Instability of Equilibria with Imperfect Private Monitoring

Heller, Yuval

University of Oxford

19 May 2015

Instability of Equilibria with Imperfect Private Monitoring

Yuval Heller*

Department of Economics and Queen's College, University of Oxford.

May 20, 2015

Abstract

Various papers have presented folk-theorem results that yield efficiency in the repeated Prisoner's Dilemma with imperfect private monitoring. I present a mild refinement that requires robustness against small perturbations in the behavior of potential opponents, and I show that only defection satisfies this refinement among all the existing equilibria in the literature.

JEL Classification: C73, D82. **Keywords**: Belief-free equilibrium, evolutionary stability, imperfect private monitoring, repeated Prisoner's Dilemma, communication.

1 Introduction

The theory of repeated games provides a formal framework to explore the possibility of cooperation in long-term relationships, such as collusion between firms. The various folk theorem results (e.g., Fudenberg & Maskin, 1986) have established that efficiency can be achieved under fairly general conditions. The classical literature assumes that players perfectly observe the past actions. Several papers in the 90-s (e.g., Abreu *et al.*, 1990) have extended these results to the case of imperfect but commonly shared information about the past action profiles.

In many real life situations players privately observe imperfect signals about the past actions. For example, each firm in a cartel privately observes its own sales, which contain imperfect information about secret prices cuts that its competitors offer to some of the customers. Formal analysis of imperfect private monitoring began with the pioneering work of Sekiguchi (1997). Since then, several papers have presented various folk theorem results that have shown that

*Email: yuval.heller@economics.ox.ac.uk. I would like to express our deep gratitude to Mehmet Ekmekci, Peter Eso, and Erik Mohlin, for many useful comments.

efficiency can be achieved also with imperfect private monitoring (see Kandori, 2002; Mailath & Samuelson, 2006 for surveys of this literature).

The current paper presents the equilibrium refinement of weak stability, which requires robustness against small perturbations in the behavior of potential opponents, and shows that only defection satisfies this mild refinement in the repeated Prisoner's Dilemma among all the equilibria in the existing literature on imperfect monitoring.[1]

Weak Stability One of the leading justifications for using a Nash equilibrium to predict behavior is its interpretation as a stable convention in a population of potential players. Suppose that individuals in a large population are repeatedly drawn to play a game, and that initially all individuals play the strategy s^* but occasionally a small group of agents may experiment with a a different strategy s'. If this induces the experimenting agents to gain more than the incumbents, than the population would move away from s^* towards s'. Thus, strategy s^* is *evolutionarily stable* (Maynard Smith & Price, 1973) if (1) it is a best-reply to itself (i.e., it is a symmetric Nash equilibrium),[2] and (2) it achieves a strictly higher payoff against any other best-reply strategy s': $u(s^*, s') > u(s', s')$. One example for an evolutionarily stable strategy in the repeated Prisoner's Dilemma is the strategy of always defecting regardless of the history.

Evolutionarily stability is arguably a too-strong refinement as demonstrated in the Rock-Scissors-Paper game (see Section 2.2) that admits a unique Nash equilibrium, which is not evolutionarily stable, but it is a plausible prediction for the long run average behavior in the population (see, e.g., Benaïm *et al.*, 2009). Motivated by this, I say that a symmetric Nash equilibrium s^* is *weakly stable* if it achieves a *weakly* higher payoff against any *evolutionarily stable* strategy s' that is a best-reply against s^*.[3] The definition implies that any symmetric game admits a weakly stable strategy, and that if s^* is not weakly stable, then it is not a plausible prediction to the long run behavior. This is because as soon as a small group of agents experiments with playing s', the population diverges away towards s', and it will remain in s' in the long run (due to s' being evolutionarily stable). A typical example for a non-weakly stable equilibrium is the mixed equilibrium in a coordination game, for which every small perturbation takes the population to one of the pure equilibria.

[1] Except the equilibria that rely on sufficient correlation between the private signals of different players (Mailath & Morris, 2002, 2006; Kandori, 2011), and the equilibria that rely on communication between the players at each round of the game (Compte 1998; Kandori & Matsushima 1998; Obara 2009), as discussed in Section 1.1.

[2] To simplify the analysis we focus in most of the paper only on symmetric equilibria, and we extend the analysis to asymmetric equilibria in Section 5.1.

