Belief-free Equilibria in Repeated Games*

Jeffrey C. Ely† Johannes Hörner‡ Wojciech Olszewski§

March 28, 2004

Abstract

We introduce a class of strategies which generalizes examples constructed in two-player games under imperfect private monitoring. A sequential equilibrium is *belief-free* if, after every private history, each player's continuation strategy is optimal independently of his belief about his opponents' private histories. We provide a simple and sharp characterization of equilibrium payoffs using those strategies. While such strategies support a large set of payoffs, they are not rich enough to generate a folk theorem in most games besides the prisoner's dilemma, even when noise vanishes.

1 Introduction

(Infinitely) repeated games have widespread application in economics as simple and tractable models of ongoing strategic relationships between agents. The tractability of the repeated game model is due to its recursive nature: the game that begins at date t is identical to the game that was begun at date 0. A powerful set of analytical techniques have been developed to characterize

*The present paper develops some ideas from Ely and Välimäki (2000) and Hörner and Olszewski (2002). We are indebted to Juuso Välimäki for his work on the project in early stages. We are also grateful to Marcin Peski for clarifying the proof of Matsushima (2002), to Ichiro Obara for pointing out an important error in an early draft of Hörner and Olszewski (2002), and to Rakesh Vohra for his suggestions on dynamic programming literature. The paper has been presented at the Canadian Economic Theory Conference in 2003, at ESEM 2003 and in workshops at Harvard-MIT, Northwestern University, University of Chicago, University of Minnesota, University of Essex, ECARES, INSEAD and HEC. We are grateful to those audiences for comments.

†Department of Economics, Boston University. ely@bu.edu. Financial support from NSF grant #9985462 is gratefully acknowledged.

‡Department of Managerial Economics and Decision Sciences, Kellogg Graduate School of Management. j-horner@kellogg.northwestern.edu.

§Department of Economics, Northwestern University. wo@northwestern.edu.

behavior in repeated games. These methods exploit a recursive property of *equilibrium* made possible by the recursive structure of the repeated game.

Abreu, Pearce, and Stachetti (1990), Fudenberg, Levine, and Maskin (1994) and others have applied these techniques to a special class of equilibria, referred to as "public equilibria," in which the behavior of all players is contingent only on information that is publicly known. While this restriction rules out some sequential equilibria, there is an important class of economic environments in which the restriction entails little loss of generality. In particular, if all of the information player i obtains about the behavior of his rivals is also *public* information (these are games with *public monitoring*), then all pure-strategy sequential equilibria are observationally equivalent to public equilibria, see Abreu, Pearce, and Stachetti (1990). Moreover, it is an implication of Fudenberg, Levine, and Maskin (1994) that in games with public monitoring, all (pure or mixed) sequential equilibrium payoffs can be obtained by public equilibria, provided information is sufficiently rich and the players are sufficiently patient (this is the "folk theorem".)

Still, the restriction to games with public monitoring leaves out many potential applications.[1] A well-known example is a repeated oligopoly model in which firms compete in prices and neither these prices nor the firms' sales are public information. This is the "secret price-cut" model of Stigler (1964). For these repeated games with "private monitoring," public strategies accomplish very little, and so to determine the equilibrium possibilities it is necessary to investigate strategies in which continuation play can be contingent on information held privately by the players. The difficulty appears to be that the recursive structure of equilibria is then lost, see Kandori (2002).

Recently however, some advances in the analysis of repeated games with private monitoring have made this obstacle appear less severe than at first glance. In the context of a two-player repeated prisoners' dilemma, Piccione (2002) and Ely and Välimäki (2002) (hereafter VPE) identified a new class of sequential equilibrium strategies that *can* be analyzed using recursive techniques and showed that this class is sufficiently rich to establish a version of the folk theorem for that game. The key characteristic of these strategies is that when they are used, the optimal continuation strategy for a given player i is independent of the prior history of play. This means that the history of the other player is a sufficient statistic for player i's payoffs and can thus be used as a state variable in a dynamic programming formulation of player i's optimization.

In this paper, we look at two-player repeated games with private monitoring and consider strategies with exactly this property. We call the property "belief-free" because it implies that a player's belief about his opponent's history is not needed for computing a best-reply. Thus, the daunting complexity of tracking a player's beliefs over time to ensure that his strategy remains a best-reply is avoided for equilibria involving belief-free strategies.

An important feature of belief-free equilibria is that in each period t there is a subset of

[1] Moreover, the above results leave open the possibility that for a fixed discount factor or for information structures that are not sufficiently rich, some mixed-strategy sequential equilibrium may achieve more than any public equilibrium. In fact this possibility has been demonstrated by Fudenberg and Tirole (1991) (exercise 5.10) and Kandori and Obara (2003). See also Mailath, Matthews, and Sekiguchi (2002) for similar examples in finitely repeated games.

actions $\mathcal{A}_i^t$ for each player i such that regardless of the prior history of play the set of optimal period t actions belongs to $\mathcal{A}_i^t$. We refer to $\mathcal{A}^t = \mathcal{A}_1^t \times \mathcal{A}_2^t$ as the *regime* prevailing at date t. Belief-free equilibria can be classified according to their regime sequences $\{\mathcal{A}^t\}$.

We fully characterize the set of payoffs that are achieved by belief-free equilibria in the presence of a public randomization device. The public randomization device suggests the regime $\mathcal{A}^t$ and the strategies ensure that all actions belonging to $\mathcal{A}_i^t$ are best-replies for player i. We demonstrate that these equilibria can be analyzed using recursive techniques which build on a variation of the concepts of self-generation and factorization due to Abreu, Pearce, and Stachetti (1990). Suppose that p is the distribution of the public randomization device so that $p(\mathcal{A})$ is the probability of regime $\mathcal{A}$. We say that a payoff vector (v_1, v_2) is *strongly* p-generated by a set of vectors $W_1 \times W_2$ if for each possible regime $\mathcal{A}$, and for each player i there is a mixed action profile α_i from the stage game whose support is included in $\mathcal{A}_i$, continuation values can be selected from W_i so that α_i is a best reply for player i and results in total p-expected payoff v_i. Let $B_p(W)$ be the set of all vectors strongly p-generated by W. We say that a set W for which $W \subset B_p(W)$ is strongly self-generating and show that all members of a strongly self-generating set are belief-free equilibrium payoffs using *i.i.d.* public randomization device p. Furthermore, the set of all belief-free equilibrium payoffs of a given game using public randomization p is itself a strongly self-generating set, indeed the largest such set.

Finally, we characterize the structure of this largest strongly self-generating set W^*. In particular we show that iteratively applying the set-valued mapping $B_p(\cdot)$ beginning with the feasible set of payoffs results in a shrinking sequence of sets whose intersection is W^*. Furthermore, we show a version of the bang-bang principle for belief-free equilibrium payoffs. Any strongly self-generating set W of payoffs can be supported by belief-free strategies that are implementable by 2 state automata whose only continuation payoffs are the extreme points of W, such as the strategies used in Ely and Välimäki (2002).

Next we consider two limiting cases: increasing patience and increasing monitoring accuracy. These are also the limits considered by VPE. For increasing patience ($\delta \to 1$), we show that the limiting set of equilibria can be easily characterized by a family of linear programs. This characterization is a version of the methods introduced by Fudenberg and Levine (1994) and adapted by Kandori and Matsushima (1998) to characterize the limiting set of public equilibria.[2] As an example of this method, the maximum payoff for player i can be found by considering an auxiliary contracting problem where player $-i$ chooses a mixed action profile and gives utility penalties to player i as a function of observed signals in order to induce i to play a mixed action that results in as high as possible a net utility for i. We use this characterization to discover the boundaries of the belief-free equilibrium payoff set under the second limit, as the players monitor one another with increasing accuracy. We find a simple formula that can easily be computed by

[2]Fudenberg and Levine (1994) analyzed public equilibria in games with long-run and short-run players, and Kandori and Matsushima (1998) looked at equilibria of games with private monitoring in which all payoff-relevant behavior was conditioned on public announcements made by the players.

linear programming methods and apply it to a series of examples. These examples show that belief-free equilibria can support a large set of payoffs, but the prisoner's dilemma considered by VPE is apparently exceptional: in general this set is not large enough to establish the folk theorem.

All of the characterizations discussed above are derived assuming a fixed *i.i.d.* public randomization device. In fact, we show (Proposition 5) that this is without loss of generality when the discount factor is close to 1: every belief free equilibrium payoff using an arbitrary sequence of public randomizations can be achieved using a fixed *i.i.d.* public randomization for discount factors close enough to 1. Furthermore, in an online appendix Ely, Hörner, and Olzsweski (2004) we show that public randomizations can be dispensed with altogether: for discount factors close enough to 1, any equilibrium payoff obtained using an *i.i.d.* public randomization can also be achieved in a belief-free equilibrium with a deterministic (cyclic) sequence of regimes. This extends the result of Fudenberg and Maskin (1991) to belief-free equilibria of repeated games with private monitoring.

As a final application of our techniques, we consider the special case of games with independent monitoring. These are games in which, conditional on the chosen action profile, the players observe statistically independent signals. Whereas the folk theorem of VPE required signals to be nearly perfectly informative, recently Matsushima (2002) demonstrated the folk theorem for the prisoner's dilemma with conditionally independent but only minimally informative signals. This was accomplished by augmenting the strategies used by Ely and Välimäki (2002) with a review phase. We apply our results to provide a sufficient condition for equilibrium payoffs in arbitrary two action games with independent monitoring. The condition generalizes the result of Matsushima (2002) and our techniques simplify the argument.

A closely related paper to ours is Kandori and Obara (2003). That paper considers repeated games with public monitoring, but analyzes *private strategies*, i.e. strategies in which a player's continuation play depends on private history as well as public history. Analysis of equilibrium involving private strategies shares many of the same complications typical of repeated games with private monitoring, and the approach used by Kandori and Obara (2003) is similar to ours. In particular they consider equilibria in "machine strategies" in which each player is made indifferent across a fixed set of actions regardless of the state of the other's machine. An important difference is that they assume that the set of actions is fixed throughout. We show that allowing for strategies that alternate among "regimes" greatly increases the set of payoffs that can be supported in equilibrium. They obtain a linear inequality characterization of equilibrium which closely resembles one of our characterizations of belief-free equilibria (see Section 4). They also obtain some results for the case of near-perfect monitoring.

The remainder of this paper is organized as follows. In section 2 we introduce the notation used in the paper, present the definition of belief-free strategies, and establish some preliminary results. Section 3 introduces the concept of strong self-generation and uses it to characterize the set of belief-free equilibrium payoffs. The bang-bang result appears here. In section 4 we present our characterization results for discount factors near 1 and near-perfect monitoring structures, and demonstrate their use with some examples. Finally, section 5 takes up the case of independent monitoring and section 6 concludes.

2 Definitions and Preliminary Results

We analyze two-player repeated games with imperfect monitoring. Each player $i = 1, 2$ has a finite action set A_i and a finite set of signals Σ_i. An action profile is an element of $A := A_1 \times A_2$. We use $\triangle W$ to represent the set of probability distributions over a finite set W, and $\mathcal{P}(W)$ to represent the collection of all non-empty subsets of W. The convex hull of W is denoted co(W). If $W \subset \mathbf{R}^2$, then W_1 and W_2 denote the projections of W. Let $\triangle A_i$ and $\triangle A := \triangle A_1 \times \triangle A_2$ represent respectively the set of mixed actions for player i and mixed action profiles.

For each possible action profile $a \in A$, the monitoring distribution $m(\cdot \mid a)$ specifies a joint probability distribution over the set of signal profiles $\Sigma := \Sigma_1 \times \Sigma_2$. When action profile a is played and signal profile σ is realized, player i observes his corresponding signal σ_i. Let $m_i(\cdot \mid a)$ denote the marginal distribution of i's signal. Letting $\tilde{u}_i(a_i, \sigma_i)$ denote the payoff to player i from action a_i and signal σ_i, we can represent stage payoffs as a function of mixed action profiles alone:

$$u_i(\alpha) = \sum_{a \in A} \sum_{\sigma_i \in \Sigma_i} \alpha(a)\, m_i(\sigma_i \mid a)\, \tilde{u}_i(a_i, \sigma_i).$$

Let V denote the convex hull of the set of feasible payoffs. Note that V is a compact subset of $\mathbf{R}^2$. Repeated game payoffs are evaluated using the discounted average criterion. The players share a common discount factor $\delta < 1$.

A t-length (private) history for player i is an element of $H_i^t := (A_i \times \Sigma_i)^t$. A pair of t-length histories (called simply a history) is denoted h^t. Each player's initial history is the null history, denoted $\emptyset$. Let H^t denote the set of all t-length histories, H_i^t the set of i's private t-length histories, $H = \cup_t H^t$ the set of all histories, and $H_i = \cup_t H_i^t$ the set of all private histories for i. A repeated-game (behavior) strategy for player i is a mapping $s_i : H_i \to \triangle A_i$. A strategy profile is denoted s. Let $U_i(s)$ denote the expected discounted average payoff to player i in the repeated game when the players use strategy profile s. For history h_i^t, let $s|_{h_i^t}$ denote the continuation strategy derived from s following history h_i^t. Specifically, if $h_i \hat{h}_i$ denotes the concatenation of the two histories h_i and $\hat{h}_i$, then $s|_{h_i}$ is the strategy defined by $s|_{h_i}(\hat{h}_i) = s(h_i \hat{h}_i)$. Given a strategy profile s, for each t and $h_{-i}^t \in H_{-i}^t$ let $B_i(s|h_{-i}^t)$ denote the set of continuation strategies for i that are best replies to $s_{-i}|_{h_{-i}^t}$.

Definition 1 *A strategy profile s is belief-free if for every h^t, $s_i|_{h_i^t} \in B_i(s|h_{-i}^t)$ for $i = 1, 2$.*

It is immediate that every belief-free strategy profile is a sequential equilibrium. We will therefore speak directly of belief-free equilibria. Observe that public strategies in games with (perfect or imperfect) public monitoring are belief-free (because private histories are trivial).

There always exists a belief-free equilibrium, since any history-independent sequence of static equilibrium action profiles is belief-free.

In the literature on private monitoring, the strategies used by VPE are belief-free. On the other hand, the strategies used by Sekiguchi (1997), Bhaskar and Obara (2002), Mailath and Morris (2002) and Matsushima (2002) are not. Proposition 2 shows that the concept of belief-free equilibrium is actually parallel to a generalization of Ely and Välimäki (2002) equilibrium (strong self-generation). It is, however, conceptually convenient to introduce belief-free profiles first; in particular, it is much easier to see that belief-free profiles are sequential equilibria.

2.1 Regimes

Suppose that s is an equilibrium belief-free strategy profile. It is convenient to describe belief-free equilibria in terms of optimal actions in a given period t. A (continuation) strategy z_i is a *belief-free sequential best-reply* to s_{-i} beginning from period t if

$$z_i|_{h_i^{\tilde{t}}} \in B_i(s|h_{-i}^{\widetilde{t}}) \qquad \text{for all } \widetilde{t} \geq t, \text{ and } h^{\widetilde{t}} \in H^{\widetilde{t}};$$

the set of belief-free sequential best-replies beginning from period t is denoted by $B_i^t(s)$. Let

$$\mathcal{A}_i^t = \{a_i \in A_i : \exists z_i \in B_i^t(s),\ \exists h_i^t \text{ such that } z_i(h_i^t)[a_i] > 0\};$$

note that $\exists h_i^t$ can be replaced with $\forall h_i^t$, because if z_i is a belief-free sequential best-reply to s_{-i} and every continuation strategy $z_i|_{h_i^t}$ gets replaced with the strategy $z_i|_{\widetilde{h}_i^t}$ for a given $\widetilde{h}_i^t$, then so obtained strategy $\widetilde{z}_i$ is also a belief-free sequential best-reply to s_{-i}. We refer to $\mathcal{A}^t = \mathcal{A}_i^t \times \mathcal{A}_{-i}^t$ as the *regime* that prevails at date t. Denote the set of all regimes by $\mathcal{J} := \mathcal{P}(A_1) \times \mathcal{P}(A_2)$. Every belief-free equilibrium gives rise to a sequence of (non-empty) regimes $\{\mathcal{A}^t\}$.

