Reputation and the Flow of Information in Repeated Games*

Eduardo Faingold  
*Insper Institute of Education and Research  
São Paulo, Brazil  
efaingold@insper.edu.br

January 2020

Abstract

Equilibrium payoff bounds from reputation effects are derived for repeated games with imperfect public monitoring in which a long-run player interacts frequently with a population of short-run players and the monitoring technology scales with the length of the period of interaction. The bounds depend on the monitoring technology through the flow of information, a measure of signal informativeness per unit of time based on relative entropy. Examples are shown where, under complete information, the set of equilibrium payoffs of the long-run player converges, as the period length tends to zero, to the set of static equilibrium payoffs, whereas when the game is perturbed by a small ex ante probability on commitment types, reputation effects remain powerful in the high-frequency limit.

Keywords: Reputation, commitment, imperfect monitoring, high-frequency repeated games.

JEL Classification: C70, C72.

---

*I thank the co-editor and three anonymous referees for exceptionally helpful comments and suggestions. I am also grateful to Mehmet Ekmecki, Drew Fudenberg, Olivier Gossner and George J. Mailath for stimulating discussions and helpful comments, and the National Science Foundation (grant SES-1227593) for financial support. This paper subsumes and improves upon the earlier analysis of reputations under frequent interactions in Faingold (2006).
1 Introduction

This paper examines reputation effects (in the sense of Kreps and Wilson (1982), Milgrom and Roberts (1982)) in repeated games in which a long-run player interacts frequently with a population of short-run players. The main result provides upper and lower bounds (akin to those of Fudenberg and Levine (1989, 1992)) on the set of Nash equilibrium payoffs of the long-run player in the limit as the discount rate and the period length tend to zero (in any order).

In the repeated games literature, it is common to interpret comparative statics with the discount factor as an exercise of varying the frequency of interaction. Thus the limit when the discount factor tends to one is interpreted as the limit when the period length tends to zero while the monitoring technology remains fixed independent of the period length. However, when players observe noisy signals about each other’s actions, as in many applications of interest, it may be more natural to assume that the informativeness of the signals is constant per unit of real time. As this requires the signal distributions to scale with the period length (Abreu, Milgrom, and Pearce, 1991), the ensuing comparative statics is different from the traditional one: as the period length shrinks to zero, the discount factor tends to one, but at the same time the informativeness of the signals over the period of interaction deteriorates. It is now well known that allowing the monitoring technology to scale with the period length in this way can have a dramatic effect on the equilibrium outcomes of repeated games: in some cases, the set of (perfect public) equilibrium payoffs of the repeated game converges, as the period length tends to zero, to the (convex hull of the) set of static Nash equilibrium payoffs (Abreu, Milgrom, and Pearce (1991), Faingold and Sannikov (2011), Fudenberg and Levine (2007, 2009), Sannikov and Skrzypacz (2007)).

How about when the repeated game is perturbed by a small ex ante probability that the long-run player may be a commitment type? Do reputation effects survive in the limit as the period length shrinks to zero and the informativeness of the signals deteriorates? The main result of this paper, Theorem 2, answers this question in the affirmative by providing upper and lower bounds on the set of Nash equilibrium payoffs of the long-run player in the limit as the period length and the discount rate tend to zero (in any order). Examples are shown in which the bounds are tight and equal the Stackelberg payoff, whereas without uncertainty over types the set of equilibrium payoffs collapses to the convex hull of stage-game Nash equilibrium payoffs.

The intuition for the dramatic difference between the complete and incomplete information results is as follows. In the perturbed game, due to the persistence of the short-run players’ posterior belief about the long-run player’s type, the relevant time frame for reputation effects is the “long run”—a reputation is not won in a day—, and in the long run the many (poorly informative) one-period signals can be aggregated to become informative. By contrast, in the unperturbed complete information game, there is no learning about types and hence no state variable linking the payoffs across periods, so that non-myopic incentives can only be created by using each period’s signal to adjust continuation pay to discourage deviations. Thus, unlike in the perturbed game, the relevant time period for incentives in the complete information game is the period of interaction itself—over which the signals are becoming noisier and noisier—, and hence the possibility of collapse to static Nash equilibrium.\footnote{Bohren (2019) also explores persistence in an alternative channel by which intertemporal incentives can be sustained}
In spite of this intuition the result is more subtle than it may appear at first glance: it cannot be proved by invoking the classic payoff bounds of Fudenberg and Levine (1992), hereafter FL. To see why, let us first review the argument of FL. Fix an arbitrarily small positive \( \varepsilon \) and a Nash equilibrium of the incomplete information game and consider a deviation in which the normal type mimics the behavior of the Stackelberg type. FL show that, under this deviation, for a set of infinite histories that has probability at least \( 1 - \varepsilon \), there is an upper bound \( K \) on the number of periods in which the short-run players are not playing a best-reply to the Stackelberg action. FL show that this bound \( K \), which depends on the monitoring technology and on \( \varepsilon \), can be chosen independent of the particular equilibrium under consideration. Due to this uniformity, in the limit as the discount factor tends to one and \( \varepsilon \) tends to zero in this order, the expected average discounted payoff of the normal type (under the deviation) must be no less than the Stackelberg payoff. It follows that the Stackelberg payoff must be a lower bound on the patient limit of the set of Nash equilibrium payoffs of the normal type. Similar ideas apply to the upper bound.

However, FL’s argument above assumes a fixed monitoring technology, and thus a fixed period of interaction \( \Delta \). The reason why taking the limit of the FL bounds as \( \Delta \to 0 \) may not lead to informative high-frequency bounds is related to how the above described integer \( K \) changes with the period length \( \Delta \) (through the monitoring technology). Section 4 contains an in-depth analysis and a detailed discussion of this issue, with a running example of a game with Poisson signals for which \( K \) fails to scale homogeneously with \( 1/\Delta \). In that example the monitoring structure has a well defined continuous time limit, and so we might expect the number of non-best-reply periods to be roughly proportional to the reciprocal of \( \Delta \) (so the real time of non-best-reply behavior is approximately independent of the frequency of interactions). But this turns out not to be the case for \( K \), which is only an upper bound on the number of non-best reply periods.

To further elaborate, recall that FL obtain the bound \( K \) by counting the number of surprises, that is, periods in which the distance between the short-run players’ equilibrium forecast of the public signal, and what their forecast would be if they knew they were facing the Stackelberg action, is greater than a threshold. Such “surprise threshold” is optimally chosen so the short-run players find it optimal to best-reply to the Stackelberg action whenever the distance between the above described forecasts falls below that threshold. This ensures that the number of surprises is indeed an upper bound on the number of non-best replies. But note that the surprise threshold, and hence FL’s upper bound \( K \) on the number of non-best replies, depends on the particular metric used to measure distance between forecasts. While the choice of metric is unimportant for the analysis of the patient limit of reputation bounds in the canonical discrete-time model, in the case where the informativeness of the signals scales with \( \Delta \) the choice of metric affects the rate with which \( K \) changes with \( \Delta \). The fact that the bound \( K \) is not of the order of \( 1/\Delta \), as required for the frequent interaction limit of the FL bounds to be informative, turns out to be a consequence of FL’s use of the total variation distance.

To circumvent this problem, Theorem 2 relies instead on Gossner (2011), which uses informa-

\footnote{in the high frequency limit which does not rely on uncertainty over types. In a different vein, Rahman (2014) shows, in a high-frequency oligopoly setting with Brownian signals à la Sannikov and Skrzypacz (2007), how private communication between firms (with or without a mediator) can overcome the impossibility of dynamic incentive provision due to noisy per-period information.}

\footnote{The Stackelberg type is the type committed to playing the Stackelberg action after all histories.}
tion theoretic methods to derive reputation payoff bounds that are tighter than those of FL for fixed discount factor and monitoring technology. The bounds provided in Theorem 2, which are the high-frequency limits of the Gossner bounds, depend on the flow of information, a measure of signal informativeness per unit of time based on relative entropy (Cover and Thomas, 2006) rather than on total variation distance. To illustrate the result, explicit calculations of the high-frequency bounds are provided for a few examples, including one in which the set of equilibrium payoffs of the normal type converges to the Stackelberg payoff in the patient high-frequency limit, while applying the same limit to the FL bounds yields uninformative results. The example is not pathological: a dramatic gap between the high-frequency limits of the Gossner and the FL bounds arises in any (nondegenerate) game in which signals arrive according to a Poisson process.

Reputation effects under frequent interactions are also examined in Faingold and Sannikov (2011), which studies a model in which the period length is exactly zero, that is, the interaction takes place directly in continuous time. Unlike the present paper, which focuses on bounds on equilibrium payoffs for patient long-run players under general assumptions on the monitoring structure and on the type space, Faingold and Sannikov (2011) provides characterizations of equilibrium behavior for fixed discount rates, by restricting attention to games with Brownian signals and assuming that there is a single commitment type. Another related paper is Faingold (2013), which extends the reputation bounds of Fudenberg and Levine (1992) to a class of continuous-time games with imperfect public monitoring and signals driven by a Lévy process, which includes games with Brownian and Poisson signals as special cases. A pair of examples in Section 3 demonstrates that the reputation result for high-frequency games of the present paper is not an implication of the reputation result for continuous-time games.

The rest of the paper is organized as follows. Section 2 sets up the baseline discrete-time model and reviews the reputation result of Fudenberg and Levine (1992). Section 3 embeds the baseline model in continuous time, states the main result concerning reputation bounds for high-frequency games, and illustrates the result in four examples. Section 4 provides an in-depth discussion of the gap between the payoff bounds based on the entropic flow of information and the high-frequency limit of the FL bounds. Section 5 concludes by stating, and briefly discussing, an open question for future research.

2 Baseline Model

2.1 Discrete-time repeated game. We recall the canonical reputation model of Fudenberg and Levine (1992), in which a long-run player faces a sequence of short-run players in an infinitely repeated game with imperfect monitoring in discrete time. At the beginning of period $n = 0, 1, 2, \ldots$,
the long-run player and the current short-run player simultaneously, and privately, choose actions $a^n_1 \in A_1$ and $a^n_2 \in A_2$ respectively, where $A_1$ and $A_2$ are finite action spaces. A publicly observable signal $y^n$ is then drawn from a measurable space $Y$ according to a probability distribution $q(\cdot | a^n_1, a^n_2)$, where $q : A \to \Delta(Y)$ and $A := A_1 \times A_2$. At the end of the period, the current short-run player departs from the game and is replaced by a new short-run player at the beginning of the following period. The new short-run player knows the entire past history of public signals.

The long-run player is privately informed of his type $\omega \in \Omega$, where $\Omega$ is a countable type space. The short-run players are uncertain as to which type of long-run player they face and share a common prior $\mu \in \Delta(\Omega)$, which satisfies $\mu(\omega) > 0$ for all $\omega \in \Omega$. The type space $\Omega$ contains strategic (or normal) types and behavioral (or commitment) types. The set of strategic types is denoted $\Omega^s \subset \Omega$ and the overall payoff of a strategic type $\omega$ is

$$
\sum_{n=0}^{\infty} (1 - \delta)^n g_1(a^n_1, a^n_2, \omega),
$$

where $g_1 : A_1 \times A_2 \times \Omega^s \to \mathbb{R}$ is the stage-game payoff function of the long-run player and $\delta \in (0, 1)$ is his discount factor. A behavioral type is a mechanistic type of long-run player who is committed to playing some mixed action $a_1 \in \Delta(A_1)$ in every period, irrespective of history. The set of behavioral types, denoted $\Omega^b := \Omega \setminus \Omega^s$, is thus naturally identified with a countable subset of $\Delta(A_1)$. Finally, the payoff function of the short-run players is common knowledge and denoted $g_2 : A_1 \times A_2 \to \mathbb{R}$.