[3] The formal definition (see Section 2.2) also allows a two-step movement towards the evolutionarily stable strategy s' (first from s^* to some strategy $\tilde{s}$, and then from $\tilde{s}$ to s'). To simplify the exposition, we ignore this slight complication in the introduction. Note, that all the equilibria in the existing literature are destabilized by a single step; the two-step part of the formal definition is only required against hypothetical belief-free equilibria.

Summary of Results The most commonly used equilibrium in the literature on imperfect private monitoring is the *belief-free equilibrium* in which the continuation strategy of each player is a best reply to his opponent's strategy at every private history. These equilibria are called "belief-free" because a player's belief about his opponent's history is not needed for computing a best-reply. Piccione (2002) and Ely & Välimäki (2002) have presented folk-theorem results for the repeated Prisoner's Dilemma using belief-free equilibria. In particular, they show that any payoff between zero and one is a symmetric belief-free equilibrium payoff if the monitoring technology is almost perfect and the players are sufficiently patient.

The main result of this paper shows that under the mild refinement of weak stability, the repeated Prisoner's Dilemma admits a unique belief-free equilibrium in which everyone defects. The intuition is that in belief-free equilibria players are always indifferent between the two actions, but this implies that always defecting is a best-reply, and as soon as a small group of agents starts defecting, the entire population diverges away towards defection.

Next I extend the above main result to the set of belief-free review-strategy equilibria (Matsushima, 2004; Hörner & Olszewski, 2006; Yamamoto, 2007, 2012) in which (1) the game is divided to a sequence of review phases in which each player plays a constant action, and (ii) at the beginning of each review phase, a player's continuation strategy is a best reply regardless of the history.

Bhaskar & Obara (2002) have presented a folk-theorem result for the repeated Prisoner's Dilemma that does not rely on belief-free equilibria. These "belief-based" equilibria divide the repeated game into several subgames, such that each player is indifferent in the beginning of each subgame between defecting in the entire subgame and playing a more cooperative strategy. My last result shows that none of these equilibria are weakly stable. The intuition is based on the observation that always defecting is a best-reply to these equilibria, and that it implies that a small perturbation would diverge the entire population towards defection.

1.1 Related Literature and Contribution

Belief-Free equilibria Ely *et al.* (2005), Miyagawa *et al.* (2008), and Yamamoto (2009, 2013) have extended the folk-theorem results that rely on belief-free equilibria to general repeated games and to costly observability. Kandori & Obara (2006) study a setup of imperfect *public* monitoring and show that belief-free private strategies can improve the efficiency relative to the maximal efficiency obtained by public strategies. Takahashi (2010, Section 3.2) adapts the notion of belief-free equilibria to obtain a folk-theorem result to a setup in which the players are randomly matched with a new opponent at each round. My main result shows that none of the equilibria in these papers satisfies weak stability in the repeated Prisoner's Dilemma.

Conditionally-Correlated Signals A few papers in the literature yield stable cooperation if the private signals are sufficiently correlated conditional on the action-profile. Mailath & Morris (2002, 2006) show that when the private signals are almost perfectly correlated conditional on the action-profile (i.e., *almost public monitoring*), then any sequential equilibrium of the nearby public monitoring game with bounded memory remains an equilibrium also with almost public monitoring. Some of these equilibria are evolutionarily stable, and in particular, cooperation can be the outcome of an evolutionarily stable strategy.

Kandori (2011) presents the notion of weakly belief-based equilibria, in which the strategy of each player is a best-reply for any private history of the opponent up to the actions of the previous round. Unlike standard belief-based equilibria, players need to form the correct beliefs about the signal obtained by the opponent in the last round. Kandori (2011) demonstrates that sufficient correlation between the private signals (conditional on the action-profile) allow to achieve a strict weakly belief-based equilibrium that yields substantial cooperation. The strictness of the equilibrium implies that it satisfies the refinement of evolutionary stability.

The current paper shows that all the mechanisms in the existing literature can only yield defection as the outcome of a weakly-stable equilibrium in the repeated Prisoner's Dilemma with conditionally-independent imperfect monitoring. I leave for future research the question whether any new mechanism may yield cooperation as a stable outcome with conditionally-independent private monitoring. One promising direction to solve this question might rely on the methods developed in Heller & Mohlin (2015) for the related setup of random matching and partial observation of the partner's past behavior. In that setup, Heller & Mohlin (2015) characterize conditions under which only defection is stable, and construct novel mechanisms to sustain stable equilibria whenever these conditions are not satisfied.

