Two-Sided Learning and Moral Hazard

Gonzalo Cisternas*
MIT Sloan
May 2015

Abstract

I study games of ex-ante symmetric uncertainty about a payoff-relevant state where a long-run player can manipulate the belief of a market by “jamming” a public signal in an additively separable way (as in Holmström, 1999), and where shocks are Brownian. A major contribution is analyzing cases where the long-run player’s flow payoff from this manipulation is nonlinear. I obtain verifiable sufficient conditions for a solution to an ordinary differential equation to characterize behavior in any equilibrium in which (i) the market holds a correct belief at all times, and (ii) actions are Markov in the common belief. Using these conditions, I show the existence of pure-strategy equilibria for a wide range of economic environments. The presence of nonlinearities gives rise to rich dynamics, which include fully endogenous ratchet-like forces. Applications include models of reputation, of earnings management, and of dynamic oligopoly.

Keywords: Learning, imperfect monitoring, ratchet effect, continuous time, Brownian motion.

JEL codes: C73, D82, D83.

*Email: gcistern@mit.edu. I would like to thank Yuliy Sannikov for his invaluable advice, and Dilip Abreu, Alessandro Bonatti, Hector Chade, Anna Costello, Wouter Dessein, Eduardo Faingold, Bob Gibbons, Sambuddha Ghosh, Leandro Gorno, Andrey Malenko, Iván Marinovic, Stephen Morris, Marcin Peski, Tavneet Suri, Juuso Toikka, Mike Whinston and all seminar participants at Columbia, 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 studies dynamic incentives in settings where agents learn about the economic environment. In financial markets for instance, because earnings convey noisy information about firms’ unobserved fundamentals, a manager may have the incentive to manipulate earnings when his compensation is tied to the value of the firm he runs (Stein, 1989). Similarly, in labor markets, because output partly carries noisy information about workers’ unobserved skills, a worker may have the incentive to exert effort when his wage depends on the market’s perception regarding his ability (Holmström 1982/1999). In both settings, the impossibility of observing an agent’s actions (e.g., reporting policy, effort choices) creates incentives for manipulating a counterparty’s belief about a payoff-relevant variable (e.g., a firm’s fundamentals, a worker’s skills). The purpose of this paper is to offer a general framework for studying such manipulation incentives.

In the class of games I analyze, a long-run player (e.g., a manager) and a population of small individuals (e.g., a pool of investors or analysts), or market, share a common prior regarding the initial value of a continuous-time Gaussian process (e.g., a firm’s fundamentals). As the interaction evolves over time, all agents learn about the hidden state from a public signal that is distorted by Brownian shocks. The long-run player can nevertheless influence the market’s belief by taking unobserved actions that affect the evolution of the public signal. Using continuous-time techniques, I characterize the actions that the long-run player takes in any equilibrium in which the market perfectly anticipates his behavior; that is, he becomes trapped into the market’s “demand for manipulation,” which leads the market to hold a correct belief about the hidden state at all times.

Quantifying the incentives that arise in these types of environments is challenging. Consider the earnings management example. If an equilibrium in which beliefs coincide exists, the payoff that the manager obtains by manipulating earnings as conjectured by the market must dominate the payoff under any alternative manipulation policy. Thus, the manager’s equilibrium behavior is precisely determined by the potential benefit of acting differently from the investors’ expectations, thereby inducing them to hold a wrong belief about the firm’s value. This belief divergence off the equilibrium path boils down to the manager holding a private belief about the firm’s fundamentals, and such form of private information is, in almost all instances, payoff-relevant (i.e., likely to affect his future behavior). In a nutshell, performing equilibrium analysis requires tackling a difficult simultaneous moral-hazard and adverse-selection problem off the equilibrium path.

In order to simplify this belief-divergence challenge, the existing literature on games of two-sided learning has relied on key modeling assumptions. For instance, a critical modeling
choice in the manipulation model of Stein (1989) (based on the reputation model of Holmström) is that the manager’s compensation is linear in the market’s belief about the firm’s profitability, which results in a constant degree of manipulation across all possible values of the firm.\textsuperscript{1} However, compensation schemes do exhibit nonlinearities in practice. Even more so, there is strong evidence suggesting that earnings management is more intense at particular thresholds.\textsuperscript{2} The presence of different forms of nonlinearities therefore leads to a number of unanswered questions: will the manager be trapped into matching analysts’ forecasts of earnings (existence of equilibrium)? What are the dynamics of earnings manipulation (economic forces)? What are the corresponding implications on the observed distribution of earnings (testable predictions)? A central methodological contribution of this paper is to offer a general framework in which these types of questions can be answered. Crucially, this can be done only through finding a tractable way to perform equilibrium analysis for situations in which this belief divergence matters for incentives. Such analysis becomes extremely non-trivial precisely when the economic environment exhibits nonlinearities.

In order to isolate the incentives for belief manipulation, I borrow from Holmström (1999) his signal-jamming technology, which washes away any experimentation motive (i.e., affecting the speed of learning). I then explore belief-manipulation dynamics in settings where (i) the long-run player’s payoffs are nonlinear, where (ii) it is costly for him to manipulate the signal, and where (iii) information arrives continuously over time. Inspired by the literature on dynamic contracting, I follow a first-order approach to perform equilibrium analysis: I first construct a necessary condition for equilibrium behavior, and then provide conditions under which such necessary condition is also sufficient. I elaborate next on why this “relaxed” approach is necessary, and then on the relevance of my methods for applied theory.

To illustrate the complexity behind this potential belief divergence, consider a manager who currently manipulates earnings, and suppose that investors are less optimistic about the firm’s prospects. Will the manager engage in more or less manipulation over time? On the one hand, driving up the market’s belief about the firm’s prospects can be beneficial for the manager, as the firm’s market value increases. However, what if investors expect more intense manipulation as the firm’s profitability increases? If the manager anticipates that he will have to manipulate earnings more actively in the future, without being able to affect the firm’s stock price, his incentives to manipulate today may actually decay. Moreover, this effect can be exacerbated by the fact that his private information suggests to him that the stock price is likely to drift up in the short-run anyway. This difficulty in identifying, at

\textsuperscript{1}Another convenient technical device is the Poisson learning technology, typically used to model experimentation effects (Hörner and Samuelson, 2014; Bonatti and Hörner, 2011 and 2014).

\textsuperscript{2}For example, at the zero-earnings threshold (Burgstahler and Dichev, 1997; Degeorge et al., 1999).
any point in time, which incentive constraints bind, implies that we have to study the payoff associated with any level of belief discrepancy. This paper shows that when information arrives frequently, the standard approach of computing such payoffs using value functions is virtually useless: the possibility of belief divergence implies that the long-run player’s value function no longer satisfies a traditional Hamilton-Jacobi-Bellman (HJB) equation. The resulting partial differential equation (PDE) is considerably more complex. To the best of my knowledge, no existence results for this equation exist.

The first-order approach that I propose bypasses this difficulty. I first show that in any equilibrium in which actions are Markov (and sufficiently differentiable) in the common belief, and learning is stationary, behavior is characterized by an ordinary differential equation (ODE): the ratcheting equation. Descriptively, this equation encapsulates how the equilibrium degree of manipulation varies across different levels of beliefs. Its main novelty is a form of nonlinearity that captures how the ratcheting implicit in the market’s demand for manipulation affects the long-run player’s incentives. More specifically, as the steepness of this demand schedule increases, the returns from manipulating the market’s belief—hence, current incentives—drop, as manipulating the signal results in a more demanding level of manipulation that has to be fulfilled tomorrow. Notably, this force is fully endogenous, and absent in settings where payoffs are linear or where information is coarse (e.g., Poisson learning with fully informative signals).

The ratcheting equation offers enormous advantages over the HJB approach: (i) conceptually, it resembles classic Euler equations, but with the novelty of incorporating dynamic ratchet-like forces, and (ii) numerically, it is easy to compute. Standing alone, however, it is powerless: it becomes relevant only in the presence second-order conditions that validate its use. For this purpose, I derive a verification theorem (Theorem 1), which states verifiable sufficient conditions under which a solution to the ratcheting equation constitutes an equilibrium. This theorem is akin to classic dynamic-programming results using HJB equations, but it is considerably simpler. Rather than solving the belief-divergence PDE, it involves a system of two ODEs—the ratcheting equation (capturing equilibrium behavior) and another ODE capturing equilibrium payoffs—that must satisfy a second-order condition that bounds the value of becoming privately informed about the hidden state. Using this result, I prove the existence—and provide a full characterization—of equilibria in which actions are a non-trivial function of the public belief, for a wide range of economic environments (Theorems 2 and 3). These three theorems successfully tackle the belief-divergence challenge, and the continuous-time approach is critical for their derivation.\(^3\)

\(^3\)To the best of my knowledge, the question of sufficiency—i.e., of existence of equilibria—has not been answered in settings incorporating both Gaussian learning and nonlinearities.
The methods that I develop open a venue for studying a diversity of applications where the presence of nonlinearities is an essential feature of the environment at hand. In one of the applications developed in the paper, I study manipulation dynamics for a manager who has a strong short-term reputational incentive to exceed a zero-earnings threshold. I show that ratchet effect can lead to firms that are expected to exceed the threshold to actually manipulate earnings more actively, despite their managers having weaker myopic incentives, and being unable to affect firms’ market values. Even more so, the ratcheting created completely breaks down the symmetry of the environment, leading to firms with better prospects to manipulate earnings more intensively on average. In another application motivated by earnings management, I show that the presence of nonlinearities can give rise to multiplicity of equilibria. Interestingly, these expectation traps arise despite the absence of complementarities between actions and the unobserved state in the signal technology. Finally, in application to dynamic Cournot duopoly where firms learn about a demand’s moving intercept (and the market price is a noisy signal of the firms’ quantities), I illustrate how the methods also apply to settings where several long-run players jam a public signal. In this context, I show that, because of inter-temporal smoothing, firms may have the incentive to incur in significant losses in anticipation of periods where overproduction is profitable.

To conclude, I discuss two key assets of this paper. First, its generality. The issue of belief divergence appears in all settings in which learning and with hidden actions interplay with each other. In this line, the tools that I develop can be used in areas beyond the applications developed here (in macroeconomics, to study monetary policy in the presence of hidden inflation trends; in labor markets, to study reputational incentives when market frictions compress wages; etc.). Second, the framework makes nontrivial predictions on observables. More specifically, in settings where the equilibrium manipulation policy is nonlinear, the distribution of the public signal gets altered; in particular, it ceases to be Gaussian. Testable predictions about these observables (e.g., observed earnings, market prices, etc.), based on the primitives of the model, can be derived. While this exercise is not pursued at this point, the results presented provide all the tools that are required for such analysis.

The paper is organized as follows. In Section 2 I introduce the general model. Section 3 sets up the long-run player’s problem as one of stochastic control. Section 4 derives necessary and sufficient conditions for Markov equilibria, and develops three applications. Section 6 states existence results. Section 7 concludes. All proofs are relegated to the Appendix.

---

4 Multiplicity of equilibria was shown to exist by Dewatripont et al. (1999) in a model of career concerns where actions and talent enter multiplicatively in the output technology.
1.1 Literature

Regarding the reputation literature, Holmström (1999) develops a linear and Gaussian model of career concerns in which a worker’s effort is history-independent. A similar phenomenon occurs in the Poisson model of Board and Meyer-ter-Vehn (2013), where a firm’s private information regarding its technology is irrelevant for its investment policy. Board and Meyer-ter-Vehn (2014) allow for belief divergence affecting a firm’s investment decisions, and they prove the existence of (potentially, mixed-strategy) equilibria.

This paper is also connected to the literature on experimentation, typically modeled via a risky arm of two possible types with a Poisson-signal technology that reveals if the arm is good. Because beliefs evolve deterministically in this case, the belief-divergence analysis gains tractability: with linear payoffs, Bergemann and Hege (2005) show that the agent’s (off path) value function satisfies intuitive properties, whereas Bonatti and Hörner (2011, 2014) apply standard optimal control techniques. Also in the context of linear payoffs, Hörner and Samuelson (2014) provide a comprehensive characterization of the set of equilibria in a principal-agent problem similar to Bergemann and Hege’s model, showing that pure-strategy equilibria (hence, Markov) may fail to exist generically.

Regarding the literature on the ratchet effect, Weitzman (1980) studies effort provision in planning economies where ratcheting is modeled explicitly in an exogenous and linear incentive scheme. Freixas et al. (1985) and Laffont and Tirole (1988) instead study information revelation when a principal can revise the incentive scheme in place. Martinez (2009) identifies the presence of ratchet forces when studying first-order conditions for effort in a model of career concerns with piecewise linear wages. The framework that I study instead permits nonlinear incentive schemes with endogenous revision rules, and it delivers general sufficient conditions for the existence of equilibria in fully dynamic setting.

To conclude, this paper contributes to the literature that analyzes dynamic incentives using continuous-time methods. Faingold and Sannikov (2011) and Bohren (2014) study games between a long-run player and a population and small players; the former allows for one-sided learning, while the latter analyzes a class of games of pure investment. Williams (2011), Prat and Jovanovic (2014) and Sannikov (2014) take a first-order approach to contracting problems where actions can have high persistence, and they are able to verify their sufficient conditions in isolated environments only. In contrast, I show that limited commitment allows the set of environments under study to be enlarged considerably.

5Cisternas (2012) studies deterministic equilibria in environments that allow for skills accumulation.
2 The General Model

In the sequel, a long-run player and a population of small players (or market) learn about a hidden state \((\theta_t)_{t \geq 0}\) from observing a public signal \((\xi_t)_{t \geq 0}\) continuously over time.

2.1 Information Structure

The pair \((\theta_t, \xi_t)_{t \geq 0}\) evolves according to

\[
\begin{align*}
    d\theta_t &= -\kappa(\theta_t - \eta) dt + \sigma_\theta dZ^\theta_t, \quad t \geq 0, \\
    d\xi_t &= (a_t + \theta_t) dt + \sigma_\xi dZ^\xi_t, \quad t \geq 0,
\end{align*}
\]

where \(Z^\theta := (Z^\theta_t)_{t \geq 0}\) and \(Z^\xi := (Z^\xi_t)_{t \geq 0}\) are independent Brownian motions. The process \((\theta_t)_{t \geq 0}\) is Gaussian in mean-reverting form, with \(\kappa \geq 0\) capturing the rate at which \(\theta\) reverts to its long-run mean, \(\eta \in \mathbb{R}\).\(^6\) The public signal follows an Ito process that carries information about \((\theta_t)_{t \geq 0}\) in its drift. The term \(a_t\) corresponds to the degree of manipulation exerted by the long-run player at time \(t \geq 0\). The volatility parameters \(\sigma_\theta\) and \(\sigma_\xi\) are strictly positive.

The actions of the long-run player can take values in an interval \(A \subseteq \mathbb{R}\), with \(0 \in A\) (no manipulation is feasible), and they are never observed by the market. Notice that the actions of the long-run player and the hidden state enter in an additively separable way in the drift of \((\xi_t)_{t \geq 0}\) (“signal-jamming” technology; Holmström, 1999). This has two implications. First, \((\xi_t)_{t \geq 0}\) satisfies the full-support assumption with respect to the long-run player’s actions;\(^7\) hence, the market can never detect deviations from equilibrium behavior. Second, the speed at which all agents learn about \((\theta_t)_{t \geq 0}\) is exogenous (Lemma 1 in the next section).

Let \(\mathcal{F}_t^\xi\) denote the information generated by the observations \((\xi_s : 0 \leq s \leq t)\). I denote by \(\mathbb{F}^\xi := (\mathcal{F}_t^\xi)_{t \geq 0}\) the public information structure. It is clear that the long-run player also observes the signal \(Y_t := \xi_t - \int_0^t a_s ds, \quad t \geq 0\). By definition, \((Y_t)_{t \geq 0}\) is also an Ito process:

\[
    dY_t = \theta_t dt + \sigma_\xi dZ^\xi_t, \quad t \geq 0.
\]

The information contained in \((\xi_s : 0 \leq s \leq t)\) and \((Y_s : 0 \leq s \leq t), \quad t \geq 0\), need not coincide. This will occur, for instance, when the long-run player manipulates the signal differently from the market’s expectation, thus leading the market to misinterpret the public

\(^6\)Also known as Ornstein-Uhlenbeck process, and it corresponds to the continuous-time analog of the traditional AR(1) process in discrete time. Unless otherwise stated, all the results extend to homogenous Gaussian diffusions; i.e., stochastic differential equations with drift linear in the state, and constant volatility.

\(^7\)This is a consequence of Girsanov’s theorem, which states that changing the drift in the public signal induces equivalent distribution over the set of paths of \((\xi_t)_{t \geq 0}\) (see, for instance, Pham, 2009).
signal. Because in this case \((Y_t)_{t \geq 0}\) becomes private information to the long-run player, I denote by \(\mathbb{F}^{\xi,Y} := (\mathcal{F}^{\xi,Y}_t)_{t \geq 0}\) his information structure (\(\mathcal{F}^{\xi,Y}_t\) being the product sigma-algebra generated by the observations \((\xi_s, Y_s : 0 \leq s \leq t))\), thus emphasizing that deviations lead to an informational advantage.