It will be convenient to distinguish belief-free equilibria such that $\mathcal{A}_i^t \supset \mathcal{A}_i$ for each t and $i = 1, 2$; we say that such a belief-free equilibrium is *bounded* by $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$.

2.2 Exchangeability

As we show in this section, belief-free equilibria satisfy an exchangeability property, similar to the exchangeability of Nash equilibria in two-person zero sum games. In particular, given two distinct belief-free equilibria, each governed by the same sequence of regimes, we obtain a new belief-free equilibrium by pairing player 1's strategy in the first equilibrium with player 2's strategy in the second. This property will be used to show that for any given sequence of regimes, the set of belief-free equilibrium payoffs has a product structure.

Proposition 1 *Let $\{\mathcal{A}^t\}$ be a sequence of regimes and let s, $\tilde{s}$ be two belief-free equilibria with regime sequence $\{\mathcal{A}^t\}$. Then the profiles $(s_1, \tilde{s}_2)$ and $(\tilde{s}_1, s_2)$ are also belief-free equilibria with regime sequence $\{\mathcal{A}^t\}$.*

Proof: We will show a stronger result. Any strategy z_2 which adheres to the regime sequence $\{\mathcal{A}^t\}$, (i.e. $z_2(h_2^t) \in \triangle \mathcal{A}_2^t$ for every $h_2^t \in H_2^t$) is a belief-free sequential best-reply to s_1. It suffices to consider pure strategies z_2.

For each $t = 0, 1, \ldots$ there exists a belief-free sequential best-reply to s_1 (beginning from period t) which plays the pure action $z_2(h_2^t)$ with positive probability after history h_2^t. This is because $z_2(h_2^t) \in \mathcal{A}_2^t$ and $\mathcal{A}_2^t$ is the regime governing s in period t. We define a (continuation) strategy $\tilde{z}_2|_{h_2^t}$ to be the strategy which begins by playing $z_2(h_2^t)$ and thereafter reverts to $s_2 \,|_{h_2^t}$. Note that

$$\tilde{z}_2|_{h_2^{\widetilde{t}}} \in B_2(s|h_1^{\widetilde{t}}) \qquad \text{for all } \widetilde{t} \geq t, \text{ and } h^{\widetilde{t}} \in H^{\widetilde{t}}.$$

This is because $s_2|_{h_2^{\widetilde{t}}} \in B_2(s|h_1^{\widetilde{t}})$ and $\tilde{z}_2|_{h_2^t}$ differs from $s_2|_{h_2^t}$ only in the initial period in which it plays one of the actions assigned positive probability by a belief-free sequential best-reply to s_1.

Now we construct a sequence of strategies for player 2, z_2^t for $t = 0, 1, \ldots$. First we set $z_2^0 = z_2|_{\emptyset}$. Next, we inductively define z_2^t by $z_2^t(h_2^\tau) = z_2^{t-1}(h_2^\tau)$ if $\tau < t$ and $z_2^t|_{h_2^t} = \tilde{z}_2|_{h_2^t}$. Observe that

$$z_2^t|_{h_2^t} = \tilde{z}_2|_{h_2^t} \in B_2(s|h_1^t) \qquad \text{for all } t, \text{ and } h^t \in H^t.$$

Now, $z_2^{t+1}|_{h_2^t}$ differs from $z_2^t|_{h_2^t}$ only by replacing its continuation strategies with $\tilde{z}_2|_{h_2^{t+1}}$. Since this cannot reduce the payoff, we have $z_2^{t+1}|_{h_2^t} \in B_2(s|h_1^t)$ and by induction

$$z_2^{t+k}|_{h_2^t} \in B_2(s|h_1^t) \qquad \text{for all } t \geq 0,\ k \geq 0,\ h^t. \tag{1}$$

By construction, for all $k \geq 0$, $z_2^{t+k}(h_2^t) = z_2(h_2^t)$ and thus for any fixed h_2^t, the sequence of continuation strategies $z_2^{t+k}|h_2^t$ converges, as $k \to \infty$ to $z_2|_{h_2^t}$, history-by-history, i.e. in the product topology. Because discounted payoffs are continuous in the product topology, (1) implies

$$z_2|_{h_2^t} \in B_2(s|h_1^t) \qquad \text{for all } t, \text{ and } h^t \in H^t.$$

which is what we set out to show. ■

As a corollary, we have that the set of all belief-free equilibrium payoffs for a given sequence of regimes is a product set.

Corollary 1 *Let $W^*(\{\mathcal{A}^t\})$ be the set of all payoffs arising from belief-free equilibria using regime sequence $\{\mathcal{A}^t\}$. Then $W^*(\{\mathcal{A}^t\}) = W_1 \times W_2$ for some subsets $W_1 \subset \mathbf{R}$ and $W_2 \subset \mathbf{R}$.*

2.3 Public Randomization

Characterizing belief-free equilibrium payoffs is considerably simplified when the players have access to a public randomization device. We will henceforth suppose that in each period, all players observe a public signal y from a set of possible public signal realizations Y.[3] Let $\nu(\hat{y}|y^t) \in \triangle Y$ be the probability in period $t+1$ of realization $\hat{y} \in Y$ conditional on the past sequence of realizations $y^t \in Y^t$. The public randomization device is independent of any other (private) history. Below we will give special attention to public randomizations which are *i.i.d.* draws from the same distribution $p \in \triangle Y$. Abuse notation slightly and write $h^t = (h_1^t, h_2^t, y^t)$ for the private/public history pair after time t.

A strategy now depends on private history as well as the sequence of realizations y^t. The set of histories for i is now $H_i = \cup_t (H_i^t \times Y^t)$ and the set of histories is $H = \cup_t (H_1^t \times H_2^t \times Y^t)$. A strategy is now a mapping $s_i : H_i \times Y \to \triangle A_i$ which specifies the mixed action to play for each private/public history pair. Continuation strategies $s_i|_{(h_i^t, y^t)}$ and best-reply continuation strategies $B_i(s|(h_{-i}^t, y^t))$ are defined as before. The definition of belief-free equilibrium must also be appropriately modified. In particular, the set of best-replies can depend on the regime, but not on the private history.

Definition 2 *In the presence of a public randomization device, a strategy profile s is belief-free if for all t, $y^t \in Y^t$, and $h_i^t \in H_i^t$, $s_i|_{(h_i^t, y^t)} \in B_i(s|(h_{-i}^t, y^t))$ for every $h_{-i}^t \in H_{-i}^t$.*

A (continuation) strategy z_i is a *belief-free sequential best-reply* to s_{-i} beginning from period t if

$$z_i|_{h_i^{\tilde{t}}, y^{\tilde{t}}} \in B_i(s|h_{-i}^{\tilde{t}}, y^{\tilde{t}}) \qquad \text{for all } \tilde{t} \geq t, \text{ and } h^{\tilde{t}} \in H^{\tilde{t}}.$$

In a belief-free equilibrium with a public randomization device, the regime in a given period depends on the current realization $\widehat{y}$:

$$\mathcal{A}_i^t(\hat{y}) = \{a_i \in A_i : \exists z_i \in B_i^t(s),\ \exists\ (h_i^t, y^t) \text{ such that } z_i(h_i^t, y^t, \hat{y})[a_i] > 0\};$$

$\exists (h_i^t, y^t)$ can be replaced with $\forall (h_i^t, y^t)$. A belief-free equilibrium is *direct* if $Y = \mathcal{J}$ and for each t and $\hat{\mathcal{A}}$, we have $\mathcal{A}_i^t(\hat{\mathcal{A}}) = \hat{\mathcal{A}}_i$. That is, in a direct equilibrium, the public randomization device simply suggests a regime, and all actions in the suggested regime are best-responses. By the revelation principle, we can restrict attention to direct belief-free equilibria.

We say that a belief-free equilibrium is *bounded* by $p \in \triangle \mathcal{J}$ if the public randomization is *i.i.d.* according to p and $\mathcal{A}_i^t(\hat{\mathcal{A}}) \supset \hat{\mathcal{A}}_i$ for each t, $\hat{\mathcal{A}}$ and $i = 1, 2$.

3 Strong Self-Generation

In this section we develop a generalization of the Abreu, Pearce, and Stachetti (1990) concept of self-generation to characterize belief-free equilibrium payoffs.

[3] At first glance, it may seem that public randomization and private monitoring do not go together, but it will be clear that public randomizations are used as a substitute for sequences of regimes in a manner similar to the case of perfect monitoring where public randomizations substitute for transitions among mixed action profiles, see Fudenbeg and Maskin (1990) and Fudenberg and Maskin (1991). We show this formally in our online appendix Ely, Hörner, and Olzsweski (2004).

Definition 3 *Let $p \in \triangle \mathcal{J}$ be a public randomization over regimes and $W \subset V$ a set of continuation payoffs. Say that v_i is p-generated by W_i if for each regime $\mathcal{A}$, there exist a mixture $\alpha_{-i}^{\mathcal{A}} \in \triangle \mathcal{A}_{-i}$ and continuation value function $w_i^{\mathcal{A}} : A_{-i} \times \Sigma_{-i} \to \mathrm{co}(W_i)$ such that for each $a_i : \mathcal{J} \to A_i$*

$$v_i \geq \sum_{\mathcal{A} \in \mathcal{J}} p(\mathcal{A}) \left[(1-\delta) u_i(a_i(\mathcal{A}), \alpha_{-i}^{\mathcal{A}}) + \delta \sum_{a_{-i} \in A_{-i}} \sum_{\sigma_{-i} \in \Sigma_{-i}} \alpha_{-i}^{\mathcal{A}}(a_{-i}) m_{-i}(\sigma_{-i} | a_i(\mathcal{A}), a_{-i}) w_i^{\mathcal{A}}(a_{-i}, \sigma_{-i}) \right] \tag{2}$$

with equality if $a_i(\mathcal{A}) \in \mathcal{A}_i$ for each $\mathcal{A}$.

We write $v_i \in B_{p,i}(W_i)$ if v_i is p-generated by W_i, and let $B_p(W) = B_{p,1}(W_1) \times B_{p,2}(W_2)$. We will also say that the pair $\{\alpha_{-i}^{\mathcal{A}}\}_{\mathcal{A} \in \mathcal{J}}$ and $\{w_i^{\mathcal{A}}\}_{\mathcal{A} \in \mathcal{J}}$ together *enforce* the public randomization p and p-generate v_i. The set W is strongly self-generating if $W \subset B_p(W)$ for some public randomization p.

Proposition 2 *If W is strongly self-generating using public randomization $p \in \triangle \mathcal{J}$, then each element of W is the payoff of a belief-free equilibrium bounded by the* i.i.d *public randomization p. Conversely, the set of all payoffs in belief-free equilibria bounded by* i.i.d *public randomization p is itself a strongly self-generating set.*

Proof: Say that a strategy for player i in the repeated game conforms to the public randomization device if for each regime $\mathcal{A}$, in any period when $\mathcal{A}$ is the realization of p, the strategy plays a randomization over $\mathcal{A}_i$. Let $v = (v_i, v_{-i})$ belong to W. We will show that player i has a strategy which conforms to the public randomization such that player $-i$'s maximum payoff is v_{-i}, and this payoff is achieved by any strategy for $-i$ which conforms to the public randomization device. Since the symmetric argument implies the same conclusion with the roles reversed, these strategies form a belief-free profile. Obviously, we have $\hat{\mathcal{A}}_i \subset \mathcal{A}_i^t(\hat{\mathcal{A}})$ for each $i = 1, 2$ and $\hat{\mathcal{A}}$.

Since $W \subset B_p(W)$, for each $u \in W_{-i}$, for each regime $\mathcal{A}$, there is a mixture, call it $\alpha_i^{\mathcal{A},u} \in \triangle \mathcal{A}_i$ and a continuation value function $w_{-i}^{\mathcal{A},u}$ which satisfy (2). Construct a Markovian strategy for player i as follows. The "state" of the strategy will be the continuation value for player $-i$. In any stage in which the state is u and the realization of the public randomization is $\mathcal{A}$, player i will play mixed action $\alpha_i^{\mathcal{A},u}$. Player i will then randomly determine the next state depending on the realization of his own mixture and the observed private signal as follows. The continuation value $w_{-i}^u(a_i, \sigma_i)$ is an element of $\mathrm{co}(W_{-i})$. Thus, there exists a pair of elements $\{\underline{u}, \bar{u}\}$ of W_{-i} such that for some $s \in [0,1]$, $w_{-i}^u(a_i, \sigma_i) = s\underline{u} + (1-s)\bar{u}$. Then, having played action a_i, and observed signal σ_i, player i will transit to state $\underline{u}$ with probability s and to $\bar{u}$ with probability $(1-s)$.

The initial state will be v_{-i}. It now follows from equation (2) and the one-stage deviation property that when the marginal distribution of i's signal is given by m_i, and the public randomization is $i.i.d$ with distribution p, any strategy for player $-i$ which conforms to the public randomization is a best-reply and achieves payoff v_{-i}.

To prove the converse, consider any belief-free equilibrium s bounded by $i.i.d$ public randomization p. Write $w_i(h) = U_i(s|_h)$ for the continuation payoff to player i after history h. Let W_i be the set of all possible continuation values for i in the equilibrium s. Formally $W_i = \{w_i(h) : h \in H\}$; note that W_i consists of payoffs in belief-free equilibria bounded by $i.i.d$ public randomization p.

Let us first observe that in a belief-free equilibrium, $w_i(h^t)$ depends only on (h^t_{-i}, y^t). To see why, suppose $w_i(h^t_i, h^t_{-i}, y^t) > w_i(\tilde{h}^t_i, h^t_{-i}, y^t)$. Then $s_i|_{\tilde{h}_i, y^t}$ is not a best-reply to $s_{-i}|_{h_{-i}, y^t}$ implying s is not belief-free. We can thus write $w_i(h^t) = w_i(h^t_{-i}, y^t)$.

Consider any date $t+1$ and history h^t. Because s conforms to the public randomization, for each $\mathcal{A}$, the mixed action $\alpha^{\mathcal{A}}_{-i} := s_{-i}(h^t_{-i}, y^t, \mathcal{A})$ played by i's opponent in regime $\mathcal{A}$ belongs to $\triangle\mathcal{A}_{-i}$. Let $a_i(\mathcal{A}) \in \mathcal{A}_i$. Since $\mathcal{A}^t_i(\mathcal{A}) \supset \mathcal{A}_i$, there is a belief-free sequential best-reply continuation strategy $\hat{s}_i$ for i which plays a_i after history $(h^t_i, y^t, \mathcal{A})$; after each possible subsequent history (h^{t+1}_i, y^{t+1}), $\hat{s}_i$ is a best-reply at (h^{t+1}_i, y^{t+1}). Because the equilibrium is belief-free, $U_i(\hat{s}_i, s_{-i}|_{h^{t+1}_{-i}, y^{t+1}}) = w_i(h^{t+1}_{-i}, y^{t+1})$ for every (h^{t+1}_{-i}, y^{t+1}).