2.2. Strategies and equilibrium. The space of $n$-period histories of the long-run player is denoted $H^n_1 := \Omega \times (A_1 \times Y)^n$ and is equipped with the product $\sigma$-field, where $H^n_2 := \Omega$. Likewise, the space of $n$-period histories of the short-run players is denoted $H^n := Y^n$ and is endowed with the product $\sigma$-field, where $H^n_2$ is the one-point set. A (behavior) strategy of player $i$ ($i = 1, 2$) is a sequence of measurable maps $\sigma_i^n : H^n_i \to \Delta(A_i)$, $n \geq 0$, such that $\sigma_1(\omega, \cdot) = \omega$ for every behavioral type $\omega \in \Omega^b$. The set of strategies of player $i$ is denoted $\Sigma_i$ and the set of strategy profiles is denoted $\Sigma := \Sigma_1 \times \Sigma_2$. Given the prior $\mu$, each strategy profile $\sigma \in \Sigma$ induces a unique probability measure $P_{\mu, \sigma}$ over the space of plays $H^\infty := \Omega \times (A_1 \times A_2 \times Y)^\infty$ endowed with the product $\sigma$-field. For each $\omega \in \Omega$, the symbol $P_{\omega, \sigma}$ designates the induced probability measure over plays conditional on $\omega$. Expectations with respect to $P_{\mu, \sigma}$ and $P_{\omega, \sigma}$ are denoted $E_{\mu, \sigma}$ and $E_{\omega, \sigma}$, respectively. A profile $\sigma \in \Sigma$ is a (Bayes-)Nash equilibrium of the repeated game $\Gamma_\delta = ((\Omega, \mu), (A_i, g_i)_{i=1,2}, (Y, q), \delta)$ if for every $\sigma' \in \Sigma_1$,

$$
E_{\mu, \sigma} \left[ \sum_{n=0}^{\infty} (1 - \delta)^n g_1(a^n_1, a^n_2) \right] \geq E_{\mu, (\sigma' \mid A, \sigma_2)} \left[ \sum_{n=0}^{\infty} (1 - \delta)^n g_1(a^n_1, a^n_2) \right],
$$

and for every $n \geq 0$ and $\sigma'_2 \in \Sigma_2$ such that $\sigma'_{2m} = \sigma_{2m}$ for every $m \neq n,

$$
E_{\mu, \sigma} [g_2(a^n_1, a^n_2)] \geq E_{\mu, (\sigma_1, \sigma'_2)} [g_2(a^n_1, a^n_2)].
$$

For each $\omega \in \Omega^s$, let

$$
N_1(\omega, \delta) := \left\{ E_{\omega, \sigma} \left[ \sum_{n=0}^{\infty} (1 - \delta)^n g_1(a^n_1, a^n_2) \right] : \sigma \text{ is a Nash equilibrium of } \Gamma_\delta \right\}.
$$

\footnote{The payoff functions of behavioral types are not explicitly modeled.}
the set of Nash equilibrium interim payoffs of type $\omega$. Finally, define

$$N_1(\omega, \delta) := \inf N_1(\omega, \delta) \quad \text{and} \quad \bar{N}_1(\omega, \delta) := \sup N_1(\omega, \delta).$$

2.3. Reputation effects. We review the classical result of Fudenberg and Levine (1992), which provides upper and lower bounds on the set of Nash equilibrium payoffs of patient strategic types. Recall that a mixed action $\alpha_2 \in \Delta(A_2)$ is a self-confirmed best-reply to $\alpha_1 \in \Delta(A_1)$, written $\alpha_2 \in B_2(\alpha_1)$, if $\alpha_2$ is not weakly dominated and $\alpha_2 \in \arg\max_{\alpha_2} g_2(\alpha_1', \alpha_2')$ for some $\alpha_1' \in \Delta(A_1)$ with $q(\cdot|\alpha_1', \alpha_2') = q(\cdot|\alpha_1, \alpha_2)$. For each $\omega \in \Omega^b$ define

$$g_1(\omega) := \sup_{\alpha_1 \in \Delta(A_1)} \inf_{\alpha_2 \in B_2(\alpha_1)} g_1(\alpha_1, \alpha_2, \omega)$$

and

$$\bar{g}_1(\omega) := \sup_{\alpha_1 \in \Delta(A_1)} \sup_{\alpha_2 \in B_2(\alpha_1)} g_1(\alpha_1, \alpha_2, \omega).$$

The latter is called the generalized Stackelberg payoff, which can be greater than the standard Stackelberg payoff due to imperfect observability. Indeed the set of self-confirmed best-replies can be greater than the set of best-replies, but the two concepts coincide when the long-run player’s actions are identified, i.e. when $q(\cdot|\alpha_1, \alpha_2) = q(\cdot|\alpha_1', \alpha_2')$ implies $\alpha_1 = \alpha_1'$ for all $\alpha_1, \alpha_1' \in \Delta(A_1)$ and all $\alpha_2 \in \Delta(A_2)$ that are not weakly dominated. Finally, the stage game is called non-degenerate if there is no undominated pure action $\alpha_2 \in A_2$ with $g_1(\cdot, \alpha_2) = g_1(\cdot, \alpha_2')$ for some $\alpha_2 \in \Delta(A_2) \setminus \{\alpha_2\}$.

Theorem 1 (Fudenberg and Levine (1992)). Suppose that behavioral types have full support, i.e. $\Omega^b$ is a dense subset of $\Delta(A_1)$. Then, for every strategic type $\omega$ with $\mu(\{\omega\}) > 0$,

$$g_1(\omega) \leq \liminf_{\delta \to 1} N_1(\omega, \delta) \leq \limsup_{\delta \to 1} \bar{N}_1(\omega, \delta) \leq \bar{g}_1(\omega).$$

Moreover, if the stage game is non-degenerate and the long-run player’s action are identified, then $\bar{g}_1(\omega) = g_1(\omega)$ and hence,

$$\lim_{\delta \to 1} N_1(\omega, \delta) = \{\bar{g}_1(\omega)\}.$$

3 Reputation under Frequent Interactions

Turning to the analysis of reputation effects in games with frequent interactions, embed the discrete-time model of the previous section in continuous time, so each period of interaction has length $\Delta > 0$. The long-run player discounts his flow payoffs exponentially at rate $r > 0$, so his discount factor across periods is $\delta = e^{-r\Delta}$. As in Abreu, Milgrom, and Pearce (1991), the monitoring structure $(Y^\Delta, q^\Delta)$ is allowed to vary with $\Delta$. In particular, the informativeness of each period’s signals may deteriorate as the periods shrink, as when the signals are sampled from an underlying continuous-time process, such as when the noise is driven by a Poisson process (as in Abreu, Milgrom, and Pearce (1991)), a Brownian motion (as in Sannikov and Skrzypacz (2007) and Fudenberg and Levine (2007, 2009)) or, more generally, a Lévy process (as in Sannikov and Skrzypacz (2010)).
The ensuing analysis relies on Gossner (2011), who uses methods from information theory (Cover and Thomas, 2006) to derive payoff bounds from reputation effects that improve on the FL bounds for fixed discount factor and monitoring structure. Given a pair of probability measures \( P \) and \( Q \) on a measurable space \( X \), the relative entropy (also called Kullback-Leibler divergence) of \( Q \) with respect to \( P \) is:

\[
H(Q \| P) := \begin{cases} 
\int \log \frac{dQ}{dP} \; dQ, & \text{if } Q \text{ is absolutely continuous w.r.t. } P, \\
\infty, & \text{otherwise},
\end{cases}
\]

where \( dQ/dP \) denotes the Radon-Nikodym derivative. The relative entropy is non-negative, equal to zero if \( Q = P \), and lower semi-continuous when viewed as a function defined on \( \Delta(X) \times \Delta(X) \) mapping into \( \mathbb{R} \cup \{+\infty\} \).

For each \( \alpha_1, \alpha'_1 \in \Delta(A_1) \) and \( \alpha_2 \in \Delta(A_2) \) define

\[
h(\alpha_1, \alpha'_1; \alpha_2) := \liminf_{\Delta \to 0} \frac{1}{\Delta} H(q^\Delta(\cdot|\alpha_1, \alpha_2) \| q^\Delta(\cdot|\alpha'_1, \alpha_2)),
\]

which inherits the properties of the relative entropy and is thus non-negative, equal to zero if \( \alpha_1 = \alpha'_1 \) for any \( \alpha_2 \), and lower semi-continuous when viewed as a function defined on \( \Delta(A_1) \times \Delta(A_1) \times \Delta(A_2) \) taking values in \( \mathbb{R} \cup \{+\infty\} \). The function \( h \) quantifies the flow of information that the short-run players receive about the actions of the long-run player in the limit when the interactions become arbitrarily frequent.

A mixed action \( \alpha_2 \in \Delta(A_2) \) is an information-flow confirmed best-reply to \( \alpha_1 \in \Delta(A_1) \), written \( \alpha_2 \in B_2^*(\alpha_1) \), if \( \alpha_2 \) is not weakly dominated and \( \alpha_2 \leq \arg \max_{\alpha'_2} g_2(\alpha'_1, \alpha'_2) \) for some \( \alpha'_1 \in \Delta(A_1) \) with \( h(\alpha_1, \alpha'_1; \alpha_2) = 0 \). The long-run player’s actions are information-flow identified if \( h(\alpha_1, \alpha'_1; \alpha_2) = 0 \) implies \( \alpha_1 = \alpha'_1 \) for all \( \alpha_1, \alpha'_1 \in \Delta(A_1) \) and undominated \( \alpha_2 \in \Delta(A_2) \). For each \( \omega \in \Omega^b \) define

\[
g^*_1(\omega) := \sup_{\alpha_1 \in \Delta(A_1)} \inf_{\alpha_2 \in B^*_2(\alpha_1)} g_1(\alpha_1, \alpha_2, \omega)
\]

and

\[
g^*_1(\omega) := \sup_{\alpha_1 \in \Delta(A_1)} \sup_{\alpha_2 \in B^*_2(\alpha_1)} g_1(\alpha_1, \alpha_2, \omega).
\]

Finally, denote by \( N_1(\omega, r, \Delta) \) and \( \tilde{N}_1(\omega, r, \Delta) \) the infimum and supremum, respectively, of the set of Nash equilibrium payoffs of type \( \omega \) in the repeated game with discount rate \( r \) and period length \( \Delta \) (i.e. the repeated game of the previous section with discount factor \( e^{-r\Delta} \) and monitoring structure \( (Y^\Delta, q^\Delta) \)).

The following result is the analogue of Theorem 1 for games with frequent interactions. The proof, relegated to Appendix A.1, invokes the main result of Gossner (2011).

