Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 168 (2017) 261–286

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Instability of belief-free equilibria

Yuval Heller [1]

Department of Economics, Bar Ilan University, Israel

Received 8 August 2015; final version received 17 December 2016; accepted 3 January 2017

Available online 6 January 2017

Abstract

Various papers have presented folk theorem results for repeated games with private monitoring that rely on belief-free equilibria. I show that these equilibria are not robust against small perturbations in the behavior of potential opponents. Specifically, I show that essentially none of the belief-free equilibria is evolutionarily stable, and that in generic games none of these equilibria is neutrally stable. Moreover, in a large family of games (which includes many public good games), the belief-free equilibria fail to satisfy even a very mild stability refinement.

JEL classification: C73; D82

Keywords: Belief-free equilibrium; Evolutionary stability; 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 and Maskin, 1986; Fudenberg et al., 1994) have established that effi-

E-mail address: yuval.heller@biu.ac.il.

[1] A previous version of this manuscript was entitled "Instability of Equilibria with Private Monitoring." I would like to express my deep gratitude to Mehmet Ekmekci, Peter Eso, Michihiro Kandori, Erik Mohlin, Thomas Norman, Ron Peretz, Satoru Takahashi, Jorgen Weibull, Yuichi Yamamoto, the associate editor, and the referees, for many helpful comments. I am grateful to the European Research Council for its financial support (starting grant #677057).

http://dx.doi.org/10.1016/j.jet.2017.01.001

0022-0531/

ciency can be achieved under fairly general conditions when players observe commonly shared information about past action profiles.

In many real-life situations players privately observe imperfect signals about past actions. For example, each firm in a cartel privately observes its own sales, which contain imperfect information about secret price cuts that its competitors offer to some of their customers. Formal analysis of private monitoring began with the pioneering work of Sekiguchi (1997). Since then, several papers have presented various folk theorem results that have shown that efficiency can be achieved also with private monitoring (see Kandori, 2002; Mailath and Samuelson, 2006, for surveys of this literature).

The most commonly used equilibrium in the literature on 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 to compute a best reply. Piccione (2002) and Ely and Välimäki (2002) present folk theorem results for the repeated Prisoner's Dilemma using belief-free equilibria under the assumptions that the monitoring technology is almost perfect and the players are sufficiently patient. Ely et al. (2005), Miyagawa et al. (2008), and Yamamoto (2009, 2014) extend the folk theorem results that rely on belief-free equilibria to general repeated games and to costly observability. Kandori and Obara (2006) study a setup of imperfect *public* monitoring and show that belief-free private strategies can improve efficiency relative to the maximal efficiency obtained by public strategies. Takahashi (2010) applies the belief-free equilibria to obtain folk theorem results for repeated games in which the players are randomly matched with a new opponent in each round.

The results of the present paper show that belief-free equilibria are not robust against small perturbations in the behavior of potential opponents, and that this instability is extreme in a family of games that include many public good games, the Prisoner's Dilemma, and coordination games.

Instability of belief-free equilibria. One of the leading justifications for using a Nash equilibrium to predict behavior is its interpretation as being 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 different strategy s'. If this induces the experimenting agents to gain more than the incumbents, then the population will move away from s^* toward s'. Thus, strategy s^* is *evolutionarily (neutrally) stable* (Maynard-Smith and 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 (weakly) higher payoff against any other best-reply strategy s': $U\left(s^*, s'\right) > U\left(s', s'\right)$. For example, the strategy of always playing a strict symmetric equilibrium of the one-shot game regardless of the history is neutrally stable, and, moreover, it is evolutionarily stable if the signal distribution has full support.

A belief-free equilibrium is *trivial* if it induces the play of a Nash equilibrium in all periods. My first result (Proposition 1) shows that only trivial belief-free equilibria may satisfy evolutionary stability. My second result (Proposition 2) makes two mild assumptions on the environment: (1) the underlying game is generic, and (2) the signal a player observes in each round is not

[2] To simplify the exposition I focus in the body of the paper on symmetric equilibria in symmetric games, and I extend the analysis to general equilibria and asymmetric games in the appendix.

completely uninformative about the partner's action. Under these mild assumptions, I show that only trivial belief-free equilibria may satisfy neutral stability.

The intuition of these results is as follows. As observed by Ely et al. (2005, Section 2.1), in each period t the set of optimal actions in a belief-free equilibrium is independent of the private history. This implies that mutants who play a symmetric Nash equilibrium in an auxiliary game in which players are allowed to choose only from the set of optimal actions weakly outperform the incumbents. Moreover, if the signal of each player contains some information about the partner's action, the players can use the actions each of them played and the private signals that each of them observed in some period in the past, to induce a correlation between their mixed actions in a later period. In a generic game, inducing either a negative or a positive correlation in the mixed action profile of the later round allows the mutants to strictly outperform the incumbents.

Refinement of weak stability. The existing notions of stability, namely, evolutionary and neutral stability, are arguably too-strong refinements, as demonstrated in the rock–paper–scissors game (see Section 2.3) that admits a unique Nash equilibrium that is not neutrally stable, but that is a plausible prediction of the long-run average behavior in the population (see, e.g., Benaïm et al., 2009). Motivated by this, I present a novel, and very mild, notion of stability. I say that a strategy s is *vulnerable* to strategy s' if agents who follow strategy s' achieve a strictly higher payoff in any heterogeneous population in which some agents follow strategy s and some follow strategy s'. The definition implies that a small group of mutants who play strategy s' will take over a population that initially plays strategy s. I say that a symmetric Nash equilibrium s^* is *weakly stable* if there does not exist a finite sequence of strategies $(s_1, ..., s_K)$, such that: (1) strategy s^* is vulnerable to s_1, (2) each strategy s_k is vulnerable to s_{k+1}, and (3) strategy s_K is evolutionarily stable.[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 of long-run behavior. This is because as soon as a small group of agents experiments with playing s_1, the population diverges to s_1. If this is followed by an invasion of a small group of agents who play s_2, then the population diverges to s_2, and after a finite number of such sequential invasions, the population diverges to s_K, and it will remain in s_K in the long run (due to s_K being evolutionarily stable).[4] A simple example of a non-weakly stable equilibrium is a mixed equilibrium in a coordination game, for which every small perturbation takes the population to one of the pure equilibria.

Weak stability of belief-free equilibria. I say that a symmetric game is *recursively strict*, if, for any subset of actions, the game in which each player is restricted to choosing an action from the subset admits a strict symmetric equilibrium. Examples of this family of games include the Prisoner's Dilemma, the Traveler's Dilemma, symmetric coordination games, and many public good games. My next result (Proposition 3) focuses on this family of games, and shows that only trivial belief-free equilibria satisfy the mild refinement of weak stability. The intuition for the Prisoner's Dilemma is that any belief-free equilibrium is vulnerable to a deterministic strategy s' in which the players defect in each period in which defection is an optimal action with respect to

[3] Remark 6 discusses the relation between weak stability and the structurally similar notion of "robustness against indirect invasions" of Van Veelen (2012).

[4] I assume that these experimentations are infrequent enough that strategies that are outperformed following the entry of a group of experimenting agents become sufficiently rare before a new group of agents starts experimenting with a different behavior.

the belief-free equilibrium, and this strategy s' is vulnerable to the evolutionarily stable strategy of always defecting. Remark 3 sketches how to extend this result to the larger set of belief-free review-strategy equilibria (Matsushima, 2004; Yamamoto, 2007; Deb, 2012; Yamamoto, 2012).

The Hawk–Dove game, which is a common application of belief-free equilibria, does not admit a strict symmetric equilibrium, and thus the results so far only show that non-trivial belief-free equilibria are not neutrally stable. The main difficulty in analyzing weak stability in Hawk–Dove games is that, in general, it is an open question whether a repeated game with private monitoring admits an evolutionarily stable strategy when the underlying game does not admit a strict symmetric equilibrium. My next result (Proposition 4) shows that a belief-free equilibrium in the repeated Hawk–Dove game is weakly stable if and only if the monitoring structure is such that the repeated game does not admit evolutionarily stable strategies. The "only if" side of the result shows that if an evolutionarily stable strategy exists, then there must be a sequence of strategies, each of which is vulnerable to its successor, that starts with the belief-free equilibrium and ends in an evolutionarily stable strategy. The "if" side of this result is trivial: if the repeated game does not admit any evolutionarily stable strategy, then there cannot be a sequence of strategies ending in an evolutionarily stable strategy, and, as a result, any Nash equilibrium is weakly stable.

An important alternative approach to belief-free equilibria in the literature on private monitoring is the "belief-based" equilibrium. Bhaskar and Obara (2002) define these equilibria and apply them to the repeated Prisoner's Dilemma. My final result (Claim 1) shows that the particular "belief-based" equilibria that are presented in Bhaskar and Obara (2002) do not satisfy weak stability.

1.1. Related literature and contribution

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 and Morris (2002, 2006), Hörner and Olszewski (2009), and Mailath and Olszewski (2011) show that when the private signals are almost perfectly correlated conditional on the action profile (i.e., when there is *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-free equilibria, in which the strategy of each player is a best reply to any private history of the opponent up to the actions of the previous round. Unlike standard belief-free equilibria, players need to form the correct beliefs about the signal obtained by the opponent in the previous round. Kandori (2011) demonstrates that if there is sufficient correlation between private signals (conditional on the action profile), then the game admits a strict, weakly belief-based equilibrium that yields substantial cooperation. The strictness of the equilibrium implies that it satisfies the refinement of evolutionary stability. In the discussion paper version of his paper Kandori (2009) points out that the specific non-trivial belief-free equilibria of Ely and Välimäki (2002) do not satisfy evolutionary stability in the repeated Prisoner's Dilemma. The present paper substantially strengthens Kandori's observation in at least two important ways: (1) I show that any non-trivial belief-free equilibrium of any underlying game is not evolutionarily stable, and, moreover, it is not neutrally stable under the mild assumptions that the game is generic and the monitoring structure has a grain of informativeness, and (2) I show that in the large family of recursively strict games, any non-trivial belief-free equilibrium fails to satisfy the very mild refinement of weak stability.

Communication and conditionally independent signals. Compte (1998), Kandori and Matsushima (1998), and Obara (2009) present 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 obtain. These equilibria are constructed such that the players have strict incentives while playing, and such that they are always 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 hence also weakly stable.[5]

The present paper shows that all the mechanisms in the existing literature can yield only 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 open question whether any new mechanism may yield cooperation as a stable outcome with conditionally independent private monitoring. This open question has interesting implications for antitrust laws. If the answer to this question is negative, then it would suggest that communication between players is critical to obtaining collusive behavior whenever the private imperfect monitoring between the firms is such that the conditional correlation between the private signals is sufficiently low.[6,7]

One promising direction toward the solution of this open question might rely on the methods developed in Heller and Mohlin (2016) for the related setup of random matching and partial observation of the partner's past behavior. In that setup, Heller and Mohlin (2016) characterize conditions under which only defection is stable, and construct novel mechanisms for sustaining stable cooperative equilibria whenever these conditions are not satisfied.

Robustness. Sugaya and Takahashi (2013) show that "generically" only belief-free equilibria are robust against small perturbations in the monitoring structure. Our main result shows that belief-free equilibria (except for defection) are not robust against 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 against both kinds of perturbations.[8]

---

5 The argument for neutral stability is sketched as follows. Having strict incentives while playing implies that any best-reply strategy induces the same play on the equilibrium path, and differs from the incumbent strategy only by sending false reports. The fact that players are always indifferent between reporting the truth and lying implies that any such best-reply strategy yields the same payoff as the incumbent strategy (both when the opponent is an incumbent as well as when he is a mutant who follows a best-reply strategy).