For fixed h^t_{-i} and y^t, we can view $w_i(h^{t+1}_{-i}, y^{t+1})$ as a continuation value function depending on the realizations of $(a_{-i}, \sigma_{-i}, \mathcal{A})$ and taking values in W_i. The payoff to i from using $\hat{s}_i$ against $s_{-i}(h^t_{-i}, y^t)$ when the realized regime in period $t+1$ is $\mathcal{A}$ can thus be written

$$(1-\delta)u_i(a_i(\mathcal{A}), \alpha^{\mathcal{A}}_{-i}) + \delta \left[\sum_{\sigma_{-i} \in \Sigma_{-i}} \sum_{a_{-i} \in A_{-i}} \alpha^{\mathcal{A}}_{-i}(a_{-i}) m_{-i}(\sigma_{-i} | a_i(\mathcal{A}), a_{-i}) w^{\mathcal{A}}_i(h^t_{-i}, y^t; a_{-i}, \sigma_{-i}) \right] \quad (3)$$

and the expected payoff before the realization of the regime in period $t+1$ is the expected value of this expression with respect to p, i.e. the right-hand side of (2).

Since $\hat{s}_i$ is a best-reply against $s_{-i}|_{h^{\tilde{t}}_{-i}, y^{\tilde{t}}}$, this is equal to $w_i(h^t, y^t)$. Moreover, since we selected i's action from each $\mathcal{A}_i$ arbitrarily, this equality holds for any selection $\{a_i(\mathcal{A})\}_{\mathcal{A} \in \mathcal{J}}$. Finally since s is direct and belief-free, player i cannot achieve a greater continuation value with a strategy that does not conform to the regime. Thus, $w_i(h^t)$ must be greater than or equal to the p-expected value of (3) when $a_i(\mathcal{A}) \notin \mathcal{A}_i$ for some $\mathcal{A}$.

This shows that the pair $(s_{-i}(h^t_{-i}, y^t; \cdot), w_i(h^t_{-i}, y^t; \cdot, \cdot))$ enforce p and generate $w_i(h^t)$. Since h^t was arbitrary, every element of W_i can be so generated. Applying the same argument for player $-i$ shows that the set $W = \prod_{i=1}^{2} W_i$ is strongly self-generating.

Now let W^* be the union of all continuation values occurring along histories of all belief-free equilibria bounded by the public randomization p. The set W^* is the union of strongly self-generating sets and is therefore strongly self-generating. ■

The preceding proposition shows that strong self-generation characterizes all belief-free equilibria bounded by an *i.i.d* public randomization. It is a consequence of our Proposition 5 below that when characterizing the limit set of belief-free equilibria as the discount factor approaches 1, it is without loss of generality to focus on belief-free equilibria bounded by *i.i.d* public randomizations. The following proposition provides an algorithm which can be used for any discount factor to compute the set of all strongly self-generating payoffs for a given *i.i.d* randomization p.

Proposition 3 *For each p, there exists a maximal strongly self-generating set W^*. W^* is the product of closed intervals. Set $W_i^0 = B_{p,i}(V_i)$, and inductively $W_i^\tau = B_{p,i}(W_i^{\tau-1})$. Then $W_i^\tau \subset W_i^{\tau-1}$ for each τ and*

$$W_i^* = \bigcap_{\tau \geq 0} W_i^\tau$$

Proof: The following observation will be used repeatedly in the proof. For any two sets W_i, W_i' such that $W_i \subset W_i'$, if v_i is p-generated by W_i then v_i is also p-generated by W_i'. Thus $B_{p,i}(\cdot)$ is monotonic in the sense that $B_{p,i}(W_i) \subset B_{p,i}(W_i')$ for $W_i \subset W_i'$.

The union of strongly self-generating sets is itself strongly self-generating. This shows the existence of a maximal set W^*. Note that $W^* = B_p(W^*)$. This follows because by monotonicity $W^* \subset B_p(W^*)$ implies $B_p(W^*) \subset B_p(B_p(W^*))$ so that $B_p(W^*)$ is strongly self-generating. By the maximality of W^*, we have $B_p(W^*) \subset W^*$ implying that these sets are equal.

We will first show that $B_{p,i}(W_i)$ is convex. Let v, v' be elements of $B_{p,i}(W_i)$. By the definition of $B_{p,i}(W_i)$, for each $\mathcal{A}$ there are associated mixed actions $\alpha_{-i}^{\mathcal{A}}(v)$, $\alpha_{-i}^{\mathcal{A}}(v')$ and continuation value functions $w_i^{\mathcal{A}}(v)$ and $w_i^{\mathcal{A}}(v')$ used to generate v_i and v_i' respectively. By the linearity of the inequalities in the definition of strong self-generation, any convex combination of v_i and v_i' is generated by the corresponding convex combinations of $\alpha_{-i}^{\mathcal{A}}(v)$ and $\alpha_{-i}^{\mathcal{A}}(v')$ and $w_i^{\mathcal{A}}(v)$ and $w_i^{\mathcal{A}}(v')$.

Next we show that the operator $B_{p,i}(\cdot)$ preserves compactness. Obviously $B_{p,i}(W_i)$ is bounded whenever W_i is. To show that $B_{p,i}(W_i)$ is closed for compact W_i, let v_i^r be a sequence of elements of $B_{p,i}(W_i)$ with $\lim_r v_i^r = v_i$. Then for each regime $\mathcal{A}$, there are corresponding sequences $(\alpha_{-i}^{\mathcal{A}})^r$ and continuation value functions $(w_i^{\mathcal{A}})^r$ taking values in $\mathrm{co}(W_i)$ used to generate v_i^r. By the compactness of $\triangle A_{-i}$ and $\mathrm{co}(W_i)$ there is a subsequence of r's such that $(\alpha_{-i}^{\mathcal{A}})^r \to \alpha_{-i}^{\mathcal{A}}$ and $(w_i^{\mathcal{A}})^r(a_{-i}, \sigma_{-i}) \to w_i^{\mathcal{A}}(a_{-i}, \sigma_{-i}) \in \mathrm{co}(W_i)$ for each of the finitely many pairs (a_{-i}, σ_{-i}). By continuity, $\alpha_{-i}^{\mathcal{A}}$ and $w_i^{\mathcal{A}}$ generate v_i, and therefore $v_i \in B_{p,i}(W_i)$.

Obviously $W_i^0 \subset V_i$. Using monotonicity we conclude that $W_i^1 \subset W_i^0$ and inductively that $W_i^\tau \subset W_i^{\tau-1}$. Since V_i is compact and $B_{p,i}(\cdot)$ preserves compactness, W_i^0 and by induction each W_i^τ are compact. Moreover, for each $\tau \geq 1$, W_i^τ is convex. Thus W_i^τ is a nested sequence of compact intervals and the set

$$W_i^\infty := \bigcap_{\tau \geq 0} W_i^\tau$$

is a closed interval.

We now show that $W^\infty := W_1^\infty \times W_2^\infty$ is strongly self-generating. Let $v_i \in W_i^\infty$. (If $W_i^\infty = \emptyset$, then $W^\infty = \emptyset$ and W^∞ is trivially strongly self-generating.) Then for every τ there exist $(\alpha_{-i}^{\mathcal{A}})^\tau$ and $(w_i^{\mathcal{A}})^\tau : A_{-i} \times \Sigma_{-i} \to \text{co}(W_i^\tau)$ which generate v_i. We claim that there exist limit continuation values $w_i^{\mathcal{A}}(a_{-i}, \sigma_{-i}) \in W_i^\infty$ such that (passing to a subsequence if necessary) $(w_i^{\mathcal{A}})^\tau(a_{-i}, \sigma_{-i}) \to w_i^{\mathcal{A}}(a_{-i}, \sigma_{-i})$ for each pair (a_{-i}, σ_{-i}). If not, then there exist a pair (a_{-i}, σ_{-i}) and an open neighborhood U of W_i^∞ such that $(w_i^{\mathcal{A}})^\tau(a_{-i}, \sigma_{-i}) \notin U$ for infinitely many τ. This implies that $\neg U \cap W_i^\tau \neq \emptyset$ for these τ. But $\neg U \cap W_i^\tau$ is an infinite family of nested, non-empty, compact sets and hence,

$$\emptyset \neq \bigcap_{\tau \geq 0} [\neg U \cap W_i^\tau] = \neg U \cap \bigcap_{\tau \geq 0} W_i^\tau = \neg U \cap W_i^\infty$$

which is a contradiction.

Again, by continuity v_i is generated by $\lim (\alpha_{-i}^{\mathcal{A}})^\tau$ (passing to a further subsequence if necessary) and $w_i^{\mathcal{A}}$. Thus $v_i \in B_p(W_i^\infty)$ and we have shown that W^∞ is strongly self-generating. It remains only to show that W^∞ includes W^*, which by the maximality of W^* would imply that $W^\infty = W^*$ and conclude the proof. But $W^* \subset V^*$, so we have $W^* = B_p(W^*) \subset W^0$ and by induction $W^* \subset W^\tau$ for every τ. Thus $W^* \subset W^\infty$. ■

3.1 Bang-Bang

In this section we show that belief-free equilibrium payoffs can always be obtained by simple strategies that can be represented by a two-state automaton. As a corollary we obtain a version of the traditional "bang-bang" result that a set is strongly self-generating if and only if the set generates its extreme points.

A *machine strategy* for player i is defined as follows. There is a set of states Θ, and for each state $\theta \in \Theta$ there is a behavior rule $\alpha^\theta : \mathcal{P}(A) \to \triangle A_i$ and a transition rule $\phi^\theta : \mathcal{P}(A) \times A_i \times \Sigma_i \to \triangle\Theta$. The interpretation is as follows. When in state θ, if the outcome of the public randomization is regime $\mathcal{A}$, the strategy plays $\alpha^\theta(\mathcal{A})$. Then, if the action a_i is realized and the signal σ_i is observed, a new state is drawn from the distribution $\phi^\theta(\mathcal{A}, a_i, \sigma_i)$ and the strategy begins the next period in that state.

Imagine that i were playing such a machine strategy and player $-i$ was informed each period of the state. Then we could compute the value $v(\theta)$ to player $-i$ of being in state θ. Furthermore, the transition rule would imply for each state the continuation value function $w_{-i}^{\mathcal{A}}$ as follows

$$w_{-i}^{\mathcal{A}}(a_i, \sigma_i) = \sum_{\theta' \in \Theta} v(\theta') \cdot \phi^\theta(\mathcal{A}, a_i, \sigma_i)[\theta']$$

Note that in the proof of Proposition 2 we construct such a machine strategy whose continuation value functions replicate the continuation value functions used in the definition of strong self-generation. In fact as we now show it is always possible to implement the payoff of a belief-free equilibrium bounded by p using machine strategies which have only two states.[4] The values of the two states in i's machine correspond to the maximum and minimum payoffs of player i in belief-free equilibria bounded by p.

Proposition 4 *Consider intervals $V_i = [\underline{v}_i, \bar{v}_i]$ for $i = 1, 2$. Let p be a public randomization over regimes. Suppose for each i that $\underline{v}_i$ is p-generated by V_i using $\underline{\alpha}^{\mathcal{A}}_{-i}$ and $\underline{w}^{\mathcal{A}}_i$ and similarly for $\bar{v}^{\mathcal{A}}_i$ using $\bar{\alpha}^{\mathcal{A}}_{-i}$ and $\bar{w}^{\mathcal{A}}_i$. Then every element v of $V_1 \times V_2$ is the payoff to a belief-free equilibrium in which each i plays a machine strategy with two states $\{\underline{\theta}, \bar{\theta}\}$ whose behavior rules are $\underline{\alpha}^{\mathcal{A}}_i$, $\bar{\alpha}^{\mathcal{A}}_i$ respectively and whose derived continuation value functions are $\underline{w}^{\mathcal{A}}_{-i}$, $\bar{w}^{\mathcal{A}}_{-i}$, respectively.*

Proof: Adapt the Markovian strategy used in the proof of Proposition 2 as follows. There are only two states, corresponding to values $\bar{v}_{-i}$ and $\underline{v}_{-i}$. Now, when then continuation value should be $w \in [\underline{v}_{-i}, \bar{v}_{-i}]$, player i will randomly transit to states $\underline{v}_{-i}$ and $\bar{v}_{-i}$ with probabilities q and $1 - q$ where $w = q\underline{v}_{-i} + (1 - q)\bar{v}_{-i}$. Finally, player i determines his initial state by similarly randomizing over the two states with probabilities calculated to provide initial value v_{-i}. ■

The equilibria provided by proposition 4 satisfy a stronger property than belief-free alone. To see this suppose that $\bar{\alpha}^{\mathcal{A}}_i$ and $\underline{\alpha}^{\mathcal{A}}_i$ are pure actions for each regime. In the initial period i is determining his state, and hence his initial action by randomizing. Regardless of the realization of this randomization, player $-i$ is indifferent over his actions in the regime. Thus, the equilibrium strategy of player $-i$ remains a best-reply even if before play $-i$ could observe the outcome of i's private randomization.

4 Characterizing Belief-Free Equilibrium Payoffs for δ close to 1

4.1 General Information Structure

The preceding results can be used to determine the set of belief-free equilibrium payoffs for any given discount factor δ. Given a fixed distribution over regimes for which the set of belief-free equilibrium payoffs has nonempty interior, it is possible to give a simple characterization of its boundary as δ approaches 1. The key insight is that, when the discount factor is large, some constraints on continuation values can be ignored, as small variations in the continuation values are sufficient to generate appropriate incentives.

To get a feel for the techniques presented here, consider a fixed regime $\mathcal{A}$, and for the moment a given discount factor δ. Let $W^*(\delta)$ be the set of belief-free equilibrium payoffs given δ, bounded

[4] In a related context, Kandori and Obara (2003) also examine conditions under which two-state automata are sufficient.

by regime $\mathcal{A}$. From previous results, we know that $W^*(\delta)$ is the product of closed intervals, so $W^*(\delta) = W_1^*(\delta) \times W_2^*(\delta)$. Let us write $W_1^*(\delta) = [\underline{w}_1, \bar{w}_1]$.

Because $W^*(\delta)$ is strongly self-generating (by Proposition 2), both $\bar{w}_1$ and $\underline{w}_1$ are generated by $W_1^*(\delta)$. Furthermore, $\bar{w}_1$ is the maximum value w_1 generated by the interval $[\underline{w}_1, w_1]$, for if $w_1 > \bar{w}_1$ were generated by $[\underline{w}_1, w_1]$, then by Proposition 4, the set $[\underline{w}_1, w_1] \times W_2^*(\delta)$ would be strongly self-generating, contradicting the definition of $W^*(\delta)$. Thus, $\bar{w}_1$ solves the following optimization problem

$$\bar{w}_1 = \max w_1$$
$$\text{such that for some } \alpha_2 \in \triangle\mathcal{A}_2 \text{ and } z_1\colon A_2 \times \Sigma_2 \to \mathbf{R}$$
$$w_1 \geq (1-\delta)u_1(a_1, \alpha_2) + \delta \left[\sum_{\sigma_2 \in \Sigma_2} \sum_{a_2 \in A_2} \alpha_2(a_2) m_2(\sigma_2 | a_1, a_2) z_1(a_2, \sigma_2) \right]$$
$$\text{for each } a_1 \in A_1, \text{ with equality for each } a_1 \in \mathcal{A}_1.$$
$$w_1(a_2, \sigma_2) \leq w_1 \quad \forall\ a_2, \sigma_2$$
$$w_1(a_2, \sigma_2) \geq \underline{w}_1 \quad \forall\ a_2, \sigma_2.$$

Similarly, $\underline{w}_1$ is the minimum value w_1 generated by the interval $[w_1, \bar{w}_1]$, and is thus the solution to the corresponding minimization problem. Conversely, the solutions to these problems characterize the boundaries of $W_1^*(\delta)$ whenever it is non-empty.