**Theorem 2 (Reputation under frequent interactions).** Suppose that behavioral types have full support, i.e. \( \Omega^b \) is a dense subset of \( \Delta(A_1) \). Then, for every strategic type \( \omega \) with \( \mu(\{\omega\}) > 0 \),

\[
g^*_1(\omega) \leq \liminf_{r \to 0} \liminf_{\Delta \to 0} N_1(\omega, r, \Delta) \leq \limsup_{r \to 0} \limsup_{\Delta \to 0} \tilde{N}_1(\omega, r, \Delta) \leq g^*_1(\omega).
\]

\(^7\)The lower semi-continuity is relative to the norm topology on \( \Delta(X) \).
Moreover, if the stage game is non-degenerate and the long-run player’s action are information-flow identified, then $g^*_1(\omega) = g^*_1(\omega)$, and hence

$$\lim_{r \to 0} \lim_{\Delta \to 0} N_1(\omega, r, \Delta) = \{g^*_1(\omega)\}.$$ 

To illustrate the result and discuss its implications, four examples of high-frequency information structures are examined below. In Example 1 the public signal is a continuous-time counting process in which the Poisson arrival rate is controlled by the players. An expression for the flow of information is derived and, applying Theorem 2, it is shown that the high-frequency repeated game features strong reputation effects when the prior has full support on behavioral types. By contrast, when it is common knowledge that long-run player is a normal type, for a wide class of stage-game payoffs only static equilibria can be sustained as equilibria of the repeated game when the frequency of interaction is high enough. In Example 2 the underlying continuous-time signal follows a diffusion process. Similar to the Poisson case, here too the incomplete information game features reputation effects that are robust to frequent interactions, while the complete information game only possesses degenerate equilibria when the interactions are frequent. In Example 3 the public signal follows a binomial random walk where the probability of an up-tick is controlled by the players. The size of the up- and down-ticks, as well as their probabilities, scale with the period length in such a way that the process approximates the continuous time diffusion of Example 2. It is shown that, for this particular discrete-time approximation, the information flow takes exactly the same form as in Example 2. However, this is not a general phenomenon. In Example 4 it is shown that reputation bounds that apply to a continuous-time limit game may be inadequate for high-frequency repeated games.

**Example 1 (Reputation under Poisson monitoring).** As in Abreu, Milgrom, and Pearce (1991) assume there is an underlying continuous-time counting process, $N(t), t \in [0, \infty)$, whose Poisson arrival rate is a function of the action profile. The players can adjust their actions at times $t = 0, \Delta, 2\Delta, \ldots,$ and at any time $t$ in this grid the value of $N(t)$ is publicly observed before the players get to choose their time-$t$ actions. Thus, in the formalism of Section 2, the public signals are the increments $y^n = N(n\Delta) - N((n-1)\Delta), n = 0, 1, 2, \ldots$, so that

$$q^\Delta(y|a) = \frac{(\lambda(a)\Delta)^y}{y!} \exp(-\lambda(a)\Delta), \quad y \in Y^\Delta = \{0, 1, 2, \ldots\},$$

where $\lambda : A \to (0, \infty)$ is the arrival rate as a function of the action profile played. Then, as shown in Appendix A.2, the information flow takes the form:

$$h(\alpha_1, \alpha'_1; \alpha_2) = \lambda(\alpha'_1, \alpha_2) - \lambda(\alpha_1, \alpha_2) - \lambda(\alpha_1, \alpha_2) \log \left( \frac{\lambda(\alpha'_1, \alpha_2)}{\lambda(\alpha_1, \alpha_2)} \right),$$

for all $\alpha_1, \alpha'_1 \in \Delta(A_1)$ and $\alpha_2 \in \Delta(A_2)$. By the concavity of the logarithm $h(\alpha_1, \alpha'_1; \alpha_2)$ is nonnegative, and it can be readily verified that $h(\alpha_1, \alpha'_1; \alpha_2) = 0$ if, and only if, $\lambda(\alpha'_1, \alpha_2) = \lambda(\alpha_1, \alpha_2)$. In particular, when the long-run player has two pure actions in the stage game, the game is information-flow identified if, and only if, those two actions induce different Poisson arrival rates regardless of the short-run player’s mixed action. In this case, by Theorem 2, the set of Nash equilibrium payoffs of a strategic type will converge to his Stackelberg payoff as the period length and the discount rate tend to zero (assuming a non-degenerate stage game and a prior with full support).
By contrast, when it is common knowledge that the long-run player is a normal type, for a wide class of frequent interaction games with Poisson signals in which the long-run player’s action is identified, the set of equilibrium payoffs of the long-run player collapses to the convex hull of his set of stage-game Nash equilibrium payoffs, irrespective of patience. For example, consider a repeated product-choice game where the long-run player is a firm choosing whether to exert (E)ffort or (S)hirk, and the short-run players are consumers choosing whether to buy the (H)igh-end version or the (L)ow-end version of the firm’s product. The stage-game payoff matrix is displayed in Figure 1 below. Suppose the arrival of a signal is “good news” about firm effort (e.g. “likes” in the firm’s Facebook page), so that \( \lambda (E, a_2) > \lambda (S, a_2) \) for all \( a_2 \). Then, if there is no uncertainty over types, the only equilibrium of the complete information game (when \( \Delta \) is sufficiently small) is the repetition of the stage-game Nash equilibrium \((S, L)\) after every history regardless of patience, by an argument similar to that in Abreu, Milgrom, and Pearce (1991).^8

The stark contrast between the complete and information scenarios is due to the difference in the relevant time frame for incentives. Under complete information, the relevant time frame for incentives is the period of interaction: a one-shot deviation by the long-run player will affect only his current continuation value, and thus incentives will be determined by how this continuation value varies with the public signals. By contrast, under incomplete information, the relevant time frame for incentives is the “long-run”: a one-shot deviation affects all future continuation values through the inference performed by the short-run players about the long-run player’s type—learning has persistence. The benefits from this additional persistence outweigh the losses due to the inefficient provision of incentives that arises when the informativeness of the signals deteriorates as the period length shrinks to zero.

This example will be carefully re-examined in Section 4, which provides an in-depth discussion of the relation between the reputation bounds from Theorem 2 and the high-frequency limit of the Fudenberg-Levine bounds. There it will be shown that the high-frequency limit of the Fudenberg-Levine bounds turns out to be completely uninformative in all games with Poisson signals, thus highlighting the importance of the entropic information flow approach.

\textbf{Example 2 (Reputation under Brownian monitoring).} As in Faingold and Sannikov (2011), assume there is an underlying diffusion process, \( X(t) \), which evolves according to

\[
dX(t) = \mu (a(t)) \, dt + dZ(t), \quad t \in [0, \infty),
\]

^8Abreu, Milgrom, and Pearce (1991) examine a symmetric partnership game between two long-run players. Nonetheless, their proof that strongly symmetric equilibria collapse to static Nash equilibrium when signals are good news and the period length is short enough carries over \textit{mutatis mutandis} to the case of a single long-run player in the product choice game (with no symmetry restriction, as there is a single long-run player).
where $Z$ is a standard Brownian motion and $\mu : A \to \mathbb{R}$ determines the drift of $X$ as a function of the profile of actions. As in the previous example, players can adjust their actions only at times $t = 0, \Delta, 2\Delta, \ldots$, and at any time $t$ in this grid the value of $X(t)$ is publicly observed before the players get to choose their actions $a(t) = (a_1(t), a_2(t))$, so that

$$Y^\Delta = \mathbb{R}, \quad q^\Delta(\cdot | a) \sim \mathcal{N}(\mu(a)\Delta, \Delta).$$

Thus, for any $\Delta > 0$ and any mixed-action profile, the induced distribution over public signals is a mixture of Gaussian distributions. Unfortunately, the relative entropy of a pair of mixtures of Gaussians does not admit a representation in terms of the relative entropies of the pairs of Gaussians in the support, and no closed-form expression is known. But our interest is not in the relative entropy for a fixed $\Delta$, but rather in the flow of information, which is the limit as $\Delta \to 0$ of the quotient between the relative entropy and the period length $\Delta$. As shown in Appendix A.3, this limit turns out to have a simple expression:

$$h(\alpha_1, \alpha'_1, \alpha_2) = \frac{1}{2} \left( \mu(\alpha'_1, \alpha_2) - \mu(\alpha_1, \alpha_2) \right)^2, \quad \alpha_1, \alpha'_1, \alpha_2 \in \Delta(A_1), \alpha_2 \in \Delta(A_2).$$

Thus, $h(\alpha_1, \alpha'_1; \alpha_2) = 0$ if and only if $\mu(\alpha'_1, \alpha_2) = \mu(\alpha_1, \alpha_2)$, and hence, if the long-run player has two pure actions in the stage game, then the game is information-flow identified if, and only if, the drifts induced by those two actions are different regardless of the mixed action of the short-run player.

We conclude that, similar to Example 1, the incomplete information game features strong reputation effects that are robust to frequent interactions, despite the fact that the equilibrium of the complete information game becomes degenerate as the period length tends to zero, as shown in Fudenberg and Levine (2007, 2009).

Example 3 (Reputation under binomial random walk monitoring). Consider the monitoring technology given by

$$Y^\Delta = \{ \sqrt{\Delta}, -\sqrt{\Delta} \}, \quad q^\Delta \left( \pm \sqrt{\Delta} \mid a \right) = \frac{1}{2} \left( 1 \pm \mu(a) \sqrt{\Delta} \right),$$

for a given function $\mu : A \to \mathbb{R}$. This relates to Brownian monitoring via Donsker’s invariance principle: for each $a \in A$, the family of stochastic processes $(X_\Delta)_{\Delta>0}$,

$$X_\Delta(t) = \sum_{n=0}^{[t/\Delta]} y^n_{\Delta}, \quad t \in [0, \infty),$$

where $y^n_{\Delta}, y^2_{\Delta}, \ldots$ are i.i.d. random variables with distribution $q^\Delta(\cdot | a)$, weakly converges to a diffusion process with drift $\mu(a)$ and volatility 1 as $\Delta \to 0$. Here, as expected, the information flow $h$ takes exactly the same form as in the previous example, as can be readily verified. However, this is not true for all discrete-time approximations of Brownian motion, as shown in the next example.

Example 4 (Inadequacy of continuous-time reputation bounds). Here we consider a pair of monitoring structures approximating a continuous-time diffusion. In both examples the equilibria of high frequency repeated games will feature strong reputation effects, while the equilibria of the continuous-time limit game will either feature weaker reputation effects, or will feature none whatsoever. The
examples thus reinforce the value of the information flow approach to reputations in high frequency games, and show that reputation bounds based on a continuous time limit game may be inadequate for high-frequency games. Consider the first example:

\[
Y^\Delta = \{\mu(a)\Delta \pm \sqrt{\Delta} : a \in A}\}, \quad q^\Delta(\mu(a)\Delta \pm \sqrt{\Delta} | a) = \frac{1}{2},
\]

for given \(\mu : A \to \mathbb{R}\). Here, as in the previous example, the cumulative process \(X_\Delta\) satisfies the conditions of Donsker’s invariance principle and so it converges weakly to a diffusion with drift \(\mu\) and unit volatility. If the long-run player has two actions in the stage game, and those two actions induce different values of \(\mu\) irrespective of the short-run player’s action, then, for every positive \(\Delta\), the short-run players can perfectly monitor the long-run player (so that \(h(\alpha_1, \alpha'_1; \alpha_2) = \infty\) whenever \(\alpha_1 \neq \alpha'_1\)). In particular, the long-run player’s actions are identified for every positive \(\Delta\). Thus, if the stage game is non-degenerate and the prior has full support on behavioral types, the set of Nash equilibrium payoffs of any strategic type converges to his Stackelberg payoff as \(\Delta\) tends to 0, irrespective of patience (i.e., for any fixed positive discount rate \(r\)). By contrast, the continuous-time limit game has “full support” imperfect monitoring: changes of drift induce absolutely continuous changes in the probability measure over the paths of the diffusion, over any finite horizon. Faingold (2013) has shown that in such a continuous time game the set of Nash equilibrium payoffs of a strategic type must converge to his Stackelberg payoff as the discount rate \(r\) tends to 0. However, for fixed discount rates, the equilibrium payoffs in the continuous-time game are lower than the Stackelberg payoff (Faingold and Sannikov, 2011), unlike the frequent interaction limit discussed above, which does not require \(r \to 0\) to yield a reputation effect.

