underbrace{[\phi(p_t,0) + \beta(\mathtt{a}(p_t,0) - \mathtt{a}(p_t,0))]}_{\equiv 0} V_{\Delta\Delta}(p_t,0) + \mu(p)V_{p\Delta}(p_t,0) + \frac{1}{2}\sigma^2 V_{pp\Delta}(p_t,0) = \mu(p)V_{p\Delta}(p_t,0) + \frac{1}{2}\sigma^2 V_{pp\Delta}(p_t,0)$. Hence, this effect connects manipulation choices across time. Bonatti and Horner (2011) also find that effort decisions are inter-temporally connected, although in a model of career concerns in which information is coarse.

[13] When time is discrete, ratchet forces appear two periods ahead, as effort today impacts the public belief tomorrow, and the latter determines the next period conjecture of equilibrium play. Martinez (2009) finds this force in a career concerns model of job assignment with piecewise linear wages. In continuos time, this effect is contemporaneous to current effort choices, through the slope of the small players' conjecture of equilibrium play, $\frac{\partial \mathtt{a}}{\partial p}(p,0)$.

The incentives equation is a powerful object. In the next section I develop applications that exploit the versatility of this equation as a tool for computing Markov perfect equilibria. The reader interested in the (i) existence of solutions to the incentives equation and/or in (ii) sufficient conditions that ensure that a solution to the incentives equation is in fact a MPE can jump to Section 4.

3 The Incentives Equation: Applications

This section presents applications that use, and sometimes extend, the insights and methods previously presented. The analysis starts with the structure of incentives in linear environments.

3.1 Reputation in Labor Markets: Career Concerns

Recall the classic model of reputation in labor markets introduced by Holmstrom (1999). A pool of competitive firms compete for the labor of a risk-neutral worker of unknown ability. In his formulation, the skills of the worker are Gaussian, either fixed over time or evolving as a random walk. More generally, in continuous-time, ability θ can be modeled as a mean-reverting process

$$d\theta_t = -\kappa(\theta_t - \eta)dt + \sigma_\theta dZ_t^\theta,\ t \geq 0,$$

which encompasses the other two specifications. Output $\xi := (\xi_t)_{t\geq 0}$ evolves according to

$$d\xi = (a_t + \theta_t)dt + \sigma_\xi dZ_t^\xi,\ t \geq 0,$$

where a_t denotes effort at $t \geq 0$, and Z^ξ shocks to output beyond the worker control.

Since the pool employers is competitive and no output-contingent contracts can be written, the flow wage that the worker receives must corresponds to the expected output flow from the market's perspective. Using our notation, the market's action takes the form

$$\mathtt{b}(a_t^*, p_t^*) = \frac{\mathbb{E}_t^{a^*}[d\xi_t]}{dt} = a_t^* + p_t^*,$$

where a_t^* is the agent's equilibrium effort decision at time t, and $p_t^* = \mathbb{E}^{a^*}[\theta_t|\mathcal{F}_t^\xi]$, $t \geq 0$.

Away from the steady state level of learning γ^* the posterior variance evolves deterministically, and the incentives equation becomes a partial differential equation on $V_\Delta(p, 0, t)$. It is easy to see that it takes the form:

$$\begin{aligned} [r + \delta_t]\, V_\Delta(p, 0, t) &= 1 + \frac{d}{dp} g'^{-1}(\beta_t V_\Delta(p, 0, t)) \\ &\quad + V_{\Delta p}(p, 0, t)[-\kappa(p - \eta)] + \frac{(\beta_t \sigma_\xi)^2}{2} V_{\Delta pp}(p, 0, t) \\ &\quad + V_{\Delta t}(p, 0, t),\ (p, t) \in \mathbb{R} \times \mathbb{R}_+ \end{aligned} \tag{26}$$

where $\beta_t := \gamma_t/\sigma_\xi^2$, $\delta_t := \beta_t + \kappa + \beta_t g'^{-1}(\beta_t V_\Delta(p, 0, t))$, and the new term $V_{\Delta t}(p, 0, t)$ captures how the value attached to inducing belief asymmetry varies as information accumulates over time. A guess of the form $V_\Delta(p,0,t)=\alpha(t)$, reduces the previous PDE to an standard time-dependent ODE

$$[r+\beta_t+\kappa]\alpha(t)=1+\dot{\alpha}(t),\ t\geq 0.$$

The solution to this ODE is $\alpha(t)=\int_t^\infty e^{-\int_t^s(r+\kappa+\beta_u)du}ds$, which implies that effort takes the form

$$g'(a_t^*)=\beta_t\int_t^\infty e^{-\int_t^s(r+\kappa+\beta_u)du}ds,\ t\geq 0.$$

Consequently, there is an equilibrium in which effort depends only on the worker's tenure, and not on his reputation.[14]

The intuition for the result is as follows. Observe that since ξ is additively separable in effort and skills, the impact of the worker's output is independent of the worker's true ability. Also, since learning is Gaussian, the posterior belief p_t^* is a linear function past output observations. These two observations imply that the impact of the worker's effort on all future levels of public beliefs is independent of the worker's reputation. But since the wage profile is also a linear function of reputation p^*, the impact of the worker's effort on *all future levels of wages* will be independent the worker's reputation as well. Finally, because the extent to which public beliefs respond to new information is captured by the sensitivity process β, the worker's incentives will be time-dependent only.

The linearity of Holmstrom's set up generates particularly simple incentives for belief manipulation. In particular, since this equilibrium depends on calendar time only, the worker will find it optimal to exert effort according to the market's current conjecture of equilibrium play, *even if he had deviated in the past.* Hence, the possibility of acquiring private information about his type has no value for the worker in the model. When all this linearity is relaxed, the incentives for belief manipulation will vary across the state space. Moreover, since beliefs have persistence, these incentives will be interconnected across time as opposed to what occurs in the linear benchmark. Finally, off the equilibrium path, the long-run player will find it optimal to exploit his private information. But when an equilibrium exists, any such deviation must be unprofitable.

3.2 Manipulation of Official Statistics

In this application I analyze a government's incentives to manipulate an official inflation statistic. A basic requirement for this type of manipulation is the lack of independence of a country's monetary authority (or the institution in charge of the creation of such statistic). When this is the case, incentives to misreport aggregate price data are particularly attractive in high-inflation economies, as governments may lower inflation statistics in an attempt to reduce or *anchor* inflation expectations.[15] The purpose of this model is to understand the size of the inefficiencies created by this type of manipulation, and explain some observed dynamics.

[14]Cisternas (2012) determines conditions under which a deterministic equilibrium will exist in environments that allow for endogenous human capital accumulation. In the particular case of exogenous skills, it can be shown that the above effort profile is indeed an equilibrium through pointwise maximization.

[15]Evidence of this type of manipulation has been documented by Cavallo (2012) for the case of Argentina.

Suppose that the aggregate price index of an economy is given by $\exp(Y_t),\ t \geq 0$, where $Y := (Y_t)_{t\geq 0}$ – the economy's *realized inflation* – evolves according to

$$dY_t = \theta_t dt + \sigma_\xi dZ_t,\ t \geq 0,$$

with $Z := (Z_t)_{t\geq 0}$ a Brownian motion and $\theta := (\theta_t)_{t\geq 0}$ an unobserved inflation component. A first interpretation is that θ represents the component of inflation that is affected by monetary policy, whereas Z captures shocks beyond the central bank's control. A second interpretation is that θ is the actual inflation rate of the economy, and Z captures measurement errors (lags in data collection, how representative the basket of goods used to construct the statistic is, etc.). In any case, in order to make good economic decisions the population must have a good assessment of the future evolution of Y, which reduces to having an accurate estimate of the current value of θ. I assume that θ is given by

$$d\theta_t = -\kappa(\theta_t - \eta)dt + \sigma_\theta dZ_t^\theta,\ t \geq 0,$$

where $\eta > 0$ corresponds to the long-run average mean of unobserved inflation. In this specification, θ fluctuates around η with high probability.

The government (long-run player) can *manipulate* the data Y gathered by the central bank, before releasing it to the public. Suppose that the public signal about realized inflation is given $(\xi_t)_{t\geq 0}$, defined as $d\xi_t = dY_t + a_t dt$, where $(a_t)_{t\geq 0}$ is the government's manipulation strategy. Thus, *official inflation* follows

$$d\xi_t = (a_t + \theta_t)dt + \sigma_\xi dZ_t,\ t \geq 0.$$

The government has preferences given by

$$\mathbb{E}\left[\int_0^\infty e^{-rt}\left(-(p_t^*)^2 - \frac{\psi}{2}a_t^2\right)dt\right],$$

where $p^* := (p_t^*)_{t\geq 0}$ corresponds to the public belief about the unobserved inflation component θ. Thus, the government would like to induce people to believe that the *perceived inflation trend* p^* is as close as possible to a target which has been normalized to zero: having some degree of inflation is beneficial, as it signals economic growth, but high inflation is harmful for well-known reasons. More interestingly, since the target is below η, the government's goal is to drive perceived inflation below its natural rate of growth η.

Observe that in this specification manipulation is costly for the government even if the population anticipate the government's actions, and thus hold correct expectations about the future evolution of prices. For instance, in economies in which variables are indexed to inflation (e.g. wages, pensions, bonds or real estate) different degrees of manipulation will incentivize agents to adjust prices periodically, thus incurring in menu costs. Or when workers and firms decide on nominal salaries, official statistics can be used strategically by the party with the largest bargaining power, thus increasing negotiation costs and leading to ex-post inefficiencies. Similarly, investment can be severely affected if this lack of commitment is expected to spread to regulatory agencies. Hence, the cost term $\psi\frac{a^2}{2}$ represents the weight that the government attaches to introducing all these inefficiencies in the economy.[16] For simplicity, I assume that the size of this inefficiencies do not depend

[16] I exclude from the analysis information costs – a type of cost that naturally arises in this type of environments – as information is public in this paper.

on the actual level of realized inflation, but only on the magnitude of the size of the manipulation.[17]

The methods presented in the previous section allow us to quantify the size of the inefficiencies created by the government. For quadratic preferences (see Section 6), the incentives equation has solution that is linear in p^*, so the degree of belief manipulation is given by

$$a^*(p^*) = \frac{1}{\psi} V_\Delta(p^*, 0) = \frac{1}{\psi}[\alpha_0 + \alpha_1 p^*],$$

with $\alpha_1 < 0$, and $\alpha_0 \leq 0$ with equality if and only if $\kappa\eta = 0$.[18] Unlike the linear case, actions are now stochastic and depend on the stochastic history of the game. Graphically:

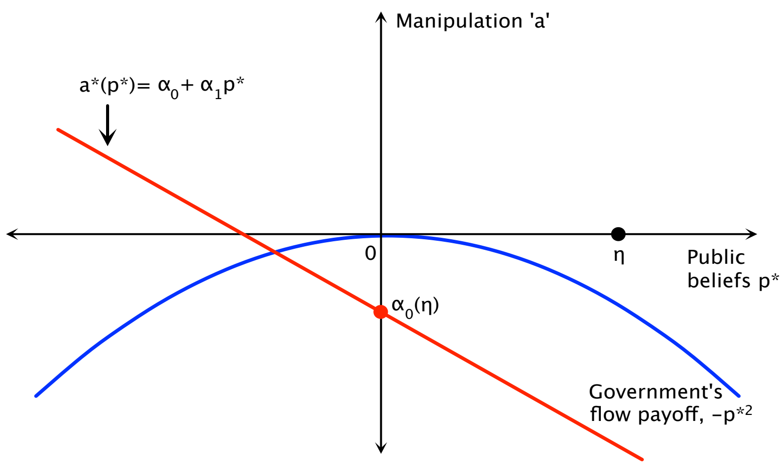

Figure 1: The equilibrium manipulation choices as a function of perceived inflation p^*.

Cavallo (2012) compares official statistics of inflation in Argentina (ξ) versus a price index that he constructs using online data from a large supermarket (a proxy for Y) over a period of four years. He shows that the gap between these two measures (captured by a^*) is always negative (official inflation is always below his index) and its magnitude fluctuates around two and three times the online statistic. This is consistent in which a model in which unobserved inflation fluctuates around a mean η with high probability, and this mean is larger than the government's target. More interestingly, he documents that manipulation of a similar magnitude and in the *same direction* also took place in 2009, year in which the country was going through a recession. Such behavior is in fact predicted by the model: from the previous picture, a government will have the incentive to reduce perceived inflation even when the latter is at, or sometimes below, its target.

[17]An extra term $-(\Delta)^2$ capturing the efficiency costs of people having the wrong (off-equilibrium) belief about inflation can be easily incorporated, but it delivers no additional economic insights.

[18]In Section 4 I derive sufficient conditions on the primitives of signal-jamming games with *linear-quadratic* structure (linear learning, quadratic payoffs) that ensure the existence of a linear equilibrium. This example satisfies those conditions.

The reason why a government may have the incentive to further reduce a price index during a recession is due of cost smoothing. Observe that in the absence of both actual and conjectured manipulation, perceived inflation p^* mean reverts toward $\eta > 0$. Consequently, low levels of inflation are likely to be short lived, and incentives for manipulation will be strong in the near future as perceived inflation goes back to η. But local predictability of public beliefs ($\kappa > 0$) and convex costs of adjustment imply that the government, in anticipation of future price manipulation, will have the incentive to manipulate prices when the economy is in a downturn ($p^* \leq 0$), as doing so effectively minimizes the total costs imposed on society. Once people recognize this possibility, they will construct beliefs using a manipulation policy that allows for manipulation during downturns ($\alpha_0 < 0$). In equilibrium, the government will become trapped into people's expectations of manipulation.

Finally, notice that manipulation is purely wasteful: in equilibrium the government does not control the evolution of beliefs (they evolve exogenously) and yet it still introduces inefficiencies. The government would like to *commit* to not manipulating prices, but such commitment is not credible.

3.3 Monetary Policy and Unobserved Inflation

In the Online Appendix I extend the previous methods to study games of investment and learning. In these games, the long-run player affects the evolution of the fundamental itself, rather than the public signal that conveys information about it. Although these are not exactly games of belief manipulation, the long-run player's incentives can still benefit from inducing belief asymmetry. In this section I study a simple model of monetary policy in the presence of unobserved components of inflation, and show how the inflationary bias induced by a central bank who lacks commitment depends on the characteristics of the market's learning process.

The price index of a two-sector economy takes the form $\exp(Y_t^1 + Y_t^2),\ t \geq 0$, where $Y^i := (Y_t^i)_{t\geq 0}$ $i = 1, 2$, evolve according to

$$\begin{aligned} dY_t^1 &= a_t dt + \sigma_1 dZ_t^1 \\ dY_t^2 &= \theta_t dt + \sigma_2 dZ_t^2,\ t \geq 0. \end{aligned} \tag{27}$$

In this specification, a_t corresponds to the rate of money growth in the economy, and $\theta := (\theta_t)_{t\geq 0}$ is an unobserved inflation trend. The Brownian motions $Z^i := (Z_t^i)_{t\geq 0}$, $i = 1, 2$, represent sectorial shocks beyond the control of the central bank (the long-run player), and are independent from each other. All the agents in the economy observe the pair (Y^1, Y^2) and hence they can learn about θ from the observation of Y^2. I refer to θ as the economy's *core inflation.*[19]

Observe that in sector 1 prices are fully flexible and money has permanent effects on the price level. In contrast, I assume that prices in sector 2 are partially rigid. More specifically, I model θ as a mean-reverting process

$$d\theta_t = (a_t - \kappa\theta_t)dt + \sigma_\theta dZ_t^\theta,\ t \geq 0,$$

[19]Measures of core inflation typically exclude volatile goods such as food and energy. Consequently, I am implicitly assuming that the volatility of shocks to prices in sector 1 is large relative to the volatility of shocks in sector 2.

reflecting the idea that any change in the stock of money is slowly internalized by core inflation. The Brownian motion $Z^\theta := (Z^\theta_t)_{t\geq 0}$ captures the policymaker's limited ability to control fundamental inflation. The volatility term σ_θ captures the degree of control that the central bank has over core inflation.

The central bank cares about long-run price stability, and thus controls the evolution of core inflation θ using money growth as an instrument. This comes at the expense of permanent inflationary effects in sector 1, which also hurt the economy. Also, because money growth is imperfectly observed, the policymaker can also exploit this informational advantage to generate inflation surprises and thus reduce unemployment (Kydland and Prescott (1977)).[20] All these features are summarized in the central bank's payoffs

$$\mathbb{E}\left[\int_0^\infty e^{-rt}\left(-k_2(\theta_t-\overline{\theta})^2dt-\frac{\psi}{2}a_t^2dt+k_1(dY_t^2-p_t^*dt)\right)\right].$$

The first term is the policymaker's loss from core inflation having deviated away from a target $\overline{\theta}$. The second term captures the loss from inducing non-zero inflation in sector 1 (whose corresponding target has been normalized to zero). Finally, the the third term is a Phillips curve capturing the costs and benefit of inflation surprises: if realized inflation in sector 2 is above market expectations, $p_t^* := \mathbb{E}^*[\theta_t|\mathcal{F}_t^{Y^2}]$, employment will increase.[21]

Suppose as a benchmark that θ is observable. Since in this case the central bank cannot surprise the economy, the monetary authority's problem becomes

$$\begin{aligned}&\max_{a\in\mathcal{A}}\ \ \mathbb{E}\left[\int_0^\infty e^{-rt}\left(-k_2(\theta_t-\overline{\theta})^2-\frac{\psi}{2}a_t^2\right)dt\right]\\ &s.t.\quad d\theta_t=(a_t-\kappa\theta_t)dt+\sigma_\theta dZ_t^\theta.\end{aligned}$$

This results in the following

Proposition 3.1. *If core inflation is observed, the central bank's optimal policy is of the form* $a^o(\theta)=\frac{\alpha_1^o+2\alpha_2^o\theta}{\psi}$, *where*

$$\alpha_1^o=\frac{-2k_2\overline{\theta}}{\frac{2\alpha_2^o}{\psi}-(r+\kappa)}>0\quad and\quad \alpha_2^o=\frac{\psi}{2}\left[(r+2\kappa)-\sqrt{(r+2\kappa)^2+\frac{8k_2}{\psi}}\right]<0.$$

Also, $a^o(\overline{\theta})=\frac{\alpha_2^o\kappa}{2\alpha_2^o-r\psi}>0$ *if* $\kappa>0$. *Consequently, the monetary policy rule is decreasing in the level of core inflation (countercyclical).*

Proof: See the Appendix.

