Equilibrium Degeneracy and Reputation Effects
in Continuous Time Games[*]

Eduardo Faingold[†]
Yuliy Sannikov[‡]

November, 2005

Abstract

We study a continuous-time dynamic game between a large player and a population of small players in which the actions of the large player are imperfectly observable. We explore two versions of the game. In the complete information game, in which it is common knowledge that the large player is a strategic normal type, we show that intertemporal incentives have no power: Irrespective of players' patience and signal informativeness, the set of equilibrium payoffs of the large player coincides with (the convex hull of) the set of static Nash equilibrium payoffs.

In the game with reputation effects, the small players believe that the large player may be a commitment type, who plays the same action at all times irrespective of the past history of play. With this perturbation, nontrivial intertemporal incentives arise. The set of sequential equilibrium payoffs of the large player can be characterized by means of a pair of ordinary differential equations. Furthermore, in the equilibrium that maximizes the large player's payoff, continuation play is determined by a single state variable: the small players' posterior belief. We apply our results to several examples from Industrial Organizations, Macroeconomics and Corporate Finance.

1 Introduction

In many economic environments a large player can benefit from committing to a course of actions to influence the behavior of a population of small players. A firm may wish to commit to fight potential entrants, to provide high quality to its customers, to honor implicit labor contracts, and to generate good returns to investors. Governments can benefit from commitment to a non-inflationary monetary policy, low capital taxation and efforts to fight corruption. Often the actions of the large player are imperfectly observable. For example, the quality of a firm's products may be a noisy outcome of a firm's hidden effort to maintain quality standards. The actual inflation rate can be a noisy signal of money supply.

We study a dynamic game in continuous time between a large player and a population of small players to gain insight behind the possibility of commitment in these situations. This game is a continuous-time analogue of the repeated game of Fudenberg and Levine (1992), hereafter FL. We assume that there is a continuum of small players and that the distribution of small players' actions is publicly observed, but not

---

the actions of any individual small player. Hence, as in FL, the small players behave myopically in every equilibrium, acting to maximize their instantaneous expected payoffs.

First we consider the complete information version of this dynamic game. We find that, due to imperfect monitoring, the large player cannot achieve better payoffs than in static Nash equilibria. When the small players perfectly observe the actions of the large player, higher payoffs than in static equilibria are sustainable in an equilibrium with two regimes: a high-payoff commitment regime and a punishment regime. Players are in the commitment regime on the equilibrium path, i.e. the large player chooses a desired action and the small players believe in commitment. Any deviation off the equilibrium path triggers the punishment regime. If that happens, the small players do not believe in commitment and play according to a static Nash equilibrium that yields low payoffs to the large player. Hence the commitment regime is enforced by the threat of a punishment in case the large player violates the commitment to obtain short-run gains. If the large player is sufficiently patient, this threat credibly enforces commitment.

With imperfect information about the large player's actions, one can try to construct an analogous equilibrium. In this equilibrium imperfect signals about the large player's actions determine the transitions between the commitment and punishment regimes. As shown in Fudenberg and Levine (1994), when monitoring has full-support, only limited commitment can be attained by such an equilibrium in a discrete-time repeated game. We show that in a continuous-time setting commitment becomes completely impossible. To explain this result we borrow intuition from Abreu, Milgrom, and Pearce (1991) and Sannikov and Skrzypacz (2005), who study a different class of repeated games with frequent actions. Commitment becomes impossible with frequent actions because players see very little information per period and so the statistical tests that trigger punishments give false positives too often.

We prove that commitment is impossible using a more general argument that does not rely on a specific equilibrium structure. The intuition is as follows: one must be able to both reward and punish the large player to provide incentives for the desired action. If only punishments are involved in the provision of incentives, those punishments destroy too much value. Because in the best equilibrium for the large player incentives can be provided only via punishments, that equilibrium cannot be better than the Nash equilibrium of a stage game.

The possibility of commitment arises if the small players are uncertain about the type of the large player. Specifically, they believe that the large player could be a *commitment type* who always plays the same action or a *normal type* who acts strategically. Then it is attractive for the normal type to imitate the commitment type because the payoff of the normal type increases when he is perceived to be a commitment type with greater probability. In equilibrium this imitation is imperfect: if it were perfect, the public signal would not be informative about the large player's type, so imitation would have no value. The normal type obtains his maximal payoff when the population is certain that he is a commitment type. Then the population's beliefs do not change and the the normal type "gets away" with any action. This feature of the equilibrium is consistent with the fact that it is impossible to provide incentives to the normal type of the large player when his payoff is maximized.

We derive an ordinary differential equation that can be used to characterize the best equilibrium payoff of the large player as a function of the population's belief. When this equation has a solution with appropriate boundary conditions, the best equilibrium is unique and takes a clean form. In this equilibrium a single state variable, the population's belief, determines the equilibrium actions and the expected future payoff of the large player after any history. The equilibrium characterization is much cleaner in continuous time than in

discrete time, where the large player's continuation payoff is necessarily nonunique for any given belief.

The incomplete information approach to reputations has its roots in the works of Kreps and Wilson (1982) and Milgrom and Roberts (1982), in their study of Selten's chain-store paradox, and of Kreps, Milgrom, Roberts, and Wilson (1982), in their analysis of cooperation in the finitely repeated Prisoner's Dilemma. Uncertainty over types, particularly over types that behave as automata committed to certain strategies, gives rise to phenomena that could not be explained by the equilibria of the underlying complete information games: entry deterrence in the Chain-store game and cooperation in (almost every period of) the Prisoner's Dilemma game.

Fudenberg and Levine (1992) study reputation effects in discounted repeated games with imperfect monitoring played by a long-run player and a sequence of short-lived players. As in the current paper, the short-run players believe that, with positive probability, the long-run player is a type committed to a certain strategy. In this environment, Fudenberg-Levine derive upper and lower bounds to the set of Nash equilibrium payoffs of the long-run player. These bounds hold asymptotically as the long-run player's discount factor tends to one. When the set of commitment types is sufficiently rich and the monitoring technology satisfies an identification condition, the upper and lower bounds coincide with the long-run player's Stackelberg payoff, i.e., the payoff he obtains from credibly committing to the strategy to which he would like to commit the most. Moreover this result obtains regardless of how small the prior probability on the commitment types is.

In a similar vein, Faingold (2005) investigates the role of reputation effects in repeated games with imperfect monitoring in which players take actions very frequently. By studying an auxiliary limit game in continuous time, reputation bounds similar to those in FL are obtained, which hold uniformly over all discrete-time games with sufficiently short periods of time. The reputation bounds obtained in Faingold (2005) do hold for the games studied in this paper. In particular, for any fixed prior probability on the commitment type, if the long-run player is sufficiently patient, then in every Nash equilibrium his payoffs are approximately bounded below by the commitment payoff.

As exemplified by the aforementioned papers, the reputation literature typically focuses on studying the set of Nash equilibrium payoffs in games where the long-run player is patient relatively to a fixed prior. In contrast, our goal in this paper is to characterize the set of public sequential equilibria for all discount rates and all priors, not only in terms of payoffs but also in terms of equilibrium behavior. Our characterization exploits the recursive nature of the set of public sequential equilibria. Such recursive structure also exists in the discrete-time world, but there is no obvious way of exploring it there. In effect, in discrete-time, constructing even one equilibrium seems to be a very difficult task[1]. The difficulty arises from the fact that with full-support imperfect monitoring, the learning process of the short-run players never stops in finite time. Even if the short-run players are almost convinced about the type of the long-run player, with positive probability histories arise which drive beliefs away from certainty.

The methods we employ are very similar to those developed in Sannikov (2004). The latter paper develops recursive methods that apply to complete information games played in continuous time. As in the current paper, the games studied in Sannikov (2004) are games with imperfect monitoring, where the observation of the players' actions is distorted by a Brownian motion. A key insight in that paper is that the dynamics of the perfect public equilibria (PPE) that attain extreme payoffs is such that the continuation values never

---

[1]The only case in which equilibria can be trivially constructed in discrete-time is when the commitment action and the short-run players' best-reply to it form a Nash equilibrium of the stage game.

leave the boundary of the set of PPE payoffs. This property, in conjunction with a characterization of the incentive constraints that must be satisfied in a PPE, makes it possible to fully characterize, by means of an ordinary differential equation, the boundary set of PPE payoffs in those games. In the present paper, we borrow many techniques and insight from Sannikov (2004), but we are naturally led to develop new methods to deal with games with incomplete information. In particular, we provide a useful characterization of the law of motion of beliefs in terms of a stochastic differential equation.

The paper is organized as follows. Section 2 presents several applications to show that the problem of policy commitment appears widely in economics. Section 3 presents the abstract model in continuous time. Sections 4, 5 and 6 carry the analysis and characterize public sequential equilibria. Section 7 summarizes the results, and explores the issues of uniqueness and multiplicity of equilibria. Section 8 concludes.

2   Examples.

In this section we present three examples to illustrate possible applications of our results. The common theme of our examples is that there is a large player, who would like to commit to a strategy to influence the behavior of a population of small players. The small players act myopically to maximize their static payoffs. The actions of the large player are imperfectly observable: the public signal about the large player's actions is distorted by a Brownian motion. With imperfect monitoring, non-trivial intertemporal incentives have power only if the population admits the possibility that the large player may be a *commitment* type, who is not able to cheat. The *normal* type of the large player acts strategically to partially imitate the commitment type, because if he cheats, the population punishes him by its beliefs. We designed three examples to span a variety of fields in economics where our methods can be used.