### 2.2 Learning Dynamics

In order to form a correct belief about the hidden state, the market must correct the bias introduced by the long-run player in the public signal. Consequently, the market’s belief will depend on its conjecture of equilibrium play, \((a^*_t)_{t \geq 0}\). Because the market observes the public signal only, this conjecture must be (progressively) measurable with respect to \(\mathbb{F}\).

Starting from a normal prior about the initial value \(\theta_0\), the systems (1)-(2) (defined by \((\theta_t)_{t \geq 0}\) and \((\xi_t)_{t \geq 0}\)) and (1)-(3) (defined by \((\theta_t)_{t \geq 0}\) and \((Y_t)_{t \geq 0}\)) are conditionally Gaussian and Gaussian learning processes, respectively (Liptser and Shiryaev, 1977). The next result states laws of motion for posterior means and variances:

**Lemma 1 (Evolution of beliefs).** Suppose that the long-run player and the market share a (normal) common prior \(\theta_0 | \mathcal{F}_0 \sim \mathcal{N}(\theta^o, \gamma^o)\), \((\theta^o, \gamma^o) \in \mathbb{R} \times \mathbb{R}_+\). Then, at any time \(t \geq 0\):

(i) the long-run player’s posterior belief, \(\theta_t | \mathcal{F}^\xi_t\), is normally distributed. The posterior mean process \(p_t := \mathbb{E}[\theta_t | \mathcal{F}^\xi_t]\), \(t \geq 0\), evolves according to

\[
dp_t = -\kappa(p_t - \eta)dt + \frac{\gamma_t}{\sigma^2} \frac{dY_t - p_tdt}{\sigma^2}, \quad t > 0, \quad p_0 = p^o, \tag{4}
\]

where \(Z_t := \frac{1}{\sigma^2} \left( Y_t - \int_0^t p_sds \right), \quad t \geq 0\), is a Brownian motion with respect the information structure \(\mathbb{F}^\xi\) (also called an innovation process);

(ii) if the market believes that \((a^*_t)_{t \geq 0}\) arises in equilibrium, its posterior belief, \(\theta_t | \mathcal{F}^\xi_t\), is normally distributed. The posterior mean, \(p^*_t := \mathbb{E}[\theta_t | \mathcal{F}^\xi_t]\), \(t \geq 0\), evolves according to

\[
dp^*_t = -\kappa(p^*_t - \eta)dt + \frac{\gamma_t}{\sigma^2} \frac{d\xi_t - (p^*_t + a^*_t)dt}{\sigma^2}, \quad t > 0, \quad p_0 = p^o, \tag{5}
\]

---

8Recall that, a stochastic process \((X_t)_{t \geq 0}\) is said to be adapted to the filtration \((\mathcal{F}_t)_{t \geq 0}\) if, for each \(t\), the random variable \(X_t : \Omega \rightarrow \mathbb{R}\) (\(\Omega\) the sample space) is measurable with respect to \(\mathcal{F}_t\). In continuous time, this notion is too weak, as it imposes no measurability requirements across time, which is crucial for integrals over time to be random variables (and hence, to be able to be taken expectations). Progressive measurability then requires, for all \(t \geq 0\), the measurability of \(X : [0, t] \times \Omega \rightarrow \mathbb{R}\) with respect to \(\mathcal{B}([0, t]) \otimes \mathcal{F}_t\), the product sigma-algebra on \([0, t] \times \Omega\). Under this joint measurability across time and events, adaptability is preserved and expectations of integrals are well defined. In the sequel, I omit the term “progressively.”

9Conditionally Gaussian means that the conditional distribution \(\theta_t | \mathcal{F}^\xi_t\) is Gaussian, despite \((\xi_t)_{t \geq 0}\) not being normally distributed (Theorem 11.1 in Liptser and Shiryaev, 1977). A nonlinear analog of the Kalman-Bucy filter still applies nonetheless.
where \( Z_t^a := \frac{1}{\sigma_\xi} \left( \xi_t - \int_0^t (p_t^\ast + a_t^\ast) ds \right), \ t \geq 0, \) is a Brownian motion with respect to the public information structure \( \mathbb{F}^\xi \) (also called an innovation process).

(iii) in both (4) and (5), \( \gamma := (\gamma_t)_{t \geq 0} \) denotes the (common) posterior variance of each learning process, which evolves according to the ODE

\[
\dot{\gamma}_t = -2\kappa \gamma_t + \sigma_\theta^2 \left( \frac{\gamma_t}{\sigma_\xi} \right)^2, \quad t > 0, \quad \gamma_0 = \gamma^0. \tag{6}
\]


In the case of the long-run player, all the relevant information for learning purposes is captured in \( Y \); hence, it suffices to condition only on \( \mathcal{F}_t^Y \) in \( \mathbb{E}[\theta_t|\mathcal{F}_t^Y], \ t \geq 0 \). On the other hand, \( \mathbb{E}^a[\cdot|\mathcal{F}_t^\xi] \) emphasizes that, in order to form its belief, the market must conjecture a stochastic process \( (a_t^\ast)_{t \geq 0} \) driving the evolution of the public signal \( (\xi_t)_{t \geq 0} \).

Observe that both \( (p_t)_{t \geq 0} \) and \( (p_t^\ast)_{t \geq 0} \) are mean-reverting processes given their corresponding information structures. In the case of the long-run player, the process \( (Z_t)_{t \geq 0} \) captures surprises in the signal \( (Y_t)_{t \geq 0} \) that move the private belief away from \( \eta \). Similarly for the market, with \( (Z_t^a)_{t \geq 0} \) capturing unanticipated changes in \( (\xi_t)_{t \geq 0} \). The intensity of the reaction of each belief process to the corresponding innovation is given by \( (\gamma_t)_{t \geq 0} \), which, at any point in time, summarizes how much information about \( (\theta_t)_{t \geq 0} \) has been accumulated. This process is exogenous, as, by the additive separability assumption, the long-run player cannot affect his (nor the market’s) speed of learning.

I am mostly interested in environments where the uncertainty about it never gets fully revealed. In this sense, the dynamic (1) is particularly suitable, as the shocks \( (Z_t^0)_{t \geq 0} \) imply that there is always some residual uncertainty about the current value of the hidden state. The model that I study is one of stationary learning:

Assumption 1 (Ex-ante symmetric uncertainty; stationary learning). The long-run player and the market share a common prior \( \theta_0|\mathcal{F}_0 \sim \mathcal{N}(p^0, \gamma^*) \), where \( \gamma^* := \sigma_\xi^2 (\sqrt{\kappa^2 + \sigma_\theta^2 / \sigma_\xi^2} - \kappa) > 0 \) (i.e., the unique strictly positive and stationary solution of (6)). Hence, \( \gamma_t = \gamma^*, \ t \geq 0 \).

While posterior means can change over time, the total information remains constant. I refer to \( (p_t)_{t \geq 0} \) (respectively, \( (p_t^\ast)_{t \geq 0} \)) as the private (respectively, public) belief process.

\footnote{For instance, at any time \( t \geq 0 \) the long-run player expects \( (Y_t)_{t \geq 0} \) to increase by \( p_t dt \) over \( [t, t + dt) \). If the realization \( dY_t - p_t dt \) is larger than zero, this is good news, and beliefs respond positively.}
2.3 Payoffs, Strategies and Equilibrium Concept

The market consists of a large number of anonymous and fully forward-looking agents, each of them negligible relative to the market size (i.e., individual actions do not affect the population average observed by the long-run player). Consequently, market participants behave myopically: each agent takes actions in an attempt to maximize an ex-ante flow payoff given (i) the current public information, and given (ii) his/her current conjecture of equilibrium play. I summarize the market’s aggregate behavior in the best-response function \( \chi(p^*_t,a^*_t), \ t \geq 0 \), with \( \chi: \mathbb{R} \times A \to \mathbb{R} \) a continuously differentiable function.

Given a conjecture \((a^*_t)_{t \geq 0}\), a strategy \((a_t)_{t \geq 0}\) of the long-run player is feasible if it corresponds to a \( \mathcal{F}^g_t \) - progressively measurable with the property that \( \mathbb{E} \left[ \int_0^t (a_s)^2 ds \right] < \infty, \ t \geq 0. \) In this case, the long-run player’s continuation payoff at time \( t \) takes the form

\[
\mathbb{E}^a \left[ \int_t^\infty e^{-r(s-t)}(u(\chi(p^*_s,a^*_s))) - g(a_s))ds \right| \mathcal{F}^g_t, \ t \geq 0. \tag{7}
\]

Here, \( u: \mathbb{R} \to \mathbb{R} \) is a continuously differentiable function that represents the component of the long-run player’s flow utility that is affected by the market’s action (e.g., flow utility from consumption). Manipulation is costly according to a twice continuously differentiable function \( g: A \to \mathbb{R} \) that satisfies \( g(0) = 0, \ g'(\lvert a \rvert) > 0 \) for \( \lvert a \rvert > 0, \) and \( g''(a) > 0 \) for all \( a \in A. \) I have also made explicit the dependence of \( \mathbb{E}^a[\cdot] \) on \((a_t)_{t \geq 0} \) to emphasize that the choice of strategy determines the likelihood of different realizations of the public signal, which ultimately determines the realization of the public belief process \((p^*_t)_{t \geq 0}.\)

**Definition 1 (Equilibrium in Pure Strategies).** A process \((a^*_t)_{t \geq 0}\) is an equilibrium in pure strategies if: (i) it corresponds to an \( A \) - valued and \( \mathcal{F}^g \) - progressively measurable process; (ii) it satisfies \( \mathbb{E} \left[ \int_0^t (a_s)^2 ds \right] < \infty, \ t \geq 0, \) (feasibility); and (iii) for all \( t \geq 0, \) \((a^*_s)_{s \geq t}\) maximizes (7) when the market constructs beliefs using \((a^*_s)_{s \geq t}\) and \( p_t = p^*_t. \) An equilibrium in pure strategies is Markov if there exists a Lipschitz function \( a^*: \mathbb{R} \to A, \) such that equilibrium actions take the form \((a^*(p^*_t))_{t \geq 0}.\)

In an equilibrium in pure strategies, the market must construct beliefs using a conjecture such that it is optimal for long-run player to follow it when beliefs are aligned. But because the market only observes the public signal, such equilibrium is constrained to depend on

---

\(^{11}\)This integrability condition ensures that, when the long-run player controls the public process (5), the latter has a solution measurable with respect to his information structure, and so the optimization problem of the long-run player is well-defined. See Section 3.1.

\(^{12}\)I will dispense with this notation when formulating the long-run player’s problem as one of optimal control, as the impact of any strategy on the public belief will be captured explicitly in its law of motion.
the public information only. Starting from a common prior therefore, beliefs coincide at all future times on the equilibrium path.

If an equilibrium in pure strategies exists, histories in which beliefs do not coincide have zero probability. Nevertheless, these off-path histories are precisely the ones that determine what happens on the equilibrium path. Put differently, in order for \( a^* \) to be optimal for the long-run player, the payoff of following \( a^* \) must dominate the payoff of all alternative strategies; but under an alternative strategy, the market will have misinterpreted the public signal, and hence beliefs need not coincide. As a result, the problem of finding equilibria requires allowing for the possibility of belief divergence, as off-equilibrium payoffs ultimately pin down equilibrium behavior. The combination of learning and imperfect monitoring therefore leads to a complex joint moral-hazard and adverse-selection problem off the equilibrium path.

In the sequel, I focus on the case of Markov equilibrium; the Lipschitz condition just ensures that long-run player’s stochastic control problem is well defined (see Section 3.1).

Remark 1. The equilibrium concept introduced in Definition 1 does not specify off-path behavior. This is not a problem: a key insight of this paper is that, in the Markov case, it is not necessary to exactly specify off equilibrium actions (or payoffs) in order to pin down equilibrium behavior. The equilibrium concept can be strengthened to Markov Perfect Equilibrium provided the long-run player has an optimal Markov strategy off the equilibrium path. While this is likely to be the case, in the next section it is shown that it is very challenging to prove such claim using value functions. The methodological contribution of this paper is therefore to offer a tractable alternative method to compute Markov equilibria that bypasses the exact computation of off-equilibrium payoffs.

3 The Long-Run Player’s Best-Response Problem and the HJB Approach

In this section I show that the long-run player’s problem of manipulating the public belief can be seen as one of investment in the presence of convex costs (Dixit and Pindyck, 1994). Formulating the long-run player’s problem in this way is particularly useful for understanding his incentives. The section concludes by arguing why following the traditional HJB approach is virtually useless for performing tractable equilibrium analysis.

3.1 The Long-Run Player’s “Investment” Problem

The precise way in which the long-run player can affect the market’s belief is as follows. From the definition of the innovation process \((Z_t)_{t \geq 0}\) (part (i) in Definition 1), we have that
$dY_t = p_t dt + \sigma_t dZ_t$ from the long-run player’s perspective. Hence,

$$d\xi_t = dY_t + a_t dt = (a_t + p_t) dt + \sigma_t dZ_t, \quad t \geq 0,$$

where $(a_t)_{t \geq 0}$ is feasible. That is, given the long-run player’s information, the public signal $(\xi_t)_{t \geq 0}$ evolves as in (2), but replacing, $\theta_t$, by his current belief about it, $p_t$, $t \geq 0$.\(^{13}\)

Plugging (8) into the dynamic (5) of the public belief, yields the law of motion of $(p^*_t)_{t \geq 0}$ given the long-run player’s information structure: letting $\beta := \gamma^*/\sigma^2$ and $\sigma := \beta \sigma_\xi$ (both depending on $(\kappa, \sigma_\theta)$ as well as $\gamma^*$) we have that, under any feasible $(a_t)_{t \geq 0}$,

$$dp^*_t = \left[ -\kappa (p^*_t - \eta) + \beta (a_t - a^*_t (p^*_t)) + \beta (p_t - p^*_t) \right] dt + \sigma dZ_t, \quad t \geq 0.$$  

(9)

Compared to (5), two extra terms appear. First, current deviations positively affect the market’s belief only when they exceed the market’s current conjecture ($a_t - a^*_t$ term). Second, the history of past deviations is summarized in the difference $p_t - p^*_t$. In particular, if the long-run player is more optimistic than the market about the fundamental, he will expect the public belief to drift up, everything else equal. This form of private information regarding $(p^*_t)_{t \geq 0}$ is likely to be payoff relevant (i.e., it is likely to affect his incentives).\(^{14}\)

Because deviations off the equilibrium path generate belief divergence, it is natural to introduce a state variable that captures its dynamic:

**Proposition 1. (Belief-Asymmetry Process).** Suppose that the long-run player follows $(a_t)_{t \geq 0}$, while the market conjectures $(a^*_t)_{t \geq 0}$. Then, $\Delta_t := p_t - p^*_t, \ t \geq 0$, evolves according to

$$d\Delta_t = \left[ -(\beta + \kappa) \Delta_t + \beta (a_t - a^*_t) \right] dt, \ t > 0.$$  

(10)

Hence, starting from a common prior ($\Delta_0 = 0$), if the long-run player follows $(a^*_t)_{t \geq 0}$, beliefs will remain aligned. That is, $\Delta \equiv 0$ on the equilibrium path.

**Proof:** Subtract (4) from (9) \(\blacksquare\)

The dynamic (10) shows that any degree of belief divergence has persistence. For instance, if a deviation at $[t, t + dt)$ takes place, and $a_s = a^*_s$ afterwards, the stock of belief asymmetry generated disappears at a rate of $\beta + \kappa > 0$, as the learning puts less weight to past observations as time progresses. Crucially, it is precisely this type of persistence what

---

\(^{13}\)Formally, this is a representation theorem: representing $(\xi_t)_{t \geq 0}$ under both a different Brownian motion and filtration. See Theorem 7.12 in Liptser and Shiryaev (1977).

\(^{14}\)Lipschitz guarantees that, for any feasible $(a_t)_{t \geq 0}$, there exists a unique $\mathbb{F}^Y$-measurable solution to (9) that also has a finite second moment. The long-run player’s optimization problem is thus well-defined.
makes incentives difficult to evaluate. More precisely, deviations in any direction lead to a persistent wedge between the long-run player’s and the market’s perception of the distribution of the public signal, and the long-run player can exploit this informational advantage (he knows both \((p_t)_{t \geq 0}\) and \((p^*_t)_{t \geq 0}\)) over long periods of time.

When the market expects the long-run player to follow a Markov strategy, all the payoff-relevant history of the game is summarized in the current value that \(p\) and \(\Delta\) take. Suppose that at \(s > 0\) the current history is \((p_s, \Delta_s) = (p, \Delta) \in \mathbb{R}^2\). Then, the long-run player’s best-response problem to the Markov conjecture \(a^* : \mathbb{R} \to A\) is given by

\[
\begin{align*}
\max_{\substack{(a_t)_{t \geq s} \quad \text{s.t.} \quad \begin{align*}
&dp_t = -\kappa(p_t - \eta)dt + \sigma dZ_t, \ t > s, \ p_s = p, \\
&d\Delta_t = [-(\beta + \kappa)\Delta_t + \beta(a_t - a^*_t(p_t + \Delta_t))]dt, \ t > s, \ \Delta_s = \Delta.
\end{align*}
\end{align*}
\]

The previous optimization problem is well-known dynamic decision problem: a problem of investment under convex costs of adjustment; namely, the long-run player “invests in belief asymmetry \(\Delta\).” There are adjustment costs both in the convex cost of manipulation \(g(\cdot)\), and in the law of motion of \(\Delta\), which resembles generalized dynamics modeling the accumulation of physical capital (Uzawa, 1969; Hayashi, 1982). In particular, an extra unit of manipulation maps into more belief asymmetry only when the size of the manipulation exceeds \(a^*\). The private belief \((p_t)_{t \geq 0}\) in turn acts as an exogenous “price” which, as it fluctuates, varies the rate of return of \(\Delta\) on the long-run player’s flow payoff, \(u \circ \chi\).