Robustness Sugaya & Takahashi (2013) show that "generically" only belief-free equilibria are robust to small perturbations in the monitoring structure. Our main result shows that belief-free equilibria (except defection) are not robust to small perturbations in the behavior of the potential opponents. Taken together, the two results suggest that defection is the unique equilibrium outcome of the repeated Prisoner's Dilemma that is robust to both kinds of perturbations.

Communication Compte (1998), Kandori & Matsushima (1998), and Obara (2009) have presented folk-theorem results that rely on (noiseless) communication between the players at each stage of the repeated game. The players use this communication to publicly report (possibly with some delay) the private signals they obtained. These equilibria are constructed such that the players have strict incentives while playing, but they are indifferent between reporting the truth and lying (regardless of the reporting strategy of the opponent). One can show that this property implies that these equilibria are neutrally stable (and thus also weakly stable).

The current paper has important implications to antitrust laws. If one agrees that weak stability is necessary for a collusive equilibrium to be plausible, then it implies that communication between the players is critical to obtain collusive behavior whenever the private imperfect monitoring between the firms is such that the conditional correlation between the private signals is sufficiently low.

Structure The model is described in Section 2. Section 3 presents the main result (instability of belief-free equilibria). Sections 4 shows the instability of the other equilibria in the literature. Section 5 extends the analysis to asymmetric equilibria and to monitoring without full-support.

2 Model

2.1 Repeated Prisoner Dilemma with Imperfect Monitoring

I analyze a symmetric two-player $\delta-$discounted repeated interaction with imperfect private monitoring. Table 1 presents the matrix payoff of the stage game: a Prisoner's Dilemma game that depends on two positive parameters g and l. When both players play action C (*cooperate*) they both get a high payoff (normalized to one), and when they both play D (*defect*) they get a low payoff (normalized to zero). When a player defects and the opponent cooperates, the defector gets $1+g$, while the opponent gets $-l$. The players share a common discount factor $0 \leq \delta < 1$.

Table 1: Matrix Payoff of a Prisoner's Dilemma Game ($g, l > 0$)

	C	D
C	1, 1	$-l$, $1+g$
D	$1+g$, $-l$	0, 0

Let $i \in \{1,2\}$ be an index denoting one of the players, and let $-i$ denote the other player. We use $\Delta(W)$ to represent the set of probability distributions over a finite set W. Let Σ_i be a finite set of signals (or messages) for each player. For each possible action profile $a \in \{C, D\}^2$, the monitoring distribution $m(\cdot|a)$ specifies a symmetric joint probability distribution over the finite set of signal profiles $\Sigma := \Sigma_1 \times \Sigma_2$. When action profile a is played and signal profile σ is realized, each player i privately observes his corresponding signal σ_i.

Let $m_i(\cdot|a)$ denote the marginal distribution over the signals of player i. To simplify the analysis, we assume that the monitoring structure has full support, i.e., that each signal profile is observed with positive probability after each action profile. Formally:

Assumption 1. *The monitoring structure has* full support: $m(\sigma|a) > 0$ *for each action profile* $a \in \{C, D\}^2$ *and each signal* $\sigma \in \Sigma$.

One example for a monitoring structure with full support is the *conditionally-independent $\epsilon-$perfect monitoring* in which each player privately observes his opponent's last action with probability $1 - \epsilon$ and observes the opposite action with the remaining probability ϵ. In Section 5.2 we show that our results are essentially the same also without Assumption 1.

A t-length (private) history for player i is a sequence that includes the action played by the player and the observed signal in each of the previous t rounds of the game. Let $H_i^t := (\{C, D\} \times \Sigma_i)^t$ denote the set of all such histories for player i. Each player's initial history is the null history, denoted ϕ. A pair of t-length histories (called simply a history) is denoted h^t. Let H^t denote the set of all 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 $s : H_i \to \Delta(A_i)$. Let s_D be the strategy that always defect ($s_D(\cdot) = D$).

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 $s|_{h_i^t}$ is the strategy defined by $s|_{h_i^t}\left(\hat{h}_i\right) = s\left(h_i\hat{h}_i\right)$. Given a strategy s', let $B(s')$ denote the set of strategies that are best replies to s', and for each t and $h_{-i}^t \in H_{-i}^t$ let $B\left(s'|h_{-i}^t\right)$ denote the set of continuation strategies for i that are best replies to $s'_{|h_{-i}^t}$.