Observe that for δ close to 1, the differences in continuation values required to satisfy the incentive constraints in the above problem can be made arbitrarily small. Thus, when $\bar{w}_1 > \underline{w}_1$, the last constraint can always be satisfied when δ is close enough to 1. As a further simplification, let us define $x_1(a_2, \sigma_2) = \frac{\delta}{1-\delta}[z_1(a_2, \sigma_2) - w_1] \leq 0$, substitute out for $z_1(\cdot, \cdot)$, and rewrite the maximization as follows.

$$\bar{w}_1 = \max w_1$$
$$\text{such that for some } \alpha_2 \in \triangle\mathcal{A}_2 \text{ and } x_1 : A_2 \times \Sigma_2 \to \mathbf{R}_-$$
$$w_1 \geq u_1(a_1, \alpha_2) + \sum_{a_2 \in A_2} \sum_{\sigma_2 \in \Sigma_2} \alpha_2(a_2)\, m_2(\sigma_2 \mid a_1, a_2)\, x_1(a_2, \sigma_2),$$
$$\text{with equality if } a_1 \in \mathcal{A}_1.$$

We can interpret this characterization as follows. Player 2 will select a mixed action $\alpha_2 \in \triangle\mathcal{A}_2$ and levy fines $-x_1(a_2, \sigma_2)$ on player 1 depending on the realized action and signal. The objective is to induce player 1 to select any action in $\mathcal{A}_1$ and to provide the maximum total utility to player 1 in the process. That maximum will turn out to be the maximum payoff of player 1 in a belief-free equilibrium using a sequence of regimes bounded by $\mathcal{A}$ when δ is close to 1. To find the minimum, we analyze the corresponding minimization problem with the difference that player 2 will offer bonuses rather than fines.

In what follows, we formalize and extend this analysis to the case of an arbitrary public randomization over regimes and use it to characterize the set of all belief-free equilibrium payoffs for δ close to 1.

For given $\mathcal{A} = \mathcal{A}_i \times \mathcal{A}_{-i}$, define $M_i^{\mathcal{A}}$ as follows:

$$
\begin{aligned}
M_i^{\mathcal{A}} = \sup v_i \qquad & \text{such that for some} \\
& \alpha_{-i} \in \triangle \mathcal{A}_{-i} \\
& x_i : \mathcal{A}_{-i} \times \Sigma_{-i} \rightarrow \mathbf{R}_- \\
v_i \geq u_i (a_i, \alpha_{-i}) + & \sum_{a_{-i} \in A_{-i}} \sum_{\sigma_{-i} \in \Sigma_{-i}} \alpha_{-i} (a_{-i}) \, m_{-i} (\sigma_{-i} \mid a_i, a_{-i}) \, x_i (a_{-i}, \sigma_{-i}), \\
& \text{for all } a_i \text{ with equality if } a_i \in \mathcal{A}_i.
\end{aligned} \tag{4}
$$

Similarly, define $m_i^{\mathcal{A}}$ as follows:

$$
\begin{aligned}
m_i^{\mathcal{A}} = \inf v_i \qquad & \text{such that for some} \\
& \alpha_{-i} \in \triangle \mathcal{A}_{-i} \\
& x_i : \mathcal{A}_{-i} \times \Sigma_{-i} \rightarrow \mathbf{R}_+ \\
v_i \geq u_i (a_i, \alpha_{-i}) + & \sum_{a_{-i} \in A_{-i}} \sum_{\sigma_{-i} \in \Sigma_{-i}} \alpha_{-i} (a_{-i}) \, m_{-i} (\sigma_{-i} \mid a_i, a_{-i}) \, x_i (a_{-i}, \sigma_{-i}), \\
& \text{for all } a_i \text{ with equality if } a_i \in \mathcal{A}_i.
\end{aligned} \tag{5}
$$

Here and in what follows, $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$. The solutions to these linear programs will be used to provide tight bounds on the sets of belief-free equilibrium payoffs.[5] The set of payoffs inside the bounds will be shown to be strongly self-generating and therefore belief-free. On the other hand, it will be shown that any belief-free equilibrium value will be in the set. The following preliminary result is useful in the sequel.

Lemma 1 *Every $v_i < M_i^{\mathcal{A}}$ is feasible for (4) and every $v_i > m_i^{\mathcal{A}}$ is feasible for (5).*

Let us order the regimes as $\mathcal{J} = \{1, \ldots, J\}$. For $i = 1, 2$, let $M_i = \left(M_i^1, \ldots, M_i^J\right)$, and $m_i = \left(m_i^1, \ldots, m_i^J\right)$. Let $\triangle M_i^{\mathcal{A}} = M_i^{\mathcal{A}} - m_i^{\mathcal{A}}$ and $\triangle M_i = M_i - m_i$. If the same regime $\mathcal{A}$ is used in every period, we can interpret $M_i^{\mathcal{A}}$ as the largest value for player i that can be enforced by player $-i$ through an appropriate choice of an immediate action and a future punishment, and $m_i^{\mathcal{A}}$ as the smallest value for player i that can be enforced by player $-i$ through an appropriate choice of an immediate action and a future reward. Clearly, it is then necessary that the difference $\triangle M_i^{\mathcal{A}}$ be nonnegative for both players. Similarly, if instead of a fixed regime, a distribution p

[5] As formulated, these programs are not linear in (α_{-i}, x_{-i}); they are however linear in (α_{-i}, y_{-i}), where $y_{-i} = \alpha_{-i} x_{-i} : \mathcal{A}_{-i} \times \Sigma_{-i} \rightarrow \mathbf{R}_-$ or $\mathbf{R}_+$.

over regimes is determined by the randomization device, then the weighted average $\triangle M_i p$ must be nonnegative, and intuition suggests that the corresponding equilibrium payoff set should be $\prod_{i=1,2}[m_i p, M_i p]$. However, the case where $\triangle M_i p$ equals 0 for some player i requires special attention. We therefore distinguish the following three cases.

1. (The positive case) There exists $p \geq 0$ such that $\triangle M_i p > 0$, $i = 1, 2$.

2. (The negative case) There exists no $p \geq 0$ such that $\triangle M_i p \geq 0$, $i = 1, 2$, with at least one strict inequality.

3. (The abnormal case) There exists no $p \geq 0$ such that $\triangle M_i p > 0$, $i = 1, 2$, but there exists $p \geq 0$ such that $\triangle M_i p \geq 0$, $i = 1, 2$, with one strict inequality.

Observe that which case obtains depends both on the stage game payoffs and on the monitoring structure. If there exists a regime $\mathcal{A}$ such that $\triangle M_i^{\mathcal{A}} > 0$ for both $i = 1, 2$, then the game is positive (consider the distribution p that assigns probability 1 to regime $\mathcal{A}$), but this is not a necessary condition. Necessary and sufficient conditions are given at the end of this subsection, using the theorem of the alternative. For a given stage game, let V be the limit of the set of (belief-free) equilibrium payoffs when $\delta \to 1$.

We first concentrate on the positive and negative case. The following proposition confirms the intuition sketched above.

Proposition 5 *V is a convex polytope. In*

1. *the positive case,*

$$V = \cup_{\{p \geq 0: \triangle M_i p \geq 0, i=1,2, p\mathbf{1}=1\}} \prod_{i=1,2} [m_i p, M_i p]\,; \tag{6}$$

2. *the negative case, V is the convex hull of the Nash equilibria of the bimatrix game.*

Proof: Consider the positive case. We first show that the right-hand side of (6) is included in V. We pick for each $\mathcal{A}$ payoffs $\bar{v}_1^{\mathcal{A}}, \bar{v}_2^{\mathcal{A}}, \underline{v}_1^{\mathcal{A}}, \underline{v}_2^{\mathcal{A}}$ and a public randomization p over regimes such that

$$m_i^{\mathcal{A}} < \underline{v}_i^{\mathcal{A}} \qquad M_i^{\mathcal{A}} > \bar{v}_i^{\mathcal{A}}$$

and

$$p \cdot (\bar{v}_i - \underline{v}_i) > 0 \qquad i = 1, 2 \tag{7}$$

where e.g. $\underline{v}_i = (\underline{v}_i^{\mathcal{A}})_{\mathcal{A} \in \mathcal{J}}$. Existence is guaranteed by the positive case.

Define $\bar{z}_i = p \cdot \bar{v}_i$, $\underline{z}_i = p \cdot \underline{v}_i$. We will show that there exists $\bar{\delta} < 1$ such that for all $\delta \in (\bar{\delta}, 1)$, the set U defined by

$$U = \text{co} \prod_{i=1,2} \{\underline{z}_i, \bar{z}_i\}$$

is strongly self-generating. Because the right-hand side of (6) is the closure of the union of all such sets U, this will prove its inclusion in V.

By Proposition 4, it is enough to show that each of the extreme values is generated by the convex hull. Consider $\bar{v}_i^{\mathcal{A}}$. Since $M_i^{\mathcal{A}} > \bar{v}_i^{\mathcal{A}}$, by Lemma 1, $\bar{v}_i^{\mathcal{A}}$ is a feasible value for (4) and hence there exists $\bar{\alpha}_{-i}^{\mathcal{A}} \in \triangle\mathcal{A}_{-i}$ and $\bar{x}_i^{\mathcal{A}} : A_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$ such that for every $a_i(\mathcal{A}) \in A_i$

$$\bar{v}_i^{\mathcal{A}} \geq u_i\left(a_i(\mathcal{A}), \bar{\alpha}_{-i}^{\mathcal{A}}\right) + \sum_{a_{-i}\in A_{-i}} \sum_{\sigma_{-i}\in\Sigma_{-i}} \bar{\alpha}_{-i}^{\mathcal{A}}(a_{-i}) m_{-i}\left(\sigma_{-i} \mid a_i(\mathcal{A}), a_{-i}\right) \bar{x}_i^{\mathcal{A}}\left(a_{-i}, \sigma_{-i}\right), \tag{8}$$

$$\text{with equality if } a_i(\mathcal{A}) \in \mathcal{A}_i. \tag{9}$$

For each $\mathcal{A}$ we multiply the above inequality by $p(\mathcal{A})$ and then sum across regimes $\mathcal{A}$ to find that for all "strategies" $\{a_i(\mathcal{A})\}_{\mathcal{A}\in\mathcal{J}}$,

$$\bar{z}_i \geq \sum_{\mathcal{A}\in\mathcal{J}} p(\mathcal{A}) \left[u_i\left(a_i(\mathcal{A}), \bar{\alpha}_{-i}^{\mathcal{A}}\right) + \sum_{a_{-i}\in A_{-i}} \sum_{\sigma_{-i}\in\Sigma_{-i}} \bar{\alpha}_{-i}^{\mathcal{A}}(a_{-i}) m_{-i}\left(\sigma_{-i} \mid a_i(\mathcal{A}), a_{-i}\right) \bar{x}_i^{\mathcal{A}}\left(a_{-i}, \sigma_{-i}\right) \right], \tag{10}$$

$$\text{with equality if } a_i(\mathcal{A}) \in \mathcal{A}_i \text{ for each } \mathcal{A}.$$

Define

$$\bar{z}_i^{\mathcal{A}}(a_{-i}, \sigma_{-i}) = \bar{z}_i + \frac{1-\delta}{\delta} \bar{x}_i^{\mathcal{A}}(a_{-i}, \sigma_{-i}) \tag{11}$$

substitute into (10), and re-arrange to obtain

$$\bar{z}_i \geq \sum_{\mathcal{A}\in\mathcal{J}} p(\mathcal{A}) \left[(1-\delta) u_i\left(a_i(\mathcal{A}), \bar{\alpha}_{-i}^{\mathcal{A}}\right) + \delta \sum_{a_{-i}\in A_{-i}} \sum_{\sigma_{-i}\in\Sigma_{-i}} \bar{\alpha}_{-i}^{\mathcal{A}}(a_{-i}) m_{-i}\left(\sigma_{-i} | a_i(\mathcal{A}), a_{-i}\right) \bar{z}_i^{\mathcal{A}}(a_{-i}, \sigma_{-i}) \right],$$

$$\text{with equality if } a_i(\mathcal{A}) \in \mathcal{A}_i \text{ for each } \mathcal{A}.$$

Because $\bar{x}_i^{\mathcal{A}}(\cdot,\cdot) \leq 0$, it follows from (11) and (7) that for all δ exceeding some $\bar{\delta} < 1$, $\bar{z}_i^{\mathcal{A}}(\cdot,\cdot)$ belongs to U_i. We have therefore shown that the extreme point $\bar{z} = (\bar{z}_1, \bar{z}_2)$ is generated by U. The symmetric derivation shows that $\underline{z}$ is generated by U. Let $\underline{\alpha}_i^{\mathcal{A}}$ and $\underline{z}_i^{\mathcal{A}}$ be the corresponding strategies and continuation value functions. Now it is easily verified that e.g. the value $(\underline{z}_i, \bar{z}_{-i})$ is generated using $\underline{\alpha}_i^{\mathcal{A}}$, $\underline{z}_i^{\mathcal{A}}$ and $\bar{\alpha}_{-i}^{\mathcal{A}}$, $\bar{z}_{-i}^{\mathcal{A}}$.