The previous formulation is very useful: we can immediately conclude that the long-run player’s actions are given by the marginal value of belief asymmetry (Tobin’s marginal \(q\) in the investment literature). That is, letting \(V^{a^*}(p, \Delta)\) denote the value of the previous problem, the optimal manipulation policy is characterized by \((p, \Delta) \mapsto V^{a^*}_\Delta(p, \Delta)\). The traditional approach to finding such policy is through solving HJB equations (next section).

To conclude, notice that, after a deviation takes place, the market not only holds a biased belief about the unobserved state \((\Delta \neq 0)\): it also holds a wrong conjecture of equilibrium play \((a^*(p + \Delta)\) term), so it actually constructs beliefs in an incorrect way.\textsuperscript{15} As it will be seen in the next section, this imposes severe difficulties to performing equilibrium analysis.

### 3.2 The HJB Approach to Equilibrium Analysis

Consider the best-response problem to a Markov conjecture \(a^* : \mathbb{R} \to A\) first. Traditional verification theorems from dynamic programming tell us that if we are able to find a solution

\textsuperscript{15}This is a consequence of the full-support property of the public signal (2).
to the HJB equation (which takes \( a^*(\cdot) \) as an input)

\[
\begin{align*}
  rV^a(p, \Delta) &= \sup_{a \in A} \left\{ u(\chi(p + \Delta, a^*(p + \Delta))) - g(a) + \mathbb{E}[dV^a(p_t, \Delta_t)/d t | (p_t, \Delta_t) = (p, \Delta)] \right\} \\
  &= \sup_{a \in A} \left\{ u(\chi(p + \Delta, a^*(p + \Delta))) - g(a) - \kappa(p - \eta)V^a_p(p, \Delta) + \frac{1}{2}\sigma^2V^a_{pp}(p, \Delta) \right\} + \left[ -\beta + \kappa \right] \Delta_t + \beta(a - a^*(p + \Delta)) \right\} \right\}, \quad (p, \Delta) \in \mathbb{R}^2, \quad (11)
\end{align*}
\]

then \( V^a \) is value function of the best-response problem.\(^{16}\) Furthermore, if

\[
\alpha(p, \Delta) := \arg \max_{a \in A} \{a\beta V^a(\Delta_t + \beta(a - a^*(p + \Delta))) V^a_\Delta(p, \Delta) \}, \quad (p, \Delta) \in \mathbb{R}^2
\]  

induces a well-defined dynamic for \((\Delta_t)_{t \geq 0}\), then \( \alpha(p, \Delta) \) is an optimal control (i.e., a best-response to \( a^* \)) in Markov form.

Consequently, if we are able to find a solution to the partial differential equation (PDE)

\[
\begin{align*}
  rV(p, \Delta) &= \sup_{a \in A} \left\{ u(\chi(p + \Delta, a^*(p + \Delta))) - g(a) - \kappa(p - \eta)V^a_p(p, \Delta) + \frac{1}{2}\sigma^2V^a_{pp}(p, \Delta) \right\} + \left[ -\beta + \kappa \right] \Delta_t + \beta(a - a^*(p + \Delta)) \right\} \right\}, \quad (p, \Delta) \in \mathbb{R}^2, \quad (13) \\
  s.t. \quad a^*(p) &\in \arg \max_{a \in A} \{\beta V_\Delta(p, 0) a - g(a)\}, \quad p \in \mathbb{R}, \quad (14)
\end{align*}
\]

it is clear that the resulting function \( p \mapsto V(p, 0) \) will correspond to the long-run player’s payoff in an equilibrium in which his actions are Markov and beliefs remain aligned. This is because the fixed-point condition (14) ensures that the market perfectly anticipates the long-run player’s (optimal) actions. In any such equilibrium, the derivative \( V_\Delta(p, 0) \) captures the long-run player’s equilibrium actions. Because in equilibrium no belief asymmetry is created, I refer to this derivative as the value of a (hypothetical) unit of belief asymmetry.

The main difficulty with the HJB-approach is that the PDE (13)-(14) is no longer a standard PDE from dynamic programming: it is **non-local**, in the sense that the behavior of the value function around a point \((p, \Delta)\) depends on information away from this point; namely, at the point \((p + \Delta, 0)\). More specifically, when actions are interior, \( V(p, \Delta) \) depends on the derivative \( V_\Delta(p + \Delta, 0) \) through the market’s (potentially wrong) conjecture of equilibrium play, \( a^*(p + \Delta) = g^{-1}(\beta V_\Delta(p + \Delta, 0)) \). This particular type of non-localness—and, consequently, the type of PDE that arises due to the belief-divergence problem—seems to

\(^{16}\) Transversality and growth conditions must be verified too. See Theorem 3.5.3 in Pham (2009).
be new in the theory of PDEs. To the best of my knowledge, no existence results are known. This important shortcoming imposes a severe limitation to the use of the HJB approach.\footnote{This technical complexity is present in virtually all environments in which learning interacts with unobserved actions. There are exceptions, of course. When payoffs are linear (Holmström, 1999; Stein, 1989), there is an additively separable (in $p$ and $\Delta$) solution to this equation, and the problem disappears. Belief divergence is also present in models of experimentation (e.g., controlling $(\gamma_t)_{t \geq 0}$). A convenient modeling device in this area has been to assume Poisson learning and coarse information structures (observing a signal reveals the true state of the world). Because in these settings equilibrium actions are deterministic, standard optimal control techniques are still tractable, as Bonatti and Hörner (2011, 2014) show.}

Before moving towards a tractable solution to this problem, it is important to stress that, although an interesting object, the PDE (13)-(14) is not our main object of interest. This PDE is useful only to the extent that it provides a method (namely, a verification theorem) for finding Markov equilibria. This method is complex because it requires an \textit{exact} computation of payoffs when beliefs do not coincide. Tackling this full program is far from necessary to characterize equilibrium behavior, as I show next.

\section{The First-Order Approach}

In this section I derive necessary and sufficient conditions for Markov equilibrium that rely on a simple ordinary differential equation. These conditions are much simpler and intuitive than the standard dynamic-programming approach of solving (13)-(14), both from the point of view of economics and of computation. The section concludes discussing some economic environments where the methods developed in this paper can be applied.

\subsection{Necessary Conditions for Markov Equilibria}

While the long-run player’s incentives in the best-response problem are given by the derivative $V^a_\Delta(p, \Delta)$, we are only interested in on-path behavior: namely on $q(p) := V^a_\Delta(p, 0)$. In order to obtain necessary conditions for Markov equilibria, I proceed now in a heuristic fashion; namely, assuming enough differentiability on $a^*(\cdot)$ and $V^a(\cdot, \cdot)$.

Applying the envelope theorem with respect to $\Delta$ in the HJB equation of the best-response problem (eqn. (11)), and setting $\Delta = 0$, we obtain the \textit{Euler equation}

\begin{equation}
    rV^a_\Delta(p, 0) = \frac{d}{dp}(u \circ \chi)(p, a^*(p)) - \kappa(p - \eta)V^a_{\Delta p}(p, 0) + \frac{1}{2}\sigma^2 V^a_{\Delta p p}(p, 0) \\
    - \left[\beta + \kappa + \beta \frac{da^*}{dp}(p)\right] V^a_\Delta(p, 0) + \beta [g^{t-1}(V^a_\Delta(p, 0)) - a^*(p))]V^a_{\Delta \Delta}(p, 0)
\end{equation}

where I have assumed that incentives are interior. While this equation depends only on
one state variable \((p)\), it hardly simplifies the computation of optimal policies compared to the original HJB equation (11). This is because (15) does not correspond to an ODE in \(p \mapsto V^{a^*}_{\Delta}(p,0)\); it also depends on \(V^{a^*}_{\Delta}(p,0)\) (last term in (15)).

As any Euler equation, (15) is a recursive expression that the long-run player’s marginal utility along the optimal path of \((p,\Delta)\) must satisfy: if the long-run player cannot obtain a first-order gain from a one-shot deviation from a specific strategy, his marginal utility must satisfy (15). Thus, as a necessary condition for optimality of actions only, (15) lacks the additional information that, in equilibrium, the optimal trajectory of \(\Delta\) must satisfy \(\Delta \equiv 0\). This additional information can be incorporated by constraining (15) to the fixed-point condition (14), i.e.,

\[
\begin{align*}
    rV^{a^*}_{\Delta}(p,0) &= \frac{d}{dp}(u \circ \chi)(p, a^*(p)) - \kappa(p-\eta) V^{a^*}_{\Delta p}(p,0) + \frac{1}{2} \sigma^2 V^{a^*}_{\Delta pp}(p,0) \\
    &- \left[ \beta + \kappa + \beta \frac{d a^*}{dp}(p) \right] V^{a^*}_{\Delta}(p,0) + \beta [g^{-1}(V^{a^*}_{\Delta}(p,0)) - a^*(p)] V^{a^*}_{\Delta}(p,0) \\
    \text{s.t.} \quad a^*(p) &= g^{-1}(\beta V^{a^*}_{\Delta}(p,0)).
\end{align*}
\]

Notably, this yields an ODE for \(p \mapsto q(p) := V^{a^*}_{\Delta}(p,\cdot)\):

**Proposition 2 (Necessary Conditions: Equilibrium Actions and Payoffs).** Assume that a Markov equilibrium \(a^* : \mathbb{R} \to A\) exists, and that both \(a^*(\cdot)\) and the long-run player’s value function exhibit enough differentiability. Then, if incentives are interior, \(a^*(p) = g^{-1}(\beta q(p))\), where \(p \mapsto q(p)\) satisfies the following nonlinear ODE:

\[
\left[ r + \beta + \kappa + \beta \frac{d}{dp} g^{-1}(\beta q(p)) \right] q(p) = \frac{d}{dp} (u \circ \chi)(p, \beta g^{-1}(q(p)))
\]

\[
-\kappa(p-\eta)q'(p) + \frac{1}{2} \sigma^2 q''(p), \quad p \in \mathbb{R}.
\]

In this equilibrium, the long-run player’s on-path payoff satisfies the ODE

\[
rU(p) = (u \circ \chi)(p, g^{-1}(\beta q(p))) - g(g^{-1}(\beta q(p))) - \kappa(p-\eta)U'(p) + \frac{1}{2} \sigma^2 U''(p), \quad p \in \mathbb{R}.
\]

**Proof:** The ODE (17) is obtained by evaluating (13)-(14) in \(\Delta = 0\). \(\square\)

The previous result reduces the search of Markov equilibria from solving the complex PDE (13)-(14), to solving a system of ODEs given by (16)-(17).\(^{19}\) Equation (16) corresponds to

---

\(^{18}\)This makes it a section of a PDE along the region \(\Delta = 0\).

\(^{19}\)The choice of notation \((q(\cdot), U(\cdot))\) rather than \((V_{\Delta}(\cdot,0), V(\cdot,0))\) is because there could be many solutions to (16)-(17). The question of sufficiency (section 4.2) establishes conditions under which a solution
a recursive expression for how the equilibrium degree of manipulation varies across different levels of public beliefs. The ODE (17) is in turn a recursive expression for the long-run player’s equilibrium payoff when he follows the strategy prescribed by \( q(\cdot) \), and no belief asymmetry is created. The latter is a standard (linear) ODE from dynamic programming. In fact, if the equilibrium payoff under \( q(\cdot) \), as a function of the current (common) belief, 
\[
U(p_t) := \mathbb{E} \left[ \int_0^\infty e^{-r(s-t)} [(u \circ \chi)(p_s) - g(g^{-1}(\beta q(p_s)))] ds \right] p_t,
\]
is of class \( C^2 \), we can apply Ito’s rule to obtain
\[
U(p_t) = [(u \circ \chi)(p_t) - g(g^{-1}(\beta q(p_t)))] dt + e^{-r dt} \mathbb{E}[U(p_{t+dt})] p_t - \kappa(p - \eta)U'(p) + \frac{1}{2} \sigma^2 U''(p) dt + U(p_t) + o(dt^2),
\]
from where it is easy to conclude that (17) holds.

Equation (16) is novel. As I explain below, constraining the Euler equation of the original problem (eqn. (15)) to the condition that beliefs must coincide on the equilibrium path, leads to (16) having a different structure compared to standard Euler equations from dynamic programming. The distinctive feature of (16) is the presence of endogenous ratchet-like forces affecting incentives. I refer to (16) as the ratcheting equation.

Equation (16) states that, in equilibrium, three economic forces must be optimally balanced. The right-hand side summarizes the benefits from inducing belief asymmetry; the first term captures myopic incentives, and the second and third terms capture cost-smoothing motives: if the value of inducing belief asymmetry is expected to increase, then, because \( g(\cdot) \) is convex, it is optimal to manipulate more today. The left-hand side instead captures the dynamic costs from engaging in this manipulation. Inducing belief divergence is costly for two reasons. Firstly, any stock of belief asymmetry (hence, the forward-looking value of it, \( q \)) naturally decays over time as new information arrives—the bigger \( \beta + \kappa \), the faster this decay. Secondly, by controlling the evolution of the public belief, the long-run player induces

\cite{q(\cdot),U(\cdot)} indeed corresponds to \((V_\Delta(\cdot,0), V(\cdot,0))\) from (13)-(14).

\textit{20}U(p_t) in both sides cancel out; then divide by \( dt \) and let \( dt \to 0 \). In the \textit{arbitrage interpretation} to continuation values, the return that the long-run player earns from following the market’s conjecture of equilibrium play (left side of (17)), must equal the flow payoff, plus the “expected capital gains.”

\textit{21}This is because, unlike standard problems in optimal control, the equilibrium problem requires the primitive of the controlled state—namely, the drift in \( \Delta \), via its dependence on \( a^* \)—to be determined simultaneously with the solution to the optimization problem itself.

\textit{22}By Ito’s rule \( dq(p_t) = -\kappa(p_t - \eta)q'(p_t) dt + q'(p_t) \sigma dZ_t + \frac{1}{2} \sigma^2 q''(p_t) dt \), so \( \mathbb{E}[dq(p_t)/dt|p_t = p] = -\kappa(p - \eta)q'(p) + \frac{1}{2} \sigma^2 q''(p). \)
the market to revise its current conjecture of manipulation. If this conjecture increases, this depresses incentives today, as the long-run player must match a higher level of manipulation in order to maintain the stock of belief asymmetry generated. To see this, observe that any stock of belief asymmetry depreciates, locally (and in monetary units) at a rate of

\[ r + \beta + \kappa + \beta \frac{da^*(p)}{dp} = r + \beta + \kappa + \beta \frac{d}{dp}q^{-1}(\beta q(p)). \]

Consequently, the steeper \( a^*(\cdot) \) is, the lower the return on inducing belief asymmetry; i.e., manipulation becomes more costly.

This ratcheting—i.e., that an agent’s incentives can be depressed (intensified) in anticipation of stronger (weaker) incentives in the future—is absent in standard Euler equations associated with decision problems in dynamic programming. In fact, given any standard optimal control problem in which the cost of controlling a state is convex, if a decision maker expects his incentives to be stronger in the future this can only strengthen his incentives today (everything else equal), as it is optimal to smooth out these costs across time.\(^{23}\) Also, in contrast with most of the existing literature studying ratchet effects, the ratcheting created in the model is both (i) fully endogenous, and (ii) fully dynamic. Depending on the context, it can play in favor or against the long-run player’s incentives.

**Remark 2.** The ratchet effect have been studied in contexts of incentives for effort provision and information revelation. Weitzman (1980) studies effort provision in planning economies in the presence of an exogenous linear incentive scheme that depends on a production target. He shows that effort is inefficiently low in a setting where the production target is updated via an exogenous (and linear) rule based on past performance, thus modeling ratcheting explicitly. Freixas et al. (1985) and Laffont and Tirole (1988) instead use adverse selection and no commitment to show that the possibility of the principal endogenously revising an incentive scheme can lead to considerable pooling across types.\(^{24}\) The framework that I study can be seen as a generalization of Weitzman (1980): it is also concerned with effort levels,

\(^{23}\)To see this, suppose that a decision-maker positively affects a one-dimensional state in an additively separable way: \( dp_t = (a_t + \mu(p_t))dt + \sigma dZ_t, \quad t \geq 0. \) For simplicity, assume that his flow payoff takes the form \( \chi(p_t) - a^2/2. \) Letting \( q(p) := V_{p}(p), \) the optimal control is exactly \( q(p), \) and the envelope theorem yields

\[ [r - \mu'(p) - q'(p)]q(p) = \chi'(p) + q'(p)\mu(p) + \frac{1}{2}\sigma^2q''(p). \]

From the left-hand side we can see that the larger \( q'(p) \) is, the larger the optimal control \( q(p), \) everything else equal. Furthermore, if \( \chi \) depends on \( a^*, \) the interaction term between \( q(\cdot) \) and the marginal flow payoff is never present in Euler equations.

