Building a Reputation under Frequent Decisions*

Eduardo Faingold †

July 2008

Abstract

I study reputation games with frequent decisions and *persistently imperfect monitoring*. In these games, as the period length tends to zero, the monitoring structure approaches a *continuous-time limit*, and, further, the limit monitoring is non-trivially imperfect. I show that for any prior probability on the long-run player's types that assigns positive probability to *commitment types*, if the long-run player's *instantaneous discount rate* is close enough to zero, there is a lower bound on the set of Nash equilibrium payoffs of the long-run player which converges, as the period length tends to zero, to a number arbitrarily close to the long-run player's *commitment payoff*.

1 Introduction

A common approach to reputation phenomena involves the study of *repeated games with incomplete information*, that is, repeated games in which, prior to the beginning of the game, some players possess private information about some payoff-relevant characteristics, their *types*. Intertemporal incentives are created by the possibility that a player of one type can mimic the behavior of another type. Of particular interest are *commitment types*. These are types that play a certain strategy in a *mechanistic way*, like "irrational" automata. The *normal types*, on the other hand, are types that behave strategically. If the prior probability on commitment types is high enough, then the equilibrium behavior of the normal types might well be very different from what it would be were commitment types given zero prior. Remarkably, as shown by Kreps, Milgrom, Roberts, and Wilson (1982) and Fudenberg and Levine (1992), even an arbitrarily small prior probability on the commitment types can exert a great impact on the behavior of the normal types. This phenomenon has come to be known as the "reputation effect."

These reputation effects, however, are operative only when the player with reputational concerns is *patient*. A common interpretation of patience in discounted games is that the interaction takes place in *continuous time* and the *length of the period is short*, that is, players are capable of adjusting their actions frequently. However, in *imperfect monitoring games* with noisy signals, as the period length shrinks to zero, the number of signals observed in any given interval of *real time* increases without bound. If signals are statistically informative, as the period length tends to zero, the degree of monitoring imperfection

*I am very grateful to George Mailath for his invaluable advice and encouragement. I also thank Dirk Bergemann, Eddie Dekel, Jan Eeckhout, Jeffrey Ely, Drew Fudenberg, John Geanakoplos, Ehud Kalai, David K. Levine, Steven Matthews, Roger Myerson, Nabil al-Najjar, Philip Reny, Rafael Rob, Larry Samuelson, Yuliy Sannikov, Marco Scarsini and William Zame for insightful comments and helpful conversations.

†Department of Economics, Yale University, 30 Hillhouse Ave., New Haven, CT, 06511, USA. E-mail: eduardo.faingold@yale.edu

becomes asymptotically negligible. The comparative statics, thus, is in tension with the very spirit of the theory of repeated games with imperfect monitoring.

I study reputation phenomena in games with frequent decisions and *persistently imperfect monitoring*. In these games, as the period length tends to zero, the monitoring structure approaches a *continuous-time limit*, and, further, the limit monitoring is non-trivially imperfect (i.e. has full support). Unlike standard imperfect monitoring games, repeated games with persistently imperfect monitoring allow a meaningful comparative statics with the period length, accommodating the continuous-time interpretation of discounting with the requirement that the overall informativeness of the monitoring be bounded.

An early example of persistently imperfect monitoring games was studied by Abreu, Milgrom, and Pearce (1991) in their paper on collusion in partnership games. I discuss a related example in Section 2. In that paper, they show that increasing the players' instantaneous discount rate and shortening the period length have markedly different effects on the structure of intertemporal incentives. In their setting, decreasing the players' instantaneous discount rate expands the scope of collusion, whereas decreasing the period length has precisely the opposite effect. This is due to an adverse *informational effect*: in games with persistently imperfect monitoring, over shorter and shorter periods, the monitoring technology becomes increasingly noisy. Indeed, when this effect is too strong, it may lead to a complete collapse of intertemporal incentives.

I consider games in which a *long-run player* faces a sequence of *short-run players*. The long-run player can be one of many types and the short-run players are uncertain as to which type of long-run player they face. I show that powerful reputation effects akin to those found in Fudenberg and Levine (1992) emerge. Specifically, for any prior probability on the long-run player's types which assigns positive probability to *commitment types*, if the long-run player's *instantaneous discount rate* is close enough to zero, there is a lower bound on the set of Nash equilibrium payoffs of the long-run player which converges, as the period length tends to zero, to a number arbitrarily close to the long-run player's *commitment payoff*. The latter is the payoff the long-run player would receive if he could credibly commit to play any strategy in the support of the short-run players' prior.

The significance of this result is that it establishes that reputation phenomena satisfy a *robustness property*, namely that the adverse informational effect of shortening the period length is not strong enough to undermine the power of reputations. While Fudenberg and Levine's (1992) analysis yields the result for *fixed period length*, it leaves open the possibility that the required bound on the firm's instantaneous discount rate becomes increasingly stringent as the period length shrinks to zero. However, it is an implication of the analysis in this paper that the power of reputation effects depends on the long-run player's patience (discount rate), but *not* on the frequency with which he moves.

To put the result in context, consider the underlying *complete information game*, that is, the game in which the long-run player' type is common knowledge. In Faingold and Sannikov (2005), we show that if the noise in the monitoring technology is driven by a Brownian motion, there is a full *collapse of intertemporal incentives* : Irrespective of the long-run players' discount rate, if players move frequently, the set of Nash equilibrium payoffs of the long-run player coincides with (the convex hull of) the set of *static* Nash equilibrium payoffs.[1] Hence, for the monitoring technology studied by Faingold and Sannikov (2005), the main result of the current paper means that *reputation effects are the only mechanism*

[1] This result is close in spirit to Abreu, Milgrom and Pearce (1991), but unlike in the latter paper, the collapse of incentives is not driven by a failure of statistical identification. What causes the collapse here is the combination of Brownian motion noise *and* the presence of short-run players. See Section 3.5.

through which non-trivial intertemporal incentives can be created.

Games with persistently imperfect monitoring offer an attractive model to study the impact of increasing the frequency of decisions. First, the setting allows a direct comparison with discrete-time models. As the period length tends to zero, the signal structure converges to a (controlled) Lévy process, that is, a continuous-time controlled Markov process, taking values in an Euclidean space, such that the increments of the process are stationary and independent whenever players use stationary strategies. This class of monitoring structures includes the Poisson model studied by Abreu, Milgrom, and Pearce (1991), as well as the Brownian motion setting examined by Sannikov (2005), but is not exhausted by these two examples. In fact, a general Lévy process has three components: a deterministic drift, a continuous term driven by a Brownian motion and a jump term.

Second, unlike standard repeated games with imperfect monitoring, in games with persistently imperfect monitoring *arbitrarily frequent decisions are not arbitrarily informative*: Even in the limit game in which players can adjust their actions at every instant of time, the monitoring structure will have "full support." More precisely, in the limit game, the controlled Lévy process that characterizes the monitoring structure will have the following property: For any pair of strategy profiles that differ only in the behavior of the long-run player, the induced probability measures over finite-horizon signal paths are *mutually absolutely continuous.* Thus, in the limit game, the long-run player has no effect over the "support" of the signal distributions. In particular, the *real time* that it takes for the short-run players' Bayesian learning process to settle around the long-run player's true behavior is bounded away from zero uniformly as the period length shrinks. In contrast, in standard repeated games, if the signal distributions satisfy a statistical identification condition, as the period length tends to zero, the monitoring structure approaches the perfectly informative one, so the learning of the short-run players settles around the truth in ever shorter intervals of real time.

Turning to the logic behind the main result of the paper, note that for each *fixed* period length, the payoff bounds from reputation effects obtained by Fudenberg and Levine (1992) are valid. Thus, a natural approach to study the robustness of reputation effects with respect to shortening the period length is to examine the limit behavior of the Fudenberg-Levine bounds as the period length converges to zero. This requires the examination of the *speed of convergence* of the short-run players' Bayesian learning. Indeed, I will argue (Section 3.4), based on sharp estimates of these convergence rates, that if information arrives in a continuous fashion and the noise in the monitoring technology is driven by a Brownian motion, then the robustness of reputation effects with respect to the comparative statics with the period length follows directly from the Fudenberg-Levine bounds. Surprisingly, when information arrives discontinuously, such as in the Poisson case studied by Abreu, Milgrom and Pearce (1991), I find that the Fudenberg-Levine reputation bounds are not sharp enough to yield the robustness result.

To obtain the payoff bounds for the *complete class* of games with persistently imperfect monitoring, I study directly the continuous-time limit game. Accordingly, the first step of the analysis derives payoff bounds from reputation effects in the limit game in which players can adjust their actions continuously over time. Such payoff bounds can be viewed as "infinitesimal" versions of the Fudenberg-Levine bounds. Owing to the convergence of the monitoring structure to the continuous-time limit, the bounds of the limit game can be extended to all games with sufficiently short periods.

2 Example

The Quality Game is a classical example in the literature. In the variation I present here, the monitoring structure is *persistently imperfect*, in the sense that, as the period length shrinks to zero, the monitoring structure approaches a continuous-time limit and, further, the limit monitoring structure has full support.

A monopolist provides a service to a continuum of identical consumers. The service is offered in two levels, $\bar{s}$ and $\underline{s}$, which are price differentiated. At every time $t \in [0, \infty)$, each consumer chooses a service level $s_t \in \{\bar{s}, \underline{s}\}$ and pays a price $p(s_t)$ for it. Prices are exogenous and satisfy $p(\bar{s}) > p(\underline{s}) > 0$. Consumers are *small anonymous players*: the monopolist, at any time, observes the aggregate distribution over the consumers' purchases, but not the choices of individual consumers.

At random times, failures in the service provision occur. The *quality of the service* up to time t is measured by the number of failures, N_t, that occurred in the time interval $[0, t]$. As in Abreu, Milgrom, and Pearce (1991), the random process $(N_t)_{t \geqq 0}$ is a *counting process* whose sample paths are *publicly observable*. Each consumer's payoff flow is given by:

$$dU_t \;=\; (s_t - p(s_t))\, dt \;-\; c(s_t)\, dN_t \,, \tag{2.1}$$

where $c(s_t)$ is the instantaneous disutility of one failure in the service provision. The overall payoff of each consumer is the continuously discounted average of his payoff flow:[2]

$$r \int_0^\infty e^{-rt}\, dU_t \,, \tag{2.2}$$

where $r > 0$ is the *instantaneous discount rate*.

At each time $t \in [0, \infty)$, the monopolist makes a costly *investment in quality* $e_t \in \{\bar{e}, \underline{e}\}$, where $\bar{e} > \underline{e} > 0$. The firm's payoff flow is given by:

$$d\Pi_t = (\bar{p}(S_t) - e_t)\, dt \,,$$

where S_t is the mass of consumers who chose the high-level service at time t, and $\bar{p}(S_t) = S_t\, p(\bar{s}) + (1 - S_t)\, p(\underline{s})$ is the firm's aggregate revenues. The firm's overall payoff in the dynamic game is the continuously discounted average of its payoff flow,

$$r \int_0^\infty e^{-rt}\, d\Pi_t \,.$$

The monopolist's investment in quality is *unobservable* to the consumers, and the quality process $(N_t)_{t \geqq 0}$ is a *noisy signal* of the firm's investment. Specifically, the firm's investment level e_t determines the *arrival rate* of failures in the service provision, $\mu(e_t)$. The higher the investment level, the lower the failure arrival rate: $\mu(\underline{e}) > \mu(\bar{e}) > 0$.

Remark 2.1 (Myopic consumers)**.** Note that consumers are *myopic players*. In effect, the choices of an individual consumer have no influence on continuation play, for each consumer is a *small anonymous player*. Hence, given an arbitrary strategy profile, any best reply of an individual consumer must prescribe a myopically optimal action following every history not precluded by the consumer's previous moves. □

[2]The sample paths of the counting process $(N_t)_{t \geqq 0}$ are step functions, so the integral in (2.2) is simply $\int_0^\infty e^{-rt}(s_t - p(s_t))\, dt \;-\; \sum_{n=1}^\infty e^{-r\tilde{t}_n} c(s_{\tilde{t}_n})$, where $0 < \tilde{t}_1 < \tilde{t}_2 < \ldots < \tilde{t}_n < \ldots$ are the (random) failure times.

Turning attention to the consumers' expected payoffs, note that the expectation of the right-hand side of equation (2.1), conditional on e_t, is

$$(s_t - p(s_t) - c(s_t)\mu(e_t))\ dt\ ,$$

since the quality signal $(N_t)_{t \geqq 0}$ is a counting process with intensity $\mu(e_t)$. Therefore, the expected flow payoffs of the monopolist and the consumers are given by functions $\pi(\cdot)$ and $u(\cdot)$, respectively, where

$$\begin{aligned} \pi(e, s) &= p(s) - e\ , \text{ and} \\ u(e, s) &= s - p(s) - c(s)\mu(e)\ , \end{aligned}$$

for all $(e, s) \in \{\bar{e}, \underline{e}\} \times \{\bar{s}, \underline{s}\}$.

Let the parameter values be $\bar{e} = 1$, $\underline{e} = 0$, and

	$\bar{s}$	$\underline{s}$
s	$3(2\mu(\underline{e}) - \mu(\bar{e}))/(\mu(\underline{e}) - \mu(\bar{e}))$	$(3\mu(\underline{e}) - \mu(\bar{e}))/(\mu(\underline{e}) - \mu(\bar{e}))$
$c(s)$	$3/(\mu(\underline{e}) - \mu(\bar{e}))$	$1/(\mu(\underline{e}) - \mu(\bar{e}))$
$p(s)$	3	1

The expected flow-payoffs π and u, then, are given by the following matrix:

	$\bar{s}$	$\underline{s}$
$\bar{e}$	2, 3	0, 2
$\underline{e}$	3, 0	1, 1

Hence, in the static game, low investment is a strictly dominant strategy for the firm; and the firm is uniformly better off when the consumers choose the high service level. The consumers, on the other hand, prefer high service level only when they believe that the firm will make high investment with probability greater than 1/2. The unique Nash equilibrium of the static game is $(\underline{e}, \underline{s})$.

While the interaction takes place in continuous time, neither the monopolist nor the consumers are capable of adjusting their actions continuously. I assume that the players can adjust their investment and consumption decisions only at discrete random times,

$$0 = \tau_0 < \tau_1 < \ldots < \tau_n < \ldots\ ,$$

which are determined by a *Poisson clock*: the durations $\tau_{n+1} - \tau_n$ are independent and exponentially distributed random variables with mean $h > 0$. Within each semi-open interval $[\tau_n, \tau_{n+1})$, actions are held constant. The parameter h is the *average length of the inaction period*, or *period length*, for short.

In the game presented above, the information flow takes place in continuous time. However, since players make decisions at discrete times, the game has an equivalent *repeated-game representation*:

(i) The stage game, given by the payoff matrix above, is repeatedly played in stages $n = 0, 1, 2, \ldots$.

(ii) At the end of each stage n, the service quality, y^n, is realized and publicly announced to the players. The signal y^n is a Poisson *random variable* with parameter $\mu(e^n)h$, where e^n is the firm's *hidden investment* at stage n.[3] The aggregate distribution over the consumers' actions is public information, but not the actions of any individual consumer.

(iii) The players' overall payoffs in the repeated game are given by the average discounted sum of their stage-game payoffs, with discount factor given by:[4]

$$\delta = \frac{1}{1 + rh}\,.$$

Let the period length $h > 0$ be fixed. For all $r > 0$, the i.i.d. play of the stage-game Nash equilibrium $(\underline{e}, \underline{s})$ is, trivially, an equilibrium of the repeated game. Moreover, because consumers are small anonymous players, the highest Nash equilibrium payoff of the firm is *bounded below 2*, uniformly for all $r > 0$.[5]

On the other hand, if the monopolist could *commit* to high investment in every period, the consumers would choose the high service level in every period, yielding the monopolist a payoff of 2. Therefore, the monopolist is (strictly) better-off under this commitment than in *any* Nash equilibrium of the repeated game. The interest in understanding when such commitments are credible is exactly what motivates the study of *reputation effects*.

Suppose there is *incomplete information*, namely the firm can be of two *types*: either a *normal type*, with payoffs as described above, or a *commitment type*, an automaton which makes high investment in every period, irrespective of history. Consumers are uncertain as to which type of service provider they face and assign prior probability $\phi_0 \in (0, 1)$ to the commitment type.

For a fixed period length $h > 0$, as the discount rate $r \to 0$, the effective discount factor $\delta \to 1$. Hence, a straightforward application of Fudenberg and Levine's (1992) result on payoff bounds from reputation effects yields:[6]

Proposition 2.1. *Fix the period length $h > 0$ and the prior probability $\phi_0 \in (0, 1]$ on the commitment type. For all $\varepsilon > 0$, there exists $\bar{r} > 0$, such that for all $r \in (0, \bar{r}]$, in every Nash equilibrium of the incomplete information game with prior ϕ_0 and discount rate r, the normal type of monopolist receives a payoff greater than $2 - \varepsilon$.*

The basic reasoning behind the Fudenberg-Levine result is the following. Fix a Nash equilibrium of the incomplete information game and consider a deviation in which the normal type makes high investment in every period, thereby mimicking the commitment type's behavior. By the firm's incentive

[3] While in the repeated-game representation the consumers observe only the number of failures within a period, and not the times at which failures occurred, this entails no loss of generality, for the number of failures in a period is a *sufficient statistic* for the failure arrival rate μ.

[4] This expression gives the value of the discount factor δ such that, given any continuous-time payoff stream $(\pi(t))_{t \geqq 0}$, constant within the intervals $[\tau_n, \tau_{n+1})$, one has $\mathsf{E}\left[r \int_0^\infty e^{-rt}\pi(t)\,dt\right] = (1-\delta)\sum_{n=0}^\infty \delta^n \pi'_n$, where $\pi'_n = \pi(\tau_n)$ for all $n \geqq 0$, and the expectation is with respect to the distribution of $(\tau_n)_{n \geqq 0}$.

[5] Since the consumers are myopic (Remark 2.1), the game is equivalent, in the best-reply sense, to a game between a *long-run player*, the firm, and a sequence of *short-run players*, the "representative consumer". The fact that the firm's Nash equilibrium payoffs are bounded below 2, then, follows from the inefficiency of *perfect public equilibria* in *moral hazard mixing games* with short-run players (Fudenberg and Levine (1994), Theorem 6.1, part (iii)). In fact, because in our game the monitoring has a *product structure* and there are *no observable deviations*, the set of Nash equilibrium payoffs coincides with the set of perfect public equilibrium payoffs, and the aforementioned result yields the desired bound.