6 This empirical prediction can be tested experimentally by comparing how subjects play the repeated Prisoner's Dilemma with private monitoring and conditionally independent signals with and without the ability to communicate by exchanging "cheap talk" messages. Matsushima et al. (2013) experimentally study this setup without communication, and their findings suggest that the subjects' behavior is substantially different from the predictions of the belief-free equilibria (in particular, subjects retaliate more severely when monitoring is more accurate). I am not aware of any experiment that studies this setup with communication.

7 See also the recent related result of Awaya and Krishna (2016), which deals with sequential equilibria of oligopolies under some plausible private monitoring structures, and shows that cheap talk communication allows one to achieve a higher level of collusion relative to the maximal level that one can achieve without communication.

8 Two existing papers present related anti–folk theorem results. Matsushima (1991) shows that defection is the unique *pure* equilibrium in the repeated Prisoner's Dilemma in which signals are conditionally independent and Nash equilibria are restricted to being independent of payoff-irrelevant private histories. As demonstrated by the "belief-based" equilibria

Structure. The model is described in Section 2. Section 3 presents the results for symmetric games. The appendix extends the analysis to asymmetric games.

2. Model

2.1. Games with private monitoring

I analyze a two-player δ-discounted repeated game with private monitoring. I use the index $i \in \{1, 2\}$ to refer to one of the players, and $-i$ to refer to the opponent. Each player i has a finite action set A_i and a finite set of signals Σ_i. An action profile is an element of $A_1 \times A_2$. I use ΔW to represent the set of probability distributions over a finite set W. Let ΔA_i and $\Delta A_1 \times \Delta A_2$ represent respectively the set of mixed actions for player i and mixed action profiles. For each player i let $u_i : A_1 \times A_2 \to \mathbb{R}$ denote the payoff function, which is extended to mixed actions in the standard (linear) way.

For each possible action profile $(a_1, a_2) \in A_1 \times A_2$, the monitoring distribution $m(\cdot|a_1, a_2)$ specifies a joint probability distribution over the set of signal profiles $\Sigma_1 \times \Sigma_2$. When action profile a is played and signal profile (σ_1, σ_2) is realized, each player i privately observes his corresponding signal σ_i. Let $m_i(\cdot|a_1, a_2)$ be the marginal probability distribution over the signal of player i: $m_i(\sigma_i|a_1, a_2) = \sum_{\sigma_{-i} \in \Sigma_{-i}} m(\sigma_i, \sigma_{-i}|a_1, a_2)$. Letting $\tilde{u_i}(a_i, \sigma_i)$ denote the payoff to player i from action a_i and signal σ_i, I can represent stage payoffs as a function of mixed action profiles only:

$$u_i(\alpha_1, \alpha_2) = \sum_{(a_1, a_2) \in A_1 \times A_2} \sum_{\sigma_i \in \Sigma_i} \alpha_1(a_1) \cdot \alpha_2(a_2) \cdot m_i(\sigma_i|a_1, a_2) \cdot \tilde{u}(a_i, \sigma_i).$$

To simplify the presentation of the results, I assume that the marginal distribution of signals of each player has a full support, i.e., that each signal is observed with a positive probability after each action profile. Formally[9]:

Assumption 1. *The monitoring structure has* full support: $m_i(\sigma_i|a_1, a_2) > 0$ *for each action profile* $(a_1, a_2) \in A_1 \times A_2$*, each player* i*, and each signal* $\sigma_i \in \Sigma_i$*.*

One example of a monitoring structure with full support is the *conditionally independent ϵ-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.

of Bhaskar and Obara (2002), the uniqueness result does not hold for *mixed* equilibria (the mixed "belief-based" equilibria achieve cooperation even though the behavior of the players is independent of payoff-irrelevant private histories, and signals may be conditionally independent). Peski (2012) studies repeated games with private monitoring. He assumes that strategies have a finite past, in each period players' preferences over actions are modified by smooth idiosyncratic shocks, the monitoring structures includes infinitely many signals, and the signals are sufficiently connected. Under these assumptions, Peski (2012) shows that all equilibria of the repeated game are trivial, in the sense that each period's play is an equilibrium of the stage game.

[9] The results can be adapted to a setup in which the monitoring structure does not have full support. The adaptation requires changing two definitions (and related minor adaptations to the proofs): (1) extending the set of trivial belief-free equilibria in Definition 2, such that it relates only to histories that occur with positive probability, and (2) refining Definition 6 of weak stability by allowing the strategy s' to be neutrally stable, rather than evolutionarily stable (because if the monitoring structure does not have full support, then no strategy is evolutionarily stable).

A t-length private history of player i (abbr., history) 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. Each player's initial history is the null history, denoted by ϕ. Let $H_i^t := (A_i \times \Sigma_i)^t$ denote the set of all t-length histories of player i, and let $H_i = \cup_t H_i^t$ the set of all histories of player i. A history profile, $\left(h_1^t, h_2^t\right) \in H_1^t \times H_2^t$, is a pair of t-length histories, one belonging to each player.

2.2. Belief-free equilibria

A repeated-game (behavior) strategy of player i is a mapping $s_i : H_i \to \Delta(A_i)$. Let S_i denote the set of all strategies of player i. For history h_i^t, let $s_i|_{h_i^t}$ denote the continuation strategy derived from s_i 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 $s_i|_{h_i^t}$ is the strategy defined by $s_i|_{h_i^t}\left(\hat{h}_i\right) = s_i\left(h_i \hat{h}_i\right)$. Given a strategy profile $\overrightarrow{s} = (s_1, s_2)$, let $B_i\left(\overrightarrow{s}|h_{-i}^t\right)$ denote the set of continuation strategies of i that are best replies to $s_{-i}|_{h_{-i}^t}$.

Definition 1 *(Ely et al., 2005)*. A strategy profile $\overrightarrow{s}^* = \left(s_1^*, s_2^*\right)$ is *belief-free* if for every history profile $\left(h_1^t, h_2^t\right)$, $s_i^*|_{h_i^t} \in B_i\left(\overrightarrow{s}^*|h_{-i}^t\right)$ for $i \in \{1, 2\}$.

The condition characterizing a belief-free strategy profile is stronger than that characterizing a 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.

A simple kind of a belief-free equilibrium, is a strategy profile in which the players play a Nash equilibrium of the underlying game in all periods, and this equilibrium is independent of the history of play. I call such belief-free equilibria trivial. Formally, let $NE((A_1, A_2), (u_1, u_2))$ denote the set of Nash equilibria of the underlying game. Let $\pi_{s_i, s_{-i}}^t \in \Delta\left(H_i^t\right)$ denote the probability that a player who follows strategy s_i observes history h_i^t, conditional on the opponent following strategy s_{-i} (and the monitoring structure m).

I say that history $h_i^t \in H_i^t$ is *feasible* given strategy s_i if there exists strategy s_{-i} such that $\pi_{s_i, s_{-i}}^t\left(h_i^t\right) > 0$. For example, if s_{a^*} is the strategy that induces player i to always play action a^* regardless of the history, then a history of player i is feasible iff all the actions of player i in the previous rounds have been a^*. I say that history profile $\left(h_1^t, h_2^t\right)$ is *feasible* given strategy profile (s_1, s_2) if each h_i^t is feasible given strategy s_i.

Definition 2. A belief-free equilibrium (s_1, s_2) is *trivial* if for every two feasible history profiles $\left(h_1^t, h_2^t\right)$, $\left(\tilde{h}_1^t, \tilde{h}_2^t\right)$ of length t:

$$\left(s_i\left(h_i^t\right), s_{-i}\left(h_{-i}^t\right)\right) = \left(s_i\left(\tilde{h}_i^t\right), s_{-i}\left(\tilde{h}_{-i}^t\right)\right) \in NE\left((A_1, A_2), (u_1, u_2)\right).$$

A trivial equilibrium is pure if the Nash equilibrium played in each round is pure (i.e., $\left|supp\left(s_i\left(h_i^t\right)\right)\right| = 1$ for each player i, period t, and feasible history profile $\left(h_1^t, h_2^t\right)$).

2.3. Evolutionary stability in symmetric games

In what follows, I study evolutionary stability in symmetric games. I focus on symmetric games because they are the most popular setup in the evolutionary game theory literature. Appendix A extends the analysis to asymmetric games.

In the setup of symmetric games I omit the index i (e.g., $A := A_i$, $u := u_i$, $m := m_i$, $H^t := H_i^t$, and $h^t := h_i^t$). I say that a strategy s is a symmetric Nash (belief-free) equilibrium if the symmetric strategy profile (s, s) is a Nash (belief-free) equilibrium.

I present a refinement of a symmetric Nash equilibrium that requires robustness against 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 (technically, a continuum) are repeatedly drawn to play a two-person symmetric 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 will 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 matched against a mutant opponent. If the incumbents achieve a higher payoff when matched against the mutants, then the mutants are outperformed. Otherwise, the mutants outperform the incumbents, and their strategy gradually takes over the population.

The formal definitions are as follows. I say that two strategies are outcome-equivalent if they always induce the same behavior regardless of the opponent's strategy. Arguably, two outcome-equivalent strategies should be considered as two different ways to represent of the same strategy.

Definition 3. Strategies s, s' are *outcome-equivalent* if: (1) their sets of feasible histories coincide (i.e., h^t is feasible given s iff it is feasible given s'), and (2) they coincide after each feasible history (i.e., $s\left(h^t\right) = s'\left(h^t\right)$ for each feasible history h^t). Given a strategy s, let $[s]$ denote its equivalent set (i.e., the set of strategies that are outcome-equivalent to s).[10]

Remark 1. Observe that:

1. In a game in which each player acts once, any equivalence set is a singleton.
2. In infinitely repeated games the equivalence set $[s]$ is a singleton iff strategy s is totally mixed (i.e., it assigns a positive probability to each action after each history).
3. Let s_a be the strategy that plays action a after any history. The equivalence set $[s_a]$ is the set of strategies that induce a player (Alice) to play action a in the first round, and after any history in which Alice has always played a.

Let $U\left(s, s'\right)$ denote the expected discounted payoff to a player following strategy s and facing an opponent who plays strategy s'.

[10] The equivalence set $[s]$ is the set of all strategies that have the same reduced strategy à la Osborne and Rubinstein (1994, p. 94).

Table 1
Examples of symmetric games.

	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	$0,\ 0$	$1,\ 1$

2×2 coordination game

	c	d
c	$1,\ 1$	$l,\ 1+g$
d	$1+g,\ l$	$0,\ 0$

Prisoner's dilemma ($g > 0 > l$)
Hawk–Dove game ($g, l > 0$)

Definition 4 *(Maynard-Smith and Price, 1973; Maynard-Smith, 1982).* A symmetric Nash equilibrium s^* is neutrally (evolutionarily) stable if $U\left(s^*, s'\right) \geq U\left(s', s'\right)$ $\left(U\left(s^*, s'\right) > U\left(s', s'\right)\right)$ for each strategy $s' \in B\left(s^*\right) \setminus \left[s^*\right]$.

Remark 2. It is more common in the evolutionary game theory literature to define an evolutionarily stable strategy as a strategy that satisfies the above inequality for any $s' \in B\left(s^*\right) \setminus \{s^*\}$. Both definitions coincide when dealing with one-shot games. This alternative definition is arguably too strict when dealing with repeated games, as it can never be satisfied unless the strategy is totally mixed. Observe that strategy s is an evolutionarily stable strategy (according to Definition 4) iff its equivalence set $[s]$ is an evolutionarily stable set à la Thomas (1985), which implies that such a strategy is asymptotically stable in the standard replicator dynamics.

Example 1. Consider an underlying game $G = (A, u)$ and a subset of actions $A' \subseteq A$ that satisfy that $\left(a', a'\right)$ is a strict equilibrium for each $a' \in A'$. Let $\left(a^t\right)_t$ be an arbitrary sequence of actions in A' (i.e., $a^t \in A'$ for each period t). Observe that the pure strategy that plays action a^t in each period t is neutrally stable for any monitoring structure, and it is evolutionarily stable if the monitoring structure has full support.