2.1   Maintaining Quality Standards.

A service provider and a mass of consumers uniformly distributed on $[0, 1]$ play the following game in continuous time. At each moment of time $t \in [0, \infty)$ the service provider makes an investment in quality $a_t \in [0, 1]$ and each consumer $i \in [0, 1]$ chooses a service level $b_t^i \in [0, 3/4]$. Consumers do not see the firm's investment, but observe the actual quality X_t of the service, which is a noisy public signal of the firm's hidden investment. The law of motion of (X_t) is given by the stochastic differential equation:

$$dX_t = 4\sqrt{a_t}(1 - \bar{b}_t)\,dt + 4(1 - \bar{b}_t)\,dZ_t,$$

where Z is a standard Brownian motion, $\bar{b}_t$ is the small players' average service level, the drift $4\sqrt{a_t}(1-\bar{b}_t)$ is the expected quality flow at time t, and $4(1 - \bar{b}_t)$ is the magnitude of the noise. The noise is decreasing with usage: the more customers use the service the better they learn its quality. The drift term captures congestion externalities: the quality of the service deteriorates with greater usage.

Consumer i pays the price equal to his service level b_t^i. The payoff of consumer i is

$$r \int_0^\infty e^{-rt}(b_t^i\,dX_t - b_t^i\,dt),$$

where $r > 0$ is a discount rate. The payoff of the service provider is

$$r \int_0^\infty e^{-rt}(\bar{b}_t - a_t)\,dt.$$

4

What if the service provider were able to commit to an investment level a^* and credibly convey this commitment to consumers? Then each consumer would choose a service level b^i that maximizes

$$b^i \left(4\sqrt{a^*}(1 - \bar{b}) - 1 \right).$$

For a given level of a^* all customers will choose the same service level $b^i = 1 - \frac{1}{4\sqrt{a^*}}$, and the service provider earns $1 - \frac{1}{4\sqrt{a^*}} - \sqrt{a^*}$. The best commitment outcome for the service provider is when he commits to $a^* = 1/4$ and all the customers choose service level $1/2$.
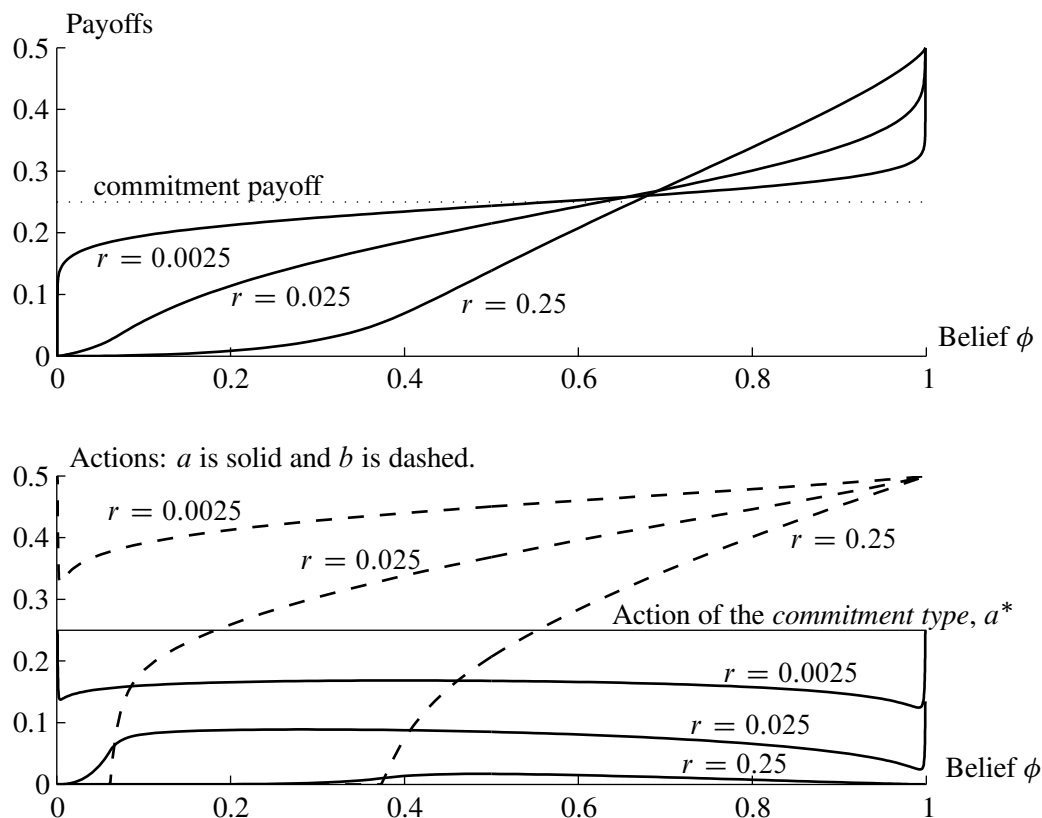
PSfrag replacements

Figure 1: Equilibrium payoffs and actions in the quality commitment game.

Let us describe the game with reputation effects. At time 0 the consumers believe that with probability p the service provider is a *commitment type*, who always chooses investment a^*. With probability $1 - p$ the service provider is a *normal type*, who chooses an investment level to maximize his expected profit. What happens in equilibrium?

Our first result is that when $p = 0$, i.e., when it is common knowledge that the large player is normal, the large player obtains a payoff of 0, which is the unique Nash equilibrium payoff of the static game. Hence intertemporal incentives completely collapse: the service provider does not invest and the consumers do not buy the service, i.e. $a_t = b^i_t = 0$, for all $t \geq 0$.

When $p \neq 0$, the top panels of Figure 1 show the (unique) equilibrium payoff of the large player as a function of the population's belief p for $a^* = 1/4$ and different discount rates r. In equilibrium the population constantly updates its belief ϕ_t, the probability assigned to the commitment type of the large player, based on the public signal (X_t). The equilibrium actions of the population and the large player are uniquely determined by the population's belief. Those actions are shown on the bottom panel of Figure 1.

Consistent with the asymptotic results in Faingold (2005), computation shows that as $r \to 0$, the large player's payoff converges to his best commitment payoff of 1/4. The customer usage level b increases towards the commitment level 1/2 as their belief ϕ increases towards 1 or the discount rate r decreases towards 0. The normal type of the large player comes closer to imitating the commitment type when r is closer to 0, but for any fixed r, the imitation is always partial.

2.2 Government Policy.

Since the seminal work of Kydland and Prescott (1977) economists have recognized that many optimal policies require commitment due to their intertemporal inconsistency. We see the credibility of government policies as the most important potential application of our paper. In many situations it is appropriate to model the government's actions to be imperfectly observable. Even when rules like tax rates are explicit, it may be suitable to model them as imperfectly observable if the actual tax that the population faces depends on corruption and the government's enforcement efforts. Even when the actions of the government are transparent, a policy game with imperfect monitoring may be the right model if the government's actions depend on private information. Here we present a simple example in which the policy is explicit, but the authority's enforcement efforts are imperfectly monitored.

There is a population $[0, 1]$ of individuals. Every individual in a population chooses the extent $b_t \in [0, 1]$ to which he will engage in a certain illegal activity (e.g. using drugs). The local authority decides how tough to be at enforcing the policy. The action of the authority is denoted by $a_t \in [0, 2]$. There is imperfect information about the authority's enforcement efforts, which is reflected in an aggregate public signal

$$dX_t = a_t \, dt + \sigma \, dZ_t,$$

where σ is a parameter that reflects the transparency of the enforcement policy. The aggregate public signal comes from the news stories about individuals who got caught, the authority's reports, etc. Individuals like engaging in this illegal activity but dislike getting caught. Also, the average level of illegal activity $\bar{b}_t$ hurts everyone. The payoff of an individual who chooses activity level b_t is given by

$$\int_0^\infty e^{-rt} (2b_t - b_t^2 - a_t b_t - 4\bar{b}_t) \, dt,$$

where $2b_t - b_t^2$ is the utility from the illegal activity, $a_t b_t$ is the expected punishment from getting caught and $4\bar{b}_t$ is the social cost from the overall level of the illegal activity. The authority wants to maximize compliance with the law, but it has to pay a cost to enforce the law, which is increasing with the level of illegal activity of the population. The payoff of the authority is given by

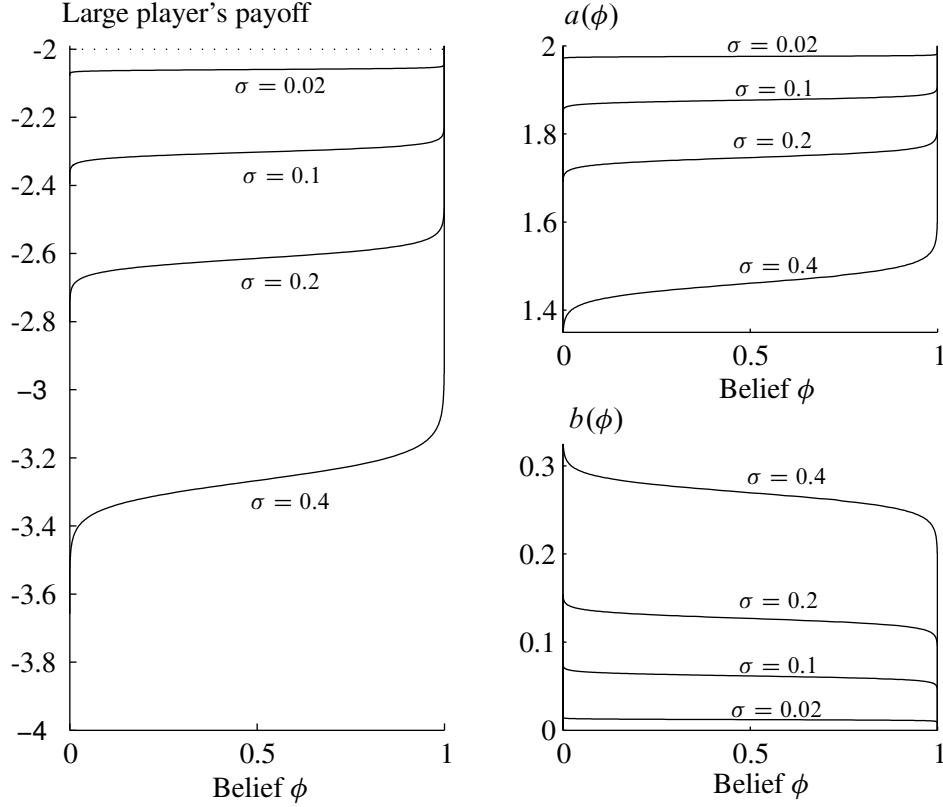$$-\int_0^\infty e^{-rt} (a_t b_t + a_t + 4\bar{b}_t) \, dt.$$

6

Figure 2: Equilibrium payoffs and actions in the law enforcement game.