[6] This is Fudenberg and Levine's (1992) Corollary 3.2, adapted to the case in which not all commitment types are in the support of the short-run players' prior.

constraint, the payoff it receives under this deviation is a lower bound for the payoff it obtains under the equilibrium strategy. Since the consumers assign positive prior probability to the commitment *type*, they must also assign positive prior probability to the event that the commitment *action*, high investment, is played in every period. Since the signals are informative ($\mu(\bar{e}) \neq \mu(\underline{e})$), with high probability (under the deviation), in all but finitely many periods the consumers' posterior belief will assign probability greater than $1/2$ to the commitment action. Thus, with high probability, the consumers choose the high service level in all but finitely many periods. Hence, if the firm is sufficiently patient, the deviation yields a payoff of approximately 2, or higher.

In discounted repeated games, a common interpretation of patience ($\delta \approx 1$) is that the period length is short ($h \approx 0$). The question arises, then, as to whether or not reputation effects emerge when the firm and the consumers make frequent decisions. Consider the following proposition, whose proof, in Appendix A, uses a continuity argument from Section 4.2. It shows that if the firm's *discount rate* $r > 0$ *is fixed*, then, unsurprisingly, reputation effects fail to emerge, even as the period length shrinks to zero.

Proposition 2.2. *Let the discount rate $r > 0$ be fixed. There exists $\bar{\phi} \in (0, 1)$ such that for all $\phi_0 \in (0, \bar{\phi}]$, if h is close enough to zero, every Nash equilibrium of the game with prior ϕ_0 and period length h yields the monopolist a payoff smaller than $2 - \kappa$, where $\kappa > 0$ is a constant that depends only on the signal structure.*

The source of the difference between the comparative statics with r and h is basic. While the effective discount factor tends to one both when the discount rate r tends to zero (with h fixed) and when the period length h tends to zero (with $r > 0$ fixed), in the former case the signal distribution is not affected by the comparative statics, whereas in the latter case it is. In fact, letting the period length go to zero brings about an *adverse informational effect*. Recall that each period's quality signal is a Poisson *random variable* with parameter $\mu(e)h$, where e is the firm's investment in the period. Hence, as $h \to 0$, the signal distribution converges to a mass point at 0, and this is so irrespective of the monopolist investment level e. Over shorter and shorter periods, signals become *increasingly noisy*: the distance between the signaling distributions induced by any two mixed actions converges to zero as the period length shrinks to zero.[7]

While the ability of making frequent decisions is not sufficient for reputation effects to emerge (Proposition 2.2), if the firm is patient *in the sense that the preference parameter r is close to zero*, then powerful reputation effects arise. The main result of the paper, expressed in the context of the Quality Game, is the following:

Proposition 2.3. *Fix a prior probability $\phi_0 \in (0, 1)$ on the commitment type. Then, for all $\varepsilon > 0$ there exists $\bar{r} > 0$, such that for all $r \in (0, \bar{r}]$, there is a lower bound on the set of Nash equilibrium payoffs of the firm in the game with discount rate r and prior ϕ_0, which converges to $2 - \varepsilon$ as the period length tends to zero.*

Thus, reputation effects satisfy a basic *robustness property*. The adverse *informational effect* that arises when players move frequently is not strong enough to undermine the power of reputations. While Fudenberg and Levine's (1992) analysis yields the result for *fixed period length*, it leaves open the possibility that the required bound on the firm's instantaneous discount rate becomes increasingly stringent

[7]This distance measure of informativeness is equivalent to the one based on the likelihood ratio (i.e. the relative entropy), because the log-likelihood ratio is bounded uniformly as $h \to 0$.

as the period length shrinks to zero. Finally, to put the result in context, it is worth mentioning that the robustness property established for reputation effects is not necessarily fulfilled in the underlying complete information game. In effect, in the complete information game, shortening the period length often leads intertemporal incentives to collapse (Section 3.5).

3 The Model

3.1 The Basic Repeated Game

This section presents the canonical dicrete-time model of reputations with imperfect monitoring. A repeated game is played by a *long-run player* (player 1) and a sequence of *short-run players* (players 2), each of whom is called to play once. At each period $n = 0, 1, 2, \ldots$, the long-run player and the current short-run player choose, simultaneously, actions $a_1^n \in A^1$ and $a_2^n \in A_2$, respectively, where the action spaces A_1 and A_2 are finite sets. The Cartesian product $A_1 \times A_2$ is denoted A.

At the end of each period $n \geqq 0$, a noisy *signal* y^n is drawn from a discrete countable space Y, according to a probability distribution $\rho(a^n) \in \Delta(Y)$, which is jointly controlled by the players through their choice of the action profile $a^n \in A$, but is otherwise independent of the past history of actions and signals. The signal y^n is then publicly announced to all players and this is the only information they receive about each other's moves. The pair (Y, ρ) is called the *signal structure* of the game.

For each $i = 1, 2$ let $\Delta(A_i)$ be the space of *mixed actions* of player i and write $\Delta(A) = \Delta(A_1) \times \Delta(A_2)$ for the set of mixed-action profiles. Denote by $\rho(\alpha)$ the probability distribution over signals induced by a mixed-action profile $\alpha \in \Delta(A)$, that is, extend the map $\rho : A \to \Delta(Y)$ from A to $\Delta(A)$ by bi-linearity.

The payoff of each short-run player depends on his own action and on the public signal, and is given by a bounded function $u_2 : A_2 \times Y \to \mathbb{R}$. The short-run players' expected payoff $\pi_2(\alpha)$ from an action profile $\alpha = (\alpha_1, \alpha_2) \in \Delta(A)$ is then given by

$$\pi_2(\alpha) = \sum_{y \in Y} u_2(\alpha_2, y)\, \rho(\alpha)[y] \, ,$$

where u_2 is extended to $\Delta(A_2) \times Y$ by linearity. The short-run players' payoff function is *common knowledge*.

The short-run players are uncertain about the payoff function of the long-run player and this uncertainty is modeled via the notion of *types*. Accordingly, assume that the long-run player privately learns, prior to the beginning of the game, the value of a payoff-relevant variable θ, his type. The short-run players possess a prior belief p over the type space Θ, a complete separable metric space. The stage-game payoff of the long-run player is given by a bounded measurable function $u_1 : A_1 \times Y \times \Theta \to \mathbb{R}$, which gives rise to an expected payoff function $\pi_1 : \Delta(A) \times \Theta \to \mathbb{R}$ in the natural way:

$$\pi_1(\alpha, \theta) = \sum_{y \in Y} u_1(\alpha_1, y, \theta)\, \rho(\alpha)[y] \, , \quad \forall \alpha \in \Delta(A), \theta \in \Theta \, ,$$

where, as usual, u_1 is extended to $\Delta(A_1) \times Y \times \Theta$ by linearity. The long-run player's overall payoff in the repeated game is the average discounted payoff of his payoff stream, with discount factor $\delta \in (0, 1)$.

A *play path* is a complete description of play, i.e., it is an infinite sequence of the form $h^\infty = (\theta, (a^0, y^0), (a^1, y^1), \ldots, (a^n, y^n), \ldots)$. Write $H^\infty = \Theta \times (A \times Y)^\infty$ for the space of play paths and endow H^∞ with the product σ-algebra $\mathcal{H}^\infty$. The action a_i^n of player $i = 1, 2$ at stage $n \geqq 0$ can be viewed as a random variable defined on the measurable space $(H^\infty, \mathcal{H}^\infty)$, and likewise for the long-run player's type θ and the public signal y^n. For $n \geqq 1$, write $H_1^n = \Theta \times (A_1 \times Y)^n$ for the set of n-period histories of the long-run player and let $H_1^0 = \Theta$ denote the set of initial histories. The information available to the long-run player at the beginning of stage $n \geqq 0$ is represented by the sub-σ-algebra $\mathcal{H}_1^n$ of $\mathcal{H}^\infty$, generated by the cylinders over H_1^n. Similarly, let $H_2^n = Y^n$ be the set of n-period *public* histories, with H_2^0 denoting the one-point set (i.e. the empty history). The public histories comprise all the information available to the short-run players. In particular, a short-run player knows neither the past moves of the long-run player nor the past moves of any previous short-run player, unless such information is conveyed by the signaling structure. Write $\mathcal{H}_2^n$ for the σ-algebra on H^∞ generated by the cylinders over H_2^n.

A behavior strategy of player $i = 1, 2$ is a sequence of maps $\sigma_i^n : H^\infty \to \Delta(A_i)$ such that σ_i^n is $\mathcal{H}_i^n$-measurable for all $n \geqq 0$. A prior probability $p \in \Delta(\Theta)$, together with a behavior-strategy profile $\sigma = (\sigma_1, \sigma_2)$, induce a probability distribution $\mathsf{P}_{p,\sigma}$ over $(H^\infty, \mathcal{H}^\infty)$. For a bounded random variable X, measurable with respect to $\mathcal{H}^\infty$, write $\mathsf{E}_{p,\sigma}(X)$ for the expectation relative to probability measure $\mathsf{P}_{p,\sigma}$.

For a given a prior probability $p \in \Delta(\Theta)$, a (Bayesian) Nash equilibrium of the repeated game is an equilibrium point of the corresponding normal form game in which players choose behavior strategies and the payoffs induced by a behavior strategy profile σ are: $\mathsf{E}_{p,\sigma}(1-\delta)\sum_{n=0}^{\infty} \delta^n \pi_1(a^n, \theta)$, for the long-run player, and $\mathsf{E}_{p,\sigma}\left[\pi_2(a^n) \mid \mathcal{H}_2^n\right]$, for the period-$n$ short-run player.

In the subsequent analysis, I single out a particular type $\theta_0 \in \Theta$ of the long-run player, which will be called the *normal type*. I assume that the short-run players' prior belief p has a *grain of truth*, that is, $p(\{\theta_0\}) > 0$. As is standard in the literature, the incomplete information game presented above is we interpreted as a *perturbation* of the complete information game in which it is common knowledge that player 1 is of type θ_0. In particular, the case of interest is that in which the prior p is such that $p(\{\theta_0\})$ is fixed arbitrarily close to 1, so that it is *almost* common knowledge that player 1 is the normal type.

An arbitrary prior probability need not give rise to reputation effects. Reputations are more powerful when *commitment types* are assigned positive prior probability. Such commitment types are types of player 1 which play, mechanistically, the same (possibly mixed) action in every period, irrespective of history.[8] The set of commitment types is identified with $\Delta(A_1)$ and we require that *commitment types have full support*: every neighborhood of any commitment type has positive probability under p.

A mixed action $\alpha_2 \in \Delta(A_2)$ is a *self-confirmed best-reply to* $\alpha_1 \in \Delta(A_1)$ if:

(a) α_2 is not weakly dominated,

(b) $\alpha_2 \in \arg\max_{\alpha_2' \in \Delta(A_2)} \pi_2(\alpha_1', \alpha_2')$ for some $\alpha_1' \in \Delta(A_1)$ such that $\rho(\alpha_1, \alpha_2) = \rho(\alpha_1', \alpha_2)$.

[8] Fudenberg and Levine (1992) observe that types committed to mixed actions cannot be expected-utility maximizers and introduce a technical device, which involves allowing player 1 to have non-stationary preferences, to overcome this undesirable feature. As this approach requires more cumbersome use of notation, I do not pursue it here. Rather, I idealize commitment types as automata which are programmed to play certain strategies.

For each $\alpha_1 \in \Delta(A_1)$, write $\mathsf{B}(\alpha_1)$ for the set of self-confirmed best responses to α_1. Define:

$$\underline{\pi}_1(\theta_0) := \sup_{\alpha_1 \in \Delta(A_1)} \inf_{\alpha_2 \in \mathsf{B}(\alpha_1)} \pi_1(\alpha_1, \alpha_2, \theta_0), \quad \bar{\pi}_1(\theta_0) := \sup_{\alpha_1 \in \Delta(A_1)} \sup_{\alpha_2 \in \mathsf{B}(\alpha_1)} \pi_1(\alpha_1, \alpha_2, \theta_0).$$

Payoffs $\underline{\pi}_1(\theta_0)$ and $\bar{\pi}_1(\theta_0)$ are called the *lower and upper Stackelberg payoffs* of type θ_0, respectively.

We say that *the game is identified* if for all $\alpha_2 \in \Delta(A_2)$ that is not weakly dominated, and for all α_1, $\alpha_1{}' \in \Delta(A_1)$, one has that $\rho(\alpha_1, \alpha_2) = \rho(\alpha_1', \alpha_2)$ implies $\alpha_1 = \alpha_1'$. The game is called *non-degenerate* if for every undominated pure action $a_2 \in A_2$ and for all $\alpha_2 \in \Delta(A_2)$, one has that $\pi_2(\cdot, \alpha_2) = \pi_2(\cdot, a_2)$ implies $\alpha_2 = a_2$.

Denote by $\mathcal{E}_1(\delta, \theta_0)$ the set of Nash equilibrium payoffs of the long-run player of type θ_0 in the game with discount factor δ, and let $\underline{\mathcal{E}}_1(\delta, \theta_0)$ and $\overline{\mathcal{E}}_1(\delta, \theta_0)$ denote its infimum and supremum, respectively. The following proposition is the central result of Fudenberg and Levine (1992):

Proposition 3.1 (Fudenberg and Levine (1992))**.** *Let the prior probability $p \in \Delta(\Theta)$ be such that commitment types have full support. Then, for every $\theta_0 \in \Theta$ with $p(\{\theta_0\}) > 0$, one has:*

$$\underline{\pi}_1(\theta_0) \leqq \varliminf_{\delta \to 1} \underline{\mathcal{E}}_1(\delta, \theta_0) \leqq \varlimsup_{\delta \to 1} \overline{\mathcal{E}}_1(\delta, \theta_0) \leqq \bar{\pi}_1(\theta_0)$$

Moreover, if the game is non-degenerate and identified, then $\underline{\pi}_1(\theta_0) = \bar{\pi}_1(\theta_0)$ and, consequently,

$$\lim_{\delta \to 1} \mathcal{E}_1(\delta, \theta_0) = \{\bar{\pi}_1(\theta_0)\}.$$

As discussed in Section 2, the above result may be interpreted as a statement about payoff bounds in the limit as the *period length* tends to zero, with a fixed discount rate for the long-run player. This interpretation, however, is problematic. For instance, suppose the game is identified. While the effective discount factor converges to 1 as the players' inertia disappears, shortening the period length also has an important informational effect: *the degree of monitoring imperfection becomes asymptotically negligible*, for the monitoring structure is fixed independent of δ. In effect, as the period length shrinks to zero, the number of signals the short-run players observe in any given interval of *real time* increases without bound. This increasing sample of statistically informative signals is drawn from probability distributions that are independent of the period length, rendering the monitoring asymptotically perfect.

A *persistently imperfect* monitoring structure is one in which the overall informativeness of the signals is *bounded* as the period length shrinks to zero. *This, however, requires the signaling structure to depend on the period length.* In particular, as the frequency of decisions increases, the signal distribution ought to adjust, in such way that the informativeness over any fixed interval of *real time* is bounded.

More precisely, a game has *persistently imperfect monitoring* if, as the period length shrinks to zero, the monitoring structure approaches a *continuous-time limit*, and, further, the limit monitoring is non-trivially imperfect. An example is the Quality Game of Section 2: increasing the frequency with which the firm and the consumers move *does not* increase the monitoring informativeness. Even in the limit game in which firm and consumers can adjust their actions continuously, the monitoring structure (a counting process with intensity controlled by the firm) has *full support*: every neighborhood of any given sample path of the signal process has positive probability under every behavior strategy of the firm.

The property that the monitoring be persistently imperfect is an essential requirement for a meaningful theory of imperfect monitoring games with frequent decisions. In a model in which it doesn't hold, the comparative statics exercise with the period length is difficult to interpret. Such exercise is equivalent to a comparative statics of information structures, and, in particular, to one in which the information structures approach the perfectly informative one.[9] This, however, is contrary to the very spirit of the theory of games with imperfect monitoring.

In the the next section, the concept of games with persistently imperfect monitoring is developed more formally.

3.2 Games with Frequent Decisions

Consider a family of games (Γ_h) indexed by the *period length* $h > 0$. Each Γ_h is a repeated game, as presented in section 3.1, with basic data $\langle(\Theta, p), (A_i, \pi_i), (Y_h, \rho_h), \delta_h\rangle$. The type space (Θ, p) and the stage game $(A_i, \pi_i)_{i=1,2}$ are independent of h and will be fixed hereafter.

While each Γ_h is formally a discrete-time repeated game, it can be viewed as a continuous-time game in which players are subject to *inertia*. Fix the period length $h > 0$ and define a continuous-time game Γ'_h as follows. Time $t \geqq 0$ runs continuously, but players are capable of adjusting their actions only at the "ticks" of a Poisson clock. More precisely, the decision times are random variables

$$0 = \tau^0 < \tau^1 < \ldots < \tau^n < \ldots ,$$

where the durations $\tau^n - \tau^{n-1}$ are independent and exponentially distributed random variables with mean $h > 0$.[10] At each time $t = \tau_n$ a new short-run player arrives and the previous short-run player departs from the game.

To each play path $h^\infty \in H^\infty$ in the repeated game Γ_h there corresponds a trajectory $(h'(t))_{t \geqq 0}$ in the continuous-time game, where $h'(t) = h^n$ for all $t \in [\tau^n, \tau^{n+1})$. Likewise, every repeated-game strategy σ_i of player i is mapped into a strategy in the continuous-time game, as follows: at each time $t = \tau_n$, $n \geqq 0$, player i plays the period-n mixed action σ_i^n prescribed by his strategy in the repeated game; in the interval $[\tau^n, \tau^{n+1})$ actions are constant and are determined by the realization of the players' randomization at $t = \tau_n$.

Turning to the definition of payoffs, fix a play path in the continuous-time game Γ'_h and consider the associated payoff flow $(\pi'_1(t))_{t \geqq 0}$ of the long-run player. The overall payoff of the long-run player on this given path is the continuously discounted average of $(\pi'_1(t))_{t \geqq 0}$,

$$r \int_0^\infty e^{-rt} \pi'_1(t)\, dt ,$$

where the long-run player's *discount rate* $r > 0$ is such that the expected value of the above expression, with respect to the distribution of $(\tau^n)_{n \geqq 0}$, coincides with the correspondent repeated-game payoff of

$$(1 - \delta_h) \sum_{n=0}^{\infty} \delta_h^n \pi_1^n ,$$

[9] More precisely, the monitoring structure approaches the perfectly informative one only when the game is identified. When identification fails, as the period length shrinks, the signal structure approaches a *partitional* one: for each *mixed* action of player 2, there is a partition on player 1's *mixed-action* space, such that in any period player 2 is informed only in which element of the partition player 1's mixed action is. (In particular, in perfect monitoring games, in the limit as $h \to 0$, mixed-actions become effectively observable).