\(^{24}\)Freixas et al. (1985) restrict to linear incentive schemes and two types, while the analysis of Laffont and Tirole (1988) is more general in both dimensions, both papers studying two-period models. Hart and Tirole (1988) study ratchet effects in a finite-horizon model, and also find considerable pooling.
yet in the presence of a nonlinear incentive scheme ($\chi$), and with a “target” ($a^\ast(\cdot)$; i.e., the level that has to be exceeded in order to generate more belief asymmetry) that is updated via an endogenous and nonlinear rule.\footnote{The updating is done via Ito’s rule: $da^\ast(p_t) = (a^\ast)'(p_t) [-\kappa(p_t - \eta) + \frac{1}{2} \sigma^2 (a^\ast)'(p_t)] dt + (a^\ast)'(p_t) dZ_t.}$ Unlike Freixas et al. (1985) and Laffont and Tirole (1988), the framework that I study uses ex-ante symmetric uncertainty—instead of on-path adverse selection—as the key modeling device. This learning-driven ratchet effect depends on the possibility of affecting the market’s belief. Consequently, it is absent in settings where information is coarse (e.g., Poisson learning with only one positive arrival rate), as in this case the market’s belief is deterministic.

\section*{4.2 Sufficiency: The Verification Theorem}

In this section I derive conditions on any solution to the system (16)-(17) such that, if the corresponding solution $q(\cdot)$ is used by the market to construct beliefs about the hidden state, the long-run player will be worse off by deviating to any alternative strategy. When this is the case, $q(\cdot)$ indeed corresponds to the value of creating a hypothetical unit of belief asymmetry.

In order to understand the power of Theorem 1 below, recall that the ratcheting equation was derived using the envelope theorem. Thus, the variational argument behind this incentive constraint relates to one-shot deviations. That is, when beliefs are aligned, if the long-run player cannot obtain a first-order gain from a strategy that differs from the market’s conjecture of equilibrium play ($a^\ast(p^\ast_t))_{t\geq 0}$ only for a short interval of time $[t, t + dt)$, the marginal value of belief asymmetry $V^\ast(\Delta^\ast, p, 0)$ must satisfy (16). Under this particular deviation, the stock of belief asymmetry generated depreciates exponentially at rate $\beta + \kappa$.

In contrast to standard dynamic decision problems, however, in the current context it is far less evident that deterring only these deviations suffices for ensuring that any deviation is unprofitable. The reason is that after a deviation takes place, the private belief appears, thus providing the long-run player with private information about the evolution of ($p^\ast_t)_{t\geq 0}$ (recall eqn. (9)). Consequently, after a deviation takes place, the long-run player might find it optimal to condition his actions explicitly on the current value that ($p_t, \Delta_t$) takes, which means that conforming to ($a^\ast(p^\ast_t))_{t\geq 0}$ thereafter is, in all likelihood, suboptimal.\footnote{An alternative explanation for the difference between the ratcheting equation and classic Euler equations is as follows: because after a one-shot deviation from ($a^\ast(p^\ast_t))_{t\geq 0}$ the long-run player does not re-optimize, he bears the extra cost (or benefit) of moving along the demand schedule $p \mapsto a^\ast(p)$; hence the ratcheting term in the right-hand side of (16) that modifies the rate of return on belief asymmetry. This cost (or benefit) is absent in standard Euler equations, as the continuation strategy in those cases is optimal. Extending the idea of unprofitable one-shot deviations to off-path histories does not simplify the problem of checking global incentive-compatibility: by Bellman’s principle of optimality, this is equivalent to solving (13)-(14).}

Recall that the necessary conditions (16)-(17) were derived in a heuristic fashion. Namely, I assumed (i) considerable differentiability on the long-run player’s value function, and that
(ii) the long-run player’s best-response was Markov off the equilibrium path. The next result dispenses with these assumptions, and takes the system (16)-(17) as the primitive object. That is, starting from a solution \((q, U)\) to this system, and, assuming that the market constructs beliefs about the hidden state using \(q\), it determines conditions under which all alternative strategies (Markov and non-Markov ones) are unprofitable:

**Theorem 1 (Verification Theorem).** Suppose that there exists \(\psi > 0\) such that \(g''(a) \geq \psi > 0\) for all \(a \in A\) (i.e., \(g\) is strongly convex). Let \(U \in C^2(\mathbb{R})\) and \(q \in C^2(\mathbb{R})\) denote a solution to the system (16)-(17) with the following properties:

(i) There exists \(C_1, C_2, C_3 > 0\) such that:
   \[ |U(p)| \leq C_1(1 + |p|^2) \text{ (quadratic growth)}, \]
   \[ |U'(p)| \leq C_2(1 + |p|) \text{ (linear growth)}, \]
   \[ |q(p_1) - q(p_2)| \leq C_3|p_1 - p_2| \text{ (Lipschitz)}; \]

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

(iii) \(U''\) and \(q'\) satisfy
   \[ |U''(p) - q'(p)| \leq \frac{\psi(r + 4\beta + 2\kappa)}{4\beta^2}, \text{ for all } p \in \mathbb{R}. \] (18)

Then the process \(a^*_t := g^{-1}(\beta q(p_t)), t \geq 0\) corresponds to a Markov equilibrium.

*Proof:* See the Appendix. \(\square\)

Theorem 1 is a *one-shot deviation principle* for the class of games under study. It takes the form of a verification theorem (i.e., it requires finding a solution to a specific equation) for finding Markov equilibria, akin to results in dynamic programming that rely on HJB equations for finding optimal policies in Markov form. The advantages of this result over the HJB approach based on solving (13)-(14) are enormous: the theory of ODEs provides existence results for (16)-(17); it offers much cleaner economic insights (next); and the numerical computation of a system of ODEs is much simpler. Moreover, this result is particularly weak in its requirements—the class of one-shot deviations behind the ratcheting equation is small.

The theorem is, except for condition (18), technical. It is based on two important insights nonetheless. First, the conditions on \(U\) ensure that its unique solution is given by

\[ U(p) := E\left[\int_0^\infty e^{-rt}[u(\chi)(p_t) - g(g^{-1}(\beta q(p_t)))]dt\right| p_0 = p, \]
i.e., the payoff from following the market’s conjecture of equilibrium play \( a^*_t = g^{t-1}(\beta q(p_t)) \), \( t \geq 0 \). Second, the rest of the assumptions are used to construct an upper bound to the long-run player’s payoff under any alternative strategy, with the important property that it coincides with \( U(\cdot) \) on the equilibrium path. Under this type of approximation to the solution to the nonlocal PDE (13)-(14), \( U \) is, by construction, an upper bound to the long-run player’s payoff when beliefs are aligned. But since \( U(\cdot) \) can be achieved by following \( (a^*_t)_{t\geq0} \), it follows that inducing no belief asymmetry is optimal.

Condition (18) is economically meaningful. While \( q(p) \) corresponds to the on-path value of a marginal boost in the public belief (recall that \( \Delta := p^* - p \)), \( U'(p) \) corresponds to the long-run player’s marginal utility along the equilibrium path. Hence, in both \( q \) and \( U' \) the public belief is boosted marginally. However, \( q \) encapsulates a strictly positive stock of belief asymmetry that decays exponentially, whereas in \( U' \) no belief asymmetry is created (as the private belief also changes marginally). Consequently, \( q(p) - U'(p) \) captures the sensitivity of payoffs to private information: it isolates the change in payoffs that is the result of the presence of private information only. As \( q'(p) - U''(p) \) grows, the long-run player’s payoff becomes more convex in the potential benefit of private information, and hence engaging in double deviations (i.e., deviations from \( a^* \) after a deviation from it has occurred) is more attractive. Condition (18) states that an equilibrium exists if \( q'(p) - U''(p) \), the value of private information, is not too large. The absolute value appears because private information can arise due to upward deviation (which leads to \( p^* > p \), and hence to a sensitivity of \( q - U' \)), or due to a downward deviation (which leads to \( p > p^* \), and hence to a sensitivity of \( U' - q \)).

Given any solution \((q,U)\) to (16)-(17), (18) can always be verified ex-post. In Section 5, however, I study specific environments for which I am able to find conditions on the primitives of the model ensuring that this second-order condition holds. In the particular case of linear environments (i.e., \( u(\chi(p)) \) is linear in \( p \)—the main case studied in the existing literature—there is always an equilibrium in which actions are constant (the ratcheting equation admits a constant solution). In fact, the additive separability in the public signal, coupled with the linearity of the flow payoff, yields expected discounted payoffs that are additively separable in \( p \) and \( \Delta \), which implies that the benefit from inducing belief asymmetry is independent from the private information that the long-run player may have. In linear settings therefore, having private information about the public belief has no value for incentives, and thus \( U'' - q' \equiv 0 \). Beyond linear settings, the possibility of becoming privately informed about the unobserved state will affect incentives non-trivially (through cost-smoothing and ratcheting), as inducing belief asymmetry is valued differently across different levels of the public belief.

Remark 3. This first-order approach to analyzing incentives has been successful in the...
literature on optimal contracts in environments where payoff-relevant variables have high persistence. The approximation of payoffs as a method to obtain sufficient conditions for optimality is an insight first developed by Williams (2011) in a reporting problem with adverse selection, and later followed by Sannikov (2014) and Prat and Jovanovic (2014). However, the additional complexity of embedding the agent’s optimization problem into the principal’s one has allowed them to verify their sufficient conditions in isolated environments only. Instead, in Section 5 I show that limiting the degree of commitment of the “principal” (here, a market) leads to a tractable characterization of incentives for a wide range of economic settings.

4.3 Applications

In this Section I show how the methods developed in this paper can be used to address interesting questions from applied-theory work.

4.3.1 Reputation and the Ratchet Effect

A firm’s (cumulative) earning process is given by

\[ dY_t = \theta_t dt + \sigma \theta_t dZ_t, \quad t \geq 0, \]

where \((\theta_t)_{t \geq 0}\) corresponds to the firm’s fundamentals, which I interpret as managerial ability. The manager can misreport performance in the report \((\xi_t)_{t \geq 0}\) released to the public:

\[ d\xi_t := a_t dt + dY_t = (a_t + \theta_t) dt + \sigma dZ_t, \quad t \geq 0, \]

where \(a_t\) denotes the degree of manipulation exerted at time \(t \geq 0\). Assume further that \(d\theta_t = \sigma \theta_t dZ_t, \quad t \geq 0\), so managerial ability is a Brownian martingale.

In order to isolate the incentives for belief manipulation, I assume that (i) manipulation is based on accounting methods (e.g., discretionary accruals—typically hard to detect), and that (ii) the firm pays its dividends far ahead in the future. Since in this case manipulating earnings does not impose real costs on the firm, the public belief \(p^*_t := \mathbb{E}[\theta_t | \mathcal{F}^t]\) is a measure of the manager’s market value (i.e., his reputation). In particular, the market expects true earnings over the next period \([t, t + dt)\) to take value

\[ \mathbb{E}^{a^*}[d\xi_t - a^*_t dt | \mathcal{F}^t] = p^*_t dt. \]

The manager’s flow payoff is increasing in his reputation:

\[ h(p^*_t) := u(\chi(p^*_t)), \quad h'(\cdot) > 0 \]

(i.e., the manager always has a myopic incentive to inflate earnings). Also, manipulation entails a private cost to the manager: \(\psi a^2_t / 2, \quad \psi > 0\) (e.g., effort disutility associated with finding intricate accounting techniques, or a reluctance to engage in “creative accounting”).

As a benchmark, suppose that the manager’s payoff is linear, i.e., \(h(p^*_t) = \alpha p^*_t\), some
\( \alpha > 0 \). It is easy to verify that the ratcheting equation (16) admits a constant solution

\[
q(p) = \frac{\alpha}{r + \beta} \Rightarrow a^* = \frac{\beta}{\psi} q(p) = \frac{\alpha \beta}{\psi (r + \beta)}.
\]  

(19)

In this equilibrium, both cost-smoothing motives and ratcheting forces are absent \((q' \equiv q'' \equiv 0)\). The intensity of manipulation increases both with the size of the manager’s myopic incentives \((\alpha)\) and with the sensitivity of beliefs to new information \((\beta)\); it decays as manipulation becomes more costly \((\psi)\) and as the manager’s impatience grows \((r)\).

Most importantly, the linearity assumption predicts that all managers exert the same degree of manipulation, regardless of the individual performance of the firms they run.\(^{29}\)

However, there is evidence that suggests that this type of manipulation is stronger in the proximity of some key thresholds or benchmarks. In particular, it has been documented that managers face strong market pressures to avoid (i) reporting losses, (ii) reporting negative earnings growth, and (iii) failing to meet analysts’ forecasts (Degeorge et al. 1999).\(^{30}\)

In light of this evidence, suppose that the manager’s short-term incentives are no longer uniform: while the manager, as in the linear benchmark, always has a local incentive to boost reported earnings \((h' > 0)\), his myopic incentives are more acute when the market expects the firm to generate zero earnings over the next period (i.e., when \(p_t^* = 0)\).

**Assumption 2.** \(h \in C^3(\mathbb{R}; \mathbb{R}), h'(\cdot) > 0 \) and symmetric around zero, \(h'(p^*) \rightarrow 0 \) as \(p^* \rightarrow \pm \infty\), \(h'(\cdot) \) has a unique local maximum (by symmetry, at zero), and \(h'''(0) < 0\).\(^{31}\)

The ratcheting equation can be used to characterize the equilibrium manipulation policy, \(a^*(p) = \beta q(p)/\psi\). Given the analysis from the linear case, a natural question that arises is whether those managers who run firms that stand exactly at the threshold indeed manipulate earnings more actively. At a first glance, one would expect this to be the case. This is because, in equilibrium, the public belief \((p_t^*)_{t \geq 0} \) (i) evolves exogenously, and, as any martingale, (ii) its future evolution is unpredictable. In the class of games studied, however, the equilibrium behavior of payoff-relevant variables gives no guidance as to what actions are the ones that precisely arise in equilibrium. Using results from Section 5.2 (bounded marginal flow payoffs) we obtain the following:

**Proposition 3.** In equilibrium: \(q \in (0, h'(0)/(r + \beta)) \) and \(q(p^*) \rightarrow 0 \) as \(p^* \rightarrow \pm \infty\); \(q' > 0 \) in \((-\infty, 0)\), \(q'(0) > 0 \) and \(q''(0) < 0\); and \(h(\cdot) \) has unique local maximum located to the right of zero. Furthermore, \(q \) is skewed to the right, i.e., \(q^*(p) > q^*(-p) \) for all \(p > 0\).

\(^{29}\)The same occurs in Holmström (1982/1999) and in Stein (1989) (albeit in slightly different environments), as both rely on the linearity assumption on the manager’s payoff for tractability reasons.

\(^{30}\)See also Burgstahler and Dichev (1997), Burgstahler and Chuk (2012) and Dichev et al. (2013) for statistical and survey-based methods that try to identify this manipulation.

\(^{31}\)Symmetry is assumed just to illustrate distortions more clearly.
Proposition 3 uncovers two interesting distortions. First, equilibrium incentives are depressed at zero relative to the corresponding linear benchmark of slope $\alpha = h'(0)$ (i.e., $q(0) < \chi'(0)/(r + \beta)$). Second, the manipulation policy is maximized to the right zero, despite those managers having weaker myopic incentives, and also being unable to effectively affect the value of the firms they run. Managers turn out to have, on average, stronger incentives to maintain high earnings, rather than to build them up. Graphically:

![Figure 1: The equilibrium manipulation policy $a^*$ and the linear benchmark $\beta h'(p)/\psi(r + \beta)$](image)

The source of these two distortions is the ratcheting created in the model. To see this, observe first that, since the manager’s myopic incentives become stronger as $p$ approaches zero from the left, the market will conjecture a strictly increasing manipulation profile in this region. The ratcheting imposed by the market then yields incentives that fall below the corresponding linear benchmark in a neighborhood to the left of zero. Suppose now that the market’s conjecture is actually maximized at zero. In this case, anticipating that his myopic incentives and the market’s conjecture will both drop to the right of zero, the manager would be incentivized to engage in more manipulation either at, or before, the zero-earnings threshold; in other words, by acting in the opposite direction (i.e., encouraging more manipulation), the ratcheting creates a profitable deviation. The market must therefore ratchet up the demand for manipulation at zero ($q'(0) > 0$) in order to value the firm correctly, which results in a manipulation profile that is skewed to the right of the threshold.\(^{32}\)

To conclude, observe that the model generates non-trivial implications for the distribution of earnings: in particular, because $a^*(\cdot)$ is nonlinear, $(\xi_t)_{t \geq 0}$ ceases to be Gaussian. The model can be used to answer applied questions such as teasing out firm-individual manipulation profiles, or testing the normality assumption on the fundamentals. Importantly, these

\(^{32}\)Cost-smoothing affects the level, yet not the skewness of incentives: letting $\psi \to \infty$ the ratcheting term disappears and the ratcheting equation becomes $(r + \beta)q(p) = h'(p) + \sigma^2 q''(p)/2$, which admits a symmetric solution around zero. It is easy to see that, to the right of zero, both the ratcheting and cost smoothing generate incentives are stronger than the linear benchmark.
exercises can be performed beyond the area of earnings management, using the primitives of the model and the predicted distribution of the relevant observables.