In the absence of commitment, the outcome of the dynamic game would be identical to the static Nash equilibrium: the authority would choose enforcement level $a = 0$, and the population would be fully engaged in the illegal activity (i.e. $b = 1$.) As a result, the authority gets a payoff of -4. If the authority was able to commit perfectly to an enforcement level, it would choose $a = 2$. This causes the level of illegal activity of 0, and the authority's payoff is -2. If the authority took no enforcement effort, but the population believed that the law was fully enforced (i.e. $a = 2$), then there would be no illegal activity and the authority would get a payoff of 0.

Now let us describe the reputational equilibrium. Assume that the population is uncertain whether the authority is a commitment type, who always chooses $a^* = 2$, or a normal strategic type. Then in equilibrium the normal type imperfectly imitates the strategic type because its payoff is increasing in the probability ϕ that the population assigns to the the large player being a commitment type. In equilibrium the payoff of the normal type of the large player, as well as everyone's actions are uniquely determined by ϕ, as shown in Figure 2. The population's belief evolves as

$$\phi_0 = 0, \quad d\phi_t = \phi_t(1 - \phi_t)\frac{(a^* - a(\phi_t))(dX_t - (\phi_t a^* + (1 - \phi_t)a(\phi_t))\,dt)}{\sigma^2},$$

and the large player chooses the action $a_t = a(\phi_t)$.

2.3 Attracting Investors.

A firm that has a profitable production technology is trying to attract investors. The firm's profit flow is

$$dY_t = f(\bar{b}_t)\, dt + \sigma(\bar{b}_t)\, dZ_t,$$

where $\bar{b}$ is the aggregate outside investment and f is an increasing and concave production function with $f(0) = 0$ and $f'(0) = \infty$. Assume that $f(\bar{b}) - R\bar{b}$ achieves a maximal value of S when $\bar{b} = 1$, where R is the fair return to outside investors. If by the firm's rules the manager's salary is $S\bar{b}$, then the investors choose $\bar{b} = 1$ efficiently and get a market return of R.

Suppose there is an agency problem because the firm's manager can divert cash flows for personal consumption. However, unlike in the settings of DeMarzo and Fishman (2003) and DeMarzo and Sannikov (2004), the investors are dispersed and they cannot write an explicit contract with the firm. Therefore, the firm must rely on its reputation.

After diversion and salary, the outside investors get

$$dX_t = (f(\bar{b}_t) - (S + a_t)\bar{b}_t)\, dt + \sigma(\bar{b}_t)\, dZ_t,$$

where a_t is the manager's diversion decision. The outside investors do not see Y_t, but only see X_t. Assume that cash diversion is inefficient: the manager's payoff flow is given by

$$(S + \lambda(a_t))\bar{b}_t,$$

where λ is weakly increasing and concave, with $\lambda(0) = 0$, $\lambda'(0) = 1$ and $\lambda'(a) = 0$ for all $a \geq \bar{a}$. The payoff of an individual investor i is given by

$$b^i \left(\frac{dX_t}{\bar{b}_t} - R \right).$$

As in the previous examples, the best outcome in the dynamic version of this game is the same as the static Nash equilibrium. The manager diverts cash at the maximal level $\bar{a}$. The investment level is less than efficient and is defined by the investors' indifference condition $f(\bar{b}_t) = \bar{b}(S + \bar{a} + R)$. Both the Nash equilibrium and the efficient commitment solution are illustrated in Figure 3.

If the manager has positive reputation, i.e. the investors believe that he may be a commitment type who does not divert cash, then the equilibrium allows great improvement. As $r \to 0$ or the noise in cash flows decreases, the outcome converges to efficiency.

3 The Game.

A *large player* participates in a dynamic game with a continuum of *small players* uniformly distributed in $I = [0, 1]$. At each time $t \in [0, \infty)$, the large player chooses an action $a_t \in A$ and each small player $i \in I$ chooses an action $b_t^i \in B$ based on their current information. Action spaces A and B are compact, convex subsets of an Euclidean space. The small players' moves are *anonymous*: at each time t, the large player observes the aggregate distribution $\bar{b}_t \in \Delta(B)$ of the small players' actions, but does not observe the action of any individual small player. There is *imperfect monitoring*: the large player's moves are not observable
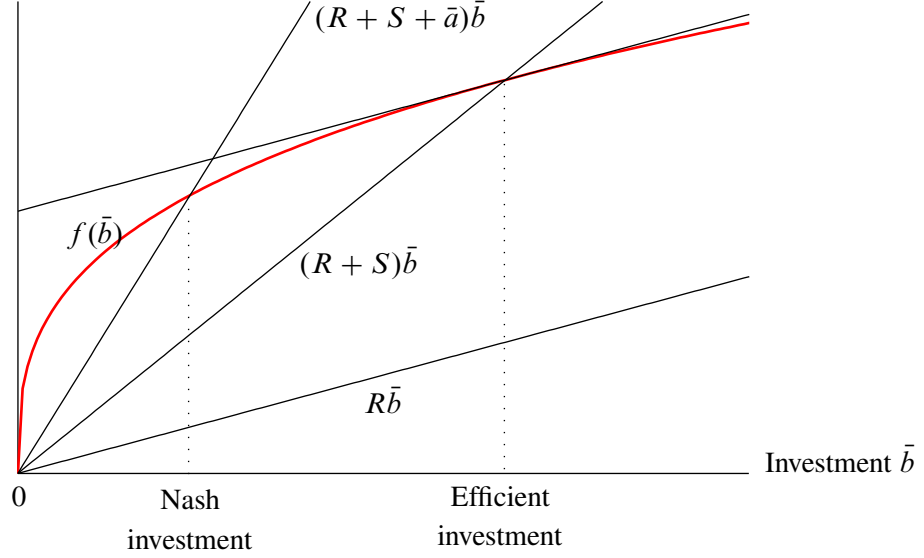
Figure 3: The Nash equilibrium and the efficient solution in the investment game.

to the small players. Instead, the small players see a noisy public signal $(X_t)_{t\geq 0}$ that depends on the actions of the large player, the aggregate distribution of the small players' actions and noise. Specifically,

$$dX_t = \mu(a_t, \bar{b}_t)\,dt + \sigma(\bar{b}_t)\cdot dZ_t,$$

where (Z_t) is a d-dimensional Brownian motion, and the drift and the volatility of the signal are defined via continuously differentiable functions $\mu : A \times B \to \mathbb{R}^d$ and $\sigma : B \to \mathbb{R}^{d\times d}$, which are linearly extended to $A \times \Delta(B)$ and $\Delta(B)$ respectively.[2] For technical reasons, assume that there is $c > 0$ such that $\|\sigma(b)\cdot y\| \geq c\,\|y\|\ \forall y \in \mathbb{R}^d$, $\forall b \in B$. Denote by $(\mathcal{F}_t)_{t\geq 0}$ the filtration generated by (X_t).

Small players have symmetric preferences. The payoff of each small player depends only on his own action, on the distribution over the small players' actions and on the sample path of the signal (X_t). A small player's payoff is

$$r\int_0^\infty e^{-rt}\left(u(b_t^i, \bar{b}_t)\,dt\,+\,v(b_t^i,\bar{b}_t)\cdot dX_t\right)$$

where $u : B \times \Delta(B) \to \mathbb{R}$ and $v : B \times \Delta(B) \to \mathbb{R}^d$ are bounded measurable functions. Then the expected payoff flow of the small players $h : A \times B \times \Delta(B) \to \mathbb{R}$ is given by

$$h(a,b,\bar{b}) = u(b,\bar{b}) + v(b,\bar{b})\cdot\mu(a,\bar{b}).$$

The small players' payoff functions are common knowledge.

The small players are uncertain about the type θ of the large player. At time 0 they believe that with probability $p \in [0,1]$ the large player is a *commitment type* ($\theta = $ c) and with probability $1 - p$ he is a *normal type* ($\theta = $ n). The commitment type mechanistically plays a fixed action $a^* \in A$ at all times. The normal type plays strategically to maximize his expected payoff. The payoff of the normal type of the large

---

[2]Functions μ and σ are extended to distributions over B by $\mu(a,\bar{b}) = \int_B \mu(a,b)\,d\bar{b}(b)$ and $\sigma(\bar{b}) = \int_B \sigma(b)\,d\bar{b}(b)$.

player is

$$r \int_0^\infty e^{-rt} \, g(a_t, \bar{b}_t) \, dt \,,$$

where the payoff flow is defined through a continuously differentiable function $g : A \times B \rightarrow \mathbb{R}$ that is extended linearly to $A \times \Delta(B)$.

The small players update their beliefs about the type of the large player by Bayes rule based on their observations of X. Denote by ϕ_t the probability that the small players assign to the large player being a commitment type at time $t \geq 0$.