[10] If instead I had assumed that the decision times lie in the grid $\{0, h, 2h, \ldots\}$, then the signals, when embedded in continuous time, would not follow a continuous-time Markov chain, for the transition times would be deterministic (see footnote 11 below). Yet, *all* results in this paper would carry over to that case.

where $(\pi_1^n)_{n \geqq 0}$ is the discrete-time payoff flow correspondent to $(\pi_1'(t))_{t \geqq 0}$, that is, $\pi_1^n = \pi_1'(\tau^n)$ for all $n \geqq 0$. This requires

$$\delta_h = \frac{1}{1+rh}\,, \tag{3.1}$$

for all $h > 0$. For this continuous-time embedding to be meaningful at all, a minimal requirement is that r be independent of h, so that the preference parameter r can be interpreted as an *instantaneous discount rate*. From here on, I assume that for some $r > 0$ the family of discount factors $(\delta_h)_{h>0}$ satisfies (3.1). In particular, note that $\delta_h \to 1$ as $h \to 0$.

Before presenting the formal definition of games with *persistently imperfect monitoring*, I provide some motivation. The definition will have two parts. First, I require that, as the period length shrinks to zero, the monitoring technology approach a continuous time limit. Second, that the limit monitoring have "full support". In the continuous-time limit, the signals are the sample paths of a controlled Markov process of Lévy type (Ethier and Kurtz (1986), pg. 379), i.e., a random process $(X_t)_{t \geqq 0}$ with values in an Euclidean space $\mathbb{R}^d$, with evolution governed by the following stochastic differential equation:

$$dX_t \;=\; b(a_t)\,dt \;+\; \gamma(a_t)\,dW_t \;+\; dJ_t\,. \tag{3.2}$$

Here, W is a standard d-dimensional *Brownian motion*, J is a *pure-jump process* and the maps $b : A \to \mathbb{R}^d$ and $\gamma : A \to \mathbb{R}^{d\times d}$ determine the *drift*, $b(a_t)$, and the *dispersion matrix*, $\gamma(a_t)$, as functions of the current action profile $a_t \in A$. The pure-jump process J is jointly controlled by the players and is characterized by a transition kernel $K : A \times \mathcal{R}^d \to \mathbb{R}$, from the action set A into the set of *finite* positive measures over $\mathcal{R}^d$ (the Borel σ-algebra on $\mathbb{R}^d$), with $K(a,\{0\}) = 0$ for all $a \in A$. The jumps arrive at rate $\mu(a_t) = K(a_t, \mathbb{R}^d) < \infty$ and, conditional on the event that a jump occurs at time t, the probability that the jump size is in a set $B \in \mathcal{R}^d$ is $K(a_t, B)/\mu(a_t)$.

Turning to the family $(Y_h, \rho_h)_{h>0}$, I assume that the signal space Y_h is a discrete subset of the Euclidean space $\mathbb{R}^d$. Fix the parameter $h > 0$ and consider an infinite sequence $(a^n)_{n\geqq 0}$ of action profiles. Conditional on $(a^n)_{n\geqq 0}$, the public signals $(y^n)_{n\geqq 0}$ form a sequence of independent random variables with marginal distributions given by $\rho_h[a^n]$. Define a *cumulative process* $(x^n)_{n\geqq 0}$ by setting $x^0 = 0$ and $x^{n+1} = \sum_{k=0}^n y^k$ for all $n \geqq 0$. The sequence $(x^n)_{n\geqq 0}$ is a Markov chain on the state space S_h, of all sums of the form $\sum_{k=0}^n y^k$ with $y^k \in Y_h$ for all $k = 1, \dots, n$. The transition probabilities are given by

$$\Pr\{x^{n+1} = x' | x^n = x\} = \rho_h(a^n)[x' - x]\,,\ \text{for all } x, x' \in S_h\,.$$

Consider the *continuous-time interpolation* $(a_t', x_t')_{t\geqq 0}$ of the sequence $(a^n, x^n)_{n\geqq 0}$. More precisely, let $(\tau^n)_{n\geqq 0}$ be the jump times of the Poisson clock (whose parameter is $1/h$) and define $(a_t', x_t') = (a^n, x^n)$ for all $t \in [\tau^n, \tau^{n+1})$. The first part of the definition of persistently imperfect monitoring will express the idea that, for h near zero, the *continuous-time Markov chain* $(x_t')_{t\geqq 0}$ "behaves like" the stochastic process $(X_t)_{t\geqq 0}$.[11]

The second part of the definition of persistently imperfect monitoring concerns the requirement that the limit monitoring have "full support". More precisely, in the continuous-time limit the signals are the sample paths of the process $(X_t)_{t\geqq 0}$, so the signal space is "large" and the concept of full support must be replaced by that of *mutual absolute continuity*. (Two probability distributions are mutually absolutely continuous if their null events coincide.[12]) Accordingly, the second part of the definition of games with

[11] Because of the Poisson clock assumption, the interpolated process $(x_t')_{t\geqq 0}$ is a continuous-time Markov chain.

[12] For discrete distributions, this simply means that the two probabilities have the same support.

persistently imperfect monitoring will be a requirement that, in the continuous-time limit, over any *finite horizon*, the long-run players' actions affect the probability law of the signal process X in an absolutely continuous way.[13]

Before turning to the definition, I impose the last restriction on the limit monitoring structure. Hereafter, I assume that K, the transition kernel associated with the jump term of equation (3.2), admits the following decomposition:

$$K(a,B) = \int_B f(a,\xi)G(d\xi)\,, \quad \text{for all } a \in A \text{ and } B \in \mathscr{R}^d\,,$$

where G is a finite positive measure on $\mathscr{R}^d$, independent of a, with support in a compact set Ξ and with $G(\{0\}) = 0$; and the density $f : A \times \Xi \to \mathbb{R}$ is a bounded measurable function such that, for some $\kappa > 0$,

$$f(a,\xi) \geqq \kappa \quad \text{for all } \xi \in \operatorname{supp} f(a,\cdot)\,. \tag{3.3}$$

For ease of reference, I call the parameters (b,γ,f,G) of equation (3.2) a *quadruplet of characteristics*.

Given any behavior strategy profile in some Γ'_h, a unique (weak sense) solution of equation (3.2) exists.[14] For our main result, however, more is needed. From here on, I will only consider characteristics (b,γ,f,G) for which a unique solution of equation (3.2) exists for all *public strategy profiles in the limit game*, including some in which players adjust their actions continuously (see Remark 3.1 below). I call such characteristics *regular*. Because a rigorous definition of regularity must be preceded by a description of the limit game, the formal definition is deferred to Section 4.1 and, for now, we shall content ourselves with this informal discussion.

Remark 3.1. A precise statement of the property that the Markov chain $(x'_t)_{t\geqq 0}$ "behaves like" $(X_t)_{t\geqq 0}$ for h near zero requires the definition of Γ'_0, the *limit game*. In Γ'_0, defined in Section 4.1, signals are determined by equation (3.2) and players' behavior strategies allow actions to be adjusted continuously over time. The players' strategy spaces in Γ'_0 are endowed with a topology under which they are compact sets. Moreover, the strategy sets in Γ'_h are naturally embedded as subsets of the strategy sets in Γ'_0. The first part of the definition of persistently imperfect monitoring games is the requirement that the signaling structure (Y_h,ρ_h) satisfy a sufficient condition for the following *consistency property*. Let the family $(\sigma'_h)_{h\geqq 0}$ be such that each σ'_h is a behavior strategy profile in Γ'_h and $\sigma'_h \to \sigma'_0$ as $h \to 0$. Then, the probability law of the process $(x'_t)_{t\geqq 0}$ under σ'_h weakly converges to that of $(X_t)_{t\geqq 0}$ under σ'_0. □

Definition 3.1. *A family of signal structures* $(Y_h,\rho_h)_{h>0}$ *displays* persistently imperfect monitoring *if, for some regular characteristics* (b,γ,f,G)*, one has:*

(a) ***Consistency****:* (Y_h,ρ_h) *is* consistent with (b,γ,f,G)*, i.e., it admits the factorization:*

$$\rho_h(a)[y] \;=\; (1-\mu(a)h) \times \breve{\rho}_h(a)[y] \;+\; \mu(a)h \times Q(\iota_h^{-1}(y)|a) \;+\; o(h)\,, \tag{3.4}$$

where

[13] The qualifier *finite horizon* is important here. Suppose that X follows $dX_t = b(a_t)\,dt + dW_t$. Then, for any pair of strategy profiles σ', σ'' in some Γ'_h, the prob. measures induced by σ' and σ'' over the finite horizon paths $(X_t)_{0\leqq t\leqq T}$ are mutually absolutely continuous for all $0 \leqq T < \infty$. However, for $T = \infty$ they are not so.

[14] This is because strategies in Γ'_h, with $h > 0$, are piecewise constant, so that existence/uniqueness can be established by appealing to the characterization of Lévy processes by their infinitesimal generators (Ethier and Kurtz (1986), pg. 380, Theorem 3.4).

(i) $\breve{\rho}^h$ *is a signaling structure with support on a subset* $\breve{Y}_h \subseteq Y_h$*, such that* $\breve{Y}_h \to \{0\}$ *as* $h \to 0$*, and for all* $a \in A$*,*

$$\begin{aligned} \breve{\mathsf{E}}_h(y|a) &= b(a)h + o(h), \\ \breve{\mathsf{Var}}_h(y|a) &= \gamma(a)\gamma(a)^\top h + o(h), \end{aligned}$$

(ii) $\mu(a) = \int_\Xi f(a,\xi)\, G(d\xi)$ *and* $Q(B|a) = \int_B f(a,\xi) G(d\xi)/\mu(a)$ *for all* $a \in A$ *and* $B \in \mathcal{R}^d$*,*

(iii) $\iota_h : \Xi \to Y_h$ *is a measurable function such that* $\|\iota_h(\xi) - \xi\| \to 0$ *uniformly as* $h \to 0$*.*

(b) ***Full support****: The support of* $\rho_h(a)$ *is independent of the long-run player's action* a_1*, for each* $h > 0$*. Furthermore, for any pair of constant pure-strategy profiles* σ' *and* σ'' *with* $\sigma'_2 = \sigma''_2$*, and any* finite *horizon* $T > 0$*, the probability measures induced by* σ' *and* σ'' *over the paths of* $(X_t)_{0 \leqq t \leqq T}$ *(via equation (3.2)) are mutually absolutely continuous.*

The meaning of the factorization in (3.4) can be read directly from the formula. With probability of order $1 - \mu(a)h$ there is no jump and the transition is given by the "diffusion-like" probabilities $\breve{\rho}_h(a)$. Indeed, condition (i) means that, for h near zero, the transition $\breve{\rho}_h(a)$ is close to that of a continuous diffusion with drift $b(a)$ and dispersion matrix $\gamma(a)$. Continuing with the interpretation, with probability of order $\mu(a)h$ a jump occurs, which is random and distributed according to $Q(\cdot|a)$. The role of the function ι_h is to ensure that the jumps lie in the discrete set Y_h.

Remark 3.2. The second part of the Full Support condition has an equivalent formulation purely in terms of the characteristics (b, γ, f, G) (Proposition B.1, Appendix B). For instance, the dispersion matrix $\gamma(a)$ must be independent of the long-run player's actions (see discussion in Example 2 below). □

I now illustrate the definition through four examples of persistently imperfect monitoring structures:

(a) *The Poisson process*. The signal structure (Y_h, ρ_h) is as in the Quality Game (Section 2), with $Y_h = \{0, 1, 2, \ldots\}$, $\mu(a) > 0$ for all $a \in A$, and

$$\rho_h(a)[k] = \frac{(\mu(a)h)^k}{k!}\, e^{-\mu(a)h}\,.$$

The characteristics of (Y_h, ρ_h) are $(0, 0, \mu, \delta_{\{1\}})$, where $\delta_{\{1\}}$ is a mass point at 1. Alternatively, one can consider a first-order approximation with $Y_h = \{0, 1\}$ and

$$\rho_h(a)[1] = 1 - e^{-\mu(a)h}\,.$$

Such signal structure is also consistent with $(0, 0, \mu, \delta_{\{1\}})$.

(b) *The random walk approximation of a diffusion*. The signal structure (Y_h, ρ_h), with $Y_h = \{\pm\gamma(a)h^{1/2} : a \in A\}$, $\gamma(a)$ independent of a_1, and

$$\rho_h(a)[\pm\gamma(a)\sqrt{h}] = \frac{\gamma(a) \pm b(a)\sqrt{h}}{2\gamma(a)}\,,$$

is consistent with $(b, \gamma, 0, 0)$. A pictorial representation of this signal structure is the binomial tree of Figure 1.

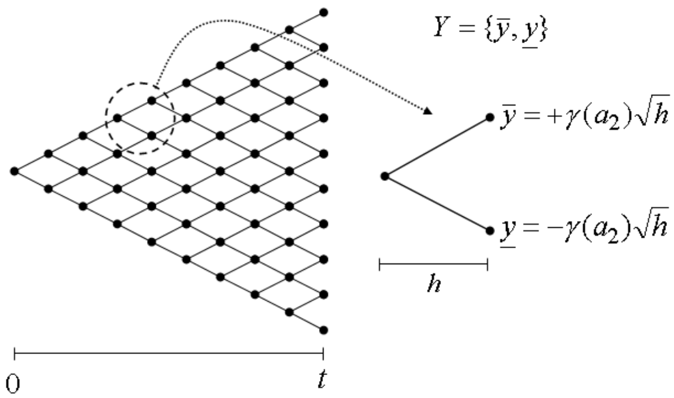

Figure 1: Random walk approximation of one-dimensional diffusion.

This example illustrates the role of indexing the signal space Y_h by the period length.

If the long-run player is allowed to control the coefficient γ, then the signal structure above will still be consistent, but will fail to satisfy the Full Support condition. Consider a constant pure-strategy profile $\sigma' = (\sigma'_1, \sigma'_2)$ that prescribes action profile $a \in A$ at all times. The probability measure induced by σ' over the paths of $(X_t)_{t \geqq 0}$ (via equation (3.2)) assigns probability 1 to the event that the *quadratic variation of* X is $(\gamma(a)^2 t)_{t \geqq 0}$.[15] Therefore, if the long-run player controls γ, Full Support is clearly violated.

(c) *Pure-jump process with discrete jump distribution.* Let $Y_h = Y$ be a discrete subset of $\mathbb{R}^d$ independent of h and containing the origin. Define

$$\rho_h(a)[y] = \begin{cases} 1 - \mu(a)h & : \quad y = 0 \\ \mu(a)h \times Q(y|a) & : \quad y \in Y \setminus \{0\} \end{cases},$$

where $\mu(a) > 0$ is the *jump rate* and $Q(\cdot|a) \in \Delta(Y \setminus \{0\})$ is the *jump distribution*. Then (Y, ρ_h) is consistent with characteristics $(0, 0, f, G)$, where $f(a, \xi) = \mu(a)Q(\xi|a)$ and $G(\xi) = 1$ for all $\xi \in \Xi = Y \setminus \{0\}$.

For instance, suppose that $Y = \{e_j : j = 1, \ldots, d\} \cup \{0\}$, where e_j is the j-th vector of the canonical basis of $\mathbb{R}^d$. Let $\mu_j(a) > 0$ be given for all $a \in A$ and $1 \leqq j \leqq d$. Define $\mu(a) = \sum_{j=1}^d \mu_j(a)$, $Q(e_j|a) = \mu_j(a)/\mu(a)$ and $\rho_h(a)$ as above. Then (Y, ρ_h) is a first-order approximation of the d-dimensional Poisson process, where $\mu_j(a)$ represents the jump rate of the j-th coordinate.

(d) *Pure-jump process with non-atomic jump distribution.* The signal space is given by $Y_h = \{-1, 0, 1\} \cup \{\pm jh : j \in \mathbb{N}, jh < 1\}$, a discretization of the interval $\Xi = [-1, 1]$. Function $\iota_h : \Xi \to Y_h$ is

[15] The *quadratic variation* of X is defined as $[X]_t = X_t^2 - 2\int_0^t X_{s-}\, dX_s$. In particular, $[X]_t$ is determined by the history $(X_s)_{s \leqq t}$, so player 2, in the limit game, effectively knows the history $([X]_s)_{0 \leqq s < t}$ at time t.

defined by:

$$\iota_h(\xi) = \begin{cases} \xi & : & \xi \in \{0, \pm 1\} \\ jh & : & \xi \in ((j-1)h, jh], j \geqq 1 \\ -jh & : & \xi \in [-jh, -(j-1)h), j \geqq 1 \end{cases} .$$

Let $A_1 = \{1, 2\}$, $A_2 = \{-1/2, 1/2\}$ and let $g(a, \cdot)$ be the density of a beta distribution with both parameters equal to $(a_1 + a_2)/3$ and with support in the closed interval with extremes $a_2 \pm 1/4$.[16] Finally, let $\mu(a) = a_1 + a_2$ and $Q(I|a) = \int_I g(a, \xi) d\xi$ for every interval $I \subset \Xi$. Define

$$\rho_h(a)[y] = \begin{cases} 1 - \mu(a)h & : & y = 0 \\ \mu(a)h \times Q(\iota_h^{-1}(y)|a) & : & y \in Y_h \setminus \{0\} \end{cases} .$$

Signal structure (Y_h, ρ_h) is consistent with characteristics $(0, 0, f, G)$, where $f(a, \cdot) = \mu(a)g(a, \cdot)$ and G is Lebesgue measure on $[-1, 1]$. The role of function ι_h is to ensure that the signal takes values in a discrete set. Because $\iota_h(\xi) \to \xi$ as $h \to 0$, the scheme above approximates a pure-jump process with non-atomic jump distribution. Note that both the jump rate and the jump distribution are controlled by both players. Also, observe that the support of $f(a, \cdot)$ depends only on player 2's actions. This ensures that the Full Support condition of Definition 3.1 is satisfied (Appendix B, Proposition B.1).

3.3 The Main Result

Fix a family of games with persistently imperfect monitoring consistent with characteristics (b, γ, f, G). Define the *infinitesimal self-confirmed best-reply correspondence* $\mathsf{B}' : \Delta(A_1) \rightrightarrows \Delta(A_2)$ as follows:[17] For each mixed action profile $(\alpha_1, \alpha_2) \in \Delta(A)$, one has $\alpha_2 \in \mathsf{B}'(\alpha_1)$ if, and only if,

(a) α_2 is not weakly dominated, and

(b) $\alpha_2 \in \arg\max_{\alpha_2' \in \Delta(A_2)} \pi_2(\alpha_1', \alpha_2')$ for some $\alpha_1' \in \Delta(A_1)$ such that $b(\alpha_1, \alpha_2) = b(\alpha_1', \alpha_2)$ and $f(\alpha_1, \alpha_2, \cdot) = f(\alpha_1', \alpha_2, \cdot)$ G-a.e.

