Cooperation, but no reciprocity: Individual strategies in the repeated Prisoner's Dilemma

Breitmoser, Yves

4 October 2012

Cooperation, but no reciprocity: Individual strategies in the repeated Prisoner's Dilemma

Yves Breitmoser*
EUV Frankfurt (Oder)

October 4, 2012

Abstract

A recent advance in our understanding of repeated PDs is the detection of a threshold $\delta^\star$ at which laboratory subjects start to cooperate predictively. This threshold is substantially above the classic threshold "existence of Grim equilibrium" and has been characterized axiomatically by Blonski, Ockenfels, and Spagnolo (2011, BOS). In this paper, I derive its behavioral foundations. First, I show that the threshold is equivalent to existence of a "Semi-Grim" equilibrium $\sigma_{cc} > \sigma_{cd} = \sigma_{dc} > \sigma_{dd}$. It is cooperative ($\sigma_{cc} > 0.5$), non-reciprocal ($\sigma_{cd} = \sigma_{dc}$), and robust to imperfect monitoring ("belief-free"). Next, I show that the no-reciprocity condition $\sigma_{cd} = \sigma_{dc}$ also follows from robustness to random-utility perturbations (logit equilibrium). Finally, I re-analyze strategies in four recent experiments and find that the majority of subjects indeed plays Semi-Grim when it is an equilibrium strategy, which explains $\delta^\star$'s predictive success.

JEL–Codes: C72, C73, C92
Keywords: Repeated Prisoner's Dilemma, experiment, equilibrium selection, cooperative behavior, reciprocity, belief-free equilibria, robustness

*I thank Friedel Bolle for many helpful comments. Address: Postfach 1786, 15207 Frankfurt(Oder), Germany, email: breitmoser@euv-frankfurt-o.de, Telephone/Fax: +3355534 2291/2390

1 Introduction

The infinitely repeated Prisoner's Dilemma has received much interest in economic research. It is a proto-typical model of cooperation between self-interested agents, but it is notoriously resistant to equilibrium refinement, which severely obstructs reliable predictions. Early experimental work, e.g. Rapoport and Mowshowitz (1966), showed that subjects indeed cooperate, and the results of Axelrod (1980a,b) show that tit-for-tat (TFT) is a theoretically successful strategy in response to opponents with unknown strategies. However, TFT is an not a subgame-perfect equilibrium, and the related "perfect tit-for-tat" (PTFT) strategy $(\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}) = (1,0,0,1)$,[1] which tends to be even more effective in response to cooperative strategies (Nowak et al., 1993; Imhof et al., 2007), exists only under the strict condition that one round of punishment suffices to deter defection. Experimental research has shown that subjects cooperate robustly even if one round of punishment does not suffice and that $\sigma_{dd} \approx 0$ actually holds on average (see e.g. Table 1). Thus, only few, if any, subjects seem to play PTFT, and existence of PTFT equilibrium is not necessary for cooperation to be sustained in laboratory experiments.

In turn, existence of the Grim equilibrium $(1,0,0,0)$ is not sufficient for cooperation to be sustained (Dal Bo, 2005), and overall, puzzling strategy estimates such as $(\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}) = (0.81, 0.43, 0.37, 0.22)$ by Rapoport and Mowshowitz (1966), which relates to neither Grim, TFT, nor PTFT, suggested that cooperation between human players is as fuzzy a concept as predicted by Folk theorems (see e.g. Stahl et al., 1991). Surprisingly, however, cooperation between human players is not fuzzy, as two recent studies, Blonski et al. (2011, BOS) and Dal Bo and Fréchette (2011), showed. They analyzed play in a large variety of experimental treatments, and both concluded that the $\delta^\star$-criterion defined axiomatically by BOS is a *predictive* threshold for cooperation. In addition, the puzzling strategy relation $\sigma_{cc} > \sigma_{dc} \approx \sigma_{cd} > \sigma_{dd}$ of Rapoport and Mowshowitz (1966) can be found implicitly in the binary regression models of Bruttel and Kamecke (2012, Table 4), where it obtains for three ways of eliciting strategies (hot play, strategy method, and a Moore

[1] $\sigma_{s',s''}$ is the probability of cooperation if one's previous choice was $s' \in \{c,d\}$ and the opponent's choice was $s'' \in \{c,d\}$

Table 1: On average, $\sigma_{cc} > \sigma_{dc} \approx \sigma_{cd} > \sigma_{dd}$ obtains in most treatments of four recent experiments, and individually, at least 50% play Semi-Grim in treatments where Semi-Grim MPEs exist (otherwise, the Semi-Grim share is zero).