A *public strategy* of the normal type of large player is a process $(a_t)_{t \geq 0}$ with values in A and progressively measurable with respect to $(\mathcal{F}_t)$. Similarly, a public strategy of small player $i \in I$ is a process $(b_t^i)_{t \geq 0}$ with values in B and progressively measurable with respect to $(\mathcal{F}_t)$. We assume that jointly the strategies of the small players and the aggregate distribution satisfy appropriate measurability properties.

Definition. *A public sequential equilibrium consists of a public strategy $(a_t)_{t \geq 0}$ of the normal type of large player, public strategies $(b_t^i)_{t \geq 0}$ of small players $i \in I$, and a progressively measurable belief process $(\phi_t)_{t \geq 0}$, such that at all times t and after all public histories:*

1. *the strategy of the normal type of large player maximizes his expected payoff*

$$E_t \left[r \int_0^\infty e^{-rt} \, g(a_t, \bar{b}_t) \, dt \mid \theta = \mathrm{n} \right]$$

2. *the strategy of each small player maximizes his expected payoff*

$$(1 - \phi_t) \, E_t \left[r \int_0^\infty e^{-rt} \, h(a_t, b_t^i, \bar{b}_t) \, dt \mid \theta = \mathrm{n} \right] \; + \; \phi_t \, E_t \left[r \int_0^\infty e^{-rt} \, h(a^*, b_t^i, \bar{b}_t) \, dt \mid \theta = \mathrm{c} \right]$$

3. *the common prior is $\phi_0 = p$ and beliefs $(\phi_t)_{t > 0}$ are determined by Bayes rule.*

A strategy profile satisfying conditions 1 and 2 is called *sequentially rational*. A belief process (ϕ_t) that satisfies condition 3 is called *consistent*.

Remark. Although the aggregate distribution of small players' actions is publicly observable, our requirement that public strategies depend only on the sample paths of X is without loss of generality. In fact, for a given strategy profile, the public histories along which there are observations of $\bar{b}_t$ that differ from those on-the-path-of-play correspond to deviations by a positive measure of small players. Therefore our definition of public strategies does not alter the set of public sequential equilibrium outcomes.

We are interested in the set of equilibrium payoffs that the normal type of the large player can achieve for a given prior p. Also, we would like to characterize the strategies in the equilibrium that achieves the best payoff for the large player for a given prior. In the next section we derive an ordinary differential equation that helps solve these problems. When the equation has a solution with appropriate boundary conditions, we fully characterize the best equilibrium for the large type for any prior. If the equation fails to have a solution, it can still be used to find the best and the worst equilibrium payoffs, as shown in Section 5. The equilibrium strategies, however, are no longer unique.

4 The Structure of Sequential Equilibria

This section derives our main results. We show that when the population is convinced that the large player is normal, i.e. $p = 0$, then the outcome cannot be better than the static Nash equilibrium. Also, we derive an ordinary differential equation that characterizes the best and the worst equilibrium payoff of the large player for any prior $p \in [0, 1]$. In a wide range of settings the ordinary differential equation has a solution with appropriate boundary conditions and the best equilibrium for the large player turns out to be unique. In that case, the equilibrium actions of all players are uniquely determined by the current belief of the small players.

Our analysis is based on the following descriptors of the equilibrium play at any moment of time: the public signal X, the small players' beliefs, the large player's continuation value, his incentives and actions. As the play of a public sequential equilibrium unfolds, these variables will interact. We characterize equilibria by deriving the laws of this interaction.

We proceed as follows. First, for a given strategy of the normal type of the large player, we find how the small players must update their beliefs from the observations of the public signal (X_t). Second, we derive a representation between the large player's continuation value and the public signal (X_t). This representation is important to formulate incentive compatibility conditions that tell us when the strategy of the normal type of the large player is optimal. We derive these results in Propositions 1, 2 and 3, and summarize them in Theorem 1. Ultimately, Theorem 1 characterizes public sequential equilibria in terms of the stochastic properties of signals X, beliefs ϕ and the large player's continuation value W.

In subsection 5, we use this characterization to derive our two main results. Theorem 3 is the main result, which characterizes the large player's equilibrium payoff as a function of the small players' beliefs via an ordinary differential equation.

We start with a proposition that explains how the small players use Bayes rule to update their beliefs based on the observations of the public signals.

Proposition 1 (Beliefs)**.** *Fix a public strategy $(a_t)_{t \geq 0}$ of the normal type of large player and an aggregate public strategy $(\bar{b}_t)_{t \geq 0}$ of the small players. Belief process $(\phi_t)_{t \geq 0}$ is consistent with strategies $(a_t, \bar{b}_t)_{t \geq 0}$ if and only if it satisfies equation*

$$d\phi_t = \gamma(a_t, \bar{b}_t, \phi_t) \cdot dZ_t^\phi \tag{1}$$

with initial condition $\phi_0 = p$, where

$$\gamma(a, \bar{b}, \phi) = \phi(1 - \phi)\sigma(\bar{b})^{-1}\left(\mu(a^*, \bar{b}) - \mu(a, \bar{b})\right), \tag{2}$$

$$dZ_t^\phi = \sigma(\bar{b}_t)^{-1}(dX_t - \mu^{\phi_t}(a_t, \bar{b}_t)\, dt) \tag{3}$$

$$\text{and} \quad \mu^\phi(a, \bar{b}) = \phi\mu(a^*, \bar{b}) + (1 - \phi)\mu(a, \bar{b}) \tag{4}$$

Proof. The strategies of the two types of large player induce two different probability measures over the paths of the signal (X_t). From Girsanov's Theorem we can find the ratio ξ_t between the likelihood that a path $(X_s : s \in [0, t])$ arises for type c and the likelihood that it arises for type n. This ratio is characterized by

$$d\xi_t = -\xi_t \rho_t \cdot dZ_s^n, \quad \xi_0 = 1, \tag{5}$$

where $\rho_t = \sigma(\bar{b}_t)^{-1} \left(\mu(a^*, \bar{b}_t) - \mu(a_t, \bar{b}_t) \right)$ and (Z_t^n) is a Brownian motion under the probability measure generated by type n's strategy.

Suppose that belief process (ϕ_t) is consistent with $(a_t, \bar{b}_t)_{t \geq 0}$. Then, by Bayes rule, the posterior after observing a path $(X_s : s \in [0, t])$ is

$$\phi_t = \frac{p\xi_t}{p\xi_t + (1-p)} \tag{6}$$

From Ito's formula,

$$
\begin{aligned}
d\phi_t &= \frac{p(1-p)}{(p\xi_t + (1-p))^2} d\xi_t - \frac{2p^2(1-p)}{(p\xi_t + (1-p))^3} \frac{\xi_t^2 \rho_t \cdot \rho_t}{2} dt \\
&= \phi_t(1-\phi_t)\rho_t \cdot dZ_t^n - \phi_t^2(1-\phi_t)(\rho_t \cdot \rho_t) \, dt \\
&= \phi_t(1-\phi_t)\rho_t \cdot dZ_t^\phi
\end{aligned} \tag{7}
$$

Conversely, suppose that beliefs (ϕ_t) satisfy equation (1) with initial condition $\phi_0 = p$. Define ξ_t using expression (6), i.e.,

$$\xi_t = \frac{1-p}{p} \frac{\phi_t}{1-\phi_t}.$$

By another application of Ito's formula, we conclude that (ξ_t) satisfies equation (5). This means that ξ_t is the ratio between the likelihood that a path $(X_s : s \in [0, t])$ arises for type c and the likelihood that it arises for type n. Hence, ϕ_t is determined by Bayes rule and the belief process is consistent.

$\square$

Coefficient γ in equation (1) is the volatility of beliefs: it reflects the speed with which the small players learn about the type of the large player. The definition of γ is important for our main ordinary differential equation that characterizes large player's equilibrium payoffs. The intuition behind equation (1) is as follows. If the small players are convinced about the type of the large player, then $\phi_t(1-\phi_t) = 0$, so they never change their beliefs. When $\phi_t \in (0, 1)$ then $\gamma(a_t, \bar{b}_t, \phi_t)$ is larger and learning is faster when the noise $\sigma(\bar{b}_t)$ is smaller or the drifts produced by the two types differ more. From the perspective of the small players, (Z_t^ϕ) is a Brownian motion and their belief (ϕ_t) is a martingale. From (7) we see that, conditional on the large player being the normal type, the drift of ϕ_t is negative: the small players eventually learn the true type of the large player.

We now proceed to analyze the second important state descriptor of the interaction between the large and the small players, the continuation value of the normal type of the large player. A player's continuation value is his future expected payoff in equilibrium after a given history. We derive how the large player's incentives arise from the law of motion of his continuation value. We will find that the large player's strategy is optimal if, and only if, a certain incentive compatibility condition holds at all times $t \geq 0$.

For a given strategy profile $(a_t, b_t)_{t \geq 0}$ define the *continuation value* W_t of the normal type at time $t \geq 0$ as

$$W_t := E_t \left[r \int_t^\infty e^{-r(s-t)} g(a_s, \bar{b}_s) \, ds \mid \theta = n \right]$$

In order to be able to isolate the large player's instantaneous incentives, we must represent the law of motion of (W_t) in terms of its drift and its sensitivity to X.

12

Proposition 2 (Continuation Values). *Let $(a_t, \bar{b}_t)$ be a public-strategy profile. The corresponding continuation values (W_t) of the normal type satisfy*

$$dW_t = r(W_t - g(a_t, \bar{b}_t)) \, dt \; + \; r\beta_t \cdot (dX_t - \mu(a_t, \bar{b}_t) \, dt) \tag{8}$$

where (β_t) is some progressively measurable process with values in $\mathbb{R}^d$.