The key difference between evolutionary stability and neutral stability is whether the mutants are allowed to obtain the same payoff as incumbents in the post-entry population. As a result neutrally stable strategies (which are not evolutionarily stable) may be vulnerable to a random drift of the population away from the initial state. The existing literature typically uses evolutionary stability as a strong refinement of stability, and neutral stability as a mild refinement.

2.4. *Weak stability in symmetric games*

One may argue that neutral stability is still "too strong" a 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 1 (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 $U\left(\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right), R\right) = -\frac{1}{3} < U(R, R) = 0$.) One can show that although $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is not neutrally stable, still, 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 much weaker stability refinement. Strategy s^* is vulnerable to strategy s' if strategy s' achieves a weakly better payoff against both s^* and s', and a strictly better payoff against one of these strategies. Formally:

Definition 5. Strategy s^* is *vulnerable* to strategy s' if $U\left(s', s^*\right) \geq U\left(s^*, s^*\right)$, $U\left(s', s'\right) \geq U\left(s^*, s'\right)$, and at least one of these inequalities is strict.

Definition 5 is equivalent to requiring that for any $0 < \beta < 1$ and any heterogeneous population in which β of the agents follow strategy s' and $1 - \beta$ of the agents follow strategy s^*, the agents following strategy s' achieve a strictly higher payoff. The definition implies that ϵ mutants who follow strategy s' will take over a population that initially plays s^* under any dynamic process in which more successful strategies become more frequent. Observe that a neutrally stable strategy is not vulnerable to any other strategy.

A symmetric Nash equilibrium s^* is weakly stable if there does not exist a finite sequence of strategies that starts at s^*, that ends in an evolutionarily stable strategy, and each of whose strategies is vulnerable to its successor. Formally:

Definition 6. A symmetric Nash equilibrium s^* is *weakly stable* if there does not exist a finite non-empty sequence of strategies $\left(s^1, ..., s^K\right)$ such that: (1) strategy s^* is vulnerable to s^1, (2) for each $1 \leq k < K$ strategy s^k is vulnerable to s^{k+1}, and (3) strategy s^K is evolutionarily stable.

I conclude this section with a few observations on Definition 6:

1. Any neutrally stable strategy is weakly stable.
2. Any game admits a weakly stable strategy.
3. The notion of weak stability is able to strictly refine Nash equilibrium only if the game admits an evolutionarily stable strategy.
4. If strategy s^* is not weakly stable, then it is not a plausible prediction of long-run behavior in the population. Even if the population initially plays s^*, as soon as a small group of agents experiments with playing $\tilde{s}$, the population will diverge to $\tilde{s}$. 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. Note that our argument relies on the assumption that these experimentations are infrequent enough that strategies that are outperformed following the entry of a group of experimenting agents become sufficiently rare before a new group of agents starts experimenting with a different behavior.
5. Definition 6 allows vulnerability to an evolutionarily stable strategy through an arbitrary number of sequential invasions (denoted by K). As shown in the proof of our main result on weak stability (Proposition 3), the maximal number of required invasions is $K \leq |A|$. Moreover, if we focus on the existing belief-free equilibria for the repeated Prisoner's Dilemma in the literature (e.g., Ely and Välimäki, 2002; Piccione, 2002), then most of them are seen to be directly vulnerable to an invasion by players who always defect (i.e., $K = 1$).
6. Definition 6 is structurally similar to Van Veelen's (2012) notion of *robustness against indirect invasions*. A strategy s^* is robust against indirect invasions if there does not exist a sequence of strategies $(s_1, ..., s_n)$, such that s^* is *weakly* vulnerable to s_1 (i.e., $s_1 \in B\left(s^*\right)$ and $U\left(s^*, s_1\right) \leq U\left(s_1, s_1\right)$), each s_k is weakly vulnerable to s_{k+1}, and s_{K-1} is (strictly) vulnerable to s_K. Note that Van Veelen's notion of robustness refines neutral stability (i.e., it

is between evolutionary stability and neutral stability), while weak stability weakens neutral stability (i.e., weak stability is between neutral stability and a symmetric Nash equilibrium).

3. Results

Ely et al. (2005) characterize the set of belief-free equilibrium payoffs, and show that such strategies support a large set of payoffs. In what follows, I show that only trivial belief-free equilibria may satisfy: (1) evolutionary stability in all games, (2) neutral stability in generic games, and (3) weak stability in the large family of recursively strict games. Next, I strengthen the instability result for Hawk–Dove games, and I sketch why belief-based equilibria (à la Bhaskar and Obara, 2002) do not satisfy weak stability.

3.1. Evolutionary stability in all games

My first result shows that any evolutionarily stable belief-free equilibrium must be trivial. The sketch of the proof is as follows. Ely et al. (2005, Section 2.1) show that the set of optimal actions in each period t is independent of the history. This implies that mutants who play a symmetric Nash equilibrium in an auxiliary game in which players are only allowed to choose from the set of optimal actions weakly outperform the incumbents. If the belief-free equilibrium is non-trivial, then the mutants' play differs from the incumbents' play, which implies that the belief-free equilibrium is not evolutionarily stable.

Proposition 1. *Let s^* be a symmetric belief-free equilibrium that is also evolutionarily stable. Then s^* is trivial.*

Proof. A continuation strategy z_i is a *belief-free sequential best reply* to s^* starting 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 $B_i^t(s^*)$. Following Ely's et al. (2005) definition, let

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

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

For each period t, let $\alpha^t \in \Delta\left(\mathcal{A}^t\right)$ be a symmetric Nash equilibrium in the symmetric game $\left(\mathcal{A}^t, u\right)$ in which players are restricted to choosing actions only in $\mathcal{A}^t \subseteq A$. Let s' be the strategy in which each player plays the mixed action α^t in each period t. The definition of the regimes $\left(\mathcal{A}^t\right)_t$ implies that a mutant player who follows strategy s' best-replies to an incumbent who follows s^*, i.e., $U\left(s', s^*\right) = U\left(s^*, s^*\right)$. The definition of α^t implies that a mutant achieves a weakly higher payoff relative to the incumbents when facing another mutant: $U\left(s', s'\right) \geq U\left(s^*, s'\right)$. This implies that s^* can be evolutionarily stable only if $s' = s^*$, which implies that s^* is trivial. □

3.2. Neutral stability in generic games

As evolutionary stability is a strong refinement, it is desirable to show that belief-free equilibria also fail to satisfy weaker notions of stability. In this subsection, I show that non-trivial belief-free equilibria fail to satisfy the weaker notion of neutral stability under two mild assumptions: (1) the underlying game is generic, and (2) the monitoring structure has a grain of informativeness.

I begin by defining the notions of a generic game and a grain of informativeness. Fix a set of actions A. Consider a random process in which each payoff $u\left(a,a'\right)$ for each pair of actions $a,a'\in A$ is independently chosen at random from an arbitrary continuous (atomless) distribution. In what follows I require two properties, both of which, hold with probability one in such a process. The first requirement is that the same payoff not appear twice in the payoff matrix. The second requirement is that for each two actions a,a' in the support of a mixed equilibrium, the average payoff conditional on both players playing the same action in $\{a,a\}'$ should not be exactly the same as the average payoff conditional on each player playing a different action in $\{a,a\}'$. I say that games that satisfy these two properties are generic games. Formally:

Definition 7. Symmetric normal-form game $G=(A,u)$ is *generic* if it satisfies the following two properties:

1. $u\left(a,a'\right)\neq u\left(\hat{a},a'\right)$ for any actions $a\neq\hat{a},a'\in A$.
2. For each non-empty subset of actions $A'\subseteq A$, each symmetric equilibrium $\alpha\in\Delta\left(A'\right)$ of the restricted game $\left(A',u\right)$, and each two different actions $a\neq a'\in supp\left(\alpha\right)$, the following inequality holds:

$$\frac{\left(\alpha\left(a\right)\right)^{2}\cdot u\left(a,a\right)+\left(\alpha\left(a'\right)\right)^{2}\cdot u\left(a',a'\right)}{\left(\alpha\left(a\right)\right)^{2}+\left(\alpha\left(a'\right)\right)^{2}}\neq 0.5\cdot\left(u\left(a,a'\right)+u\left(a',a\right)\right). \tag{1}$$

I say that a monitoring structure has a grain of informativeness if for any mixed action played by the players, the joint distribution of action played and signal observed by each player can be used as a (possibly weak) correlation device between the players. Formally:

Definition 8. Fix a symmetric game $G=(A,u)$. A symmetric monitoring structure m has *a grain of informativeness* if for each mixed action $\alpha\in\Delta\left(A\right)$ with a non-trivial support ($\left|supp\left(\alpha\right)\right|>1$), there exist functions $f^{+},f^{-}:A\times\Sigma\rightarrow\{0,1\}$, such that if each player i chooses action a_i according to the distribution α, and at the end of the round observes signal σ_i, and calculates the values of $f^{+}\left(a_i,\sigma_i\right)$ and $f^{-}\left(a_i,\sigma_i\right)$, then the players' values of f^{+} (f^{-}) are positively (negatively) correlated, i.e.,

$$\begin{aligned}
&\Pr\left(f^{+}\left(a_1,\sigma_1\right)=f^{+}\left(a_2,\sigma_2\right)=1\right)\\
&=\sum_{(a,a')\in A^2}\alpha\left(a\right)\cdot\alpha\left(a'\right)\cdot\sum_{(\sigma,\sigma')\in\Sigma^2}m\left(\sigma,\sigma'|a,a'\right)\cdot f^{+}\left(a,\sigma\right)\cdot f^{+}\left(a',\sigma'\right)\\
&>\left(\sum_{(a,a')\in A^2}\alpha\left(a\right)\cdot\alpha\left(a'\right)\cdot\sum_{(\sigma,\sigma')\in\Sigma^2}m\left(\sigma,\sigma'|a,a'\right)\cdot f^{+}\left(a,\sigma\right)\right)^{2}\\
&=\Pr\left(f^{+}\left(a_1,\sigma_1\right)\right)\cdot\Pr\left(f^{+}\left(a_2,\sigma_2\right)\right),
\end{aligned}$$

and

$$\begin{aligned}
&\Pr\left(f^{-}(a_1,\sigma_1)=f^{-}(a_2,\sigma_2)=1\right)\\
&=\sum_{(a,a')\in A^2}\alpha(a)\cdot\alpha\left(a'\right)\cdot\sum_{(\sigma,\sigma')\in\Sigma^2}m\left(\sigma,\sigma'|a,a'\right)\cdot f^{-}(a,\sigma)\cdot f^{-}\left(a',\sigma'\right)\\
&<\left(\sum_{(a,a')\in A^2}\alpha(a)\cdot\alpha\left(a'\right)\cdot\sum_{(\sigma,\sigma')\in\Sigma^2}m\left(\sigma,\sigma'|a,a'\right)\cdot f^{-}(a,\sigma)\right)^2\\
&=\Pr\left(f^{-}(a_1,\sigma_1)\right)\cdot\Pr\left(f^{-}(a_2,\sigma_2)\right).
\end{aligned}$$

Intuitively, the mild requirement of a grain of informativeness is satisfied whenever the signal a player obtains (combined with his own action) is not completely uninformative about the partner's action. The following example shows how to explicitly construct f^{+} *and* f^{-} for conditionally independent signals.