	Standard. parameters			Average behavior							Share	Individual classification (Section 5)			
Treatment	b	a	δ	$\hat{\sigma}_{cc}$		$\hat{\sigma}_{dc}$		$\hat{\sigma}_{cd}$		$\hat{\sigma}_{dd}$	$\hat{\sigma}_{cc} > 0.25$	Alw-Def	Grim	Semi-G	TFT
Blonski et al. (2011)															
1	3	2	0.75	0.902	$\gg$	0.288	$<$	0.399	$\gg$	0.02	0.475	0.31	0.14	0.52	0.03
2	1.25	1.12	0.75	1	$\gg$	0.333	$\approx$	0.238	$\gg$	0.003	0.075	0	0.7	0	0.3
3	2.5	1.5	0.5	0.917	$\gg$	0.19	$\approx$	0.095	$\gg$	0.011	0.3	1	0	0	0
4	2.5	1.5	0.75	0.947	$\gg$	0.188	$\approx$	0.211	$\gg$	0.027	0.333	0.54	0.24	0	0.21
5	2.5	1.5	0.875	0.989	$\gg$	0.282	$\approx$	0.323	$\gg$	0.026	0.525	0.29	0	0.59	0.12
6	1.43	1.29	0.5	0	$\approx$	0	$\approx$	0.118	$\approx$	0.007	0	0.8	0.2	0	0
7	1.43	1.29	0.75	0.977	$\gg$	0.372	$\approx$	0.279	$\gg$	0.004	0.2	0	0.46	0.54	0
8	1.43	1.29	0.875	0.967	$\gg$	0.205	$\approx$	0.289	$\gg$	0.017	0.4	0.4	0	0.6	0
9	2.4	1.8	0.75	0.927	$\gg$	0.196	$\approx$	0.196	$\gg$	0.021	0.7	0.22	0.31	0.47	0
10	4.67	3	0.75	0.88	$\gg$	0.277	$\approx$	0.192	$\gg$	0.042	0.725	0.17	0	0.73	0.1
Dal Bo and Fréchette (2011)															
1	2.92	1.54	0.5	0.665	$\gg$	0.46	$\gg$	0.252	$\gg$	0.037	0.364	1	0	0	0
2	2.92	1.54	0.75	0.732	$\gg$	0.384	$\approx$	0.406	$\gg$	0.057	0.682	0.47	0.15	0	0.38
3	2.92	2.15	0.5	0.553	$\gg$	0.273	$\approx$	0.316	$\gg$	0.092	0.5	0.45	0.17	0	0.37
4	2.92	2.15	0.75	0.927	$\gg$	0.51	$\approx$	0.443	$\gg$	0.124	0.921	0.08	0	0.87	0.06
5	2.92	2.77	0.5	0.828	$\gg$	0.227	$\ll$	0.431	$\gg$	0.074	0.696	0.35	0	0.52	0.13
6	2.92	2.77	0.75	0.943	$\gg$	0.311	$\approx$	0.383	$\gg$	0.146	1	0	0	0.85	0.15
Duffy and Ochs (2009), "random rematching" treatment															
	3	2	0.9	0.964	$\gg$	0.36	$\approx$	0.337	$\gg$	0.11	0.929	0	0	0.89	0.11
Fudenberg et al. (2012), treatments 1–5 are "noisy" PDs (actions are perturbed), treatment 6 is "no-noise"															
1	2.5	1.5	0.875	0.842	$\gg$	0.33	$\gg$	0.245	$\gg$	0.064	0.833				
2	3	2	0.875	0.872	$\gg$	0.417	$\approx$	0.42	$\gg$	0.108	0.885				
3	3.5	2.5	0.875	0.887	$\gg$	0.513	$\approx$	0.473	$\gg$	0.161	0.875				
4	5	4	0.875	0.911	$\gg$	0.453	$\approx$	0.469	$\gg$	0.159	0.8				
5	5	4	0.875	0.93	$\gg$	0.602	$\approx$	0.664	$\gg$	0.252	0.948				
6	5	4	0.875	0.971	$\gg$	0.425	$\approx$	0.5	$\gg$	0.074	0.917	0	0	0.61	0.39