Proof. Following Sannikov (2004), let V_t denote the average discounted payoff of the normal type conditional on $\mathcal{F}_t$, i.e.,

$$V_t = E_t \left[r \int_0^\infty e^{-rs} g(a_s, \bar{b}_s) \, ds \; \mid \; \theta = \mathrm{n} \right]$$

With respect to the probability measure induced by the normal type, (V_t) is a bounded martingale. Therefore it has a representation of the form:

$$dV_t = r e^{-rt} \beta_t \cdot \sigma(\bar{b}_t) dZ_t^{\mathrm{n}} \tag{9}$$

where $dZ_t^{\mathrm{n}} = \sigma(\bar{b}_t)^{-1} (dX_t - \mu(a_t, \bar{b}_t) \, dt)$ is a Brownian motion from the point of view of the normal type of the large player, $\beta_t = r^{-1} e^{rt} \frac{d}{dt} \langle V, Z^{\mathrm{n}} \rangle_t$ and the brackets operation $\langle \cdot, \cdot \rangle$ is called the *cross-variation*.

From the definition of (W_t) it follows that

$$V_t = r \int_0^t e^{-rs} g(a_s, \bar{b}_s) \, ds \; + \; e^{-rt} W_t$$

Differentiating on both sides yields:

$$dV_t = r e^{-rt} g(a_t, \bar{b}_t) \, dt \; - \; r e^{-rt} W_t \, dt \; + \; e^{-rt} dW_t \tag{10}$$

Comparing equations (9) and (10) yields the desired result. $\qquad\square$

Next, we derive conditions for sequential rationality. The condition for the small players is straightforward: they maximize their static payoff because a deviation of any individual small player does not affect the future equilibrium play. The situation of the normal type of the large player is more complicated: he acts optimally if he maximizes the sum of the current payoff flow and the expected change in his continuation value.

Proposition 3 (Incentive Compatibility). *A public-strategy profile $(a_t, \bar{b}_t)_{t \geq 0}$ is sequentially rational with respect to a belief process (ϕ_t) if, and only if, for all times $t \geq 0$ and after all public histories:*

$$a_t \; \in \; \underset{a' \in A}{\arg \max} \; g(a', \bar{b}_t) \; + \; \beta_t \cdot \mu(a', \bar{b}_t) \tag{11}$$

$$b \; \in \; \underset{b' \in B}{\arg \max} \; u(b', \bar{b}_t) + v(b', \bar{b}_t) \cdot \mu^{\phi_t}(a_t, \bar{b}_t), \quad \forall b \in \mathrm{supp} \, \bar{b}_t \tag{12}$$

Proof. Let (a_t, b_t^i) be a strategy profile and $(\tilde{a}_t)$ a strategy of the normal type. Denote by (W_t) the continuation values of the large player when profile (a_t, b_t^i) is played. If the normal type of large player plays strategy $(\tilde{a}_t)$ up to time t and then switches back to (a_t), his expected payoff conditional on $\mathcal{F}_t$ is given by:

$$\tilde{V}_t = r \int_0^t e^{-rs} g(\tilde{a}_s, \bar{b}_s) \, ds \; + \; e^{-rt} W_t$$

13

By Proposition 2 and the expression above,

$$
\begin{aligned}
d\tilde{V}_t &= re^{-rt}\left(g(\tilde{a}_t,\bar{b}_t) - W_t\right)dt \ + \ e^{-rt}dW_t \\
&= re^{-rt}\left((g(\tilde{a}_t,\bar{b}_t) - g(a_t,\bar{b}_t))\,dt + \beta_t\cdot(dX_t - \mu(a_t,\bar{b}_t)\,dt)\right)
\end{aligned}
$$

where the $\mathbb{R}^d$-valued process (β_t) is given by Proposition 2.

Hence the profile $(\tilde{a}_t,\bar{b}_t)$ yields the normal type the following expected payoff:

$$
\begin{aligned}
\tilde{W}_0 &= E[\tilde{V}_\infty] = E\left[\tilde{V}_0 \ + \ \int_0^\infty d\tilde{V}_t\right] \\
&= W_0 \ + \ E\left[r\int_0^\infty e^{-rt}\left(g(\tilde{a}_t,\bar{b}_t) - g(a_t,\bar{b}_t) \ + \ \beta_t\cdot(\mu(\tilde{a}_t,\bar{b}_t) - \mu(a_t,\bar{b}_t))\right)dt\right]
\end{aligned}
$$

where we used the fact that $\tilde{V}_0 = W_0$ and that (X_t) has drift $\mu(\tilde{a}_t,\bar{b}_t)$ under the probability measure induced by $(\tilde{a}_t,\bar{b}_t)$.

Suppose that strategy profile $(a_t,\bar{b}_t)$ and belief process (ϕ_t) satisfy the IC conditions (11) and (12). Then, for every $(\tilde{a}_t)$, one has $W_0 \geq \tilde{W}_0$, and the normal type is sequentially rational at time 0. By a similar argument, the normal type is sequentially rational at all times t, after all public histories. Also, note that small players are maximizing their instantaneous expected payoffs. Since the small players' actions are anonymous, no unilateral deviation by a small player can affect the future course of play. Therefore each small player is also sequentially rational.

Conversely, suppose that IC condition (11) fails. Choose a strategy $(\tilde{a}_t)$ such that $\tilde{a}_t$ attains the maximum in (11) for all $t \geq 0$. Then $\tilde{W}_0 > W_0$ and the large player is not sequentially rational. Likewise, if condition (12) fails, then a positive measure of small players is not maximizing their instantaneous expected payoffs. By the anonymity of small player's actions, this implies that a positive measure of small players is not sequentially rational. $\qquad\square$

Denote by $\mathcal{E} : [0,1] \rightrightarrows \mathbb{R}$ the correspondence that maps a prior probability $p \in [0,1]$ on the commitment type into the set of public sequential equilibrium payoffs of the normal type in the game with prior p. Below we summarize the previous results and state the recursive characterization of $\mathcal{E}$.

Theorem 1 (Sequential Equilibrium). *For a given prior $p \in [0,1]$ on the commitment type, we have that $w \in \mathcal{E}(p)$ if, and only if, there exist a public strategy profile (a_t,b_t) and a progressively measurable process (β_t) such that:*

1. *For all $t \geq 0$ and all public histories, $W_t \in \mathcal{E}(\phi_t)$, where (ϕ_t) and (W_t) denote the corresponding solutions of equations (1) and (8) with initial conditions $\phi_0 = p$ and $W_0 = w$, respectively.*

2. *the Incentive Compatibility conditions (11) and (12) hold for profile (a_t,b_t), beliefs (ϕ_t) and process (β_t)*

Therefore, correspondence $\mathcal{E}$ is the largest correspondence[3] such that a controlled process (ϕ,W), defined by (1) and (8), can be kept in $\mathrm{Graph}(\mathcal{E})$ by controls (a_t,b_t) and (β_t) that satisfy (11) and (12).

---

[3]This means that there is no other correspondence with this property whose graph contains the graph of $\mathcal{E}$ as a proper subset

5 Equilibrium Degeneracy under Complete Information

In this section we examine the structure of the set of equilibrium payoffs of the large player in the complete information game ($p = 0$), i.e., in the game in which it is common knowledge that the large player is the normal type.

Theorem 2. *Suppose that the population of small players is convinced that the large player is normal, i.e. $p = 0$. Then in any public sequential equilibrium the large player cannot achieve a payoff outside the convex hull of his stage-game Nash equilibrium payoff set, i.e.*

$$\mathcal{E}(0) = \text{co} \left\{ g(a, \bar{b}) \ : \ a \in \arg\max_{a' \in A} g(a', \bar{b}) \ , \ b \in \arg\max_{b' \in B} h(a, b', \bar{b}) \ , \ \forall b \in \text{supp } \bar{b} \right\}$$

Proof. Let $\bar{v}$ be the highest pure-strategy Nash equilibrium payoff of the large player in the static game. We show that it is impossible to achieve a payoff higher than $\bar{v}$ in any public equilibrium. (A proof for the lowest Nash equilibrium payoff is similar). Suppose there was a public equilibrium in which the large player's continuation value W_0 was greater than $\bar{v}$. By Proposition 3, for some progressively measurable process (β_t), the large player's continuation value must follow the SDE

$$dW_t = r(W_t - g(a_t, \bar{b}_t)) \, dt + r\beta_t \cdot (dX_t - \mu(a_t, \bar{b}_t) \, dt),$$

where a_t maximizes $g(a', \bar{b}_t) + \beta_t \mu(a', \bar{b}_t)$. Denote $\bar{D} = W_0 - \bar{v}$. Let us show that as long as $W_t \geq \bar{v} + \bar{D}/2$, either the drift of W_t is greater than $r\bar{D}/2$ or the volatility of W_t is uniformly bounded away from zero. If $g(a_t, \bar{b}_t) < \bar{v} + \bar{D}/2$ then the drift of W_t is greater than $r\bar{D}/2$. If $g(a_t, \bar{b}_t) \geq \bar{v} + \bar{D}/2$, then, by Lemma (1) in the Appendix (making $\phi = 0$ identically), $|\beta_t| \geq \epsilon$ for some strictly positive ϵ (independent of t and on the sample path). Therefore, W_t becomes arbitrarily large with positive probability, a contradiction. ∎

The intuition behind this result is as follows. In order to give incentives to the large player to take an action that results in a payoff better than static Nash, his continuation value must respond to the public signal X. When his continuation value reaches its upper bound, such incentives cannot be provided. In effect, if at the upper bound the large player's continuation value were sensitive to the public signal process (X_t), then, with positive probability, the process of continuation value would escape above this upper bound, which is not possible. Therefore, at the upper bound, continuation values cannot depend on the public signal and so, in the best equilibrium, the normal type must be playing a myopic best response.