□

[20] Imperfect monitoring of money growth has been assumed extensively in the literature. See for example Cukierman and Meltzer (1986).

[21] For our purposes, defining a Philips that excludes sector 1 is without loss of generality. If we had instead defined it as $d\xi_t-(a_t^*+p_t^*)dt$, the central bank's preferences would remain linear in Δ (recall that $\mathbb{E}_t[d\xi_t-(a_t^*+p_t^*)dt]=[a_t-a_t^*+\Delta_t]dt)$. Consequently, such model would also have an equilibrium in which the value attached to inducing belief asymmetry is constant across different levels of private beliefs.

If θ is hidden, but the central bank has commitment power, then $p \equiv p^*$. Since no economic stimulus is possible, the central bank's problem becomes

$$\max_{a\in\mathcal{A}} \ \mathbb{E}\left[\int_0^\infty e^{-rt}\left(-k_2(p_t-\overline{\theta})^2-\frac{\psi}{2}a_t^2\right)dt\right]$$
$$s.t. \quad dp_t = (a_t-\kappa p_t)dt+\beta\sigma_2 d\overline{Z}_t^0,$$

where $\overline{Z}^0$ is a Brownian motion from the central bank's perspective. Observe that this problem has the same structure as the one just solved in Proposition 3.1, except for the volatility term $\beta\sigma_2 \neq \sigma_\theta$. Since the volatility term does not affect the form of the policy rule found in Proposition 3.1, such rule is also optimal in this case.[22]

3.3.1 Optimal Policy: No Commitment

When the central bank lacks commitment, the relevant state variables correspond to the policymaker's private belief about core inflation p, and the belief-asymmetry process $\Delta := p^* - p$. If learning is stationary, their dynamics are given by

$$\begin{aligned} dp_t &= (a_t-\kappa p_t)dt+\beta\sigma_2 d\overline{Z}_t^0 \\ d\Delta_t &= [-(\beta+\kappa)\Delta_t+(a_t^*-a_t)]dt \end{aligned}$$

with $\beta := \gamma^*/\sigma_2^2 = \sqrt{\kappa^2+(\sigma_\theta/\sigma_2)^2}-\kappa$. Observe that when the long-run player affects the evolution of the fundamental the effect of both a and a^* on Δ has the exact opposite sign than the case in which he is affecting the evolution of the public signal instead. This is because by $a > a^*$ leads him to be more optimistic about the current level of θ.

The central bank's ex-ante flow payoffs take the form $-k_1\Delta_t - k_2(p_t-\overline{\theta})^2-\frac{\psi}{2}a^2$, $t\geq 0$, (up to an additive constant). Observe that from the central bank's perspective, the gains from reducing unemployment are linear in Δ. Consequently, we would expect to find an equilibrium in which the *inflationary bias* – the gap between the full commitment rule and the time-consistent one under no commitment – is *uniform across all levels of public beliefs*. This is confirmed in the next proposition:

Theorem 3.2. *When θ is unobserved, the central bank's optimal policy takes the form $\alpha^*(p) = \frac{\alpha_1^o+2\alpha_2^o p}{\psi}+\frac{\alpha_3^*}{\psi}$ with α_1^o and α_2^o as in the observable case, and*

$$\alpha_3^* = \frac{k_1}{r+\sqrt{\kappa^2+(\sigma_\theta/\sigma_2)^2}-\frac{2\alpha_2^o}{\psi}} > 0.$$

Proof: See the Appendix.

□

The size of the shift is given by the term

$$-V_\Delta(p,0) = \alpha_3^* = -\frac{k_1}{r+\beta+\kappa-\frac{2\alpha_2^o}{\psi}} = \frac{k_1}{r+\sqrt{\kappa^2+(\sigma_\theta/\sigma_2)^2}-\frac{2\alpha_2^o}{\psi}} > 0,$$

[22] Because the payoff function is quadratic, the volatility term only affects the *level* of the central bank's value function.

which is increasing in k_1 and decreasing in r, ψ, κ and σ_θ/σ_2. In particular, the incentives of the central bank to surprise the economy with inflation are high when the impact on employment is prolonged. The latter effect is captured by the rate at which belief asymmetry decays over time, $\beta + \kappa = \sqrt{\kappa^2 + (\sigma_\theta/\sigma_2)^2}$. The higher this rate, the shorter the effect on employment. When monetary policy has loose control (σ_θ large), or prices Y_2 are very informative (low σ_2), or prices in sector 2 assimilate money growth faster (high κ), people will discount past information more heavily (beliefs will have lower persistence), and this will generate more commitment.

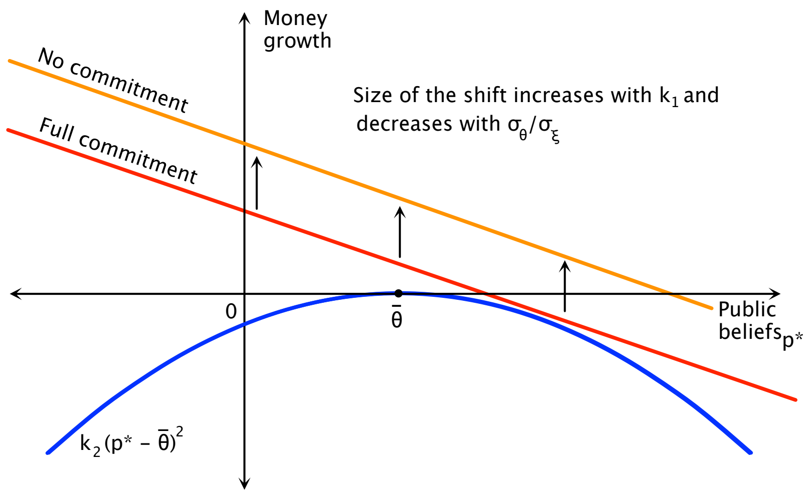

Figure 2: Non-commitment rule and the inflationary bias.

In Cukierman and Meltzer (1986) a central bank has private information about its (exogenous) current preference for stimulating the economy over controlling inflation, and prices are a noisy signal of the central bank's current target of money growth (imperfect control). The central bank thus chooses a rate of money growth based on its private information, taking into account that the economy learns about its preferences through the observation of realized prices. An important message in that paper the lower the degree of control over its objective that the central bank has, the higher the incentives to surprise the economy with high inflation. In contrast, the current example shows that having low control over an unobserved component of inflation (as measured by high levels of σ_θ), induce a policymaker to exhibit more commitment. The example presented in this section thus highlights the importance of understanding the nature of the shocks that limit a central bank's ability to achieve its goals, as loose control over unobserved variables, as opposed to observed ones, generate opposite incentives.

3.4 Nonlinear Learning: Work-Shirk Equilibria

One of the most relevant features of the incentives equation is that it can shed lights on the shape of equilibrium behavior without the need to fully solve the complex partial differential equations that characterize the long-run player's value function. This is particularly important in environments that exhibit high nonlinearities, as I illustrate below.

Suppose that the long-run player is a worker whose ability θ is drawn from a discrete random variable taking values in $\{0,1\}$. The worker dislikes effort according to the function $g(a) = \frac{a^2}{2}$, and the flow payoff from the market's actions is given by $u(p^*, a^*(p^*))$.

Let $p_t^* := \mathbb{P}^*(\theta_t = 1|\mathcal{F}_t^\xi)$ denote the public belief that the worker's ability is high given the information up to time t, and p_t the corresponding private belief, $t \geq 0$. In this case it can be easily checked that $p^* = \frac{p\Delta}{1+p(\Delta-1)}$, with (p,Δ) evolving as

$$dp_t = \frac{p_t(1-p_t)}{\sigma_\xi}dZ_t^Y \quad \text{and} \quad d\Delta_t = \frac{\Delta_t(a_t - a_t^*)}{\sigma_\xi^2}dt, \; t \geq 0, \tag{28}$$

and $\Delta \equiv 1$ capturing the region in which beliefs are aligned. Moreover, the incentives equation takes the form

$$\left[r + \frac{V_{\Delta p}(p,0)}{\sigma_\xi^4}\right] V_\Delta(p,0) = p(1-p)\left[h'(p) + \frac{V_{\Delta p}(p,0)}{\sigma_\xi^2} + \frac{p(1-p)}{2\sigma_\xi^2}V_{\Delta pp}(p,0)\right], \tag{29}$$

where $p(1-p)h'(p) = \left[u_p + u_a \frac{da^*}{dp}\right]_{p^*=p} = u_p(p, V_\Delta(p,0)) + u_a(p, V_\Delta(p,0))V_{\Delta p}(p,0)$.

Using results from ordinary differential equations, it can be shown that there exists a non-negative solution of class C^2 to the boundary value problem defined by the above ODE and the boundary conditions $V_\Delta(0,0) = V_\Delta(1,0) = 0$.[23] One would expect equilibrium effort to vanish as public beliefs tend to 0 or 1. This is because public beliefs become unresponsive to new information asymptotically in those limits. Provided a non-negative MPE that vanishes at the extreme exists, then it should look like the figures below:

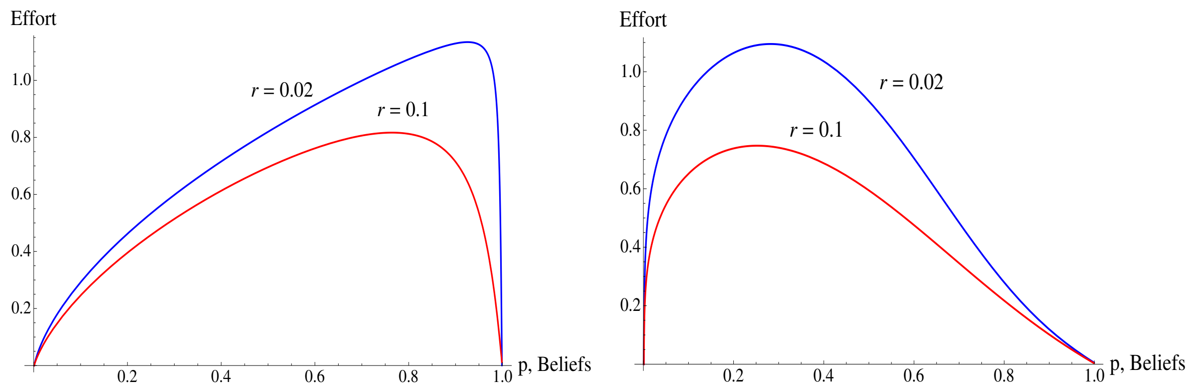

Figure 3: Left panel: $u(p^*, a^*) = p^*$. Right panel: $u(p^*, a^*) = a^* + p^*$. Parameter values: $\sigma_\xi = 1$.

In the left panel, the worker is payed according to the perceived value of his skills only, i.e. $u(p^*, a^*) = p^*$. In this case, the myopic gain from belief distortion is given by

$$\frac{d}{d\Delta}u\left(\frac{p\Delta}{1+p(\Delta-1)}\right)\Bigg|_{\Delta=1} = \frac{d}{d\Delta}\left[\frac{p\Delta}{1+p(\Delta-1)}\right]\Bigg|_{\Delta=1} = p(1-p), \; p \in (0,1).$$

[23]The proof of this result is available in the online Appendix.

Moreover, the learning dynamics are also symmetric around $p = 1/2$. Yet, incentives may not exhibit that property. This is because the worker's actions also affect the market's expectation of equilibrium play, and this ratcheting effect is not uniform across different levels of reputation.[24] For low reputations, inducing a marginal unit of belief asymmetry increases the threshold of belief manipulation that the worker will face in the near future. But since his reputation is low already, and such additional effort is not compensated, the monetary loss from being perceived as less skilled is small, which results in low effort. As reputation increases wages goes up, making it too costly for the worker to shirk and thus risking to start re-building his reputation again, as he is not compensated by this effort.

In the right panel the worker is also rewarded by effort. Now, the (on-path) myopic gain from belief distortion ceases to be symmetric around $p = 1/2$ and it is given by

$$p(1-p)[1 + V_{\Delta p}(p, 0)], \; p \in (0, 1).$$

Observe that in this case incentives will be stronger for low reputations. This is because the market compensates any additional unit of effort, offsetting the negative effect of facing a tougher standard. Hence, rewarding effort generates incentives for *building* a reputation. On the other hand a worker with a high reputation can afford to shirk, as the costs associated with building a reputation are partially covered by the compensation that he receives.

4 Sufficiency: Verification Theorems

The incentives equation, as a local incentive constraint, does not ensure that the long-run player does not benefit from inducing a large degree of belief discrepancy. In order to verify incentive compatibility globally we have to estimate the long-run player's payoff for deviations of any size. This section presents "verification" theorems (verifiable sufficient conditions) that ensure that a solution to the incentives equation is in fact a MPE.

Before stating the theorems I introduce the following concepts:

Definition 4.1.
- *A function $f : \mathbb{R} \to \mathbb{R}$ is said to satisfy a linear (respectively, quadratic) growth condition, if there exists a constant $C > 0$ such that $|f(x)| \leq C(1 + |x|)$ (respectively, $|f(x)| \leq C(1 + |x|^2)$).*
- *A differentiable function $g : Dom(g) \to \mathbb{R}$ is strongly convex if there exists a constant $\psi > 0$ such that $g(y) - g(x) - g'(x)(y - x) \geq \frac{\psi}{2}(x - y)^2$ for all $x, y \in Dom(g)$. If g is twice differentiable, strong convexity is equivalent to the existence of $\psi > 0$ such that $g''(x) \geq \psi > 0$ for all $x \in Dom(g)$.*

4.1 Verification Theorem I: General Case

Proposition 2.7 is important because it offers a way for showing the existence of MPE. In fact, if V solves the PDE (23)-(24), then it satisfies the HJB equation (20) with

$$a^*(\cdot) := \arg\max_{a \in A} \{a\beta V_\Delta(\cdot, 0) - g(a)\}$$

[24] From the incentives equation (29) it can be seen that the term $V_{\Delta p}(p, 0)V_\Delta(p, 0)$ makes the ODE non symmetric around $p = 1/2$.

taken as given in the dynamics of Δ. Hence, classic verification theorems from dynamic programming apply. Furthermore, since V satisfies the equilibrium condition (24), the corresponding policy evaluated at $\Delta = 0$ is, by construction, a MPE.

Theorem 4.2. (Verification Theorem I). *Suppose that $V \in C^{2,1}(\mathbb{R}^2)$ is a solution to the PDE (23)-(24) satisfying a quadratic growth condition. Let $a_t^* = \arg\max_{a \in A}\{a\beta V_\Delta(p_t + \Delta_t, 0)) - g(a)\}$, $t \geq 0$, and suppose that V satisfies*

$$\limsup_{t\to\infty} \mathbb{E}[e^{-rt}V(p_t, \hat{\Delta}_t^*)] \geq 0, \ \forall \textit{ feasible strategies } \hat{a}$$

where $\hat{\Delta}_t^*$ denotes the belief asymmetry process under the pair $(\hat{a}, a^*)$. Then, V is an upper bound to the long-run player's value function when the small players construct beliefs using a^*. Furthermore, if under the pair $(\alpha, a^*) := (g'^{-1}(\beta V_\Delta(p, \Delta)), g'^{-1}(\beta V_\Delta(p + \Delta, 0)))$ the ODE (17) has a unique solution $\Delta^{\alpha,*}$ that satisfies

$$\liminf_{t\to\infty} \mathbb{E}[e^{-rt}V(p_t, \Delta_t^{\alpha,*})] \leq 0,$$

and the process $\alpha_t := g'^{-1}(\beta V_\Delta(p_t, \Delta_t^{\alpha *})$ is feasible, then V is the long-run player's value function. Thus, a^* is a MPE. Moreover, if a^* is differentiable and interior, then it satisfies the incentives equation.

Proof: Since the PDE (23)-(24) is a standard HJB equation subject to an additional fixed point condition, the result follows directly from standard verification theorems in dynamic programming. See for instance Pham (2009).

□

Besides delivering an explicit expression for a MPE, Theorem 4.2 provides additional insights on the way in which incentives are structured in this class of games. In fact, in Section 6 I apply the previous theorem to a class of linear-quadratic games for which the associated PDE admits an analytic solution. In this case, it is possible to characterize the dynamics of belief asymmetry off the equilibrium path, and thus to understand how the long-run player can exploit his private information after a suboptimal actions have been undertaken.

Remark 4.3. There are very few classes of PDEs that admit analytic solutions. Furthermore, (23)-(24) imposes additional difficulties, as numerical methods would require calculating the value function and its derivative simultaneously at two different points, for all points in the domain. This paper offers a *localizing* method to the numerical approximation of markovian equilibria. More specifically, any solution $q(p)$ of the incentives equation can be used to construct a valid guess of a^*, which in turn plugged into the dynamics of belief asymmetry. The resulting PDE is therefore *local* and thus standard numerical methods apply. The resulting solution can then be compared against the initial guess along the equilibrium path. This numerical approach is a feasible way for evaluating the plausibility of existence of equilibria in highly nonlinear environments.