Note: The payoff parameters are standardized to be of the same form as the PD in Fig. 1b. $\hat{\sigma}_{cc}, \hat{\sigma}_{cd}, \hat{\sigma}_{dc}, \hat{\sigma}_{dd}$ are the relative frequencies of cooperation in the four states (across the all subjects, not counting the first rounds of each game). The relation signs indicate the p-values of Fisher tests on these relative frequencies ("$\gg, \ll$" indicate $\alpha < .01$, "$>, <$" indicate $\alpha < .05$, and "$\approx$" indicate insignificance). The relative frequencies of $\sigma_{cd,dc}$ have been pooled in tests against either σ_{cc} or σ_{dd} (as σ_{cd}, σ_{dc} mostly do not differ significantly).

Individual classification: The population weights of the various strategies preview the estimation results of Section 5. Shares of inexistent or insignificant components are set to zero. *Always Defect* is $(\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}) = (0,0,0,0)$ and *Grim* is $(1,0,0,0)$. Tit-for-tat classification contains TFT itself $(1,0,1,0)$, "weak" perfect TFT $(1,0,0,0.5)$, and Always-Cooperate $(1,1,1,1)$; below, their shares are also reported separately. Semi-Grim contains regular Semi-Grim $\approx (1,0.3,0.3,0)$ and belief-free Semi-Grim $\approx (0.9,0.5,0.5,0.1)$; both as defined below (the actual frequencies depend on the treatment parameters).

procedure following Engle-Warnick and Slonim, 2004, 2006), and similarly in many more experiments—although it had never been reported anymore. To examine its generality, I re-analyzed behavior in four recent experiments, and Table 1 shows that the average strategy satisfies $\sigma_{cc} > \sigma_{dc} = \sigma_{cd} > \sigma_{dd}$ in most treatments, including Prisoner's Dilemmas with exogenous noise. These results show that reliable point predictions concerning strategies and cooperation actually are possible, but since $\sigma_{cc} > \sigma_{dc} = \sigma_{cd} > \sigma_{dd}$ does not relate to a known strategy and $\delta^{\star}$ has been defined by means of axioms on the set of games rather than the set of strategies (see Def. 2.1 below), it is unclear whether these findings admit a behavioral interpretation.

The present paper extends this work in two ways. First, I show that the BOS axioms are equivalent to existence of what I propose to call "Semi-Grim" equilibrium, and the Semi-Grim strategy is indeed the puzzling construct $\sigma_{cc} > \sigma_{dc} = \sigma_{cd} > \sigma_{dd}$ observed by Rapoport and Mowshowitz (1966) and in Table 1. Second, I show in a latent class analysis (the results of which are previewed in Table 1) that the Semi-Grim strategy is played by the majority of subjects when it is an equilibrium. In turn, only a minority plays always-defect, always-cooperate, Grim, tit-for-tat, or similar strategies when the Semi-Grim MPE exists.

Combined, this shows that Semi-Grim is predictive at both, the aggregate level and the individual level, and since it is the Markov perfect equilibrium corresponding to the BOS axioms, this explains their predictiveness. I also show that the "no-reciprocity" condition $\sigma_{cd} = \sigma_{dc}$ underlying Semi-Grim equally follows from "Markov logit equilibrium", i.e. by requiring robustness with respect to random utility perturbations in the sense of McKelvey and Palfrey (1995), which lends further validity to this deviation from TFT. Finally, I identify two kinds of Semi-Grim equilibria, namely a regular one, which admits purification (Bhaskar et al., 2008; Doraszelski and Escobar, 2010), and a fully mixed one, which is included in the belief-free equilibria constructed by Ely and Välimäki (2002). Belief-free equilibria have received much attention in the recent theoretical literature (Ely et al., 2005; Hörner and Olszewski, 2006; Hörner and Lovo, 2009; Fudenberg and Yamamoto, 2010; Kandori, 2011), as they are robust to private monitoring, such as the possibility that the opponent forgets the current state. Thus, both kinds of Semi-Grim equilibria are normatively plausible, which seems to explain their joint occurrence in the four considered experiments.