6 Reputation Effects and the Optimality Equation

When there is a possibility that the large player is a commitment type, the normal type of the large player does not need incentives in the best equilibrium outcome. The best outcome is achieved when the population is certain that he is a commitment type who plays action a^*. Then the normal type can choose any action without altering the small players' belief, and the set of payoffs that he can achieve is defined by

$$\mathcal{E}(1) = \text{co} \left\{ g(a, \bar{b}) \ : \ a \in \arg\max_{a' \in A} g(a', \bar{b}) \ , \ b \in \arg\max_{b' \in B} h(a^*, b', \bar{b}) \ , \ \forall b \in \text{supp } \bar{b} \right\}.$$

Now we shall present our main result: the upper and lower boundaries of the correspondence $\mathcal{E}$ can be characterized by means of a pair of ordinary differential equations.

Definition. *A continuous function* $U : [0,1] \to \mathbb{R}$ *is a* solution of the Upper Optimality Equation *if it is twice continuously differentiable in the open interval* $(0,1)$ *and satisfies the* Upper Optimality Equation:

$$
\begin{aligned}
rU(\phi) \;=\; &\max_{(a,\bar{b})} \quad rg(a,\bar{b}) - \frac{\left\| \gamma(a,\bar{b},\phi) \right\|^2}{(1-\phi)} U'(\phi) + \tfrac{1}{2} \left\| \gamma(a,\bar{b},\phi) \right\|^2 U''(\phi) \\
&\text{subject to} \quad (a,\bar{b}) \in \Psi(\phi, U'(\phi))
\end{aligned}
$$

with boundary conditions $U(0) = \max \mathcal{E}(0)$ *and* $U(1) = \max \mathcal{E}(1)$, *where* γ *and* μ^{ϕ} *are defined by (2) and (4), and the correspondence* $\Psi : [0,1] \times \mathbb{R} \rightrightarrows A \times \Delta(B)$ *is defined by:*

$$
\Psi(\phi, z) := \left\{ (a,\bar{b}) : \begin{array}{l} a \in \displaystyle\arg\max_{a' \in A} \; rg(a',\bar{b}) + \gamma(a,\bar{b},\phi) \cdot \sigma(\bar{b})^{-1} \mu(a',\bar{b}) z \\[2mm] b \in \displaystyle\arg\max_{b' \in B} \; u(b',\bar{b}) + v(b',\bar{b}) \cdot \mu^{\phi}(a,\bar{b}), \quad \forall b \in \mathrm{supp}(\bar{b}) \end{array} \right\}
$$

A solution $U : [0,1] \to \mathbb{R}$ *of the Upper Optimality Equation is* regular *if it satisfies:*

1. $\lim_{\phi \to 0} \phi(1-\phi)U'(\phi) = \lim_{\phi \to 1} \phi(1-\phi)U'(\phi) = 0$

2. $\lim_{\phi \to 0} \phi^2(1-\phi)^2 U''(\phi) = \lim_{\phi \to 1} \phi^2(1-\phi)^2 U''(\phi) = 0$

A *Lower Solution* is defined in an analogous way, replacing the max operator that appears in the first line of the equation and in the boundary conditions by a min operator. Figure 4 shows a typical form of the correspondence $\mathcal{E}$ for the case when the stage game has more than one Nash equilibrium, but the small players' best response to a^* is unique. In this case $\mathcal{E}(0)$ is an interval, while $\mathcal{E}(1)$ is a single point.

Figure 4: A typical form of $\mathcal{E}$.

Remark. The role of the regularity conditions above is to ensure that the objective function on the RHS of the Optimality Equation is continuous at 0 and 1, and that the correspondence $\phi \mapsto \Psi(\phi, U'(\phi))$ is upper hemi-continuous at 0 and 1.

Our main result is:

Theorem 3. *Suppose that the Upper and Lower Optimality equations have regular solutions U and L, respectively. Then, for every prior probability $p \in [0,1]$ on the commitment type, one has:*

$$
\mathcal{E}(p) = [L(p), U(p)]
$$

Proof. First, we shall prove that for any prior probability $p \in [0,1]$, all sequential equilibria yield the normal type a payoff of $U(p)$ or lower. Suppose, towards a contradiction, that for some $p \in [0,1]$, $(a_t, \bar{b}_t)_{t \geq 0}$ is a sequential equilibrium profile that yields the normal type a payoff of $W_0 > U(p)$. Denote by (ϕ_t) and (W_t)

the corresponding beliefs and continuation values of the normal type, respectively. Also, let $U_t = U(\phi_t)$. By Itô's Lemma and Proposition 1 it follows that (U_t) has drift given by

$$-\frac{\left\|\gamma(a_t,\bar{b}_t,\phi_t)\right\|^2}{1-\phi_t}\,U'(\phi_t) + \frac{1}{2}\left\|\gamma(a_t,\bar{b}_t,\phi_t)\right\|^2 U''(\phi_t),$$

and volatility given by

$$\gamma(a_t,\bar{b}_t,\phi_t)\,U'(\phi_t).$$

Also, by Proposition 2, we have that (W_t) follows equation (8) for some progressively measurable process (β_t). In addition, equilibrium strategy profile (a_t, b_t) must satisfy, together with beliefs (ϕ_t) and process (β_t), the IC conditions (11) and (12).

Now define a process $D_t = W_t - U_t$. The volatility of D is given by

$$r\beta_t\sigma(\bar{b}_t) - \gamma(a_t,\bar{b}_t,\phi_t)\,U'(\phi_t),$$

and the drift of D is

$$rD_t + rU_t - \left(-\frac{\left\|\gamma(a_t,\bar{b}_t,\phi_t)\right\|^2}{1-\phi_t}\,U'(\phi_t) + \frac{1}{2}\left\|\gamma(a_t,\bar{b}_t,\phi_t)\right\|^2 U''(\phi_t)\right).$$

In the expression above, notice that the term in parentheses is the objective function of the program on the RHS of the Upper Optimality Equation, evaluated at $a_t, \bar{b}_t$ and ϕ_t.

We claim that for all $t \geq 0$, either:

(a) the drift of D is greater than or equal to $rD_0/2$ at time t; or

(b) the volatility of D is bounded away from zero uniformly (in t and sample path)

The proof of this claim is postponed and can be found in Lemma 1 in the Appendix. Here we provide a crude intuition: when the volatility of D is exactly zero, the large player's IC condition (11) coincides with the corresponding constraint in the Upper Optimality equation. We reproduce these constraints below:

$$(11): \qquad a_t \in \arg\max_{a' \in A} \; rg(a',\bar{b}_t) + r\beta_t \cdot \mu(a',\bar{b}_t)$$
$$(\text{Optim. Eq.}): \quad a \in \arg\max_{a' \in A} \; rg(a',\bar{b}) + \gamma(a,\bar{b},\phi)\cdot\sigma(\bar{b})^{-1}\mu(a',\bar{b})\,U'(\phi)$$

The IC for the myopic type is always the same as his constraint in the Optimality equation. Therefore the pair $(a_t, \bar{b}_t)$ is feasible for the program on the RHS of the Upper Optimality Equation associated to ϕ_t and, consequently, the drift of D at t is no smaller than rD_t. The claim then follows from a continuity argument (spelled out in Lemma 1 in the Appendix).

By (a) and (b) above it follows that (D_t) is unbounded with positive probability, which is a contradiction since (W_t) and (U_t) are bounded processes. The contradiction shows that for any prior $p \in [0, 1]$, there cannot be an equilibrium that yields the normal type of large player a payoff larger than $U(p)$. In a similar way, it can be shown that no equilibrium yields payoffs below $L(p)$. Therefore,

$$\mathcal{E}(p) \subset [L(p), U(p)].$$

We shall now prove that for all $p \in [0, 1]$, there is a public sequential equilibrium that yields the large player a payoff of $U(p)$. For each $\phi \in [0, 1]$, let $a(\phi)$ and $\bar{b}(\phi)$ be values of a and $\bar{b}$ that attain the maximum in the RHS of the upper optimality equation. For a given prior probability $p \in [0, 1]$, let (ϕ_t) be the solution of equation:

$$d\phi_t = \gamma \left(a(\phi_t), \bar{b}(\phi_t), \phi_t \right) \cdot \sigma(\bar{b}(\phi_t))^{-1} \left(dX_t - \mu^{\phi_t}(a(\phi_t), \bar{b}(\phi_t)) \, dt \right)$$

with initial condition $\phi_0 = p$. Define strategies $a_t = a(\phi_t)$ and $b_t = b(\phi_t)$. By Proposition 1, belief process $(\phi_t)_{t \geq 0}$ is consistent with strategy profile $(a_t, \bar{b}_t)_{t \geq 0}$.