Definition 1 (Ely *et al.* 2005)**.** A strategy profile (s, s') is *belief-free* if for every h^t, $s|_{h_i^t} \in B\left(s'|h_{-i}^t\right)$ for $i \in \{1, 2\}$.

The condition characterizing belief-free strategy profile is stronger than that characterizing sequential equilibrium. In a sequential equilibrium, a player's continuation strategy is the player's best reply given his belief about his opponent's continuation strategy, that is, given a unique probability distribution over the opponent's private histories. In a belief-free strategy profile, a player's continuation strategy is his best reply to his opponent's continuation strategy at every private history. In other words, a sequential equilibrium is a belief-free strategy profile if it has the property that a player's continuation strategy is still the player's best reply when he secretly learns about his opponent's private history.

We will therefore speak directly of belief-free equilibria. We say that a strategy s is a symmetric Nash (belief-free) equilibrium if the profile (s,s) is a Nash (belief-free) equilibrium. One example for a symmetric belief-free equilibrium is the strategy s_D.

2.2 Evolutionary Stability

I present a refinement of a symmetric Nash equilibrium that requires robustness to a small group of agents who experiment with a different behavior (see, Weibull, 1995 for an introductory textbook).

Suppose that individuals in a large population are repeatedly drawn to play a symmetric two-person game, and that there is an underlying dynamic process of social learning in which more successful strategies (which induce higher average payoffs) become more frequent. Suppose that initially all individuals play the equilibrium strategy s^*. Now consider a small group of agents (called, *mutants*) who play a different strategy s'. If s' is not a best reply to s^*, then if the mutants are sufficiently rare they would be outperformed. If s' is a best reply to s^*, then the relative success of the incumbents and the mutants depends only on the average payoff they achieve when being matched against a mutant opponent. If the incumbents achieve a higher payoff when being matched against mutants, then the mutants are outperformed. Otherwise, the mutants outperform the incumbents, and their strategy gradually takes over the population.

The formal definition is as follows. Let $U(s,s')$ denote the expected discounted payoff to a player following strategy s and facing an opponent who plays strategy s' .

Definition 2 (Maynard Smith & Price, 1973; Maynard-Smith, 1982)**.** A symmetric Nash equilibrium s^* is neutrally (evolutionarily) stable if $U(s^*,s') \geq U(s',s')$ $(U(s^*,s') > U(s',s'))$ for each strategy $s^* \neq s' \in B(s^*)$.

Note that any strict symmetric equilibrium, such as the strategy s_d (always defecting), is evolutionarily stable.

Neutral stability is arguably a "too strong" refinement because: (1) some games do not admit any neutrally stable strategies, and (2) some equilibria that are not neutrally stable are plausible predictions of the time-average behavior in the game. This is demonstrated in the Rock-paper-scissors game in Table 2 (left side). The unique symmetric equilibrium is $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$, which is not neutrally stable (because $R \in B\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ and $U\left(\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),R\right) = -0.5 < U(R,R) = 0$. One can show that although $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ is not neutrally stable, then under mild assumptions on the dynamics, the time average of the aggregate play converges to $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ (Benaïm *et al.*, 2009).

This motivates me to present a weaker stability refinement that focuses only on vulnerability to mutant strategies that are evolutionarily stable. Strategy s^* is vulnerable to strategy s' if strategy s' can take over a population that initially plays s^*. Formally:

Table 2: Auxiliary Games to Demonstrate Notions of Stability

	R	P	S
R	0, 0	-2, 1	1, -2
P	1, -2	0, 0	-2, 1
S	-2, 1	1, -2	0, 0

Rock-Paper-Scissors.

	a	b
a	1, 1	0, 0
b	0, 0	1, 1

Coordination Game

Definition 3. Strategy s^* is *vulnerable* to strategy s' if $s' \in B(s^*)$ and $U(s', s') > U(s^*, s')$.