Implications of these results are discussed in Section 6. Aside from this, Section 2 introduces notation and basic definitions, Section 3 reviews belief-free equilibria, Section 4 theoretically analyzes Semi-Grim equilibria, and Section 5 estimates the individual strategies in recent experiments. Proofs are relegated to the appendix.

2 Preliminary remarks

Notation The players are denoted as $i, j \in N = \{1,2\}$. They play an infinitely repeated Prisoner's dilemma (PD) as defined in Figure 1a. Any such PD can be transformed into the "standardized" form Figure 1b, and the best-known examples of PDs are even "simple" in the sense of Figure 1c, with $a = 2$ or $a = 3$. These particular PDs are simple, as (i) playing (c,c) is socially efficient (i.e. it maximizes the sum of payoffs), and (ii) the gain from defecting in response to a cooperating opponent exactly offsets the loss when cooperating in response to a defecting opponent (which simplifies many algebraic expressions on repeated PDs).

The set of actions of the constituent game is $S = \{c,d\}$, and the set of states of the repeated game is $S \times S$. A Markov strategy $\sigma \in [0,1]^{S\times S}$ maps each state to a probability of choosing c in that state.[2] For example, $\sigma_{s',s''}$ denotes the probability that the respective player cooperates conditional on $s' \in S$ being his previous action and $s'' \in S$ being his opponent's previous action. I focus on strategy profiles that are symmetric between players, and those will be denoted as (σ,σ). This assumption is standard in analyses of experimental data, and it allows me to drop the player index in denoting strategies. Note the implied inversion of indices, however, i.e. if player i cooperates with probability $\sigma_{s',s''}$ in state (s',s''), then his opponent $j \neq i$ considers the same state to be (s'',s') and thus cooperates with probability $\sigma_{s'',s'}$.

[2]Such strategies are also known as 1-memory strategies (Barlo et al., 2009) and are special cases of strategies with bounded recall (Sabourian, 1998). Bruttel and Kamecke (2012) show that choices other than the previous ones (i.e. lag 2 and higher) are often insignificant, although they are not entirely irrelevant.

Figure 1: Prisoner's dilemma (PD) games (with $p_{dc} > p_{cc} > p_{dd} > p_{cd}$ and $b > a > 1$)

(a) The "General" PD

	c	d
c	p_{cc}, p_{cc}	p_{cd}, p_{dc}
d	p_{dc}, p_{cd}	p_{dd}, p_{dd}

(b) "Standardized" PD

	c	d
c	a, a	$0, b$
d	$b, 0$	$1, 1$

(c) The "Simple" PD

	c	d
c	a, a	$0, a+1$
d	$a+1, 0$	$1, 1$

Payoffs and equilibria Given strategy profile (σ, σ), the expected payoff of choosing c in state $(s', s'') \in S \times S$ is denoted as $\pi_{s',s''}(c)$, the expected payoff of choosing d is denoted as $\pi_{s',s''}(d)$, and the expected payoff overall is $\pi_{s',s''}$ in state (s', s''). These payoffs satisfy, for all $(s', s'') \in S \times S$,

$$\pi_{s',s''} = \sigma_{s',s''} \cdot \pi_{s',s''}(c) + (1 - \sigma_{s',s''}) \cdot \pi_{s',s''}(d) \tag{1}$$

$$\pi_{s',s''}(c) = \sigma_{s'',s'} \cdot \big(\delta\pi_{cc} + (1-\delta)\, p_{cc}\big) + (1 - \sigma_{s'',s'}) \cdot \big(\delta\pi_{cd} + (1-\delta)\, p_{cd}\big) \tag{2}$$

$$\pi_{s',s''}(d) = \sigma_{s'',s'} \cdot \big(\delta\pi_{dc} + (1-\delta)\, p_{dc}\big) + (1 - \sigma_{s'',s'}) \cdot \big(\delta\pi_{dd} + (1-\delta)\, p_{dd}\big). \tag{3}$$

Solving the Equation system (1)–(3) for $(\pi_{s',s''})$ over all states yields the expected payoffs as functions of σ. This is algebraically straightforward, but the resulting expressions are fairly cumbersome (see e.g. Lemma A.1 in the appendix for the respective solutions of simple PDs). Finally, a strategy profile (σ, σ) is a Markov perfect equilibrium (MPE) if for all $(s', s'') \in S \times S$,