4.2 Verification Theorem II: Gaussian Learning

When learning is Gaussian it is possible to exploit the linearity of the learning process in order to establish sufficient conditions for the existence of MPE without fully solving the previous PDE. More specifically, when a MPE exists and it exhibits enough differentiability, the long-run player's payoff on the equilibrium path $V(\cdot,0)$, and the derivative of the value function with respect to the stock of belief asymmetry $V_\Delta(\cdot,0)$, solve the ODEs

$$rU(p) = u(p,q(p)) - g(g'^{-1}(\beta q(p))) + U'(p)\mu(p) + \frac{1}{2}\sigma^2 U''(p) \quad (30)$$

$$\tilde{r}(p,q(p))q(p) = u_p(p,g'^{-1}(\beta q(p))) + u_a(p,g'^{-1}(\beta q(p)))\frac{d}{dp}g'^{-1}(\beta q(p)) + q'(p)\mu(p) + \frac{1}{2}\sigma^2 q''(p),\ p\in\mathbb{R}, \quad (31)$$

provided incentives are interior. Theorem 4.4 below states conditions on any solution (U,q) to the previous system that ensure that $p\mapsto a^*(p) := g'^{-1}(\beta q(p))$ is in fact a MPE.

Intuitively, the long-run player will engage in large deviations whenever he can exploit the advantage of having access to private information about the fundamentals. Condition (iii) in the Theorem establishes a uniform bound on the long-run players *information rent* that ensures that "double deviations" are never profitable.

Theorem 4.4. (Verification Theorem II). *Suppose that learning is linear and that the manipulation cost function $g:\mathbb{R}\to\mathbb{R}$ is strongly convex. Let $U\in\mathcal{C}^2(\mathbb{R})$ and $q\in\mathcal{C}^2(\mathbb{R})$ denote a solution to the system (30)-(31) with the following properties:*

(i) U satisfies a quadratic growth condition; U' and q satisfy a linear growth condition,

(ii) For any feasible strategy $\hat{a}$, $\lim_{t\to\infty}\mathbb{E}[e^{-rt}U(p_t+\hat{\Delta}^q_t)] = \lim_{t\to\infty}\mathbb{E}[e^{-rt}q(p_t+\hat{\Delta}^q_t)\hat{\Delta}^q_t] = \lim_{t\to\infty}\mathbb{E}[e^{-rt}U'(p_t+\hat{\Delta}^q_t)\hat{\Delta}^q_t] = 0$, where $dp_t = -\kappa(p_t-\eta)dt + \sigma dZ_t$, and $\hat{\Delta}^q := (\hat{\Delta}^q_t)_{t\geq 0}$ denotes the belief asymmetry process under $(\hat{a}, g'^{-1}(\beta q(\cdot)))$,[25]

(iii) U'' and q' satisfy

$$|q'(p) - U''(p)| \leq \frac{\psi(r+4\beta+2\kappa)}{4\beta^2},\ \textit{for all}\, p\in\mathbb{R}. \quad (32)$$

Then the process $a^*_t := g'^{-1}(\beta q(p_t))$, $t\geq 0$ corresponds to a MPE.

Proof: See the Appendix.

□

The growth and transversality conditions on U ensure that there exists a unique solution to (30) which corresponds to the long-run player's utility if he follows a^*, whereas the growth conditions on U' and q are purely technical. Strong convexity, the transversality

[25]This condition can be weakend to $\limsup_{t\to\infty}\mathbb{E}[e^{-rt}[U(\hat{p}^q_t) + [q(\hat{p}^q_t) - U'(\hat{p}^q_t)]\hat{\Delta}^q_t] \geq 0$.

conditions on U' and q, and condition (iii) allow me to find an upper bound to the long-run player's payoff (under all feasible strategies) that coincides with U on the equilibrium path. Hence, it is never optimal for him to deviate from a^* when beliefs are aligned.

Condition (iii) is economically meaningful. It is easy to see that $U'(p) - q(p)$ is the derivative of long-run player's payoff with respect to his private type on the equilibrium path, keeping the level of public beliefs fixed at $p^* = p$. Analogously, $q(p) - U'(p)$ is the corresponding directional derivative in the opposite direction. Hence, $|U''(p) - q'(p)|$ measures the value for the long-run player of becoming privately informed about his type. If this *information rent* (in either direction) is large enough, the long-run player may find it optimal to deviate off the equilibrium path.

Similar bounds have been found in the literature of optimal contracting in settings where the asymmetric information present in the environment has high persistence. Both Williams (2011), in a context of persistent private information, and more recently Sannikov (2014), in a setting where actions have long-term impact on performance, derive similar bounds that make their optimal contracts fully incentive compatible. Since their conditions are bounds on the volatility of a "marginal utility" process, these are effectively bounds on the second derivative of the agent's value function (under the optimal contract), as it is the case in the class of games analyzed here. The bounds found here are nonetheless two-sided as the long-run player can benefit from under or over manipulating the public signal, whereas their bounds are instead one-sided, as deviations only in one direction matter for sufficiency.[26]

5 Uniformly Bounded Marginal Flow Payoffs

In this Section I apply Theorem 4.4 to show the existence of equilibria in environments where the marginal impact of public beliefs on the long-run player's payoff is uniformly bounded across all levels of public beliefs. For simplicity I restrict attention to *commitment models*: environments in which the small players' action is of the form $b_t := \mathtt{b}(p_t^*)$ (i.e. independent of a^*).[27] In these settings, when a MPE exists, the long-run player exerts costly effort in equilibrium, yet he cannot affect his flow playoff (as $p^* = p$ evolves exogenously on the equilibrium path). Hence, the long-run player would like to commit to not to manipulate the public signal, but such announcement is not incentive compatible.

Recall that when learning is Gaussian the key state variables take the form

$$\begin{aligned} dp_t &= -\kappa(p_t - \eta)dt + \sigma dZ_t^Y \\ d\Delta_t &= [-(\beta+\kappa)\Delta_t + \beta(a_t - a_t^*)]dt \end{aligned}$$

where $\sigma = \beta\sigma_\xi$ and $\beta = \beta(\kappa) := \frac{\gamma^*}{\sigma_\xi^2} = \sqrt{\kappa^2 + \sigma_\theta^2/\sigma_\xi^2} - \kappa$ is the *sensitivity of public beliefs to new information.* Moreover, from Section 3.1 we know that when $h(p) = \alpha p$, some $\alpha \in \mathbb{R}$, the model becomes fully linear, so the incentives equation admits a *constant* solution $q(p) \equiv \frac{\alpha}{r+\beta(\kappa)+\kappa}$. Consequently, whenever marginal payoffs become asymptotically linear, we can exploit the simplicity of the linear case to pin down natural asymptotic conditions that incentives equation must satisfy.

[26] In particular, Williams (2011) allows for only downward deviations in his report-dependent contract.

[27] Most of the results carry over to the case in which a^* also affects the long-run player's payoff, as long as the boundedness conditions on derivatives are extended appropriately.

The following assumptions are imposed in this section:

Assumption 5.1. *(i) Learning is Gaussian.*

(ii) Let $h(p) := u(\mathtt{b}(p))$*. Then* $m := \inf_{p\in\mathbb{R}} h'(p) > -\infty$ *and* $M := \sup_{p\in\mathbb{R}} h'(p) < \infty$*.*

(iii) Both $\lim_{p\to\infty} h'(p)$ *and* $\lim_{p\to-\infty} h'(p)$ *exist.*

(iv) $g : A \to \mathbb{R}$ *is twice differentiable and strongly convex, and* $g^{-1}(J) \subset A$*, where* $J := [m/(r+\beta(\kappa)+\kappa), M/(r+\beta(\kappa)+\kappa)]$*.*

One of the remarkable advantages of continuous-time stochastic control is its natural connection with partial and ordinary differential equations. Assumption (ii) allows us to apply classic results on the existence of bounded solutions to ODEs over unbounded domains to the system of ordinary differential equations (30)-(31). Condition (iii) states that flow payoffs become asymptotically linear, suggesting that the solution to the incentives equation must look like its corresponding analog in the linear case. Condition (iv) ensures that A is large enough so that incentives are always interior.

Theorem 5.2. (Existence of Bounded Solutions to the Incentives Equation). *Suppose that Assumption 5.1 holds. Then, there exists a solution* $q \in \mathcal{C}^2(\mathbb{R})$ *to the incentives equation such that*

$$q(p) \in \left[\frac{m}{r+\beta(\kappa)+\kappa}, \frac{M}{r+\beta(\kappa)+\kappa}\right], \text{ for all } p \in \mathbb{R}, \tag{33}$$

Moreover, any solution satisfying (33) also verifies that

$$\lim_{p\to-\infty} q(p) = \frac{\lim_{p\to-\infty} h'(p)}{r+\beta(\kappa)+\kappa} \text{ and } \lim_{p\to\infty} q(p) = \frac{\lim_{p\to\infty} h'(p)}{r+\beta(\kappa)+\kappa}. \tag{34}$$

When $\kappa > 0$*, any solution satisfying (33) and (34) also has a uniformly bounded derivative. When* $\kappa = 0$ *there exists a* $\mathcal{C}^2$ *solution that, in addition to satisfying (33) and (34), it also has a uniformly bounded derivative.*

Proof: See the Appendix.

□

Any solution satisfying (33)-(34) and that also has a uniformly bounded derivative will be referred to as a *bounded solution* of the incentives equation.

Having information about the derivative of a solution to the incentives equation is useful for three reasons. First, a uniform bound on q' guarantees that, given any feasible strategy, there will exist a unique strong solution for the dynamic of p^* (eqn. (16)) and consequently, for the dynamic of Δ. Thus, the long-run player's optimization problem is well-posed. Second, Theorem 4.4 imposes growth and transversality conditions on U', which will in turn depend on q'. Third, when the bounds on q' can be derived explicitly (as in Section 5.1 for the case $\kappa = 0$ and quadratic cost of effort), such conditions can be verified ex-post on any numerical solution of an "approximate" incentives equation that is defined over a bounded domain.

The next result deals with the existence and uniqueness of a solution to the ODE (30). It states that the unique solution to it is in fact the payoff that the long-run player would obtain if the small players construct beliefs using a bounded solution to the incentives equation, and he does not deviate off the equilibrium path.

Proposition 5.3. (The Long-Run Player's Equilibrium Payoff as a Solution to an ODE.) *Suppose that Assumption 5.1 holds and let q denote a bounded solution of the incentives equation. Then, there exists a unique $U \in \mathcal{C}^2(\mathbb{R})$ solution to the ODE (30). This solution is given by*

$$U(p) = \mathbb{E}\left[\int_0^\infty e^{-rt}[h(p_t) - g((g')^{-1}(\beta q(p_t)))]dt\right] \tag{35}$$

where $dp_t = -\kappa(p_t - \eta)dt + \sigma dZ_t$ for $t > 0$ and $p_0 = p$. Furthermore, U satisfies a linear growth and U' is uniformly bounded.

Proof: See the Appendix.

□

Finally, I establish conditions on the primitives $(r, m, M, \psi, \kappa, \sigma_\theta, \sigma_\xi)$ that ensure that a solution (U, q) to (30)-(31) satisfies the conditions of Theorem 4.4:

Theorem 5.4. (Existence of MPE) *Suppose that Assumption 5.1 holds and that $\kappa = 0$. Let $q : \mathbb{R} \to \mathbb{R}$ denote any bounded solution to the incentives equation. Then, the long-run player's information rent takes the form*

$$U''(p) - q'(p) = -\mathbb{E}\left[\int_0^\infty e^{-rs}\beta q(p_s)ds\right], \text{ with } dp_t = \sigma dZ_t,\ p_0 = p. \tag{36}$$

Moreover, if

$$\frac{M-m}{\psi} \leq \frac{\sqrt{2r\sigma_\xi^2}(r+\beta)^2}{4\beta^2} = \frac{\sqrt{2r\sigma_\xi^2}(r\sigma_\xi + \sigma_\theta)^2}{4\sigma_\theta^2}, \tag{37}$$

the process $(g'^{-1}(\beta q(p_t)))_{t\geq 0}$ is a MPE.

Proof: See the Appendix.

□

Equation (36) expresses the long-run player's information rent as a expected discounted payoff of (equilibrium) marginal utilities with respect to belief asymmetry. Intuitively, since beliefs have persistence, acquiring private information will affect all future continuation values.[28] In the specific case of beliefs evolving as a Brownian martingale ($\kappa = 0$), a change of any size in the initial level of private beliefs will affect all future beliefs by the same amount, so the continuation value at time s changes by $q(p_s)$, $s > 0$.

Condition (37) is found by obtaining a bound for (36), and imposing that (iii) in Theorem 4.4 holds. Condition (37) is relaxed when manipulation becomes more costly

[28] Similar expressions have been derived by Sannikov (2014) in a contracting environment.

(ψ increases) and when $1/\sigma_\xi$ and σ_θ decrease: in the latter two cases, beliefs respond less strongly to new information, reducing the benefits of belief manipulation.[29] Notice that while the condition is violated for $M > m$ and $r = 0$, it is always possible to find parameters $(\sigma_\theta, \sigma_\xi, \psi)$ under which the condition holds in the $r > 0$ case. Finally, the condition is relaxed when $M - m$ decays, and trivially satisfied when payoffs are linear ($M = m$), as private information ceases to have any value in this case (Section 3.1).

To conclude, observe that the case $\kappa = 0$ corresponds to the environment that offers the *largest returns* from belief manipulation. In fact, since $\beta(\kappa)$ is decreasing in κ and $\beta(\kappa) + \kappa = \sqrt{\kappa + \sigma_\theta^2/\sigma_\xi^2}$ increases in κ, public beliefs become (i) less responsive to new information and they (ii) decay more rapidly over time, as κ increases. Hence, Markovian equilibria are also expected to arise in the mean reverting case.

6 Unbounded Marginal Flow Payoffs: The Linear-Quadratic Case

This section introduces a subclass of *linear-quadratic* games for which on- and off-path incentives can be fully characterized via the verification theorem 4.2. These games have a linear-quadratic structure because learning is Gaussian (linear) and the long-run player's flow utility is a quadratic loss function of public beliefs and the cost of manipulation.

I show below that, when a condition on the curvature on the long-run player's flow payoff holds, the PDE (23)-(24) that summarizes global behavior (Proposition 2.7) admits an analytic solution. In this case, a the MPE found is linear in the public belief, yet exhibits all the forces present in the incentives equation. Interestingly, this curvature condition is also *necessary* for the existence of a linear MPE: when this condition is violated, the small players cannot discipline the long-run player using a linear conjecture of signal manipulation.

6.1 Linear-Quadratic Games: Existence Result

Definition 6.1. *A signal-jamming game is said to be of linear-quadratic form if*

(i) *Fundamentals θ are a mean reverting process:* $d\theta_t = -\kappa(\theta_t - \eta)dt + \sigma_\theta dZ_t^\theta$, $t \geq 0$;

(ii) $A = \mathbb{R}$ *and* $g(a) = \frac{\psi}{2}a^2$, $\psi > 0$;

(iii) $h(p^*) := u(\mathtt{b}(p^*, a^*)) = u_0 + u_1 p^* - u_2 p_t^{*2}$, *where* $u_0, u_1 \in \mathbb{R}$ *and* $u_2 \geq 0$.[30]

The next result shows the existence of a *linear* (in beliefs) equilibrium which exhibits all the forces mentioned in the previous sections.

[29] Observe that belief distortions become more persistence as $\beta(0) := \sigma_\theta/\sigma_\xi$ decrease, which provides more incentives for belief manipulation. However, the sensitivity effect is stronger. This was also found in the linear case, where incentives are proportional to $\frac{\beta(0)}{r+\beta(0)}$.

[30] The analysis can be easily extended to the case in which the market's action is also linear in a^*: $u(\mathtt{b}(p^*, a^*)) = u_0 + u_1(k_1 p^* + k_2 a^*) - u_2(k_1 p_t^* + k_2 a_t^*)^2$, $k_1, k_2 \in \mathbb{R}$.

Theorem 6.2. *Suppose that a linear-quadratic game of signal manipulation is such that*

$$u_2 \leq \frac{\psi(r+\beta(\kappa)+2\kappa)^2}{8\beta^2(\kappa)}. \tag{38}$$

Then a linear MPE exists. In this equilibrium, the long-run player's value function is given by $V(p,\Delta) = \alpha_0 + \alpha_1 p + \alpha_2\Delta + \alpha_3 p\Delta + \alpha_4 p^2 + \alpha_5\Delta^2$, *and the equilibrium degree of manipulation by* $a^*(p^*) = \frac{\beta}{\psi}V_\Delta(p,0) = \frac{\beta}{\psi}[\alpha_2 + \alpha_3 p^*]$, *where* $\alpha_0, \alpha_1 \in \mathbb{R}$, $\alpha_4, \alpha_5 < 0$, *and*

$$\alpha_2 = \frac{\eta\kappa\alpha_3 + u_1}{r+\beta(\kappa)+\kappa+\frac{\beta^2(\kappa)\alpha_3}{\psi}}, \text{ and} \tag{39}$$

$$\alpha_3 = \frac{\psi}{2\beta^2(\kappa)}\left[-(r+\beta(\kappa)+2\kappa) + \sqrt{(r+\beta(\kappa)+2\kappa)^2 - \frac{8u_2\beta^2(\kappa)}{\psi}}\right] < 0. \tag{40}$$

Off the equilibrium path, $\Delta_t = \Delta_0 e^{\rho t}$, $t \geq 0$, *with* $\rho < 0$, *so any stock of belief asymmetry vanishes asymptotically.*

Proof: See the Appendix.

□

The long-run player's on-path utility takes the form $V(p,0) = \alpha_0 + \alpha_1 p + \alpha_4 p^2$, (see the Appendix for the explicit expressions), and the equilibrium manipulation strategy is a decreasing function of public beliefs ($\alpha_3 < 0$). This is intuitive as the long-run player has the incentive to push public beliefs toward the bliss point $\frac{u_1}{2u_2}$.