The discontinuity in the example below is even more dramatic:

\[
Y^\Delta = \{-\sqrt{\Delta}, \pm \sqrt{\Delta}(1-\Delta)\},
\]

and for each \(a_1 \in A_1 := \{0, 1\}, a_2 \in A_2,\)

\[
q^\Delta(y | a) = \begin{cases} 
\frac{1}{2}((1-a_1)(1-\Delta) + a_1 \Delta) & \text{if } y = \pm \sqrt{\Delta}(1-a_1) \\
\frac{1}{2}(a_1(1-\Delta) + (1-a_1)\Delta) & \text{if } y = \pm \sqrt{\Delta}(1-(1-a_1)\Delta).
\end{cases}
\]

Here the cumulative process \(X_\Delta\) converges weakly to a driftless Brownian motion irrespective of \(a\).\(^9\) In particular, in the continuous time limit the signal is completely uninformative, and therefore only static equilibria are possible. Nonetheless, for small but positive \(\Delta\) we have almost perfect monitoring: for any pair of mixed action profiles \(\alpha, \alpha'\),

\[
H(q^\Delta(\cdot | \alpha), q^\Delta(\cdot | \alpha')) \geq 2|\alpha_1(0) - \alpha'_1(0)|^2(1-2\Delta)^2,
\]

so that \(h(\alpha_1, \alpha'_1; \alpha_2) = \infty\) whenever \(\alpha_1 \neq \alpha'_1\).\(^10\) As in the previous example, here a reputation result obtains in the limit as the period length \(\Delta\) tends to zero for any fixed discount rate. But, unlike

\(^9\)The conditions of Donsker’s invariance principle are satisfied: \(Y^\Delta \to \{0\}\) as \(\Delta \to 0\), and for any action profile \(a\) the expectation of \(y\) is zero and the variance of \(y\) is \(\Delta + o(\Delta)\).

\(^{10}\)The total variation distance between \(q^\Delta(\cdot | \alpha)\) and \(q^\Delta(\cdot | \alpha')\) equals \(|\alpha_1(0) - \alpha'_1(0)|(1 - 2\Delta)\). The lower bound on \(H(q^\Delta(\cdot | \alpha), q^\Delta(\cdot | \alpha'))\) then follows from Pinsker’s inequality, which asserts that relative entropy is bounded below by twice the square of the total variation (Cover and Thomas, 2006).
the previous example, only static equilibria are possible in the continuous-time game, irrespective of patience. The examples show that reputation bounds that apply to a continuous-time limit game may be inadequate for high-frequency repeated games.\textsuperscript{11,12}

4 Relation to the FL Bounds

Recall that Fudenberg and Levine (1992) provide bounds on the set of Nash equilibrium payoffs of strategic types, which in our setting depend both on the discount rate and the frequency of interactions (through the monitoring structure). Thus the limits of the FL bounds (as the period length and the discount rate go to zero) can be compared to the payoff bounds from Theorem 2. How do those bounds compare?

Below I argue that the difference between those bounds can be quite dramatic. In the product choice game with Poisson signals of Example 1, the frequent interaction limit of the FL bounds turns out to be completely uninformative, while the reputation bounds based on the information flow are tight and equal the Stackelberg payoff in the patient limit. By contrast, such a dramatic difference can never arise in the canonical discrete-time setting with fixed period length: for a fixed monitoring technology the improvement of the Gossner bounds over the FL bounds is known to disappear as the long-run player becomes arbitrarily patient.

Recall that the calculation of the FL bounds rely on the following statistical lemma, which is closely related to the classic result of Blackwell and Dubins (1962) on “merging” of Bayesian forecasts.

**Uniform Merging Lemma** (Fudenberg and Levine (1992)).\textsuperscript{13} Let \((\Omega, \mathcal{F})\) be a measurable space endowed with a filtration \((\mathcal{F}_n)_{n\geq 0}\). For all \(\phi_0 \in (0, 1]\), \(\varepsilon > 0\) and \(\eta > 0\) there exists a non-negative integer \(K\) such that, for all \(P, Q\) and \(Q'\) probability measures on \((\Omega, \mathcal{F})\) satisfying \(P = \phi_0 Q + (1 - \phi_0) Q'\),

\[
Q\left(\#\{n \geq 0 : d_n(Q, P) \geq \varepsilon\} \geq K\right) \leq \eta,
\]

\textsuperscript{11}In this last example, the signal distribution can be interpreted as a two-stage lottery, where the second stage lotteries have disjoint support: if \(a_1 = 0\) (resp. \(a_1 = 1\)), with probability \(1 - \Delta\) (resp. \(\Delta\)) signal \(y\) is drawn from \(\{\pm \sqrt{\Delta}\}\) with equal probabilities in the second stage, and with probability \(\Delta\) (resp. \(1 - \Delta\)) \(y\) is is drawn from \(\{\pm \sqrt{(1 - \Delta)}\}\) with equal probabilities in the second stage. A variation of the example involves a single-stage lottery: \(y\) is drawn from \(\{\pm \sqrt{\Delta}\}\) with equal probabilities if \(a_1 = 0\), and drawn from \(\{\pm \sqrt{(1 - \Delta)}\}\) with equal probabilities if \(a_1 = 1\). This variant has essentially the same features as the example above, but the monitoring does not have full support.

\textsuperscript{12}While both examples of discontinuity from discrete to continuous time above are in a continuous diffusion setting, similar examples exist in settings with pure jump processes. Such examples can be designed, for instance, by having the long-run player’s actions affect the support of the distribution of jumps slightly, but not their arrival rate.

\textsuperscript{13}An implication of Blackwell and Dubins (1962) merging result is that \(d_n(Q, P) \to 0\) \(Q\)-almost surely. The uniform merging lemma strengthens this implication under a stronger form of absolute continuity, which came to be known as the “grain of truth” condition. The uniform merging result establishes the uniformity of the rate of convergence \(K\) over all \(P, Q\) and \(Q'\) that satisfy \(P = \phi Q + (1 - \phi) Q'\) for a fixed size \(\phi\) of the “grain of truth.” Merging of forecasts plays a central role also in the literature on rational learning in games (Kalai and Lehrer (1993)).
where
\[ d_n(Q, P) := \text{ess sup}_{B \in \mathcal{F}_{n+1}} |P(B|\mathcal{F}_n) - Q(B|\mathcal{F}_n)|. \]

For concreteness, consider the product-choice game of Example 1, which has Poisson signals. To simplify exposition, assume that only the long-run player controls the arrival rate \( \lambda \). Fix the period length \( \Delta > 0 \) and consider a behavioral type \( \hat{\omega} \) committed to playing \( E \) with probability \( \hat{\alpha}_1(E) \) strictly greater than 1/2. Let \( \sigma \) be an arbitrary Nash equilibrium of the repeated game with period length \( \Delta \). Following the notation of Section 2, we write
\[ P_{\mu, \sigma} = \mu(\hat{\omega})P_{\hat{\omega}, \sigma} + (1 - \mu(\hat{\omega}))P_{\neg \hat{\omega}, \sigma}, \text{ where } P_{\neg \hat{\omega}, \sigma} = \frac{\sum_{\omega \neq \hat{\omega}} \mu(\omega)P_{\omega, \sigma}}{1 - \mu(\hat{\omega})}. \]

Consider a deviation in which the normal type mimics the behavioral type \( \hat{\omega} \), i.e. plays \( \hat{\alpha}_1 \) after every history. Since \( \hat{\alpha}_1(E) > 1/2 \), we can find \( \varepsilon(\Delta) > 0 \) small enough that for every mixed action \( \alpha_1 \),
\[ d_{TV}(q^\Delta(\cdot|\alpha_1), q^\Delta(\cdot|\hat{\alpha}_1)) \leq \varepsilon(\Delta) \implies \alpha_1(E) > 1/2, \]
where \( d_{TV} \) denotes the total variation distance. Thus, following any \( n \)-period public history \( h^n_2 \) satisfying \( d_{TV}(q^\Delta(\cdot|E_{\mu, \sigma}(\sigma_1^n | h^n_2)), q^\Delta(\cdot|\hat{\alpha}_1)) \leq \varepsilon(\Delta) \), the uninformed player expects the informed player to play \( E \) with probability greater than 1/2 and will therefore choose \( H \) with probability one. Thus, following all such histories, the normal type gets a payoff of 3\(-\hat{\alpha}_1(E)\) under the deviation.

Following FL, let us apply the uniform merging lemma to argue that, under the deviation, with high probability in all but finitely many periods the total variation distance between the short-run players’ equilibrium forecast of the public signal, and the true signal distribution under the deviation, is at most \( \varepsilon(\Delta) \). To carry out the argument, fix \( \eta > 0 \) and denote by \( K(\Delta, \eta, \hat{\omega}) \) the smallest integer \( K \) such that inequality (2) holds for \( \varepsilon = \varepsilon(\Delta) \). Let
\[ \phi = \mu(\hat{\omega}), \quad P = P_{\mu, \sigma}, \quad Q = P_{\hat{\omega}, \sigma}, \quad Q' = P_{\neg \hat{\omega}, \sigma} \text{ and } \mathcal{F}_n \]
equal to the product \( \sigma \)-field on \( H^n_2 \). Thus, under the deviation, with probability at least \( 1 - \eta \) there are at most \( K(\Delta, \eta, \hat{\omega}) \) periods in which the uninformed player is expecting the informed player to choose \( E \) with probability less than or equal to 1/2. Hence, with probability at least \( 1 - \eta \) under the deviation, there are at most \( K(\Delta, \eta, \hat{\omega}) \) periods in which the uninformed player is choosing \( H \) with probability less than one. Assuming, conservatively, that such exceptional periods are the \( K(\Delta, \eta, \hat{\omega}) \) initial periods yields the following lower bound for the payoff that the normal type receives in any Nash equilibrium:
\[ \text{FL}(r, \Delta, \eta, \hat{\omega}) := (1 - \eta)(3 - \hat{\alpha}_1(E)) \exp(-r\Delta K(\Delta, \eta, \hat{\omega})). \]
As \( K(\Delta, \eta, \hat{\omega}) \) only depends on the period length \( \Delta \) (through \( \varepsilon(\Delta) \)), on \( \eta \) and on the behavioral type \( \hat{\omega} \), letting \( r \to 0 \) first, followed by \( \eta \to 0 \) and then \( \hat{\alpha}_1(E) \to 1/2 \), we get the convergence of

14 The essential supremum of a family \( F \) of measurable functions defined on a probability space, written ess sup \( F \), is a measurable function \( g \) satisfying: (i) \( g \geq f \) a.s. for all \( f \in F \); and (ii) if \( h \) is a measurable function satisfying \( h \geq f \) a.s. for all \( f \in F \), then \( h \geq g \) a.s. See Proposition VI-1-1 in Neveu (1975) for the existence and, up to equivalence class, uniqueness of the essential supremum.

15 The total variation distance between a pair of probability measures, \( P \) and \( Q \), defined on a measurable space (\( \Omega, \Sigma \)) is \( d_{TV}(P, Q) = \sup_{A \in \Sigma} |P(A) - Q(A)| \).