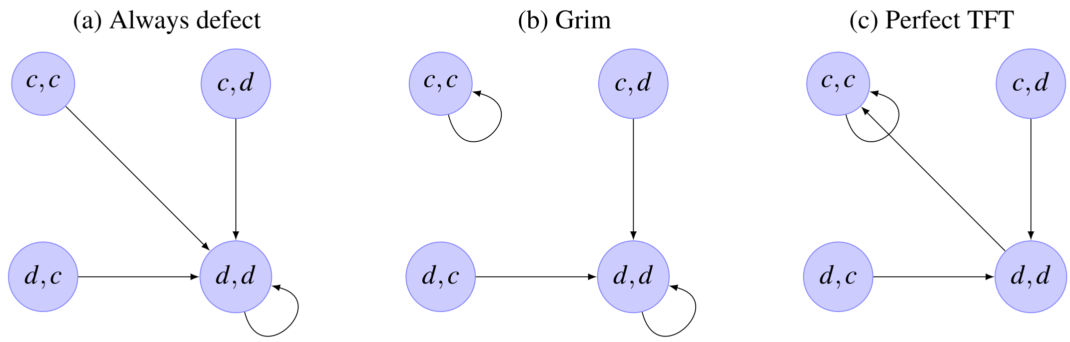
$$\sigma_{s',s''} > 0 \Rightarrow \pi_{s',s''}(c) \geq \pi_{s',s''}(d) \quad \text{and} \quad \sigma_{s',s''} < 1 \Rightarrow \pi_{s',s''}(c) \leq \pi_{s',s''}(d). \tag{4}$$

Figure 2 reviews the three well-known pure MPEs that may exist in repeated PDs. The always-defect MPE exists in general. The Grim MPE $(\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}) = (1, 0, 0, 0)$ exists if

$$p_{cc} \geq (1-\delta)\, p_{dc} + \delta\, p_{dd} \qquad \Leftrightarrow \qquad \delta \geq \frac{p_{dc} - p_{cc}}{p_{dc} - p_{dd}} =: \delta_{\text{Grim}}. \tag{5}$$

An equilibrium sustaining (at least) temporary cooperation along the path of play

Figure 2: Pure Markov perfect equilibria



exists if and only if the Grim equilibrium exists. Thus, the condition $\delta \geq \delta_{\text{Grim}}$ is theoretically necessary for cooperation to be sustained in equilibrium. Since always-defect is an equilibrium in general, $\delta > \delta_{\text{Grim}}$ is not theoretically sufficient.

A threshold for "sufficient" patience Roth and Murnighan (1978) show that cooperation increases if Grim and PTFT equilibria exist, and Murnighan and Roth (1983) as well as Dal Bo (2005) confirm that the discount rate is of major relevance with respect to the emergence of cooperative play. Applied research therefore often assumes that players select cooperative equilibria if δ exceeds a given threshold, usually the threshold for existence of Grim (see Blonski et al., 2011, for further discussion). Dal Bo and Fréchette (2011) and Blonski et al. (2011, BOS) specifically designed experiments to test the hypothesis that existence of Grim equilibrium is sufficient in practice. They found that it is not sufficient, while the cooperation rate can be predicted well based on the discount factor's relation to a higher threshold characterized axiomatically by BOS. To review the BOS axioms, let the tuple $\langle p_{cc}, p_{cd}, p_{dc}, p_{dd}, \delta \rangle$ denote a repeated "general" PD, let $\overline{\Gamma}$ denote the set of repeated PDs, and let $\overline{\Gamma}^c \subset \overline{\Gamma}$ denote the subset of repeated PDs where players "likely" cooperate. BOS characterize $\overline{\Gamma}^c$ through three "parsimonious" axioms, which they argue to be generally plausible, and two "comprehensive" axioms, which ensure uniqueness of the selection criterion, i.e. of the implied δ-threshold.

Definition 2.1 (BOS axioms)**.** $\overline{\Gamma}^c$ satisfies all of the following conditions.