Next we show that the right-hand side of (6) includes V. Observe that $(1-\delta) M_i^{\mathcal{A}} + \delta\bar{V}$

solves

$$\max_{\substack{\alpha_{-i}\in\mathcal{A}_{-i}\\ V_i:\mathcal{A}_{-i}\times\Sigma_{-i}\to\mathbf{R}}} (1-\delta)\,u_i\left(a_i,\alpha_{-i}\right)+\delta\sum_{a_{-i}\in\mathcal{A}_{-i}}\sum_{\sigma_{-i}\in\Sigma_{-i}}\alpha_{-i}\left(a_{-i}\right)m_{-i}\left(\sigma_{-i}\mid a_i,a_{-i}\right)V_i\left(a_{-i},\sigma_{-i}\right) \quad (12)$$

subject to, for all $a_i\in\mathcal{A}_i$, $a_i'\in A_i$ (with equality if $a_i'\in\mathcal{A}_i$),

$$\begin{aligned}(1-\delta)u_i\left(a_i,\alpha_{-i}\right)+\delta&\sum_{a_{-i}\in\mathcal{A}_{-i}}\sum_{\sigma_{-i}\in\Sigma_{-i}}\alpha_{-i}\left(a_{-i}\right)m_{-i}\left(\sigma_{-i}\mid a_i,a_{-i}\right)V_i\left(a_{-i},\sigma_{-i}\right)\\ &\geq(1-\delta)\,u_i\left(a_i',\alpha_{-i}\right)+\delta\sum_{a_{-i}\in\mathcal{A}_{-i}}\sum_{\sigma_{-i}\in\Sigma_{-i}}\alpha_{-i}\left(a_{-i}\right)m_{-i}\left(\sigma_{-i}\mid a_i',a_{-i}\right)V_i\left(a_{-i},\sigma_{-i}\right)\end{aligned} \quad (13)$$

$$V_i\left(a_{-i},\sigma_{-i}\right)\leq\bar{V}\text{ for all }\left(a_{-i},\sigma_{-i}\right)\in A_{-i}\times\Sigma_{-i}$$

Similarly, $(1-\delta)\,m_i^{\mathcal{A}}+\delta\underline{V}$ solves

$$\min_{\substack{\alpha_{-i}\in\mathcal{A}_{-i}\\ V_i:\mathcal{A}_{-i}\times\Sigma_{-i}\to\mathbf{R}}} (1-\delta)\,u_i\left(a_i,\alpha_{-i}\right)+\delta\sum_{a_{-i}\in\mathcal{A}_{-i}}\sum_{\sigma_{-i}\in\Sigma_{-i}}\alpha_{-i}\left(a_{-i}\right)m_{-i}\left(\sigma_{-i}\mid a_i,a_{-i}\right)V_i\left(a_{-i},\sigma_{-i}\right) \quad (14)$$

subject to, for all $a_i\in\mathcal{A}_i$, $a_i'\in A_i$ (with equality if $a_i'\in\mathcal{A}_i$),

$$\begin{aligned}(1-\delta)u_i\left(a_i,\alpha_{-i}\right)+\delta&\sum_{a_{-i}\in\mathcal{A}_{-i}}\sum_{\sigma_{-i}\in\Sigma_{-i}}\alpha_{-i}\left(a_{-i}\right)m_{-i}\left(\sigma_{-i}\mid a_i,a_{-i}\right)V_i\left(a_{-i},\sigma_{-i}\right)\\ &\geq(1-\delta)\,u_i\left(a_i',\alpha_{-i}\right)+\delta\sum_{a_{-i}\in\mathcal{A}_{-i}}\sum_{\sigma_{-i}\in\Sigma_{-i}}\alpha_{-i}\left(a_{-i}\right)m_{-i}\left(\sigma_{-i}\mid a_i',a_{-i}\right)V_i\left(a_{-i},\sigma_{-i}\right)\end{aligned} \quad (15)$$

$$V_i\left(a_{-i},\sigma_{-i}\right)\geq\underline{V}\quad\text{for all }\left(a_{-i},\sigma_{-i}\right)\in A_{-i}\times\Sigma_{-i}$$

This can be interpreted as follows: Suppose that regime $\mathcal{A}$ is played in period t, and $\bar{V}$ and $\underline{V}$ are the largest and smallest, respectively, continuation payoffs of player i following that regime. Then the payoff of player i in terms of period t cannot exceed $(1-\delta)\,M_i^{\mathcal{A}}+\delta\bar{V}$ and cannot fall below $(1-\delta)\,m_i^{\mathcal{A}}+\delta\underline{V}$ in any equilibrium such that regime $\mathcal{A}$ is played in period t.

Given a public randomization device, denote by $\nu_{t'}\left(\mathcal{A};y_t\right)$ the probability of regime $\mathcal{A}$ occurring in period $t'\geq t$ conditional on the public history y_t. A necessary condition for the set of belief-free equilibria using that randomization device to be non-empty is that, for $i=1,2$, and for any t, any public history y_t which has positive probability under the public randomization, and any regime $\mathcal{A}$ such that $\nu_t\left(\mathcal{A};y_t\right)>0$,

$$M_i^{\mathcal{A}}+\sum_{t'=t+1}^{\infty}\sum_{\mathcal{A}'}\delta^{t'-t}\nu_{t'}\left(\mathcal{A}';y_t\right)M_i^{\mathcal{A}'}\geq m_i^{\mathcal{A}}+\sum_{t'=t+1}^{\infty}\sum_{\mathcal{A}'}\delta^{t'-t}\nu_{t'}\left(\mathcal{A}';y_t\right)m_i^{\mathcal{A}'}.$$

This follows from the repeated application of (12) and (14). For the initial period (before the realization of the randomization device), this set of constraints becomes

$$(1-\delta)\sum_{t'=1}^{\infty}\sum_{\mathcal{A}}\delta^{t'-1}\nu_{t'}\left(\mathcal{A};y_0\right)\left(M_i^{\mathcal{A}}-m_i^{\mathcal{A}}\right)\geq 0. \quad (16)$$

Note finally that the payoff of player i belongs to interval

$$\left[(1-\delta)\sum_{t'=1}^{\infty}\sum_{\mathcal{A}}\delta^{t'-1}\nu_{t'}(\mathcal{A};y_0)\,m_i^{\mathcal{A}},(1-\delta)\sum_{t'=1}^{\infty}\sum_{\mathcal{A}}\delta^{t'-1}\nu_{t'}(\mathcal{A};y_0)\,M_i^{\mathcal{A}}\right].$$

Define a probability distribution over regimes p by

$$p(\mathcal{A}):=(1-\delta)\sum_{t'=1}^{\infty}\delta^{t'-1}\nu_{t'}(\mathcal{A};y_0);$$

then, for $i=1,2$,

$$\sum_{\mathcal{A}}p(\mathcal{A})\left(M_i^{\mathcal{A}}-m_i^{\mathcal{A}}\right)\geq 0, \tag{17}$$

and the equilibrium payoff of player i is an element of $\left[\sum_{\mathcal{A}}p(\mathcal{A})\,m_i^{\mathcal{A}},\sum_{\mathcal{A}}p(\mathcal{A})\,M_i^{\mathcal{A}}\right]$.[6]

Consider now the negative case. Then there exists no p such that $\sum_{\mathcal{A}}p(\mathcal{A})\left(M_i^{\mathcal{A}}-m_i^{\mathcal{A}}\right)\geq 0$, $i=1,2$, and at least one inequality holds strictly. By arguments similar to the positive case, there is a belief-free equilibrium only for p satisfying the two inequalities. Since the two inequalities must be then equalities, the continuation payoffs $V_i(a_{-i},\sigma_{-i})$ is independent of a_{-i} and σ_{-i}. Thus, each player's action in period t is a best reply to his opponent's (mixed) action, and so the set of belief-free equilibrium payoffs is then the convex hull of the set of Nash equilibria in the static game. ■

It is worth emphasizing two consequences of Proposition 5. First, there are games such that, for two regimes $\mathcal{A}_1$ and $\mathcal{A}_2$, there exists $p\in(0,1)$ such that $p\triangle M_i^{\mathcal{A}_1}+(1-p)\triangle M_i^{\mathcal{A}_2}>0$ for both $i=1,2$, yet it is neither the case that $\triangle M_i^{\mathcal{A}_1}>0$ for both i, nor that $\triangle M_i^{\mathcal{A}_2}>0$ for both i. That is, there is no belief-free equilibrium payoff corresponding to the constant sequence $\mathcal{A}_1$, nor to the constant sequence $\mathcal{A}_2$. Yet there are belief-free equilibrium payoffs for some distribution over those two regimes. Therefore, the set of belief-free equilibrium payoffs may be larger than the convex hull of the belief-free equilibrium payoffs that use a constant sequence of regimes. Among

[6] The second part of the proof resembles, and actually has been inspired by, a theorem from dynamic programming (See Theorem 3.1 of Altman (1999)). Altman studies an infinite-horizon constrained optimization problem. The agent chooses a *policy*, i.e. an action for each period, possibly mixed, from a finite set of possible actions. Altman allows for multiple states of the world and for the realization of state of the world depending on previous actions. Each policy generates an *occupation measure*, defined over the set of actions by

$$P(a)=(1-\delta)\Sigma_{t=1}^{\infty}\delta^{t-1}P_t(a),$$

where $P_t(a)$ denotes the probability that action a is taken in period t given the policy.

Altman's Theorem 3.1 asserts that, under some conditions on the kind of constraints, the set of *stationary policies*, i.e. policies where the action taken in each period is independent of the actions taken in the past, is *complete*, i.e. generates the same set of occupation measures as the set of all policies.

the examples presented below, the prisoner's dilemma and the one-sided prisoner's dilemma are games for which this inclusion is strict. Second, Proposition 5 does not assume, but implies, that attention can be restricted to public randomizations which are *i.i.d* draws from the same distribution.

Proposition 5 does not cover the abnormal case. The abnormal case is non-generic, in a sense described below. As a preliminary, the following lemma is needed.

Lemma 2 *Suppose that the stage game has the property that*

$$u_1(a, \alpha_2) = u_1(a, \alpha_2') \tag{18}$$

whenever a *is a best reply to both* α_2 *and* α_2'. *Then player 1 has a dominant action.*

Proof: Consider the family of sets $A_2' \subset A_2$ with the property that some action $a_1 \in A_1$ is a best reply to every action $a_2 \in A_2'$. Pick a maximal, with respect to inclusion, element of this family. If $A_2' = A_2$, then there exists $a_1 \in A_1$ which is a best reply to every $a_2 \in A_2'$, which is therefore a dominant action.

Suppose instead that $A_2' \neq A_2$. Pick $a_2' \in A_2 - A_2'$. Then there exists $a_1 \in A_1$ which is a best reply to every $a_2 \in A_2'$ but which is not a best reply to a_2'. Let α_2 be the mixed action that assigns probability $(1-\lambda)/|A_2'|$ to every action $a_2 \in A_2'$, where $|A_2'|$ stands for the number of elements of A_2', and probability λ to action a_2'. For $\lambda > 0$ close enough to zero, an action a_1 which is a best reply to every $a_2 \in A_2'$ is also a best reply to α_2. This yields, by (18), that $u_1(a_1, a_2') = u_1(a_1, a_2)$ for all $a_2 \in A_2'$. Since a_1 is not a best reply to a_2', $u_1(a_1', a_2') > u_1(a_1, a_2')$ for every best reply a_1'. For $\lambda > 0$ close enough to one, there exists an action a_1' which is a best reply to a_2' and a best reply to α_2. By (18), $u_1(a_1', a_2') = u_1(a_1', \alpha_2)$, and so $u_1(a_1', a_2) \geq u_1(a_1', a_2')$ for some $a_2 \in A_2'$.

This yields $u_1(a_1', a_2) \geq u_1(a_1', a_2') > u_1(a_1, a_2') = u_1(a_1, a_2)$, and so a_1 is not a best reply to a_2, a contradiction. ■

We show now that the abnormal case only obtains for non-generic stage games, independently of the monitoring structure.

Proposition 6 *In the abnormal case, one of the players has a dominant action yielding the same payoff against all actions of the other player.*

Proof: Without loss of generality, suppose that there exists a non-negative p such that $\triangle M_1 p = 0$, $\triangle M_2 p > 0$. Then $\triangle M_1^{\mathcal{A}} \leq 0$ for all $\mathcal{A} \in \mathcal{J}$. Indeed, if $\triangle M_1^{\mathcal{A}} > 0$ for some $\mathcal{A}$ the j-th element of $\mathcal{J}$, then $p^\lambda = (1-\lambda)\, p + \lambda e_j$, where e_j is the j-th unit vector, would satisfy $\triangle M_i p^\lambda > 0$, $i = 1, 2$, for $\lambda > 0$ close enough to zero. For all $a \in A_1$, define regime $\mathcal{A}(a) = a \times A_2$. It is clear that this regime maximizes $\triangle M_1^{\mathcal{A}}$ among all regimes for which $a \in \mathcal{A}_1$. Suppose that a is a best reply to both α_2 and α_2'. Since $M_1^{\mathcal{A}(a)} \geq \max\{u_1(a, \alpha_2), u_1(a, \alpha_2')\} \geq \min\{u_1(a, \alpha_2), u_1(a, \alpha_2')\} \geq m_1^{\mathcal{A}(a)}$, it follows that $u_1(a, \alpha_2) = u_1(a, \alpha_2')$, and Lemma 2 applies. It follows immediately from (18) that the dominant action yields the same payoff against all actions of the other player.■

Nevertheless, the set of (belief-free) equilibrium payoffs can be characterized in the abnormal case as well. In the abnormal case, there exists a probability vector p and a player i such that $\sum_{\mathcal{A}} p(\mathcal{A}) \left(M_j^{\mathcal{A}} - m_j^{\mathcal{A}}\right) \geq 0$ with the inequality holding strictly for player $-i$ and binding for player i. Although there may be several such probability vectors, the player i for which the corresponding inequality binds is always the same, because otherwise any strict mixture of such probability vectors, where the corresponding inequality binds for different players, satisfies both inequalities strictly, and so the stage game is positive, not abnormal. By Proposition 6, player i has a dominant action yielding a constant payoff v. Because player i cannot be given intertemporal incentives, one can restrict attention to regimes $\mathcal{A}$ in which $u_i(a_i, a_{-i}) = v$ for all $(a_i, a_{-i}) \in \mathcal{A}$. If there is no probability vector p with support restricted to this subset of regimes such that $\sum_{\mathcal{A}} p(\mathcal{A}) \left(M_{-i}^{\mathcal{A}} - m_{-i}^{\mathcal{A}}\right) > 0$, then it is not possible to provide intertemporal incentives to player $-i$ either, and V reduces to the convex hull of the stage game's Nash equilibria. Otherwise, player $-i$'s payoff is, as in the positive case and by the same argument, the union over such probability vectors p of $[pm_{-i}, pM_{-i}]$. In either case, player i's equilibrium payoff is unique and equal to v. We summarize this discussion in the following proposition:

Proposition 7 *Suppose that the stage game is abnormal, and that player* 1 *has a dominant action yielding a constant payoff* v. *Let* $\mathcal{J}' = \left\{\mathcal{A} \in \mathcal{J} : \forall_{(a_1,a_2)\in\mathcal{A}}\ u_1(a_1, a_2) = v\right\}$. *If there exists* p *with support in* $\mathcal{J}'$, *such that* $\triangle M_2 p > 0$, *then* $V = \cup_{\{p\geq 0;\ supp(p)\subseteq\mathcal{J}',\ \triangle M_2 p\geq 0, p\mathbf{1}=1\}} \{v\} \times [m_2 p, M_2 p]$. *Otherwise* V *is the convex hull of the Nash equilibria of the bimatrix game.*

The theorem of the alternatives can be used to give alternative characterizations of the positive case, or to determine whether a payoff can be achieved in equilibrium. We conclude this subsection with two such results.

Proposition 8 *The positive case obtains if and only if, for all* $(x_1, x_2) \geq 0$,

$$\max_{\mathcal{A}\in\mathcal{J}} \left\{\triangle M_1^{\mathcal{A}} x_1 + \triangle M_2^{\mathcal{A}} x_2\right\} > 0.$$

The proof is an application of a theorem of the alternative (Theorem 2.10 of Gale (1960)).

Proposition 9 *In the positive case,* (V_1, V_2) *is an equilibrium payoff if and only if*

$$\min_{(x_1,x_2,x_3,x_4,x_5,x_6)\geq 0} \max_{\mathcal{A}\in\mathcal{J}} \left\{ \begin{array}{c} \left(M_1^{\mathcal{A}} - V_1\right) x_1 + \left(M_2^{\mathcal{A}} - V_2\right) x_2 + \left(V_1 - m_1^{\mathcal{A}}\right) x_3 \\ + \left(V_2 - m_2^{\mathcal{A}}\right) x_4 + \triangle M_1^{\mathcal{A}} x_5 + \triangle M_2^{\mathcal{A}} x_6 \end{array} \right\} \geq 0$$

This is again an application of a theorem of the alternative (Theorem 2.8 of Gale (1960)). Finally, as a last application of the same theorem, we state a necessary and sufficient condition for there not to exist any equilibrium payoff that (weakly) Pareto-dominates $V = (V_1, V_2)$: there exists a nonnegative vector (x_1, x_2, x_3, x_4) such that:

$$\max_{\mathcal{A}} \left\{\left(M_1^{\mathcal{A}} - V_1\right) x_1 + \left(M_2^{\mathcal{A}} - V_2\right) x_2 + \triangle M_1^{\mathcal{A}} x_3 + \triangle M_2^{\mathcal{A}} x_4\right\} < 0.$$

4.2 Vanishing Noise

How does the set of belief-free equilibrium payoffs vary with the information structure? In particular, what is the limit, if any, of this set, when the noise is arbitrarily small? To address these issues, we define $\Gamma(a_{-i})$ as the matrix whose (k,l)-th entry is the probability that player $-i$ observes signal σ^l_{-i} given that player i's action is a^k_i and player $-i$'s action is a_{-i}, i.e. $m_{-i}\left(\sigma^l_{-i} \mid a^k_i, a_{-i}\right)$, where signals σ_{-i} and actions a_i have been indexed, respectively, by l and k. Let $\Gamma = (\Gamma(a_{-i}), \text{ all } a_{-i}, i = 1,2)$ be the information structure. We write $V(\Gamma)$, $M^{\mathcal{A}}_i(\Gamma)$ and $m^{\mathcal{A}}_i(\Gamma)$ to denote V, $M^{\mathcal{A}}_i$ and $m^{\mathcal{A}}_i$ when the information structure is emphasized. Given a set of signals, we say that monitoring is ε-perfect if, for any player i and any action profile a, there exists a set of signals $\Sigma_i(a) \subset \Sigma_i$ such that $\sum_{\sigma_i \in \Sigma_i(a)} m_i(\sigma_i \mid a) > 1 - \varepsilon$, and $\sum_{\sigma_i \in \Sigma_i(a)} m_i(\sigma_i \mid a') < \varepsilon$ for any $a' \neq a$. Convergence to 0-perfect monitoring is denoted $\varepsilon \to 0$.