A symmetric Nash equilibrium is weakly stable if (1) it is not vulnerable to any evolutionarily stable strategy, and (2) it is not "two-step" vulnerable to an evolutionarily stable strategy, i.e., it is not vulnerable to any strategy that is vulnerable to an evolutionarily stable strategy. Formally:

Definition 4. A symmetric Nash equilibrium s^* is *weakly stable* if:

1. There does not exist s' such that s^* is vulnerable to s', and s' is evolutionarily stable; and
2. There do not exist $\tilde{s}$, s' such that s^* is vulnerable to $\tilde{s}$, strategy $\tilde{s}$ is vulnerable to s', and strategy s' is evolutionarily stable

Observe that the definition of weak stability satisfies the following appealing properties:

1. Any neutrally stable strategy is weakly stable.
2. Any game admits a weakly stable strategy.
3. If s^* is not weakly stable, then it is not a plausible prediction to the long run behavior in the population. Even if the population initially plays s^*, then as soon as a small group of agents experiments with playing s', the population diverges away towards $\tilde{s}$. If this is followed by another small group of agents who play s', then the population will converge to s', and will remain there in the long run.

3 Instability of Belief-Free Equilibria

Piccione (2002) and Ely & Välimäki (2002) present folk-theorem results for the repeated Prisoner's Dilemma using belief-free equilibria. In particular, they show that any payoff between zero and one is a symmetric belief-free equilibrium payoff if the monitoring technology is almost perfect and the players are sufficiently patient. Out main result shows that under the mild refinement of weak stability, the game admits a unique symmetric belief-free equilibrium in which everyone defects. In section 5.1 we show how to extend this result to asymmetric equilibria.

The sketch of the proof is as follows. As observed by Ely *et al.* (2005, Section 2.1), the set of optimal actions at each round in a belief-free equilibrium is independent of the private history. This implies that the belief-free equilibrium is vulnerable to a deterministic strategy that always defect except in rounds in which cooperation is the unique optimal action. This deterministic strategy is vulnerable to the evolutionarily stable strategy in which everyone defects.

Proposition 1. *Strategy s_d is the unique weakly stable belief-free symmetric equilibrium of the repeated Prisoner's Dilemma game.*

Proof. Let s^* be a symmetric belief-free equilibrium. A continuation strategy z_i is a *belief-free sequential best-reply* to s^* beginning from period t if

$$z_i|h_i^{\tilde{t}} \in B_i(s^*|h_{-i}^{\tilde{t}})\ \forall \tilde{t} \geq t \text{ and } h^{\tilde{t}} \in H^{\tilde{t}};$$

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

$$\mathcal{A}_i^t = \left\{a \in \{C, D\} \,|\exists z_i \in B_i^t(s),\ \exists h_i^t \text{ such that } z_i\left(h_i^t\right)(a_i) > 0\right\};$$

denote the set of actions in the support of some belief-free sequential best-reply beginning from period t. As noted by Ely *et al.* (2005, Section 2.1) $\exists h_i^t$ 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 $z_i|\tilde{h}_i^t$ for a given $\tilde{h}_i^t$, then so obtained strategy z_i is also a belief-free sequential best-reply to s_{-i}. Note that the symmetry of the profile (s^*, s^*) implies that $\mathcal{A}^t := \mathcal{A}_i^t = \mathcal{A}_j^t$.

We complete the proof by dealing with two (exclusive and exhaustive) cases:

1. There exists time t such that: (I) both actions are best replies ($\mathcal{A}^t = \{C, D\}$), and (II) there is an history after which one of the players cooperate with positive probability, i.e., $\exists h_i^t$ s.t. $s_i^*\left(h_i^t\right)(C) > 0$. Let $\tilde{s}$ be the deterministic strategy that plays D at all periods except in those in which C is the unique best-reply to s^*:

$$\tilde{s}\left(h_i^t\right) = \begin{cases} C & \mathcal{A}_i^t = \{C\} \\ D & \text{otherwise} \end{cases}.$$

   The definition of $\tilde{s}$ implies that $\tilde{s}$ is a best reply to s^*. The fact that $\tilde{s}$ induces a deterministic play that is independent of the observed signals implies that defection is the unique best-reply against $\tilde{s}$ at all rounds (and after any history), i.e., $\{s_D\} = B_i^t(\tilde{s})$. Observe that strategy $\tilde{s}$ defects with a strictly higher probability than s^*. This implies that $\tilde{s}$ yields a strictly larger payoff against itself relative to s^* $U(s^*, \tilde{s}) < U(\tilde{s}, \tilde{s})$. Moreover, if $\tilde{s} \neq s_D$, then strategy s_D is a best reply against $\tilde{s}$ and $U(\tilde{s}, s_D) < U(s_D, s_D)$. As s_d is evolutionarily stable, it implies that s^* is not weakly stable.