*A*1 *(Positive linear payoff transformation invariance).* For all $\tau(x) = \alpha x + \beta$, $\alpha > 0$, $\langle p_{cd}, p_{dc}, p_{cc}, p_{dd}, \delta\rangle \in \overline{\Gamma}^c$ implies $\langle \tau(p_{cd}), \tau(p_{dc}), \tau(p_{cc}), \tau(p_{dd}), \delta\rangle \in \overline{\Gamma}^c$.

*A*2 *(δ-monotonicity).* If $p_{dc} > p_{cc} > p_{dd} > p_{cd}$ and $p_{dc} + p_{cd} < 2\,p_{cc}$, then there exists $\delta^\star(p_{cc}, p_{cd}, p_{dc}, p_{dd}) \in (0,1]$ such that for all δ': $\langle p_{cc}, p_{cd}, p_{dc}, p_{dd}, \delta'\rangle \in \overline{\Gamma}^c \;\Leftrightarrow\; \delta' > \delta^\star(p_{cc}, p_{cd}, p_{dc}, p_{dd})$.

*A*3 *(Boundary conditions).* $p_{cd} \to -\infty \;\Rightarrow\; \delta^\star(p_{cc}, p_{cd}, p_{dc}, p_{dd}) \to 1$ and $p_{cd} \to p_{dd} \;\Rightarrow\; \delta^\star(p_{cc}, p_{cd}, p_{dc}, p_{dd}) \to (p_{dc} - p_{cc})/(p_{dc} - p_{dd})$.

*A*4 *(Incentive independence).* There exists an additively separable function $\mu(x_1, x_2, x_3)$ such that $\langle p_{cd}, p_{dc}, p_{cc}, p_{dd}, \delta\rangle \in \overline{\Gamma}^c$ iff $\mu(p_{cc} - p_{dd}, p_{dc} - p_{cc}, p_{dd} - p_{cd}) \geq 0$.

*A*5 *(Equal weight).* $\langle p_{cd}, p_{dc}, p_{cc}, p_{dd}, \delta\rangle \in \overline{\Gamma}^c$ if and only if $\langle p'_{cd}, p'_{dd}, p_{cc}, p_{dd}, \delta\rangle \in \overline{\Gamma}^c$ with $p_{dc} - p_{cc} = p_{dd} - p'_{cd}$ and $p'_{dd} - p_{cc} = p_{dd} - p_{cd}$.

Proposition 2.2 (BOS, Proposition 2)**.** *If $\overline{\Gamma}^c$ satisfies A1–A5 from Definition 2.1, then*

$$\langle p_{cd}, p_{dc}, p_{cc}, p_{dd}, \delta\rangle \in \overline{\Gamma}^c \quad \Leftrightarrow \quad \delta \geq \frac{p_{dc} + p_{dd} - p_{cd} - p_{cc}}{p_{dc} - p_{cd}} =: \delta_{BOS}. \tag{6}$$

For a comprehensive discussion of these axioms, let me refer to Blonski et al. (2011). Briefly, the first two axioms require invariance with respect to linear payoff transformations (*A*1) and that more patient players are not less likely to cooperate (*A*2), which implies that the selection criterion can be characterized by a δ-threshold. *A*3 requires that players stop cooperating if the risk gets high ($p_{cd} \to -\infty$) and that they always cooperate if there is no risk ($p_{cd} \to p_{dd}$), conditional on the existence of Grim equilibria. The other two axioms ensure uniqueness of the δ-threshold by restricting the relevance of the long-run gains from cooperation $(p_{cc} - p_{dd})$, in relation to that of the short-run gains from defection $(p_{dc} - p_{cc})$ and to the risk of cooperation $(p_{dd} - p_{cd})$. *A*4 requires additive separability of the selection criterion with respect to these incentives, and *A*5 requires equality of the implied weights of the two short-run incentives. The δ-thresholds from existence of Grim and PTFT satisfy *A*4 but not *A*5.

The axioms do not bear an obvious relation to a strategy or an equilibrium concept. In this way, they differ from other axiomatic theories of equilibrium selection, such as Govindan and Wilson (2006, 2012). As established below, requiring $A1-A5$ is equivalent to requiring existence of belief-free Semi-Grim MPEs, i.e. of equilibria corresponding with the observations in Table 1.