The incentives generated within the class of linear-quadratic games, although linear, satisfy all the forces identified in the incentives equations. First, the size of the marginal flow payoffs drive the size of the long-run player's incentives: as the myopic gain from belief manipulation decays, equilibrium effort decreases. Second, the long-run player conditions his actions to the anticipated economic conditions: if for instance $u_1 = 0$ and $\eta < 0$, then $\alpha_2 > 0$,[31] which means that the long-run player indeed manipulates the signal at the bliss point of his preferences. This is because he anticipates that fundamentals will mean revert to η with high probability, region in which it is optimal to exert signal manipulation. But since the cost of signal manipulation is convex, it is optimal to invest in signal manipulation today. Finally, the rate at which a marginal unit of belief asymmetry depreciates over time is endogenous. This is because

$$\tilde{r}(p) := r + \phi_\Delta(p,0) + \frac{\beta^2}{\psi}V_{\Delta p}(p,0) = r + \beta + \kappa + \beta^2\alpha_3$$

with $\alpha_3 < 0$. Hence, when a MPE in linear strategies exists, the ratcheting forces encourage manipulation (as the standard of manipulation decays in the direction of manipulation) and also uniform across all the state space. All these effects can be seen in the following figure:

[31] Comes from the fact that $\alpha_3 < 0$ and $r + \beta + \kappa + \frac{\beta^2\alpha_3}{\psi} > 0$.

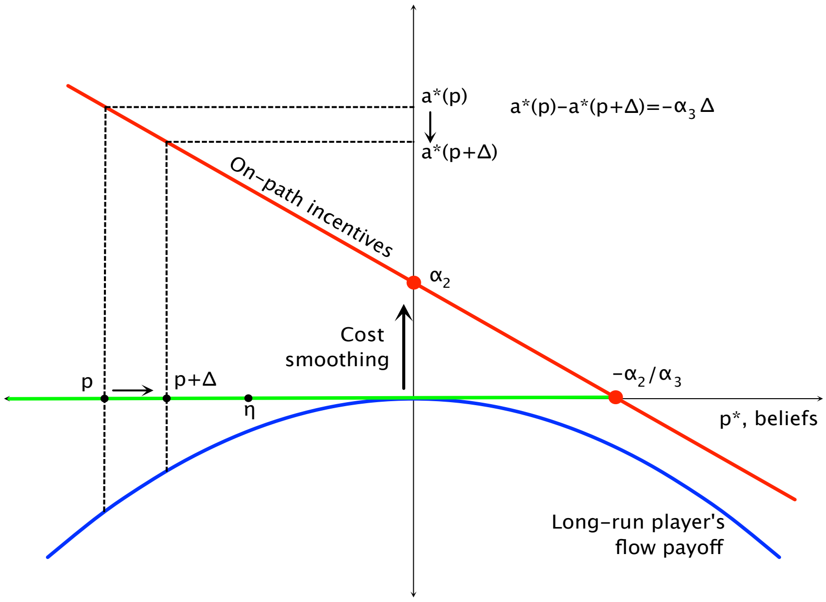

Figure 4: Two determinants of the size of incentives: cost smoothing and the effect that distorting the market's conjecture has on rate at which Δ depreciates.

The relevant parameters of the linear-quadratic model correspond to the rate of mean reversion κ, the long-run mean of fundamentals η, and the convexity parameter ψ of the effort disutility function. The sensitivity of equilibrium incentives to these parameters can be seen in the following figures:

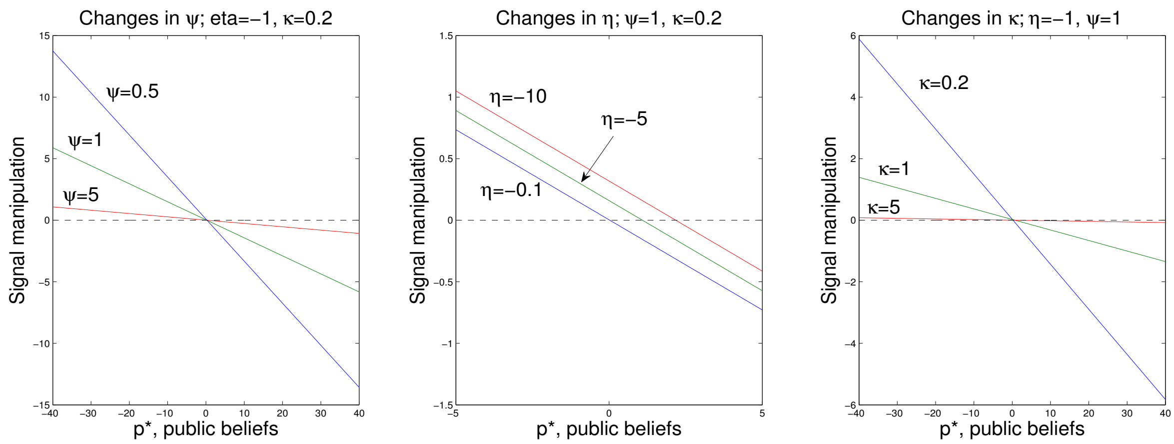

Figure 5: Sensitivity of equilibrium incentives to ψ, η and κ, respectively.

In the left panel, given any fixed volatility (slope) of the long-run player's strategy, more convex manipulation cost reduce his utility. Hence, the equilibrium strategy becomes less steep as ψ increases. In the middle panel, The middle panel, as the distance between the long-run player's consumption bliss point and the long-run average value of fundamentals (η) increases, the speed at which fundamentals will move away from the bliss point will go up, thus inducing more signal manipulation due to cost smoothing. Finally, changes in κ (third panel) can have two effects on incentives. First, as the rate of mean reversion increases there is a pressure towards more cost smoothing (numerator in

α_2). However, an increase in κ makes public beliefs less persistent and also less responsive to new information.

6.2 The Curvature Condition

Theorem 6.2 ensures the existence of a linear MPE provided the curvature condition (38) holds. In this section I show that the curvature condition is also *necessary* for the existence of such a linear equilibria.

In order to understand the intuition behind this result, observe that as the curvature of the payoff function u_2 becomes larger, the myopic benefit from inducing belief asymmetry increases. Consequently, in order to make deviations more costly, the small players have to impose a tougher manipulation standard on the long-run player. One way in which this can be done is through imposing a steeper conjecture a^*. However, steeper conjecture generate a second effect: the return from inducing belief asymmetry increases. In fact, in any MPE a^*, the rate of return on belief asymmetry

$$\tilde{r}(p) = r + \beta + \kappa + \frac{\beta^2}{\psi}\frac{da^*}{dp^*}(p^*),\ p^* \in \mathbb{R},$$

decays as a^* becomes more negatively sloped. Intuitively, by pushing public beliefs toward zero, the long-run player will face an even lower manipulation standard tomorrow, which increases the benefits from manipulating the belief of the small players. I show next that for any $u_2 > \frac{\psi(r+\beta+2\kappa)^2}{8\beta^2}$, a linear conjecture cannot control *simultaneously* both the immediate benefits from a deviation (measured by the size of marginal flow payoffs) and the long-term benefits from engaging in large deviations off the equilibrium path (measured by $\tilde{r}$).

Formally, suppose that the small players conjecture that $\hat{a}(p^*) = \frac{\beta\hat{\alpha}_3}{\psi}p^* = \frac{\beta\hat{\alpha}_3}{\psi}(p+\Delta)$, $\hat{\alpha}_3 < 0$, will arise in equilibrium.[32] Consider the deterministic control problem $\mathcal{P}(\hat{\alpha}_3)$

$$\max_{a\in\mathcal{A}} \quad \int_0^\infty e^{-rt}\left(-u_2(p_t+\Delta_t)^2 - \frac{\psi}{2}a_t^2\right)dt$$

$$s.t. \quad dp_t = -\kappa p_t dt, \tag{41}$$

$$d\Delta_t = \left[-\left(\beta+\kappa+\frac{\beta^2}{\psi}\hat{\alpha}_3\right)\Delta_t + \beta a_t - \frac{\beta^2}{\psi}\hat{\alpha}_3 p_t\right]dt, \tag{42}$$

where $u_2, \kappa, \beta, \psi > 0$. This problem corresponds to a deterministic version of the linear-quadratic game previously studied in the case in which u_1, η and the volatility term in the private beliefs process are all zero (β, however, depends on $\sigma_\xi > 0$, in the same way it does in the stochastic game). Studying this problem is without loss of generality when the goal is to find the long-run player's best response in the original stochastic game.[33]

[32] It is straightforward to argue that there is no equilibrium in which on-path effort is given by $\hat{\alpha}_3 p^*$, $p^* \in \mathbb{R}$ with $\hat{\alpha}_3 > 0$.

[33] This is true because of two reasons. First, u_1 and η do not affect the slope of the long-run player's best response, which is what at the end of the day matters for the existence of a linear equilibrium. Second, since the original problem has a linear-quadratic structure, any second order term will only affect the level (or constant term) of the long-run player's value function. Thus, when a linear best-response to $\hat{a}$ exists in the original stochastic problem, this one can be found through solving this deterministic version (this is called the *certainty equivalence principle*).

The next results show that the previous problem always admits a linear best response. Furthermore, it shows that when the curvature condition is violated, the long-run player responds *more aggressively* to $\hat{a}$. This is intuitive, as the long-run player gains from manipulating the signal toward his preference bliss point.

Proposition 6.3. *The value function associated with $\mathcal{P}(\hat{\alpha}_3)$, $\hat{\alpha}_3 < 0$, has the form $V(p,\Delta) = \alpha_3 p\Delta + \alpha_4 p^2 + \alpha_5 \Delta^2$, where α_3, α_4 and α_5 depending on $\hat{\alpha}_3$. Moreover, if $u_2 > \frac{\psi(r+\beta+2\kappa)^2}{8\beta^2}$, then $\alpha_3(\hat{\alpha}_3) < \hat{\alpha}_3$. Consequently, starting from a common prior, the long-run agent finds it optimal to engage in belief manipulation at time zero.*

Proof: See the Appendix.

□

When the curvature condition is violated, the PDE (23)-(24) does not admit a quadratic function as a solution, and hence a linear MPE ceases to exist. However, this does not mean that there are no equilibria in pure strategies. A MPE may exists provided the incentives equation (25) admits another solution that is nonlinear, and provided the information rent that the long-run player acquires after a deviation is not too large.

7 Conclusions

In this paper I developed a class of continuous-time games for studying strategic behavior in environments where agents learn about the relevant economic environment. Necessary and sufficient conditions for the existence of Markov Perfect Equilibria that are captured in a powerful nonlinear ODE are obtained. Most importantly, the methods and results presented in this paper are general enough to be applied in a wide set of environments, ranging from the determinants of workers' incentives in labor markets, to central banks' behavior in response to unobserved states of the economy.

The choice of casting the model in continuous-time is driven by its well-known advantages over traditional ones. Continuous-time methods are useful because they offer clarity, tractability and computational power. In particular, exploiting the connection between stochastic control and the theory of differential equations allowed me to characterize equilibria for environments well beyond the linear frameworks previously studied. The use of martingale methods allowed me to find intuitive sufficient conditions for the existence of equilibria. More fundamental is the fact that the continuous-time approach typically offers particularly clean insights.

The results in this paper rely on the manipulation technology having an additively-separable structure. Allowing for complementarities between actions and fundamentals creates another channel for incentives: experimentation. By studying models with an additively-separable structure, I am able to eliminate the experimentation effect and concentrate only on belief manipulation motives. The model is thus not appropriate for studying incentives in environments where affecting the speed of learning of others plays a predominant role. However, the envelope methods used to characterize incentives have a direct analog in such non-separable settings.

Finally, the characterizations obtained in this paper depend on the assumption of ex-ante symmetric uncertainty about the underlying fundamental. While it is widely understood that people make economic decisions under incomplete information about the environment, most of the economic research has dealt with the case of adverse selection only. Learning about the environment occurs everywhere, and hence it deserves more attention. This paper has offered a framework for understanding incentives in settings where agents have approximately the degree of uncertainty about the economic environment at the outset.

References

[1] Acemoglu, D. and J-S. Pischke (1999): "The Structure of Wages and Investment in General Training," *Journal of Political Economy, 107, pp. 539-572.*

[2] Board, S., and M. Meyer-ter-Vehn (2010a): "A Reputation Theory of Firm Dynamics," Working Paper, UCLA.

[3] ——. (2013): "Reputation for Quality," *Econometrica, Vol. 81, No. 6, pp. 2381-2462.*

[4] Bohren, A. (2013): "Stochastic Games in Continuous Time: Persistent Actions in Long-term Relationships," Working Paper, UC San Diego.

[5] Bonatti, A., and J. Hörner (2011): "Career Concerns with Coarse Information," Cowles Foundation Discussion Paper 1831, Cowles Foundation for Research in Economics, Yale University, revised Jan. 2012.

[6] Cavallo, A. (2012): "Online and Official Price Indexes: Measuring Argentina's Inflation," *Journal of Monetary Economics, Vol. 60, No. 2, pp. 152-165.*

[7] Cisternas, G. (2012): "Shock Persistence, Endogenous Skills and Career Concerns," Working Paper, Princeton.

[8] Cukierman, A., and A. Meltzer (1986): "A Theory of Ambiguity, Credibility and Inflation Under Discretion and Asymmetric Information," *Econometrica, Vol. 54, No. 5, pp.1099-1128.*

[9] DeMarzo, P., and Y. Sannikov (2011): "Learning, Termination and Payout Policy in Dynamic Incentive Contracts," Working Paper, Princeton.

[10] Dewatripont, M., I. Jewitt, and J. Tirole (1999a): "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies, Vol. 66, No. 1, pp. 199-217.*

[11] ——. (1999b): "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies," *Review of Economic Studies, Vol. 66, No. 1, pp. 199-217.*

[12] Dixit, A., and R. Pindyck (1994): *Investment Under Uncertainty.* New Jersey: Princeton University Press.

[13] Faingold, E., and Y. Sannikov, (2010): "Reputation in Continuous-Time Games." *Econometrica, Vol. 79, No. 3, pp. 773-876.*

[14] He, Z., B. Wei and J. Yu: "Optimal Long-Term Contracting with Learning," Working paper.

[15] Holmstrom, B. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics, Vol. 10, No. 2, pp. 74-91.*

[16] ——. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *The Review of Economic Studies, 66, pp. 169-182.*

[17] Karlin, S., and H. Taylor (1981): *A Second Course in Stochastic Processes*. London: Academic Press.

[18] Kovrijnykh, A. (2007): "Career Uncertainty and Dynamic Incentives," Working paper, University of Chicago.

[19] Kydland, F. and E. Prescott (1977). "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy, University of Chicago Press, Vol. 85, No. 3, pp 473-491.*

[20] Laffont, J.J. and J. Tirole (1988): "The Dynamics of Incentive Contracts," *Econometrica, Vol. 56, No. 5, pp. 1153-1175.*

[21] Liptser, R., and A. Shiryaev (1977): *Statistics of Random Processes I and II*. New York: Springer-Verlag.

[22] Martinez, L. (2009): "Reputation, Career Concerns, and Job Assignments" *The B.E. Journal of Theoretical Economics, Berkeley Electronic Press, Vol. 9, No. 1. (Contributions), Article 15.*

[23] Pham, H. (2009): *Continuous-time Stochastic Control and Optimization with Financial Applications*. Berlin: Springer.

[24] Sannikov, Y. (2014): "Moral Hazard and Long-Term Incentives," Working paper.

[25] Prat, J. and B. Jovanovic (2013): "Dynamic Contracts when Agent's Quality is Unknown," Working paper.

[26] Williams, N. (2011): "Persistent Private Information," *Econometrica, Vol. 79, No. 4, pp. 1233-1275.*

[27] Wonham, W. (1985): *Linear Multivariate Control: A Geometric Approach*. Berlin: Springer.

8 Appendix A: Proofs of Sections 2 and 3

Proof of Proposition 2.4: Case (L) (and the public belief dynamics (13)-(14) for the Gaussian case) is a direct consequence of Theorem 12.1 in Liptser and Shiryaev (1977).

For case (NL), Theorem 9.1 in Liptser and Shiryaev (1977) shows that $\pi_t := \mathbb{P}(\theta_t = h|\mathcal{F}_t^Y)$, $t \geq 0$, evolves according to

$$d\pi_t = (\lambda_1(1-\pi_t) - \lambda_0\pi_t)dt + \frac{\delta\pi_t(1-\pi_t)}{\sigma_\xi}\left(\frac{dY_t - \pi_t dt}{\sigma_\xi}\right),$$

where $\delta := h - \ell$, and $Z_t^Y := \frac{1}{\sigma_\xi}\left(Y_t - \int_0^t \pi_s ds\right)$ is a $\mathbb{F}^Y-$Brownian motion from the long-run player's standpoint.

Starting from any point in $(0,1)$, π never hits zero or one (Karlin and Taylor (1981)). Thus, applying Ito's rule to $p_t := \log(\pi_t/(1-\pi_t))$ we get

$$dp_t = \left(\frac{\lambda_1}{\pi_t} - \frac{\lambda_0}{1-\pi_t} - \frac{\delta^2(1-2\pi_t)}{2\sigma_\xi^2}\right)dt + \frac{\delta}{\sigma_\xi}dZ_t^Y.$$

Since $\pi_t = \frac{e^{p_t}}{1+e^{p_t}}$, the dynamics of (43) as a function of p is given by

$$dp_t = \left[\lambda_1\frac{e^{p_t}+1}{e^{p_t}} - \lambda_0(1+e^{p_t}) - \frac{\delta^2}{2\sigma_\xi^2}(1 - 2\frac{e^{p_t}}{1+e^{p_t}})\right]dt + \frac{\delta}{\sigma_\xi^2}\left(dY_t - \frac{e^{p_t}}{1+e^{p_t}}dt\right),\ t \geq 0.$$

Finally, the dynamic of the small players' belief (13)-(14) in the nonlinear case are obtained using the same Theorem, using signal ξ and the innovation process $Z_t^* := \frac{1}{\sigma_\xi}\left(\xi_t - \int_0^t (a_s^* + \pi_s^*)ds\right)$, $t \geq 0$. This concludes the proof.