Example 2. Consider a game with two actions $A=\{c,d\}$ and a monitoring structure with two signals $\Sigma=\{C,D\}$, such that player i observes signal C with probability $1-\epsilon$ (ϵ) if the partner plays c (d) for some $\epsilon<0.5$. Let the functions f^{+} and f^{-} be defined as follows: $f^{+}(c,C)=f^{+}(d,D)=0$, $f^{+}(c,D)=f^{+}(d,C)=1$, $f^{-}(c,D)=1$, $f^{-}(c,C)=f^{-}(d,D)=f^{-}(d,C)=0$. The values of f^{+} are positively correlated between the two players because these values differ only if there has been an observation error (a probability that is strictly less than 50%). The values of f^{-} are negatively correlated between the two players, because they coincide with the value of 1 only if there have been two observation errors (which happens with a small probability of $O\left(\epsilon^2\right)$).

The following result shows that if the game is generic and the monitoring structure has a grain of informativeness, then no non-trivial belief-free equilibrium satisfies neutral stability.

Proposition 2. *Assume that $G=(A,u)$ is a generic game and the monitoring structure has a grain of informativeness. Let s^* be a symmetric belief-free equilibrium that is also neutrally stable. Then s^* is trivial.*

Proof. Let $\gamma^t=\gamma^t(s^*)\in\Delta\left(\mathcal{A}^t\right)$ be the marginal distribution of actions played by each player in period t in the belief-free symmetric equilibrium s^*. Let $\mathcal{T}$ be the sequence of periods in which the support of γ^t includes at least two actions, (i.e., $\left\{t\in\mathbb{N}|\left|supp\left(\gamma^t\right)\right|>1\right\}$). If $\mathcal{T}=\emptyset$, then both players play a pure equilibrium in each period, and s^* is trivial. If $\mathcal{T}=\left\{\bar{t}\right\}$, then the fact that $\left|\gamma^t\right|=1$ for every $t\notin\mathcal{T}$, implies that both players play a pure equilibrium in each period $t\notin\mathcal{T}$, and that the players myopically best-reply to each other in round $\bar{t}$. Due to the fact that s^* is a belief-free equilibrium, this implies that each action $a\in\mathcal{A}^{\bar{t}}$ is a myopic best reply against the partner for any possible history of length $\bar{t}$, which implies that the players play a Nash equilibrium of the stage game (which is independent of the observed history) in round $\bar{t}$, and that s^* is trivial.

Next assume that there exists $\hat{t}\in\mathcal{T}$, such that the restricted normal-form game $\left(supp\left(\gamma^{\hat{t}}\right),u\right)$ admits a symmetric pure equilibrium. This equilibrium must be strict due to the game being generic. Let s' be the strategy that induces mutants to play in each period $t\neq\hat{t}$ a symmetric mixed equilibrium (which depends on the period, but not on the observed history) in the restricted normal-form game $\left(supp\left(\gamma^{t}\right),u\right)$, and to play a strict symmetric equilibrium in the restricted game $\left(supp\left(\gamma^{\hat{t}}\right),u\right)$ in period $\hat{t}$. The definition of s' and the fact that s^* is belief-free imply that $U\left(s',s^*\right)=U\left(s^*,s^*\right)$, and that $U\left(s^*,s'\right)<U\left(s',s'\right)$. The latter inequality holds because the mutants achieve a strictly higher payoff in round $\hat{t}$ and a weakly higher payoff against other mutants in all other rounds. This contradicts the assumption that s^* is neutrally stable.

Thus, we are left with the case in which there exist $t_1<t_2\in\mathcal{T}$, such that the restricted normal-form game $\left(supp\left(\gamma^{t_1}\right),u\right)$ ($\left(supp\left(\gamma^{t_2}\right),u\right)$) admits a symmetric non-pure equilibrium α_1 (α_2), i.e., $|supp\left(\alpha_1\right)|>1$ ($|supp\left(\alpha_2\right)|>1$). Assume first that the LHS of (1) is greater than the RHS. Let f^+ be the function defined in Definition 8 with respect to the mixed action α_1. Let s^+ ($\tilde{s}^+$) be the strategy that induces an agent who follows it (1) to play the mixed action α_1 in round t_1, (2) to play a symmetric equilibrium in the restricted game $\left(supp\left(\gamma^{t}\right),u\right)$ in each round $t\neq t_1,t_2$, and (3) to play on the marginal the mixed equilibrium α_2 in round t_2, but to condition his play on the values of a_1 (his own action in round t_1) and σ_1 (the signals he observed in round t_1); specifically, the agent is more (less) likely to play action a and less (more) likely to play action a' when $f^+\left(a^k,\sigma^k\right)=1$. These changes in the probabilities of playing actions a and a' are determined, such that, after each history h^{t_2} of length t_2, the mixture of the mixed action played by an agent who follows strategy s^+ and the mixed action played by an agent who follows strategy $\tilde{s}^+$ is α_2, i.e., for each action $\hat{a}\in A$, $0.5\cdot s^+\left(h^{t_2}\right)\left(\hat{a}\right)+0.5\cdot\tilde{s}^+\left(h^{t_2}\right)\left(\hat{a}\right)=\alpha_2\left(\hat{a}\right)$.

Observe that the strategies s^+ and $\tilde{s}^+$ induce the same behavior in all rounds $t\neq t_2$. Let s_{mix} be the mixture of the strategies s^+ and $\tilde{s}^+$; i.e., $s_{mix}\equiv\alpha_2$ in round t_2, and s_{mix} coincides with s^+ and $\tilde{s}^+$ in each round $t\neq t_2$. Observe that s_{mix} induces an agent who follows it to play symmetric mixed equilibria in all rounds. This implies that $U\left(s^*,s_{mix}\right)\leq U\left(s_{mix},s_{mix}\right)$. The fact that s_{mix} is a mixture of s^+ and $\tilde{s}^+$ (and that the three strategies coincide in all rounds $t\neq t_2$) implies that $U\left(s^*,s_{mix}\right)=0.5\cdot U\left(s^*,s^+\right)+0.5\cdot U\left(s^*,\tilde{s}^+\right)$. This implies that either $U\left(s^*,s^+\right)\leq U\left(s_{mix},s_{mix}\right)$ or $U\left(s^*,\tilde{s}^+\right)\leq U\left(s_{mix},s_{mix}\right)$. Assume without loss of generality that $U\left(s^*,s^+\right)\leq U\left(s_{mix},s_{mix}\right)$.

Consider a homogeneous group of mutants, each following strategy s^+. The definition of s^+ and the fact that s^* is belief-free imply that $U\left(s^+,s^*\right)=U\left(s^*,s^*\right)$, and that $U\left(s^+,s^+\right)>U\left(s_{mix},s_{mix}\right)\geq U\left(s^*,s^+\right)$. The inequality $U\left(s^+,s^+\right)>U\left(s_{mix},s_{mix}\right)$ holds because strategy s^+ coincides with strategy s_{mix} in any period $t\neq t_2$. In period t_2 agents who follow strategy s^+ achieve a higher expected payoff when being matched with other agents who follow strategy s^+ because when these agents are matched they induce a positive correlation in their random play of the actions a and a', which increases their average payoff, due to the LHS of (1) being greater than the RHS, relative to the uncorrelated profile played by agents who follow the strategy s_{mix}. This implies that s^* is not neutrally stable.

If the LHS of (1) is less than the RHS, then we define analogous strategies $s^-$and $\tilde{s}^-$ with respect to the function f^-, and use an analogous argument to the one above where $s^-(\tilde{s}^-)$ replaces s^+ ($\tilde{s}^+$) and negative correlation replaces positive correlation in the random play of the mutants in round t_2. □

3.3. Weak stability in recursively strict symmetric games

Although neutral stability is considered to be a mild evolutionary refinement, the arguments presented in Section 2.3 suggest that in some setups it may be too strong, and it would be desirable to extend the instability result to a weaker evolutionary refinement. In what follows I study the family of recursively strict games, and show that within this family any weakly stable belief-free equilibrium is trivial.

I say that a symmetric game is recursively strict, if all the symmetric games induced by restricting both players to choosing actions from a given subset of actions admit a strict symmetric equilibrium. Formally:

Definition 9. A symmetric normal-form game $G = (A, u)$ is *recursively strict* if for any non-empty subset of actions $A' \subseteq A$, the game $G = \left(A', u\right)$, in which players are restricted to choose actions from A', admits a strict symmetric equilibrium (i.e., there is $a \in A'$ such that $u(a, a) > u\left(a', a\right)$ for each $a' \neq a \in A'$).

A few examples of recursively strict games are:

1. The Prisoner's Dilemma (as described in Table 1).
2. Symmetric coordination games, which satisfy that (a, a) is a strict equilibrium for each action $a \in A$.
3. Games with an ordered set of actions $A = \{a_1, ..., a_n\}$, which satisfy that $u(a_k, a_k) > u(a_l, a_k)$ for each $1 \leq k < l \leq n$. In particular, such games include:
   (a) Traveler's Dilemma game (Basu, 1994). The set of actions is $A = \{2, ..., 100\}$ (interpreted as evaluations of the value of one of two lost identical suitcases), both players get a payoff equal to the minimal evaluation, and, in addition, if the evaluations differ, then the player who wrote the lower (higher) evaluation gets a bonus (malus) of 2 to his payoff.
   (b) Public good games. The index $1 \leq k \leq n$ is interpreted as the level of contribution to a public good. The payoff for a player who plays a_k and whose partner plays a_l is $f(k, l) - g(k)$, where the function f is symmetric, strictly supermodular, and increasing in both parameters, the function g is strictly increasing and convex, and $f(k+1, k) - g(k+1) < f(k, k) - g(k)$ for each $k < n$.

Our next result shows that only trivial and pure belief-free equilibria satisfy the mild refinement of weak stability if the underlying stage game is recursively strict. In particular, the symmetric Prisoner's Dilemma game admits a unique weakly stable belief-free equilibrium in which both players defect in all periods.

Proposition 3. *Assume that the symmetric underlying game $G = (A, u)$ is recursively strict. Let s^* be a symmetric belief-free equilibrium. If s^* is weakly stable, then it is trivial and pure.*

Proof. Let $\gamma^t = \gamma^t(s^*) \in \Delta\left(\mathcal{A}^t\right)$ be the marginal distribution of actions played by each player in period t in the belief-free symmetric equilibrium s^*. Assume first that γ^t is pure in all periods t. This implies that s^* induces a deterministic play that is independent of the observed signals. Thus a player's best reply coincides with his myopic best reply, which implies that the pure action profile played in each period must be an equilibrium of the underlying game (i.e., s^* is trivial and pure).

Otherwise, there exists time t such that $\left|supp\left(\gamma^t(s^*)\right)\right| > 1$. For each period t, let $a_1^t \in supp\left(\gamma^t(s^*)\right)$ be a strict symmetric equilibrium in the symmetric game $\left(supp\left(\gamma^t(s^*)\right), u\right)$ in which players are restricted to choosing actions only in $supp\left(\gamma^t(s^*)\right)$. Let s_1 be the strategy in which each player chooses action a_1^t in each period t. The definition of the regimes $\left(\mathcal{A}^t\right)_t$ implies that a mutant player who follows strategy s_1 best-replies to an incumbent who follows

s^*, i.e., $U(s_1, s^*) = U(s^*, s^*)$. The definition of a_1^t implies that a mutant achieves a strictly higher payoff relative to the incumbents when facing another mutant: $U(s_1, s_1) > U(s^*, s_1)$.

For each $1 \le k$, define $A_{k+1}^t = \text{argmax}_{a \in A} u\left(a, a_k^t\right)$ as the set of pure best replies against a_k^t. Let a_{k+1}^t be a strict symmetric equilibrium in the symmetric game $\left(A_{k+1}^t, u\right)$. Observe that there exists a minimal $1 \le \bar{k} \le |A|$ such that for each t, $A_{\bar{k}}^t = A_{\bar{k}+1}^t = \left\{a_{\bar{k}}^t\right\}$ is a singleton, which implies that action $a_{\bar{k}}^t$ is a strict equilibrium of the unrestricted game (A, u). This is because otherwise the sequence of actions $\left(a_1^t, ..., a_{|A|+1}^t\right)$ must include a non-trivial cycle, which contradicts the fact that there exists an action $\hat{a} \in \left\{a_1^t, ..., a_{|A|+1}^t\right\}$ that is a strict equilibrium in the restricted game $\left(\left\{a_1^t, ..., a_{|A|+1}^t\right\}, u\right)$.