4.3.2 Expectation Traps: Multiplicity of Equilibria

As in the previous example, \((\xi_t)_{t \geq 0}\) are reported earnings, \((\theta_t)_{t \geq 0}\) a Brownian martingale \((\kappa = 0)\), and \((p_t^*)_{t \geq 0}\) the market’s belief about the growth of cumulative earnings. In contrast to the previous application, manipulating earnings can have a nonlinear impact on the firm’s market value. More precisely, the firm’s (normalized) stock price is given by

\[
E_t^a \left[ r \int_t^\infty e^{-r(s-t)} (d\xi_s - (a_s^*)^2 ds) \bigg| \mathcal{F}_t^\xi \right] = p_t^* + E_t^a \left[ r \int_t^\infty e^{-r(s-t)} a_s^*(1 - a_s^*) ds \bigg| \mathcal{F}_t^\xi \right].
\]

In the left-hand side, the term \(-(a_t^*)^2\) captures the loss in market value due to the market’s awareness that reports do not truly reflect true performance.\(^{33}\) Under this specification, the market rewards moderate levels of manipulation (i.e., based on acceptable accrual-based techniques). Higher levels of manipulation \((a^* > 1)\) instead require real transactions that may reduce long-term profitability. The market also punishes downward manipulation: for instance, if agency problems are a concern, or if the firm has exhausted all its profitable investment opportunities. The firm’s value is maximized at \(a^* = 1/2\).

The manager cares about the firm’s long-term performance, but he also enjoys beating the market’s next-period earnings prediction. For simplicity, I assume that the manager attaches equal weight to both components in his flow payoff:

\[
\begin{align*}
\frac{d\xi_t - (a_t^* + p_t^*) dt}{\text{short-termism}} + \left( p_t^* + E_t^a \left[ r \int_t^\infty e^{-r(s-t)} a_s^*(1 - a_s^*) ds \bigg| \mathcal{F}_t^\xi \right] \right) dt - \left( a_t + \frac{\psi}{2} a_t^2 \right) dt.
\end{align*}
\]

Notice that the first term can lead to a ratchet effect: if showing good performance leads to a higher \(a^*\) tomorrow, it will be harder for the manager to surprise the market again. Using that \(d\xi_t = (a_t + p_t) dt + \sigma dZ_t\) from the manager’s perspective, the relevant component (for optimization purposes) of the manager’s flow payoff is given by

\[
-a_t^* dt + E_t^a \left[ r \int_t^\infty e^{-r(s-t)} a_s^*(1 - a_s^*) ds \bigg| \mathcal{F}_t^\xi \right] - \frac{\psi}{2} a_t^2.
\]

We can immediately immediately verify that \(a^* = 0\) is an equilibrium: because the manager attaches equal weight to both the firm’s short- and long-run performance, flow payoffs become

---

\(^{33}\)The equality comes from the public belief evolving as a martingale given the public information, and \((\xi_t)_{t \geq 0}\) being driven by \((a_t^*)_{t \geq 0}\) from the market’s perspective.
independent of \((p^*_t)_{t \geq 0}\), and hence there are no gains to manipulating earnings if the market conjectures that this will not happen. This is inefficient, as some degree of manipulation increases the firm’s value.

Interestingly, there is also a linear equilibrium. Guessing \(a^*(p^*) = \beta \alpha p^*/\psi\), and using that \(p^*_s = p^*_t + \sigma Z^*_{s-t}\), we obtain \(u(x(p^*_s)) - g(a_t) = -\beta^2 \alpha^2 (p^*_t)^2/\psi^2 - \psi a^2_t/2 + \text{constant}\). The ratcheting equation can be used to find \(\alpha\). Sufficiency is verified Proposition, which relies on an existence result for linear-quadratic games (Theorem 2, Section 5.1):

**Proposition 4.** For all parameters \((r, \psi, \sigma, \xi, \sigma_\theta) \in \mathbb{R}^4_+ \setminus \{0\}\), there exists a linear equilibrium \(a^*(p^*) = \beta \alpha p^*/\psi\). In this equilibrium, \(\alpha < 0\).

*Proof:* See the appendix. \(\Box\)

In this equilibrium, the market believes that the manager is likely to be manipulating earnings. The market then imposes a negatively sloped manipulation profile so as to minimize the costs of this action on the firm’s value. Interestingly, the target of this demand schedule is \(a^* \equiv 0\), and not the efficient bliss point \(1/2\). In fact, if the latter point were the target, the ratchet effect would incentivize the manager to under-manipulate earnings at that market price: by inducing the market be more pessimistic about the firm’s prospects, his short-term incentives are easier to fulfill tomorrow.

This *multiplicity* result shows how nonlinearities present in the market’s action can generate non-trivial predictions on economic outcomes. Multiple equilibria have been shown to exist only in models that allow for complementarities between actions and the unobserved state, and restricted to settings with only two periods of interaction (see Dewatripont et al. 1999 for an application to career concerns).

### 4.3.3 Dynamic Oligopoly with Symmetric Incomplete Information

The tools developed in this paper can be easily adapted to settings in which \(N\) long-run players can symmetrically affect the signal about an unobserved state. I will illustrate these ideas in the context of a dynamic Cournot duopoly with unobserved demand intercept.

Two firms produce a homogeneous product continuously over time. Firm \(i\)’s profit over \([t, t + dt]\) is given by \(a^i t [d\xi_t - cdt]\), where \(c > 0\) is the marginal cost of production (common across firms) and \(a^i t\) its current output level, \(i = 1, 2\). The process

\[
d\xi_t = (\theta_t - [a^1_t + a^2_t]) dt + \sigma \xi dZ^\xi_t, \quad t \geq 0,
\]

(20)

corresponds to the good’s cumulative price process, where \(\theta_t\) denotes the demand intercept.
at time $t$. This intercept is always hidden and follows the mean-reverting process (1) around $\eta > 0$. In order for the model to be meaningful, I assume that $\eta > c$.

I will search for symmetric equilibria in which on-path behavior depends on the common belief about the intercept. To this end, it suffices to assume that firm $-i$ follows the conjectured equilibrium strategy $a_i^* = a^*(p_i^*)$, while $i$ has unilaterally deviated. The relevant process for firm $i$'s learning purposes is

$$dY_t := d\xi_t + [a_t^i + a^*(p_t + \Delta)] dt = \theta_t dt + \sigma_t dZ_t^i.$$ 

Letting $p^* := p + \Delta$, the pair $(p, \Delta)$ will evolve according to

$$dp_t = -\kappa(p_t - \eta) dt + \sigma dZ_t$$

$$d\Delta_t = [-(\beta + \kappa) \Delta_t + \beta(-a_t^i + a^*(p_t + \Delta_t))] dt,$$

where $(Z_t)_{t \geq 0}$ is a $\mathbb{F}^Y$-Brownian motion. Observe that, by increasing output above $a_t^*$, firm $i$ leads its rival to become more pessimistic about demand.

It is easy to verify that the ratcheting equation now becomes

$$\left[ r + \beta + \kappa - \beta \frac{da^*}{dp}(p) \right] q(p) = -a^*(p) \frac{da^*}{dp}(p) - q'(p)\kappa(p - \eta) + \frac{1}{2}\sigma^2 q''(p) \quad (21)$$

s.t. $a^*(p) = \frac{p - c}{3} - \frac{\beta}{3} q(p) \quad (22)$

where $q(p)$ is the value of creating a (hypothetical) unit of belief asymmetry, and where (22) corresponds to firm $i$'s first-order condition (after imposing symmetry). In this environment, the term $da^*/dp$ plays a dual role. First, it captures firm $i$'s change in current profits due to its rival immediate reaction (first term on the right in (21)). Second, through the “ratcheting channel,” it captures how its rival’s immediate reaction affects future quantities (and hence, prices) through the persistence of the learning process (left-hand side of (21)).

The previous equation admits a linear solution:

**Proposition 5.** Let $q(p) = q_1 + q_2 p$ denote a linear Markov equilibrium. Then, $q_2 < 0$, and it is independent of $\eta > 0$ and $c > 0$. Consequently, there is $\bar{p} > 0$ such that $a^*(p) > (p - c)/3$ for all $p > \bar{p}$. A sufficient condition for $\bar{p} < c$ (i.e., $a^*(c) > 0$) is that $\eta/c > (\beta q_2 - 1)/9\kappa q_2 \geq 0$.

**Proof:** See the appendix. \(\Box\)

In any linear equilibrium, there is always a threshold above which firms over-produce relative to the static Cournot equilibrium under complete information (which is trivially

---

34 Riordan (1985) recognizes this derivative a conjectural variation—in the current model, fully rational and micro-founded through a learning process. In his model, a demand intercept also has persistence, but its current hidden value always get revealed a period after as profits gets realized.
Markov). This is intuitive: if firm \(-i\) plays myopically, then, firm \(i\) obtains first-order gain after a small output increase that makes firm \(-i\) more pessimistic about future demand, at the expense of only a second-order loss in current profits.

The incentives to produce above the static Nash can be quite strong. In fact, when \(p \leq c\), firm \(i\) knows that there is a high chance that the intercept will revert back to \(\eta > c\), region in which the market becomes profitable. This short-run predictability, coupled with flow profits that are concave in quantities, means that there are potential gains to smooth out the cost and benefits of manipulating a rival’s belief. At \(p \leq c\), these gains outweigh the cost of operating at a loss if the expected surplus—as measured by \(\eta/c\) in this setting—is sufficiently large. These inter-temporal smoothing motives are much less acute in corresponding two-period environments (e.g., Mirman et al., 1983), as the presence of a last period of interaction pins down behavior at the end of the game (i.e., playing the static Cournot equilibrium).

The choice of a linear demand function is for expositional reasons: the dynamic of \(\Delta\) preserves the linearity in \(a - a^*\) presented in the baseline model of Section 3. The methods can be extended to price processes of the form \(d\xi_t = (\theta_t - F(a^1_t + a^2_t))dt + \sigma dZ_t^\xi\) (with \(F(0) = 0\) and \(F'(\cdot) < 0\)), as in this case the model still maintains the additive separability between state and actions, which is the key assumption for an ODE-characterization of incentives.

5 Existence of Markov Equilibria

This section exploits the tractability of Theorem 1 to derive two existence results of Markov equilibria that can be easily computed using the ratcheting equation. The environments under study are (i) a class of linear-quadratic games, and (ii) a class of games of bounded marginal flow payoffs. In the former, equilibrium behavior is linear in the public belief. In the latter, equilibria is fully nonlinear.

5.1 Linear-Quadratic Games

Definition 2. A game of belief manipulation is linear-quadratic if: \(A = \mathbb{R}\); \(g(a) = \frac{\psi}{2}a^2\), \(\psi > 0\); and \(h(p^*) := u(\chi(p^*, a^*)) = u_0 + u_1p^* - u_2p^*_t\), where \(u_0, u_1 \in \mathbb{R}\) and \(u_2 \geq 0\).\(^{35}\)

The next result shows the existence of a Markov equilibrium that is linear in beliefs:

---

\(^{35}\)The analysis can be easily extended to the case in which the market’s action is also linear in \(a^*\): \(u(\chi(p^*, a^*)) = u_0 + u_1(k_1p^* + k_2a^*) - u_2(k_1p^*_t + k_2a^*_t)^2, k_1, k_2 \in \mathbb{R}\).
Theorem 2. Suppose that

\[ u_2 \leq \frac{\psi(r + \beta + 2\kappa)^2}{8\beta^2(\kappa)}. \]  

(23)

Then, a Markov equilibrium \( a^*_t = \beta[q_1 + q_2p^*_t]/\psi, t \geq 0 \), exists. The coefficients are given by

\[ q_1 = \frac{\eta\kappa q_2 + u_1}{r + \beta + \kappa + \frac{\beta^2}{\psi}q_2}, \quad \text{and} \]

(24)

\[ q_2 = \frac{\psi}{2\beta^2} \left[ -(r + \beta + 2\kappa) + \sqrt{(r + \beta + 2\kappa)^2 - \frac{8u_2\beta^2}{\psi}} \right] < 0. \]

(25)

Proof: See the Appendix. \qed

In this class of games, flow payoffs are quadratic and, except for \( a^* \), the dynamics of \( (p, \Delta) \) in the best-response problem are linear. Hence, it is natural to conjecture a quadratic value for the long-run player’s best-response problem, case in which incentives are linear. Provided (23) holds, it is easy to show analytically that a solution \((q, U)\) with these characteristics exists. Also, because the value of private information \( U''(q) - q'(p) \) is constant, a less demanding version of the second-order condition (18) can be developed, and verified to hold too. Thus, the conditions for existence in Theorem 1 are satisfied.

The intuition behind the curvature condition (23) is as follows. As \( u_2 \) grows, the myopic benefit from inducing belief asymmetry increases. Thus, in order prevent the long-run player from inducing belief asymmetry, the market must 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, a more negatively sloped conjecture also increases the rate of return on belief asymmetry \( r + \beta + \kappa + \frac{\beta^2}{\psi} \frac{da^*_t}{dp} \), thus incentivizing more manipulation. Consequently, when the curvature condition is violated, a linear conjecture cannot control simultaneously (i) the long-run player’s myopic incentives, and (ii) his incentives to manipulate due to the market ratcheting down its demand for manipulation too quickly.

To conclude, if flow payoffs are linear, the curvature condition is trivially satisfied (i.e., \( u_2 = 0 \)). Hence, there is always an equilibrium in which actions are constant (i.e., \( q_2 = 0 \)):

\[ q(p) = \frac{u_1}{r + \beta(\kappa) + \kappa} \Rightarrow q'(a^*) = \beta \frac{u_1}{r + \beta(\kappa) + \kappa}, \]

where the dependence \( \beta = \beta(\kappa) \) is made explicit. In this case, \( U(p) \) is linear, and hence \( U'' - q' \equiv 0 \) (private information has no value). The equilibrium found by Holmström (1999) in the stationary case corresponds to the specification \( u_1 = 1 \) and \( \kappa = 0 \).
5.2 Bounded Marginal Flow Payoffs

Definition 3. A game of belief manipulation is one of bounded marginal flow payoffs if:

(i) $\chi(p, a^*) = \chi(p)$ and $g : A \to \mathbb{R}$ is strongly convex ($g'' > \psi$, some $\psi > 0$).

(ii) There exist $m, M \in \mathbb{R}$ s.t. $-\infty < m := \inf_{p \in \mathbb{R}} (u \circ \chi)'(p) \leq \sup_{p \in \mathbb{R}} (u \circ \chi)'(p) := M < \infty$.

(iii) Both $\lim_{p \to -\infty} (u \circ \chi)'(p)$ and $\lim_{p \to \infty} (u \circ \chi)'(p)$ exist.

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

Condition (i) states that only the underlying fundamental is valued by the market, and hence manipulating the signal is purely wasteful (i.e., non-zero manipulation arises in equilibrium, yet the public belief evolves exogenously).\(^{36}\) Condition (ii) says that marginal flow payoffs are bounded. Condition (iii) states that flow payoffs become asymptotically linear, suggesting that the solution to the ratcheting equation must look like its corresponding analog in the linear case (Holmström, 1999). Finally, (iv) ensures that incentives are interior.

I first show that there exists a solution $(q, U)$ to (16)-(17):

Proposition 6 (Existence of Bounded Solutions to the Ratcheting Equation). There exists $q \in C^2(\mathbb{R})$ solution to the ratcheting equation such that

$$q(p) \in \left[\frac{m}{r + \beta(\kappa) + \kappa}, \frac{M}{r + \beta(\kappa) + \kappa}\right], \text{ for all } p \in \mathbb{R}. \quad (26)$$

Moreover, any solution satisfying (26) also verifies that

$$\lim_{p \to -\infty} q(p) = \frac{\lim_{p \to -\infty} (u \circ \chi)'(p)}{r + \beta(\kappa) + \kappa} \text{ and } \lim_{p \to \infty} q(p) = \frac{\lim_{p \to \infty} (u \circ \chi)'(p)}{r + \beta(\kappa) + \kappa}. \quad (27)$$

If $\kappa > 0$, such solution also has a bounded derivative. When $\kappa = 0$ there exists a $C^2$-solution that, in addition to satisfying (26) and (27), it also has a bounded derivative.

Proof: See the Appendix. \(\square\)

Any solution satisfying (26)-(27) and that also has a bounded derivative over all $\mathbb{R}$ will be referred to as a bounded solution of the ratcheting equation.\(^{37}\)

\(^{36}\)The results can be extended to allow for the long-run player’s actions having real effects (i.e., $\chi$ depending on $a^*$) by extending the condition (ii) in Definition 3 to partial derivatives appropriately.

\(^{37}\)A bounded derivative ensures that the Lipschitz condition on $a^*(p) := g^{-1}(q(p))$ is satisfied, and that the transversality and growth conditions on $U'$ in Theorem 1 hold.
Proposition 7 (Long-Run Player’s Equilibrium Payoff). Let \( q \) denote a bounded solution of the ratcheting equation. Then, there exists a unique \( U \in C^2(\mathbb{R}) \) solution to the ODE (17). This solution is given by

\[
U(p) = \mathbb{E} \left[ \int_0^\infty e^{-rt} [(u \circ \chi)(p_t) - g(g^{-1}(\beta q(p_t)))] dt \mid p_0 = p \right] 
\]

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

Proof: See the Appendix.