16 Recall that when the uninformed player assigns probability greater than \( \frac{1}{2} \) to \( E \), her unique best-reply is to choose \( H \).
the lower bound $FL(r, \Delta, \eta, \hat{\omega})$ to the Stackelberg payoff of 2.5. This summarizes the argument in Fudenberg and Levine (1992).

Turning to the frequent interaction limit, we now consider what happens to $FL(r, \Delta, \eta, \hat{\omega})$ when we send $\Delta$ to 0 before sending $r$ to 0. To carry out this limit, we need to examine the speed with which $K(\Delta, \eta, \hat{\omega}) \rightarrow \infty$ as $\Delta \rightarrow 0$. And for that purpose we need to understand both:

(i) the speed with which $\varepsilon(\Delta) \rightarrow 0$ as $\Delta \rightarrow 0$, and

(ii) the speed with which the bound $K$ from the uniform merging lemma tends to $\infty$ as $\varepsilon \rightarrow 0$.

Regarding (i), note that for all $\alpha_1 \in \Delta(A_1)$,

$$d_{TV}(q^{\Delta}(\cdot|\alpha_1), q^{\Delta}(\cdot|\hat{\alpha}_1)) \leq |\hat{\alpha}_1(E) - \alpha_1(E)||\lambda(E) - \lambda(S)|\Delta + o(\Delta).$$

Since $\hat{\alpha}_1(E) > 1/2$, it follows that for any choice of $\varepsilon(\Delta)$ for which the implication in (3) holds for all $\alpha_1 \in \Delta(A_1)$, we must have $\varepsilon(\Delta) = O(\Delta)$.

As for (ii), the following result, which is proved in Appendix A.4, provides sharp estimates:

**Proposition 1 (Speed of uniform merging).** Let $(\Omega, \mathcal{F})$ be a measurable space endowed with a filtration $(\mathcal{F}_n)_{n \geq 0}$. For each $\phi_0 \in (0, 1]$, $\varepsilon > 0$ and $\eta > 0$, denote by $K^*(\varepsilon, \eta, \phi_0)$ the smallest positive integer such that inequality (2) holds for all probability measures $P, Q$ and $Q'$ on $(\Omega, \mathcal{F})$ satisfying $P = \phi_0 Q + (1 - \phi_0) Q'$. Then, for all $\phi_0 \in (0, 1]$, and $\eta > 0$,

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^2 K^*(\varepsilon, \eta, \phi_0) < \infty.$$

Furthermore, if $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all $n$, then for all $\gamma > 0$, $\phi_0 \in (0, 1]$ and small enough $\eta > 0$,

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{2-\gamma} K^*(\varepsilon, \eta, \phi_0) = \infty.$$

Now consider the FL bound (4) with $K(\Delta, \eta, \hat{\omega}) := K^*(\varepsilon(\Delta), \eta, \mu(\hat{\omega}))$. Then, the second part of Proposition 1 above (with $\gamma = 1$) implies that $\varepsilon(\Delta) K(\Delta, \eta, \hat{\omega}) \rightarrow \infty$ as $\Delta \rightarrow 0$. Since we argued that $\varepsilon(\Delta) = O(\Delta)$, it follows that the real time $\Delta \times K(\Delta, \eta, \hat{\omega})$ tends to $\infty$ as $\Delta$ shrinks to 0, and hence that $FL(r, \Delta, \eta, \hat{\omega}) \rightarrow 0$ as $\Delta \rightarrow 0$, for all $\hat{\omega}$ with $\hat{\alpha}_1(E) > 1/2$, and all $\eta > 0$ small enough.

To sum up, we have shown that in the product-choice game with Poisson monitoring of Example 1 the FL bounds become completely uninformative in the high frequency limit. Nonetheless, using the entropic bounds in Gossner (2011), we were able to show that the reputation effect retains its power: for any $\varepsilon > 0$, if the long-run player is patient enough and the period of interaction is short enough, in any Nash equilibrium the normal type will receive a payoff within $\varepsilon$ of 2.5, the Stackelberg payoff.

* * * *

In light of the discussion above, the reader may wonder why the gap between the Gossner and the FL bounds can become so dramatic in the frequent interaction limit. To get a precise intuition, an
in-depth examination of FL’s uniform merging result is in order. The key idea behind that result is the well-known property that the martingale \( \phi_n \) of posterior beliefs (on the event that \( Q \) is the true data generating process) has a uniformly bounded quadratic variation:

\[
E_P \sum_{n=0}^{\infty} E_P (|\phi_{n+1} - \phi_n|^2 | \mathcal{F}_n) \leq 1,
\]

and thus, since \( P = \phi_0 Q + (1 - \phi_0) Q' \),

\[
E_Q \sum_{n=0}^{\infty} E_P (|\phi_{n+1} - \phi_n|^2 | \mathcal{F}_n) \leq 1/\phi.
\]

By Markov’s inequality, we conclude that:

For all \( \varepsilon > 0 \) and \( K > 0 \), with probability on the order of \( O(1/(\varepsilon K)) \) under \( Q \), there are at most \( K \) periods in which the variance of one-period ahead posterior beliefs

\[
E_P (|\phi_{n+1} - \phi_n|^2 | \mathcal{F}_n),
\]

exceeds \( \varepsilon \).

But what is the relation between the one-period variance of posteriors, \( E_P (|\phi_{n+1} - \phi_n|^2 | \mathcal{F}_n) \), and the total variation distance between the one-period forecast \( P_n = P|\mathcal{F}_{n+1} \) and the true distribution \( Q_n = Q|\mathcal{F}_{n+1} \)? Since under \( Q \) the posterior odds ratio \( (1-\phi_n)/\phi_n \) is a martingale, Doob’s maximal inequality implies that the probability that \( \phi_n \) ever falls below any given positive threshold \( \eta < \phi_0 \) is at most \( \eta/\phi_0 \). Therefore, with probability at least \( 1 - \eta/\phi_0 \) under \( Q \),

\[
E_P (|\phi_{n+1} - \phi_n|^2 | \mathcal{F}_n) \geq \eta^2 E_P (|1 - \phi_{n+1}/\phi_n|^2 | \mathcal{F}_n),
\]

and so it only remains to examine the relationship between the variance of the posterior-to-prior ratio, \( E_P (|1 - \phi_{n+1}/\phi_n|^2 | \mathcal{F}_n) \), and the total variation distance between \( P_n \) and \( Q_n \). By Bayes’ rule, the posterior-to-prior ratio \( \phi_{n+1}/\phi_n \) equals the likelihood ratio \( \frac{dQ_n}{dP_n} \) between the conditional and unconditional distributions. Hence, by Jensen’s inequality, for any \( A \in \mathcal{F}_{n+1} \),

\[
E_P (|1 - \phi_{n+1}/\phi_n|^2 | \mathcal{F}_n) \geq \int_A |1 - \frac{dQ_n}{dP_n}|^2 dP_n \geq |P_n(A) - Q_n(A)|^2,
\]

and hence, taking the supremum over \( A \in \mathcal{F}_{n+1} \) yields

\begin{equation}
E_P (|1 - \phi_{n+1}/\phi_n|^2 | \mathcal{F}_n) \geq d_{TV}(P_n, Q_n)^2.
\end{equation}

Therefore, with probability at least \( 1 - \eta/\phi_0 \) under \( Q \),

\[
E_P (|\phi_{n+1} - \phi_n|^2 | \mathcal{F}_n) \geq \eta^2 d_{TV}(P_n, Q_n)^2,
\]

which leads us to conclude:

For all \( \eta > 0 \), \( \varepsilon > 0 \) and \( K > 0 \), with probability on the order of \( O(1/(\eta^2 \varepsilon^2 K) + \eta) \) under \( Q \), there are at most \( K \) periods in which the total variation distance between \( Q_n \) and \( P_n \) exceeds \( \varepsilon \).
With the arguments above in place, we can now explain why relative entropy leads to sharper reputation bounds. The key is to note that inequality (5) is not tight. Indeed, a simple argument using the inequality \( x \log x \leq x - 1 + (x - 1)^2 \) for \( x > 0 \) yields

\[
E_P \left( |1 - \phi_{n+1}/\phi_n|^2 | F_n \right) \geq H(Q_n | P_n),
\]

which, by Pinsker’s inequality \( d_{TV}(Q_n, P_n) \leq \sqrt{\frac{1}{2} H(Q_n \| P_n)} \) (Cover and Thomas, 2006), is an improvement on (5). Mimicking the argument in the previous paragraph with inequality (6) in place of (5) leads to:

\[
\text{For all } \eta > 0, \varepsilon > 0 \text{ and } K > 0, \text{ with probability on the order of } O(1/(\eta^2 \varepsilon K) + \eta) \text{ under } Q, \text{ there are at most } K \text{ periods in which the relative entropy of } Q_n \text{ with respect to } P_n \text{ exceeds } \varepsilon.
\]

Thus, for any threshold \( \varepsilon > 0 \) we can choose an integer \( K_H \) large enough so the probability of the event that there are more than \( K_H \) surprises—periods in which the relative entropy \( H \) of the true distribution with respect to the Bayesian forecast exceeds the threshold \( \varepsilon \)—is arbitrarily small. Moreover, as we send the surprise threshold \( \varepsilon \) to 0, so long as we scale \( K_H \) homogeneously with \( 1/\varepsilon \), the probability that there are more than \( K_H \) surprising periods remains small.

This scaling property is the key to understand why reputation effects retain their power in the Poisson monitoring case with high-frequency play. Indeed, if we adopt the relative entropy measure of surprise in the Poisson case, then the surprise threshold \( \varepsilon \) required for the short-run players to play a best-reply to the commitment action is on the order of \( O(\Delta) \). Then the bound \( \Delta \times K_H \) on the real time during which the uninformed players get surprised does not blow up as \( \Delta \to 0 \). By contrast, if we measure surprise by the total variation metric, while the required threshold \( \varepsilon \) for best-reply behavior remains on the order of \( O(\Delta) \), now the associated bound \( K_{TV} \) on the number of surprising periods scales homogeneously with \( 1/\varepsilon^2 \) rather than with \( 1/\varepsilon \).\(^{17}\) Therefore, the bound \( \Delta \times K_{TV} \) on the real length of time during which the uniformed players expect a surprise blows up as \( \Delta \to 0 \).

The arguments above suggest that a dramatic difference between the Gossner and the FL reputation bounds obtains whenever the gap between the right and left sides of Pinsker’s inequality,

\[
d_{TV} \left( q^\Delta(\cdot|\hat{a}), q^\Delta(\cdot|a) \right) \leq \sqrt{\frac{1}{2} H q^\Delta(\cdot|\hat{a}) q^\Delta(\cdot|a)},
\]

is significant. We have seen that this is indeed the case under Poisson monitoring (Example 1): the left side of the inequality is on the order of \( \Delta \) while the right side is on the order of \( \sqrt{\Delta} \).

Conversely, we expect that when the left and right sides of Pinsker’s inequality converge to zero at comparable rates as \( \Delta \to 0 \), the patient frequent-interaction limits of the Gossner and FL bounds coincide. It can be shown that this is precisely what happens under Brownian monitoring (Example 2).