For each $2 \le k \le \bar{k}$, let s_k be the strategy in which each player chooses action a_k^t in each period t. The definitions of the strategies $\{s_1, ..., s_{\bar{k}}\}$ imply that (1) each strategy s_k is vulnerable to the strategy s_{k+1}, i.e., $s_{k+1} \in B(s_k)$, and $U(s_{k+1}, s_{k+1}) > U(s_k, s_{k+1})$, and (2) $s_{\bar{k}}$ is a pure strategy in which the players play a strict symmetric equilibrium of the underlying (unrestricted) game in each round t, which implies that $s_{\bar{k}}$ is evolutionarily stable, and that s^* is not weakly stable. □

Remark 3 *(Instability of belief-free review-strategy equilibria).* Matsushima (2004), Yamamoto (2007, 2012), and Deb (2012) use the notion of a *belief-free review-strategy equilibrium* (also called *block 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 (2) at the beginning of each review phase, a player's continuation strategy is a best reply regardless of the history. A simple adaptation of the proof of Proposition 3 show that defection is the unique weakly stable symmetric belief-free review-strategy equilibrium.[11]

The sketch of the adaptation of the proof is 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. The strategy s_1 is adapted such that it is defined in each round t_l that begins a review process in an analogous way to the definition given in Proposition 3, and it induces agents who follow it to play the same action up to the end of the l-th review phase. The remaining strategies $\{s_2, ..., s_{\bar{k}}\}$ are defined in the same way as in the proof of Proposition 3, and analogous arguments show that s^* is not weakly stable (unless it is trivial and pure).

3.4. Weak stability in hawk–dove games

The Hawk–Dove game (see the payoff matrix in Table 1) is a common application of belief-free equilibria. This game does not admit a strict symmetric equilibrium, and thus the general results above show only that non-trivial belief-free equilibria are not neutrally stable.[12] The main difficulty in analyzing weak stability of belief-free in Hawk–Dove games is that, in general, it is

11 Similarly, one can further adapt the proof to show the instability of Sugaya's (2015) equilibria, in which each review phase is divided into several sub-phases, and players may switch their action at the beginning of each sub-phase.

12 One can show that any Hawk–Dove game is generic according to Definition 7. Specifically, the LHS of (1) is always smaller than the RHS because the unique mixed equilibrium ($\tilde{\alpha}(c) = \frac{l}{l+g}$) yields an expected payoff of $\frac{l \cdot (1+g)}{l+g}$, which is strictly less than the average payoff of the agents conditional on playing different actions $(1 + g + l)$. Thus, the result holds for any monitoring structure with a grain of informativeness.

an open question whether a repeated game without strict symmetric equilibria admits an evolutionarily stable strategy. In this subsection I show that any belief-free equilibrium in the repeated Hawk–Dove game is weakly stable iff the monitoring structure is such that the game admits evolutionarily stable strategies.

The one-shot Hawk–Dove game admits a unique symmetric equilibrium, which is also an evolutionarily stable strategy. Analogous arguments to those appearing in the proof of Proposition 2 show that the trivial belief-free equilibrium in which the players keep playing the symmetric equilibrium of the one-shot game is not neutrally stable in the repeated games if the monitoring structure has a grain of informativeness (because mutants can use past history to induce a negative correlation between their played actions and thus outperform the incumbents).

A repeated Hawk–Dove game with imperfect public monitoring (with full support) admits evolutionarily stable strategies. One example of an evolutionarily stable strategy is the one according to which each agent mixes in the first round with some distribution $(\alpha(c), \alpha(d))$, which is chosen such that each player is indifferent between the two actions. If the public signal σ that is observed at the end of the first round is such that the action profile (c, d) is more (less) likely than (d, c), conditional on observing σ, then the players play in all the remaining rounds the deterministic sequence $((d, c), (c, d), (d, c), (c, d), ...)$ $(((c, d), (d, c), (c, d), (d, c) ...))$. If both asymmetric action profiles have the same posterior probability conditional on observing σ, then the players randomize according to $(\alpha(c), \alpha(d))$ in the next round, and repeat the same procedure described above.

It is an open problem, which is left for future research, whether a repeated Hawk–Dove game with private monitoring admits an evolutionarily (or neutrally) stable strategy. The following result shows that a belief-free equilibrium is weakly stable iff the monitoring structure is such that the repeated Hawk–Dove game does not admit an evolutionarily stable strategy.

Proposition 4. *Let the underlying game $G = (\{c, d\}, u)$ be a Hawk–Dove game. Assume that the monitoring structure has a grain of informativeness. Let s^* be a belief-free equilibrium of the repeated game. Then s^* is weakly stable iff the repeated game does not admit an evolutionarily stable strategy.*

Proof. If the repeated game does not admit an evolutionarily stable strategy, then it is immediate from the definition of weak stability that s^* is weakly stable. Otherwise, let $\hat{s}$ be an evolutionarily stable strategy of the repeated game. Let $\gamma^t = \gamma^t(s^*) \in \Delta(\mathcal{A}^t)$ be the marginal distribution of actions played by each player in period t in the belief-free symmetric equilibrium s^*.

Let $\mathcal{T}$ be the sequence of periods in which both actions are played with positive probability, (i.e., $\{t \in \mathbb{N} | supp(\gamma^t) = \{c, d\}\}$). Assume first that $\mathcal{T}$ is finite, and let $\bar{t} = max(\mathcal{T})$ be the last element in $\mathcal{T}$. Observe, that for every $t > \bar{t}$, both players play deterministically, and at each such period they choose the same action. However, this implies s^* is not a Nash equilibrium, as one of the players can achieve a strictly higher payoff by choosing the opposite action at each period $t > \bar{t}$.

Next, assume that $\mathcal{T} = \mathbb{N}$. The fact that both actions are best replies at all periods implies that $U(\hat{s}, s^*) = U(s^*, s^*)$, and because $\hat{s}$ is evolutionarily stable this equality implies that $U(s^*, \hat{s}) < U(\hat{s}, \hat{s})$. Thus, s^* is vulnerable to $\hat{s}$, which implies that s^* is not weakly stable.

Thus, we are left with the case in which $|\mathcal{T}| = \infty$ and $\mathbb{N} \backslash \hat{\mathcal{T}} \neq \emptyset$. Let $\tilde{\alpha}$ be the unique symmetric equilibrium of the (unrestricted) stage game. Given two periods $t_k < t_l \in \mathcal{T}$, let $s'_{(t_k, t_l)}$ be a strategy that induces the mutants to play in each round $t \neq t_l$ a symmetric equilibrium in the restricted game $(supp(\gamma^t), u)$. Observe that the mutants who follow $s'_{(t_k, t_l)}$ play the mixed equilibrium $\tilde{\alpha}$ in each round $t \neq t_l \in \mathcal{T}$. Let a_k be the action the agent played at time t_k and let σ_k be the signal he observed at the end of round t_k. In period t_l an agent who follows strategy s' plays on the marginal the mixed equilibrium $\tilde{\alpha}$, but he conditions his play on the values of a_k and σ_k. Specifically, the agent is more likely to play action c and less likely to play action d when $f^-\left(a^k, \sigma^k\right) = 1$.

Let $t_0 \in \mathbb{N}\backslash\mathcal{T}$. The fact that $\mathcal{T}$ is infinite implies that for any $\epsilon > 0$, there exist $t_k < t_l \in \mathcal{T}$ such that (1) the probability that an agent plays action c at time t_l changes by at most ϵ conditional on the value of $f^-\left(a^k, \sigma^k\right)$ when the agent follows strategy s^* and faces a partner who follows strategy $s'_{(t_k,t_l)}$, and (2) the myopic gain from playing α instead of $s^*(t_0)$ in period t_0 outweighs the maximal possible discounted loss in period t_l, i.e.,

$$u\left(\tilde{\alpha}, s^*(t_0)\right) - u\left(s^*(t_0), s^*(t_0)\right) > \delta^{t_l - t_0} \cdot max\,(g, l)\,.$$

Let $t_k < t_l \in \mathcal{T}$ be two periods that satisfy these conditions for a sufficiently small ϵ. The definition of $s'_{(t_k,t_l)}$ and the fact that s^* is belief-free implies that $U\left(s'_{(t_k,t_l)}, s^*\right) = U(s^*, s^*)$, and that $U\left(s'_{(t_k,t_l)}, s'_{(t_k,t_l)}\right) > U\left(s^*, s'_{(t_k,t_l)}\right)$. The latter inequality holds due to the same argument as in the proof of Proposition 2 (see also footnote 12). This implies that s^* is vulnerable to $s'_{(t_k,t_l)}$.

Let $s_{\tilde{\alpha}}$ be the strategy that plays the mixed equilibrium $\tilde{\alpha}$ at all periods (regardless of the observed history). The second condition in the definition of $s'_{(t_k,t_l)}$ above implies that $U\left(s_{\tilde{\alpha}}, s'_{(t_k,t_l)}\right) > U\left(s'_{(t_k,t_l)}, s'_{(t_k,t_l)}\right)$ because the two strategies have the same expected payoff when facing a partner who follows $s'_{(t_k,t_l)}$ in all periods in $\mathcal{T}\backslash\{t_l\}$, and the higher payoff that $s_{\tilde{\alpha}}$ yields in each period in $\mathbb{N}\backslash\hat{\mathcal{T}}$ outweighs the lower payoff in period t_l.The definition of $s_{\tilde{\alpha}}$ implies that $U(s_{\tilde{\alpha}}, s_{\tilde{\alpha}}) = U\left(s'_{(t_k,t_l)}, s_{\tilde{\alpha}}\right)$. Thus, strategy $s'_{(t_k,t_l)}$ is vulnerable to $s_{\tilde{\alpha}}$.

The definition of $s_{\tilde{\alpha}}$ implies that $U(s_{\tilde{\alpha}}, s_{\tilde{\alpha}}) = U\left(\hat{s}, s_{\tilde{\alpha}}\right)$, and because $\hat{s}$ is evolutionarily stable this equality implies that $U\left(s_{\tilde{\alpha}}, \hat{s}\right) < U\left(\hat{s}, \hat{s}\right)$. Thus, $s_{\tilde{\alpha}}$ is vulnerable to $\hat{s}$, which implies that s^* is not weakly stable. □

An immediate corollary of Proposition 4 is that repeated Hawk–Dove games with public imperfect monitoring do not admit any weakly stable belief-free equilibria (see Kandori and Obara, 2006, for an application of belief-free equilibria with public imperfect monitoring).

3.5. Instability of the belief-based equilibria of Bhaskar and Obara (2002)

Proposition 3 shows that non-trivial belief-free equilibria do not satisfy weak stability in the repeated Prisoner's Dilemma. The literature on the repeated Prisoner's Dilemma with private monitoring also includes another approach to induce cooperation, namely, the belief-based equilibria of Bhaskar and Obara (2002). In what follows I sketch why these particular belief-based equilibria also fail to satisfy weak stability.

Bhaskar and 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"). Bhaskar and 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 and Obara present for each $0 < x < 1$ a symmetric sequential equilibrium s_x that yields a 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 I sketch this equilibrium s_x, and then show that it is not weakly stable.