Finally, I establish conditions on the primitives \( (r, m, M, \psi, \kappa, \sigma, \theta, \sigma_\xi) \) that ensure that \((q, U)\) as above meets the requirements of Theorem 1. I do this for \( \kappa = 0 \): the environment that, from an ex-ante perspective, generates the largest returns from belief manipulation.\(^{38}\)

Theorem 3 (Existence of Markov Equilibrium). Let \( q : \mathbb{R} \to \mathbb{R} \) denote a bounded solution to the ratcheting equation. Then, the sensitivity \( U' - q' \), and the value of private information \( U'' - q'' \), take the form

\[
U'(p) - q(p) = \mathbb{E} \left[ \int_0^\infty e^{-(r+\kappa)t} \beta q(p_t) dt \mid p_0 = p \right] \quad \text{and} \\
U''(p) - q'(p) = \mathbb{E} \left[ \int_0^\infty e^{-(r+2\kappa)t} \beta q'(p_t) dt \mid p_0 = p \right],
\]

respectively, where \( dp_t = -\kappa(p_t - \eta)dt + \sigma dZ_t \), \( p_0 = p \). Moreover, when \( \kappa = 0 \), (30) admits an analytic solution in terms of \( q(\cdot) \). In this case, if

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

the process \((g^{-1}(\beta q(p_t)))_{t \geq 0}\) is a Markov equilibrium.

Proof: See the Appendix.

Theorem (3) proves the existence—and provides a full characterization—of Markov equilibria in which equilibrium actions are nonlinear.\(^{39}\) A key expression for understanding

\(^{38}\)The ratio \( \beta(\kappa)/\beta(\kappa + \kappa) \) is maximized at \( \kappa = 0 \). When this ratio increases, either (i) beliefs react more strongly to signal surprises (numerator), or (ii) belief asymmetry has more persistence (denominator), both effects making belief manipulation more profitable.

\(^{39}\)Kovrijnykh (2007), Martinez (2009) and Bar-Isaac and Deb (2014) study career concerns in restricted settings that also incorporate (specific) nonlinearities, and where the question of sufficiency is not addressed.
the structure of the games under study is (29), which writes the sensitivity of the long-run player’s payoff with respect to private information as a present value of (hypothetical) marginal values of belief asymmetry. The appearance of changes in continuation values (i.e., \( q(\cdot) \))—rather than marginal flows—is due to the high persistence in the environment. More concretely, consider a marginal change in \( p_0 \) (the initial value of the private belief). Then, from (9), this change affects all future values of \( p_t^* \), thus affecting all future flow payoffs; the continuation value starting at zero then gets affected. But because the private belief itself has persistence, \( p_{dt} \) also changes, which in turn affects \( (p_t^*)_{s>dt} \), and hence all future flow payoffs get affected again; the continuation payoff starting at \( dt \) changes too, and so on. The result states that each continuation value reacts by \( \beta q(p_t), t > 0 \), and that this change decays at rate \( \beta + \kappa \), the (equilibrium) rate of decay of \( \Delta \) (or \( p^* \)).

To understand the economics behind (29), suppose that \( m \geq 0 \); i.e., the long-run player’s flow payoff is non-decreasing. Now, consider a local downward deviation from the market’s current conjecture of equilibrium play. In this case, (i) the long-run player would become more optimistic than the market about the underlying fundamental, and (ii) payoffs change, by definition, by \( U'(p) - q(p) \). Because \( q(\cdot) \geq 0 \) when \( m \geq 0 \) (Proposition 6), the right-hand side in (29) states that the long-run player expects his payoff to increase. Intuitively, becoming more optimistic about the hidden state leads the long-run player to expect the public belief to drift up (see footnote (41) below), and hence to expect all future continuation values to be positively affected. The long-run player thus benefits from downward deviations in a setting where, locally, higher actions yield higher payoffs—a ratchet effect is created.41

To conclude, condition (31) is derived by obtaining a bound for the value of private information, \( U''(p) - q'(p) \), and imposing that (18) in Theorem 1 holds.42 Condition (31) is relaxed when manipulation becomes more costly (\( \psi \) increases) and when \( \sigma_\theta/\sigma_\xi \) decreases: in the latter case, beliefs respond less strongly to new information, thus reducing the benefits of belief manipulation. The condition is also relaxed when \( M - m \) decays, and trivially satisfied when payoffs are linear (\( M = m \)), as private information ceases to have any value.

6 Conclusions

This paper has analyzed a class of continuous-time games where agents learn about the economic environment. The framework is particularly useful for analyzing incentives in settings where agents roughly share the same degree of uncertainty at the outset, and where experi-

---

40Expression (9): \( dp_t^* = [-\kappa(p_t^* - \eta) + \beta(a_t - a_t^*) + \beta(p_t - p_t^*)]dt + \sigma dZ_t, t \geq 0. \)

41This also appears in the contracting environments of Prat and Jovanovic (2014) and Bhaskar (2014).

42For \( \kappa > 0 \) the bound also exists (as \( q' \) is bounded), but the hurdle to tie it back to primitives is purely technical. Refer to Remark 6 in the proof of Theorem 3 in the Appendix.
mentation motives are likely to have second-order effects on incentives. A key contribution of the techniques presented here is that they considerably expand the class economic questions that can be studied under the umbrella of signal-jamming models. In the process, I have provided a systematic treatment of the issue of private beliefs off the equilibrium path, which crucially affects equilibrium behavior when learning and nonlinearities are present.

Assuming ex-ante symmetric uncertainty is not just a natural benchmark for studying incentives in games of learning; it is crucial from the point of view of applications. First, observe that if the long-run player had ex-ante superior information than the market, his actions would naturally incorporate part of his private information. The public signal would therefore convey statistical information both of the hidden state and of the long-run player’s private information, and the market would attempt to optimally “filter” both. Beyond linear-quadratic games (linear dynamics, linear-quadratic payoffs), or settings in which the fundamental takes finite values, the long-run player’s action will be a nonlinear function of his private information, which results in posterior beliefs that can no longer be reduced to a set of finite-dimensional statistics. Second, the combination of ex-ante symmetric uncertainty and learning makes this framework a tractable one for studying ratchet effects in fully dynamic settings. Ex-ante symmetric uncertainty is therefore a convenient modeling technique if the goal is to understand strategic behavior in the presence of learning and nonlinearities, or if the goal is to analyze ratchet-like forces.

Finally, I discuss two possible extensions of the methods presented. First, by delivering nontrivial predictions for the distribution of the observable, the model can be used to address empirical questions in environments with inherent nonlinearities. One such potential area is labor markets, where market imperfections are known to affect the wage structure nonlinearly across the skills distribution (Acemoglu and Pischke, 1999). Because in this case different workers will react to this wage compression differently, the observed distribution of wages is likely to be determined endogenously by the effort choices of the workforce. Second, a theoretical extension involves studying incentives in environments where affecting the speed of learning is possible, and where information arrives frequently. While the arguments used to obtain necessary and sufficient conditions for equilibria have a direct analog in settings beyond the additively separable world, the analysis is complicated by the presence of an additional experimentation effect. These and other questions are left for future research.

References


7 Appendix: Proofs

I first prove Theorems 1, 2 and 3, and Propositions 6 and 7 which are needed for the latter. The Proofs from Section 4.3 (Applications—Propositions 3, 4 and 5) rely on the previous theorems, and relegated to the end.

*Proof of Theorem 1:* Suppose that the market constructs beliefs using \( a^*(p_t^*) := g^{-1}(\beta q(p_t^*)) \), \( t \geq 0, q \) as in the theorem. Observe first that given \((\hat{a}, a^*)\)—with \( \hat{a} \) a feasible strategy of the long-run player—the ODE (10) of \( \Delta \) admits a solution. To see this, notice that the strong convexity of \( g \) coupled with \( q \) being Lipschitz implies that \( p^* \mapsto a^*(p^*) \) is Lipschitz as well. Hence, the two-dimensional process

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

has a drift and volatility that satisfy standard Lipschitz and linear growth conditions in \( \|(p,p^*)\| \) (with a slope that is independent of \( \hat{a} \)). Also, by feasibility, \( \hat{a} \) satisfies the integrability condition \( \mathbb{E}[\int_0^t a_s^2 ds] < \infty, t \geq 0 \). Consequently, there exists a unique strong solution

36
\((p, p^*)\) to the previous SDE (Theorem 1.3.15 in Pham 2009), and hence of \(\Delta\). The long-run player’s optimization problem is thus well-defined.

Take any solution \((q, U)\) as in the theorem. Consider the function

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

(32)

I will show that, for a suitably chosen \(\Gamma\), 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 \(h(p) := (u \circ \chi)(p, a^*(p))\), and \(\hat{\Delta}\) denotes the belief asymmetry process under the pair \((a^*(p^*_t), \hat{a})\). Applying Itô’s rule to \(\hat{V}\) we obtain

\[
\frac{dV_t}{e^{-rt}} = [h(\hat{p}^*_t) - g(\hat{a}_t)]dt - r \left\{U(\hat{p}^*_t) + [q(\hat{p}^*_t) - U'(\hat{p}^*_t)]\hat{\Delta}_t + \frac{\Gamma}{2} \hat{\Delta}_t^2\right\} dt
\]

(33)

\[\begin{aligned}
&+ \left\{U'(\hat{p}^*_t)[-\kappa(\hat{p}^*_t - \eta) - \beta \hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))] + \frac{1}{2} \sigma^2 U''(\hat{p}^*_t)\right\} dt \\
&+ \hat{\Delta}_t \left\{q'(\hat{p}^*_t)[-\kappa(\hat{p}^*_t - \eta) - \beta \hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))] + \frac{1}{2} \sigma^2 q''(\hat{p}^*_t)\right\} dt \\
&- \hat{\Delta}_t \left\{U''(\hat{p}^*_t)[-\kappa(\hat{p}^*_t - \eta) - \beta \hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))] + \frac{1}{2} \sigma^2 U'''(\hat{p}^*_t)\right\} dt \]
\]

\(\begin{aligned}
&+ [q(\hat{p}^*_t) - U'(\hat{p}^*_t)](-\beta + \kappa)\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))]dt \\
&+ \Gamma \hat{\Delta}_t[-(\beta + \kappa)\hat{\Delta}_t + \beta(\hat{a}_t - a^*(\hat{p}^*_t))]dt + \text{Brownian term,}
\end{aligned}\)

where we have used that \(\hat{p}^*_t := p_t + \hat{\Delta}_t\) evolves according to \(d\hat{p}^*_t = (-\kappa(\hat{p}^*_t - \eta) + \beta(\hat{a} - \cdots\)

37
\[ a^*(\hat{p}_t^*) - \beta \hat{\Delta}_t dt + \sigma dZ_t. \] Now, using (17) and (16) we obtain

\[
\begin{align*}
(A) & = rU(\hat{p}_t^*) - h(\hat{p}_t^*) + g(\alpha^*(\hat{p}_t^*)) + U'(\hat{p}_t^*)[-\beta \hat{\Delta}_t + \beta(\hat{a}_t - \alpha^*(\hat{p}_t^*))] \\
(B) & = \left[ r + \beta + \kappa + \beta \frac{da^*(p^*)}{dp} \right] q(\hat{p}_t^*) - h'(\hat{p}_t^*) + q'(\hat{p}_t^*)[-\beta \hat{\Delta}_t + \beta(\hat{a}_t - \alpha^*(\hat{p}_t^*))] \\
(C) & = (r + \kappa)U'(\hat{p}_t^*) - h'(\hat{p}_t^*) + q(\alpha^*(\hat{p}_t^*)) \frac{da^*(p^*)}{dp} + U''(\hat{p}_t^*)[-\beta \hat{\Delta}_t + \beta(\hat{a}_t - \alpha^*(\hat{p}_t^*))]
\end{align*}
\]

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

\[
\frac{dV_t}{e^{-rt}} = [g(\alpha^*(\hat{p}_t^*)) - g(\hat{a}_t) + g'(\alpha^*(\hat{p}_t^*))(\hat{a}_t - \alpha^*(\hat{p}_t^*))]dt \\
+ \beta[\Gamma + q'(\hat{p}_t^*) - U''(\hat{p}_t^*)]\hat{\Delta}_t(\hat{a}_t - \alpha^*(\hat{p}_t^*))dt \\
- [\beta(q'(\hat{p}_t^*) - U''(\hat{p}_t^*)) + \Gamma \left( \frac{r}{2} + \beta + \kappa \right)]\hat{\Delta}_t^2 dt + \text{Stochastic integral}
\]

\[ \Rightarrow \dot{V}_t - \dot{V}_0 \leq \int_0^t e^{-rs} \left( \frac{-\psi}{2}(\hat{a}_s - \alpha^*(\hat{p}_s))^2 + \beta[\Gamma + q'(\hat{p}_s^*) - U''(\hat{p}_s^*)]\hat{\Delta}_s(\hat{a}_s - \alpha^*(\hat{p}_s^*)) \\
- \left[ \beta(q'(\hat{p}_s^*) - U''(\hat{p}_s^*)) + \Gamma \left( \frac{r}{2} + \beta + \kappa \right) \right]\hat{\Delta}_s^2 \right) ds + \text{Stochastic integral}
\]

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

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

(34)

over the set \( \mathcal{R} := \{R^*(p) | p \in \mathbb{R} \} \). It is clear that if \( \mathcal{R} \) is unbounded, no \( \Gamma \in \mathbb{R} \) satisfies (34) over the whole set \( \mathcal{R} \). Consequently, a necessary condition for the bound to hold is that \(|q'(p) - U''(p)|\) is bounded. Assuming that \( \bar{R} := \max\{|\sup(\mathcal{R})|, |\inf(\mathcal{R})|\} < \infty \), it is easy to show that (34) holds over \( \mathcal{R} \) when \( \Gamma = \bar{R} \) if \( \bar{R} \leq \psi(r + 2\kappa + 4\beta)/4\beta^2 \). Thus, a sufficient condition for \( \dot{V} \) to be a supermartingale is that \(|q'(p) - U''(p)| \leq \psi(r + 2\kappa + 4\beta)/4\beta^2 := \Gamma \).

**Remark 4.** The bound on \( R^*(p) := q'(p) - U''(p) \) can be relaxed if one has more information about the values that \( R(\cdot) \) takes. In particular, it is easy to check that

(I) if \( \mathcal{R} = \{-\bar{R}\}, \bar{R} > 0 \), (34) holds when \( \Gamma = \bar{R} \),

(II) if \( \mathcal{R} = \{\bar{R}\}, \bar{R} > 0 \), (34) when \( \Gamma = \frac{\bar{R}(r + 2\kappa - 2\beta)}{r + 2\kappa + 2\beta} \) if \( \bar{R} \leq \frac{\psi(r + 2\kappa + 4\beta)/4\beta^2}{r + 2\kappa} \); 

(III) if \( \mathcal{R} \subseteq [-\bar{R}, 0], \bar{R} > 0 \), (34) holds when \( \Gamma = \bar{R} \) if \( \bar{R} \leq \psi(r + 2\kappa + 2\beta)/\beta^2 \). \( \square \)
With this in hand, a standard localizing argument (which uses (i) in the Theorem) allows us to get rid of the stochastic integral through taking expectations, concluding that

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

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 conclude that

\[ \mathbb{E}\left[ \int_0^\infty e^{-rs} [h(\hat{p}^*_s) - g(\hat{a}_s)] ds \right] \]

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 \( U \in C^2(\mathbb{R}) \) to the ODE (17) satisfying a quadratic growth condition. Hence, there exists \( C > 0 \) such that \( |\mathbb{E}[e^{-rt}U(p_t)]| \leq e^{-rt}C(1 + \mathbb{E}[p^2_t]) \to 0 \) as \( t \to \infty \), this because \( (p_t)_{t \geq 0} \) is mean-reverting or a Brownian martingale. The Feynman-Kac representation theorem (see for instance Remark 3.5.6. in Pham (2009)) yields that \( U \) is of the form

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

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

Proof of Theorem 2: It is straightforward to verify that \( U(p) = U_0 + U_1 p + U_2 p^2 \) and \( q(p) = q_1 + q_2 p \) solve the system of ODEs (16)-(17) if and only if they solve the system:

\[
\begin{align*}
(U_0) : 0 &= r U_0 - u_0 - \eta \kappa U_1 + \frac{\beta^2}{2\psi} q_1^2 - \sigma^2 U_2 \\
(U_1) : 0 &= (r + \kappa) U_1 - u_1 + \frac{\beta^2}{\psi} q_1 q_2 - 2 \eta \kappa U_2 \\
(U_2) : 0 &= (r + 2 \kappa) U_2 + \frac{\beta^2}{2\psi} q_2^2 + u_2 \\
(q_1) : 0 &= \left( r + \kappa + \beta + \frac{\beta^2}{\psi} q_2 \right) q_1 - \eta \kappa q_2 - u_1 \\
(q_2) : 0 &= (r + \beta + 2 \kappa) q_2 + \frac{\beta^2}{\psi} q_2^2 + 2 u_2. 
\end{align*}
\]

Observe that the last quadratic admits a solution if and only if the curvature condition
(23) holds. When this is the case, all the equations can be solved in an iterative fashion. We consider
\[ q_2 = \frac{\psi}{2\beta^2} \left[-(r + \beta + 2\kappa) + \sqrt{(r + \beta + 2\kappa)^2 - \frac{8u_2\beta^2}{\psi}}\right] < 0, \]
that is, the root of \( q_2 \) with the smallest absolute value (this ensure that transversality conditions hold). In particular, \( r + \beta + \kappa + \beta^2q_2/\psi > 0 \), which yields
\[ q_1 = \frac{u_1 + \eta\kappa q_2}{r + \beta + \kappa + \beta^2q_2/\psi} \quad \text{and} \quad U_2 = -\frac{u_2 + \beta^2 q_2^2/2\psi}{r + 2\kappa} = \frac{(r + \beta + 2\kappa)q_2}{2(r + 2\kappa)}. \]

Now we verify requirements (ii) and (iii) of Theorem 1:

- **Bound on** \( q'(p) - U''(p) \): the sufficient condition (18) is too tight, as it was derived under basically no knowledge of regarding \( q' - U'' \). In linear-quadratic games, however, \( q'(p) - U''(p) \) is constant, and hence we can rely on Remark 4 in the Proof of Theorem 1. More precisely, since
\[ q'(p) - U''(p) = q_2 - 2U_2 = q_2 - \frac{(r + \beta + 2\kappa)q_2}{r + 2\kappa} = -\frac{\beta q_2}{r + 2\kappa} > 0, \]
we require that
\[ q'(p) - U''(p) = -\frac{\beta}{r + 2\kappa}q_2 \leq \frac{\psi(r + 2(\kappa + \beta))^2}{4\beta^2(r + 2\kappa)}. \]
But since \( -q_2 < \psi(r + \beta + 2\kappa)/2\beta^2 \), the previous condition will be satisfied if \( 2\psi(r + \beta + 2\kappa) < (r + 2(\beta + \kappa))^2 \), which is clearly true.