□

Proof of Proposition 3.1: It is easy to see that $V(\theta) = \alpha_0^o + \alpha_1^o\theta + \alpha_2^o\theta^2$ solves the HJB equation

$$rV(\theta) = \max_{a\in\mathbb{R}}\left\{-k_2(\theta-\overline{\theta})^2 - \frac{\psi}{2}a^2 + [a-\kappa\theta]V_\theta(\theta) + \frac{1}{2}\sigma_\theta^2 V_{\theta\theta}(\theta)\right\}$$

when $\alpha_0^o = -k_2\overline{\theta}^2 + \frac{\alpha_1^o}{2\psi} + \alpha_2\sigma_\theta^2$, $\alpha_1^o = -\frac{2k_2\overline{\theta}}{\frac{2\alpha_2^o}{\psi} - (r+\kappa)}$ and $\alpha_2^o = \frac{\psi}{2}\left[(r+2\kappa) - \sqrt{(r+2\kappa)^2 + \frac{8k_2}{\psi}}\right]$.

In order to show that V is indeed the policymaker's value function and that $a^o(\theta) = \frac{1}{\psi}V_\theta(\theta)$ an optimal policy, two things remain to be checked:

1. Any feasible policy $\hat{a}$ such that $\mathbb{E}\left[\int_0^\infty e^{-rt}\left| - k_2(\theta_t^{\hat{a}} - \overline{\theta})^2 - \frac{\psi}{2}\hat{a}_t^2\right|dt\right] < \infty$, must also satisfy $\limsup_{t\to\infty} e^{-rt}\mathbb{E}[V(\theta_t^{\hat{a}})] \geq 0$, where $\theta^{\hat{a}}$ is the dynamic of θ under the policy $\hat{a}$.
2. Second, that along the conjectured optimal strategy a^o, $\liminf_{t\to\infty} e^{-rt}\mathbb{E}[V(\theta_t^{a^o})] \leq 0$.

As it is well known, condition 1 (along with the fact that V solves the HJB equation) implies that V is an upper bound to the policymaker's utility. Part 2. yields that V is attainable under the Markov control a^o. Conditions 1. and 2. are shown to be true in the non-commitment case, which corresponds to a slightly more general environment than the one analyzed here. We refer the reader to the proof of Theorem 3.2.

□

Proof of Theorem 3.2: Suppose the the market conjectures a manipulation strategy of the form $a^*(p^*) = \frac{1}{\psi}[\alpha_1 - \alpha_2 + 2\alpha_3 p^*]$ where $\alpha_1 = -\frac{2\psi k_2 \overline{\theta}}{2\alpha_3 - \psi(r+\kappa)}$, $\alpha_2 = \frac{k_1}{2\alpha_3\psi - (r+\beta+\kappa)}$ and $\alpha_3 = \psi\left[(r+2\kappa) - \sqrt{(r+2\kappa)^2 + 8k_2/\psi}\right]/2$ (observe that $\alpha_1 = \alpha_1^o$ and $\alpha_3 = \alpha_2^o$ the parameters of the full observability case).

It is easy to the that the quadratic form $V = \alpha_0 + \alpha_1 p + \alpha_2 \Delta + \alpha_3 p^2$ with $\alpha_0 = -k_2\overline{\theta}^2 + \frac{1}{2\psi}(\alpha_1^2 - \alpha_2^2) + (\beta\sigma_2)^2\alpha_3$, solves the HJB equation

$$\begin{aligned} rV(p,\Delta) &= \sup_{a\in\mathbb{R}} \Big\{ -k_1\Delta - k_2(p-\overline{\theta}^2)^2 - \frac{\psi}{2}a^2 + [a-\kappa p]V_p(p,\Delta) + \frac{1}{2}\beta^2\sigma_2^2 V_{pp}(p,\Delta) \\ &\qquad + [-\Delta(\beta+\kappa) + \frac{1}{\psi}(\alpha_1 - \alpha_2 + 2\alpha_3(p+\Delta)) - a]V_\Delta(p,\Delta) \Big\} \end{aligned}$$

In fact, the right-hand side yields a first order condition of the form

$$\alpha(p,\Delta) = \frac{1}{\psi}[V_p(p,\Delta) - V_\Delta(p,\Delta)] = \frac{1}{\psi}[\alpha_1 - \alpha_2 + 2\alpha_3 p] = a^*(p).$$

while the market's conjecture off the equilibrium path then takes the form $a^*(p+\Delta) = \frac{1}{\psi}[\alpha_1 - \alpha_2 + 2\alpha_3(p+\Delta)]$, as expressed in the last line of the HJB equation. Inserting the above first order condition along with the corresponding expressions for V, V_p, V_Δ and V_{pp} in the HJB equation yields the system of equations

$$\begin{aligned} (\alpha_0):\ r\alpha_0 &= -k_2\overline{\theta}^2 + \frac{1}{2\psi}(\alpha_1^2 - \alpha_2^2) + (\beta\sigma_2)^2\alpha_3 \\ (\alpha_1):\ r\alpha_1 &= 2k_2\overline{\theta} - \alpha_1\kappa + \frac{2\alpha_1\alpha_3}{\psi} \\ (\alpha_2):\ r\alpha_2 &= -k_1 + \alpha_2\left[\frac{2\alpha_3}{\psi} - (\beta+\kappa)\right] \\ (\alpha_3):\ r\alpha_3 &= -k_2 - 2\alpha_3\kappa + \frac{2\alpha_3^2}{\psi}. \end{aligned}$$

The expressions for α_i, $i = 0,1,2,3$, stated above are the unique solution to this system.

In order to show that V is the policymaker's value function and a^* a MPE, two things remain to be checked:

1. Any feasible strategy $\hat{a}$ such that $\mathbb{E}\left[\int_0^\infty e^{-rt}\left| -k_1\Delta_t^{a,a^*} - k_2(p_t^{\hat{a}} - \overline{\theta})^2 - \frac{\psi}{2}\hat{a}_t^2\right| dt\right] < \infty$, must also satisfy $\limsup_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t^a, \hat{\Delta}_t^*)] \geq 0$.
2. Second, that along the strategy $\alpha_t := \frac{1}{\psi}[\alpha_1 - \alpha_2 + 2\alpha_3 p_t]$, $\liminf_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t^\alpha, \Delta_t^{\alpha,*})] \leq 0$, where p^α denote the perceived inflation process under the control α, and $\Delta^{\alpha,*}$ the belief asymmetry process under (α, a^*).[34]

[34]The existence of solutions p^α and $\Delta^{\alpha,*}$ to such linear equations is trivial, as α and a^* are linear policies.

1. Consider a feasible strategy $\hat{a}$ such that the policymaker's payoff is finite. Observe that under the pair $(\hat{a}, a^*)$, the belief-asymmetry process $\hat{\Delta}^*$ takes the form

$$\hat{\Delta}^* = e^{-\vartheta t}\Delta_o + [1 - e^{-\vartheta t}]\frac{\alpha_1 - \alpha_2}{\psi\vartheta} + \int_0^t e^{-\vartheta(t-s)}\left[\frac{2\alpha_3}{\psi}p_s^{\hat{a}} + \hat{a}_s\right]ds \tag{43}$$

with $\vartheta := \beta + \kappa - 2\alpha_3 > 0$. Moreover, since $p_t^{\hat{a}} = e^{-\kappa t}p_0 + \int_0^t e^{-\kappa(t-a)}[\hat{a}_s ds + \beta\sigma_2 d\overline{Z}_s^0]$, $t \geq 0$, we can use integration by parts and Fubini's theorem to show that the the payoff under strategy $\hat{a}$ can be written as

$$\begin{aligned} &\mathbb{E}\left[\int_0^\infty e^{-rt}\left(-k_1\hat{\Delta}_t^* - k_2(p_t^{\hat{a}} - \overline{\theta})^2 - \frac{\psi}{2}\hat{a}_t^2\right)dt\right] \\ = \;&\mathbb{E}\left[\int_0^\infty e^{-rt}\left(-k_2(p_t^{\hat{a}})^2 + C_1\hat{a}_s - \frac{\psi}{2}\hat{a}_s^2\right)dt\right] + C_2 \end{aligned}$$

for some constants C_1 and C_2. Therefore, the payoff under $\hat{a}$ will be finite if and only if

$$\mathbb{E}\left[\int_0^\infty e^{-rt}(p_t^{\hat{a}})^2 dt\right] < \infty \quad \text{and} \quad \mathbb{E}\left[\int_0^\infty e^{-rt}\hat{a}_t^2 dt\right] < \infty. \tag{44}$$

from where we conclude that $\liminf_{t\to 0}\mathbb{E}[e^{-rt}(p_t^{\hat{a}})^2] = 0$. But since $\alpha_3 < 0$, we conclude that $\limsup_{t\to 0}\mathbb{E}[e^{-rt}\alpha_3(p_t^{\hat{a}})^2] = 0$.

Now, plugging the expression for $p_t^{\hat{a}}$ into (43) and using integration by parts, we can find positive constants C_3, C_4 and C_5 such that

$$|\mathbb{E}[\hat{\Delta}_t^*]| \leq C_3 e^{-\vartheta t} + C_4\int_0^t e^{-\kappa(t-s)}\mathbb{E}[|\hat{a}_s|]ds + C_5\int_0^t e^{-\vartheta(t-s)}\mathbb{E}[|\hat{a}_s|]ds.$$

The Cauchy-Schwartz's and Jensen's inequalities then yield that

$$e^{-rt}\int_0^t e^{-\lambda(t-s)}\mathbb{E}[|a_s|]ds \leq \underbrace{\left(e^{-rt}e^{-2\lambda t}\int_0^t e^{2\lambda s}ds\right)^{1/2}}_{L_t^1 :=}\underbrace{\left(e^{-rt}\int_0^t \mathbb{E}[a_s^2]ds\right)^{1/2}}_{L_t^2 :=}.$$

for $\lambda = \kappa, \vartheta > 0$. It is easy to see that $L_t^1 \to 0$ as $t \to \infty$. For L_t^2, observe that $e^{-rt}\int_0^t \mathbb{E}[\hat{a}_s^2]ds < \int_0^t e^{-rs}\mathbb{E}[\hat{a}_s^2]ds < \int_0^\infty e^{-rs}\mathbb{E}[\hat{a}_s^2]ds < \infty$, and thus L^2 is uniformly bounded. This shows that $\lim_{t\to 0} e^{-rt}\mathbb{E}[\hat{\Delta}_t^*] = 0$. Furthermore, since the exact same argument (Cauchy-Schwartz and Jensen) applied to $\mathbb{E}[p_t^{\hat{a}}]$, it is also that $\lim_{t\to 0} e^{-rt}\mathbb{E}[p_t^a] = 0$, concluding that $\limsup_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t^a, \Delta_t^{a,a^*})] \geq 0$ for any feasible strategy satisfying under which the policymaker attains finite utility.

2. Suppose that the policymaker follows the policy $\alpha(p, \Delta) = \alpha_t := \frac{1}{\psi}[\alpha_1 - \alpha_2 + 2\alpha_3 p]$. Then, the belief-asymmetry process evolves according to

$$d\Delta_t = [-(\beta + \kappa)\Delta_t + \alpha_3\Delta]dt,\ t > 0,\ \Delta_0 = \Delta^o.$$

As a consequence, $\Delta_t^{\alpha,*} = e^{-(r+\beta+\kappa-\alpha_3)t}\Delta^o$. Moreover, since $\alpha_3 < 0$, $e^{-rt}\Delta_t^{\alpha,*} \to 0$ as $t \to \infty$. Now the posterior belief process p_t^α is the solution to the SDE

$$dp_t = (\alpha_1 - \alpha_2 + \alpha_3 p_t - \kappa p_t)dt + \beta\sigma_2 d\overline{Z}_t^0,\ t > 0,$$

i.e. $(p_t^\alpha)_{t\geq 0}$ is mean reverting around $\frac{\alpha_1-\alpha_2}{\kappa-\alpha_3}$ at rate $\kappa - \alpha_3 > 0$. As a result, this policy delivers finite utility and, moreover, $\lim_{t\to\infty} \mathbb{E}[p_t^\alpha] = \lim_{t\to\infty} \mathbb{E}[(p_t^\alpha)^2] = 0$. This concludes the proof of the theorem.

□

9 Appendix B: Proofs of Sections 4 and 5

Proof of Theorem 4.4: Take any solution (U, q) satisfying the conditions of the theorem, and suppose that the small players believe that the long-run player is following the strategy $a^*(p_t^*) := g'^{-1}(\beta q(p_t^*))$, $t \geq 0$. Consider the function

$$U(p+\Delta) + [q(p+\Delta) - U'(p+\Delta)]\Delta + \frac{\Gamma}{2}\Delta^2 \tag{45}$$

We will show that, for a suitably chosen Γ, the assumptions in the theorem ensure that this function is an upper bound to the long-run player's payoff under any feasible strategy.

More concretely, given a feasible strategy $\hat{a} := (\hat{a}_t)_{t\geq 0}$, define the process

$$\hat{V}_t := \int_0^t e^{-rs}[h(p_s+\hat{\Delta}_t) - g(\hat{a}_s)]ds + e^{-rt}\left\{U(p_t+\hat{\Delta}_t) + [q(p_t+\hat{\Delta}_t) - U'(p_t+\hat{\Delta}_t)]\hat{\Delta}_t + \frac{\Gamma}{2}\hat{\Delta}_t^2\right\},$$

where $h(p) := u(p, a^*(p))$, and $\hat{\Delta}$ denotes the belief asymmetry process under the pair $(a^*(p_t^*), \hat{a})$. Applying Ito's rule to $\hat{V}$ we obtain

$$\begin{aligned}
\frac{dV_t}{e^{-rt}} &= [h(\hat{p}_t^*) - g(\hat{a}_t)]dt - r\left\{U(\hat{p}_t^*) + [q(\hat{p}_t^*) - U'(\hat{p}_t^*)]\hat{\Delta}_t + \frac{\Gamma}{2}\hat{\Delta}_t^2\right\}dt \qquad (46)\\
&\quad + \underbrace{\left\{U'(\hat{p}_t^*)[-\kappa(\hat{p}_t^* - \eta) - \beta\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}_t^*))] + \frac{1}{2}(\beta\sigma)^2 U''(\hat{p}_t^*)\right\}}_{(i)} dt \\
&\quad + \Delta_t \underbrace{\left\{q'(\hat{p}_t^*)[-\kappa(\hat{p}_t^* - \eta) - \beta\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}_t^*))] + \frac{1}{2}(\beta\sigma)^2 q''(\hat{p}_t^*)\right\}}_{(ii)} dt \\
&\quad - \hat{\Delta}_t \underbrace{\left\{U''(\hat{p}_t^*)[-\kappa(\hat{p}_t^* - \eta) - \beta\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}_t^*))] + \frac{1}{2}(\beta\sigma)^2 U'''(\hat{p}_t^*)\right\}}_{(iii)} dt \\
&\quad + [q(\hat{p}_t^*) - U'(\hat{p}_t^*)][-(\beta+\kappa)\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}_t^*))]dt \\
&\quad + \Gamma\hat{\Delta}_t[-(\beta+\kappa)\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}_t^*))]dt + \text{Brownian term},
\end{aligned}$$

where we have used that $\hat{p}^*_t := p_t + \hat{\Delta}_t$ evolves according to $d\hat{p}^*_t = (-\kappa(\hat{p}^*_t - \eta) + \beta(\hat{a} - a^*(\hat{p}^*_t)) - \beta\hat{\Delta}_t)dt + \beta\sigma dZ_t$. Now, using the (30) and (31) we obtain

$$\begin{aligned}
(i) &= rU(\hat{p}^*_t) - h(\hat{p}^*_t) + g(a^*(\hat{p}^*_t)) + U'(\hat{p}^*_t)[-\beta\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))] \\
(ii) &= \left[r + \beta + \kappa + \beta\frac{da^*(p^*_t)}{dp}\right] q(p^*_t) - h'(\hat{p}^*_t) + q'(\hat{p}^*_t)[-\beta\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))] \\
(iii) &= (r+\kappa)U'(\hat{p}^*_t) - h'(\hat{p}^*_t) + \underbrace{g'(a^*(\hat{p}^*_t))}_{=\beta q(\hat{p}^*_t)}\frac{da^*(p^*)}{dp^*} + U''(\hat{p}^*_t)[-\beta\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))]
\end{aligned}$$

with the last equality coming from the fact that U is three times differentiable. Consequently, and using that g is strongly convex,

$$\begin{aligned}
\frac{dV_t}{e^{-rt}} &= [g(a^*(\hat{p}^*_t)) - g(\hat{a}_t) + g'(a^*(\hat{p}^*_t))(\hat{a}_t - a^*(\hat{p}^*_t))]dt \\
&\quad + \beta[\Gamma + q'(\hat{p}^*_t) - U''(\hat{p}^*_t)]\hat{\Delta}_t(\hat{a}_t - a^*(\hat{p}^*_t))dt \\
&\quad - [\beta(q'(\hat{p}^*_t) - U''(\hat{p}^*_t)) + \Gamma\left(\frac{r}{2} + \beta + \kappa\right)]\hat{\Delta}^2_t dt + \text{Stochastic integral} \\
\Rightarrow \hat{V}_t - \hat{V}_0 &\leq \int_0^t e^{-rs}\left(-\frac{\psi}{2}(\hat{a}_s - a^*(\hat{p}^*_s))^2 + \beta[\Gamma + q'(\hat{p}^*_s) - U''(\hat{p}^*_s)]\hat{\Delta}_s(\hat{a}_s - a^*(\hat{p}^*_s))\right. \\
&\quad \left. - \left[\beta(q'(\hat{p}^*_s) - U''(\hat{p}^*_s)) + \Gamma\left(\frac{r}{2} + \beta + \kappa\right)\right]\hat{\Delta}^2_s\right) ds + \text{Stochastic integral}
\end{aligned}$$

The integrand of the Lebesgue integral is a quadratic form in $(\hat{\Delta}, \hat{a} - a^*(\hat{p}^*))$. Letting $R^*(p) := q'(p) - U''(p)$, this quadratic form will be non-positive whenever Γ is such that

$$\frac{\psi}{2}\left[\beta R^*(\hat{p}^*_t) + \Gamma\left(\frac{r}{2} + \beta + \kappa\right)\right] - \frac{\beta^2[\Gamma + R^*(\hat{p}^*_t)]^2}{4} \geq 0 \tag{47}$$

over the set $\{R^*(p) |\, p \in \mathbb{R}\}$. Observe that given any $\Gamma > 0$, the previous condition will be violated when $|R^*|$ is large. Thus, in order for the previous inequality to hold, R^* must take values in an interval of finite length, and the endpoints of this interval will depend on Γ. Since R^* can potentially take both negative and positive values, we choose $\Gamma > 0$ that maximizes the length of the symmetric interval in which R^* is allowed to take values without violating inequality (47).