Let s_T be the trigger strategy: cooperate 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 set of periods into disjoint sequences (say, into n sequences, $(\mathcal{T}_1, ..., \mathcal{T}_n)$, where sequence $\mathcal{T}_k$ includes the periods that are equal to k modulo n), and the play in each sequence is independent of the other sequences. Each player mixes in the first round of each sequence: he plays s_T (trigger strategy) with probability π and plays s_d (always defect) with the remaining probability. Bhaskar and Obara show that there exist a division into sequences and a mixing probability π such that (1) the expected discounted symmetric payoff of the game is at least x, and (2) strategy s_x is a sequential equilibrium.[13]

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 sequence $\mathcal{T}_k$ implies that s_d is a best reply to 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. □

Appendix A. Analysis of asymmetric games

The main text analyzes stability of symmetric equilibria in symmetric games, as this is the setup analyzed in most of the evolutionary game theory literature. In many applications, there are observable differences between the agents (e.g., age, sex, and status), which can be perceived by both agents and upon which behavior can be conditioned. When these differences are payoff-relevant (or monitoring-relevant), the repeated game is asymmetric, and when the differences are payoff-irrelevant, the underlying game is symmetric, but agents can still condition their play on these observable differences. For brevity, we will use the notion of asymmetric games to refer to both situations. The appendix adapts the notions of stability and extends the main results to deal with asymmetric games.

A.1. Definitions of stability in asymmetric games

In this subsection I adapt the notions of stability to the setup of asymmetric games (see Weibull, 1995, Chapter 2.7 and Van Damme, 1991, Sections 9.5–9.8, for introductory textbooks). I consider a large population of agents (technically, a continuum) in which agents are drawn to play a two-person repeated game, and at the beginning of each such repeated interaction, nature randomly determines who will be player 1 and who will be player 2, such that each agent has a probability of 50% of being in each role.

Each agent in the population follows a strategy profile (s_1, s_2), where each s_i describes the behavior of the agent when he is assigned to play the role of player i. The (*ex-ante*) expected payoff of an agent who follows strategy profile (s_1, s_2) and is matched with a partner who follows strategy profile $\left(s_1', s_2'\right)$ is

[13] As observed by Bhaskar (2000) and Bhaskar et al. (2008), these belief-based equilibria can be purified à la Harsanyi (1973) in a simple way (while this is not the case for belief-free equilibria). Nevertheless, I show that they still do not satisfy weak stability.

$$\bar{U}\left(\left(s_1, s_2\right), \left(s'_1, s'_2\right)\right) := \frac{1}{2} \cdot U_1\left(s_1, s'_2\right) + \frac{1}{2} \cdot U_2\left(s'_1, s_2\right),$$

which is the average agent's payoff in each of the two possible roles.

A strategy profile $\left(s_1^*, s_2^*\right)$ is a Nash equilibrium if it is the best reply against itself, i.e.,

$$\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^*, s_2^*\right)\right) \geq \bar{U}\left(\left(s'_1, s'_2\right), \left(s_1^*, s_2^*\right)\right) \ \forall \left(s'_1, s'_2\right) \in S_1 \times S_2 \Leftrightarrow$$
$$U_1\left(s_1^*, s_2^*\right) \geq U_1\left(s'_1, s_2^*\right) \ \forall s'_1 \in S_1 \text{ and } U_2\left(s_1^*, s_2^*\right) \geq U_2\left(s_1^*, s'_2\right) \ \forall s'_2 \in S_2.$$

Let $B\left(s_1^*, s_2^*\right)$ denote the set of strategy profiles that are best replies against the strategy profile $\left(s_1^*, s_2^*\right)$, i.e.,

$$B\left(s_1^*, s_2^*\right) = \operatorname{argmax}_{(s_1, s_2) \in S_1 \times S_2}\left(\bar{U}\left(\left(s_1, s_2\right), \left(s_1^*, s_2^*\right)\right)\right).$$

Recall that $\pi^t_{s_i, s_{-i}} \in \Delta\left(H_i^t\right)$ denotes the probability that a player who follows strategy s_i observes history h_i^t, conditional on the opponent following s_{-i}, and recall that history $h_i^t \in H_i^t$ is *feasible* given strategy $\hat{s}_i$ if there exists strategy $s_{-i} \in S_{-i}$ such that $\pi^t_{\hat{s}_i, s_{-i}}\left(h_i^t\right) > 0$. Two strategy profiles are outcome-equivalent if they always induce the same behavior regardless of the opponent's strategy profile. Formally:

Definition 10. Strategy profiles (s_1, s_2) and $\left(s'_1, s'_2\right)$ are *outcome-equivalent* if: (1) their sets of feasible histories coincide (i.e., for each role i, history h_i^t is feasible given s_i iff it is feasible given s'_i), and (2) they coincide after each feasible history (i.e., $s_i\left(h_i^t\right) = s'_i\left(h_i^t\right)$ for each player i and for each feasible history h_i^t).

Given a strategy profile (s_1, s_2), let $[(s_1, s_2)]$ denote its *equivalent set* (i.e., the set of strategies that are outcome-equivalent to (s_1, s_2)).

A Nash equilibrium $\left(s_1^*, s_2^*\right)$ is evolutionarily stable if the incumbents (who follow $\left(s_1^*, s_2^*\right)$), achieve a higher payoff against any best-replying mutants (who follow strategy profile $\left(s'_1, s'_2\right)$). Formally:

Definition 11 *(Taylor, 1979; Weibull, 1995, Chapter 5.1).* A Nash equilibrium $\left(s_1^*, s_2^*\right)$ is *neutrally (evolutionarily) stable* if $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s'_1, s'_2\right)\right) \geq \bar{U}\left(\left(s'_1, s'_2\right), \left(s'_1, s'_2\right)\right)$ ($\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s'_1, s'_2\right)\right) > \bar{U}\left(\left(s'_1, s'_2\right), \left(s'_1, s'_2\right)\right)$), for each best-reply strategy profile $\left(s'_1, s'_2\right) \in B\left(s_1^*, s_2^*\right) \setminus \left[\left(s_1^*, s_2^*\right)\right]$.

In what follows I adapt the notion of weak stability to the setup of asymmetric games. Strategy profile $\left(s_1^*, s_2^*\right)$ is vulnerable to $\left(s'_1, s'_2\right)$ if the former induces a strictly higher payoff in any heterogeneous population in which some of the agents follow $\left(s_1^*, s_2^*\right)$ and the others follow $\left(s'_1, s'_2\right)$. Formally:

Definition 12. Strategy profile $\left(s_1^*, s_2^*\right)$ is *vulnerable* to $\left(s'_1, s'_2\right)$ if $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^*, s_2^*\right)\right) \leq \bar{U}\left(\left(s'_1, s'_2\right), \left(s_1^*, s_2^*\right)\right)$, $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s'_1, s'_2\right)\right) \leq \bar{U}\left(\left(s'_1, s'_2\right), \left(s'_1, s'_2\right)\right)$, and at least one of these inequalities is strict.

Note that a neutrally stable equilibrium is not vulnerable to any strategy profile.

A Nash equilibrium is weakly stable if there does not exist a sequence of strategy profiles, starting with this equilibrium and ending with an evolutionarily stable equilibrium, such that each profile in the sequence is vulnerable to its successor. Formally:

Definition 13. A Nash equilibrium $\left(s_1^*, s_2^*\right)$ is *weakly stable* if there does not exist a finite non-empty sequence of strategy profiles $\left(\left(s_1^1, s_2^1\right), ..., \left(s_1^K, s_2^K\right)\right)$ such that: (1) strategy profile $\left(s_1^*, s_2^*\right)$ is vulnerable to $\left(s_1^1, s_2^1\right)$, (2) for each $1 \leq k < K$, strategy profile $\left(s_1^k, s_2^k\right)$ is vulnerable to $\left(s_1^{k+1}, s_2^{k+1}\right)$, and (3) strategy profile $\left(s_1^K, s_2^K\right)$ is evolutionarily stable.

It is well known (Weibull, 1995, Chapter 2.7) that a strategy profile is evolutionarily stable iff it is a strict equilibrium. This immediately implies that only trivial belief-free equilibria may be evolutionarily stable.

Fact 1. Let $\left(s_1^*, s_2^*\right)$ be a belief-free equilibrium that is also evolutionarily stable. Then $\left(s_1^*, s_2^*\right)$ is trivial and pure.

A.2. *Result for generic asymmetric games*

In this subsection I show how to adapt Proposition 2 to deal with asymmetric games. I say that a game is generic if: (1) the same number does not appear twice in the same column of the payoff matrix of player i, and (2) for each two pairs of actions $\left\{(a_1, a_2), \left(a_1', a_2'\right)\right\}$ in the support of a mixed equilibrium, the average payoff conditional on the players playing either of these action profiles is not exactly the same as the average payoff conditional on the players playing either of the "crossed" action profiles $\left\{\left(a_1, a_2'\right), \left(a_1', a_2\right)\right\}$. Both properties hold with probability one if each payoff is independently distributed from a continuous (atomless) distribution. Formally, for each action profile $(a_1, a_2) \in A_1 \times A_2$ let $\bar{u}(a_1, a_2) := 0.5 \cdot (u_1(a_1, a_2) + u_2(a_1, a_2))$ denote the average payoff of two players who follow the action profile (a_1, a_2).

Definition 14. A normal-form game $G = ((A_1, A_2), u)$ is *generic* if it satisfies the following two conditions:

1. For any player i and any actions $a_i \neq a_i' \subseteq A_i$ and $a_{-i} \in A_{-i}$, the following inequality holds: $u_i(a_i, a_{-i}) \neq u_i\left(a_i', a_{-i}\right)$.
2. For any pair of non-empty subsets of actions $A_1' \subseteq A_1$ and $A_2' \subseteq A_2$, any equilibrium $(\alpha_1, \alpha_2) \in \Delta\left(A_1'\right) \times \Delta\left(A_2'\right)$ of the restricted game $\left(\left(A_1', A_2'\right), u\right)$, and any two pairs of actions $a_1 \neq a_1' \in supp(\alpha_1)$ and $a_2 \neq a_2' \in supp(\alpha_2)$, the following inequality holds:

$$\frac{\alpha_1(a_1) \cdot \alpha_2(a_2) \cdot \bar{u}(a_1, a_2) + \alpha_1\left(a_1'\right) \cdot \alpha_2\left(a_2'\right) \cdot \bar{u}\left(a_1', a_2'\right)}{\alpha_1(a_1) \cdot \alpha_2(a_2) + \alpha_1\left(a_1'\right) \cdot \alpha_2\left(a_2'\right)} \neq \frac{\alpha_1(a_1) \cdot \alpha_2\left(a_2'\right) \cdot \bar{u}\left(a_1, a_2'\right) + \alpha_1\left(a_1'\right) \cdot \alpha_2(a_2) \cdot \bar{u}\left(a_1', a_2\right)}{\alpha_1(a_1) \cdot \alpha_2\left(a_2'\right) + \alpha_1\left(a_1'\right) \cdot \alpha_2(a_2)}. \tag{2}$$

Observe that the first requirement implies that if in a Nash equilibrium, one of the players plays a pure action, then it must be that the equilibrium is pure (for both players) and strict.