\(^{17}\)Proposition 1 shows that the property that \( K_{TV} \) does not scale with \( 1/\varepsilon \) holds not only when the bound \( K_{TV} \) is chosen using the martingale argument above, but it holds generally for any choice of \( K_{TV} \) that provides a uniform upper bound on the number of surprises with high probability.
Indeed, in that case, assuming that the long-run player has only two actions $a_1$ and $a'_1$ (to simplify exposition), the total variation distance between $q^\Delta(\cdot|\alpha)$ and $q^\Delta(\cdot|\hat{\alpha})$ has a closed-form expression:

$$d_{TV}(q^\Delta(\cdot|\alpha), q^\Delta(\cdot|\hat{\alpha})) = |\alpha_1(a_1) - \hat{\alpha}_1(a_1)| \left(2\Phi(\mu(a_1) - \mu(a'_1) / \sqrt{\Delta}) - 1\right),$$

where $\Phi$ is the cdf of a zero-mean, unit-variance Gaussian random variable. Thus,

$$d_{TV}(q^\Delta(\cdot|\alpha), q^\Delta(\cdot|\hat{\alpha})) = \frac{|\alpha_1(a_1) - \hat{\alpha}_1(a_1)|}{\sqrt{2\pi}} \left|\mu(a_1) - \mu(a'_1)\right| \sqrt{\Delta} + O(\Delta).$$

On the other hand, we have seen in Example 2 that $H(q^\Delta(\cdot|\hat{\alpha})\|q^\Delta(\cdot|\alpha))$ is on the order of $\Delta$. We conclude that in the Brownian monitoring case (Example 2) the left and right sides of Pinsker’s inequality are both on the order of $\sqrt{\Delta}$, and hence the FL and Gossner bounds coincide in the patient frequent-interaction limit in this case.

Similarly, when signals follow the binomial random walk of Example 3, both sides of Pinsker’s inequality are on the order of $\sqrt{\Delta}$, and so there is no limiting gap between the FL and the Gossner bounds. Finally, for the random walks of Example 4, both sides of Pinsker’s inequality are on the order of a constant, so there is clearly no limiting gap between the FL and the Gossner bounds.

### 5 Conclusion

The insight that reputation provides a foundation for commitment in long-run relationships is a major cornerstone of dynamic game theory. Previous research has established that for a wide class of complete information repeated games between a long-run player and a population of short-run players, high-frequency play renders non-trivial intertemporal incentives impossible to sustain due to the extreme noisiness in the observation of the long-run players’ actions within short periods of interaction. This paper has shown that, unlike complete information incentives, reputation effects remain powerful in the high frequency limit.

Research on reputation in high frequency games could be fruitfully advanced by the investigation of equilibrium behavior. This paper focuses on equilibrium payoffs in the patient limit, and so it says nothing about the equilibrium dynamics of building, managing and exploiting (and hence depleting) one’s reputation. The analysis of equilibrium behavior in discrete-time reputation games with fixed discounting is complicated by the non-Markovian character of equilibria and their multiplicity. By contrast, for a class of continuous-time reputation games with Brownian signals, Faingold and Sannikov (2011) has shown that equilibrium is unique, Markovian and characterized by an ordinary differential equation, greatly simplifying the analysis of equilibrium behavior. So it would be interesting to investigate under what conditions equilibrium of high-frequency games with signals approximating a continuous-time diffusion inherit (at least approximately) that nice structure.

However, Example 4 points to a major difficulty: vanishing details of the discrete-time approximation matter, in that the equilibrium outcomes of a limiting continuous-time game can be dramatically different from the equilibrium outcomes of close-by high-frequency games. Now, the discontinuities of Example 4 are rather extreme, as they consist of high-frequency games with perfectly (or almost perfectly) revealing signals converging to a continuous-time game with full support Brownian
signals. In particular, since monitoring is perfect (or almost perfect) under those high-frequency approximations, the equilibrium dynamics of reputation does not turn out to be particularly interesting. But examples from Sadzik and Stacchetti (2015) point to a less extreme, and hence more interesting, possibility: discrete-time approximations of Brownian signals for which the high-frequency flow of information is finite and yet different from the flow of information in the limit game. A plausible conjecture is that the basic insight of Sadzik and Stacchetti (2015), which is conveyed in the context of high-frequency agency problems, carries over to high-frequency reputation games. Namely, that it is possible to characterize equilibrium behavior in a broad class of high-frequency reputation games by a version of the differential equation of Faingold and Sannikov (2011) suitably modified to incorporate the informationally relevant details of the information structure approximating the continuous-time diffusion, analogously to Sadzik and Stacchetti (2015). This seems worth pursuing in future research.

A Appendix

A.1 Proof of Theorem 2

The main result of Gossner (2011) yields the following lower bound on $N_1(\omega, r, \Delta)$:

$$\sup_{\hat{\omega} \in \Omega^b} w^\Delta_{\hat{\omega}} ((1 - e^{-r\Delta}) \log \mu(\hat{\omega})) = \sup_{\hat{\omega} \in \Omega^b} w^\Delta_{\hat{\omega}} (r\Delta \log \mu(\hat{\omega}) + o(\Delta)), $$

where $w^\Delta_{\hat{\omega}}(\cdot)$ is the pointwise supremum over all convex functions that lie below $\varepsilon \mapsto \inf \{ g_1(\hat{\omega}, \alpha_2; \omega) : \alpha_2 \in B^\Delta_{2,\varepsilon}(\hat{\omega}) \}$, and $B^\Delta_{2,\varepsilon}(\hat{\omega})$ is the set of all undominated mixed actions $\alpha_2 \in \Delta(A_2)$ such that $\alpha_2 \in \arg\max_{\alpha'_2} \{ g_2(\alpha_1, \alpha'_2) \}$ for some mixed action $\alpha_1 \in \Delta(A_1)$ satisfying $H(q^\Delta(\cdot|\hat{\omega}, \alpha_2); q^\Delta(\cdot|\alpha_1, \alpha_2)) \leq \varepsilon$. Since for all $\hat{\omega}, \alpha_1$ and $\alpha_2$,

$$H(q^\Delta(\cdot|\hat{\omega}, \alpha_2); q^\Delta(\cdot|\alpha_1, \alpha_2)) \geq h(\hat{\omega}, \alpha_1; \alpha_2) \Delta + o(\Delta),$$

we have

$$B^\Delta_{2,1 - e^{-r\Delta}} \log \mu(\hat{\omega})(\hat{\omega}) \subseteq B^*_{2,r \log \mu(\hat{\omega}) + o(1)}(\hat{\omega}),$$

and hence,

$$\sup_{\hat{\omega} \in \Omega^b} w^\Delta_{\hat{\omega}} ((1 - e^{-r\Delta}) \log \mu(\hat{\omega})) \geq \sup_{\hat{\omega} \in \Omega^b} w^*_{\hat{\omega}} (r \log \mu(\hat{\omega}) + o(1)),$$

where $o(1)$ is a function that converges to zero as $\Delta \to 0$, and $w^*_{\hat{\omega}}(\cdot)$ is the pointwise supremum over all convex functions that are below $\varepsilon \mapsto \inf \{ u_1(\hat{\omega}, \alpha_2; \omega) : \alpha_2 \in B^*_{2,\varepsilon}(\hat{\omega}) \}$. The result then follows from the lower semi-continuity of the function $h$, by taking the limit as $\Delta \to 0$ and $r \to 0$. The proof for the upper bound is analogous, and hence omitted. Also omitted is the proof that if the game is non-degenerate and the action of the long-run player is information-flow identified then $g^*_1(\omega) = \tilde{g}^*_1(\omega)$ for every strategic type $\omega$, as this is similar to the proof of Theorem 3.3 of Fudenberg and Levine (1992).

---

See Gossner (2011, Lemma 8) for details of a similar continuity argument.
A.2 Proof for Example 1

Fix $\alpha_1, \alpha'_1 \in \Delta(A_1)$ and $\alpha_2 \in \Delta(A_2)$ and write $\alpha = (\alpha_1, \alpha_2)$, $\alpha' = (\alpha'_1, \alpha_2)$. With slight abuse of notation, for each $a \in A$ we also write $\alpha(a)$ (resp. $\alpha'(a)$) to denote the probability with which the pure-action profile $a$ is played under $\alpha$ (resp. $\alpha'$). We have:

$$H(q^\Delta(\cdot|\alpha) \parallel q^\Delta(\cdot|\alpha'))$$

$$= \sum_{\alpha} \alpha(\tilde{a}) \exp(-\lambda(\tilde{a}) \Delta) \sum_{k=0}^{\infty} \frac{\lambda(\tilde{a})^k \Delta^k}{k!} \log \frac{\sum_a \alpha(a) \lambda(a)^k \exp(-\lambda(a) \Delta)}{\sum_{\alpha'} \alpha'(a') \lambda(a')^k \exp(-\lambda(a') \Delta)}$$

$$= \sum_{\alpha} \alpha(\tilde{a}) \exp(-\lambda(\tilde{a}) \Delta) \left( \log \frac{\sum_a \alpha(a) \exp(-\lambda(a) \Delta)}{\sum_{\alpha'} \alpha'(a') \exp(-\lambda(a') \Delta)} + \Delta \lambda(\tilde{a}) \log \frac{\sum_a \alpha(a) \lambda(a) \exp(-\lambda(a) \Delta)}{\sum_{\alpha'} \alpha'(a') \lambda(a') \exp(-\lambda(a') \Delta)} + o(\Delta) \right),$$

and hence,

$$h(\alpha_1, \alpha'_1; \alpha_2) = \sum_{\alpha} \alpha(\tilde{a}) \left( \lim_{\Delta \to 0} \frac{1}{\Delta} \log \frac{\sum_a \alpha(a) \exp(-\lambda(a) \Delta)}{\sum_{\alpha'} \alpha'(a') \exp(-\lambda(a') \Delta)} + \lambda(\tilde{a}) \log \frac{\lambda(a)}{\lambda(a')} \right)$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta} \log \frac{\sum_a \alpha(a) \exp(-\lambda(a) \Delta)}{\sum_{\alpha'} \alpha'(a') \exp(-\lambda(a') \Delta)} + \lambda(a) \log \frac{\lambda(a)}{\lambda(a')}$$

where, by L’Hospital rule,

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \log \frac{\sum_a \alpha(a) \exp(-\lambda(a) \Delta)}{\sum_{\alpha'} \alpha'(a') \exp(-\lambda(a') \Delta)} = \frac{1}{\sum_a \alpha(a) \exp(-\lambda(a) \Delta) \sum_{\alpha'} \alpha'(a') \exp(-\lambda(a') \Delta)}$$

$$\times \left[ - \sum_a \alpha(a) \lambda(a) \exp(-\lambda(a) \Delta) \sum_{\alpha'} \alpha'(a') \exp(-\lambda(a') \Delta) + \sum_{\alpha'} \alpha'(a') \lambda(a') \exp(-\lambda(a') \Delta) \sum_a \alpha(a) \exp(-\lambda(a) \Delta) \right] = \lambda(a') - \lambda(a),$$

which is the desired result.