More specifically, take any $\Gamma > 0$ and observe that at $R^* = -\Gamma$ the previous equality always holds. Evaluating the left-hand side of (47) at $R^* = \Gamma$, we see that equality will hold if and only if $\Gamma \geq 0$ and $\Gamma \leq \frac{\psi(r+2\kappa+4\beta)}{4\beta^2}$. This allows us to conclude that when $\Gamma = \frac{\psi(r+2\kappa+4\beta)}{4\beta^2}$ and $|U(p) - q(p)| \leq \Gamma$ for all $p \in \mathbb{R}$, $\hat{V}$ is a supermartingale.

A standard localizing argument (which uses the growth conditions (i) in the Theorem) allows us to get rid of the stochastic integral through taking expectations, concluding that

$$\mathbb{E}[e^{-rt}[U(\hat{p}^*_t) + [q(\hat{p}^*_t) - U'(\hat{p}^*_t)]\hat{\Delta}_t + \Gamma\hat{\Delta}^2_t] \leq \underbrace{U(p_0)}_{=\hat{V}_0} - \mathbb{E}\left[\int_0^t e^{-rs}[h(\hat{p}^*_s) - g(\hat{a}_s)]ds\right].$$

The limit conditions (ii) in the Theorem allow us to conclude the the lim sup of the left hand side in the previous expression is larger or equal than zero. Applying the dominated convergence theorem on the right hand side we obtain we conclude that $\mathbb{E}\left[\int_0^t e^{-rs}[h(\hat{p}_s^*) - g(\hat{a}_s)]ds\right]$ converges to $\mathbb{E}[\hat{V}_\infty] := \mathbb{E}\left[\int_0^\infty e^{-rs}[h(\hat{p}_s^*) - g(\hat{a}_s)]ds\right]$. Hence

$$\mathbb{E}[\hat{V}_\infty] = \mathbb{E}\left[\int_0^\infty e^{-rt}[h(\hat{p}_t^*) - g(\hat{a}_t)]ds\right] \leq U(p_0).$$

Now, take any solution $U \in \mathcal{C}^2(\mathbb{R})$ to the ODE (30) satisfying a quadratic growth condition. Moreover, since $(p_t)_{t\geq 0}$ is a mean reverting process or a Brownian martingale, there exists $C > 0$ such that $|U(p_t)| \leq C(1 + p_t^2)$. Therefore, $\mathbb{E}[e^{-rt}U(p_t)] \to 0$ when $t \to \infty$. Standard results yield that U admits the probabilistic (Feynman-Kac) representation

$$U(p) = \mathbb{E}\left[\int_0^\infty e^{-rt}(h(p_s) - g(g'^{-1}(\beta q(p_s))))ds\right]$$

with $dp_t = -\kappa(p_t - \eta)dt + \beta\sigma dZ_t$, $t > 0$, $p_0 = p$ (Pham (2009)). Hence, $U(p_0)$ is an upper bound to the long-run player's payoff, and is attained under a^*. This concludes the proof.

□

In order to prove the existence results in Propositions 5.2 and 5.3 we rely on the following results De Coster and Habets (2006) and Schmitt (1969):

Theorem 9.1. (De Coster and Habets (2006), Theorem II.5.6) *Consider the second order differential equation*

$$u'' = f(t, u, u') \tag{48}$$

with $f : \mathbb{R}^3 \to \mathbb{R}$ a continuous function. Let α, β of class $\mathcal{C}^2(\mathbb{R})$ such that $\alpha \leq \beta$, and consider the set $E = \{(t, u, v) \in \mathbb{R}^3 | \alpha(t) \leq u \leq \beta(t)\}$. Assume that for all $t \in \mathbb{R}$

$$\text{(C1)} \quad \alpha'' \geq f(t, \alpha, \alpha') \quad \textit{and} \quad \beta'' \leq f(t, \beta, \beta').$$

Assume also that for any bounded interval I, there exists a positive continuous function $\phi_I : \mathbb{R}^+ \to \mathbb{R}$ that satisfies

$$\int_0^\infty \frac{sds}{\varphi_I(s)} = \infty, \tag{49}$$

and for all $t \in I$, $(u, v) \in \mathbb{R}^2$ with $\alpha(t) \leq u \leq \beta(t)$, $|f(t, u, v)| \leq \varphi_I(|v|)$. Then (48) has at least one solution $u \in \mathcal{C}^2(\mathbb{R})$ such that $\alpha \leq u \leq \beta$.

Remark 9.2. The proof of this theorem in fact delivers a stronger result for the case in which α and β are uniformly bounded and φ_I is independently of I. In this case, the authors prove the existence of $u \in C^2$ solution to (48) satisfying $\alpha \leq u \leq \beta$ and also that u' is uniformly bounded. See the p. 123 in De Coster and Habets (2006) for proof of the Theorem and the discussion that addresses this remark.

Proof of Proposition 5.2: The incentives equation can be written as

$$q''(p) = \underbrace{\frac{2}{\sigma^2}\left[\left(r + \beta + \kappa + \beta^2 \frac{q'(p)}{g''((g')^{-1}(\beta q(p)))}\right) q(p) + \kappa(p - \eta)q'(p) - h'(p)\right]}_{:=f(p,q,q')}.$$

Let $m := \inf_{p\in R} h'(p)$ and $M := \sup_{p\in R} h'(p)$. Take $A, B \in \mathbb{R}$ and notice that

$$
\begin{aligned}
f(p, A, 0) \leq 0 &\Leftrightarrow (r+\beta+\kappa)A - h'(p) \leq 0 \Leftrightarrow A \leq \frac{m}{r+\beta+\kappa} \\
f(p, B, 0) \geq 0 &\Leftrightarrow (r+\beta+\kappa)B - h'(p) \geq 0 \Leftrightarrow B \geq \frac{M}{r+\beta+\kappa}.
\end{aligned} \tag{50}
$$

Hence, we can try to find a solution taking values in $J := \left[\frac{m}{r+\beta+\kappa}, \frac{M}{r+\beta+\kappa}\right]$.

Since g is twice continuously differentiable and strongly convex, there exists $\psi > 0$ such that $g''(\cdot) \geq \psi$. Hence, for bounded interval $I \subset \mathbb{R}$, if $p \in I$ and $u \in J$ we have that

$$
\left|\beta^2 \frac{q'(p)}{g''((g')^{-1}(\beta u))}\right| \leq \frac{\beta^2}{\psi}|q'(p)|.
$$

Consequently, it is say to see that for any bounded interval I we can find constants $\phi_0 > 0$ and $\phi_{1,I} > 0$ such that

$$
|f(p, u, v)| \leq \varphi_I := \phi_0 + \phi_{1,I}|v|,
$$

when $p \in I$ and $u \in J$. Since that the right-hand side satisfies the Nagumo condition (49), Theorem 9.1 ensures the existence of a solution to the incentives equation that is of class $\mathcal{C}^2(\mathbb{R})$ and that takes values in J. Finally, observe that when $\kappa = 0$ we can choose $\phi_{1,I} > 0$ independent of I, so the existence of a solution that in addition has a uniformly bounded derivative is also ensured.

Now we study the asymptotic behavior of any bounded solution to the incentives equation taking values in J. We start showing the both limits exist, and then separate divide the analysis for the cases $\kappa = 0$ and $\kappa > 0$.

Lemma 9.3. *Suppose that both $h'_\infty := \lim_{p\to\infty} h'(p)$ and $h'_{-\infty} := \lim_{p\to-\infty} h'(p)$ exist. Then $q_\infty := \lim_{p\to\infty} q(p)$ and $q_{-\infty} := \lim_{p\to-\infty} q(p)$ exist.*

Proof: Suppose that $\lim_{p\to\infty} q(p)$ does not exist. Then $(q(p))_{p\geq 0}$ has at least two different cluster points c^1 and c^2, one of them different from $\frac{h'_\infty}{r+\beta+\kappa}$. Without loss of generality, assume that $c := \max\{c^1, c^2\} > \frac{h'_\infty}{r+\beta+\kappa}$ and call the respective distance $\delta > 0$. Given $\epsilon < \delta/3$, we can find a sequence $(p_n)_{n\in\mathbb{N}}$ of local maxima of $(q(p))_{p\geq 0}$ such that $q(p_n) > c - \epsilon$ for all $n \geq \bar{N}$, some $\bar{N} \in \mathbb{N}$. But evaluating the incentives equation in the sequence p_n for large n we obtain

$$
\underbrace{q''(p_n)}_{\leq 0} = \frac{2(r+\beta+\kappa)}{(\beta\sigma)^2}\left[q(p_n) - \frac{h'(p_n)}{(r+\beta+\kappa)}\right] > \delta/3
$$

where the right-most inequality comes from the fact that for large n, $|h'(p_n) - h'_\infty| < \epsilon(r+\beta+\kappa)$. This is a contradiction. The case in which $c := \min\{c^1, c^2\} < \frac{h_\infty}{r+\beta+\kappa}$ is analogous if we construct a sequence of local minima. Consequently, $\lim_{t\to\infty} q(p)$ exists, and since the argument for the other limit is analogous, $\lim_{t\to-\infty} q(p)$ must exist as well.

□

Now we show that the limits in (34) hold:

Case $\kappa = 0$**:** Let $\beta(0)$ denote the sensitivity of beliefs to new information evaluated at $\kappa = 0$. Suppose that $q(p)$ converges to some $L \neq \frac{h'_\infty}{r+\beta(0)}$ as $p \to \infty$. If this convergence is monotone, then $q'(p)$ and $q''(p)$ must converge to zero, so

$$\frac{\beta\sigma^2}{2}q''(p) - \beta^2\frac{q(p)q'(p)}{g''((g')^{-1}(\beta q(p)))} \to 0, \text{ as } q(p) \text{ is uniformly bounded}$$

and

$$\lim_{p\to\infty} -h'(p) + (r + \beta(0))q(p) \neq 0.$$

Thus, the incentives equation would not hold for p large enough, a contradiction.

Suppose now that this convergence is not monotone, so $q(p)$ oscillates as it converges to L. If $L > \frac{h'_\infty}{r+\beta(0)}$ (which can occur only when $h'_\infty < M$), we can find a sequence of local maxima $(p_n)_{n\in\mathbb{N}}$ such that $q'(p_n) = 0$, $q''(p_n) \leq 0$ and

$$q''(p_n) = \frac{2}{\beta\sigma^2}\left[-h'(p_n) + (r + \beta(0))q(p_n)\right].$$

But since $(r+\beta(0))q(p_n)$ converges to $L(r+\beta(0)) > h'_\infty$, the incentives equation is violated for n large enough, a contradiction. Equivalently, if $L < \frac{h'_\infty}{r+\beta(0)}$ (which can occur only when $h'_\infty > m$), we can find a sequence of minima such that an analogous contradiction holds. Thus, $q(p)$ must converge to $\frac{h'_\infty}{r+\beta(0)}$. The case $p \to -\infty$ is identical.

Case $\kappa > 0$: We show that (34) holds in a sequence of steps.

Step 1: $\lim_{p\to\infty} q'(p) = \lim_{p\to-\infty} q'(p) = 0$ and q' is uniformly bounded. We show that the first limit exists (for the other limit the argument is analogous). Notice that q' cannot diverge; otherwise, q becomes unbounded, a contradiction. Instead, suppose that $(q'(p))$ has at least two cluster points c^1 and c^2, and that $c := \max\{c^1, c^2\} > 0$ (otherwise, it must be the case that $\min\{c^1, c^2\} < 0$, and the argument is identical). In this case, we can find a sequence of local maxima of $(p_n)_{n\in\mathbb{N}}$ of q' such that $q'(p_n) > c - \epsilon > 0$ for large n. Then, $q''(p_n) = 0$, so the left-hand side of the incentives equation is identically equal to zero, but the right-hand side diverges when $\kappa > 0$, as $p_n q'(p_n) \to \infty$. Hence, $q'(p)$ must converge. Clearly, it must converge to zero; otherwise, $q(p)$ becomes unbounded, a contradiction.

Finally, notice that from here we can conclude that q' is uniformly bounded, as $|q'|$ converges to zero asymptotically in either direction and it is continuous.

Step 2: $\lim_{p\to\infty} pq'(p) = \lim_{p\to-\infty} pq'(p) = 0$. Notice that $\lim_{p\to\infty} pq'(p)$ either exists or takes value $+\infty$. The latter cannot be true, as the incentives equation would imply that $\lim_{p\to\infty} q''(p) = +\infty$, implying that q' diverges, a contradiction. Suppose that $\lim_{p\to\infty} pq'(p) = L > 0$. Then, given $\epsilon > 0$ small and p_0 large enough, we have that for $p > p_0$

$$q'(p) > \frac{L-\epsilon}{p} > 0 \Rightarrow q(p) > q(p_0) + (L - \epsilon)\log(p/p_0),$$

which implies that $q(p) = O(\log(p))$, a contradiction. The case $L < 0$ is analogous, allowing us to conclude that $\lim_{p\to\infty} pq'(p) = 0$. Finally, the analysis for limit $\lim_{p\to-\infty} pq'(p) = 0$ is identical.

Step 3: $\lim_{p\to\infty} q''(p) = \lim_{p\to-\infty} q''(p) = 0$. Using Step 1 and Step 2, the incentives equation implies that $\lim_{p\to-\infty} q''(p)$ exists. But if this limits is not zero, then q' diverges as $p \to \infty$, as $q'(p) = O(p)$, a contradiction. Hence, $\lim_{p\to-\infty} q''(p) = 0$. The other limit is analogous.

Since $q'(p)$, $pq'(p)$ and $q''(p)$ converge all to zero as $p \pm \infty$, the incentives equation implies that

$$0 = \lim_{p\to\pm\infty} q''(p) = \lim_{p\to\pm\infty} [(r+\beta+\kappa)q(p) - h'(p)],$$

from where we conclude.

$\square$

Proof of Proposition 5.3: We first show that, given q a bounded solution to the incentives equation, there exists a solution to (30) satisfying a quadratic growth conditions; for this purpose we apply Theorem 9.1. Then we apply the Feynman-Kac probabilistic representation theorem to show that the unique solution to (30) satisfying a quadratic growth and a transversality condition is precisely the long-run player's on-path payoff. Finally, we show via first principles that the long-run player's payoff satisfies a linear growth condition, and that it has a uniformly bounded derivative when q' is uniformly bounded.

Let $\alpha(p) = -A - Bp^2$. It is easy to see that given any $A, B > 0$ for every bounded interval I we can find constants $\phi_{0,I}, \phi_{1,I} > 0$ such that

$$\underbrace{\frac{2}{\sigma^2}|-h(p) + g((g')^{-1}(\beta q(p))) + \kappa v(p-\eta) + ru|}_{:=f(p,u,v)} \le \phi_{0,I} + \phi_{1,I}|v| := \varphi_I(|v|)$$

$(u, v) \in \mathbb{R}^2$ is such that $|u| \le A + Bp^2$, $p \in I$. Observe that the right hand side satisfies the Nagumo condition (49).

Now, since $G := \sup_{p\in\mathbb{R}} |g((g')^{-1}(\beta q(p)))| < \infty$,

$$\underbrace{-h(p) + g((g')^{-1}(\beta q(p_t))) - \kappa\alpha'(p)(p-\eta) - r\alpha(p)}_{:=\frac{(\beta\sigma)^2}{2}f(p,-\alpha(p),-\alpha'(p))} \le C(1+|p|) + G - 2B\kappa(p-\eta) - r(A+Bp^2)$$

where we have also used that $\|h'\|_\infty < \infty$ implies that h satisfies a linear growth condition (i.e. there exists $C > 0$ such that $|h(p)| \le C(1+|p|)$ for all $p \in \mathbb{R}$). Consequently,

$$\begin{aligned} & C(1+|p|) + G - 2B\kappa(p-\eta) - r(A+Bp^2) \le -\frac{\beta\sigma^2}{2}\alpha''(p) = -B\beta\sigma^2 \\ \Leftrightarrow \quad & H(p) := \underbrace{\left(C + G + B\beta\sigma^2 + 2B\kappa\eta - rA\right)}_{(1)} + \underbrace{\left(C|p| - 2B\kappa p - rBp^2\right)}_{(2)} \le 0, \forall p \in \mathbb{R}. \end{aligned}$$

Given any $B > 0$, (1) ≤ 0 is guaranteed to hold when $rA \ge C + G + 2B\beta\sigma^2/2 + 2B\kappa\eta$. Thus, if $\kappa > 0$, (2) ≤ 0 will be automatically satisfied if $2B\kappa > C \Leftrightarrow B > C/2\kappa$, yielding

$H(\cdot) \leq 0$. If instead $\kappa = 0$, (2) will be violated for $|p|$ small, but choosing B sufficiently large, and then A satisfying the previous inequality with enough slackness ensures that $H(p) \leq 0$ for all $p \in \mathbb{R}$.

For $\nu(p) = -\alpha(p)$, notice that

$$\underbrace{-h(p) + g((g')^{-1}(\beta q(p_t))) + \kappa\nu'(p)(p-\eta) + r\nu(p)}_{:=\frac{(\beta\sigma)^2}{2}f(p,\nu(p),\nu'(p))} \geq -C(1+|p|)-G+2B\kappa(p-\eta)+r(A+Bp^2).$$

So imposing, $\frac{\beta\sigma^2}{2}\nu''(p) = B\beta\sigma^2 \leq -C(1+|p|) - G + 2B\kappa(p-\eta) + r(A+Bp^2)$ yields the exact same condition found for α. Consequently, if we choose A, B satisfying the conditions above, α and ν are lower and upper solutions, respectively. Thus, there exist a $U \in C^2(\mathbb{R})$ solution to (30) such that $|U(p)| \leq \nu(p)$, which means that U satisfies a quadratic growth condition. Finally, the fact that $\kappa \geq 0$ and that U has quadratic growth ensures that $\mathbb{E}[e^{-rt}U(p_t)] \to 0$ as $t \to 0$. Thus, the probabilistic representation follows from the Feynman-Kac formula in infinite horizon (see Pham (2009) Remark 3.5.6.).