2. In any time t in which both actions are best replies, each player defects after each history, i.e., $s_i^*(h_i^t)(D) = 1\ \forall h_i^t$ s.t $\mathcal{A}^t = \{C, D\}$. Observe that in this case s^* induces a deterministic play that is independent of the observed signals. This implies that defection is the unique best-reply against s^* at all rounds (and after any history), i.e., $\{s_D\} = B_i^t(s^*)$, and we get a contradiction unless $s^* = s_D$.

□

Remark 1. Note that in all the belief-free equilibria presented in the existing literature the set of belief-free sequential best-replies always includes the action D (i.e., $D \in \mathcal{A}_i^t$ for each t), which implies that all these equilibria are directly vulnerable to s_d (i.e., $\tilde{s} = s_d$ in the above proof) and the two-step vulnerability is required only against hypothetical belief-free equilibria in which at some rounds cooperation is the unique best-reply.

4 Instability of Other Equilibria

The previous section shows that the main kind of equilibrium used in the literature on private imperfect monitoring (the belief-based equilibrium) does not satisfy weak stability. In this section we extend this result to the other kinds of equilibria in the literature: belief-free review-strategy equilibria, and belief-based equilibria.

4.1 Instability of Belief-Free Review-Strategy Equilibria.

Matsushima (2004); Hörner & Olszewski (2006); Yamamoto (2007, 2012) have used the notion of *belief-free review-strategy equilibrium* in which (1) the infinite horizon is regarded as a sequence of review phases such that each player chooses a constant action throughout a review phase, and (ii) at the beginning of each review phase, a player's continuation strategy is a best reply regardless of the history. Simple adaptation to the proof of Proposition 1 show that defection is the unique weakly-stable symmetric belief-free review-strategy equilibrium.

The adaptation of the proof is done as follows. Let s^* be a symmetric belief-free review-strategy equilibrium. Let $(t_l)_{l=1}^{\infty}$ be the increasing sequence of starting times for the review phases. Let

$\mathcal{A}_i^l \subseteq \{C, D\}$ denote the set of actions in the support of some continuation sequential best-reply that begins from period t_l and plays the same action throughout the l-th review phase. Let $\tilde{s}$ be the deterministic strategy that always defects except in review phases in which $\mathcal{A}_i^l = \{C\}$. An analogous argument to the proof of Proposition 1 shows that s^* is vulnerable to $\tilde{s}$ (and $\tilde{s}$ is vulnerable to s_d).

4.2 Instability of Belief-Based Equilibria

Bhaskar & Obara (2002) (extending Sekiguchi, 1997) present a folk-theorem result for the repeated Prisoner's Dilemma that does not rely on belief-free equilibria. Instead, the best-reply of each player depends on his belief about the private history of the opponent ("belief-based equilibria"). In what follows, I show that these equilibria are not weakly stable.

Bhaskar & Obara (2002) consider a symmetric signaling structure with two signals $\Sigma_i = \{c, d\}$, where c (resp., d) is more likely when the opponent plays C (resp. D). Given any action profile, there is a probability of $\epsilon > 0$ that exactly one player receives a wrong signal and a probability of $\xi > 0$ that both players receive wrong signals. Bhaskar & Obara present for each $0 < x < 1$ a symmetric sequential equilibrium s_x that yields a symmetric payoff of at least x whenever ϵ and ξ are sufficiently small. This construction is the key element in their folk-theorem result. In what follows we sketch this equilibrium s_x, and then show that it is not weakly stable.

Let s_T be the trigger strategy: cooperating as long as all observed signals are c-s, and defect in the remaining game if signal d is ever observed. The strategy s_x divides the entire game into disjoint subgames (say, to n subgames, where subgame k includes the rounds that are equal to k modulo n), and the play in each subgame is independent of the other subgames. Each player mixes at the first round of each subgame: he plays s_T (trigger strategy) with probability π and plays s_D (always defecting) with the remaining probability. Bhaskar & Obara show that there exists a division into subgames and a mixing probability π such that: (1) the expected discounted symmetric payoff of the game is at least x, (2) strategy s_x is a sequential equilibrium.[4]

Claim 1. The symmetric sequential equilibrium s_x is not weakly stable.

Sketch of proof. The fact that s_x mixes between s_D and s_T at the beginning of each subgame, implies that s_D is a best reply against s_x. Recall, that s_D is evolutionarily stable and the unique best reply to itself. These observations immediately imply that s_x is not weakly stable. □

[4]As observed by Bhaskar (2000) and Bhaskar *et al.* (2008), these belief-based equilibria can be purified à la Harsanyi in a simple way (while this is not the case for belief-free equilibria). Nevertheless, we show they still do not satisfy weak stability.