3 Belief-free equilibria in relation to the BOS axioms

Given strategy profile (σ,σ), define the *cooperation incentive* in state $(s',s'') \in S \times S$ to be the difference of expected payoffs from one-time cooperation and one-time defection $\tilde{\pi}_{s',s''} := \pi_{s',s''}(c) - \pi_{s',s''}(d)$, with continuation play evolving according to σ. The player is strictly best off cooperating in state (s',s'') if $\tilde{\pi}_{s',s''} > 0$, he is best off defecting if $\tilde{\pi}_{s',s''} < 0$, and he randomizes only if $\tilde{\pi}_{s',s''} = 0$. In "simple" repeated PDs (Figure 1c), the differences of the cooperation incentives in the four states satisfy

$$\tilde{\pi}_{cc} - \tilde{\pi}_{cd} = (\sigma_{dc} - \sigma_{cc}) \cdot \mu \qquad \tilde{\pi}_{cc} - \tilde{\pi}_{dc} = (\sigma_{cd} - \sigma_{cc}) \cdot \mu \tag{7}$$

$$\tilde{\pi}_{cc} - \tilde{\pi}_{dd} = (\sigma_{dd} - \sigma_{cc}) \cdot \mu \qquad \tilde{\pi}_{cd} - \tilde{\pi}_{dc} = (\sigma_{cd} - \sigma_{dc}) \cdot \mu \tag{8}$$

with $\mu = \delta(1-\delta)(a-1)(\sigma_{dc} + \sigma_{cd} - \sigma_{cc} - \sigma_{dd})/r$ and $r > 0$. Thus, if $\mu = 0$, i.e. if $\sigma_{dc} + \sigma_{cd} = \sigma_{cc} + \sigma_{dd}$ in simple PDs, then $\tilde{\pi}_{cc} = \tilde{\pi}_{cd} = \tilde{\pi}_{dc} = \tilde{\pi}_{dd}$. If $\tilde{\pi}_{cc} = 0$ holds in addition, the underlying strategy profile (σ,σ) is a fully mixed MPE, i.e. an MPE where players are indifferent in all states. These are the "robust" equilibria derived by Ely and Välimäki (2002). As players are always indifferent, their best responses are independent of their beliefs about the opponent's history (in case the history is not common knowledge, i.e. under private monitoring). Hence, such *belief-free MPEs* (Ely et al., 2005) are robust to private monitoring, which is not the case for strict MPEs (Kandori, 2002). I refer to these equilibria also as belief-free MPEs, although the repeated game considered here is theoretically one of perfect monitoring (the subjects in the experiments may disagree, of course).

Solving Eqs. (7)–(8) for σ_{cd}, σ_{dc} yields a two-dimensional manifold of MPEs (as previously observed by Bhaskar et al., 2008). Proposition 3.1 derives these equilibrium strategies and their existence conditions for standardized PDs. It refines the results of Ely and Välimäki (2002) and Bhaskar et al. (2008), who characterize strategy and existence condition implicitly, in relation to the terms π_{cc} and π_{cd} defined in Eq. (1). Proposition 3.1 eliminates these endogenous entities, which will allow us to relate the existence condition to the BOS criterion.

Proposition 3.1. *In any standardized PD with* $1/\delta < a < b < (a-\delta)/(1-\delta)$*, a two-dimensional manifold of belief-free MPEs exists. It consists of all strategy profiles* (σ,σ) *satisfying*

$$\sigma_{cd} = \frac{(a-b)\,\delta\sigma_{dd} + (b-1)\,\delta\sigma_{cc} + a - b}{(a-1)\,\delta}, \quad \sigma_{dc} = \frac{a\delta\sigma_{dd} - \delta\sigma_{cc} + 1}{(a-1)\,\delta}. \tag{9}$$

All MPEs that are not belief-free in the above sense are locally isolated, and hence they are regular and finitely many (Doraszelski and Escobar, 2010). Thus, almost all MPEs are belief-free (if the latter exist). Now, their existence condition $1/\delta < a < b < (a-\delta)/(1-\delta)$ can be alternatively expressed as

$$\delta \geq \min\left\{\frac{1}{a}, \frac{b-a}{b-1}\right\} \qquad \Leftrightarrow \qquad \delta \geq \min\left\{\frac{p_{dd}-p_{cd}}{p_{cc}-p_{cd}}, \frac{p_{dc}-p_{cc}}{p_{dc}-p_{dd}}\right\} =: \delta_{\text{BF-MPE}}. \tag{10}$$