We conclude the proof by showing that if q' is uniformly bounded, (i) U' is uniformly bounded and that (ii) U satisfies a linear growth condition. For $p \in \mathbb{R}$ and $h > 0$ let $p_t^h := e^{-\kappa t}(p+h) + (1-e^{-\kappa t})\eta + \sigma\int_0^t e^{-\kappa(t-s)}dZ_s$, that is, the common belief process starting from $p_0 = p + h$, $h \geq 0$. Notice that $p_t^h - p_t^0 = e^{-\kappa t}h$ for all $t \geq 0$, so

$$\begin{aligned}|U(p+h) - U(p)| &\leq \mathbb{E}\left[\int_0^\infty e^{-rt}(|h(p_t^h) - h(p_t^0)| + |g(g'^{-1}(\beta q(p_t^h))) - g(g'^{-1}(\beta q(p_t^0)))|)dt\right] \\ &\leq \frac{(\|h'\|_\infty + R)h}{r}, \text{ for some } R > 0,\end{aligned}$$

where we have used that q' is uniformly bounded in $\mathbb{R}$ and that $g(g'^{-1}(\cdot))$ is Lipschitz over the set $[\frac{\beta m}{r+\beta+\kappa}, \frac{\beta M}{r+\beta+\kappa}]$. Hence, U' is uniformly bounded.

Finally, it is easy to see that if $h'(p)$ is uniformly bounded, then h' satisfies a linear growth condition. Also, since $q(\cdot)$ is uniformly bounded, $G := \sup_{p\in\mathbb{R}} g(g'^{-1}(\beta q(p))) < \infty$. When $\kappa > 0$, $p_t = e^{-\kappa t}p_0 + \kappa\eta\int_0^t e^{-\kappa(t-s)}ds + \sigma\int_0^t e^{-\kappa(t-s)}dZ_s$, so

$$|U(p_0)| \leq \mathbb{E}\left[\int_0^\infty e^{-rt}C\left(1 + \kappa\eta t + |p_0| + \left|\int_0^t e^{-\kappa(t-s)}dZ_s\right|\right) + G)dt\right]$$

But since $\int_0^t e^{-\kappa(t-s)}dZ_s \sim \mathcal{N}(0, \frac{1-e^{-2\kappa t}}{2\kappa})$, the random part in the right-hand side of the previous expression has finite value. When $\kappa = 0$ the same is true, as $Z_t = \sqrt{t}Z_1$ in distribution. Consequently, there exists $K > 0$ such that $|U(p_0)| \leq K(1 + |p_0|)$.

□

Proof of Theorem 5.4: Suppose that $\kappa = 0$ and take any bounded solution q to the incentives equation satisfying that q' is uniformly bounded. Denote by $\hat{\Delta}$ the belief asymmetry process when the long-run player follows the feasible strategy $\hat{a} := (\hat{a}_t)_{t\geq 0}$ and the small players construct beliefs using $a^*(p_t^*) := g'^{-1}(\beta q(p_t^*)), t \geq 0$. Observe that such $\hat{\Delta} := (\hat{\Delta}_t)_{t\geq 0}$ exists. In fact, since q and q' are uniformly bounded, the two-dimensional SDE with random coefficients

$$\begin{aligned}dp_t^* &= [-\kappa(p_t^* - \eta) + \beta(\hat{a}_t - a_t^*(p_t^*)) + \beta(p_t - p_t^*)]dt + \sigma dZ_t^Y \\ dp_t &= -\kappa(p_t - \eta)dt + \sigma dZ_t^Y\end{aligned}$$

has a drift and volatility that are globally Lipschitz, and both grow at most linearly in $\|(p, p^*)\|$ (with a slope that is independent of $\hat{a}$). Also, $\hat{a}$ satisfies the integrability condition $\mathbb{E}[\int_0^t a_s^2 ds] < \infty$, $t \geq 0$. Consequently, there exists a unique strong solution (p, p^*) to the previous SDE (Theorem 1.3.15 in Pham (2009)). Hence, there is a unique solution for the dynamic of Δ. Now that we know that the set of feasible strategies is non-empty, I prove the theorem in two steps:

Step 1: Conditions (i) and (ii) in Theorem 4.4 hold. Observe that from Proposition 5.3, $U(\cdot)$ has a linear growth condition (hence, quadratic growth condition), and U' is uniformly bounded, so the linear growth condition holds trivially. Also since q uniformly bounded, the linear growth condition holds as well. Thus, (i) in Theorem 4.4 holds.

As for condition (ii), we first show that $\lim_{t\to\infty} e^{-rt}\mathbb{E}[\hat{\Delta}_t] = 0$. Observe that

$$\begin{aligned} \hat{\Delta}_t &= e^{-\beta t} \underbrace{\hat{\Delta}_0}_{=0} + \beta \int_0^t e^{-\beta(t-s)}[\hat{a}_s - a^*(p_s + \hat{\Delta}_s)]ds \\ \Rightarrow |\hat{\Delta}_t| &\leq \underbrace{\beta \int_0^t e^{-\beta(t-s)}|\hat{a}_s|ds}_{I_t:=} + \underbrace{\beta \int_0^t e^{-\beta(t-s)}|a^*(p_s + \hat{\Delta}_s)|ds}_{J_t:=}. \end{aligned} \tag{51}$$

Since $q(\cdot)$ is uniformly bounded there exists $K_1 > 0$ such that $J_1 \leq K_1(1 - e^{-\beta t})$. As for I_1 notice that

$$\begin{aligned} &\int_0^t e^{-\beta(t-s)}|\hat{a}_s|ds \leq \left(\int_0^t e^{-2\beta(t-s)}ds\right)^{1/2} \left(\int_0^t \hat{a}_s^2 ds\right)^{1/2} \\ \Rightarrow \quad &|e^{-rt}\mathbb{E}[\hat{\Delta}_t]| \leq \underbrace{e^{-rt}|K_1(1 - e^{-\beta t})|}_{\to 0 \text{ as } t\to\infty} + \underbrace{\left(e^{-rt}\int_0^t e^{-2\beta(t-s)}ds\right)^{1/2}}_{\to 0 \text{ as } t\to\infty} \left(e^{-rt}\int_0^t \mathbb{E}[\hat{a}_s^2]ds\right)^{1/2} < \infty \end{aligned}$$

for all $t \geq 0$, where the right-most inequality comes form the fact that $\hat{a} \in \mathcal{L}^2$. But

$$e^{-rt}\int_0^t \mathbb{E}[\hat{a}_s^2]ds < \int_0^t e^{-rs}\mathbb{E}[\hat{a}_s^2]ds < \int_0^\infty e^{-rs}\mathbb{E}[\hat{a}_s^2]ds < \infty$$

because otherwise, by strong convexity, the total cost of manipulation would be $+\infty$, a contradiction. Hence, $e^{-rt}\mathbb{E}[\hat{\Delta}_t] \to 0$ as $t \to \infty$.

With this in hand, it is easy to show that all the limits in (ii) holds. This is because $|U(p_t + \hat{\Delta}_t)| \leq C_1(1 + |p_t| + |\hat{\Delta}_t|)$, $|q(p_t + \hat{\Delta}_t)\hat{\Delta}_t| \leq C_2|\hat{\Delta}_t|$ and $|U'((p_t + \hat{\Delta}_t)\hat{\Delta}_t| \leq C_3|\hat{\Delta}_t|$, for some constants C_1, C_2 and C_3 all larger than zero.

Step 2: Condition (iii) in Theorem 4.4 holds. Recall that the long-run player's payoff (35)

$$\mathbb{E}\left[\int_0^\infty e^{-rt}[h(p_t) - g((g')^{-1}(\beta q(p_t)))]dt\right] =: U(p)$$

with $dp_t = -\kappa(p_t - \eta)dt + \sigma dZ_t$, $t > 0$, and $p_0 = p$, is the unique C^2 solution to the ODE (30) satisfying a quadratic growth condition. Because the right-hand side of that ODE is differentiable, U is three times differentiable. Hence, U' satisfies the following ODE in $p \mapsto f(p)$:

$$f''(p) = \frac{2}{\sigma^2}\left[-h'(p) + \beta\frac{q(p)q'(p)}{g''(g'^{-1}(\beta q(p)))} + rf(p)\right],\ p \in \mathbb{R}. \tag{52}$$

Claim: The ODE (52) has a unique solution that is uniformly bounded. Hence, such solution corresponds to $U'(p)$.

Proof of the claim: Consider two bounded solutions f_1 and f_2 of the ODE (52), and suppose that there exists p_0 such that $f_1(p_0) > f_2(p_0)$. If $f_1'(p_0) \geq f_2'(p_0)$, then $f_1''(p_0) > f_2''(p_0)$, which implies that $f_1' - f_2'$ is strictly increasing at p_0. But then, the difference $f_1 - f_2$ increases (strictly) in a neighborhood to the right of p_0. Applying the same argument to the right of t_0 allows us to conclude that $f_1' - f_2$ is strictly increasing on the set (p_0, ∞), so the difference $f_1 - f_2$ becomes unbounded, a contradiction. If instead $f_1'(p_0) < f_2'(p_0)$ an analogous argument applies over the set $(-\infty, p_0)$.

□

The previous claim allows us to determine U' through finding the unique bounded solution to (52). In fact, it is easy to see that

$$U'(p) = -\frac{1}{2\sqrt{\nu}}\left[\int_{-\infty}^{p} e^{-\sqrt{\nu}(p-y)}\ell(y)dy + \int_{p}^{\infty} e^{-\sqrt{\nu}(y-p)}\ell(y)dy\right]. \tag{53}$$

where $\nu := 2r/\sigma^2$ and $\ell(y) := \frac{2}{\sigma^2}\left[-h'(y) + \frac{\beta q(y)q'(y)}{g''(g'^{-1}(\beta q(y)))}\right]$ is a bounded solution to (52) – hence the derivative of the long-run player's on-path utility. Now we derive an analytic expression for $U'' - q'$. Differentiating (53) yields

$$U''(p) = \frac{1}{2}\int_{-\infty}^{p} e^{-\sqrt{\nu}(p-y)}\ell(y)dy - \frac{1}{2}\int_{p}^{\infty} e^{-\sqrt{\nu}(y-p)}\ell(y)dy.$$

Using the incentives equation we obtain that

$$\ell(p) := \frac{2}{\sigma^2}\left[-h'(p) + \frac{\beta q(p)q'(p)}{g''(g'^{-1}(\beta q(p)))}\right] = q''(p) - \frac{2(r+\beta)}{\sigma^2}q(p).$$

and straightforward calculations yield that

$$U''(p) = q'(p) + \underbrace{\frac{2\beta}{\sigma^2}\left[-\frac{1}{2}\int_{-\infty}^{p} e^{-\sqrt{\nu}(p-y)}q(y)dy + \frac{1}{2}\int_{p}^{\infty} e^{-\sqrt{\nu}(y-p)}q(y)dy\right]}_{-x(p):=}. \tag{54}$$

It is easy to see that $x(p)$ is twice continuously differentiable, that satisfies a quadratic growth condition (in fact, it is bounded), and also that $\mathbb{E}[e^{-rt}x(p_t)] \to 0$ as $t \to \infty$. Moreover, it satisfies the ODE $x''(p) = \frac{2}{\sigma^2}[-\beta q(p) + rx(p)]$. The Feynman-Kac formula says then that

$$U''(p) - q'(p) =: -x(p) = -\mathbb{E}\left[\int_0^{\infty} e^{-rs}\beta q(p_s)ds \Big| p_0 = p\right].$$

Using (54), and recalling that $q(\cdot) \in \left[\frac{m}{r+\beta}, \frac{M}{r+\beta}\right]$, that $\nu = 2r/\sigma^2$ and that $\sigma = \beta\sigma_\xi$, we get

$$|U''(t) - q'(t)| \leq \frac{M-m}{(r+\beta)\sqrt{2r\sigma_\xi^2}},\ t \geq 0.$$

But since $\beta = \sigma_\theta/\sigma_\xi$ when $\kappa = 0$, we have that (32) in Theorem 4.4 will hold if

$$\frac{M-m}{(r+\beta)\sqrt{2r\sigma_\xi^2}} \leq \left.\frac{\psi(r+4\beta+2\kappa)}{4\beta^2}\right|_{\kappa=0} \Leftrightarrow \frac{M-m}{\psi} \leq \frac{\sqrt{2r\sigma_\xi^2}(r\sigma_\xi+4\sigma_\theta)(r\sigma_\xi+\sigma_\theta)}{4\sigma_\theta^2}.$$

Since condition (37) is tighter than the one just derived, condition (iii) in Theorem 4.4 holds. This concludes the proof.

□

10 Appendix C: Proofs of Section 6

Proof of Theorem 6.2: In order to prove Theorem 6.2, I verify that the conditions of Theorem 4.2 hold. That is, for suitably chosen coefficients:

1. $V(p,\Delta) = \alpha_0 + \alpha_1 p + \alpha_2\Delta + \alpha_3 p\Delta + \alpha_4 p^2 + \alpha_5\Delta^2$ solves the PDE (23)-(24);
2. For any feasible strategy $\hat{a} := (\hat{a}_t)_{t\geq 0}$, $\limsup_{t\to\infty} \mathbb{E}[e^{-rt}V(p_t, \hat{\Delta}_t^*)] \geq 0$, where $\hat{\Delta}_t^*$ denotes the belief asymmetry process under $\hat{a}$ and $a^*(p_t^*) = \frac{\beta}{\psi}V_\Delta(p_t^*, 0)$, $t \geq 0$.
3. Under the Markovian control $\alpha(p,\Delta) := \frac{\beta}{\psi}V_\Delta(p,\Delta)$ and the Markovian conjecture $a^*(p+\Delta) := \frac{\beta}{\psi}V_\Delta(p+\Delta, 0)$ the ODE

$$d\Delta_t = [-(\beta+\kappa)\Delta_t + \beta(\alpha(p_t,\Delta_t) - a^*(p_t+\Delta_t))]dt \tag{55}$$

has a unique solution $\Delta^{\alpha,*} := (\Delta_t^{\alpha,*})_{t\geq 0}$ a.s. satisfying $\liminf_{t\to\infty} \mathbb{E}[e^{-rt}V(p_t, \Delta_t^{\alpha,*})] \leq 0$, and the processes $\alpha_t = \frac{\beta}{\psi}V_\Delta(p_t, \Delta_t^{\alpha,*})$ and $a_t^* := \frac{\beta}{\psi}V_\Delta(p_t + \Delta_t^{\alpha,*}, 0))$ are feasible.

Before proving 1-3, I make some important observations:

(i) Given $a^*(p^*) := \frac{\beta}{\psi}[\alpha_2 + \alpha_3 p^*]$ and any feasible strategy $\hat{a}$, there exists a unique solution $\hat{\Delta}^* := (\hat{\Delta}_t^*)_{t\geq 0}$ to (55). In fact, the two-dimensional SDE with random coefficients

$$\begin{aligned} dp_t^* &= [-\kappa(p_t^*-\eta) + \beta(\hat{a}_t - a_t^*(p_t^*)) + \beta(p_t - p_t^*)]dt + \sigma dZ_t^Y \\ dp_t &= -\kappa(p_t-\eta)dt + \sigma dZ_t^Y \end{aligned} \tag{56}$$

has a drift and volatility that are globally Lipschitz, and both grow at most linearly in $\|(p,p^*)\|$ (with a slope that is independent of $\hat{a}$). Also, $\hat{a}$ satisfies the integrability condition $\mathbb{E}[\int_0^t a_s^2 ds] < \infty$, $t \geq 0$. Consequently, there exists a unique strong solution (p,p^*) to the previous SDE (Theorem 1.3.15 in Pham (2009)). Hence, there is a unique solution for the dynamic of Δ.

(ii) Given the conjecture $a^*(p^*) = \frac{\beta}{\psi}[\alpha_2 + \alpha_3 p^*]$, the set of feasible strategies under which the long-run player's payoff is finite is non-empty, so the optimization problem is well defined. Furthermore, $(a^*(p_t^*))_{t\geq 0}$ is a feasible strategy, and the long-run player attains finite utility when he follows a^*.

(iii) Since $dp_t = -\kappa(p_t - \eta)dt + \beta\sigma_\xi dZ_t^Y$, $\mathbb{E}[p_t]$ is uniformly bounded and $\mathbb{E}[p_t^2]$ grows at most linearly in t.