Remark 3.3. Note that γ does not enter the definition above. This is because the Full Support condition of Definition 3.1 implies that γ is independent of player 1's actions (Appendix B, Proposition B.1), so that player 2 effectively knows γ. □

For each $\theta_0 \in \Theta$ such that $p(\theta_0) > 0$ define:

$$\underline{\pi}_1'(\theta_0) := \sup_{\alpha_1 \in \Delta(A_1)} \inf_{\alpha_2 \in \mathsf{B}'(\alpha_1)} \pi_1(\alpha_1, \alpha_2, \theta_0) , \quad \bar{\pi}_1'(\theta_0) := \sup_{\alpha_1 \in \Delta(A_1)} \sup_{\alpha_2 \in \mathsf{B}'(\alpha_1)} \pi_1(\alpha_1, \alpha_2, \theta_0).$$

We say that the game is *identified* if for all $\alpha_2 \in \Delta(A_2)$ and $\alpha_1', \alpha_1'' \in \Delta(A_1)$ such that

$$\begin{aligned} b(\alpha_1', \alpha_2) &= b(\alpha_1'', \alpha_2) \\ f(\alpha_1', \alpha_2, \cdot) &= f(\alpha_1'', \alpha_2, \cdot) \quad G - \text{a.e.} , \end{aligned}$$

[16] Note that $(a_1 + a_2)/3 < 1$ so that the density g is bounded above 0 and condition (3.3) is satisfied.

[17] Here I used the same terminology as in Section 3.1, where correspondence B was also called self-confirmed best-reply. Nevertheless, B and B' are *different objects*: B is defined in terms of (Y_h, ρ_h), whereas B' depends only on the characteristics (b, γ, f, G). Loosely speaking, B' is an "infinitesimal" version of B.

one has $\alpha_1' = \alpha_1''$. The game is called *non-degenerate* if for every undominated pure action $a_2 \in A_2$ and for all $\alpha_2 \in \Delta(A_2)$, one has that $\pi_2(\cdot, \alpha_2) = \pi_2(\cdot, a_2)$ implies $\alpha_2 = a_2$.

For each h and $r > 0$, write $\underline{\mathcal{E}}_1(r, h, \theta_0)$ (resp. $\overline{\mathcal{E}}_1(r, h, \theta_0)$) for the infimum (resp. supremum) over the set of Nash equilibrium payoffs of the long-run player of type θ_0 in the game with period length h and discount factor $\delta_h = 1/(1 + rh)$.

The main result of the paper is:

Theorem 3.1. *Let $(\Gamma_h)_{h>0}$ be a family of games with persistently imperfect monitoring consistent with regular characteristics (b, γ, f, G). Suppose that commitment types have full support. If $\theta_0 \in \Theta$ is such that $p(\{\theta_0\}) > 0$, then for all $\varepsilon > 0$ there exists $\bar{r} > 0$ such that for any discount rate $r \in (0, \bar{r}]$ one has:*

$$\underline{\pi}_1'(\theta_0) - \varepsilon \leqq \liminf_{h \to 0} \underline{\mathcal{E}}_1(r, h, \theta_0) \leqq \limsup_{h \to 0} \overline{\mathcal{E}}_1(r, h, \theta_0) \leqq \bar{\pi}_1'(\theta_0) + \varepsilon .$$

Moreover, if the stage-game is non-degenerate and identified, then $\bar{\pi}_1(\theta_0) = \underline{\pi}_1(\theta_0)$ and, consequently,

$$\lim_{r \to 0} \lim_{h \to 0} \overline{\mathcal{E}}_1(r, h, \theta_0) = \{\bar{\pi}_1'(\theta_0)\} .$$

The result provides payoff bounds from reputation effects which hold *uniformly* over all games Γ_h with sufficiently short period length. A natural approach to establish the desired uniformity is to consider the Fudenberg and Levine (1992) bounds, which are valid for each r, $h > 0$, and argue, based directly on the consistency conditions, that they converge to the bounds of Theorem 3.1 (with the order of limits being $h \to 0$ first, and then $r \to 0$). As will be shown in the next section, this approach works only for a particular class of games. When it works, though, it provides a rather simple proof of Theorem 3.1.

To obtain the desired result for the complete class of games with persistently imperfect monitoring the proof must improve on the Fudenberg-Levine bounds. The way this is accomplished is by studying the limit game. The proof, in Section 4, derives the bounds for the limit game and shows that such bounds are valid along the approximating sequence.

3.4 The Speed of Bayesian Learning

At the core of the study of reputation effects lies the learning problem faced by the short-run players. In this section, I relate Theorem 3.1 to the *speed* with which the short-run players learn about the distribution over signals they face in a Nash equilibrium. I will show how Fudenberg and Levine's (1992) payoff bounds yield a proof of Theorem 3.1 when the monitoring structure is of diffusion type, i.e., there are no jumps and the characteristics are of the form $(b, \gamma, 0, 0)$. Interestingly, diffusion games are essentially the *only* class of games with persistently imperfect monitoring for which the conclusion of Theorem 3.1 follows from Fudenberg and Levine's payoff bounds.

First, we review a statistical result that will be instrumental in the analysis that follows. Fix a measurable space $(\Omega, \mathcal{F})$ together with a filtration $(\mathcal{F}_n)_{n \geqq 1}$, with each $\mathcal{F}_n$ generated by a countable partition of Ω into measurable sets. Given P, P' probability measures on $(\Omega, \mathcal{F})$ define:

$$d_n(P, P') = \sup_{B \in \mathcal{F}_n} \left| P(B | \mathcal{F}_{n-1}) - P'(B | \mathcal{F}_{n-1}) \right| .$$

The following merging result, due to Fudenberg and Levine (1992), is the main building block in their proof of the reputation result (see also Sorin (1999)).

Proposition 3.2. *For all ε, η and $\phi > 0$, there exists a natural number $N \geqq 1$, such that for all P, P' and P'', probability measures on $(\Omega, \mathcal{F})$ with $P = \phi P' + (1-\phi)P''$, one has:*

$$P'\left(\#\{n \geqq 1 \,:\, d_n(P, P') \geqq \varepsilon\} \geqq N\right) \leqq \eta \tag{3.5}$$

The P'-almost sure convergence $d_n \to 0$ follows from Blackwell and Dubins (1962) classical "merging of opinions" result. The Fudenberg-Levine merging result strengthens the latter under a *grain of truth condition*. It establishes the uniformity of the convergence rate over all P, P' and P'' that satisfy $P = \phi P' + (1-\phi)P''$, where the "size of the grain", ϕ, is fixed.[18]

Let us review how this result relates to reputation effects. First, let us discuss the case in which the period length $h > 0$ is fixed. For exposition clarity, the discussion will be conducted in the context of the Quality Game of Section 2.

Fix the prior probability $\phi \in (0,1)$ on the commitment type. Let σ be an arbitrary Nash equilibrium of the repeated game with discount factor δ_h and prior ϕ, and denote by P the probability measure induced by σ over the infinite histories of the public signal. Also, let P' (resp. P'') be the probability over infinite public histories induced by σ *conditional* on the commitment type (resp. normal type). Thus, $P = \phi P' + (1-\phi)P''$. For each $n \geqq 0$, $\mathcal{F}_n$ will be the σ-algebra generated by the period-n public histories, which are the histories the consumers recall at period n. Fix $\eta' > 0$ and choose $\varepsilon(h) > 0$ small enough so that $\alpha_1(\bar{e}) > 1/2$ whenever $\sup_{k\geqq 0} |\rho_h(\alpha_1)[k] - \rho_h(\bar{e})[k]| < \varepsilon(h)$. For future reference, we note that for any choice of $\varepsilon(h)$ for which the latter implication holds, one has $\limsup_{h\to 0} \varepsilon(h)/h < \infty$.[19]

If the normal firm deviates and mimics the commitment type, the payoff from this deviation is lower than the payoff the firm receives under the equilibrium strategy. Proposition 3.2 will help finding a lower bound for the firm's payoff under this deviation. Let the integer $N(h)$ be such that inequality (3.5) holds for $\varepsilon = \varepsilon(h)$ and $\eta = \eta'/2$. This means that, under the deviation, with probability at least $1 - \frac{\eta'}{2}$, there are at most $N(h)$ periods in which the consumers are not consuming the high service level. (Recall that whenever the consumers assign probability greater than $\frac{1}{2}$ to high investment, their best-reply is to consume high service). Assume, conservatively, that such exceptional periods are precisely the $N(h)$ initial periods. This yields the Fudenberg-Levine lower bound, $\mathsf{FL}^{\eta'}$, for the payoff the firm receives in a Nash equilibrium:

$$\mathsf{FL}^{\eta'}(r,h) := 2\delta_h^{N(h)} - \eta' = 2\left(\frac{1}{1+rh}\right)^{N(h)} - \eta' \,.$$

Notice that $N(h)$ depends only on the prior ϕ and on the period length h, which are fixed, and on η'. In particular, it does not depend on the equilibrium profile and, consequently, not on the discount rate r. The above expression gives a lower bound for *all* Nash equilibrium payoffs of the firm in the game with discount rate r, for each $r > 0$. Letting $r \to 0$ first, and then $\eta' \to 0$, yields the commitment payoff of 2. This is Fudenberg and Levine's (1992) argument.

Consider Theorem 3.1. It asserts that for all $\eta > 0$ there is a lower bound $\ell(r,h)$ for the set of Nash equilibrium payoffs of the firm, and a critical value $\bar{r} > 0$ of the discount rate (independent of h), such that for all $r < \bar{r}$ one has $\ell(r,h) \to 2-\eta$, as $h \to 0$. Can this lower bound be taken as the Fudenberg-Levine lower bound? In other words, does $\liminf_{r\to 0} \liminf_{h\to 0} \mathsf{FL}^{\eta}(r,h) \geqq 2-\eta$? Surprisingly, the answer is

[18] Merging also plays a central role in the literature on rational non-equilibrium learning in games (Kalai and Lehrer (1993)).
[19] This follows from the fact that $\sup_{k\geqq 0} |\rho_h(\alpha_1)[k] - \rho_h(\bar{e})[k]| = [1-\alpha_1(\bar{e})][(\mu(\underline{e}) - \mu(\bar{e}))h + o(h)]$.

negative. The following result, proved in Appendix C, is key to understand why. It gives sharp estimates of the speed of convergence in Proposition 3.2.

Proposition 3.3. *For each ε, η and $\phi > 0$, let $N^*(\varepsilon, \eta, \phi)$ be the smallest natural number such that inequality (3.5) holds for all probability measures P, P' and P'' with $P = \phi P' + (1-\phi)P''$. Then, for every η and $\phi > 0$, one has:*

$$\limsup_{\epsilon \to 0} \varepsilon^2 N^*(\varepsilon, \eta, \phi) < \infty \,.$$

Furthermore, if $\mathcal{F}_n \subsetneqq \mathcal{F}_{n+1}$ for all $n \geqq 1$, then, for all $q > 0$ and all $\phi \in (0, 1]$, there exists $\bar{\eta} > 0$ such that for all $0 < \eta < \bar{\eta}$,

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

Consider the Fudenberg-Levine bound with $N(h) = N^*(\varepsilon(h))$, where N^* is as defined in the Proposition above (omitting the arguments η and ϕ). Recall that, necessarily, $\limsup_{h \to 0} \varepsilon(h)/h < \infty$. Hence, the second part of Proposition 3.3 (with $q = 1$) yields $\varepsilon(h)N(h) \to \infty$ as $h \to 0$. This implies that the *real time* $hN(h) \to \infty$ as $h \to 0$. Thus, for h small enough,

$$\mathsf{FL}^\eta(r, h) = 2(\frac{1}{1+rh})^{N(h)} - \eta \leqq 2e^{-\frac{1}{2}rhN(h)} - \eta \,,$$

and therefore $\limsup_{h \to 0} \mathsf{FL}^\eta(r, h) \leqq -\eta$.

There is nothing special about the Quality Game or the Poisson process: for any game (with generic payoff structure) in which the long-run player controls the jump term, the Fudenberg-Levine bound is too crude to yield the conclusion of Theorem 3.1. The proof of Theorem 3.1 hence provides an improvement on the available bounds, which is *necessary* to derive the payoff bounds from reputation effects in games with persistently imperfect monitoring.

Interestingly, if one restricts attention to diffusion games, then the Fudenberg-Levine bound suffices. In fact, consider a variation of the Quality Game in which the payoff structure is unchanged but the signal structure (Y_h, ρ_h) is, as in Example 2 above, given by $Y_h = \{\pm h^{\frac{1}{2}}\}$, and for $e \in \{\bar{e}, \underline{e}\}$,

$$\rho_h(e)[\pm h^{\frac{1}{2}}] = \frac{1 \pm b(e)h^{\frac{1}{2}}}{2} \,,$$

where $b(\bar{e}) \neq b(\underline{e})$, so that (Y_h, ρ_h) is consistent with

$$dX_t = b(e_t)\, dt \; + \; dW_t \,.$$

The difference from the Poisson case is that here $\varepsilon(h)$ can be taken of order $\sqrt{h}$, so it converges to zero much faster than in the Poisson case.[20] Let $N(h) = N^*(\varepsilon(h))$. It follows from the first part of Proposition 3.3 that $\varepsilon(h)^2 N(h)$ is bounded as $h \to 0$. Since $\sqrt{h} = O(\varepsilon(h))$ we have that $\limsup_{h \to 0} hN(h) < T$ for some $0 < T < \infty$. Therefore,

$$\mathsf{FL}^\eta(r, h) = 2(\frac{1}{1+rh})^{N(h)} - \eta \geqq 2e^{-rhN(h)} - \eta \geqq 2e^{-rT} - \eta \,,$$

[20]Taking $\varepsilon(h) = \frac{1}{2}\,|b(\bar{e}) - b(\underline{e})|\,\sqrt{h}$ suffices: If α_1 is a mixed action of the firm such that $\|\rho_h(\alpha_1) - \rho_h(\bar{e})\| < \varepsilon(h)$, then $\alpha_1(\bar{e}) > 1/2$.

and thus, for all $\eta > 0$, $\liminf_{r \to 0} \liminf_{h \to 0} \mathsf{FL}^{\eta}(r, h) \geqq 2 - \eta$ and letting $\eta \to 0$ yields the desired result. Once again, nothing is special about this example and whenever the long-run player 1 *does not* control the jump term, the Fudenberg-Levine bound suffices.

The following proposition summarizes the above discussion:

Proposition 3.4. *In the Quality Game of Section 2 (with Poisson signal structure), one has for all η and $r > 0$:*

$$\limsup_{h \to 0} \mathsf{FL}^{\eta}(r, h) \leqq -\eta\,.$$

In the variation of the Quality Game in which the signal structure is given by the random walk diffusion approximation (Example 2), one has for all $\eta > 0$:

$$\liminf_{r \to 0} \liminf_{h \to 0} \mathsf{FL}^{\eta}(r, h) \geqq 2 - \eta\,.$$

3.5 Equilibrium Degeneracy

The main result of this paper, Theorem 3.1, establishes that reputation effects satisfy a robustness property, namely the adverse informational effect of shortening the period length (see Section 2) is not strong enough to interfere with the long-run player's reputation. In this section I will argue, based on a result of Faingold and Sannikov (2005) for the limit game, that the analogous robustness property for complete information games is not fulfilled when the monitoring structure is given by a Brownian motion.

Proposition 3.5. *Fix a family of* complete information *games with persistently imperfect monitoring consistent with regular characteristics of the form $(b, \gamma, 0, 0)$. Then, for all $r > 0$, the set of Nash equilibrium payoffs of the long-run player converges, as the period length h tends to 0, to the convex hull of the set of static Nash equilibrium payoffs.*

An implication of Theorem 3.1 and the above proposition is that *the intertemporal incentives created by reputation effects are the only non-trivial form of incentives that can arise in games with short-run players and Brownian motion type of monitoring.*

The intuition behind Proposition 3.5 is as follows. Fix a Nash equilibrium. As in Abreu, Milgrom, and Pearce (1991), when the period is very short, signals become too noisy and the statistical tests used to detect deviations give false positives too often. This means that punishment phases will be triggered too often on the equilibrium path, destroying too much value. In diffusion games, this effect is too extreme: the speed with which signal informativeness declines is so fast that intertemporal incentives totally collapse in the limit. The proof of the above proposition is omitted. It follows directly from Faingold and Sannikov's result for the limit game and the continuity argument of Section 4.2.

While the intuition behind this result is related to the phenomenon studied by Abreu, Milgrom and Pearce (1991), there is an important qualitative difference. Unlike in their paper, what drives the collapse of incentives in Proposition 3.5 is not a failure of statistical identification by the monitoring structure, but rather the presence of short-run players. That alone is not enough though. If signals arrive according to a Poisson process, as in Abreu, Milgrom and Pearce, the collapse of incentives may not happen. Hence, the above result is a product of both ingredients: the Brownian motion monitoring structure and the presence of short-run players. Also, it should be noted that, unless there is failure of statistical identification, the presence of short-run players is necessary. In fact, Sannikov (2005) shows that with two long-run players and Brownian motion monitoring, a very rich set of equilibria can be supported in the dynamic game.

A related collapse of incentives result was obtained in a recent paper by Sannikov and Skrzypacz (2005) on the failure of collusion in Cournot duopoly games with frequent decisions. In a setting with two long-run players in which the noise is driven by a Brownian motion, they show that collusion becomes impossible in the limit as the period length shrinks to zero. The speed with which signal informativeness deteriorates in the Brownian motion setting is central to their analysis.

4 Proof of the Main Result

The proof will have two parts. In the first part, in Section 4.1, I define the continuous-time limit game and prove a version of Theorem 3.1 for it. In the second part, in Section 4.2, a continuity result extends the payoff bounds obtained in Section 4.1 for the limit game to all games with sufficiently short periods.

Before turning to the limit game I state the following useful result, valid for each repeated game Γ_h.

Proposition 4.1. *Every Nash equilibrium in Γ_h is outcome-equivalent to a sequential equilibrium in public strategies.*

Proof. The proof involves two standard reductions and the details are omitted. First, by the full support assumption of Definition 3.1, the only (potentially) observable deviations are those of player 2. Because player 2 is short-run, the proof of Proposition 3 in Sekiguchi (1997) can be extended to game Γ_h,[21] yielding that every Nash equilibrium in Γ_h is outcome-equivalent to a sequential equilibrium.[22] Furthermore, because only one player is long-run, game Γ_h has, trivially, a *product structure* (Fudenberg and Levine (1994), Section 5). Hence, an immediate extension of Fudenberg and Levine's (1994) Theorem 5.2 to games with incomplete information can be invoked, proving that every sequential equilibrium is outcome equivalent to a sequential equilibrium in public strategies.[23] □

4.1 The Limit Game

In the limit game, Γ'_0, the type space (Θ, p) and the payoff structure $(A_i, \pi_i)_{i=1,2}$ are as in Γ_h, and the public signals $(X_t)_{t \geqq 0}$ follow the stochastic differential equation (3.2).