The next lemma shows that $M^{\mathcal{A}}_i(\Gamma)$ and $m^{\mathcal{A}}_i(\Gamma)$ are monotonic in $\mathcal{A}_i$, $\mathcal{A}_{-i}$ and in Γ, equipped with the Blackwell ordering. In particular, they have well-defined limits when the noise vanishes:

$$\bar{M}^{\mathcal{A}}_i := \max_{\alpha_{-i} \in \mathcal{A}_{-i}} \min_{a_i \in \mathcal{A}_i} u_i(a_i, \alpha_{-i}), \; \bar{m}^{\mathcal{A}}_i := \min_{\alpha_{-i} \in \mathcal{A}_{-i}} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}).$$

To understand these limits, consider the case of perfect information. Recall that $M^{\mathcal{A}}_i$ is the maximum payoff player i can get, given that he must be indifferent among all actions in $\mathcal{A}_i$, and given that his opponent is restricted to actions in $\mathcal{A}_{-i}$ and to negative transfers (fines). Therefore, player i's payoff is at most equal to his lowest payoff among all actions in $\mathcal{A}_i$, given his opponent's action $\alpha_{-i} \in \triangle\mathcal{A}_{-i}$. Hence, his maximal payoff is the maximum of this worst payoff over all mixed actions in $\triangle\mathcal{A}_{-i}$, and player $-i$, upon observing this "worst" action a_i, does not punish player i. Similarly, $\bar{m}^{\mathcal{A}}_i$ can be interpreted as the minimum payoff to which player $-i$ can hold player i to, without deviating himself from $\mathcal{A}_{-i}$ when player i is not restricted to $\mathcal{A}_i$. (Observe that, in the definition of $\bar{m}^{\mathcal{A}}_i$, the maximum is taken with respect to A_i, not $\mathcal{A}_i$).

These properties of $M^{\mathcal{A}}_i$ and $m^{\mathcal{A}}_i$ are recorded in the next lemma. Recall that: a matrix M is *Markov* if the entries of M are non-negative and the sum of entries of each row is equal to 1; a matrix A is a *garbling* of a matrix B if there exists a Markov matrix M such that $A = BM$. We say that Γ is a garbling of Γ' if $\Gamma(a_{-i})$ is a garbling of $\Gamma'(a_{-i})$, for all a_{-i} and all $i = 1,2$.

Lemma 3 *For all $i = 1,2$ and all $\mathcal{A}, \mathcal{A}' \in \mathcal{J}$,*

1. $\mathcal{A}_i = \mathcal{A}'_i$, $\mathcal{A}_{-i} \subseteq \mathcal{A}'_{-i}$ *implies that* $M^{\mathcal{A}}_i \leq M^{\mathcal{A}'}_i$ *and* $m^{\mathcal{A}}_i \geq m^{\mathcal{A}'}_i$;

2. $\mathcal{A}_i \subseteq \mathcal{A}'_i$, $\mathcal{A}_{-i} = \mathcal{A}'_{-i}$ *implies that* $M^{\mathcal{A}}_i \geq M^{\mathcal{A}'}_i$ *and* $m^{\mathcal{A}}_i \leq m^{\mathcal{A}'}_i$;

3. $M^{\mathcal{A}}_i \leq \bar{M}^{\mathcal{A}}_i$ *and* $\lim_{\varepsilon \to 0} M^{\mathcal{A}}_i = \bar{M}^{\mathcal{A}}_i$;

4. $m^{\mathcal{A}}_i \geq \bar{m}^{\mathcal{A}}_i$ *and* $\lim_{\varepsilon \to 0} m^{\mathcal{A}}_i = \bar{m}^{\mathcal{A}}_i$;

5. *If* Γ *is a garbling of* Γ', *then* $M_i^{\mathcal{A}}(\Gamma) \leq M_i^{\mathcal{A}}(\Gamma')$ *and* $m_i^{\mathcal{A}}(\Gamma) \geq m_i^{\mathcal{A}}(\Gamma')$.

Proof: 1. and 2. follow from the fact that increasing the number of constraints decreases the value of a linear program.

5. We prove the first part, the second part is analogous. Since $\Gamma(a_{-i})$ is a garbling of $\Gamma'(a_{-i})$, there exists a Markov matrix $M(a_{-i})$ such that $\Gamma(a_{-i}) = \Gamma'(a_{-i}) M(a_{-i})$. For $x_i : \mathcal{A}_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$, denote by $x_i(a_{-i})$ the column vector with coordinates $x_i(a_{-i}, \sigma_{-i}^l)$. Notice that if $\alpha_{-i} \in \triangle\mathcal{A}_{-i}$ and $x_i : \mathcal{A}_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$ achieve value v_i in the program (4) with noise $\Gamma(a_{-i})$, then α_{-i} and x_i' defined by

$$x_i'(a_{-i}) = M(a_{-i})\, x_i(a_{-i})$$

also achieve value v_i in the program (4) with noise $\Gamma'(a_{-i})$. Since $x_i : \mathcal{A}_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$ and the entries of $M(a_{-i})$ are non-negative, $x_i' : \mathcal{A}_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$.

3. Notice that $\bar{M}_i^{\mathcal{A}}$ solves the program (4) with no noise. Suppose first that $\Gamma(a_{-i})$ is a square matrix. By 5., $M_i^{\mathcal{A}} \leq \bar{M}_i^{\mathcal{A}}$. For $\varepsilon \to 0$, $\Gamma(a_{-i})$ converges to the identity matrix; in particular, $\Gamma(a_{-i})$ is invertible. Take $\alpha_{-i} \in \triangle\mathcal{A}_{-i}$ and $x_i : \mathcal{A}_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$ that achieve value v_i in the program (4) with no noise. Define x_i'' by

$$x_i''(a_{-i}) = \Gamma(a_{-i})^{-1} x_i(a_{-i}).$$

Then α_{-i} and x_i'' also achieve value v_i in the program (4) with noise $\Gamma(a_{-i})$. Of course, the coordinates of $x_i''(a_{-i})$ may be positive; say that the highest of them is equal to x. Then α_{-i} and

$$x_i'(a_{-i}) = x_i''(a_{-i}) - x$$

achieve value $v_i - x$ in the program (4) with noise $\Gamma(a_{-i})$; and $x_i' : \mathcal{A}_{-i} \times \Sigma_{-i} \to \mathbf{R}_-$. Since $\Gamma(a_{-i})^{-1}$ converges to the identity matrix, $x \to 0$ as $\varepsilon \to 0$, and so $M_i^{\mathcal{A}}$ converges to $\bar{M}_i^{\mathcal{A}}$.

Suppose now that $\Gamma(a_{-i})$ is not a square matrix. Since player $-i$ can perfectly infer a_i as $\varepsilon \to 0$, $\Gamma(a_{-i}) \to \Gamma^0(a_{-i})$ such that each column of $\Gamma^0(a_{-i})$ has at most one non-zero entry. Define matrix $\Gamma'(a_{-i})$ and vector $x_i'(a_{-i})$ as follows. Let k-th column of $\Gamma'(a_{-i})$ be the sum of all columns of $\Gamma(a_{-i})$ and $x_i'(a_{-i})$ be the sum of all entries of $x_i(a_{-i})$ such that the only non-zero entry of the corresponding column of $\Gamma^0(a_{-i})$ stands in k-th row. Apply now the argument for square $\Gamma(a_{-i})$ to $\Gamma'(a_{-i})$ and $x_i'(a_{-i})$.

4. is analogous. ■

Because $M_i^{\mathcal{A}}(\Gamma)$ and $m_i^{\mathcal{A}}(\Gamma)$ are monotonic in Γ, the set of belief-free equilibrium payoffs $V(\Gamma)$ itself is monotonic Γ. This result is formalized in the following corollary.

Corollary 2 *If* Γ *is a garbling of* Γ', *then* $V(\Gamma) \subset V(\Gamma')$.

Proof: It follows from Lemma 3, Propositions 5, 6 and 8. ■

4.3 Examples

When the noise is small enough, the positive case obtains for many common games, such as the prisoner's dilemma, or the battle of the sexes. Therefore, the procedure to determine the limit of the payoff set when $\delta \to 1$ and $\varepsilon \to 0$ is simple. First, compute for each of the (finitely many) regimes and for each player the values $\bar{M}_i^{\mathcal{A}}$ and $\bar{m}_i^{\mathcal{A}}$. Then, enumerate the vertices of the set $\{p \in \mathbf{R}_+ \mid p \cdot \mathbf{1} = 1,\ \triangle \bar{M}_i p \geq 0, i = 1, 2\}$, where $\mathbf{1}$ is the vector with all entries equal to one. Finally, for each of the vertices and each player, compute $\bar{M}_i p$ and $\bar{m}_i p$. Although it follows from the previous results that this procedure yields the set of belief-free equilibrium payoffs when the discount factor is taken to 1 first, and then the noise is taken to 0, it is clear from the previous results that the order of limits is irrelevant.

The following examples show that in general, belief-free strategies are not sufficient to establish a folk theorem for vanishing noise. In this respect, the prisoner's dilemma turns out to rather "exceptional". The crucial feature of the prisoner's dilemma is that, for any payoff v_i in $[0, 1]$, there exists an action by player $-i$ that guarantees that player i cannot get more than v_i, independently of his action, yet also another action that guarantees that player i cannot get less than v_i.

Example 1 *(The Battle of the Sexes) Consider the game:*

	L	R
T	$(2, 1)$	$(0, 0)$
B	$(0, 0)$	$(1, 2)$

To determine V, the limit (when $\delta \to 1$) equilibrium payoff set when the noise vanishes, we can derive the values in figure 1 from the formulas in section 4.2.

It follows from the table that V is spanned by $(2/3, 2/3)$, $(2, 1)$ and $(1, 2)$. The folk theorem does not obtain, but all Pareto-optimal payoffs are equilibrium payoffs. Interestingly, the payoff set is the convex hull of the Nash payoffs, yet we are in the positive case as

$$\frac{1}{2}\triangle \bar{M}_1^{\mathcal{A}_1} + \frac{1}{2}\triangle \bar{M}_1^{\mathcal{A}_2} = \left(\frac{1}{6}, \frac{1}{6}\right)$$

for $\mathcal{A}_1 = \{TB\} \times R$ and $\mathcal{A}_2 = T \times \{LR\}$. Note that with 2 replaced with $3/2$ in the payoff matrix, we would be in the negative case.

Example 2 *(One-Sided Prisoner's Dilemma) Consider the game:*

	L	R
T	$(2, 2)$	$(0, 3)$
B	$(0, 0)$	$(1, 1)$

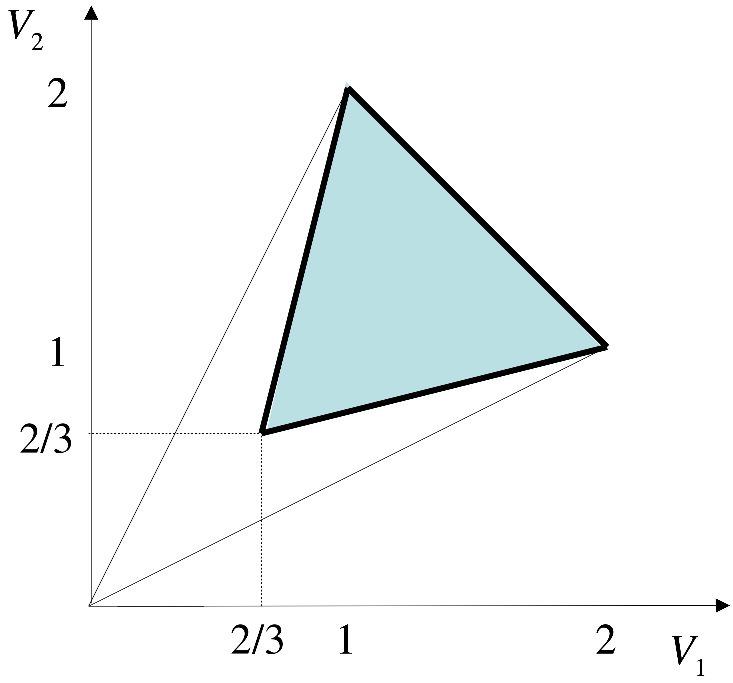

Figure 1: Battle of the Sexes

The table of values is displayed in figure 2 and the payoff set V is the convex hull of $(2/3, 1)$, $(1, 1)$, $(4/3, 2)$, $(4/3, 7/3)$, $(6/7, 18/7)$ and $(2/3, 2)$. The folk theorem fails, but some Pareto-optimal payoffs are equilibrium payoffs (but not the one maximizing the sum of payoffs).

Example 3 *Stag-Hunt game:*

	L	R
T	$(4, 4)$	$(0, 3)$
B	$(3, 0)$	$(2, 2)$

The set V is equal to the convex hull of $\{(2, 3), (3, 2), (4, 4), (2, 2)\}$, which is less than the convex hull of the feasible and individual rational payoff set.

Example 4 *(Prisoner's Dilemma). Consider*

	L	R
T	$(1, 1)$	$(-L, 1 + G)$
B	$(1 + G, -L)$	$(0, 0)$

where G, $L > 0$. The set V is the feasible and IR set (irrespective of whether $(1, 1)$ maximizes the sum of payoffs or not). If $(1, 1)$ does not maximize the sum of payoffs, then the extreme (asymmetric) payoff vector $(0, 1 + G - L)$ is obtained by playing regime $C \times D$ with probability $1 - L/(1 + G)$, regime $D \times \{CD\}$ with probability $L/(1 + G + L)$ and regime $\{CD\} \times D$ with complementary probability (probabilities are between 0 and 1 as $1 + G - L > 0$). Otherwise, the extreme asymmetric payoff $(0, 1 + G/(1 + L))$ is achieved by playing $C \times D$ with probability $1/(1+L)$, $C \times \{CD\}$ with probability $L/((1+G)(1+L))$ and $\{CD\} \times \{CD\}$ with complementary probability. The payoff $(1, 1)$ is achieved by playing the regime $\{CD\} \times \{CD\}$ with probability one.