- **Transversality conditions**: from the proof of Theorem 1 it suffices to show that
\[ \limsup_{t \to \infty} E[e^{-rt}(U(\hat{p}_{t*}) + [q(\hat{p}_{t*}) - U'(\hat{p}_{t*})]\hat{\Delta}_t + \Gamma\hat{\Delta}_t^2] \geq 0, \quad (35) \]
with \( \Gamma = \frac{[q'(\hat{p}_{t*}) - U''(\hat{p}_{t*})](r + 2\kappa - 2\beta)}{r + 2(\beta + \kappa)} = \frac{[q_2 - 2U_2](r + 2\kappa - 2\beta)}{r + 2(\beta + \kappa)} \), for any feasible strategy \( (\hat{a}_t)_{t \geq 0} \). To this end, observe that:
- it is clear that both \( \lim_{t \to 0} e^{-rt}E_t[p_t] = \lim_{t \to 0} e^{-rt}E_t[p_t^2] = 0 \), as \( (p_t)_{t \geq 0} \) is mean reverting of a Brownian martingale;
- using that (i) \( \beta^2 q_2/\psi + \beta + \nu + r > 0, \nu = \kappa, 2\kappa \), that (ii) by feasibility,
\[
\mathbb{E}[\int_0^t e^{-rs} \hat{a}_s ds] < \infty, \text{ and that (iii)}
\]
\[
\hat{\Delta}_t = \Delta^0 e^{-(\beta+\kappa+\beta^2q^2)t} + \int_0^t e^{(\beta+\kappa+\beta^2q^2)(s-t)}[\beta \hat{a}_s - \beta^2(q_1 + q_2p_s)] ds,
\]

it is easy to show that \( \lim_{t \to 0} e^{-rt} \mathbb{E}_t[\hat{\Delta}_t] = \lim_{t \to 0} e^{-rt} \mathbb{E}_t[p_t \hat{\Delta}_t] = 0. \)

As a result, (35) reduces to \( \limsup_{t \geq 0} e^{-rt} \mathbb{E}[(q_2 - 2U_2 + \Gamma - u_2) \hat{\Delta}_t^2] \geq 0. \) If, \( q_2 - 2U_2 + \Gamma - u_2 > 0, \) this is trivially true. Suppose that this is not the case. Since (i) flow payoffs are bounded by above and (ii) \( \hat{a} \) delivers finite utility (by feasibility), we have that \( \mathbb{E} \left[ \int_0^\infty e^{-rt} \left. u(\chi(p_t + \hat{\Delta}_t)) \right| dt \right] < \infty. \) Hence, \( \limsup_{t \to \infty} e^{-rt} \mathbb{E}[u(\chi(p_t + \hat{\Delta}_t))] \geq 0. \)

Using that \( \lim_{t \to \infty} e^{-rt} \mathbb{E}[p_t] = \lim_{t \to \infty} e^{-rt} \mathbb{E}[\hat{\Delta}_t] = \lim_{t \to \infty} e^{-rt} \mathbb{E}[p_t^2] = \lim_{t \to \infty} e^{-rt} \mathbb{E}[p_t \hat{\Delta}] = 0, \) and that \( u_2 < 0, \) we conclude that

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

This concludes the proof.

\[\square\]

In order to prove the existence results in Propositions 6 and 7 we rely on the following result from De Coster and Habets (2006):

**Theorem 4.** (De Coster and Habets (2006), Theorem II.5.6) Consider the second order differential equation

\[
u'' = f(t, u, u') \tag{36}
\]

with \( f : \mathbb{R}^3 \to \mathbb{R} \) a continuous function. Let \( \alpha, \beta \) of class \( C^2(\mathbb{R}) \) 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 \( t \in \mathbb{R} \)

\[(C1) \ \alpha'' \geq f(t, \alpha, \alpha') \text{ and } \beta'' \leq f(t, \beta, \beta').\]

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

\[
\int_0^\infty \frac{sd\varphi(s)}{\varphi_I(s)} = \infty, \tag{37}
\]

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 (36) has at
least one solution \( u \in C^2(\mathbb{R}) \) such that \( \alpha \leq u \leq \beta \).

**Remark 5.** The proof of this theorem in fact delivers a stronger result for the case in which \( \alpha \) and \( \beta \) are bounded and \( \varphi_I \) is independently of \( I \). In this case, the authors prove the existence of \( u \in C^2 \) solution to (36) satisfying \( \alpha \leq u \leq \beta \) and also that \( u' \) is bounded. See the page 123 in De Coster and Habets (2006) for proof of the Theorem and the discussion that addresses this remark.

**Proof of Proposition 6:** Let \( h(p) := u(\chi(p)) \). The ratcheting equation can be written as

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

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

\[
\begin{align*}
&f(p, A, 0) \leq 0 \iff (r + \beta + \kappa)A - h'(p) \leq 0 \iff A \leq \frac{m}{r + \beta + \kappa}, \\
&f(p, B, 0) \geq 0 \iff (r + \beta + \kappa)B - h'(p) \geq 0 \iff B \geq \frac{M}{r + \beta + \kappa}.
\end{align*}
\]

Hence, we aim 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 (37), Theorem 4 ensures the existence of a solution to the ratcheting equation that is of class \( C^2(\mathbb{R}) \) 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 bounded derivative is also ensured.

Now we study the asymptotic behavior of any solution to the ratcheting equation that takes values in \( J \). We start showing the both limits exist, and then separate the analysis into the cases \( \kappa = 0 \) and \( \kappa > 0 \).
Lemma 2. 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 ratcheting equation in the sequence $p_n$ for large $n$ we obtain

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

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, i.e. $\lim_{t \to -\infty} q(p)$ must exist as well. \hfill \square

Now we show that the limits in (27) 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

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

as $q(p)$ is bounded

and

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

Thus, the ratcheting 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

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

But since $(r + \beta(0))q(p_n)$ converges to $L(r + \beta(0)) > h'_\infty$, the ratcheting 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 \( \frac{k_\infty}{r + \beta(0)} \). The case \( p \rightarrow -\infty \) is identical.

**Case \( \kappa > 0 \):** We show that (27) holds in a sequence of steps.

**Step 1:** \( \lim_{p \rightarrow \infty} q'(p) = \lim_{p \rightarrow -\infty} q'(p) = 0 \) and \( q' \) is bounded. We show that the first limit exists (for the other limit the argument is analogous). Notice that \( q' \) cannot diverge; because \( q \) is of class \( C^2 \), this would imply that it 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 ratcheting equation is identically equal to zero, but the right-hand side diverges when \( \kappa > 0 \), as \( p_n q'(p_n) \rightarrow \infty \). 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 bounded, as \( |q'| \) converges to zero asymptotically in either direction and it is continuous.

**Step 2:** \( \lim_{p \rightarrow \infty} pq'(p) = \lim_{p \rightarrow -\infty} pq'(p) = 0 \). Notice that \( \lim_{p \rightarrow \infty} pq'(p) \) either exists or takes value \( +\infty \). The latter cannot be true, as the ratcheting equation would imply that \( \lim_{p \rightarrow \infty} q''(p) = +\infty \), implying that \( q' \) diverges, a contradiction. Suppose that \( \lim_{p \rightarrow \infty} pq'(p) = L > 0 \). 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 \( \lim_{p \rightarrow \infty} pq'(p) = 0 \). Finally, the analysis for limit \( \lim_{p \rightarrow -\infty} pq'(p) = 0 \) is identical.

**Step 3:** \( \lim_{p \rightarrow \infty} q''(p) = \lim_{p \rightarrow -\infty} q''(p) = 0 \). Using Step 1 and Step 2, the ratcheting equation implies that \( \lim_{p \rightarrow \infty} q''(p) \) exists. But if this limit is different from zero, then \( q' \) diverges as \( p \rightarrow \infty \), as \( q'(p) = O(p) \), a contradiction. Hence, \( \lim_{p \rightarrow -\infty} q''(p) = 0 \). The analysis for the other limit is analogous.

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

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

from where we conclude. \( \square \)

**Proof of Proposition 7:** We first show that, given \( q \) a bounded solution to the ratch-
et ing equation, there exists a solution to (17) satisfying a quadratic growth condition; for this purpose we apply Theorem 4. Then we apply the Feynman-Kac probabilistic representation theorem to show that the unique solution to (17) 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 bounded derivative when \( q' \) is bounded. As in the previous proof, let \( h(p) := u(\chi(p)) \).

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

\[
\frac{2}{\sigma^2} - h(p) + g(g^{-1}(\beta q(p))) + \kappa v(p - \eta) + ru \leq \phi_{0,I} + \phi_{1,I}v := \varphi_I(v)
\]

where \( (u, v) \in \mathbb{R}^2 \) is such that \( |u| \leq A + Bp^2 \), and \( p \in I \). Observe that the right hand side satisfies the Nagumo condition (37).

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

\[
-h(p) + g(g^{-1}(\beta q(p))) - \kappa \alpha'(p)(p - \eta) - r \alpha(p) \leq C(1 + |p|) + G - 2B\kappa(p - \eta) - r(A + Bp^2)
\]

\[
:= \frac{\sigma^2}{2} f(p,-\alpha(p),-\alpha'(p))
\]

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 \( |h(p)| \leq C(1 + |p|) \) for all \( p \in \mathbb{R} \)). Consequently,

\[
C(1 + |p|) + G - 2B\kappa(p - \eta) - r(A + Bp^2) \leq -\frac{\sigma^2}{2}\alpha''(p) = -B\sigma^2
\]

\[
\Leftrightarrow H(p) := \left( C + G + B\sigma^2 + 2B\kappa\eta - rA \right) + \left( C|p| - 2B\kappa p - rBp^2 \right) \leq 0, \forall p \in \mathbb{R}.
\]

If \( \kappa > 0 \), (2) \( \leq 0 \) will be automatically satisfied if \( B \) satisfies that \( 2B\kappa > C \Leftrightarrow B > C/2\kappa \).

Now, (1) \( \leq 0 \) is guaranteed to hold when \( A \) satisfies \( rA \geq C + G + 2B\sigma^2/2 + 2B\kappa\eta \). Hence, \( H(\cdot) \) is non-positive if we assume that \( A \) and \( B \) satisfy the conditions just stated. If instead \( \kappa = 0 \), (2) will be violated for \( |p| \) small, but choosing \( B \) sufficiently large, and then \( A \) satisfying the same condition but with enough slackness ensures that \( H(p) \leq 0 \) for all \( p \in \mathbb{R} \).

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

\[
-h(p) + g(g^{-1}(\beta q(p_t))) + \kappa \nu'(p)(p - \eta) + r\nu(p) \geq -C(1 + |p|) - G + 2B\kappa(p - \eta) + r(A + Bp^2)
\]

\[
:= \frac{\sigma^2}{2} f(p,\nu(p),\nu'(p))
\]
So imposing, \( \frac{\sigma^2}{2} \nu''(p) = B \sigma^2 \leq -C (1 + |p|) - G + 2B \kappa (p - \eta) + r (A + B p^2) \) yields the exact same condition found for \( \alpha \). Consequently, if we choose \( A, B \) satisfying the conditions above, \( \alpha \) and \( \nu \) are lower and upper solutions, respectively. Thus, there exist a \( U \in C^2(\mathbb{R}) \) solution to (17) 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 \( \mathbb{E} [e^{-r t} U(p_t)] \to 0 \) as \( t \to 0 \). Thus, the probabilistic representation follows from the Feynman-Kac formula in infinite horizon (Pham (2009) Remark 3.5.6.).

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

\[
|U(p + h) - U(p)| \leq \mathbb{E} \left[ \int_0^\infty e^{-r t} \left( |h(p^h_t) - h(p_0)| + |g(g^{\prime -1}(\beta q(p^h_t))) - g(g^{\prime -1}(\beta q(p_0)))| \right) dt \right] \\
\leq \frac{(\|h\|_\infty + R) h}{r}, \text{ for some } R > 0,
\]

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

Finally, it is easy to see that if \( h'(p) \) is bounded, then \( h' \) satisfies a linear growth condition. Also, since \( q(\cdot) \) is bounded, \( G := \sup_{p \in \mathbb{R}} g(g^{\prime -1}(\beta q(p))) < \infty \). 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

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

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 \( |U(p_0)| \leq K (1 + |p_0|) \). \( \square \)

**Proof of Theorem 3:** Let \( h(p) := u(\chi(p)) \). Take any bounded solution \( q \) to the ratcheting equation.

**Step 1: Conditions (i) and (ii) in Theorem 1 hold.** Observe that from Proposition 7, \( U(\cdot) \) has a linear growth condition (hence, quadratic growth condition), and \( U' \) is bounded, so the linear growth condition trivially holds. Also, since \( q' \) bounded, \( q \) is automatically Lipschitz. Thus, (i) in Theorem 1 holds.
As for condition (ii), we first show that \( \lim_{t \to \infty} e^{-rt} \mathbb{E}[\hat{\Delta}_t] = 0 \). Observe that

\[
\hat{\Delta}_t = e^{-(\beta + \kappa)t} \hat{\Delta}_0 + \beta \int_0^t e^{-(\beta + \kappa)(t-s)} [\dot{a}_s - a^*(p_s + \hat{\Delta}_s)] ds
\]

\[
\Rightarrow |\hat{\Delta}_t| \leq \beta \int_0^t e^{-(\beta + \kappa)(t-s)} |\dot{a}_s| ds + \beta \int_0^t e^{-(\beta + \kappa)(t-s)} |a^*(p_s + \hat{\Delta}_s)| ds.
\]

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

\[
\int_0^t e^{-(\beta + \kappa)(t-s)} |\dot{a}_s| ds \leq \left( \int_0^t e^{-2(\beta + \kappa)(t-s)} ds \right)^{1/2} \left( \int_0^t |\dot{a}_s|^2 ds \right)^{1/2}
\]

\[
\Rightarrow |e^{-rt} \mathbb{E}[\hat{\Delta}_t]| \leq \left( \int_0^t e^{-2(\beta + \kappa)(t-s)} ds \right)^{1/2} \left( \int_0^t |\dot{a}_s|^2 ds \right)^{1/2}
\]

\[
+ \left( e^{-rt} \int_0^t e^{-2(\beta + \kappa)(t-s)} ds \right)^{1/2} \left( e^{-rt} \int_0^t \mathbb{E}[|\dot{a}_s|^2] ds \right)^{1/2} < \infty
\]

for all \( t \geq 0 \), where the right-most inequality comes from the fact that \( \dot{a} \in \mathcal{L}^2 \). But

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

where the last inequality comes from strong convexity: otherwise, 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 \( |U(p_t + \hat{\Delta}_t)| \leq C_1 (1 + |p_t| + |\dot{\Delta}_t|) \), \( q(p_t + \hat{\Delta}_t) \hat{\Delta}_t |\leq C_2 |\Delta_t| \), and \( |U'(p_t + \hat{\Delta}_t) \hat{\Delta}_t| \leq C_3 |\Delta_t| \) for some constants \( C_1, C_2 \) and \( C_3 \) all larger than zero.