By Proposition 4.1, in each Γ_h, restricting attention to equilibria in public strategies does not restrict the set of equilibrium payoffs. Accordingly, in the limit game Γ'_0 only public strategies will be considered.

The *outcome set* is the measurable space $(\Omega, \mathcal{F})$, where $\Omega = \Theta \times D[0, \infty)$ and $D[0, \infty)$ is the canonical space of $\mathbb{R}^d$-valued right-continuous trajectories with left limits, endowed with the Skorohod

[21] In Sekiguchi's (1997) game there are two long-run players and *both* players' deviations are non-observable. However, if one of the players is short-run and recalls only the public histories, the assumption that the short-run player does not move the support of the signal distribution can be dispensed with.

[22] The definition of sequential equilibrium here is the standard extension of that of Kreps and Wilson (1982) to infinitely repeated games with finite action sets. Because of the full support assumption, the fact that the type space is potentially uncountable is not a problem. Since only the short-run players move the support of the signal distribution, the only information sets of player 2 that are off the path of play are those precluded by previous players 2. Since player 2 is a short-run player who conditions his play only on public information and, moreover, player 1 cannot affect the support of the signal distribution, the definition of player 2's beliefs on such info. sets is immaterial.

[23] Although Theorem 5.2 of Fudenberg and Levine (1994) assumes that the actions of the short-run player are public, with a single long-run player this assumption can be dispensed with and their proof remains valid *mutatis mutandi*.

topology (Billingsley (1968), Chapter 3). Here, $\mathcal{F}$ is the Borel σ-algebra on the product $\Omega = \Theta \times D[0,\infty)$. Let $\theta : \Omega \to \Theta$ denote the natural projection onto Θ and $X_t : \Omega \to \mathbb{R}^d$ the coordinate process. For each $t \geqq 0$, write $\mathcal{F}_t^1$ for the σ-algebra generated by $(\theta, X_s : 0 \leqq s \leqq t)$ and $\mathcal{F}_t^2$ for that generated by $(X_s : 0 \leqq s \leqq t)$. The *predictable σ-algebra of player $i = 1,2$*, $\mathcal{P}_i$, is the σ-algebra on $\Omega \times [0,\infty)$ generated by the rectangles

$$B \times (s,t], \text{ with } B \in \mathcal{F}_s^i \text{ and } 0 \leqq s \leqq t\,,$$

and $C \times \{0\}$, with $C \in \mathcal{F}_0^i$.

A behavior strategy of player $i = 1,2$ is a $\mathcal{P}_i$-measurable function $\sigma_i : \Omega \times [0,\infty) \to \Delta(A_i)$. This implies, in particular, that the time-t mixed action σ_{it} is measurable with respect to $\mathcal{F}_{t-}^i = \bigvee_{s<t} \mathcal{F}_s^i$ for all t.[24]

Fix a quadruplet of characteristics (b,γ,f,G). For a mixed action profile $\alpha \in \Delta(A)$ write $b(\alpha)$ and $f(\alpha,\cdot)$ for the bi-linear extensions of b and f. I will also use the notation $\gamma(\alpha)$, but the meaning will be different. For each $a \in A$, denote by $c(a) = \gamma(a)\gamma(a)^\top$ the *diffusion matrix* and by $c(\alpha)$ the bilinear extension of c to $\Delta(A)$. By the Kuratowsky-Ryll-Nardzewsky selection theorem (Aliprantis and Border (1999), Theorem 17.13), there is a measurable function $\overline{\gamma} : \Delta(A) \to \mathbb{R}^{d\times d}$ such that $\overline{\gamma}(\alpha)\overline{\gamma}(\alpha\,)^\top = c(\alpha)$ for all $\alpha \in \Delta(A)$ and $\overline{\gamma}|_A = \gamma$. With slight abuse of notation I will write $\gamma(\alpha)$ for $\overline{\gamma}(\alpha)$.

Definition 4.1. *The characteristics (b,γ,f,G) are called* regular *if for any prior probability p, and any behavior strategy profile σ, there exists a probability measure P on $(\Omega,\mathcal{F})$, and a filtration $(\mathcal{G}_t)_{t\geqq 0}$ (satisfying the usual conditions), such that θ is $\mathcal{G}_0$-measurable, (X_t) is adapted to $(\mathcal{G}_t)$, $\mathsf{P}\circ\theta^{-1} = p$, and*

$$X_t \;=\; \int_0^t b(\sigma_s)\,ds \;+\; \int_0^t \gamma(\sigma_s)\,dW_s \;+\; J_s\,, \quad \lambda \times \mathsf{P} - a.s.\,, \tag{4.1}$$

where, relative to $(\mathcal{G}_t)_{t\geqq 0}$, W is a Brownian motion in $\mathbb{R}^d$ and J is a pure-jump process with compensator $dt \times K(\sigma_t, d\xi) = dt \times f(\sigma_t,\xi)G(d\xi)$. Moreover, there is uniqueness in the sense of probability law, that is, for any $(\tilde{\mathsf{P}}, (\tilde{\mathcal{G}}_t), \tilde{W}, \tilde{J})$ satisfying the above, one has:[25]

$$\mathsf{P}\circ(\theta,X)^{-1} = \tilde{\mathsf{P}}\circ(\theta,X)^{-1}$$

This definition gives a meaning to equation (3.2) when players randomize continuously over time. The justification for this definition will be provided by the continuity result of Section 4.2.

Remark 4.1. It is well-known that allowing players to randomize continuously over time may pose a modeling problem: the realizations of the action paths may not be measurable functions of time. Here this problem does not arise, for the strategies are defined in terms of the public histories only and, moreover, the coefficients in equation (4.1) are averaged using the mixed actions σ_t directly, so there is no direct dependence on the realized action paths. (See Kushner (1990), Chapter 3, for a discussion about continuous-time mixtures in stochastic control problems.) □

[24] Predictability is slightly stronger a requirement than measurability with respect to $\mathcal{F}_{t-}^i$ for all t. The source of the difference is that predictability also embodies measurability with respect to the time parameter, which is necessary when working in continuous time.

[25] The choice of the measurable selection $\gamma(\alpha)$ is immaterial: the regularity of the characteristics (b,γ,f,G) depends only on b, c and K.

Given a strategy profile σ and a measurable set $\Theta' \subset \Theta$ with $p(\Theta') > 0$, let P denote the induced probability measure on $(\Omega, \mathcal{F})$ and $\mathsf{P}' = \mathsf{P}(\cdot|\theta \in \Theta')$, i.e., P' is the probability measure P *conditional* on the event that the type is in Θ'. The P-augmentation of player 2's filtration will be denoted $(\mathcal{F}_t)$.

Define an $(\mathcal{F}_t)$-predictable process D as follows:

$$D_t = \left\| \mathsf{E}(b(\sigma_t)|\mathcal{F}_{t-}) - \mathsf{E}'(b(\sigma_t)|\mathcal{F}_{t-}) \right\| + \int_\Xi \left\| \mathsf{E}(f(\sigma_t, \xi)|\mathcal{F}_{t-}) - \mathsf{E}'(f(\sigma_t, \xi)|\mathcal{F}_{t-}) \right\| G(d\xi),$$

where predictable versions of the conditional expectations above have been fixed. Denote by $\phi_t = \mathsf{P}(\theta \in \Theta'|\mathcal{F}_t)$ the martingale of posterior probabilities on the event that the type is in Θ' and write $\langle \phi \rangle$ for the *quadratic variation* of ϕ, i.e., the (almost surely) unique increasing process with $\langle \phi \rangle_0 = 0$ such that $\phi_t^2 - \langle \phi \rangle_t$ is a martingale.

The following proposition parallels the Fudenberg-Levine uniform merging result (Proposition 3.2). The proof below combines ideas found in Sorin (1999)'s proof of the Fudenberg-Levine merging result with tools from non-linear filtering of random processes.

Proposition 4.2. *For all $\varepsilon, \eta, \underline{\phi} > 0$ there exists $T > 0$ such that, given any prior p, any measurable set $\Theta' \subset \Theta$ with $p(\Theta') \geqq \underline{\phi}$, and any behavior strategy profile σ, one has*

$$\mathsf{P}'[\lambda \{t \in \mathbb{R}_+ : D_t \geqq \varepsilon\} \geqq T] \leqq \eta,$$

where λ denotes the Lebesgue measure on $\mathbb{R}_+$.

Proof. In the spirit of Sorin (1999)'s proof of the Fudenberg-Levine merging result, consider the quadratic variation $\langle \phi \rangle$ of the martingale of posteriors. Since $\phi_t^2 - \langle \phi \rangle_t$ is a P-martingale, we have $\mathsf{E}[\phi_t^2 - \langle \phi \rangle_t] = \phi_0^2$, and hence $\mathsf{E}\langle \phi \rangle_t = \mathsf{E}\phi_t^2 - \phi_0^2 \leqq \mathsf{E}\phi_t - \phi_0^2 = \phi_0(1 - \phi_0)$. Therefore, one has for all $t \geqq 0$,

$$\mathsf{E}\langle \phi \rangle_t \leqq 1, \tag{4.2}$$

which holds for any prior p, measurable set Θ' with $p(\Theta') > 0$, and strategy profile σ. Also, by Corollary B.1 in Appendix B, there is a constant $C > 0$, which depends only on the characteristics (b, γ, f, G), such that for any prior p, measurable set Θ' with $p(\Theta') > 0$, and strategy profile σ,

$$\langle \phi \rangle_t \geqq C \int_0^t \phi_{s-}^2 D_s^2 \, ds \quad \mathsf{P}-\text{a.s.}. \tag{4.3}$$

Hence, from the inequalities (4.2) and (4.3) above, one has for all $\psi > 0$,

$$\begin{aligned} \mathsf{E}[\lambda \{t \geqq 0 : \phi_{t-} D_t > \psi\}] &= \mathsf{E} \int_{\{t : \phi_{t-} D_t > \psi\}} ds \\ &\leqq \psi^{-2} \mathsf{E} \int_{\{t : \phi_{t-}^2 D_t^2 > \psi^2\}} \phi_{s-}^2 D_s^2 \, ds \\ &\leqq \psi^{-2} \mathsf{E} \int_0^\infty \phi_{s-}^2 D_s^2 \, ds \\ &\leqq C^{-1} \psi^{-2}, \end{aligned}$$

which yields, for all $T > 0$,

$$\mathsf{P}[\lambda \{t \geqq 0 : \phi_{t-} D_t > \psi\} \geqq T] \leqq \frac{1}{C \psi^2 T}.$$

Therefore, for all $T > 0$,

$$
\begin{aligned}
\mathsf{P}'[\lambda\{t \geqq 0 : D_t > \varepsilon\} \geqq T] &\leqq \mathsf{P}'\left[\{\inf_{t\geqq 0}\phi_t > \eta\phi_0/2\} \cap \{\lambda[t : D_t > \varepsilon] \geqq T\}\right] + \\
&\qquad + \mathsf{P}'\left[\inf_{t\geqq 0}\phi_t \leqq \eta\phi_0/2\right] \\
&\leqq \mathsf{P}'[\lambda\{t : \phi_{t-}D_t > \varepsilon\eta\phi_0/2\} \geqq T] + \eta/2 \\
&\leqq \frac{1}{\phi_0}\mathsf{P}[\lambda\{t : \phi_{t-}D_t > \varepsilon\eta\phi_0/2\} \geqq T] + \eta/2 \\
&\leqq \frac{4}{C\phi_0^3\varepsilon^2\eta^2 T} + \eta/2 \leqq \frac{4}{C\underline{\phi}^3\varepsilon^2\eta^2 T} + \eta/2\,,
\end{aligned}
$$

where the second inequality follows from Lemma B.2 in the Appendix and the third inequality follows from $\mathsf{P} = \phi_0\mathsf{P}' + (1-\phi_0)\mathsf{P}''$, where $\mathsf{P}'' = \mathsf{P}_{p,\sigma}(\cdot|\{\theta \notin \Theta'\})$. Choosing $T = 8/(C\underline{\phi}^3\varepsilon^2\eta^3)$ yields the desired result. □

The above proposition is the main step in the proof of the continuos-time reputation result (Theorem 4.1 below). In fact, it is the only step that is substantially different from the proof of Proposition 3.1. The main result of this section is the following:

Theorem 4.1. *Fix the prior $p \in \Delta(\Theta)$. If $\theta_0 \in \Theta$ is such that $p(\theta_0) > 0$, then*

$$
\underline{\pi}'_1(\theta_0) \leqq \liminf_{r\to 0} \underline{\mathcal{E}}_1(r,\theta_0) \leqq \limsup_{r\to 0} \overline{\mathcal{E}}_1(r,\theta_0) \leqq \overline{\pi}'_1(\theta_0)\,.
$$

Moreover, if the stage-game is non-degenerate and identified, then $\underline{\pi}'_1(\theta_0) = \overline{\pi}'_1(\theta_0)$ and, consequently,

$$
\lim_{r\to 0} \mathcal{E}_1(r,\theta_0) = \{\underline{\pi}'_1(\theta_0)\}\,.
$$

Proof. Given Proposition 4.2 above, the proof of the first part replicates, *mutatis mutandi*, the proofs of Theorem 3.1 and Corollary 3.2 of Fudenberg and Levine (1992). As for the second part, the proof follows that of Theorem 3.3 of Fudenberg and Levine (1992). □

4.2 Convergence

In this section, I define a suitable topology on the players' strategy spaces in Γ'_0 that will allow, via an upper hemi-continuity argument, the payoff bounds from Theorem 4.1 to be extended to all games Γ_h with h close enough to zero.

A behavior strategy $\sigma_i : \Omega\times[0,\infty) \to \Delta(A_i)$ of player i is *history-independent* if it does not depend on $\omega \in \Omega$. Let Σ_i^* denote the space of history-independent strategies of player i. The space Σ_i^* will be endowed with the *compact-weak topology* (see Kushner (1990), Chapter 3). A sequence $(\sigma_i^\nu)_{\nu\geqq 1}$ in Σ_i^* *weakly converges to* $\sigma_i \in \Sigma^*$ if for every continuous and bounded function $g : A\times[0,\infty) \to \mathbb{R}$ one has for all $T > 0$:

$$
\int_0^T g(\sigma_{it}^\nu, t)\,dt \to \int_0^T g(\sigma_{it}, t)\,dt\,,
$$

where we write $\Delta(A_i) \ni \alpha_i \mapsto g(\alpha_i, t)$ for the linear extension of $g(\cdot, t)$ to $\Delta(A_i)$.

To illustrate this definition consider the following example. The action set is $A_i = \{T, B\}$ and the sequence $(\sigma_i^\nu)_{\nu \geqq 1}$ is such that

$$\sigma_i^\nu = \begin{cases} T & : \quad t \in [j/\nu, (j+1)/\nu)\,, \text{for } j \text{ even} \\ B & : \quad t \in [j/\nu, (j+1)/\nu)\,, \text{for } j \text{ odd} \end{cases}$$

Because of the oscillatory behavior, this sequence of history-independent strategies does not converge in any classical sense. However, it weakly converges to the strategy σ_i that prescribes the $1/2-1/2$ mixture between T and B at every time t (recall Remark 4.1). Notice that, unlike in discrete-time settings, here a sequence of pure strategy may converge to a completely mixed one.

To relate this topology to a more familiar one note that for each T, σ_i induces a probability measure $m_T^{\sigma_i}$ over the product $A_i \times [0, T]$, namely

$$m_T^{\sigma_i}(\{a_i\} \times [s, t]) = \frac{1}{T} \int_s^t \sigma_{i t'}(a_i)\, dt' \,, \quad \text{for } a_i \in A_i \text{ and } 0 \leqq s \leqq t\,.$$

The probability measure $m_T^{\sigma_i}$ gives the "fraction of time" spent by the strategy σ_i on a specific action.

Hence, the compact-weak topology on Σ_i^* is simply the topology of weak convergence of the probability measures $m_T^{\sigma_i}$, for every $T > 0$. In particular, the compact-weak topology on Σ_i^* is metrizable by the following metric:

$$d(\sigma_i, \sigma_i') = \sum_{n=1}^{\infty} 2^{-n}\, r_n(m_n^{\sigma_i}, m_n^{\sigma_i'})\,,$$

where r_n is the Prohorov metric on the space of probability measures over $A_i \times [0, n]$.

Since $A_i \times [0, T]$ is compact, the set of probability measures over $A_i \times [0, T]$ is *tight* for each $T \geqq 0$. By Prohorov's Theorem (Billingsley (1968), Chapter 1, Theorem 6.1), Σ_i^* is a compact metric space. Write Σ^* for the Cartesian product $\Sigma_1^* \times \Sigma_2^*$ endowed with the product topology.

The next step is to define a topology on the space of all behavior strategy profiles, history-independent or otherwise. First, consider the limit game and denote the space of behavior strategy profiles in Γ_0' by Σ^0. Each strategy profile $\sigma \in \Sigma^0$ can be viewed as a mapping $\sigma : \Omega \to \Sigma^*$. This is the viewpoint I will adopt from here on.

Fix the prior probability p over the type space Θ. Each strategy profile $\sigma \in \Sigma_0$ induces, via equation (3.2), a probability measure P_σ^0 over $(\Omega, \mathcal{F})$. The mapping $\sigma : \Omega \to \Sigma^*$ induces a probability measure over Σ^*, namely $\mathsf{P}_\sigma^0 \circ \sigma^{-1}$, where the compact metric space Σ^* is endowed with the measurable structure induced by the metric d above. Every strategy profile $\sigma \in \Sigma^0$ will be *identified* with the probability measure $\mathsf{P}_\sigma^0 \circ \sigma^{-1}$ over history-independent strategy profiles.

Now I turn to the strategy profiles in the games with period length $h > 0$. Recall the definition of the family of games Γ_h' from Section 3.1. In Γ_h', the signals are given by a continuous-time controlled Markov chain $(x_t')_{t \geqq 0}$, which is the continuous-time interpolation of the discrete-time chain $(x^n)_{n \geqq 0}$, where $x^0 = 0$, $x^n = \sum_{k=0}^{n-1} y^k$ and y^k is the period-k signal from the discrete-time game Γ_h.

Denote the space of strategy profiles in Γ_h' by Σ^h. A strategy profile $\sigma \in \Sigma^h$ induces a probability measure P_σ^h over the outcome space $(\Omega, \mathcal{F})$: the probability law of the Markov chain (x_t') under σ^h. Analogously to the case $h = 0$, I will identify the strategy profile σ with the probability measure $\mathsf{P}_\sigma^h \circ \sigma^{-1}$ over Σ^*.