Now define a process (β_t) by

$$\beta_t = r^{-1} \gamma(a_t, \bar{b}_t, \phi_t) \cdot \sigma(\bar{b}_t)^{-1} U'(\phi_t)$$

Let $W_t = U(\phi_t)$. By Itô's Lemma,

$$
\begin{aligned}
dW_t &= U'(\phi_t) \, d\phi_t + \left\| \gamma(a_t, \bar{b}_t, \phi_t) \right\|^2 \frac{U''(\phi_t)}{2} \, dt \\
&= \left\| \gamma(a_t, \bar{b}_t, \phi_t) \right\|^2 \left(-\frac{U'(\phi_t)}{1 - \phi_t} + \frac{1}{2} U''(\phi_t) \right) dt + \gamma(a_t, \bar{b}_t, \phi_t) U'(\phi_t) \cdot dZ_t^{\mathrm{n}}
\end{aligned}
$$

where $dZ_t^{\mathrm{n}} = \sigma(\bar{b}_t)^{-1} \left(dX_t - \mu(a_t, \bar{b}_t) dt \right)$. By the optimality equation and the definitions of W_t and β_t above, it follows that (W_t) satisfies equation:

$$dW_t = r(W_t - g(a_t, \bar{b}_t)) \, dt + r\beta_t \cdot (dX_t - \mu(a_t, \bar{b}_t) \, dt)$$

By Proposition 2, process (W_t) is the process of continuation values associated with strategy profile $(a_t, \bar{b}_t)$, and (β_t) is its sensitivity to (X_t). In addition, it follows from the constraints of the optimality equation and the definition of β_t that profile $(a_t, \bar{b}_t)$ satisfies the IC conditions (11) and (12) with respect to beliefs (ϕ_t) and sensitivity process (β_t). By Proposition 3 we conclude that $(a_t, \bar{b}_t)$ is sequentially rational. Therefore strategy profile $(a_t, \bar{b}_t)$ together with beliefs (ϕ_t) is a public sequential equilibrium that yields a payoff of $W_0 = U(p)$ to the large player. □

7 Summary and Discussion.

In this section we summarize our main results and explain what happens during the play of an equilibrium that achieves the best payoff for the normal type of the large player. First, we consider the standard case when the optimality equation has an upper solution $U : [0, 1] \to \mathbb{R}$. At the end of the section we explain why the optimality equation may not have an appropriate solution, and what happens in that case. Assuming that U exists, the dynamics in the best equilibrium for the normal type of the large player is uniquely defined. During the play of this equilibrium, the action of the large player and the distribution of the small players' actions are uniquely determined by the current belief of the small players ϕ_t. The maximizers $a : [0, 1] \to A$ and $\bar{b} : [0, 1] \to \Delta(B)$ in the optimality equation determine the current actions of all players as follows

$$a_t = a(\phi_t), \quad \text{and} \quad b_t = b(\phi_t).$$

During the course of the play, the small players update their beliefs by Bayes rule based upon their observations of the signal X. The sensitivity of the belief ϕ_t towards X is captured by

$$\gamma(a_t, \bar{b}_t, \phi_t) = \phi_t(1 - \phi_t)\sigma(\bar{b}_t)^{-1} \left(\mu(a^*, \bar{b}_t) - \mu(a_t, \bar{b}_t) \right).$$

The small players update faster when the volatility of signal is smaller, and when the actions of the two types of the large player are far away. Eventually, the small players become more convinced about the type of the large player. As a result, the payoff of the normal type converges to a static Nash equilibrium payoff.

However, the small players can take a long time to learn about the large player's type. When the commitment payoff of the large type is greater than his static Nash equilibrium payoff, the large player's equilibrium continuation value $U(\phi)$ is strictly increasing in the small players' belief ϕ. Because he benefits from positive beliefs, the normal type shifts away from the static best response towards the commitment action a^*. Intuitively, the normal type must imitate the commitment type closer when the benefit from improved beliefs $U'(\phi)$ is greater, when the signals are more transparent (smaller σ), and when he cares more about the future (smaller r). These incentives weaken when ϕ_t is closer to 0 or 1, i.e. the small players update their beliefs slowly because they are sufficiently convinced about the large player's type. Formally, at all times the large type's action a_t maximizes

$$\max_{a'} \ rg(a', \bar{b}_t) + \frac{\phi_t(1 - \phi_t)\left(\mu(a^*, \bar{b}_t) - \mu(a_t, \bar{b}_t)\right) \cdot \mu(a', \bar{b}_t) U'(\phi_t)}{\sigma(\bar{b}_t)^2}, \tag{13}$$

where μ and σ were taken to be one-dimensional for greater clarity. The first term captures the large player's current payoff, and the second term reflects the expected effect of his current action on the small players' beliefs. Note that the benefit from imitating decreases as the normal type takes an action closer to a^*, as reflected by the factor $\mu(a^*, \bar{b}_t) - \mu(a_t, \bar{b}_t)$ in the numerator of the right hand side. This happens because with greater imitation the signal X becomes less informative about the large player's type.

The incentives of the small players are much more straightforward: they play a static best response to the expected action of the large player. Intuitively, at any moment the strategic interaction between the large and the small players can be interpreted as a static game, in which the payoffs of the large player are perturbed because he also cares about the beliefs of the small players. When the large player is patient, the beliefs matter more and he imitates the commitment type better. Figure 5 shows sample paths of the small players' beliefs in an equilibrium of a quality commitment game from Section 2 for different discount rates. We see that as $r \to 0$, the small players take longer to learn the type of the large player due to greater imitation.

Figure 5: Sample paths of beliefs and actions in the quality commitment game.

In this example the equilibrium is unique. The signal X determines fully how the beliefs evolve and the beliefs completely determine the current actions of all players. The reason for uniqueness is that the best momentary payoff of the large player is continuous in how much he cares about the beliefs of the small players.

Non-uniqueness may arise if the payoff of the large player is discontinuous. To illustrate what we mean, consider the following variation of the game of quality commitment. Suppose that the payoff of the small players is given by $b^i dX_t - b^i dt$ as before, where

$$dX_t = \sqrt{a_t}(1 + \bar{b}_t) \, dt + (2 - \bar{b}_t) \, dZ_t.$$

The noise is still decreasing with usage as before. However, now there are positive network externalities reflected in the drift of X: customers benefit from others using the same service. As before, the payoff flow of the large player is $\bar{b}_t - a_t$. Suppose that the sets of stage-game actions are $A = B = [0, 1]$.

This example involves a discontinuity in the large player's payoff because the buyers require an expected investment level of at least $1/4$ in order to fully participate by choosing $b = 1$. If the buyers expect an investment level of less than 1/4, nobody participates and the seller's payoff drops discontinuously. To see what happens in equilibrium, assume that the commitment type chooses $a^* = 1/2$ and note that the static best response of the large player is always 0. Figure 6 shows the maximal equilibrium payoff of the large player as a function of the population's belief. The curve U has three portions: the black portion that satisfies the optimality equation with $\bar{b} = 1$, the blue portion that satisfies the optimality equation with $\bar{b} = 0$ and the red portion in between at which the optimality equation is discontinuous. The slope of the red portion is the critical slope where the possibility that the small players fully use the service is barely available.
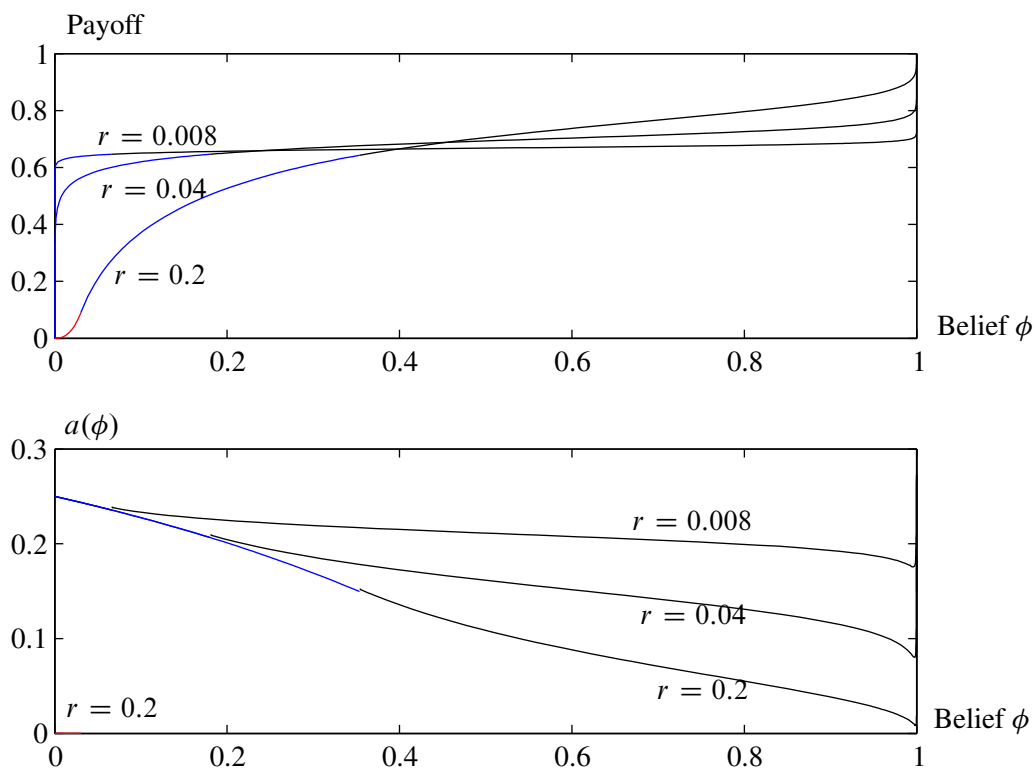

Figure 6: Equilibrium of the quality commitment game with network effects.

The equilibrium play is determined uniquely in the best equilibrium for the large player as long as the belief of the small players is belongs to the black or blue portion of U. As soon as the belief reaches the red portion, the equilibrium play is no longer unique. In a typical equilibrium, the pair (ϕ_t, W_t) drifts underneath the curve U. Therefore, even if the play starts from the best equilibrium for the large player for a given belief, the equilibrium is not necessarily the best after a sufficient amount of time.

To conclude, in a quality commitment game with network effects, the large player must possess a critical reputation level in order to convince the small players to coordinate on full participation. Once the full participation breaks down, the equilibrium becomes non-unique. In the next section, we extend our characterization to the case of multiple equilibria.