5 Extensions

To simplify the presentation of our results we focused only on symmetric equilibria (for which the definition of evolutionary stability is simpler), and that the monitoring has full support. In this section we show that both assumption can be relaxed.

5.1 Instability of Asymmetric Equilibria

In this section we show how to extend our model to deal with asymmetric monitoring structures and asymmetric equilibria.

Adaptations to the model We relax the assumption that the monitoring structure is symmetric. That is, the alphabet of each player may be different (i.e. $\Sigma_1 \neq \Sigma_2$), and the distribution of signals $m(\cdot|a)$ is not necessarily symmetric. Following Selten (1980), we consider a single population of agents that are randomly matched into pairs to play the repeated Prisoner Dilemma (possibly with asymmetric monitoring structure), and the agents in each pair are randomly assigned into the two roles: player 1, and player 2.

A complete-strategy of an agent $\sigma = (s_1, s_2)$ is a pair describing his strategy in each of the two roles (i.e., for each $i \in \{1,2\}$ $s_i : H_i \to \Delta(A_i)$ is the strategy he plays when he is player i). Any equilibrium of the underlying game (s_1, s_2) corresponds to a symmetric equilibrium in the population in which all agents play the complete-strategy $\sigma_{s_1,s_2} := (s_1, s_2)$. Finally, we say that the Nash equilibrium (s_1, s_2) is weakly stable if the complete-strategy σ_{s_1,s_2} is weakly stable.

Adaptations to the proofs Proposition 1 also holds for asymmetric equilibria. Let (s_1^*, s_2^*) be a belief-free equilibrium, and let $\sigma_{s_1^*,s_2^*}$ be the corresponding complete strategy. The proof should be adapted as follows. The sets of best-reply actions at each round, $\mathcal{A}_i^t$ and $\mathcal{A}_j^t$, remain the same (though, they are not necessarily equal sets in this case). The strategy $\tilde{s}_1$ $(\tilde{s}_2)$ is defined with respect to $\mathcal{A}_2^t$ $(\mathcal{A}_1^t)$ in an analogous way to the definition of $\tilde{s}$ in the original proof. The definition implies that the complete strategy $\tilde{\sigma} = (\tilde{s}_1, \tilde{s}_2)$ is a best reply to $\sigma_{s_1^*,s_2^*}$, that $U\left(\sigma_{s_1^*,s_2^*}, \tilde{\sigma}\right) < U(\tilde{\sigma}, \tilde{\sigma})$, and that $\sigma_d = (s_d, s_d)$ is the unique best reply against $\tilde{\sigma}$, which implies that $\sigma_{s_1^*,s_2^*}$ is not weakly stable (unless it coincides with σ_d) using the same arguments as in the original proof.

Claim 1 can be adapted in a similar manner to show that also all the asymmetric sequential equilibria in Bhaskar & Obara (2002) and Chen (2010) are not weakly stable.

5.2 Instability without Full-Support Monitoring

In this section we show how to extend our model and results to deal with monitoring structures without full support.

We slightly refine Definition 4 of weak stability by allowing the strategy s' to be neutrally stable, rather than evolutionary stable (because if the monitoring structure does not have full support, then no strategy is evolutionarily stable, and, in particular, the strategy s_D is only neutrally stable). Claim 1 holds with this new definition with minor adaptations to the proof that are implied by the adapted definition of weak stability.

Proposition 1 should be slightly weakened in this setup, to state that any weakly stable belief-free equilibrium is realization equivalent to s_D (i.e., the players always defect on the equilibrium path). Case 1 of the proof is adapted by assuming that there an history of length t that is reached with positive probability in which the players cooperate with positive probability. Similarly, case 2 is adapted to assume that in any time in which both actions are best replies, the players defect after each history that is reached with positive probability. All the other arguments in the proof remain essentially the same.

References

Abreu, Dilip, Pearce, David, & Stacchetti, Ennio. 1990. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 1041–1063.

Benaïm, Michel, Hofbauer, Josef, & Hopkins, Ed. 2009. Learning in games with unstable equilibria. *Journal of Economic Theory*, **144**(4), 1694–1709.

Bhaskar, V. 2000. *The robustness of repeated game equilibria to incomplete payoff information*. Unversity of Essex.