Definition 4.2. *Let $\sigma \in \Sigma$ be a limit-game strategy profile and $(\sigma^\nu)_{\nu \geqq 1}$ a sequence of strategy profiles with $\sigma^\nu \in \Sigma^{h(\nu)}$ for all $\nu \geqq 1$, where $h(\nu)$ is a sequence converging to 0. The sequence (σ^ν)* converges to σ *(and we write $\sigma^\nu \to \sigma$) if*

$$\mathsf{P}^{h(\nu)}_{\sigma^\nu} \circ (\sigma^\nu)^{-1} \longrightarrow \mathsf{P}^0_\sigma \circ \sigma^{-1} ,$$

where the above convergence is in the sense of weak convergence of probability measures over the metric space Σ^*.

The following result is the basis of the continuity argument.

Proposition 4.3. *Let $(\sigma^\nu)_{\nu \geqq 1}$ be an arbitrary sequence of strategy profiles with $\sigma^\nu \in \Sigma^{h(\nu)}$ for all $\nu \geqq 1$, where $h(\nu)$ is a sequence converging to 0. Then:*

(a) A subsequence $(\sigma^{\nu'})_{\nu'}$ exists which converges to some $\sigma \in \Sigma^0$.

(b) Suppose that $\sigma^\nu \to \sigma \in \Sigma^0$. Given any $\hat{\sigma}_i \in \Sigma_i$, there exists a sequence $\hat{\sigma}^\nu_i \in \Sigma^{h(\nu)}_i$ such that $\hat{\sigma}^\nu \to \hat{\sigma}$, where $\hat{\sigma}^\nu = (\hat{\sigma}^\nu_i, \sigma^\nu_{-i})$ and $\hat{\sigma} = (\hat{\sigma}_i, \sigma_{-i})$.

Proof idea. The proof is technical and the reader is referred to Chapters 10 and 13 of Kushner and Dupuis (2001) (Section 10.4) for detailed arguments. I will sketch the logic of the proof. For part (a), notice that, by the tightness of the probability measures $\mathsf{P}^{h(\nu)}_{\sigma^\nu} \circ (\sigma^\nu)^{-1}$, there exists a subsequence $(\sigma^{\nu'})$ converging to some σ. However, there is no guarantee, a priori, that σ satisfies the appropriate measurability conditions of Section 4.1. To establish this fact, one exploits the consistency conditions of Definition 3.1 and criteria for tightness in $D[0,\infty)$, proceeding in three steps: First, one shows that the probability measures $(\mathsf{P}^{h(\nu')}_{\sigma^{\nu'}})_{\nu'}$, over $(\Omega, \mathcal{F})$, are tight. Second, one shows that the limit of every convergent subsequence corresponds to the law of the process $(X_t)_{t \geqq 0}$ under σ with a possibly augmented filtration. Finally, one argues from the characterization of the limit process that σ must satisfy the appropriate non-anticipativity conditions of Section 4.1. As for part (b) this is a standard construction in Control Theory, known as the Chattering Lemma. (See Theorem 5.2 of Kushner (1990)). □

Fix a family $(\Gamma_h)_{h>0}$ of games with persistently imperfect monitoring game and, for each $h > 0$, let $\mathcal{E}_1(h)$ denote the set of Nash equilibrium payoffs of the long-run player in the repeated game Γ_h. In addition, $\mathcal{E}_1(0)$ will denote the set of Nash equilibrium payoffs of the long-run player in the limit game Γ'_0. The continuity result below follows directly from the above proposition.

Proposition 4.4. *The correspondence $\mathcal{E}_1$ is upper hemi-continuous.*

Proof. Let $(h(\nu))_{\nu \geqq 1}$ be a sequence of period lengths converging to zero and (v^ν_1) a sequence of payoffs, with $v^\nu_1 \in \mathcal{E}_1(h^\nu)$ for all $\nu \geqq 1$, which converges to some $v_1 \in \mathbb{R}$. For each $\nu \geqq 1$, choose a Nash equilibrium profile $\sigma^\nu \in \Sigma^{h(\nu)}$ of the game $\Gamma'_{h(\nu)}$ that yields the payoff v^ν_1. By part (a) of the Proposition above, there is a subsequence, still denoted (σ^ν), converging to some $\sigma \in \Sigma^0$.

I claim that σ yields an expected payoff of v_1 to the long-run player. Fix $\epsilon > 0$ and choose $T > 0$ such that $2e^{-rT} M < \varepsilon$ where M is an upper bound on expected flow payoffs. Hence,

$$\left| v_1 - \mathsf{E}^0_\sigma \, r \int_0^\infty e^{-rt} \pi_1(\sigma_t) \, dt \right| \leqq \left| v_1 - v^\nu_1 \right| + \left| v^\nu_1 - \mathsf{E}^0_\sigma \, r \int_0^\infty e^{-rt} \pi_1(\sigma_t) \, dt \right| \leqq$$

$$\leqq \left| v_1 - v^\nu_1 \right| + \left| \mathsf{E}^{h(\nu)}_{\sigma^\nu} \, r \int_0^T e^{-rt} \pi_1(\sigma^\nu_t) \, dt - \mathsf{E}^0_\sigma \, r \int_0^T e^{-rt} \pi_1(\sigma_t) \, dt \right| + \varepsilon$$

By $v_1^\nu \to v_1$ and the definition of the convergence $\sigma^\nu \to \sigma$, letting $\nu \to$ yields:

$$\left| v_1 - \mathsf{E}_\sigma^0 \, r \int_0^\infty e^{-rt} \pi_1(\sigma_t)\, dt \right| \leqq \varepsilon \, .$$

Since ε is arbitrary, this proves the claim.

Let $\hat{\sigma}_1 \in \Sigma^0$ be an arbitrary strategy of player 1. By part (b) of the Proposition above, there exists a sequence $\hat{\sigma}_1^\nu \in \Sigma_1^{h(\nu)}$ such that $\hat{\sigma}^\nu \to \hat{\sigma}$, with the same notation as in the statement of Proposition 4.4. Since for each $\nu \geqq 1$, σ^ν is a Nash equilibrium of $\Gamma'_{h(\nu)}$, one has for all $T > 0$ and $\nu \geqq 1$:

$$\mathsf{E}_{\sigma^\nu}^{h(\nu)} \, r \int_0^T e^{-rt} \pi_1(\sigma_t^\nu)\, dt \quad \geqq \quad \mathsf{E}_{\hat{\sigma}^\nu}^{h(\nu)} \, r \int_0^T e^{-rt} \pi_1(\hat{\sigma}_t^\nu)\, dt - 2Me^{-rT}$$

Letting $\nu \to \infty$ first and then $T \to \infty$ yields:

$$v_1 = \mathsf{E}_\sigma^0 \, r \int_0^\infty e^{-rt} \pi_1(\sigma_t)\, dt \quad \geqq \quad \mathsf{E}_{\hat{\sigma}}^0 \, r \int_0^\infty e^{-rt} \pi_1(\hat{\sigma}_t)\, dt$$

Since $\hat{\sigma}_1$ is arbitrary, I have shown that σ_1 is a best reply for the long-run player.

To show that σ_2 is a best reply for player 2, fix $t \geqq 0$ and pick an arbitrary strategy $\hat{\sigma}_2 \in \Sigma_2$ with $\hat{\sigma}_{2s} = \sigma_{2s}$ for all $s \leqq t$. Since player 2 is myopic, one has for all $t' > t$ and $\nu \geqq 1$:

$$\mathsf{E}_{\sigma^\nu}^{h(\nu)} \int_t^{t'} \pi_2(\sigma_s^\nu)\, ds \quad \geqq \quad \mathsf{E}_{\hat{\sigma}^\nu}^{h(\nu)} \int_t^{t'} \pi_2(\hat{\sigma}_s^\nu)\, ds$$

Letting $\nu \to \infty$ yields:

$$\mathsf{E}_\sigma^0 \int_t^{t'} \pi_2(\sigma_s)\, ds \quad \geqq \quad \mathsf{E}_{\hat{\sigma}}^0 \int_t^{t'} \pi_2(\hat{\sigma}_s)\, ds \, .$$

Since the above inequality holds for all $t' > t$, σ_2 must be myopically superior than $\hat{\sigma}_2$, which concludes the proof. □

In contrast to the proposition above, in the *perfect monitoring* continuous-time games studied by Simon and Stinchcombe (1987) upper hemi-continuity fails to hold. The above result is closer in spirit to the upper hemi-continuity found in Bergin and MacLeod (1993), though in the latter the strategy sets fail to be compact.

I conclude with the proof of Theorem 3.1.

Proof of Theorem 3.1. Fix $\varepsilon > 0$ and choose $r > 0$ low enough so that the bounds of Theorem 4.1 hold for the limit game with discount rate $r > 0$ with an ε slack. By the upper hemi-continuity above, they must hold for Γ_h, provided h is close enough to zero. □

Appendices

A Proof of Proposition 2.2

The proof proceeds in three steps, outlined below:

(a) Study the complete information continuous-time limit game and show that $2-\kappa$ is an upper bound for the set of Nash equilibrium payoffs for the firm, where $\kappa = \frac{\mu(\overline{e})}{\mu(\overline{e})-\mu(\underline{e})}$.

(b) Argue that, with a fixed discount rate, the Nash equilibrium payoff correspondence (of the limit game) is upper hemi-continuous with respect to the prior at $\phi_0 = 0$. Conclude that a critical belief level $\overline{\phi} > 0$ exists, such that if the prior probability ϕ_0 on the commitment type is below $\overline{\phi}$, then, in the incomplete information game with prior ϕ_0, the Nash equilibrium payoffs of the normal type of monopolist are bounded above by $2-\kappa+\varepsilon$.

(c) Fix $\phi_0 < \overline{\phi}$. And consider the games with positive period length h now. Argue that the Nash equilibrium payoff correspondence is upper hemi-continuous with respect to h. Conclude that if h is close enough to zero, then the Nash equilibrium payoffs of the normal type are bounded above by $2-\kappa+2\varepsilon$.

As for (a), an explanation is in order. Note that the $2-\kappa$ bound can be derived directly by applying Abreu, Milgrom and Pearce's argument (adapted to the case in which player 2 is short-run), which does not require the study of the limit game. This would show that $2-\kappa$ is a payoff bound on the *limit of equilibria*, as opposed to a bound on the *equilibria of the limit-game*. This approach, however, is not sufficient here. A priori, the Nash equilibrium correspondence is only upper hemi-continuous with respect to the period length h and there is, in principle, no guarantee that we could find the bound on ϕ_0 independent of h, if we proceeded as in Abreu, Milgrom and Pearce.[26] The study of the limit game is what gives the uniformity.

Given the topology on strategy sets introduced in Section 4.1, part (b) follows a standard argument: Consider an arbitrary sequence of priors converging to zero and a corresponding sequence of Nash equilibrium payoffs of the firm. Associated to the sequence of Nash equilibrium payoffs is a sequence of equilibrium profiles. Extract a convergent subsequence of the latter, in the sense defined in Section 4.1. Since each term of this sequence is a Nash equilibrium and the mapping $(\phi,\sigma) \mapsto \mathsf{P}_{\phi,\sigma}$ is continuous in the topology of weak convergence of probability measures, the limit profile must be a Nash equilibrium of the complete information game. As for part (c), it is a particular case of Proposition 4.4.

Only part (a) remains to be proven. The proof will be carried in several steps.

Lemma A.1. *Fix a strategy profile σ of the limit game and let $(V_t)_{t\geqq 0}$ denote the corresponding continuation-value process of the firm. Then, there is a predictable process $(\beta_t)_{t\geqq 0}$ such that:*

$$dV_t = r(V_t - \pi(\sigma_t))\,dt + \beta_t\,(dN_t - \mu(\sigma_{1t})\,dt)$$

Proof. Let $\overline{V}_t$ denote the average discounted payoff of the monopolist conditional on $\mathcal{F}_t$, i.e.,

$$\overline{V}_t = \mathsf{E}\left[r\int_0^\infty e^{-rs}\pi_1(\sigma_s)\,ds \;\middle|\; \mathcal{F}_t \right]$$

Since $\overline{V}$ is a bounded $(\mathcal{F}_t)$-martingale, it has a representation of the form:[27]

$$\overline{V}_t = \overline{V}_0 + \int_0^t re^{-rs}\beta_s\,d\overline{N}_s\,, \tag{A.1}$$

[26] It is an *implication* of the study of the limit game that the Nash equilibrium payoff correspondence is also lower hemi-continuous, but this cannot be assumed a priori.

[27] This step follows from the representation theorem for martingales adapted to counting processes (Bremaud (1981), Theorem 9, pg. 64).

where the martingale $\bar{N}$ is the *compensated counting process*, defined as $\bar{N}_t = N_t - \int_0^t \mu(\sigma_{1s})\, ds$, and the predictable process β is given by $\beta_t = r^{-1} e^{rt} \frac{d}{dt} \langle \bar{V}, \bar{N} \rangle_t$, with the angle brackets $\langle \cdot, \cdot \rangle$ denoting the *cross-variation* of square-integrable martingales.

From the definition of the continuation-value process $(V_t)_{t \geqq 0}$ it follows that

$$\bar{V}_t = r \int_0^t e^{-rs} \pi_1(\sigma_s)\, ds \ + \ e^{-rt} V_t \,,$$

which, in differential form, is:

$$d\bar{V}_t = re^{-rt} \pi_1(\sigma_t)\, dt \ - \ re^{-rt} V_t\, dt \ + \ e^{-rt} dV_t \tag{A.2}$$

A direct comparison of equation (A.2) and the differential form of equation (A.1) yields the desired result. □

The proposition below gives a characterization, in terms of the basic data of the game, of the incentive constraints in public perfect equilibria. It is analogous to Proposition 3 in Faingold and Sannikov (2005). I include a proof for completeness.

Lemma A.2. *A public-strategy profile $(\sigma_t)_{t \geqq 0}$ is a public perfect equilibrium if, and only if, for almost every time $t \geqq 0$ and after almost every public history:*

$$\sigma_{1t}[e] > 0 \quad \Rightarrow \quad e \ \in \ \underset{e' \in \{\bar{e}, \underline{e}\}}{\arg\max} \ \pi_1(e', \sigma_{2t}) \ + \ \beta_t\, \mu(e', \sigma_{2t}) \,, \tag{A.3}$$

$$\sigma_{2t}[s] > 0 \quad \Rightarrow \quad s \ \in \ \underset{s' \in \{\bar{s}, \underline{s}\}}{\arg\max} \ u(\sigma_{1t}, s') \,, \tag{A.4}$$

where β is the predictable process of Lemma A.1.

Proof. Fix a strategy profile σ. Denote by (V_t) the continuation-value of the long-run player under σ and by β the predictable process of Lemma A.1. Let σ_1' be an arbitrary strategy of the long-run player. Fix a time $t \geqq 0$ and consider the strategy ${}^t[\sigma_1']$ which plays like σ' up to t and like σ_1 from time t onwards. The overall expected payoff of the firm from strategy profile $({}^t[\sigma_1']_s, \sigma_{2s})_{s \geqq 0}$, conditional on $\mathcal{F}_t$, is given by:

$$\bar{V}_t' = r \int_0^t e^{-rs} \pi_1(\sigma_{1s}', \bar{b}_s)\, ds \ + \ e^{-rt} V_t$$

By Proposition A.1 and the expression above,

$$\begin{aligned} d\bar{V}_t' \ &= \ re^{-rt} \left(\pi_1(\sigma_{1t}', \sigma_{2t}) - V_t \right) dt \ + \ e^{-rt} dV_t \\ &= \ re^{-rt} \left[\left(\pi_1(\sigma_{1t}', \sigma_{2t}) - \pi_1(\sigma_t) \right) dt + \beta_t \cdot d\bar{N}_t \right] . \end{aligned}$$

Hence the profile (σ_1', σ_2) yields the firm the following expected payoff:

$$\begin{aligned} V_0' \ &= \ E[\bar{V}_\infty'] \ = \ E\left[\bar{V}_0' \ + \ \int_0^\infty d\bar{V}_t' \right] \\ &= \ V_0 \ + \ E\left[r \int_0^\infty e^{-rt} \left(\pi_1(\sigma_{1t}', \sigma_{2t}) - \pi_1(\sigma_t) \ + \ \beta_t \cdot (\mu(\sigma_{1t}', \sigma_{2t}) - \mu(\sigma_t) \right) dt \right] , \end{aligned}$$

where I used the fact that $\bar{V}_0' = V_0$ and that $N_t - \int_0^t \mu(\sigma_{1s}', \sigma_{2s})\, ds$ is a martingale under (σ_1', σ_2).

Suppose that strategy profile σ satisfies conditions (A.3) and (A.4). Then, for every σ_1', one has $V_0 \geqq V_0'$, and the firm is sequentially rational at time 0. A slight modification of the argument above proves that the firm is sequentially rational at (almost) all times t and after (almost) all public histories. The consumers are maximizing their instantaneous expected payoffs, therefore each consumer is also sequentially rational and σ is a PPE.

Conversely, suppose that condition (A.3) fails at a positive measure time-history set. Choose a strategy (σ_{1t}') such that σ_{1t}' attains the maximum in (A.3) for almost all $t \geqq 0$ and after almost every public history. Then $V_0' > V_0$ and the firm is not sequentially rational. Likewise, if condition (A.4) fails, then a positive measure of consumers is not maximizing their instantaneous expected payoffs. Since the consumers are small anonymous players, a positive measure of small players is not sequentially rational. □

Lemma A.3. *Consider the auxiliary one-shot game below, which depends on a parameter $\beta \in \mathbb{R}$:*

	$\bar{\mathfrak{z}}$	$\underline{\mathfrak{z}}$
$\bar{e}$	$2 + \beta\mu(\bar{e}), 3$	$\beta\mu(\bar{e}), 2$
$\underline{e}$	$3 + \beta\mu(\underline{e}), 0$	$1 + \beta\mu(\underline{e}), 1$

If $\beta \leqq 1/(\mu(\underline{e})-\mu(\bar{e}))$, the maximum Nash equilibrium payoff of the firm is no greater than $2-\frac{\mu(\bar{e})}{\mu(\underline{e})-\mu(\bar{e})}$.