I say that a monitoring structure has a grain of informativeness if for any mixed action played by the players, the joint distribution of the action played and the signal observed by each player can be used as a (possibly weak) correlation device between the players. Formally:

Definition 15. Fix a normal-form game $G = ((A_1, A_2), u)$. A monitoring structure m has *a grain of informativeness* if for each profile of mixed actions $(\alpha_1 \in \Delta(A_1), \alpha_2 \in \Delta(A_2))$ with a non-trivial support ($|supp(\alpha_i)| > 1$ for each i), there exists a pair of functions $\left(f_1^+ : A_1 \times \Sigma_1 \to \{0,1\}, f_2^+ : A_2 \times \Sigma_2 \to \{0,1\}\right)$, such that when each player i chooses action a_i according to distribution α_i, observes signal σ_i, and calculates the values of $f_i^+(a_i, \sigma_i)$, then the players' values of f_1^+ and f_2^+ are positively correlated, i.e.,

$$\begin{aligned}
&\Pr\left(f_1^+(a_1,\sigma_1) = f_2^+(a_2,\sigma_2) = 1\right) \\
&= \sum_{(a_1,a_2)\in A_1\times A_2} \alpha_1(a_1)\cdot\alpha_2(a_2)\cdot \sum_{(\sigma_1,\sigma_2)\in\Sigma_1\times\Sigma_2} m(\sigma_1,\sigma_2|a_1,a_2)\cdot f_1^+(a_1,\sigma_1)\cdot f_2^+(a_2,\sigma_2) \\
&> \prod_{i\in\{1,2\}} \left(\sum_{(a_1,a_2)\in A_1\times A_2} \alpha_1(a_1)\cdot\alpha_2(a_2)\cdot \sum_{(\sigma_1,\sigma_2)\in\Sigma_1\times\Sigma_2} m(\sigma_1,\sigma_2|a_1,a_2)\cdot f_i^+(a_i,\sigma_i) \right) \\
&= \Pr\left(f_1^+(a_1,\sigma_1) = 1\right)\cdot \Pr\left(f_2^+(a_2,\sigma_2) = 1\right).
\end{aligned}$$

Intuitively, the mild requirement of grain of informativeness is satisfied whenever the signal each player obtains (combined with his own action) is not completely uninformative about the partner's action.

The following result shows that if the game is generic and the monitoring structure has a grain of informativeness, then no non-trivial belief-free equilibrium satisfies neutral stability.

Proposition 5. *Assume that $G = (A, u)$ is a generic game, and the monitoring structure has a grain of informativeness. Let $\left(s_1^*, s_2^*\right)$ be a belief-free equilibrium that is also neutrally stable. Then $\left(s_1^*, s_2^*\right)$ is trivial.*

Proof. Let $\gamma_i^t = \gamma_i^t(s^*) \in \Delta\left(\mathcal{A}^t\right)$ be the marginal distribution of actions played by each player i at period t in the belief-free symmetric equilibrium $\left(s_1^*, s_2^*\right)$. Let $\mathcal{T}$ be the sequence of periods in which the support of either players includes at least two actions, i.e., $\left\{t \in \mathbb{N} | max\left(\left|supp\left(\gamma_1^t\right), \left|\left|supp\left(\gamma_2^t\right)\right|\right) > 1\right\}$.

If $\mathcal{T} = \emptyset$, then both players play a pure equilibrium in each period, and $\left(s_1^*, s_2^*\right)$ is trivial. If $\mathcal{T} = \left\{\bar{t}\right\}$, then the fact that $\left|\gamma_i^t\right| = 1$ for every $t \notin \mathcal{T}$, implies that both players play a pure equilibrium in each period $t \notin \mathcal{T}$, and that the players myopically best-reply to each other in round $\bar{t}$. Due to the fact that $\left(s_1^*, s_2^*\right)$ is a belief-free equilibrium, this implies that each action $a_i \in \mathcal{A}_i^{\bar{t}}$ is a myopic best reply against the partner for any possible history of length $\bar{t}$, which implies that the players play a Nash equilibrium of the stage game (which is independent of the observed history) in round $\bar{t}$, and that $\left(s_1^*, s_2^*\right)$ is trivial.

Next, assume that there exists $\hat{t} \in \mathcal{T}$ such that the restricted game $\left(\left(supp\left(\gamma_1^{\hat{t}}\right), supp\left(\gamma_2^{\hat{t}}\right)\right), u\right)$ admits an equilibrium in which either of the players plays a pure action. Due to the game being generic, this implies that this equilibrium, $\left(a_1^{\hat{t}}, a_2^{\hat{t}}\right)$, is pure (for both players) and strict. For each period $t \neq \hat{t}$, let $\left(\alpha_1^t, \alpha_2^t\right) \in \Delta\left(supp\left(\gamma_1^t\right)\right) \times \Delta\left(supp\left(\gamma_2^t\right)\right)$ be an equilibrium of the restricted game $\left(\left(supp\left(\gamma_1^{\hat{t}}\right), supp\left(\gamma_2^{\hat{t}}\right)\right), u\right)$. Let $\left(s_1', s_2'\right)$ be the strategy profile in which each player i plays mixed action α_i^t in each period $t \neq \hat{t}$ (regardless of the observed history) and plays action $a_i^{\hat{t}}$ in period $\hat{t}$. The fact that $\left(s_1^*, s_2^*\right)$ is belief-free and the definition of $\left(s_1', s_2'\right)$ imply that (1) $\left(s_1', s_2'\right)$ is a best reply against $\left(s_1^*, s_2^*\right)$, and (2) $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1', s_2'\right)\right) < \bar{U}\left(\left(s_1', s_2'\right), \left(s_1', s_2'\right)\right)$. These implications contradict the assumption that $\left(s_1^*, s_2^*\right)$ is neutrally stable.

Thus, we are left with the case in which there exist $t_1 < t_2 \in \mathcal{T}$, such that the restricted normal-form game $\left(\left(supp\left(\gamma_1^{t_1}\right), supp\left(\gamma_2^{t_1}\right)\right), u\right)$ $\left(\left(\left(supp\left(\gamma_1^{t_2}\right), supp\left(\gamma_2^{t_2}\right)\right), u\right)\right)$ admits equilibrium $\beta = \left(\beta_1, \beta_2\right)$ $\left(\alpha = \left(\alpha_1, \alpha_2\right)\right)$ in which both players mix. Assume first that the LHS of (2) is greater than the RHS. Let $f_1^+$and f_2^+ be the functions defined in Definition 15 with respect to the mixed actions α_1 and α_2. Let s_i^+ $(\tilde{s}_i^+)$ be the strategy that induces an agent who follows it (when acting in the role of player i) (1) to play the mixed action β_i in round t_1, (2) to play his part of an arbitrary equilibrium in the restricted game $\left(\left(supp\left(\gamma_1^t\right), supp\left(\gamma_2^t\right)\right), u\right)$ in each round $t \notin \{t_1, t_2\}$, and (3) to play on the marginal the mixed equilibrium α_i in round t_2, but to condition his play on the values of a_i (his own action in round t_1) and σ_i (the signal he observed in round t_1); specifically, the agent is more (less) likely to play action a_i and less (more) likely to play action a_i' when $f_i^+\left(a^k, \sigma^k\right) = 1$. These changes in the probabilities of playing actions a_i and a_i' are adjusted, such that, after each history $h_i^{t_2}$, the mixture of the mixed action played by an agent who follows strategy s_i^+ and the mixed action played by an agent who follows strategy $\tilde{s}_i^+$ is α_i, i.e., for each action $\hat{a} \in A$, $0.5 \cdot s_i^+\left(h_i^{t_2}\right)\left(\hat{a}\right) + 0.5 \cdot \tilde{s}_i^+\left(h_i^{t_2}\right)\left(\hat{a}\right) = \alpha_i\left(\hat{a}\right)$.

Observe that the strategies s_i^+ and $\tilde{s}_i^+$ induce the same behavior in all rounds $t \neq t_2$. Let s_i^{mix} be the mixture of the strategies s_i^+ and $\tilde{s}_i^+$; i.e., $s_i^{mix} \equiv \alpha_i$ in round t_2, and s_i^{mix} coincides with s_i^+ and $\tilde{s}_i^+$ in each round $t \neq t_2$. Observe that s_i^{mix} induces an agent who follows it to play mixed equilibria in all rounds. This implies that $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^{mix}, s_2^{mix}\right)\right) \leq \bar{U}\left(\left(s_1^{mix}, s_2^{mix}\right), \left(s_1^{mix}, s_2^{mix}\right)\right)$. The fact that s_i^{mix} is a mixture of s_i^+ and $\tilde{s}_i^+$ (and that the three strategies coincide in all rounds $t \neq t_2$) implies that $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^{mix}, s_2^{mix}\right)\right) = 0.5 \cdot \bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^+, s_2^+\right)\right) + 0.5 \cdot \bar{U}\left(\left(s_1^*, s_2^*\right), \left(\tilde{s}_1^+, \tilde{s}_2^+\right)\right)$. This implies that either $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^+, s_2^+\right)\right) \leq \bar{U}\left(\left(s_1^{mix}, s_2^{mix}\right), \left(s_1^{mix}, s_2^{mix}\right)\right)$ or $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(\tilde{s}_1^+, \tilde{s}_2^+\right)\right) \leq \bar{U}\left(\left(s_1^{mix}, s_2^{mix}\right), \left(s_1^{mix}, s_2^{mix}\right)\right)$. Assume without loss of generality that $\bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^+, s_2^+\right)\right) \leq \bar{U}\left(\left(s_1^{mix}, s_2^{mix}\right), \left(s_1^{mix}, s_2^{mix}\right)\right)$.

Consider a homogeneous group of mutants, each following strategy $\left(s_1^+, s_2^+\right)$. The definition of $\left(s_1^+, s_2^+\right)$ and the fact that $\left(s_1^*, s_2^*\right)$ is belief-free imply that $\bar{U}\left(\left(s_1^+, s_2^+\right), \left(s_1^*, s_2^*\right)\right) = \bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^*, s_2^*\right)\right)$, and that $\bar{U}\left(\left(s_1^+, s_2^+\right), \left(s_1^+, s_2^+\right)\right) > \bar{U}\left(\left(s_1^{mix}, s_2^{mix}\right), \left(s_1^{mix}, s_2^{mix}\right)\right) \geq \bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_1^+, s_2^+\right)\right)$.

The inequality $\bar{U}\left(\left(s_1^+, s_2^+\right), \left(s_1^+, s_2^+\right)\right) > \bar{U}\left(\left(s_1^{mix}, s_2^{mix}\right), \left(s_1^{mix}, s_2^{mix}\right)\right)$ holds because strategy s_i^+ coincides with strategy s_i^{mix} in any period $t \neq t_2$. In period t_2 agents who follow strategy $\left(s_1^+, s_2^+\right)$ achieve a higher expected payoff when being matched with other agents who follow strategy s^+ because when the former agents are matched they induce a positive correlation in their random play of the actions a_i and a_i', which increases their average payoff, due to the LHS of (2) being greater than the RHS, relative to the uncorrelated profile played by agents who follow the strategy $\left(s_1^{mix}, s_2^{mix}\right)$. This implies that $\left(s_1^*, s_2^*\right)$ is not neutrally stable.

If the LHS of (2) is less than the RHS, then we define analogous strategies $s_i^-$and $\tilde{s}_i^-$ with respect to the functions f_1^- and f_2^-, and we use an analogous argument to the one above where $s_i^-(\tilde{s}_i^-)$ replaces $s_i^+(\tilde{s}_i^+)$ and negative correlation replaces positive correlation in the random play of mutants in round t_2. □

A.3. Result for recursively strict asymmetric games

I conclude by extending Proposition 3 to the asymmetric setup. I say that a game is recursively strict, if all the games induced by restricting each player to choosing actions from a given subset of actions admit a strict equilibrium. Formally:

Definition 16. A normal-form game $G = ((A_1, A_2), u)$ is *recursively strict* if for any non-empty subset of actions $A_1' \subseteq A_1$ and $A_2' \subseteq A_2$, the game $G = \left(\left(A_1', A_2'\right), u\right)$, in which each player i is restricted to choosing actions from A_i', admits a strict equilibrium.

A couple of examples of recursively strict games are: (1) the (possibly asymmetric) Prisoner's Dilemma, (2) the (possibly asymmetric) public good game, (3) the (possibly asymmetric) Hawk–Dove game. Observe that a symmetric Hawk–Dove game is recursively strict in the current setup (in which players can condition their play on their role in the game), even though it is not recursively strict in the setup of symmetric games in which players cannot condition their play on their role in the game (see Section 3.4 above).

My last result shows that if the underlying game is recursively strict, then any belief-free equilibrium that satisfies weak stability is trivial and pure.