Step 1: It is easy to see that $V(p,\Delta) = \alpha_0 + \alpha_1 p + \alpha_2\Delta + \alpha_3 p\Delta + \alpha_4 p^2 + \alpha_5\Delta^2$ satisfies the PDE (23)-(24)

$$\begin{aligned} rV(p,\Delta) &= \max_a \Big\{ u_0 + u_1(p+\Delta) - u_2(p+\Delta)^2 - \frac{\psi}{2}a^2 \\ &\quad + -\kappa[p-\eta]V_p(p,\Delta) + \frac{1}{2}\beta^2\sigma_\xi^2 V_{pp}(p,\Delta) \\ &\quad + \left[-\Delta(\beta+\kappa) + \beta(a - a^*(p+\Delta))\right] V_\Delta(p,\Delta) \Big\} \\ s.t. &\quad a^*(p) \in \arg\max_{a\in A} \{\beta V_\Delta(p,0)a - \frac{\psi}{2}a^2\}. \end{aligned} \tag{57}$$

if and only if

$$\begin{aligned} (0) : 0 &= r\alpha_0 - u_0 - \eta\kappa\alpha_1 + \frac{1}{2\psi}\beta^2\alpha_2^2 - \beta^2\sigma_\xi^2\alpha_4 \\ (1) : 0 &= r\alpha_1 - u_1 + \kappa\alpha_1 + \frac{\beta^2\alpha_3}{\psi}\alpha_2 - 2\eta\kappa\alpha_4 \\ (2) : 0 &= \left(r + \kappa + \beta + \frac{\beta^2\alpha_3}{\psi}\right)\alpha_2 - \eta\kappa\alpha_3 - u_1 \\ (3) : 0 &= (r + \beta + 2\kappa)\alpha_3 + \frac{\beta^2}{\psi}\alpha_3^2 + 2u_2. \\ (4) : 0 &= r\alpha_4 + \frac{1}{2\psi}\beta^2\alpha_3^2 + 2\kappa\alpha_4 + u_2 \\ (5) : 0 &= \left(r + 2\left[\kappa + \beta + \frac{\beta^2\alpha_3}{\psi}\right]\right)\alpha_5 - \frac{2\beta^2}{\psi}\alpha_5^2 + u_2. \end{aligned}$$

Notice that equations (2) and (3) are satisfied by α_2 and α_3 as in (39) and (40), respectively. Moreover, given α_3, equations (0), (1),(2) and (4) have a unique solution. For equation (5), we choose its unique negative root

$$\alpha_5 = \frac{r + 2(\beta+\kappa) + \frac{2\alpha_3\beta^2}{\psi} - \sqrt{\left(r + 2(\beta+\kappa) + \frac{\alpha_3\beta^2}{\psi}\right)^2 + \frac{8\beta^2 u_2}{\psi}}}{4\beta^2}.$$

Step 2: It suffices to show that $\limsup_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t,\hat{\Delta}_t^*)] = 0$ for any feasible strategy $\hat{a}$ that attains finite utility. This is done in two Lemmas and concludes in Corollary 10.3.

Lemma 10.1. *Let $\hat{a}$ be a feasible strategy under which the the long-run player attains finite utility. Then, there are positive constants C_1 and $C_2(\hat{a})$ such that*

$$|\mathbb{E}[\hat{\Delta}_t^*]| < C_1[1 + e^{-(\frac{\beta^2\alpha_3}{\psi}+\beta+\kappa)t}] + C_2(a)[e^{rt}(1 + e^{-2(\frac{\beta^2\alpha_3}{\psi}+\beta+\kappa)t})]^{1/2}. \tag{58}$$

As a result, $\lim_{t\to\infty} e^{-rt}\mathbb{E}[\hat{\Delta}_t^*] = 0.$

Proof: Take any such strategy. Under the small players' conjecture a^* we can write

$$\hat{\Delta}^*_t = \Delta^o e^{-(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})t} + \int_0^t e^{(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})(s-t)}[\beta\hat{a}_s - \beta^2(\alpha_2+\alpha_3 p_s)]ds, t \geq 0.$$

Using this and the fact that $\mathbb{E}[p_t]$ is uniformly bounded, we find C_1 s.t.

$$|\mathbb{E}[\hat{\Delta}^*_t]| \leq C_1[1+e^{-(\frac{\beta^2\alpha_3}{\psi}+\beta+\kappa)t}] + \int_0^t e^{(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})(s-t)}\beta\mathbb{E}[|\hat{a}_s|]ds.$$

Now,

$$I := \int_0^t e^{(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})(s-t)}\mathbb{E}[|\hat{a}_s|]ds < \left(e^{rt}\int_0^t e^{2(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})(s-t)}ds\right)^{1/2}\left(e^{-rt}\int_0^t \mathbb{E}[\hat{a}_s^2]ds\right)^{1/2}$$

where in the last inequality we used Cauchy-Schwarz's and Jensen's inequalities. But $e^{-rt}\int_0^t \mathbb{E}[\hat{a}_s^2]ds < C(\hat{a}) := \int_0^\infty e^{-rs}\mathbb{E}[\hat{a}_s^2]ds$, which is finite since flow payoffs are bounded by above and a attains finite utility. Therefore $I \leq C_2(a)[e^{rt}(1+e^{-2(\frac{\beta^2\alpha_3}{\psi}+\beta+\kappa)t})]^{1/2}$ for some positive constant $C_2(\hat{a})$. This proves (58).

Finally, from α_3's definition it is easy to see that $\beta+\kappa+\frac{\beta^2\alpha_3}{\psi}+r>0$. This inequality and the previously found bound, yield $\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[\hat{\Delta}^*_t]=0$.

□

Lemma 10.2. *Let $\hat{a}$ be a feasible strategy under which the long-run player attains finite utility. Then, $\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[p_t\hat{\Delta}^*_t]=0$. As a consequence, $\liminf\limits_{t\to\infty} e^{-rt}\mathbb{E}[(\hat{\Delta}^*_t)^2]=0$.*

Proof: Applying Ito's rule to $e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)t}p_s\hat{\Delta}^*_t$ we obtain the expression

$$\begin{aligned} p_t\hat{\Delta}^*_t &= e^{-(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)t}p^o\Delta^o + \underbrace{\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}\hat{\Delta}^*_s[\kappa\eta ds+\beta\sigma_\xi dZ^Y_s]}_{I_t:=} \\ &+ \underbrace{\beta\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}p_s\hat{a}_s ds}_{J_t:=} - \frac{\beta^2}{\psi}\underbrace{\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}[\alpha_2 p_s+\alpha_3 p_s^2]ds}_{K_t:=} \end{aligned}$$

Because $\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa+r>0$ the first term in the right-hand side of the previous expression goes to zero when discounted at rate r. The same occurs with the last term, as $\mathbb{E}[\alpha_2 p_s+\alpha_3 p_s^2]]$ grows at most linearly in t.

Now, since $\hat{a}$ and a^* satisfy the integrability condition $\mathbb{E}\left[\int_0^t |x_s|^2 ds\right]<\infty$, it is also the case that $\mathbb{E}\left[\int_0^t |\hat{\Delta}^*_s|^2 ds\right]<\infty$ for all $t\geq 0$.[35] Hence, the stochastic integral has zero mean, yielding $\mathbb{E}[I_t]=\eta\kappa\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}\mathbb{E}[\hat{\Delta}^*_s]ds,\ t\geq 0$.

[35] Recall that $\hat{\Delta}^*_t = \Delta^o e^{-(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})t} + \int_0^t e^{(\beta+\kappa+\frac{\beta^2\alpha_3}{\psi})(s-t)}[\beta\hat{a}_s - \beta^2(\alpha_2+\alpha_3 p_s)]ds$, from where it is easy to conclude.

Using the bound (58) there exists $C_3(\hat{a}) < \infty$ such that

$$\begin{aligned}\frac{|\mathbb{E}[I_t]|}{C_3(\hat{a})} &\leq \int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}[1+e^{-(\frac{\beta^2\alpha_3}{\psi}+\beta+\kappa)s}]ds \\ &\quad + \int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}[e^{rs}(1+e^{-2(\frac{\beta^2\alpha_3}{\psi}+\beta+\kappa)s})]^{1/2}ds \\ &\leq \int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}ds + e^{-(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)t}\frac{e^{\kappa t}-1}{\kappa} \\ &\quad + e^{-(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)t}\int_0^t [e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa+r/2)s}+e^{(\kappa+r/2)s}]ds.\end{aligned}$$

Observing that $\frac{\beta^2\alpha_3}{\psi}+\beta+\nu+r>0$ and $\frac{\beta^2\alpha_3}{\psi}+\beta+\nu+r>\beta^2\alpha_3+\beta+\nu+r/2>0$ for $\nu=\kappa,2\kappa$, we conclude that $e^{-rt}\mathbb{E}[I_t]\to 0$.

As for J_t, applying the Cauchy-Schwartz inequality twice:

$$e^{-rt}|\mathbb{E}[J_t]| \leq \left(e^{-rt}\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}\mathbb{E}[\hat{a}_s^2]ds\right)^{1/2}\left(e^{-rt}\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}\mathbb{E}[p_s^2]ds\right)^{1/2}.$$

Since $(\mathbb{E}[p_t^2])_{t\geq 0}$ is at most linear in t and $\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa+r>0$, the last term in the right-hand side of the previous expression goes to zero as $t\to\infty$. But observe that

$$e^{-rt}\int_0^t e^{(\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa)(s-t)}\mathbb{E}[\hat{a}_s^2]ds < \int_0^\infty e^{-rs}\mathbb{E}[\hat{a}_s^2]ds < \infty,$$

where we used that $\frac{\beta^2\alpha_3}{\psi}+\beta+2\kappa>0$ and the fact that $\hat{a}$ yields finite utility. Thus $\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[p_t\hat{\Delta}^*]=0$.

Finally, since flow payoffs are bounded by above and $\hat{a}$ delivers finite utility, we must have that $\mathbb{E}[\int_0^\infty e^{-rt}u(p_t+\hat{\Delta}_t^*)dt]<\infty$. Hence, $\limsup\limits_{t\to\infty} e^{-rt}\mathbb{E}[u(p_t+\hat{\Delta}_t^*)]\geq 0$. Using that $\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[p_t]=\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[\hat{\Delta}_t^*]=0$, we obtain that

$$\limsup_{t\to\infty} e^{-rt}\mathbb{E}[u(p_t+\hat{\Delta}_t^*)^2]\geq 0 \Rightarrow \liminf_{t\to\infty} e^{-rt}\mathbb{E}[(p_t+\hat{\Delta}_t^*)^2]=0.$$

But since $\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[p_t\hat{\Delta}^*]=0$, $0=\liminf\limits_{t\to\infty} e^{-rt}\mathbb{E}[(p_t+\Delta_t^{a,a^*})^2]\geq \liminf\limits_{t\to\infty} e^{-rt}\mathbb{E}[(\Delta_t^{a,a^*})^2]$.

□

With these 3 lemmas we obtain the following:

Corollary 10.3. *Suppose that the function V is such that $\alpha_5<0$. Then, for any $\hat{a}$ that delivers finite utility,* $\limsup\limits_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t,\hat{\Delta}_t^*)]=0$.

Proof: Using Lemmas 10.1 and 10.2, we have that $\lim\limits_{t\to\infty} e^{-rt}\mathbb{E}[\chi_t]=0$, for $\chi=p,\hat{\Delta}^*,p\hat{\Delta}^*$ and p^2. Thus,

$$\limsup_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t,\hat{\Delta}_t^*)]=\limsup_{t\to\infty} e^{-rt}\alpha_5\mathbb{E}[(\hat{\Delta}_t^*)^2]=\alpha_5\liminf_{t\to\infty} e^{-rt}\mathbb{E}[(\hat{\Delta}_t^*)^2]=0.$$

□

Step 3: It is easy to see that given $\alpha(p,\Delta) := \frac{\beta}{\psi}(\alpha_2 + \alpha_3 p + 2\alpha_5\Delta)$ and $a^*(p,\Delta) := \frac{\beta}{\psi}[\alpha_2 + \alpha_3(p+\Delta)]$, the resulting ODE for Δ has as unique solution the function $\Delta_t^{\alpha,*} = \Delta^o e^{\rho t}$, $t \geq 0$, with $\rho := \frac{(2\alpha_5-\alpha_3)\beta^2}{\psi} - \kappa - \beta$. Using α_5's definition we see that

$$\rho = \frac{r - \sqrt{\left(r + 2(\beta+\kappa) + \frac{2\alpha_3\beta^2}{\psi}\right)^2 + \frac{8\beta^2 u_2}{\psi}}}{2}, \tag{59}$$

yielding that both $\rho - r$ and $2\rho - r$ are strictly less than zero. Consequently, $\lim_{t\to 0} e^{-rt}\mathbb{E}[\Delta_t^{\alpha,*}] = \lim_{t\to 0} e^{-rt}\mathbb{E}[(\Delta_t^{\alpha,*})^2] = 0$, which results in $\lim_{t\to\infty} e^{-rt}\mathbb{E}[V(p_t,\Delta_t^{\alpha,*})] = 0$. Finally, it is straightforward to prove that the processes $\alpha_t := \frac{\beta}{\psi}(\alpha_2 + \alpha_3 p_t + 2\alpha_5\Delta^o e^{\rho t})$ and $a_t^* := \frac{\beta}{\psi}(\alpha_2 + \alpha_3[p_t + \Delta^o e^{\rho t}])$ satisfy the integrability condition $\mathbb{E}[\int_0^t x_s^2 ds]$, and that the long-run player attains finite utility under α, making α a feasible control. This concludes Step 3.

Finally, we must show that any stock of belief asymmetry vanishes asymptotically. Notice that $\rho < 0$ if and only if

$$\begin{aligned}
0 &\leq 4(\beta+\kappa)^2 + \frac{4\beta^4\alpha_3^2}{\psi^2} + 4r(\beta+\kappa) + \frac{4r\beta^2\alpha_3}{\psi} + \frac{8\beta^2(\beta+\kappa)\alpha_3}{\psi} + \frac{8\beta^2 u_2}{\psi} \\
\Leftrightarrow 0 &\underbrace{\leq}_{(*)} (\beta+\kappa)^2 + r(\beta+\kappa) + \frac{\beta^3\alpha_3}{\psi} + \frac{\beta^2}{\psi}\underbrace{[\beta^2\alpha_3^2 + (r+\beta+2\kappa)\alpha_3 + 2u_2]}_{=0,\text{ by definition of } \alpha_3}
\end{aligned}$$

But using the definition of α_3

$$\frac{\beta^3\alpha_3}{\psi} = \beta\frac{-(r+\beta+2\kappa) + \sqrt{(r+\beta+2\kappa)^2 - \frac{8\beta^2 u_2}{\psi}}}{2}$$

we can see that (*) is true.

□

Proof of Proposition 6.3: It is easy to see that the system (p,Δ) is (i) *stabilizable* (the belief-asymmetry process is controllable and private beliefs decay to zero) and that (ii) the system is *detectable* (in the (p,p^*) coordinate system, the "unobserved" component p (i.e. the state variable that does not contribute to the flow payoff) decays to zero). Consequently, the solution of this linear-quadratic regulator problem exists and it is unique, and the value function is quadratic (Theorem 12.3. in Wonham (1985)).

Guessing a solution of the form $V(p,\Delta) = \alpha_3 p\Delta + \alpha_4 p^2 + \alpha_5\Delta^2$, $\alpha_i = \alpha_i(\hat{\alpha}_3)$, $i = 3,4,5$ must satisfy

$$\alpha_3(\hat{\alpha}_3) = \frac{-2u_2 - \frac{2\beta^2}{\psi}\alpha_5(\hat{\alpha}_3)\hat{\alpha}_3}{r + \beta + 2\kappa + \frac{\beta^2}{\psi}\hat{\alpha}_3 - \frac{2\beta^2}{\psi}\alpha_5(\hat{\alpha}_3)} \tag{60}$$

$$\alpha_4(\hat{\alpha}_3) = \frac{-u_2 + \frac{\beta^2\alpha_3(\hat{\alpha}_3)}{2\psi}(\alpha_3(\hat{\alpha}_3) - 2\hat{\alpha}_3)}{r + 2\kappa} \tag{61}$$

$$\alpha_5(\hat{\alpha}_3) = \frac{r + 2\left(\beta+\kappa+\frac{\beta^2}{\psi}\hat{\alpha}_3\right) \pm \sqrt{\left(r + 2\left(\beta+\kappa+\frac{\beta^2}{\psi}\hat{\alpha}_3\right)\right)^2 + \frac{8\beta^2}{\psi}u_2}}{\frac{4\beta^2}{\psi}}. \tag{62}$$

Choose the negative root of $\alpha_5(\hat{\alpha_3})$. Then, $\alpha_3(\hat{\alpha}_3)$ and $\alpha_4(\hat{\alpha}_3)$ are uniquely defined. Moreover, it is easy to show that

$$\rho(\hat{\alpha}_3) := \beta + \kappa + \frac{\beta^2}{\psi}\hat{\alpha}_3 - \frac{2\beta^2}{\psi}\alpha_5(\hat{\alpha}_3) = \frac{1}{2}\left(\sqrt{\left(r + 2\left(\beta + \kappa + \frac{\beta^2}{\psi}\hat{\alpha}_3\right)\right)^2 + \frac{8\beta^2}{\psi}u_2} - r\right),$$

yielding that $r + \rho(\hat{\alpha}_3) > 0$ and $r + 2\rho(\hat{\alpha}_3) > 0$. In particular, the denominator of $\alpha_3(\hat{\alpha}_3)$, $r + \kappa + \rho(\hat{\alpha}_3) > 0$, so $\alpha_3(\hat{\alpha}_3) < 0$.

Under the control $\alpha(p, \Delta) = \alpha_3(\hat{\alpha}_3)p + \alpha_5(\hat{\alpha}_3)\Delta$, the belief-asymmetry process becomes

$$\Delta_t = \Delta_0 e^{-\rho(\hat{\alpha}_3)t} + \frac{\beta^2(\alpha_3(\hat{\alpha_3}) - \hat{\alpha}_3)}{\psi}\int_0^t e^{-\rho(\hat{\alpha}_3)(t-s)}p_s ds$$

with $p_s = p_0 e^{-\kappa s}$, $s \geq 0$. Consequently, $e^{-rt}\Delta_t^2$ converges to zero as $t \to \infty$. Since $e^{-rt}p_t^2$ also converges to zero as $t \to \infty$, the conjectured value function satisfies the transversality conditions. It follows that V as above must be the long-run player's value function.

Using that the denominator of $\alpha_3(\hat{\alpha}_3)$ is strictly positive, we conclude that

$$\alpha_3(\hat{\alpha}_3) < \hat{\alpha}_3 \Leftrightarrow \frac{\beta^2}{\psi}(\hat{\alpha}_3)^2 + (r + \beta + 2\kappa)\hat{\alpha}_3 + 2u_2 > 0,$$

which holds for all $\hat{\alpha}_3 \in \mathbb{R}$ if the curvature condition is violated. This concludes the proof.

$\square$