A.3 Proof for Example 2

Fix $\alpha_1, \alpha'_1 \in \Delta(A_1)$ and $\alpha_2 \in \Delta(A_2)$ and write $\alpha = (\alpha_1, \alpha_2)$, $\alpha' = (\alpha'_1, \alpha_2)$. With slight abuse of notation, for each $a \in A$ we also write $\alpha(a)$ (resp. $\alpha'(a)$) to denote the probability with which the pure-action profile $a$ is played under $\alpha$ (resp. $\alpha'$). We have:

$$H \left( q^\Delta(\cdot|\alpha) \parallel q^\Delta(\cdot|\alpha') \right) = \int_{-\infty}^{\infty} \log \frac{f^\Delta(y|\alpha)}{f^\Delta(y|\alpha')} f^\Delta(y|\alpha) \, dy,$$
where, for each $a \in A$, $f^\Delta(\cdot|a)$ denotes the density function of a Gaussian random variable with mean $\mu(a)\Delta$ and variance $\Delta$. For any $\tilde{a} \in A$,

$$
\frac{f^\Delta(y|a)}{f^\Delta(y|a')} = \frac{\sum_a \alpha(a) \exp \left[ - (y - \mu(\tilde{a})\Delta + (\mu(\tilde{a}) - \mu(a))\Delta)^2/(2\Delta) \right]}{\sum_{a'} \alpha'(a') \exp \left[ - (y - \mu(\tilde{a})\Delta + (\mu(\tilde{a}) - \mu(a'))\Delta)^2/(2\Delta) \right]} = \frac{\sum_a \alpha(a) \exp \left[ (\mu(a) - \mu(\tilde{a}))(y - \mu(\tilde{a})\Delta) - (\mu(a) - \mu(\tilde{a}))^2\Delta/2 \right]}{\sum_{a'} \alpha'(a') \exp \left[ (\mu(a') - \mu(\tilde{a}))(y - \mu(\tilde{a})\Delta) - (\mu(a') - \mu(\tilde{a}))^2\Delta/2 \right]} = \frac{\sum_a \alpha(a) \exp \left[ \sqrt{\Delta}s(a, \tilde{a})(y - \mu(\tilde{a})\Delta)/\sqrt{\Delta} - \Delta s(a, \tilde{a})^2/2 \right]}{\sum_{a'} \alpha'(a') \exp \left[ \sqrt{\Delta}s(a', \tilde{a})(y - \mu(\tilde{a})\Delta)/\sqrt{\Delta} - \Delta s(a', \tilde{a})^2/2 \right]},
$$

where $s(a, a') := \mu(a) - \mu(a')$ for all $a, a' \in A$. Therefore,

$$H \left( q^\Delta(\cdot|a) \| q^\Delta(\cdot|a') \right)
= \sum_{\tilde{a}} \alpha(\tilde{a}) \int_{-\infty}^{\infty} \log \left[ \sum_a \alpha(a) \exp \left[ \sqrt{\Delta}s(a, \tilde{a})(y - \mu(\tilde{a})\Delta)/\sqrt{\Delta} - \Delta s(a, \tilde{a})^2/2 \right] \right] f^\Delta(y|\tilde{a}) \, dy
- \sum_{\tilde{a}} \alpha(\tilde{a}) \int_{-\infty}^{\infty} \log \left[ \sum_{a'} \alpha'(a') \exp \left[ \sqrt{\Delta}s(a', \tilde{a})(y - \mu(\tilde{a})\Delta)/\sqrt{\Delta} - \Delta s(a', \tilde{a})^2/2 \right] \right] f^\Delta(y|\tilde{a}) \, dy,
$$

hence the change of variables $z = (y - \mu(\tilde{a})\Delta)/\sqrt{\Delta}$ yields

$$
(7) \quad H \left( q^\Delta(\cdot|a) \| q^\Delta(\cdot|a') \right)
= \sum_{\tilde{a}} \alpha(\tilde{a}) \int_{-\infty}^{\infty} \log \left[ \sum_a \alpha(a) \exp \left[ \sqrt{\Delta}s(a, \tilde{a})z - \Delta s(a, \tilde{a})^2/2 \right] \right] f(z) \, dz
- \sum_{\tilde{a}} \alpha(\tilde{a}) \int_{-\infty}^{\infty} \log \left[ \sum_{a'} \alpha'(a') \exp \left[ \sqrt{\Delta}s(a', \tilde{a})z - \Delta s(a', \tilde{a})^2/2 \right] \right] f(z) \, dz,
$$

where $f$ denotes the density of a Gaussian random variable with mean 0 and variance 1.

To get a tractable expression for the integral

$$I_{a, \tilde{a}}^\Delta := \int_{-\infty}^{\infty} \log \left[ \sum_a \alpha(a) \exp \left[ \sqrt{\Delta}s(a, \tilde{a})z - \Delta s(a, \tilde{a})^2/2 \right] \right] f(z) \, dz,$$

consider a second-order expansion of the integrand

$$g_{a, \tilde{a}}^\Delta(z) := \log \left[ \sum_a \alpha(a) \exp \left[ \sqrt{\Delta}s(a, \tilde{a})z - \Delta s(a, \tilde{a})^2/2 \right] \right]$$

around $z = 0$. Using Taylor’s formula with Lagrange’s remainder and the fact that $z$ has zero mean and unit variance under $f$, we get

$$
I_{a, \tilde{a}}^\Delta = g_{a, \tilde{a}}^\Delta(0) + \frac{1}{2} (g_{a, \tilde{a}}^\Delta)'(0) + \frac{1}{6} \int_{-\infty}^{\infty} (g_{a, \tilde{a}}^\Delta)''(0) \left( \theta_{a, \tilde{a}}^\Delta(z) z^3 \right) f(z) \, dz.
$$
Finally, differentiating \( g^\Delta_{a,\bar{a}} \) twice and evaluating at \( z = 0 \) gives

\[
(g^\Delta_{a,\bar{a}})''(0) = \left[ \frac{\sum_a \alpha(a)s(a, \bar{a})^2 \exp(-\Delta s(a, \bar{a})^2/2)}{\sum_a \alpha(a) \exp(-\Delta s(a, \bar{a})^2/2)} - \left( \frac{\sum_a \alpha(a)s(a, \bar{a}) \exp(\Delta s(a, \bar{a})^2/2)}{\sum_a \alpha(a) \exp(-\Delta s(a, \bar{a})^2/2)} \right)^2 \right] \Delta,
\]

and thus,

\[
\lim_{\Delta \to 0} \frac{(g^\Delta_{a,\bar{a}})''(0)}{\Delta} = \sum_a \alpha(a)s(a, \bar{a})^2 - \left( \sum_a \alpha(a)s(a, \bar{a}) \right)^2.
\]

Finally, differentiating \( g^\Delta_{a,\bar{a}} \) three times yields

\[
(g^\Delta_{a,\bar{a}})'''(z) = \Delta^{3/2} \times \frac{\sum_a \sum_{a'} \sum_{a''} \sum_{a'''} \Phi(a, a', a'', a''') \exp \left( \sqrt{\Delta} \gamma(a, a', a'', a''') z - \frac{\Delta}{2} \delta(a, a', a'', a''') \right)}{\sum_a \sum_{a'} \sum_{a''} \sum_{a'''} \psi(a, a', a'', a''') \exp \left( \sqrt{\Delta} \gamma(a, a', a'', a''') z - \frac{\Delta}{2} \delta(a, a', a'', a''') \right)},
\]

where

\[
\gamma(a, a', a'', a''') := s(a, \bar{a}) + s(a', \bar{a}) + s(a'', \bar{a}) + s(a''', \bar{a}),
\]

\[
\delta(a, a', a'', a''') := s(a, \bar{a})^2 + s(a', \bar{a})^2 + s(a'', \bar{a})^2 + s(a''', \bar{a})^2,
\]

\[
\Phi(a, a', a'', a''') := \alpha(a)\alpha(a')\alpha(a'')\alpha(a''') \left[ s(a, \bar{a})^3 - s(a', \bar{a})s(a'', \bar{a})^2 - s(a, \bar{a})s(a'', \bar{a})^2 + s(a', \bar{a})s(a'', \bar{a})s(a''', \bar{a}) \right],
\]

\[
\psi(a, a', a'', a''') := \alpha(a)\alpha(a')\alpha(a'')\alpha(a''').
\]

This expression is bounded in \( z \): as \( z \to \pm \infty \) the numerator and denominator both grow unbounded, or decay to zero, at the same rate. In particular,

\[
(g^\Delta_{a,\bar{a}})'''(\theta^\Delta(z)) = O(\Delta^{3/2}) \text{ uniformly in } z.
\]

---

19The measurability of \( \theta^\Delta \) follows from the upper hemi-continuity of the correspondence that maps each \( z \) to the set of all \( y \in [0, 1] \) satisfying \( g(z) = g(0) + g'(0)z + \frac{1}{2} g''(0)z^2 + \frac{1}{6} g'''(y)z^3 \), and the fact that nonempty-valued upper hemi-continuous correspondences admit measurable selections.
which implies

\[ \frac{1}{\Delta} \int_{-\infty}^{\infty} (g_{a,\hat{a}}^{\Delta})''(\theta_{a,\hat{a}}^{\Delta}(z)) z^3 f(z) \, dz = O(\sqrt{\Delta}) \]

since \( \int_{-\infty}^{\infty} |z|^3 f(z) \, dz < \infty \).

By Lebesgue’s dominated convergence, it follows from (7)–(11) that

\[
\lim_{\Delta \to 0} \frac{\sum_{a} \alpha(\hat{a})(I_{a,\hat{a}}^{\Delta} - I_{a',\hat{a}}^{\Delta})}{\Delta} = \frac{1}{2} \sum_{a} \alpha(\hat{a}) \left[ (\sum_{a} \alpha'(a)s(a, \hat{a}))^2 - (\sum_{a} \alpha(a)s(a, \hat{a}))^2 \right] \\
= \frac{1}{2} \sum_{a} \alpha(\hat{a}) [ (\mu(a') - \mu(\hat{a}))^2 - (\mu(\alpha) - \mu(\hat{a}))^2 ] \\
= \frac{1}{2} \sum_{a} \alpha(\hat{a})(\mu(a') - \mu(\alpha))(\mu(a') + \mu(\alpha) - 2\mu(\hat{a})) \\
= \frac{1}{2}(\mu(a') - \mu(\alpha))^2, \\
\]

as was to be shown.

A.4 Proof of Proposition 1

Lemma 1. Let \( A_1, A_2, \ldots \) be a sequence of events on a probability space \( (\Omega, \mathcal{F}, P) \) satisfying \( \lim \inf_{n \geq 1} P(A_n) > c > 0 \). Then,

\[
\lim \inf_{N \to \infty} P \left( \frac{\# \{ n \leq N : A_n \text{ obtains} \}}{N} > \frac{c}{2} \right) > \frac{c}{2}.
\]

Proof. For any event \( E \) we denote by \( 1_E \) the indicator function of \( E \). Let \( N_0 \) be such that \( P(A_n) > c \) for all \( n > N_0 \). By Markov’s inequality, for all \( N > N_0 \),

\[
P \left( \frac{1}{N} \sum_{n=1}^{N} 1_{A_n} > \frac{c}{2} \right) = 1 - P \left( \frac{1}{N} \sum_{n=1}^{N} 1_{A_n^c} \geq 1 - \frac{c}{2} \right) \\
\geq 1 - \frac{2}{(2-c)} \frac{1}{N} \sum_{n=1}^{N} P(A_n^c) \\
> 1 - \frac{2}{(2-c)} \left( \frac{N_0}{N} + \frac{(N-N_0)(1-c)}{N} \right),
\]

and therefore,

\[
\lim \inf_{N \to \infty} P \left( \frac{1}{N} \sum_{n=1}^{N} 1_{A_n} > \frac{c}{2} \right) \geq 1 - \frac{2}{2-c} (1-c) > \frac{c}{2}.
\]

\( \square \)
Lemma 2. If \( p_k = \frac{1}{2}(1 + 1/k^\alpha) \) with \( \alpha > 1/2 \), then

\[
\sum_{k=1}^{\infty} \log\left(\frac{1}{4p_k(1-p_k)}\right) < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \left(\log\left(\frac{p_k}{1-p_k}\right)\right)^2 < \infty .
\]