8 Conclusion.

We explore the issue of commitment in a setting where the actions of a large player who would like to commit are imperfectly observable. In this setting it is very difficult to construct equilibria in a repeated discrete-time model. On the other hand, in a continuous-time model we find a clean and tractable characterization of equilibria. The boundary of the set of equilibrium belief-value pairs is described by an ordinary differetial equation. The choice of equilibrium actions is conveniently determined by the population's belief about the type of the large player and the rate at which his payoff changes with the population's belief.

We aim to bring the theoretical contributions of this paper closer to applications. The situation in which a large player would like to commit to a policy to influence the behavior of small players is very common in economics. We presented three different examples from three fields in which our results apply. In these examples, we can study how the commitment constraints of the large player distort the equilibrium outcome away from the optimum, and to what extent. In our example of a firm that is trying to attract investors, there are two sources of inefficiency: the manager diverts funds from investors, and there is underinvestment. The investment decreases with volatility due to the agency problem. Also, we can explore what features of a model ease commitment. In the investment example, we can study the effect of a requirement that the manager holds his company's stock. In the example with quality commitment, we can study how the service provider can signal with his pricing schedule that he is determined to carry out commitment.

Appendix

The following parametrization of the Upper Optimality Equation is useful for the proof that will follow:

$$rU(\phi) = \max_{\varrho \in \mathbb{R}^d} \max_{(a, \bar{b}) \in A \times \Delta(B)} rg(a, \bar{b}) + \phi^2 (1 - \phi)^2 \left(\tfrac{1}{2} U''(\phi) - \tfrac{U'(\phi)}{1 - \phi} \right) \|\varrho\|^2$$

$$\text{subject to}$$

$$(a, \bar{b}) \in \Phi(\phi, \varrho, r^{-1}\phi(1 - \phi)U'(\phi)\varrho)$$

where

$$\varrho(a, \bar{b}) = \sigma(\bar{b})^{-1}(\mu(a^*, \bar{b}) - \mu(a, \bar{b})),$$

and correspondence $\Phi : [0, 1] \times \mathbb{R}^d \times \mathbb{R}^d \rightrightarrows A \times \Delta(B)$ is defined as

$$\Phi(\phi, \varrho, \beta) = \left\{ (a, \bar{b}) \in A \times \Delta(B) \left| \begin{array}{rcl} a & \in & \arg\max_{a' \in A} g(a', \bar{b}) + \beta^\top \sigma(\bar{b})^{-1}\mu(a', \bar{b}) \\ \bar{b} & \in & \arg\max_{\bar{b}' \in \Delta(B)} \phi\, h(a^*, \bar{b}', \bar{b}) + (1 - \phi)\, h(a, \bar{b}', \bar{b}) \\ \varrho & = & \varrho(a, \bar{b}) \end{array} \right. \right\}$$

for each $(\phi, \varrho, \beta) \in [0, 1] \times \mathbb{R}^{d \times d}$.

Recall that a function $U : [0, 1] \to \mathbb{R}$ is a *regular solution of the Upper Optimality Equation* if it is continuous, C^2 in the open interval $(0, 1)$, satisfies the equation above at each $\phi \in (0, 1)$ and satisfies the boundary conditions $U(0) = \max \mathcal{E}(0)$, $U(1) = \max \mathcal{E}(1)$ and the regularity conditions below:

1. $\lim_{\phi \to 0} \phi(1 - \phi)U'(\phi) = \lim_{\phi \to 1} \phi(1 - \phi)U'(\phi) = 0$

2. $\lim_{\phi \to 0} \phi^2 (1-\phi)^2 U''(\phi) = \lim_{\phi \to 1} \phi^2 (1-\phi)^2 U''(\phi) = 0$

For future reference, given a regular solution U, let v and w denote the continuous functions defined by:

$$v(\phi) := \begin{cases} \phi(1-\phi)U'(\phi) & : \quad \phi \in (0,1) \\ 0 & : \quad \phi = 0 \text{ or } 1 \end{cases}$$

and

$$w(\phi) := \begin{cases} \phi^2(1-\phi)^2 U''(\phi)/2 & : \quad \phi \in (0,1) \\ 0 & : \quad \phi = 0 \text{ or } 1 \end{cases}$$

Let $(a_t, \bar{b}_t, \phi_t)_{t \geq 0}$ be a public sequential equilibrium and let (β_t) denote the corresponding sensitivity process. The incentive constraints that $(a_t, \bar{b}_t)$ must satisfy (Proposition 3) can be written as:

$$(a_t, \bar{b}_t) \in \Phi(\phi_t, \varrho(a_t, \bar{b}_t), \beta_t)$$

Let (W_t) be the continuation-value process associated with this equilibrium and let $U_t = U(\phi_t)$, where ϕ is the belief process. Recall that $D = W - U$ satisfies:

$$dD_t = (rD_t + rU_t - F_t)\, dt \ + \ (r\beta_t - v(\phi_t)\varrho_t) \cdot dZ_t$$

where

$$\begin{aligned} \varrho_t &= \varrho(a_t, \bar{b}_t)\,, \\ F_t &= f(a_t, \bar{b}_t, \phi_t, v(\phi_t), w(\phi_t), \varrho(a_t, \bar{b}_t))\,, \\ f(a, \bar{b}, \phi, v, w, \varrho) &= rg(a, \bar{b}) + (w - \phi v)\, \|\varrho\|^2\,. \end{aligned}$$

Lemma 1 (Continuity). *Suppose the Upper Optimality Equation has a regular solution $U : [0,1] \to \mathbb{R}$. Let $(a_t, \bar{b}_t, \phi_t)$ be a public sequential equilibrium. Then, for all $\epsilon > 0$, there exists $\delta > 0$ such that, almost everwhere, if the absolute value of the volatility of D is less than δ, then the drift of D is greater than $rD - \epsilon$.*

Proof. Let $f^* : [0,1] \times \mathbb{R}^2 \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ denote the map defined by:

$$f^*(\phi, v, w, \varrho, \beta) = \max_{(a,\bar{b}) \in \Phi(\phi, \varrho, \beta)} f(a, \bar{b}, \phi, v, w, \varrho)$$

Since f is continuous and Φ has a closed graph, f^* is upper semi-continuous. Also, note that, by the definition of f^*, by the optimality equation and by the boundary conditions, one has:

$$rU(\phi) \geq f^*(\phi, v(\phi), w(\phi), \varrho, v(\phi)\varrho) \tag{14}$$

for all $\varrho \in \mathbb{R}^d$ and $\phi \in [0,1]$.

Fix $\epsilon > 0$ and consider the sets F and K defined as:

$$\begin{aligned} F &= \{(\phi, \varrho, \beta) \in [0,1] \times \mathbb{R}^d \times \mathbb{R}^d \mid f^*(\phi, v(\phi), w(\phi), \varrho, \beta) \geq rU(\phi) + \epsilon\}\,, \\ K &= \{(\phi, \varrho, \beta) \in [0,1] \times \mathbb{R}^d \times \mathbb{R}^d \mid \beta = r^{-1}v(\phi)\varrho, \ \|\varrho\| \leq R\}\,, \end{aligned}$$

where $R = \max_{(a,\bar{b})} \|\varrho(a, \bar{b})\|$.

By the continuity of U, v and w, and by the upper semi-continuity of f^*, F is a closed set. Also, by the continuity of v, K is compact. Moreover, it follows from (14) that $K \cap F = \emptyset$. Hence $\eta := \texttt{dist}(K, F) > 0$.

In sum, we have shown that $\forall \epsilon > 0 \; \exists \eta > 0$ such that $\forall (\phi, \varrho, \beta) \in [0, 1] \times \bar{B}_R(0) \times \mathbb{R}^d$,

$$|\beta - r^{-1} v(\phi) \varrho| < \eta \quad \Rightarrow \quad r U(\phi) - f^*(\phi, v(\phi), w(\phi), \varrho, \beta) > -\epsilon \; . \tag{15}$$

Fix $\epsilon > 0$ and let $\eta > 0$ be such that the above implication holds. Let $\delta = r\eta$. Fix a time $t \geqq 0$ and suppose that the volatility of D at t is less than δ, i.e.,

$$\left\| r\beta_t - v(\phi_t)\varrho(a_t, \bar{b}_t) \right\| < \delta \; .$$

It follows from (15) that:

$$
\begin{aligned}
rD_t - \epsilon \quad &< \quad rD_t + rU(\phi_t) - f^*(\phi_t, v(\phi_t), w(\phi_t), \varrho(a_t, \bar{b}_t), r\beta_t) \\
&\leqslant \quad rD_t + rU_t - F_t \\
&= \quad \text{drift of } D \text{ at } t \; ,
\end{aligned}
$$

where the last inequality follows from the fact that $(a_t, \bar{b}_t) \in \Phi(\phi_t, \varrho(a_t, \bar{b}_t), \beta_t)$. $\qquad\qquad\square$

Remark. Notice the role played by the regularity conditions: they guarantee that v and w are continuous at the boundary points, ensuring that F and K are closed and compact sets, respectively.

References

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

DEMARZO, P., AND M. FISHMAN (2003): "Optimal Long-Term Financial Contracting with Privately Observed Cash Flows," Working Paper, Northwestern University.

DEMARZO, P., AND Y. SANNIKOV (2004): "A Continuous-Time Agency Model of Optimal Contracting and Capital Structure," Working Paper, Stanford University.

FAINGOLD, E. (2005): "Building a Reputation Under Frequent Decisions," Working Paper, 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.

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): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

KYDLAND, F., AND E. PRESCOTT (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3), 473–492.

MILGROM, P., AND J. ROBERTS (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*.

SANNIKOV, Y. (2004): "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.