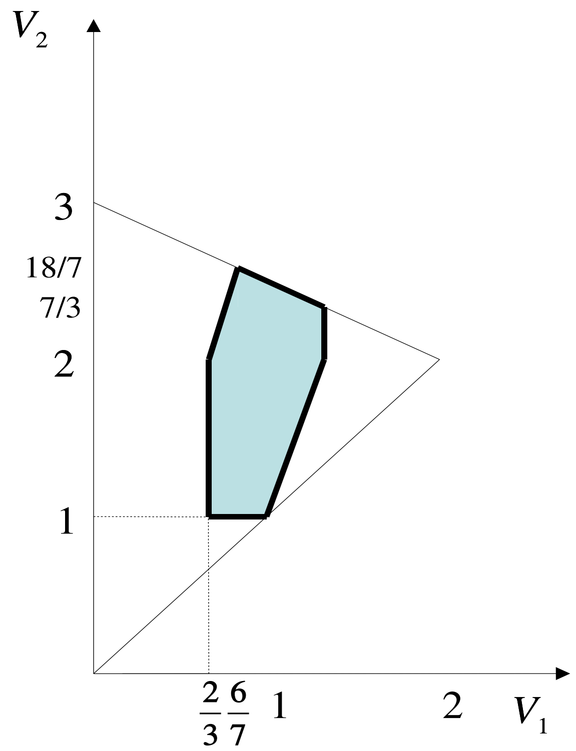

Figure 2: One-Sided Prisoner's Dilemma

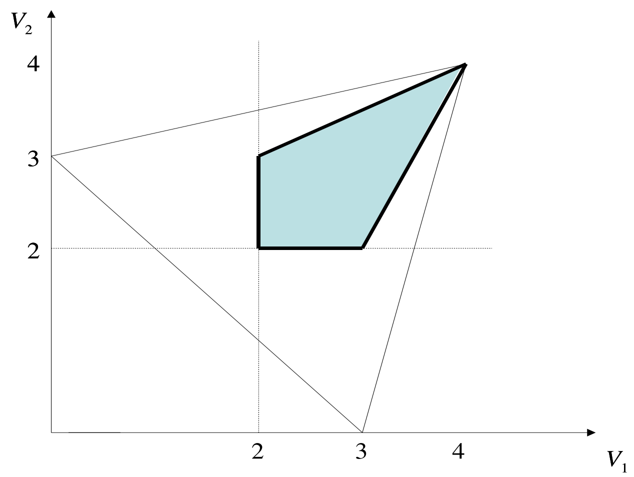

Figure 3: Stag Hunt

In all the previous examples, the limit set of belief-free equilibrium payoffs always included Pareto-optimal payoffs. This is not always the case: the following stage game admits a unique Nash equilibrium payoff vector $(1/5, 1/5)$, which is not Pareto-optimal. It is easy to verify that this game is negative (when the noise vanishes in particular), so that this payoff vector is also the unique belief-free equilibrium payoff vector in the infinitely repeated game, no matter how patient players are.

	L	R
T	$(2, -1)$	$(-1, 2)$
B	$(-1, 1)$	$(1, -1)$

5 Conditionally Independent Signals

Following VPE, in section 4 we examined the limiting set of belief-free equilibrium payoffs when monitoring is nearly perfect. We showed how to characterize this limit set and applied the characterization to some examples. In a recent paper, Matsushima (2002) demonstrated a new approach to constructing equilibria in repeated games with *conditionally independent* monitoring. Monitoring is conditionally independent if, given the realized action profile, the signals of the two players are statistically independent. Matsushima's construction augments the simple two-state machine strategies of Ely and Välimäki (2002) with a "review phase" similar to those introduced by Radner (1985). Applied to the Prisoner's Dilemma, Matsushima (2002) showed that these strategies can be used to establish the Folk Theorem even when the noise in monitoring is far from perfect.

The logic of the Matsushima (2002) construction is roughly as follows. Rather than switching between "reward" and "punishment" states in each period as the Ely and Välimäki (2002) strategies do, the Matsushima (2002) strategies remain in the current state for a T-stage review period during which information is collected and acted upon only at the end. The strategy plays a constant action throughout the review period: cooperation in the reward state, defection in the punishment state. Transition between states occurs at the end of the review period and depends (probabilistically) on the sequence of signals observed. The key is to construct these transitions so that in any state the most profitable deviation is also to a constant action. For example, in the reward state, the most profitable deviation is to defect in every stage of the review period. Given this, to detect a deviation by the opponent it is enough to test the hypothesis that the opponent deviated in *all* of the last T stages. No matter how weak the monitoring technology, T can be chosen large enough so that this test has arbitrarily high power. In this way, the T-stage review period can be treated as a single stage of the Ely and Välimäki (2002) strategies but with nearly perfect monitoring. Conditional independence of the monitoring is used to ensure that, within a review period, neither player obtains any information revealing whether or not he is likely to pass the test.

In this section, we generalize this scheme. We consider games with two actions and show

how to apply the Matsushima (2002) construction to equilibria with an arbitrary set of regimes (as opposed to the two-regime equilibria of Ely and Välimäki (2002)). The result is that some payoffs which we obtained in section 4 for small noise can be obtained under any conditionally independent monitoring technology.

Throughout this section, we assume that signals are conditionally independent. That is,

$$m(\sigma \mid a) = m_1(\sigma_1 \mid a)\, m_2(\sigma_2 \mid a)$$

for every $a \in A$ and $\sigma = (\sigma_1, \sigma_2) \in \Sigma$. We also assume that A_1 and A_2 consist of two actions; we do so because only under this assumption we are able to construct transitions between states so that in any state the most profitable deviation is to a constant action.

Finally we assume that for every action of player $-i$, the distribution of signals of player $-i$ depends on player i's action. That is,

$$m_{-i}\left(\cdot \mid a_i^1, a_{-i}\right) \neq m_{-i}\left(\cdot \mid a_i^2, a_{-i}\right)$$

whenever $a_i^1 \neq a_i^2$. This implies that there exists signals σ_{-i}^1 and σ_{-i}^2 such that:

$$m_{-i}\left(\sigma_{-i}^1 \mid a_i^1, a_{-i}\right) > m_{-i}\left(\sigma_{-i}^1 \mid a_i^2, a_{-i}\right) \tag{19}$$

and

$$m_{-i}\left(\sigma_{-i}^2 \mid a_i^1, a_{-i}\right) < m_{-i}\left(\sigma_{-i}^2 \mid a_i^2, a_{-i}\right); \tag{20}$$

of course, σ_{-i}^1 and σ_{-i}^2 may depend on a_{-i}.

Given a_{-i}, let $f_i^j(r, T, \tau)$ denote the probability that player $-i$ receives signal σ_{-i}^j exactly r times in T periods when player $-i$ plays a_{-i} in each out of the T periods and player i plays a_i^j in exactly τ out of the T periods. Let

$$F_i^j(r, T, \tau) := \sum_{s=1}^{r} f_i^j(r, T, \tau).$$

That is, $F_i^j(r, T, \tau)$ denotes the probability that player $-i$ receives signal σ_{-i}^j at most r times during the T periods when player $-i$ plays a_{-i} and player i plays a_i^j in exactly τ periods. Both $f_i^j(r, T, \tau)$ and $F_i^j(r, T, \tau)$ depend on a_{-i}, but when there is no confusion we will suppress this dependence to economize on notation.

The following two results are slightly reformulated versions of Matsushima (2002), Lemmas 1 and 2.

Lemma 4 *For every real number $C \geq 0$ and action a_{-i}, there exists a sequence of integers $(r_i^j(T))_{T=1}^{\infty}$ such that:*

$$\lim_{T\to\infty} F_i^j(r_i^j(T), T, 0) = 1, \tag{21}$$

$$\lim_{T\to\infty} F_i^j(r_i^j(T), T, T) = 0, \tag{22}$$

$$\liminf_{T\to\infty} T f_i^j(r_i^j(T), T-1, 0) \geq C \tag{23}$$

A function $h(\tau)$ for $\tau = 1, \dots, T$ is *single-peaked* if $h(\tau) \geq h(\tau+1)$ implies $h(\tau+1) \geq h(\tau+2)$.

Lemma 5 *For every $T = 1, 2, ...,$ and every $r = 0, ..., T-1$, $f_i^j(r, T-1, \tau-1)$, as a function of $\tau = 1, ..., T$, is single-peaked.*

For a given regime $\mathcal{A}$, define

$$N_i^{\mathcal{A}} := \max_{a_{-i} \in \mathcal{A}_{-i}} \min_{a_i \in \mathcal{A}_i} u_i(a_i, a_{-i}),$$

$$n_i^{\mathcal{A}} := \min_{a_{-i} \in \mathcal{A}_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i}),$$

and $\triangle N_i^{\mathcal{A}} = N_i^{\mathcal{A}} - n_i^{\mathcal{A}}$ and $\triangle N_i = N_i - n_i$; note that, compared to $\overline{M}_i^{\mathcal{A}}$ and $\overline{m}_i^{\mathcal{A}}$ defined in Section 4.2, the choice of player $-i$ is restricted to pure actions. This is used to ensure that, within a review period, neither player obtains any information revealing whether or not he is likely to pass the test. Remember that player $-i$ will play a_{-i} throughout the review period. If she were randomizing between playing the two actions, then player i could obtain, within the review period, information revealing whether or not he is likely to pass the test.

Proposition 10 *If there exists $p \geq 0$ such that $\triangle N_i p > 0$, $i = 1, 2$, then the limit of the set of equilibrium payoffs when $\delta \to 1$ contains*

$$\cup_{\{p \geq 0 : \triangle N_i p \geq 0, i=1,2, p\mathbf{1}=1\}} \prod_{i=1,2} [n_i p, N_i p].$$

Note that, for the Prisoner's Dilemma, $N_i^{\mathcal{A}} = \overline{M}_i^{\mathcal{A}}$ and $n_i^{\mathcal{A}} = \overline{m}_i^{\mathcal{A}}$ for all regimes $\mathcal{A}$, and so Proposition 10 indeed generalizes Matsushima (2002).

Proof: Fix a direct public randomization p. For a given regime $\mathcal{A}$, take $\overline{a}_{-i}^{\mathcal{A}} \in \mathcal{A}_{-i}$ with the property that

$$\min_{a_i \in \mathcal{A}_i} u_i\left(a_i, \overline{a}_{-i}^{\mathcal{A}}\right) = N_i^{\mathcal{A}}; \tag{24}$$

and let's suppose without loss of generality that $N_i^{\mathcal{A}} = u_i\left(a_i^1, \overline{a}_{-i}^{\mathcal{A}}\right)$. For a given T, we define define $\overline{x}_i^{\mathcal{A}}$ to be zero if $u_i(a_i^2, \bar{a}_{-i}^{\mathcal{A}}) < N_i^{\mathcal{A}}$ and otherwise the solution to

$$u_i\left(a_i^2, \overline{a}_{-i}^{\mathcal{A}}\right) - u_i\left(a_i^1, \overline{a}_{-i}^{\mathcal{A}}\right) = \bar{x}_i^{\mathcal{A}}\left[F_i^2(r_i^2(T), T, T) - F_i^2(r_i^2(T), T, 0)\right]. \tag{25}$$

and let

$$\overline{z}_i = \sum_{\mathcal{A} \in \mathcal{J}} p(\mathcal{A})\left[N_i^{\mathcal{A}} + (1 - F_i^2(r_i^2(T), T, 0))\overline{x}_i^{\mathcal{A}}\right].$$

Similarly, take $\underline{a}_{-i}^{\mathcal{A}} \in \mathcal{A}_{-i}$ with the property that

$$\max_{a_i \in A_i} u_i\left(a_i, \underline{a}_{-i}^{\mathcal{A}}\right) = n_i^{\mathcal{A}};$$

say that $n_i^{\mathcal{A}} = u_i\left(a_i^2, \underline{a}_{-i}^{\mathcal{A}}\right)$. For a given T, define $\underline{x}_i^{\mathcal{A}}$ by

$$u_i\left(a_i^2, \underline{a}_{-i}^{\mathcal{A}}\right) - u_i\left(a_i^1, \underline{a}_{-i}^{\mathcal{A}}\right) = \underline{x}_i^{\mathcal{A}}\left[m_{-i}\left(\sigma_{-i}^1 \mid a_i^1, \underline{a}_{-i}^{\mathcal{A}}\right)^T - m_{-i}\left(\sigma_{-i}^1 \mid a_i^2, \underline{a}_{-i}^{\mathcal{A}}\right)^T\right] \tag{26}$$

and let

$$\underline{z}_i = \sum_{\mathcal{A}\in\mathcal{J}} p(\mathcal{A})\left[n_i^{\mathcal{A}} + m_{-i}\left(\sigma_{-i}^1 \mid a_i^2, \underline{a}_{-i}^{\mathcal{A}}\right)^T \underline{x}_i^{\mathcal{A}}\right].$$

Note that $\bar{x}_i^{\mathcal{A}}$ is non-positive and $\underline{x}_i^{\mathcal{A}}$ is non-negative. It follows from (21) and (19) that

$$\lim_{T\to\infty} \overline{z}_i = N_i p \quad \text{and} \quad \lim_{T\to\infty} \underline{z}_i = n_i p.$$

For a given T, divide the time horizon into T-period blocks: $\{0, 1, ..., T-1\}, \{T, T+1, ...2T-1\}, ...$, and consider the auxiliary game in which each player is constrained to play the same pure action in all T periods of a given block. We can equivalently think of this as a repeated game whose stage game consists of T plays of the original stage game with the constraint that each player must play the same pure action in each of the T plays. In such a game, the effective discount factor is δ^T. At the end of each stage, the players observe the T signals generated by the plays within the stage. We will begin by constructing for this constrained game (for δ close enough to 1) and for every vector of the set

$$U = co\prod_{i=1,2}\{\underline{z}_i, \overline{z}_i\},$$

a belief-free equilibrium which achieves this vector. To do this, we will show that each extreme value is generated by U, and apply Proposition 4.

To generate $\bar{z}_i$, player $-i$ plays $\bar{a}_{-i}^{\mathcal{A}}$ in regime $\mathcal{A}$. Player i is then promised continuation values to begin the next block as a function of T-length sequences of signals defined as follows. First, player $-i$ partitions the set of T-sequences into two subsets $\Sigma_{-i}^{\text{Pass}}$ and $\Sigma_{-i}^{\text{Fail}}$. A sequence of signals belongs to $\Sigma_{-i}^{\text{Pass}}$ if at most $r_i^2(T)$ of the signals in the sequence are equal to σ_{-i}^2, otherwise the sequence is assigned to $\Sigma_{-i}^{\text{Fail}}$. The continuation value for i in regime $\mathcal{A}$ will depend only on which of these sets the sequence of signals observed by $-i$ belongs to. In $\Sigma_{-i}^{\text{Pass}}$, the continuation value is $\bar{z}_i$, in $\Sigma_{-i}^{\text{Fail}}$, the continuation value is

$$\overline{z}_i + \frac{1-\delta^T}{\delta^T}\overline{x}_i^{\mathcal{A}}$$

These continuation values belong to U when δ is close enough to 1.