Proposition 6. *Assume that the underlying game $G = ((A_1, A_2), u)$ is recursively strict. Let $\left(s_1^*, s_2^*\right)$ be a belief-free equilibrium. If $\left(s_1^*, s_2^*\right)$ is weakly stable, then it is trivial and pure.*

Proof. Let $\gamma_i^t = \gamma_i^t\left(s_1^*, s_2^*\right) \in \Delta(A_i)$ be the marginal distribution of actions played by player i in period t in the belief-free equilibrium $\left(s_1^*, s_2^*\right)$. Assume first that $\gamma_i^t(s^*)$ is pure for each player i and each period t. This implies that $\left(s_1^*, s_2^*\right)$ induces a deterministic play that is independent of the observed signals. Thus a player's best reply coincides with his myopic best reply, which implies that the pure action profile played in each period must be an equilibrium of the underlying game (i.e., $\left(s_1^*, s_2^*\right)$ is trivial and pure).

Otherwise, there exists period t such that $\left|supp\left(\gamma_i^t\left(s_1^*, s_2^*\right)\right)\right| > 1$ for some player i. For each period t, let $\left(a_{1,1}^t, a_{2,1}^t\right) \in A_1 \times A_2$ be a strict equilibrium in the game $\left(\left(supp\left(\gamma_1^t\left(s_1^*, s_2^*\right)\right), supp\left(\gamma_2^t\left(s_1^*, s_2^*\right)\right)\right), u\right)$. Let $\left(s_{1,1}, s_{2,1}\right)$ be the mutant strategy profile in which a mutant agent in the role of player i chooses action $a_{i,1}^t$ in each period t. The fact that $\left(s_1^*, s_2^*\right)$ is belief-free equilibrium implies that mutants who follow strategy profile $\left(s_{1,1}, s_{2,1}\right)$ best reply against $\left(s_1^*, s_2^*\right)$. The fact that each $\left(a_{1,1}^t, a_{2,1}^t\right)$ is a strict equilibrium in $supp\left(\gamma_1^t\left(s_1^*, s_2^*\right)\right) \times supp\left(\gamma_2^t\left(s_1^*, s_2^*\right)\right)$ implies that a mutant achieves a strictly higher expected payoff relative to the incumbents when facing another mutant, i.e., $\bar{U}\left(\left(s_{1,1}, s_{2,1}\right), \left(s_{1,1}, s_{2,1}\right)\right) > \bar{U}\left(\left(s_1^*, s_2^*\right), \left(s_{1,1}, s_{2,1}\right)\right)$, which implies that $\left(s_1^*, s_2^*\right)$ is vulnerable to $\left(s_{1,1}, s_{2,1}\right)$.

For each odd $k \geq 1$, let $a_{k+1,2}^t$ be the unique best reply against $a_{k,1}^t$ (the best reply is unique due to the assumption that the game is recursively strict), and let $a_{k+1,1}^t = a_{k,2}^t$. For each even $k \geq 2$, let $a_{k+1,1}^t$ be the unique best reply against $a_{k,2}^t$ and let $a_{k+1,2}^t = a_{k,1}^t$. Observe that there exists a minimal $1 \leq \bar{k} \leq n^2 + 1$, such that for each period t and each player i, $\left(a_{\bar{k},1}^t, a_{\bar{k},2}^t\right)$ is a strict equilibrium of the unrestricted game $((A_1, A_2), u)$. The proof of this observation is as follows. If such a minimal $\bar{k}$ does not exist, then there is a period t and a sequence of actions $\left(\left(a_{1,1}^t, a_{1,2}^t\right), ..., \left(a_{n^2+1,1}^t, a_{n^2+1,2}^t\right)\right)$ that includes a non-trivial cycle. Let $\left(a_1', a_2'\right)$ be a strict equilibrium in the restricted game $\left(\left(\left\{a^t_{1,1}, ..., a^t_{n+1,1}\right\}, \left\{a^t_{1,2}, ..., a^t_{n^2+1,2}\right\}\right), u\right)$. The definition of the sequence $\left(\left(a^t_{1,1}, a^t_{1,2}\right), ..., \left(a^t_{n^2+1,1}, a^t_{n^2+1,2}\right)\right)$ implies that either there is an odd k such that $a^t_{k,1} = a'_1$ or there is an even k such that $a^t_{k,2} = a'_2$. In both cases, the definition of $\left(a'_1, a'_2\right)$ implies that the sequence must continue to action profile $\left(a'_1, a'_2\right)$ and hence cannot move from $\left(a'_1, a'_2\right)$ to any other action profile in the domain $\left\{a^t_{1,1}, ..., a^t_{n+1,1}\right\} \times \left\{a^t_{1,2}, ..., a^t_{n+1,2}\right\}$, which contradicts the fact that there is a non-trivial cycle.

For each $2 \leq k \leq \bar{k}$, let $\left(s_{k,1}, s_{k,2}\right)$ be the strategy profile in which each agent in the role of player i chooses action $a^t_{k,i}$ in each period t. The definitions of the strategies $\left\{\left(s_{1,1}, s_{1,2}\right), ..., \left(s_{\bar{k},1}, s_{\bar{k},2}\right)\right\}$ imply that (1) each strategy profile $\left(s_{k,1}, s_{k,2}\right)$ is vulnerable to the strategy profile $\left(s_{k+1,1}, s_{k+1,2}\right)$, and (2) strategy profile $\left(s_{\bar{k},1}, s_{\bar{k},2}\right)$ is a pure strategy profile in which the players play a strict equilibrium of the underlying (unrestricted) game in each period t, which implies that $\left(s_{\bar{k},1}, s_{\bar{k},2}\right)$ is evolutionarily stable, and that $\left(s^*_1, s^*_2\right)$ is not weakly stable. □

References

Awaya, Y., Krishna, V., 2016. On communication and collusion. Am. Econ. Rev. 106 (2), 285–315.

Basu, K., 1994. The traveler's dilemma: paradoxes of rationality in game theory. Am. Econ. Rev. 84 (2), 391–395.

Benaïm, M., Hofbauer, J., Hopkins, E., 2009. Learning in games with unstable equilibria. J. Econ. Theory 144 (4), 1694–1709.

Bhaskar, V., 2000. The Robustness of Repeated Game Equilibria to Incomplete Payoff Information. University of Essex.

Bhaskar, V., Mailath, G.J., Morris, S., 2008. Purification in the infinitely-repeated prisoners' dilemma. Rev. Econ. Dyn. 11 (3), 515–528.

Bhaskar, V., Obara, I., 2002. Belief-based equilibria in the repeated prisoner's dilemma with private monitoring. J. Econ. Theory 102 (1), 40–69.

Compte, O., 1998. Communication in repeated games with imperfect private monitoring. Econometrica 66 (3), 597–626.

Deb, J., 2012. Cooperation and community responsibility: a folk theorem for repeated matching games with names. Available at SSRN 1213102.

Ely, J.C., Hörner, J., Olszewski, W., 2005. Belief-free equilibria in repeated games. Econometrica 73 (2), 377–415.

Ely, J.C., Välimäki, J., 2002. A robust folk theorem for the prisoner's dilemma. J. Econ. Theory 102 (1), 84–105.

Fudenberg, D., Levine, D., Maskin, E., 1994. The folk theorem with imperfect public information. Econometrica, 997–1039.

Fudenberg, D., Maskin, E., 1986. The folk theorem in repeated games with discounting or with incomplete information. Econometrica, 533–554.

Harsanyi, J.C., 1973. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. Int. J. Game Theory 2 (1), 1–23.

Heller, Y., Mohlin, E., 2016. Observations on cooperation. Unpublished.

Hörner, J., Olszewski, W., 2009. How robust is the folk theorem? Q. J. Econ. 124 (4), 1773–1814.

Kandori, M., 2002. Introduction to repeated games with private monitoring. J. Econ. Theory 102 (1), 1–15.

Kandori, M., 2009. Weakly Belief-Free Equilibria in Repeated Games with Private Monitoring. CIRJE F-Series CIRJE-F-491. CIRJE, Faculty of Economics, University of Tokyo.

Kandori, M., 2011. Weakly belief-free equilibria in repeated games with private monitoring. Econometrica 79 (3), 877–892.

Kandori, M., Matsushima, H., 1998. Private observation, communication and collusion. Econometrica 66 (3), 627–652.

Kandori, M., Obara, I., 2006. Efficiency in repeated games revisited: the role of private strategies. Econometrica 74 (2), 499–519.

Mailath, G.J., Morris, S., 2002. Repeated games with almost-public monitoring. J. Econ. Theory 102 (1), 189–228.

Mailath, G.J., Morris, S., 2006. Coordination failure in repeated games with almost-public monitoring. Theor. Econ. 1 (3), 311–340.

Mailath, G.J., Olszewski, W., 2011. Folk theorems with bounded recall under (almost) perfect monitoring. Games Econ. Behav. 71 (1), 174–192.

Mailath, G.J., Samuelson, L., 2006. Repeated Games and Reputations, vol. 2. Oxford University Press.

Matsushima, H., 1991. On the theory of repeated games with private information: part I: anti-folk theorem without communication. Econ. Lett. 35 (3), 253–256.

Matsushima, H., 2004. Repeated games with private monitoring: two players. Econometrica 72 (3), 823–852.

Matsushima, H., Tanaka, T., Toyama, T., 2013. Behavioral Approach to Repeated Games with Private Monitoring. CIRJE-F-879. University of Tokyo.

Maynard-Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press.

Maynard-Smith, J., Price, G.R., 1973. The logic of animal conflict. Nature 246, 15.

Miyagawa, E., Miyahara, Y., Sekiguchi, T., 2008. The folk theorem for repeated games with observation costs. J. Econ. Theory 139 (1), 192–221.

Obara, I., 2009. Folk theorem with communication. J. Econ. Theory 144 (1), 120–134.

Osborne, M.J., Rubinstein, A., 1994. A Course in Game Theory. MIT Press.

Peski, M., 2012. An anti-folk theorem for finite past equilibria in repeated games with private monitoring. Theor. Econ. 7 (1), 25–55.

Piccione, M., 2002. The repeated prisoner's dilemma with imperfect private monitoring. J. Econ. Theory 102 (1), 70–83.

Sekiguchi, T., 1997. Efficiency in repeated prisoner's dilemma with private monitoring. J. Econ. Theory 76 (2), 345–361.

Sugaya, T., 2015. Folk theorem in repeated games with private monitoring. Unpublished.

Sugaya, T., Takahashi, S., 2013. Coordination failure in repeated games with private monitoring. J. Econ. Theory 148 (5), 1891–1928.

Takahashi, S., 2010. Community enforcement when players observe partners' past play. J. Econ. Theory 145 (1), 42–62.

Taylor, P.D., 1979. Evolutionarily stable strategies with two types of player. J. Appl. Probab., 76–83.

Thomas, B., 1985. On evolutionarily stable sets. J. Math. Biol. 22 (1), 105–115.

Van Damme, E., 1991. Stability and Perfection of Nash Equilibria, vol. 339. Springer.

Van Veelen, M., 2012. Robustness against indirect invasions. Games Econ. Behav. 74 (1), 382–393.

Weibull, J.W., 1995. Evolutionary Game Theory. The MIT Press.

Yamamoto, Y., 2007. Efficiency results in N player games with imperfect private monitoring. J. Econ. Theory 135 (1), 382–413.

Yamamoto, Y., 2009. A limit characterization of belief-free equilibrium payoffs in repeated games. J. Econ. Theory 144 (2), 802–824.

Yamamoto, Y., 2012. Characterizing belief-free review-strategy equilibrium payoffs under conditional independence. J. Econ. Theory 147 (5), 1998–2027.

Yamamoto, Y., 2014. Individual learning and cooperation in noisy repeated games. Rev. Econ. Stud. 81 (1), 473–500.