Bhaskar, V, & Obara, Ichiro. 2002. Belief-based equilibria in the repeated prisoners' dilemma with private monitoring. *Journal of Economic Theory*, **102**(1), 40–69.

Bhaskar, V, Mailath, George J, & Morris, Stephen. 2008. Purification in the infinitely-repeated prisoners' dilemma. *Review of Economic Dynamics*, **11**(3), 515–528.

Compte, Olivier. 1998. Communication in repeated games with imperfect private monitoring. *Econometrica*, 597–626.

Ely, Jeffrey C, & Välimäki, Juuso. 2002. A robust folk theorem for the prisoner's dilemma. *Journal of Economic Theory*, **102**(1), 84–105.

Ely, Jeffrey C, Hörner, Johannes, & Olszewski, Wojciech. 2005. Belief-free equilibria in repeated games. *Econometrica*, **73**(2), 377–415.

Fudenberg, Drew, & Maskin, Eric. 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 533–554.

Heller, Yuval, & Mohlin, Erik. 2015. *Stable Observable Behavior*. Unpublished.

Hörner, Johannes, & Olszewski, Wojciech. 2006. The Folk Theorem for Games with Private Almost-Perfect Monitoring. *Econometrica*, **74**(6), 1499–1544.

Kandori, Michihiro. 2002. Introduction to repeated games with private monitoring. *Journal of Economic Theory*, **102**(1), 1–15.

Kandori, Michihiro. 2011. Weakly Belief-Free Equilibria in Repeated Games With Private Monitoring. *Econometrica*, **79**(3), 877–892.

Kandori, Michihiro, & Matsushima, Hitoshi. 1998. Private observation, communication and collusion. *Econometrica*, 627–652.

Kandori, Michihiro, & Obara, Ichiro. 2006. Efficiency in repeated games revisited: The role of private strategies. *Econometrica*, **74**(2), 499–519.

Mailath, George J, & Morris, Stephen. 2002. Repeated games with almost-public monitoring. *Journal of Economic Theory*, **102**(1), 189–228.

Mailath, George J, & Morris, Stephen. 2006. Coordination failure in repeated games with almost-public monitoring. *Theoretical Economics*, **1**(3), 311–340.

Mailath, George J, & Samuelson, Larry. 2006. *Repeated games and reputations*. Vol. 2. Oxford university press Oxford.

Matsushima, Hitoshi. 2004. Repeated games with private monitoring: Two players. *Econometrica*, **72**(3), 823–852.

Maynard-Smith, John. 1982. *Evolution and the theory of games*. Cambridge University Press.

Maynard Smith, John, & Price, George R. 1973. The logic of animal conflict. *Nature*, **246**, 15.

Miyagawa, Eiichi, Miyahara, Yasuyuki, & Sekiguchi, Tadashi. 2008. The folk theorem for repeated games with observation costs. *Journal of Economic Theory*, **139**(1), 192–221.

Obara, Ichiro. 2009. Folk theorem with communication. *Journal of Economic Theory*, **144**(1), 120–134.

Piccione, Michele. 2002. The repeated prisoner's dilemma with imperfect private monitoring. *Journal of Economic Theory*, **102**(1), 70–83.

Sekiguchi, Tadashi. 1997. Efficiency in repeated prisoner's dilemma with private monitoring. *Journal of Economic Theory*, **76**(2), 345–361.

Sugaya, Takuo, & Takahashi, Satoru. 2013. Coordination failure in repeated games with private monitoring. *Journal of Economic Theory*, **148**(5), 1891–1928.

Takahashi, Satoru. 2010. Community enforcement when players observe partners' past play. *Journal of Economic Theory*, **145**(1), 42–62.

Weibull, Jörgen W. 1995. *Evolutionary game theory*. The MIT press.

Yamamoto, Yuichi. 2007. Efficiency results in N player games with imperfect private monitoring. *Journal of Economic Theory*, **135**(1), 382–413.

Yamamoto, Yuichi. 2009. A limit characterization of belief-free equilibrium payoffs in repeated games. *Journal of Economic Theory*, **144**(2), 802–824.

Yamamoto, Yuichi. 2012. Characterizing belief-free review-strategy equilibrium payoffs under conditional independence. *Journal of Economic Theory*, **147**(5), 1998–2027.

Yamamoto, Yuichi. 2013. Individual learning and cooperation in noisy repeated games. *The Review of Economic Studies*, **81**(1), 473–500.