Proof. Follows directly from the definition of \( p_k \), the inequality \( \log(1+x) < x \) for \( x > 0 \), and the fact that \( \sum_{k=1}^{\infty} k^{-v} < \infty \) for all \( v > 1 \). \( \square \)

Lemma 3. Let \( P, Q \) and \( Q' \) be probability measures on a given measurable space \((\Omega, \mathcal{F},(\mathcal{F}_n))\) such that \( P = \phi_0 Q + (1 - \phi_0) Q' \) for some \( \phi_0 \in (0, 1] \). Denote by \( \phi_n \) be the posterior probability given \( \mathcal{F}_n \) on the event that “the true underlying measure is \( Q' \”), i.e.

\[
\phi_n = \phi_0 \frac{dQ^n}{dP^n},
\]

where \( Q^n \) and \( P^n \) designate the restrictions of \( Q \) and \( P \) to \( \mathcal{F}_n \), respectively. Then, for all \( \eta \in (0, 1] \),

\[
Q(\phi_n \leq \eta \phi_0) \leq \eta .
\]

Proof. Since the odds ratio \((1 - \phi_n)/\phi_n\) is a martingale under \( Q \), we have:

\[
Q(\phi_n \leq \eta \phi_0) = Q\left(\frac{1 - \phi_n}{\phi_n} \geq \frac{1 - \eta \phi_0}{\eta \phi_0}\right) \\
\leq E_Q[(1 - \phi_n)/\phi_n] \\
= (1 - \phi_0)/\phi_0 \\
= (1 - \eta \phi_0)/(\eta \phi_0) \leq \eta,
\]

as was to be shown. \( \square \)

Lemma 4. Let \( (\mathcal{F}_n)_{n \geq 0} \) be a strictly increasing sequence of sub-\( \sigma \)-algebras on a measurable space \((\Omega, \mathcal F)\). For all \( \phi_0 \in (0, 1] \) and \( \nu > 1/2 \) there exist probability measures \( P, Q \) and \( Q' \) on \((\Omega, \mathcal{F})\) with \( P = \phi_0 Q + (1 - \phi_0) Q' \) such that for all \( C > 0 \),

\[
\liminf_{N \to \infty} Q \left( \#\{n \geq 0 : d_n(P, Q) \geq C/N^\nu \} \geq N \right) \geq 1/8.
\]

Proof. First, we prove the result for the case in which the measurable space is the set of all infinite sequences of Heads and Tails and the \( \sigma \)-algebra \( \mathcal{F}_n \) is the one generated by the histories of length \( n \). Let \( Q \) be the probability measure corresponding to i.i.d. draws of an unbiased coin. Let \( Q' \) be the probability corresponding to independent draws of a coin such that the probability of Heads in period \( n \) is \( p_n = 1/2(1 + 1/n^\alpha) \), where \( 1/2 < \alpha < \nu \). Let \( P = \phi_0 Q + (1 - \phi_0) Q' \) and \( \phi_n \) be the posterior probability that the stochastic process follows \( Q \) conditional on period \( n \) information.

Denote by \( \chi_n \) the \( \{0, 1\} \)-valued random variable that takes value 1 if and only if the coin turns up Heads at time \( n \). Noting that \( d_n(P, Q) = (p_n - 1/2)(1 - \phi_n) \), explicit calculation of Bayes rule
yields for any $C > 0$:

\[
(12) \quad Q \left( \frac{d_n(P, Q)}{\phi_n} \geq \frac{4C/\phi_0}{n^v} \right) = Q \left( \frac{1 - \phi_n}{\phi_n} \geq \frac{8C/\phi_0}{n^{v-\alpha}} \right) = Q \left( \prod_{k=1}^{n} (2p_k)^{\chi_k} (2(1 - p_k))^{(1-\chi_k)} \geq \frac{8C/(1 - \phi_0)}{n^{v-\alpha}} \right)
\]

\[
= Q \left( \sum_{k=1}^{n} \chi_k \log(2p_k) + \sum_{k=1}^{n} (1 - \chi_k) \log(2(1 - p_k)) \geq \log(8C/(1 - \phi_0)) - (v - \alpha) \log n \right) = Q \left( \sum_{k=1}^{n} \log \left( \frac{p_k}{1 - p_k} \right) \chi_k \geq \sum_{k=1}^{n} \log \left( \frac{1}{2(1 - p_k)} \right) + B - \gamma \log n \right)
\]

\[
= Q \left( \sum_{k=1}^{n} \log \left( \frac{p_k}{1 - p_k} \right) (\chi_k - \frac{1}{2}) \geq \frac{1}{2} \sum_{k=1}^{n} \log \left( \frac{1}{4p_k(1 - p_k)} \right) + B - \gamma \log n \right)
\]

\[
\geq Q \left( \left| \sum_{k=1}^{n} \log \left( \frac{p_k}{1 - p_k} \right) (\chi_k - \frac{1}{2}) \right| \leq \gamma \log n - \frac{1}{2} \sum_{k=1}^{n} \log \left( \frac{1}{4p_k(1 - p_k)} \right) - B \right)
\]

By Chebyshev’s inequality, for any $M > 0$,

\[
Q \left( \left| \sum_{k=1}^{n} \log \left( \frac{p_k}{1 - p_k} \right) (\chi_k - \frac{1}{2}) \right| \geq M \right) \leq \frac{\sum_{k=1}^{\infty} \left( \log \left( \frac{p_k}{1 - p_k} \right) \right)^2}{4M^2}.
\]

By Lemma 2, $\sum_{k=1}^{\infty} \left( \log \left( \frac{p_k}{1 - p_k} \right) \right)^2 < \infty$, hence we can choose $M > 0$ large enough that

\[
(13) \quad Q \left( \left| \sum_{k=1}^{n} \log \left( \frac{p_k}{1 - p_k} \right) (\chi_k - \frac{1}{2}) \right| \geq M \right) \leq \frac{1}{2}.
\]

On the other hand, by Lemma 2, $\sum_{k=1}^{\infty} \log \left( \frac{1}{4p_k(1 - p_k)} \right) < \infty$. Therefore, for all $n$ large enough,

\[
\gamma \log n - \frac{1}{2} \sum_{k=1}^{n} \log \left( \frac{1}{4p_k(1 - p_k)} \right) - B > M.
\]

Therefore, by the inequalities (12) and (13) above, for all $n$ large enough,

\[
Q \left( \frac{d_n(P, Q)}{\phi_n} \geq \frac{4C/\phi_0}{n^v} \right) \geq \frac{1}{2}
\]

By Lemma 3, $Q \left( \phi_n \leq \frac{1}{4} \phi_0 \right) \leq \frac{1}{4}$. Thus, for $n$ large enough,

\[
Q \left( \frac{d_n(P, Q)}{\phi_n} \geq \frac{C}{n^v} \right) + \frac{1}{4} \geq Q \left( \frac{d_n(P, Q)}{\phi_n} \geq \frac{C}{n^v} \right) + Q \left( \phi_n \leq \frac{1}{4} \phi_0 \right) \geq \frac{1}{2}.
\]
We have thus proved that
\[ \liminf_{n \to \infty} Q \left( d_n(P, Q) \geq C / n^\eta \right) \geq \frac{1}{4}. \]
Therefore, the events \( \{d_n(P, Q) \geq C / n^\eta\} \) satisfy the assumption of Lemma 1, and so
\[ \liminf_{N \to \infty} Q \left( \frac{\#\{n \leq N : d_n(P, Q) \geq C / n^\eta\}}{N} \geq \frac{1}{8} \right) \geq \frac{1}{8}. \]
Since
\[ Q \left( \frac{\#\{n \leq N : d_n(P, Q) \geq C / n^\eta\}}{N} \geq \frac{1}{8} \right) \geq Q \left( \frac{\#\{n \leq N : d_n(P, Q) \geq C / n^\eta\}}{N} \geq \frac{1}{8} \right), \]
it follows that
\[ \liminf_{N \to \infty} Q \left( \#\{n \geq 0 : d_n(P, Q) \geq C / n^\eta\} \geq \frac{1}{8} N \right) \geq \frac{1}{8}, \]
which implies the desired result for the case in which \((\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0})\) is the canonical coin tossing space.

If \((\Omega, \mathcal{F})\) is an arbitrary space with a strictly increasing filtration \((\mathcal{F}_n)\), then, for every \(n\), there are at least two disjoint events \(A_n, B_n\) such that \(A_n, B_n \in \mathcal{F}_n \setminus \mathcal{F}_{n-1}\). Proceed as in the previous construction, identifying the event \(A_n\) with “Heads” and \(B_n\) with “Tails”.

With the lemmas above in place we are now ready to prove Proposition 1. First, a close look at S. Sorin’s proof of the Fudenberg-Levine uniform merging lemma (Sorin, 1999, Lemma 2.5) reveals that the integer \(K\) of inequality (2) can be taken as \(C = \epsilon^2\) for some constant \(C > 0\) that depends only on \(\eta\) and \(\phi\). Hence, \(\limsup_{\epsilon \to 0} \epsilon^2 K^* (\epsilon, \eta, \phi) < \infty\), which is the first claim.

Turning to the second claim assume that the filtration is strictly increasing. Suppose, towards a contradiction, that for some \(q > 0\) and \(\phi_0 \in (0, 1]\) there exist a function \((\epsilon, \eta) \mapsto N(\epsilon, \eta)\) and a sequence \(\eta_k \to 0\) with
\[ \limsup_{\epsilon \to 0} \epsilon^{2-q} N(\epsilon, \eta_k) < \infty, \]
and such that for all \(P, Q\) and \(Q'\) with \(P = \phi_0 Q + (1 - \phi_0) Q'\),
\[ Q \left( \#\{n \geq 0 : d_n(P, Q) \geq \epsilon \geq N(\epsilon, \eta_k) \} \leq \eta_k \right) \]
for all \(k \geq 1\) and for all \(\epsilon > 0\) sufficiently small. Let \(M_k = 1 + \limsup_{\epsilon \to 0} \epsilon^{2-q} N(\epsilon, \eta_k)\) and define a double sequence \(\epsilon_{m,k} = (M_k / m)^1/(2-q)\), for \(k, m \geq 1\). Hence, for all \(k \geq 1\),
\[ \limsup_{m \to \infty} N(\epsilon_{m,k}, \eta_k) / m < 1. \]
By Lemma 4, there exist probability measures \(P, Q\) and \(Q'\), with \(P = \phi_0 Q + (1 - \phi_0) Q'\), such that for all \(k \geq 1\),
\[ \liminf_{m \to \infty} Q \left( \#\{n \geq 0 : d_n(P, Q) \geq \epsilon_{m,k} \} \geq N(\epsilon_{m,k}, \eta_k) \right) \geq 1/8. \]
Since $\eta_k < 1/8$ for all $k$ large enough, it follows that for all $\delta > 0$ and all $k$ sufficiently large there exists $\varepsilon \in (0, \delta]$ such that:

$$Q (\#\{n \geq 0 : d_n(P, Q) \geq \varepsilon\} \geq N(\varepsilon, \eta_k)) > \eta_k,$$

which is a contradiction. The contradiction demonstrates that

$$\lim_{\varepsilon \to 0} \epsilon^{2-d} K^* (\varepsilon, \eta, \phi) = \infty$$

for all $q > 0$, $\phi_0 \in (0, 1]$ and $\eta$ sufficiently small, as was to be shown.

**References**