Proof. When $\beta > -1/(\mu(\underline{e}) - \mu(\bar{e}))$, $\underline{e}$ is a strictly dominant strategy for the firm and the only Nash equilibrium of this game yields the firm a payoff of $1 + \beta\mu(\underline{e})$, which is no greater than $2 - \kappa$ since $\beta \leqq 1/(\mu(\underline{e}) - \mu(\bar{e}))$. If $\beta \leqq -1/(\mu(\underline{e}) - \mu(\bar{e}))$, then $(\bar{e}, \bar{\mathfrak{z}})$ is a Nash equilibrium which yields the firm a payoff of $2 + \beta\mu(\bar{e}) < 2 - \frac{\mu(\bar{e})}{\mu(\underline{e})-\mu(\bar{e})}$. □

Proposition A.1. *In every public perfect equilibrium, the firm receives a payoff no greater than $2 - \kappa$, where $\kappa = \mu(\underline{e})/(\mu(\underline{e}) - \mu(\bar{e}))$.*

Proof. Fix a public perfect equilibrium σ and consider the associated continuation-value process V. By Lemma A.1, for some predictable process β_t, one has:

$$dV_t = r(V_t - \tilde{\pi}_{1t})\, dt + \beta_t\, dN_t\,,$$

where $\tilde{\pi}_{1t} = \pi_1(\sigma_t) + \beta_t\, \mu(\sigma_{1t})$. Moreover, by Lemma A.2, $\tilde{\pi}_{1t}$ is a Nash equilibrium payoff of the one-shot game of Lemma A.3, for all $t \geqq 0$.

Let $\bar{v}$ be the highest Nash equilibrium payoff of the firm in the dynamic game. From the equation above, it follows that whenever $V_t = \bar{v}$, necessarily $\beta_t \leqq 0$ and $V_t - \tilde{\pi}_{1t} \leqq 0$, for otherwise the value process would escape above $\bar{v}$ with positive probability, a contradiction. Hence, by Lemma A.3, $\bar{v} \leqq \tilde{\pi}_{1t} \leqq 2 - \kappa$. □

B Proofs for Section 4.1

The following Proposition gives a useful characterization of the Full Support condition purely in terms of the characteristics (b, γ, f, G).

Proposition B.1. *Let (b, γ, f, G) be a quadruplet of regular characteristics. The following statements are equivalent:*

(a) For every pair of constant pure-strategy profiles σ' and σ'', with $\sigma'_2 = \sigma''_2$, the probability measures induced by σ' and σ'' are mutually locally absolutely continuous.

(b) The quadruplet (b, γ, f, G) satisfies the following conditions:

(i) $\gamma(a)$ is independent of a_1,

(ii) For all $a', a'' \in A$ with $a'_2 = a''_2$, one has:

$$b(a') - b(a'') \in \text{span}\; c(a'),$$

with $c(a') = \gamma(a')\gamma(a')^\top$,

(iii) The support of $f(a, \cdot)$ is independent of a_1.

(c) For every pair of behavior strategy profiles σ' and σ'', with $\sigma'_2 = \sigma''_2$, the probability measures induced by σ' and σ'' are mutually locally absolutely continuous.

Proof. (a)⇒(b): Given the assumption in (3.3), this implication follows directly from Girsanov's Theorem for Semimartingales (Jacod and Shiryaev (2002), Theorem 3.24, pg. 172). (b)⇒(c): The regularity of (b, γ, f, G), the boundedness of f and the assumption in (3.3) imply that the conditions of Theorem 5.34 on pg. 200 of Jacod and Shiryaev (2002) are satisfied, which yields the desired result. (c)⇒(a): Trivial. □

For the next lemma, we need the following definition.

Definition B.1. *An* integer-valued random measure *is a family $\mathcal{J}(\omega; dt \times d\xi)_{\omega\in\Omega}$ of positive measures on $(\mathbb{R}_+ \times \Xi, \mathcal{R}_+ \otimes \mathcal{R}^d)$ such that:*

(a) $\mathcal{J}(\cdot; \{0\} \times \Xi) = 0$ and $\mathcal{J}(\cdot, \{t\} \times \Xi) \leqq 1$ for all $t \in \mathbb{R}_+$,

(b) for each $B \in \mathcal{R}_+ \otimes \mathcal{R}^d$, $\mathcal{J}(\cdot; B)$ takes values in the set of non-negative integers, and

(c) for each progressively measurable function $(\omega, t, \xi) \mapsto H_t(\omega, \xi)$, the process $H * \mathcal{J}$, defined as

$$(H * \mathcal{J})_t(\omega) = \int_{[0,t]\times\Xi} H_s(\omega, \xi)\mathcal{J}(\omega; ds \times d\xi),$$

is progressively measurable.[28]

The reader is referred to Chapter II of Jacod and Shiryaev (2002) for details about integer-valued random measures.

If J is a pure-jump process then the random measure $\mathcal{J}$, defined by:

$$\mathcal{J}([0,t] \times B) = \sum_{s\in[0,t]} I_{\{\Delta J_s \in B\setminus\{0\}\}},$$

is an integer-valued random measure, where $\Delta J_s = J_s - J_{s-}$. Moreover, by construction,

$$J_t = \int_{[0,t]\times\Xi} \xi\, \mathcal{J}(ds, d\xi).$$

[28]Here *progressively measurable* is with respect to the filtration $(\mathcal{G}_t \otimes \mathcal{B}(\Xi))_{t\geqq 0}$, where $\mathcal{B}(\Xi)$ is the Borel σ-algebra on Ξ.

The *compensator* of an integer-valued random, relative to a probability measure Q, is the (almost surely unique) predictable random measure $\nu(\omega; dt, d\xi)$ such that:

$$\mathsf{E}_{\mathsf{Q}}(\int_{[0,t]\times\Xi} H_s(\xi)\, \mathcal{J}(ds, d\xi)) = \mathsf{E}_{\mathsf{Q}}(\int_{[0,t]\times\Xi} H_s(\xi)\, \nu(ds, d\xi))$$

for every non-negative predictable process H.

Lemma B.1. *Given any prior p, measurable set $\Theta' \subset \Theta$ with $p(\Theta') > 0$, and behavior strategy profile σ, the martingale of posteriors $\phi_t = \mathsf{P}(\theta \in \Theta'|\mathcal{F}_t)$ satisfies:*

$$\phi_t = \phi_0 + \int_0^t \phi_{s-}\, \varrho_s \cdot dM_s + \int_{[0,t]\times\Xi} \phi_{s-}\vartheta_s(\xi) \left(\mathcal{J}(ds\times d\xi) - ds\times \bar{K}_s(d\xi)\right) ,$$

where, relative to P and $(\mathcal{F}_t)$, M is a continuous square-integrable martingale with quadratic variation $\langle M\rangle_t = \int_0^t c(\sigma_s)\, ds$, $\mathcal{J}$ is an integer-valued random measure with predictable compensator $dt\times\bar{K}_t(d\xi)$, with $\bar{K}_t(d\xi) = \mathsf{E}(f(\sigma_t)|\mathcal{F}_{t-})G(d\xi)$, and ϱ, ϑ are predictable processes such that:

$$\begin{aligned} c(\sigma_t)\cdot\varrho_t &= \mathsf{E}'\left(b(\sigma_t)|\mathcal{F}_{t-}\right) - \mathsf{E}\left(b(\sigma_t)|\mathcal{F}_{t-}\right) \\ \vartheta_t(\xi) &= \frac{\mathsf{E}'\left(f(\sigma_t,\xi)|\mathcal{F}_{t-}\right) - \mathsf{E}\left(f(\sigma_t,\xi)|\mathcal{F}_{t-}\right)}{\mathsf{E}\left(f(\sigma_t,\xi)|\mathcal{F}_{t-}\right)} . \end{aligned}$$

Proof. Recall that $\mathsf{P}' = \mathsf{P}(\cdot|\theta \in \Theta)$ and $\mathsf{P}'' = \mathsf{P}(\cdot|\theta \notin \Theta)$. By the Full Support condition, P' is locally absolutely continuous with respect to P''. Let $(Z_t)_{t\geqq 0}$ denote the *density process*, i.e., Z_t is the Radon-Nikodym derivative of $\mathsf{P}'|_{\mathcal{F}_t}$ w.r.t. $\mathsf{P}''|_{\mathcal{F}_t}$. Bayes' formula yields:

$$\phi_t = \frac{\phi_0 Z_t}{\phi_0 Z_t + 1 - \phi_0} \quad \mathsf{P}-\text{a.s.}.$$

By the regularity of the characteristics and Theorem 4.29 on pg. 187 of Jacod and Shiryaev (2002), all P''-martingales have the *representation property with respect to X* (for a definition, see Jacod and Shiryaev (2002), pg. 185). Hence the hypothesis of Theorem 5.19 on pg. 194 of Jacod and Shiryaev (2002) are satisfied and we conclude that the density process follows:

$$Z_t = Z_0 + \int_0^t Z_{s-}\beta_s \cdot dM''_s + \int_0^t Z_{s-}(Y_s(\xi) - 1)\left(\mathcal{J}(ds, d\xi) - ds\, K''_s(d\xi)\right),$$

where M'' is the continuous martingale part of X relative to P'', β is a predictable process such that $c(\sigma_s)\cdot\beta_s = b'_s - b''_s$ and

$$\begin{aligned} b''_s &= \mathsf{E}'(b(\sigma_s)|\mathcal{F}_{s-}), \\ b''_s &= \mathsf{E}''(b(\sigma_s)|\mathcal{F}_{s-}), \\ f'_s(\cdot) &= \mathsf{E}'(f(\sigma_s,\cdot)|\mathcal{F}_{s-}), \\ f''_s(\cdot) &= \mathsf{E}''(f(\sigma_s,\cdot)|\mathcal{F}_{s-}), \\ Y_s(\xi) &= \frac{f'_s(\xi)}{f''_s(\xi)}, \\ K''_s(d\xi) &= f''_s(\xi)G(d\xi). \end{aligned}$$

Let $F : \mathbb{R}_+ \to [0, 1]$ be the smooth function defined by:

$$F(z) = \frac{\phi_0 z}{\phi_0 z + 1 - \phi_0} \quad \text{for all } z \in \mathbb{R}_+ .$$

and note that

$$\phi_t = F(Z_t)$$

for all $t \geqq 0$.

From here on, the proof is purely algebraic and consists of applying a version of Itô's formula, namely Corollary B.2 below, to the expression above, which yields:

$$\begin{aligned} d\phi_t &= \underbrace{F'(Z_{t-})Z_{t-}\beta_t \cdot dM''_t + \frac{1}{2}F''(Z_{t-})Z^2_{t-}\beta_t^\top \cdot c_t \cdot \beta_t dt}_{=:d\Lambda_t} + \\ &\quad + [F(Z_{t-}Y_t(\xi)) - F(Z_{t-})]\,\mathcal{J}(dt, d\xi) - F'(Z_{t-})Z_{t-}(Y_t(\xi)-1)\,dt\,K''_t(d\xi)\,, \end{aligned} \tag{B.1}$$

where F' and F'' denote the first and second derivatives of F:

$$\begin{aligned} F'(z) &= \frac{\phi_0(1-\phi_0)}{(\phi_0 z + 1 - \phi_0)^2}\,, \\ F''(z) &= -\frac{2\phi_0^2(1-\phi_0)}{(\phi_0 z + 1 - \phi_0)^3}\,. \end{aligned}$$

One has:

$$\begin{aligned} F'(Z_{t-})Z_{t-} &= \phi_{t-}(1-\phi_{t-})\,, \\ F''(Z_{t-})Z^2_{t-} &= -2\phi^2_{t-}(1-\phi_{t-})\,, \end{aligned}$$

and, consequently,

$$d\Lambda_t = \phi_{t-}(1-\phi_{t-})\beta_t \underbrace{(dM''_t - \phi_{t-}c_t\beta\,dt)}_{=:dM_t}\,.$$

Also,

$$\begin{aligned} F(Z_{t-}Y_t) - F(Z_{t-}) &= \frac{\phi_0(1-\phi_0)Z_{t-}(Y_t-1)}{(\phi Z_{t-}Y_t + 1 - \phi_0)(\phi Z_{t-} + 1 - \phi)} \\ &= \frac{\phi_{t-}(1-\phi_0)(Y_t-1)}{\phi_0 Z_{t-}Y_t + 1 - \phi_0} \\ &= \frac{\phi_{t-}(1-\phi_{t-})(Y_t-1)}{\phi_{t-}Y_t + 1 - \phi_{t-}}\,, \end{aligned}$$

and

$$F'(Z_{t-})(Y_t - 1) = \phi_{t-}(1-\phi_{t-})(Y_t-1)\,.$$

Plugging the expressions above into B.1 yields:

$$\begin{aligned} d\phi_t &= \phi_{t-}\left[(1-\phi_{t-})\beta_t\,dM_t + \right. \\ &\quad \left. + \frac{(1-\phi_{t-})(Y_t-1)}{\phi_{t-}Y_t + 1 - \phi_{t-}}\big(\mathcal{J}(dt,d\xi) - (\phi_{t-}Y_t + 1 - \phi_{t-})(Y_t-1)\,dt\,K''_t(d\xi)\big)\right]\,, \end{aligned} \tag{B.2}$$

which is the desired result with

$$\begin{aligned} \varrho_t &= (1-\phi_{t-})\beta_t\,, \\ \vartheta_t &= \frac{(1-\phi_{t-})(Y_t-1)}{\phi_{t-}Y_t + 1 - \phi_{t-}}\,. \end{aligned}$$

That M is an $(\mathcal{F}_t)$-martingale under P follows from Girsanov's Theorem. □

Corollary B.1. *There is a constant $C > 0$, which depends only on the characteristics (b, γ, f, G), such that for any prior p, measurable set $\Theta' \subset \Theta$ with $p(\Theta') > 0$, and behavior strategy profile σ, one has*

$$\langle \phi \rangle_t \geqq C \int_0^t \phi_{s-}^2 D_s^2 \, ds \quad \mathsf{P}-a.s.$$

Proof. By Lemma B.1, the martingale of posteriors ϕ follows:

$$d\phi_t = \phi_{t-}\varrho_t \, dM_t + \phi_{t-}\vartheta_t(\xi) \, (\mathcal{J}(dt, d\xi) - dt \times \bar{K}_t(d\xi))$$

where M has quadratic variation $\int_0^t c_s \, ds$, with $c_s = c(\sigma_s)$. Therefore,

$$d\langle \phi \rangle_t = \phi_{t-}^2 \varrho_t^\top c_t \varrho_t \, dt + \phi_{t-}^2 \vartheta_t(\xi)^2 f_t^\phi(\xi) \, G(d\xi) \, dt$$

where

$$f_t^\phi(\cdot) = \phi_{t-} f_t'(\cdot) + (1 - \phi_{t-}) f_t''(\cdot) \, .$$

Define $b_t^\phi = \phi_{t-} b_t' + (1 - \phi_{t-}) b_t''$ and recall that:

$$\begin{aligned} \varrho_t &= (1 - \phi_{t-})\beta_t \quad \therefore \quad c_t \varrho_t = b_t' - b_t^\phi \\ \vartheta_t &= \frac{(1 - \phi_{t-})(Y_t - 1)}{\phi_{t-} Y_t + 1 - \phi_{t-}} = \frac{f_t' - f_t^\phi}{f_t^\phi} \, . \end{aligned}$$

In particular,

$$\varrho_t^\top c_t \varrho_t = \varrho_t^\top (b_t' - b_t^\phi) = (b_t' - b_t^\phi)^\top c_t^{-1} (b_t' - b_t^\phi)$$

where c_t^{-1} is the right inverse, i.e., $c_t c_t^{-1} = \mathrm{id}$ which is well-defined over $\operatorname{span} c_t$. (By Proposition B.1, $b_t' - b_t^\phi \in \operatorname{span} c_t$ so there is no loss of generality in assuming that $c(\alpha)$ is invertible for all $\alpha \in \Delta(A)$). Let $C' > 0$ be such that $v^\top \cdot c(a) \cdot v \leqq C' \, \|v\|^2$ for all $a \in A$ and $v \in \mathbb{R}^d$. Hence,

$$w^\top \cdot c(\alpha)^{-1} \cdot w \geqq 1/C' \, \|w\|^2$$

for all $w \in \operatorname{span} c(\alpha)$ and all $\alpha \in \Delta(a)$.

Finally, choosing $C'' > 0$ such that $f(a, \xi) \leqq C''$ for all $a \in A$ and $\xi \in \Xi$, and letting $C = 1/C' + 1/C''$ yields:

$$\begin{aligned} d\langle \phi \rangle_t &\geqq 1/C' \phi_{t-}^2 \left\| b_t' - b_t^\phi \right\|^2 dt + 1/C'' \phi_{t-}^2 \left\| f_t'(\xi) - f_t^\phi(\xi) \right\|^2 G(d\xi) \, dt \\ &\geqq C \, \phi_{t-}^2 \left(\left\| b_t' - b_t^\phi \right\|^2 + \left\| f_t'(\xi) - f_t^\phi(\xi) \right\|^2 G(d\xi) \right) dt \, , \end{aligned}$$

which concludes the proof. □

Itô's formula is standard tool in stochastic calculus. The version I need allows the process to have jumps. The reader is referred to Jacod and Shiryaev (2002), Theorem 4.57, on pg. 57.

Proposition B.2 (Itô's Formula)**.** *Let M be a square integrable martingale and g a C^2 map. Then,*

$$\begin{aligned} g(M_t) = g(M_0) + \int_0^t g'(M_s) \, dM_s + \frac{1}{2} \int_0^t g''(M_{s-}) \, d\langle M \rangle_s + \\ + \sum_{0 \leqq s \leqq t} \left(g(M_s) - g(M_{s-}) - g'(M_{s-}) \Delta M_s \right) \quad \text{(B.3)} \end{aligned}$$

Corollary B.2. *Let M be a P''-martingale. Suppose that $M_t = M_t^c + \int_0^t H_s(\xi)\cdot(\mathcal{J}(ds,d\xi) - ds\,K''(ds,d\xi))$, where M^c and $\int_0^t H_s(\xi)\cdot(\mathcal{J}(ds,d\xi) - ds\,K''(ds,d\xi))$ are the continuous and discontinuous martingale parts of M, respectively, where H is a predictable integrand. If g is a C^2 function, then*

$$g(M_t) = g(M_0) + \int_0^t g'(M_s)\,dM_s^c + \frac{1}{2}\int_0^t g''(M_{s-})\,d\langle M\rangle_s + \\ + \int_{[0,t]\times\Xi} (g(M_{s-} + H_s(\xi)) - g(M_{s-}))\,\mathcal{J}(ds,d\xi) - \int_{[0,t]\times\Xi} g'(M_{s-})H_s(\xi)\,ds\,K''(d\xi) \quad \text{(B.4)}$$

Proof. This follows directly from Proposition B.2, the definition of $\mathcal{J}$ and the definition of the integral $\int_0^t H_s(\xi)\cdot(\mathcal{J}(ds,d\xi) - ds\,K''(ds,d\xi))$. □

The lemma below was used in the proof of Proposition 4.2.