Against this, when player i plays a_i^1, his payoff is

$$(1-\delta^T)u_i(a_i^1, \bar{a}_{-i}^{\mathcal{A}}) + \delta^T\left[\bar{z}_i + (1 - F(r_i^2(T), T, 0))\frac{1-\delta^T}{\delta^T}\bar{x}_i^{\mathcal{A}}\right] \tag{27}$$

and when he plays a_i^2, his payoff is

$$(1-\delta^T)u_i(a_i^2,\bar{a}^{\mathcal{A}}_{-i})+\delta^T\left[\bar{z}_i+(1-F(r_i^2(T),T,T))\frac{1-\delta^T}{\delta^T}\bar{x}_i^{\mathcal{A}}\right]$$

By our construction of $\bar{x}_i^{\mathcal{A}}$, i weakly prefers a_i^1 and is indifferent when $u_i(a_i^2,\bar{a}^{\mathcal{A}}_{-i})\geq N_i^{\mathcal{A}}$. By (24), the latter case must hold when $a_i^2\in\mathcal{A}_i$. Thus, each $\mathcal{A}_i$ is enforced, and i's payoff in regime $\mathcal{A}$ is given by (27). It can be rewritten as

$$(1-\delta^T)\left[N_i^{\mathcal{A}}+(1-F(r_i^2(T),T,0))\bar{x}_i^{\mathcal{A}}\right]+\delta^T\bar{z}_i$$

Multiplying by $p(\mathcal{A})$ and summing over regimes shows that i's expected payoff is $\bar{z}_i$. We show that $\underline{z}_i$ is generated by U by a similar argument. In this case we use the action $\underline{a}^{\mathcal{A}}_{-i}$ and the continuation values are $\underline{z}_i+\frac{1-\delta^T}{\delta^T}\underline{x}_i^{\mathcal{A}}$ when *all* of the signals are equal to σ^1_{-i} and $\underline{z}_i$ itself if at least one signal is different from σ^1_{-i}.

We now apply Proposition 4 to obtain for all δ close enough to 1, for each value $u\in U$, a belief free equilibrium s (with the structure given by Proposition 4) which achieves value u. We can extend s to a fully specified strategy profile of the unconstrained game by assuming that each player ignores his own deviations. We will show that for T large and δ close to 1, (this extended) s is a Nash equilibrium of the unconstrained game.

By the principle of optimality, it is enough to show that player i cannot profit by deviating to an alternative T-period strategy within some block and then return to the continuation strategy dictated by s at the beginning of the next block. Because signals are conditionally independent, i's beliefs about the continuation strategy of $-i$ are independent of i's own history of signals, conditional on his own history of actions. Thus it suffices to consider deviations to within-block strategies whose play does not depend on the signals observed within the block. Because the payoff to any such strategy depends only on the sequence of actions it induces, it is enough to check that i cannot improve his payoff by deviating in some block to a sequence of actions different than those prescribed by s.

Consider any regime $\mathcal{A}$. Player $-i$ will either play $\bar{a}^{\mathcal{A}}_{-i}$ or $\underline{a}^{\mathcal{A}}_{-i}$. We will show that deviations are unprofitable in either case. Suppose first that player $-i$ plays $\bar{a}^{\mathcal{A}}_{-i}$. If $u_i\left(a_i^2,\bar{a}^{\mathcal{A}}_{-i}\right)<u_i\left(a_i^1,\bar{a}^{\mathcal{A}}_{-i}\right)$ and $\mathcal{A}=\{a_i^1\}$, then the continuation value beginning in the next block is independent of history and hence player i prefers playing the static best-reply a_i^1 in all T periods to playing any other sequence of actions. Otherwise, $u_i\left(a_i^2,\bar{a}^{\mathcal{A}}_{-i}\right)\geq u_i\left(a_i^1,\bar{a}^{\mathcal{A}}_{-i}\right)$ and, by construction, player i is indifferent between playing a_i^1 in all T periods and playing a_i^2 in all T periods. We have to show that player i prefers this to playing a_i^1 in some periods and a_i^2 in other periods. Because of discounting we can without loss of generality restrict attention to sequences such that a_i^2 is played in the first τ periods and a_i^1 is played in the remaining $T-\tau$ periods of the block.

Let $V(0)$ denote player i's payoff to playing a_i^1 in all T periods, and let $V(\tau)$ denote player i's payoff to playing a_i^2 in the first τ periods and a_i^1 in the remaining $T-\tau$ periods. Let

$y := u_i\left(a_i^2, \overline{a}_{-i}^{\mathcal{A}}\right) - u_i\left(a_i^1, \overline{a}_{-i}^{\mathcal{A}}\right)$. Then

$$V(\tau) - V(0) = (1-\delta^\tau)y + \delta^T\left[F_i^2(r_i^2(T), T, 0) - F_i^2(r_i^2(T), T, \tau)\right]\left(\frac{1-\delta^T}{\delta^T}\overline{x}_i^{\mathcal{A}}\right),$$

which by (25) can be rewritten

$$V(\tau) - V(0) = y(1-\delta^T)\left[\frac{1-\delta^\tau}{1-\delta^T} - g(\tau)\right]$$

where

$$g(\tau) := \frac{F_i^2(r_i^2(T), T, 0) - F_i^2(r_i^2(T), T, \tau)}{F_i^2(r_i^2(T), T, 0) - F_i^2(r_i^2(T), T, T)}$$

Note that $g(0) = 0$, $g(T) = 1$. We wish to show that T can be chosen large enough and δ close enough to 1 so that $V(\tau) - V(0) \leq 0$. Since

$$\lim_{\delta \to 1}\left(\frac{1-\delta^\tau}{1-\delta^T}\right) = \frac{\tau}{T},$$

it suffices to show that we can choose T sufficiently large so that

$$h(\tau) := g(\tau) - \frac{\tau}{T} > 0 \qquad \text{for } \tau = 1, ..., T-1. \tag{28}$$

Write $\triangle g(\tau) = g(\tau+1) - g(\tau)$. We have

$$\triangle g(\tau) = \frac{[F_i^2(r_i^2(T), T, \tau) - F_i^2(r_i^2(T), T, \tau+1)]}{[F_i^2(r_i^2(T), T, 0) - F_i^2(r_i^2(T), T, T)]}$$

Observe that

$$\begin{aligned} F_i^2(r_i^2(T), T, \tau) - F_i^2(r_i^2(T), T, \tau+1) &= f_i^2(r_i^2(T), T-1, \tau) \cdot \\ &\cdot \left[m_{-i}\left(\sigma_{-i}^2 \mid a_i^2, a_{-i}\right) - m_{-i}\left(\sigma_{-i}^2 \mid a_i^1, a_{-i}\right)\right]. \end{aligned} \tag{29}$$

so that by Lemma 5, the function $\triangle g(\tau)$ is a constant times a single-peaked function and is therefore itself single-peaked as a function of $\tau = 0, \ldots, T-1$.

Now by (23) we can choose T large enough to satisfy

$$f_i^2(r_i^2(T), T-1, 0) > \frac{1}{T\left[m_{-i}\left(\sigma_{-i}^2 \mid a_i^2, a_{-i}\right) - m_{-i}\left(\sigma_{-i}^2 \mid a_i^1, a_{-i}\right)\right]}$$

and by (21), (22) and (29),

$$\triangle g(0) := g(1) > \frac{1}{T} \tag{30}$$

which further implies $h(1) > 0$. Finally, we show that for such large T, the function $h(\tau)$ is single-peaked. Since $h(1) > 0$ and $h(T) = 0$, this will establish (28).

Suppose $h(\tau) \geq h(\tau+1)$. This is equivalent to

$$\triangle g(\tau) := g(\tau+1) - g(\tau) \leq \frac{\tau+1-\tau}{T} = \frac{1}{T} < \triangle g(0)$$

Since $\triangle g(\tau)$ is single-peaked, we have then

$$\begin{aligned}\triangle g(\tau+1) &\leq \triangle g(\tau)\\ g(\tau+2) - g(\tau+1) &\leq \frac{1}{T} = \frac{\tau+2}{T} - \frac{\tau+1}{T}\\ g(\tau+1) - \frac{\tau+1}{T} &\geq g(\tau+2) - \frac{\tau+2}{T}\end{aligned}$$

and the latter is equivalent to $h(\tau+1) \geq h(\tau+2)$. Thus, $h(\tau)$ is single-peaked.

Suppose now that player $-i$ plays $\underline{a}^{\mathcal{A}}_{-i}$. Re-using notation from the previous step, let $V(0)$ denote player i's payoff to playing a^1_i in all T periods, and let $V(\tau)$ denote player i's payoff to playing a^2_i in the first τ periods and a^1_i in the remaining $T-\tau$ periods. Let $y := u_i\left(a^2_i, \underline{a}^{\mathcal{A}}_{-i}\right) - u_i\left(a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)$. Then $V(\tau) - V(0)$ is equal to

$$(1-\delta^\tau)y - \delta^T m_{-i}\left(\sigma^1_{-i} \mid a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)^{T-\tau} \cdot \left[m_{-i}\left(\sigma^1_{-i} \mid a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)^\tau - m_{-i}\left(\sigma^1_{-i} \mid a^2_i, \underline{a}^{\mathcal{A}}_{-i}\right)^\tau\right]\left(\frac{1-\delta^T}{\delta^T}\underline{x}^{\mathcal{A}}_i\right) =$$

$$= (1-\delta^T)y\left\{\frac{(1-\delta^\tau)}{(1-\delta^T)} - g(\tau)\right\}$$

where

$$g(\tau) = \frac{m^{-i}\left(\sigma^1_{-i} \mid a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)^{T-\tau}\left[m_{-i}\left(\sigma^1_{-i} \mid a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)^\tau - m_{-i}\left(\sigma^1_{-i} \mid a^2_i, \underline{a}^{\mathcal{A}}_{-i}\right)^\tau\right]}{\left[m_{-i}\left(\sigma^1_{-i} \mid a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)^T - m_{-i}\left(\sigma^1_{-i} \mid a^2_i, \underline{a}^{\mathcal{A}}_{-i}\right)^T\right]}$$

Since

$$\lim_{\delta\to 1}\frac{(1-\delta^\tau)}{(1-\delta^T)} = \frac{\tau}{T},$$

it suffices to show that

$$g(\tau) > \frac{\tau}{T} \tag{31}$$

for $\tau = 1, ..., T-1$.

To this end observe that

$$g(\tau) = \frac{1-c^\tau}{1-c^T},$$

where

$$c = \frac{m_{-i}\left(\sigma^1_{-i} \mid a^2_i, \underline{a}^{\mathcal{A}}_{-i}\right)}{m_{-i}\left(\sigma^1_{-i} \mid a^1_i, \underline{a}^{\mathcal{A}}_{-i}\right)} < 1,$$

and therefore $g(\tau)$ is an increasing and concave function of τ, which takes value 0 at $\tau = 0$ and takes value 1 at $\tau = T$. This implies (31). ■

6 Concluding Comments

This paper has defined and analyzed a class of strategies, *belief-free strategies*, for two-player games. This class of strategies has the appealing feature that, in equilibrium, the support of optimal strategies is history-independent. We provide a recursive characterization of the equilibrium payoff set supported by such strategies, *strong self-generation*, which builds on and extends the concept of self-generation, of Abreu, Pearce, and Stachetti (1990). This payoff set includes the equilibrium payoffs explicitly computed by Ely and Välimäki (2002) and Piccione (2002). As for self-generation, there exists a version of value iteration which holds for this set. Similarly, attention can be restricted to bang-bang strategies. Finally, we offer a two-stage linear programming procedure to compute this payoff set when the discount factor tends to one.

Attention has been restricted to two-player games. It is straightforward to generalize the definition of belief-free strategies and of regimes to more than two players. However, it is no longer true that, for a fixed sequence of regimes, the payoff set has a product structure. Therefore, we cannot decompose the computation of the payoff set into distinct linear programs any longer, and determining the equilibrium payoff set becomes significantly harder.

By definition, belief-free strategies do not require keeping track of beliefs. As we provide several examples in which the equilibrium payoff set is a strict subset of the feasible and individually rational payoff set, it is therefore clear that any attempt to generalize folk theorems to games with private monitoring must use more complicated, belief-based strategies. Such strategies are the subject of ongoing research.

	$T \times L$	$T \times R$	$B \times L$	$B \times R$	$\{TB\} \times L$	$\{TB\} \times R$	$T \times \{LR\}$	$B \times \{LR\}$	$\{TB\} \times \{LR\}$
$\bar{M}_1^{\mathcal{A}}$	2	0	0	1	0	0	2	1	2/3
$\bar{M}_2^{\mathcal{A}}$	1	0	0	2	1	2	0	0	2/3
$\bar{m}_1^{\mathcal{A}}$	2	1	2	1	2	1	2/3	2/3	2/3
$\bar{m}_2^{\mathcal{A}}$	1	1	2	2	2/3	2/3	1	2	2/3
$\triangle\bar{M}_1^{\mathcal{A}}$	0	−1	−2	0	−2	−1	4/3	1/3	0
$\triangle\bar{M}_2^{\mathcal{A}}$	0	−1	−2	0	1/3	4/3	−1	−2	0

Table 1: Table of Values for Battle of the Sexes

	$T \times L$	$T \times R$	$B \times L$	$B \times R$	$\{TB\} \times L$	$\{TB\} \times R$	$T \times \{LR\}$	$B \times \{LR\}$	$\{TB\} \times \{LR\}$
$\bar{M}_1^{\mathcal{A}}$	2	0	0	1	0	0	2	1	2/3
$\bar{M}_2^{\mathcal{A}}$	2	3	0	1	2	3	2	0	2
$\bar{m}_1^{\mathcal{A}}$	2	1	2	1	2	1	2/3	2/3	2/3
$\bar{m}_2^{\mathcal{A}}$	3	3	1	1	1	1	3	1	1
$\triangle\bar{M}_1^{\mathcal{A}}$	0	−1	−2	0	−2	−1	4/3	1/3	0
$\triangle\bar{M}_2^{\mathcal{A}}$	−1	0	−1	0	1	2	−1	−1	1

Table 2: Table of Values for One-Sided Prisoner's Dilemma

References

Abreu, D., D. Pearce, and E. Stachetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

Altman, E. (1999): *Constrained Markov Decision Processes.* Chapman Hall/CRC, Boca Raton.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 40–69.

Ely, J. C., J. Hörner, and W. Olzsweski (2004): "Dispensing with Public Randomization in Belief-Free Equilibria," `http://www.kellogg.northwestern.edu/faculty/horner/ehoappendix.pdf`.

Ely, J. C., and J. Välimäki (2000): "Notes on Private Monitoring with Non-Vanishing Noise," Northwestern University.

——— (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102, 84–105.

Fudenbeg, D., and E. Maskin (1990): "Nash and Perfect Equilibria of Discounted Repeated Games," *Journal of Economic Theory*, 51, 194–206.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), 997–1040.

Fudenberg, D., and D. K. Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103–135.

Fudenberg, D., and E. Maskin (1991): "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*, 53, 428–438.

Fudenberg, D., and J. Tirole (1991): *Game Theory.* MIT Press, Cambridge.

Gale, D. (1960): *The Theory of Linear Economic Models.* The University of Chicago Press, Chicago.

Hörner, J., and W. Olszewski (2002): "The "Belief-Free" Approach in the Repeated Prisoner's Dilemma with Private Monitoring," Northwestern University.

Kandori, M. (2002): "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102, 1–15.

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66, 627–652.

Kandori, M., and I. Obara (2003): "Efficiency in Repeated Games Revisited: The Role of Private Strategies," UCLA.

Mailath, G. J., S. A. Matthews, and T. Sekiguchi (2002): "Private Strategies in Finitely Repeated Games with Imperfect Public Monitoring," *Contributions to Theoretical Economics*, 2(1), `http://www.bepress.com/bejte/contributions/vol2/iss1/art2`.

Mailath, G. J., and S. Morris (2002): "Repeated Games with Almost Public Monitoring," *Journal of Economic Theory*, 102, 189–228.

Matsushima, H. (2002): "The Folk Theorem for the Prisoner's Dilemma with Imperfect Private Monitoring," University of Tokyo.

Piccione, M. (2002): "The Repeated Prisoners' Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70–83.

Radner, R. (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53(5), 1173–1198.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 76(2), 345–361.

Stigler, G. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 72, 44–61.