**Step 2: Condition (iii) in Theorem 1 holds.** Recall that the long-run player’s payoff (28), i.e.

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

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 (17) 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
\( p \mapsto U'(p): \)
\[
U''(p) = \frac{2}{\sigma^2} \left[ -h'(p) + \beta^2 \frac{q(p)q'(p)}{g''(g^{-1}(\beta q(p)))} + (r + \kappa)U'(p) + \kappa(p - \eta)U''(p) \right], \quad p \in \mathbb{R}. \tag{40}
\]

Moreover, from the ratching equation (16)
\[
-h'(p) + \beta^2 \frac{q(p)q'(p)}{g''(g^{-1}(\beta q(p)))} = -(r + \beta + \kappa)q(p) - \kappa(p - \eta)q'(p) + \frac{1}{2} \sigma^2 q''(p).
\]

Replacing this into (40) yields that \( U' - q \) satisfies the ODE
\[
(U''' - q'')(p) = \frac{2}{\sigma^2} \left[ -\beta q(p) + (r + \kappa)(U' - q)(p) + \kappa(p - \eta)(U'' - q')(p) \right], \quad p \in \mathbb{R}. \tag{41}
\]

But since \( U' - q \) is bounded, it automatically satisfies both a quadratic growth condition and the transversality condition \( \lim_{t \to \infty} \mathbb{E}[e^{-(r + \kappa)t}(U'(p_t) - q(p_t))] = 0 \). The Feynman-Kac formula (Pham 2009, Remark 3.5.6.) tells us that the solution to the previous ODE is unique—hence, given by \( (U' - q)(\cdot) \)—and that it has the probabilistic representation
\[
U'(p) - q(p) = \mathbb{E} \left[ \int_0^\infty e^{-(r + \kappa)t} \beta q(p_t) dt \bigg| p_0 = p \right], \quad p \in \mathbb{R}. \tag{42}
\]

Now, notice that (41) implies that \( U'' - q' \) is twice continuously differentiable. Also, using that \( q' \) is bounded and that \( (p_t)_{t \geq 0} \) is mean-reverting we can replicate the argument that shows that \( U' \) is bounded (proof of Proposition 7) to show that \( U''' - q'' \) is bounded as well; hence, it trivially satisfies a quadratic growth condition and the transversality condition \( \lim_{t \to \infty} \mathbb{E}[e^{-(r + 2\kappa)t}(U''(p_t) - q'(p_t))] = 0 \). Furthermore, differentiating (41) we obtain that \( U''' - q''' \) satisfies the ODE
\[
(U''''' - q'''')(p) = \frac{2}{\sigma^2} \left[ -\beta q'(p) + (r + 2\kappa)(U' - q)(p) + \kappa(p - \eta)(U''' - q'')(p) \right], \quad p \in \mathbb{R}. \tag{43}
\]

The Feynman-Kac formula allows us to conclude that
\[
U''(p) - q'(p) = \mathbb{E} \left[ \int_0^\infty e^{-(r + 2\kappa)t} \beta q'(p_t) dt \bigg| p_0 = p \right], \quad p \in \mathbb{R}. \tag{44}
\]

When \( \kappa = 0 \), the right-hand side of the previous expression—or, equivalently, the solution to (43)—admits an analytic representation in terms of \( q \). In fact, it is easy to see that
\[
\frac{\beta}{\sigma^2 \sqrt{\nu}} \left[ \int_{-\infty}^p e^{-\sqrt{\nu}(p - y)} q'(y) dy + \int_p^\infty e^{-\sqrt{\nu}(y - p)} q'(y) dy \right]. \tag{45}
\]

48
where \( \nu := 2r/\sigma^2 \), is a solution to (43) when \( \kappa = 0 \). Because it is of class \( C^2 \) and bounded (hence, its satisfies a quadratic growth condition and the last transversality condition), Feynman-Kac tell us that its must coincide with \( U'' - q' \). Integrating by parts yields
\[
U''(p) - q'(p) = \frac{\beta}{\sigma^2} \left[ -\int_{-\infty}^{p} e^{-\sqrt{\nu}(y-p)} q(y) dy + \int_{p}^{\infty} e^{-\sqrt{\nu}(y-p)} q(y) dy \right].
\] (46)

Recalling that \( q(\cdot) \in \left[ \frac{m}{r+\beta}, \frac{M}{r+\beta} \right] \), that \( \nu = 2r/\sigma^2 \) and that \( \sigma = \beta\sigma_\xi \), it is easy to see that
\[
|U''(p) - q'(p)| \leq \frac{M - m}{(r + \beta)^3} p \in \mathbb{R}.
\]

But since \( \beta = \sigma_\theta / \sigma_\xi \) when \( \kappa = 0 \), we have that (18) in Theorem 1 will hold if
\[
\frac{M - m}{(r + \beta)^3} \psi(r + 4\beta + 2\kappa) \leq \frac{\sqrt{2r\sigma^2_\xi(r\sigma_\xi + 4\sigma_\theta)(r\sigma_\xi + \sigma_\theta)}}{4\sigma^2_\theta}.
\]

Since condition (31) is tighter than the one just derived, condition (iii) in Theorem 1 holds. This concludes the proof.

Remark 6. When \( \kappa > 0 \), the ODE (43) also has a solution of the form \( p \mapsto \int_{\mathbb{R}} g(p, y)q'(y)dy \), but the kernel \( g(p, y) \) admits no closed-form solution. This is because all solutions to (43) are constructed using the corresponding ones for the homogenous problem (i.e. \( q' \equiv 0 \)), which take the form of confluent hypergeometric functions; see Abramowitz and Stegun 1964. □

Proofs of Propositions 3, 4 and 5:

Proof of Proposition 3: Throughout this proof, \( h(p) := u(\chi(p)) \), \( p \in \mathbb{R} \) and \( q \) is a bounded solution to the ratcheting equation as in Proposition 6. The proof consists of 6 steps.

Step 1: \( q'(0) \geq 0 \). Consider the ode
\[
\left[ r + \beta - \frac{\beta^2}{\psi} \tilde{q}'(p) \right] \tilde{q}'(p) = h'(p) + \frac{1}{2} \sigma^2 \tilde{q}''(p), \quad p > 0.
\]

Notice that a solution \( \tilde{q} \) to it in fact satisfies \( \tilde{q}(p) = q(-p) \) for \( p \geq 0 \). In particular, \( \tilde{q}'(0) = -q'(0) \) and \( \tilde{q}''(0) = q''(0) \). In what follows, \( \tilde{q} \) denotes a solution to this ODE, whereas \( q \) denote a solution to the original ratcheting equation over \( \mathbb{R}_+ \).

Suppose that \( q'(0) < 0 \). Then, \( \tilde{q} > q \) locally to the right of zero. Notice that \( \tilde{q} - q \) is bounded and that \( \tilde{q}(p) - q(p) \to 0 \) as \( p \to \infty \). Hence, there exists a \( \hat{p} \) at which \( \tilde{q} - q \) is
maximized. In particular, \( \tilde{q}(\hat{p}) = q'(\hat{p}) \) and \( \tilde{q}'(\hat{p}) - q''(\hat{p}) \leq 0 \). The latter is equivalent to

\[
\left[ r + \beta - \frac{\beta^2}{\psi} q'(\hat{p}) \right] \tilde{q}(\hat{p}) \leq \left[ r + \beta + \frac{\beta^2}{\psi} q'(\hat{p}) \right] q(\hat{p}).
\]

Observe first that since \( \tilde{q}(\hat{p}) > q(\hat{p}) \) and \( \tilde{q}'(\hat{p}) = q'(\hat{p}) \) it must be the case that \( \tilde{q}'(\hat{p}) = q'(\hat{p}) > 0 \) for the previous inequality to hold. By continuity, there exist \( \tilde{p} \in (0, \hat{p}) \) at which \( q \) is minimized, with \( q \) being locally increasing to the right of \( \tilde{p} \). Since \( q''(\tilde{p}) \geq 0 \) and \( h' \) is non-increasing in \( \mathbb{R}_+ \), we have that

\[
\frac{1}{2} \sigma^2 q''(p) = \left( \frac{r + \beta}{\psi} q(p) - h(p) \right) + \frac{\beta^2}{\psi} q'(p)q(p) > 0
\]

slightly to the right of \( \tilde{p} \), which implies that \( q' \) is strictly increasing in the same region. Thus, the previous inequality holds over \( \mathbb{R}_+ \). Because \( q \) is of class \( C^2 \), \( q' \) grows indefinitely, which means that \( q \) grows indefinitely as well, a contradiction. We conclude that \( q'(0) \geq 0 \).

**Step 2:** \( q'(0) > 0 \). Suppose that \( q'(0) = 0 \). Then,

\[
q(0) = \frac{h'(0) + \frac{1}{2}(\sigma)q''(p)}{r + \beta}.
\]

We conclude that \( q \) can’t be strictly convex at zero, as this implies that \( q > h'(0)/(r + \beta) \), a contradiction. Hence, \( q''(0) \leq 0 \). If \( q''(0) < 0 \), we have that

\[
q'''(0) = \frac{2}{\sigma^2} q(0)q''(0) < 0 \quad \text{and} \quad q'''(0) = -\frac{2}{\sigma^2} \tilde{q}(0)\tilde{q}''(0) > 0,
\]

where we used that \( q \) is \( C^3 \) at 0, that \( h''(0) = q'(0) = 0 \), and that \( q(0) > 0 \) (otherwise, 0 is a minimum, a contradiction with \( q'' < 0 \)). From the previous inequalities we conclude that \( \tilde{q} > q \) in a neighborhood of zero, which boils down to the analysis performed in (i), thus yielding a contradiction. Hence, \( q''(0) = 0 \). In particular, \( q'''(0) = 0 \).

Notice that since \( q'(0) = q''(0) = 0 \), it must be the case that \( q(0) = \frac{h'(0)}{r + \beta} \), i.e., \( q \) achieves its maximum value. Because \( h' \) is twice continuously differentiable at zero, and \( h'''(0) < 0 \), we have that \( q \) is of class \( C^4 \) at zero, and hence

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

But this implies that \( q \) must grow locally to the right of zero, a contradiction. We conclude that \( q'(0) \neq 0 \) and thus that \( q'(0) > 0 \). In particular, \( q(0) > 0 \) (otherwise, zero is a local
minimum).

**Step 3.** $q''(0) < 0$. It is clear that $q''(0) \leq 0$. Otherwise, $q'$ is strictly increasing at zero and, since $h'$ decays in $\mathbb{R}_+$, $q'' > 0$ everywhere, which means that (applying the same used before) $q$ grows without bound. Suppose that $q''(0) = 0$. Then,

$$\frac{\sigma^2}{2} q''(0) = (r + \beta)q'(0) + \frac{\beta^2}{\psi} q'(0)q(0) + \frac{\beta^2}{\psi} q(0)q''(0) - h''(0) > 0$$

Then, $q'' > 0$ slightly to the right of zero, which means that $q'$ keeps growing locally. Because $h'$ decreases over $\mathbb{R}_+$, $q''$, $q'$ and $q$ grow indefinitely over the same interval, which is a contradiction. We conclude that $q''(0) < 0$.

**Step 4.** $q$ is non-decreasing in $\mathbb{R}_{-}$. Using $\tilde{q}$, suppose that $\tilde{q}' > 0$ at some point in $\mathbb{R}_+$. Since $\tilde{q}'(0) < 0$, the must exist $\hat{p}$ such that $\tilde{q}'(\hat{p}) = 0$ and $\tilde{q}$ grows locally to the right of $\hat{p}$. Using that $\tilde{q}''(\hat{p}) \geq 0$ and that $h'$ is decreasing, we conclude that $\tilde{q}'' > 0$ to the right of $\hat{p}$, and hence $\tilde{q}'$ must be growing. Consequently, there must be a point $\bar{p}$ such that $\beta + r = \beta^2 \tilde{q}'(\bar{p})/\psi$. We conclude that

$$0 = h'(\bar{p}) + \frac{\sigma^2}{2} \tilde{q}''(\bar{p}) > 0,$$

which is a contradiction. Hence, $\tilde{q}' = -q' \leq 0$ over $(-\infty, 0)$. Notice that if there is a flat portion over a region in which $h'$ is strictly increasing, then $q' = q'' = 0$. Hence the ratcheting equation is violated.

**Step 5:** Uniqueness of local maximum. Existence of a global maximum is ensured by the fact that $q(p) \to 0$ as $p \to \pm$ and that $q$ is bounded. By the previous point, such maximum is to the right of zero. If there is another local maximum, then there must be a local minimum in between. Convexity at that point and $h'$ decreasing allows us to conclude that $q$ must grow indefinitely to the right of that minimum, a contradiction. Hence, $q$ must be monotone to the right of that global maximum. We conclude that it must be strictly decreasing, as (i) if $q$ increases, then the definition of global maximum is violated, and (ii) if $q$ is flat, the ratcheting equation is violated.

**Step 6.** $q(p) > q(-p)$, $p > 0$. Suppose that there is $p > 0$ at which $\tilde{q}(p) > q(p)$. Using the same arguments from (i), we conclude that there is a point $\hat{p}$ at which $\tilde{q}'(\hat{p}) = q'(\hat{p})$, $\tilde{q}(\hat{p}) > q(\hat{p})$ and

$$\left[ r + \beta - \frac{\beta^2}{\psi} \tilde{q}'(\hat{p}) \right] \tilde{q}(\hat{p}) \leq \left[ r + \beta + \frac{\beta^2}{\psi} q'(\hat{p}) \right] q(\hat{p}).$$

This implies that $\tilde{q}'(\hat{p}) > 0$, but this contradicts the fact that $\tilde{q}$ is decreasing in $\mathbb{R}_+$. Hence, $\tilde{q} \leq q$. If there is $p > 0$ such that $q(p) = \tilde{q}(p)$, then, $q - \tilde{q}$ achieves a local maximum, and
hence $q'(p) = q'(p)$ and $q''(p) - q''(p) \geq 0$. Because $\tilde{q}$ is strictly decreasing,

$$\frac{h'(p) + \frac{1}{2} \sigma^2 q''(p)}{r + \beta + \frac{\beta^2}{\psi} q'(p)} > \frac{h'(p) + \frac{1}{2} \sigma^2 \tilde{q}''(p)}{r + \beta - \frac{\beta^2}{\psi} \tilde{q}'(p)},$$

when $r + \beta + \beta^2 q'(p)/\psi > 0$, yielding a contradiction. If instead $r + \beta + \beta^2 q'(p)/\psi \leq 0$, using the ratcheting equation we conclude that $h'(p) + \frac{1}{2} \sigma^2 q''(p) \leq 0$. Using that $\tilde{q}'(p) < 0$ and that $q''(p) - \tilde{q}''(p) \geq 0$

$$\tilde{q}(p) = h'(p) + \frac{1}{2} \sigma^2 \tilde{q}''(p) r + \beta - \frac{\beta^2}{\psi} \tilde{q}'(p) < 0,$$

which is a contradiction. Hence $q > \tilde{q}$ everywhere. \hfill \square

\textbf{Proof of Proposition 4:} Consider a linear-quadratic game with parameters $u_0 = u_1 = 0$, $u_2 > 0$ and $\kappa = 0$. Theorem 2 tells us that en equilibrium exists provided the curvature condition (23) holds, in which case the linear Markov equilibrium is given by

$$a^*(p^*) = \frac{\beta}{\psi} q_2 p^*, \text{ where } q_2 = \frac{\psi}{2\beta^2} \left[-(r + \beta) + \sqrt{(r + \beta)^2 - 8u_2\beta^2\psi}\right] < 0.$$

Consequently, suppose that the market conjectures $a^* = \beta \alpha^*_p / \psi$, $\alpha < 0$, as the equilibrium policy. Define $u_2 := \beta^2 \alpha^2 / \psi^2$. We must show that we can find a fixed point $q_2 = \alpha$, and that under such fixed point the curvature condition holds. Simple algebra shows that after plugging $u_2 = \beta^2 \alpha^2 / \psi^2$ into the expression for $q_2$, and imposing $q_2 = \alpha$, we obtain

$$\alpha \left[r + \beta + \alpha \frac{\beta^2}{\psi} \left(1 + \frac{2}{\psi}\right) \right] = 0 \Rightarrow \alpha = -\frac{\psi^2 (r + \beta)}{\beta^2 (2 + \psi)}.$$

Consequently, the curvature condition will hold if and only if

$$u_2 = \frac{\beta^2}{\psi^2} \alpha^2 = \frac{\beta^2}{\psi^2} \left(\frac{\psi^2 (r + \beta)}{\beta^2 (2 + \psi)}\right)^2 \leq \frac{\psi (r + \beta)^2}{8\beta^2} \quad \Rightarrow \frac{\psi}{(2 + \psi)^2} \leq \frac{1}{8} \quad \Rightarrow 0 \leq (2 - \psi)^2,$$

which is always true. This concludes the proof. \hfill \square

\textbf{Proof of Proposition 5:} It is easy to verify that $q(p) = q_1 + q_2 p$, $q_1, q_2 \in \mathbb{R}$, is a solution
to (21)-(22) if and only if

\[ 9(r + 2\kappa)q_2 + 4\beta^2 q_2^2 + 1 = 0 \]
\[ [9(r + \beta + \kappa) + 2\beta(1 - \beta q_2)]q_1 - (1 - \beta q_2)c - 9\kappa \eta q_2 = 0 \]

The first quadratic admits two solutions, both of them strictly negative. In either case

\[ q_1 = \frac{(1 - \beta q_2)c + 9\kappa \eta q_2}{9(r + \beta + \kappa) + 2\beta(1 - \beta q_2)} \leq 0 \iff (1 - \beta q_2)c \leq -9\kappa \eta q_2. \]
\[ \iff \frac{\eta}{c} \geq \frac{1 - \beta q_2}{9\kappa q_2} > 0 \] (47)

The latter is sufficient for \( a^*(c) = -\beta(q_1 + q_2c)/3 > 0. \) □