Lemma B.2. *Given any prior p, measurable set $\Theta' \subset \Theta$ with $p(\Theta') > 0$, and behavior strategy profile σ, one has for all $\eta > 0$,*

$$\mathsf{P}'\left[\inf_{t\geqq 0}\phi_t \leqq \eta\phi_0\right] \leqq \eta\,.$$

Proof. Since the odds ratio $(1-\phi_t)/\phi_t$ is a martingale under P', the maximal inequality for non-negative supermartingales yields:

$$\begin{aligned}\mathsf{P}'\left[\inf_{t\geqq 0}\phi_t \leqq \eta\phi_0\right] &= \mathsf{P}'\left[\sup_{t\geqq 0}(1-\phi_t)/\phi_t \geqq (1-\phi_0\eta)/(\phi_0\eta)\right] \\ &\leqq (1-\phi_0)\eta/(1-\phi_0\eta) \leqq \eta \quad \square\end{aligned}$$

C The Speed of Uniform Merging

Fix a measurable space $(\Omega,\mathcal{F})$ together with a filtration $(\mathcal{F}_n)_{n\geqq 1}$. Assume that for every probability measure P on $(\Omega,\mathcal{F})$, there exists a regular version of the conditional probability $P(\cdot|\mathcal{F}_n)$, for all $n \geqq 1$. The latter is true if, for example, $(\Omega,\mathcal{F})$ is a Polish space with the Borel σ-algebra.

Given P, Q probability measures on $(\Omega,\mathcal{F})$ define:

$$d_n(P,Q) = \operatorname{ess\,sup}_{B\in\mathcal{F}_n} |P(B|\mathcal{F}_{n-1}) - Q(B|\mathcal{F}_{n-1})|$$

Whenever Q is locally absolutely continuous with respect to P, the essential supremum above is well defined and $d_n(P,Q)$ is $\mathcal{F}_{n-1}$-measurable.

The following result is due to Fudenberg and Levine (1992). An alternative proof, based on a merging argument, is found in Sorin (1999), where the result is referred to as "uniform weak merging", in allusion to the uniformity of the convergence rates under the "grain of truth" condition.

Proposition C.1. *For all ε, η and $\phi > 0$, there exists a natural number $N \geqq 1$, such that for all P, Q and Q', probability measures on $(\Omega,\mathcal{F})$ with $P = \phi Q + (1-\phi)Q'$, one has:*

$$Q\left(\#\{n \geqq 1 \,:\, d_n(P,Q) \geqq \varepsilon\} \geqq N\right) \leqq \eta\,. \quad \text{(C.1)}$$

The following proposition provides sharp estimates of the speed of convergence for uniform weak merging. See Sandroni and Smorodinsky (1999) for a related result.

Proposition C.2. *For each ε, η and $\phi > 0$, let $N^*(\varepsilon, \eta, \phi)$ be the smallest natural number such that inequality (3.5) holds for all probability measures P, Q and Q' with $P = \phi Q + (1-\phi)Q'$. Then, for every η and $\phi > 0$, one has:*

$$\limsup_{\epsilon \to 0} \varepsilon^2 N^*(\varepsilon, \eta, \phi) < \infty \,.$$

Furthermore, if each $\mathcal{F}_n$ is atomic and $\mathcal{F}_n \subsetneqq \mathcal{F}_{n+1}$ for all $n \geqq 1$, then, for all $q > 0$ and all $\phi \in (0, 1]$, there exists $\bar{\eta} > 0$ such that for all $0 < \eta < \bar{\eta}$,

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

Proof. A close look at Sorin's proof of Proposition C.1 (Lemma 2.5 in Sorin (1999)) reveals that the natural number N_ε of inequality (3.5) can be taken as c/ε^2, where c is a positive real number which depends on η and $\underline{\phi}$. Hence $\limsup_{\varepsilon \to 0} \varepsilon^2 N^*_\varepsilon \leqq \limsup_{\varepsilon \to 0} \varepsilon^2 N = c < \infty$.

Consider the second claim. Assume the filtration is strictly increasing. Suppose, towards a contradiction, that for some $q > 0$ and $\phi \in (0, 1]$ there exist a function $(\varepsilon, \eta) \mapsto N(\varepsilon, \eta)$ and a sequence $\eta_k \to 0$ with

$$\limsup_{\varepsilon \to 0} \varepsilon^{2-q} N(\varepsilon, \eta_k) < \infty \,,$$

and such that for all P, Q and Q' with $P = \phi Q + (1-\phi)Q'$ one has:

$$Q\left(\#\{n \geqq 1 \,:\, d_n(P, Q) \geqq \varepsilon\} \geqq N(\varepsilon, \eta_k)\right) \leqq \eta_k$$

for all $k \geqq 1$ and for all $\varepsilon > 0$ sufficiently small.

For each $k \geqq 1$, let $M_k = 1 + \limsup_{\varepsilon \to 0} \varepsilon^{2-q} N(\varepsilon, \eta_k)$. Define a double sequence $\varepsilon_{m,k} = (M_k/m)^{1/(2-q)}$ for all $k, m \geqq 1$. Hence we have for all $k \geqq 1$,

$$\limsup_{m \to \infty} N(\varepsilon_{m,k}, \eta_k)/m < 1$$

By Lemma C.2 below, there exist probability measures P, Q and Q', with $P = \phi Q + (1-\phi)Q'$, such that for all $k \geqq 1$,

$$\liminf_{m \to \infty} Q\left(\#\{n \geqq 1 \,:\, d_n(P, Q) \geqq \varepsilon_{m,k}\} \geqq N(\varepsilon_{m,k}, \eta_k)\right) \geqq 1/2$$

Since $\eta_k < 1/2$ eventually, we have shown that for all k sufficiently large and $\forall \bar{\varepsilon} > 0$ $\exists \varepsilon \in (0, \bar{\varepsilon}]$ such that:

$$Q\left(\#\{n \geqq 1 \,:\, d_n(P, Q) \geqq \varepsilon\} \geqq N(\varepsilon, \eta_k)\right) > \eta_k \,,$$

which is a contradiction. This contradiction shows that $\forall q > 0$ and $\forall \phi \in (0, 1]$, one has

$$\lim_{\varepsilon \to 0} \varepsilon^{2-q} N^*(\varepsilon, \eta, \phi) = \infty$$

for all η sufficiently small. □

The following lemma is needed in the proof Lemma C.2. It is, essentially, a refinement of the Borel-Cantelli Lemma. Under a stronger assumption than the latter, it asserts that, in addition to the events happening infinitely often with probability one, there is positive probability that the upper density of the indices for which the event obtains is bounded above 0.

Lemma C.1. *Let $A_1, A_2, \ldots$ be a sequence of independent events on a probability space $(\Omega, \mathcal{F}, P)$. Suppose that $c = \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} PA_k > 0$. Then $P(A_n \text{ i.o.}) = 1$ and, in addition,*

$$P\left(\limsup_{N\to\infty} \frac{\#\{n \leqq N \;:\; A_n \text{ obtains}\}}{N} \geqq c\right) \geqq \liminf_{N\to\infty} P\left(\frac{\#\{n \leqq N \;:\; A_n \text{ obtains}\}}{N} \geqq c\right) \geqq \frac{1}{2}.$$

Proof. Since the assumption implies that $\sum_n PA_n = \infty$, it follows from Borel-Cantelli that $P(A_n \text{ i.o.}) = 1$. Let us turn to the statement concerning the upper density. By the Central Limit Theorem,

$$\frac{\sum_{k=1}^{n} I_{A_k} - PA_k}{\sqrt{\sum_{k=1}^{n} PA_k}} \longrightarrow \mathcal{N}(0,1)\,,$$

where the convergence is in distribution. Let $0 < c < \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} PA_k$. We have:

$$\begin{aligned} P\left(\frac{1}{n}\sum_{k=1}^{n} I_{A_k} \geqq c\right) &= P\left(\frac{\sum_{k=1}^{n} I_{A_k} - PA_k}{\sqrt{\sum_{k=1}^{n} PA_k}} \geqq \frac{cn - \sum_{k=1}^{n} PA_k}{\sqrt{\sum_{k=1}^{n} PA_k}}\right) \\ &\geqq P\left(\frac{\sum_{k=1}^{n} I_{A_k} - PA_k}{\sqrt{\sum_{k=1}^{n} PA_k}} \geqq 0\right) \quad \text{for } n \text{ sufficiently large}\,. \end{aligned}$$

Therefore, denoting by Φ the standard normal distribution function, we have

$$\liminf_{n} P\left(\frac{1}{n}\sum_{k=1}^{n} I_{A_k} \geqq c\right) \geqq 1 - \Phi(0) = 1/2\,.$$

□

Lemma C.2. *Suppose that $(\mathcal{F}_n)_{n\geqq 1}$ is a strictly increasing sequence of atomic σ-algebras. Then, $\forall \nu > 1/2$ and $\forall \phi \in (0,1]$, $\exists P, Q$ and Q' with $P = \phi Q + (1-\phi)Q'$, such that $\forall c > 0$,*

$$\liminf_{N\to\infty} Q\left(\#\{n \geqq 1 \;:\; d_n(P,Q) \geqq c/N^{\nu}\} \geqq N\right) \geqq 1/2\,,$$

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 σ-algebra $\mathcal{F}_n$ is that generated by all 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 $P = \phi Q + (1-\phi)Q'$ and ϕ_n be the posterior probability that the stochastic process follows Q conditional on period n information.

Notice that $d_n(P,Q) = (p_n - 1/2)(1-\phi_n)$. Denote by χ_n the $\{0,1\}$-valued random variable that takes value 1 if and only if the coin turns out Heads at time n. Explicit computation of Bayes rule yields for all $K > 0$:

$$
\begin{aligned}
Q\left(\frac{d_n(P,Q)}{\phi_n} \geqq \frac{4K/\phi_1}{n^\nu}\right) &= Q\left(\frac{1-\phi_n}{\phi_n} \geqq \frac{8K/\phi_1}{n^{\nu-\alpha}}\right)\\
&= Q\left(\prod_{k=1}^n (2p_k)^{\chi_k}(2(1-p_k))^{(1-\chi_k)} \geqq \frac{8K/(1-\phi_1)}{n^{\nu-\alpha}}\right)\\
&= Q\left(\sum_{k=1}^n \chi_k \ln(2p_n) + \sum_{k=1}^n (1-\chi_k)\ln(2(1-p_k)) \geqq \underbrace{\ln(8K/(1-\phi_1))}_{=B} - \underbrace{(\nu-\alpha)}_{=\gamma>0}\ln n\right)\\
&= Q\left(\sum_{k=1}^n \ln(\frac{p_k}{1-p_k})\chi_k \geqq \sum_{k=1}^n \ln(\frac{1}{2(1-p_k)}) + B - \gamma \ln n\right)\\
&= Q\left(\sum_{k=1}^n \ln(\frac{p_k}{1-p_k})(\chi_k - \frac{1}{2}) \geqq \frac{1}{2}\sum_{k=1}^n \ln(\frac{1}{4p_k(1-p_k)}) + B - \gamma \ln n\right)\\
&\geqq Q\left(\left|\sum_{k=1}^n \ln(\frac{p_k}{1-p_k})(\chi_k - \frac{1}{2})\right| \leqq \gamma \ln n - \frac{1}{2}\sum_{k=1}^n \ln(\frac{1}{4p_k(1-p_k)}) - B\right) \quad \text{(C.2)}
\end{aligned}
$$

Chebishev's inequality yields for all $M > 0$:

$$
Q\left(\left|\sum_{k=1}^n \ln(\frac{p_k}{1-p_k})(\chi_k - \frac{1}{2})\right| \geqq M\right) \leqq \frac{\sum_{k=1}^\infty (\ln(\frac{p_k}{1-p_k}))^2}{4M^2}
$$

By Lemma C.3 below $\sum_{k=1}^\infty (\ln(\frac{p_k}{1-p_k}))^2 < \infty$, hence we can choose $M > 0$ large enough so that

$$
Q\left(\left|\sum_{k=1}^n \ln(\frac{p_k}{1-p_k})(\chi_k - \frac{1}{2})\right| \geqq M\right) \leqq \frac{1}{2}\,. \quad \text{(C.3)}
$$

On the other hand, by Lemma C.3, also $\sum_{k=1}^\infty \ln(\frac{1}{4p_k(1-p_k)}) < \infty$. Therefore, for n sufficiently large,

$$
\gamma \ln n - \frac{1}{2}\sum_{k=1}^n \ln(\frac{1}{4p_k(1-p_k)}) - B > M\,.
$$

Therefore, by the inequalities (C.2) and (C.3) above, for all n sufficiently large,

$$
Q\left(d_n(P,Q)/\phi_n \geqq \frac{4K/\phi_1}{n^\nu}\right) \geqq \frac{1}{2}
$$

By Lemma C.4, $Q\left(\phi_n \leqq \frac{1}{4}\phi_1\right) \leqq \frac{1}{4}$. Thus, for n large enough,

$$
\begin{aligned}
Q\left(d_n(P,Q) \geqq K/n^\nu\right) + \frac{1}{4} &\geqq Q\left(d_n(P,Q) \geqq K/n^\nu\right) + Q\left(\phi_n \leqq \frac{1}{4}\phi_1\right)\\
&\geqq Q\left(d_n(P,Q)/\phi_n \geqq \frac{4K/\phi_1}{n^\nu}\right) \geqq \frac{1}{2}\,.
\end{aligned}
$$

We have thus proved that,

$$
\liminf_{n\to\infty} Q\left(d_n(P,Q) \geqq K/n^\gamma\right) \geqq \frac{1}{4}\,.
$$

Therefore, the events $\{d_n(P,Q) \geqq K/n^\nu\}$ satisfy the assumption of Lemma C.1 and we have

$$\liminf_{N\to\infty} Q\left(\frac{\#\{n \leqq N \,:\, d_n(P,Q) \geqq K/n^\nu\}}{N} \geqq \frac{1}{4}\right) \geqq \frac{1}{2}\,.$$

Since

$$Q\left(\frac{\#\{n \leqq N \,:\, d_n(P,Q) \geqq K/N^\nu\}}{N} \geqq \frac{1}{4}\right) \geqq Q\left(\frac{\#\{n \leqq N \,:\, d_n(P,Q) \geqq K/n^\nu\}}{N} \geqq \frac{1}{4}\right),$$

it follows that

$$\liminf_{N} Q\left(\#\{n \,:\, d_n(P,Q) \geqq K/N^\nu\} \geqq \frac{1}{4}N\right) \geqq \frac{1}{2}\,,$$

which proves the desired result for the case in which $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geqq 1})$ is the canonical coin tossing space.

If $(\Omega, \mathcal{F})$ is an arbitrary space with a strictly increasing atomic filtration $(\mathcal{F}_n)$, then, for every n, there are at least two disjoint atomic 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". □

Lemma C.3. *If $p_k = \frac{1}{2}(1 + 1/k^\alpha)$ with $\alpha > 1/2$, then*

$$\sum_{k=1}^{\infty} \ln(\frac{1}{4p_k(1-p_k)}) < \infty \quad \textit{and} \quad \sum_{k=1}^{\infty}(\ln(\frac{p_k}{1-p_k}))^2 < \infty\,.$$

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

Lemma C.4. *Let P, Q and Q' be such that $P = \phi Q + (1-\phi)Q'$. Let ϕ_n be the posterior probability on the event that the process follows Q, conditional on $\mathcal{F}_n$. Then $\forall c > 0$,*

$$Q(\phi_n \leqq c\phi_1) \leqq c\,.$$

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

$$\begin{aligned} Q(\phi_n \leqq c\phi_1) &= Q\left(\frac{1-\phi_n}{\phi_n} \geqq \frac{1-c\phi_1}{c\phi_1}\right) \\ &\leqq \frac{E_Q[(1-\phi_n)/\phi_n]}{(1-c\phi_1)/(c\phi_1)} \\ &= \frac{(1-\phi_1)/\phi_1}{(1-c\phi_1)/(c\phi_1)} \leqq c \quad \square \end{aligned}$$

References

ABREU, D., P. MILGROM, AND D. PEARCE (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59(6), 1713–1733.

ALIPRANTIS, C. D., AND K. C. BORDER (1999): *Infinite Dimensional Analysis : A Hitchhiker's Guide*. Springer-Verlag, 2nd edn.

BERGIN, J., AND W. B. MACLEOD (1993): "Continuous Time Repeated Games," *International Economic Review*, 34, 21–37.

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. John Wiley & Sons.

BLACKWELL, D., AND L. DUBINS (1962): "Merging of Opinions with Increasing Information," *Ann. Math. Statist.*, 38, 882–886.

BREMAUD, P. (1981): *Point Processes and Queues: Martingale Dynamics*. Springer-Verlag.

ETHIER, S. N., AND T. G. KURTZ (1986): *Markov Processes: Characterization and Convergence*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc.

FAINGOLD, E., AND Y. SANNIKOV (2005): "Equilibrium Degeneracy and Reputation Effects," University of Pennsylvania.

FUDENBERG, D., AND D. K. LEVINE (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed," *Review of Economic Studies*, 59, 561–579.

——— (1994): "Efficiency and Observability with Long-run and Short-run Players," *Journal of Economic Theory*, 62(1), 103–135.

JACOD, J., AND A. N. SHIRYAEV (2002): *Limit Theorems for Stochastic Processes*, vol. 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 2nd edn.

KALAI, E., AND E. LEHRER (1993): "Rational Learning Leads to Nash Equilibrium," *Econometrica*, 61, 1019–1045.

KREPS, D., P. MILGROM, J. ROBERTS, AND R. WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoners-Dilemma," *Journal of Economic Theory*, 27(2), 245–252.

KREPS, D., AND R. WILSON (1982): "Sequential Equilibria," *Econometrica*, 50, 863–894.

KUSHNER, H. J. (1990): *Weak Convergence Methods and Singularly Perturbed Stochastic Control and FIltering Problems*. Birkhauser.

KUSHNER, H. J., AND P. DUPUIS (2001): *Numerical Methods for Stochastic Control Problems in Continuous Time*, vol. 24 of *Applications of Mathematics*. Springer-Verlag, 2nd edn.

SANDRONI, A., AND R. SMORODINSKY (1999): "The Speed of Rational Learning," *International Journal of Game Theory*, 28, 199–210.

SANNIKOV, Y. (2005): "Games with Imperfectly Observed Actions in Continuous Time," Working Paper, University of California at Berkeley.

SANNIKOV, Y., AND A. SKRZYPACZ (2005): "Impossibility of Collusion under Imperfect Monitoring with Flexible Production," Working Paper, University of California at Berkeley.

SEKIGUCHI, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory* ,, 76, 345–361.

SIMON, L. K., AND M. STINCHCOMBE (1987): "Extensive Form Games in Continuous Time: Pure Strategies," *Econometrica*, 57, 1171–1214.

SORIN, S. (1999): "Merging, Reputation and Repeated Games with Incomplete Information," *Games and Economic Behavior*, 29, 274–308.