This condition satisfies the "parsimonious" BOS axioms $A1, A2, A3$, but it violates $A4$, since the minimum function is not additively separable, and as can be verified easily, it also violates weight equality $A5$. Thus, requiring existence of belief-free MPEs violates both comprehensive axioms of BOS, whereas existence of Grim or PTFT violated only $A4$. The latter violation is comparably minor, though, as each branch of the criterion in (10) is additively separable in the way it is required. Thus, $A4$ holds locally for almost all games if the criterion was "existence of belief-free MPEs". Axiom $A5$, equality of weights of defection gains $(p_{dc}-p_{cc})$ and cooperation risk $(p_{dd}-p_{cd})$, continues to be violated though, as for the Grim condition $\delta \geq \delta_{\text{Grim}}$. This violation is obvious in the latter case, since p_{cd} is strategically irrelevant in Grim equilibria, and the violation obtains similarly in all equilibria with strict cooperation in any state. Then, p_{cd} is irrelevant in that state, and the implicit weights of $p_{dc}-p_{cc}$ and $p_{dd}-p_{cd}$ cannot be equal. In belief-free equilibria, in turn, players randomize in all states, i.e. all possible outcomes are relevant. Due to the asymmetry $\sigma_{cd} \neq \sigma_{dc}$, however, the implicit weights of "defection gains" and "co-operation risks" still do not equate exactly. This will change in belief-free equilibria where the two players are equally likely to cooperate in all states, as shown next.

4 Semi-Grim equilibria

I call a strategy "Semi-Grim" if it satisfies $\sigma_{cc} > \sigma_{cd} = \sigma_{dc} > \sigma_{dd}$, as observed in most treatments in Table 1. First, I show that the central sub-condition $\sigma_{cd} = \sigma_{dc}$ follows from robustness to random utility perturbations. To be precise, conditional on $\sigma_{cc} > 0.5 > \sigma_{dd}$, logit equilibrium implies $\sigma_{cd} = \sigma_{dc}$. Logit equilibrium is a special case of quantal response equilibrium (McKelvey and Palfrey, 1995) and extends to dynamic games as "Markov logit equilibrium" (as defined in Breitmoser et al., 2010). Logit equilibrium has been shown to explain experimental observations in many circumstances, including the centipede game (Fey et al., 1996), traveler's dilemma (Capra et al., 1999), auctions (Goeree et al., 2002b), public goods games (Goeree et al., 2002a), monotone contribution games (Choi et al., 2008), and beauty contests (Breitmoser, 2012). Thus, it is a plausible starting point for explaining behavior also in repeated games. Formally, a strategy profile (σ, σ) is a Markov logit equilibrium (MLE) if there exists $\lambda \in \mathbb{R}_+$ such that for all $(s', s'') \in S \times S$,

$$\sigma_{s',s''} = \frac{\exp\{\lambda \cdot \pi_{s',s''}(c)\}}{\exp\{\lambda \cdot \pi_{s',s''}(c)\} + \exp\{\lambda \cdot \pi_{s',s''}(d)\}}. \tag{11}$$

Rearranging Eq. (11) yields the alternative expression $\log\big((1 - \sigma_{s',s''})/\sigma_{s',s''}\big) = \lambda \cdot \big(\pi_{s',s''}(d) - \pi_{s',s''}(c)\big)$. Thus, in MLE, differing cooperation rates, e.g. $\sigma_{cc} > \sigma_{dd}$, require corresponding differences in cooperation incentives, e.g. $\tilde{\pi}_{cc} > \tilde{\pi}_{dd}$. To establish the aforementioned claim on MLEs, it therefore suffices to show that $\tilde{\pi}_{cc} \neq \tilde{\pi}_{dd}$ and $\tilde{\pi}_{cd} \neq \tilde{\pi}_{dc}$ cannot be satisfied simultaneously. The following proposition does that, and thus the general observation $\sigma_{cc} > 0.5 > \sigma_{dd}$ theoretically implies $\sigma_{dc} = \sigma_{cd}$ (as observed). In addition, it shows that the alternative class of "alternating" equilibria $\sigma_{dc} \neq \sigma_{cd}$ require $\sigma_{dc} > \sigma_{cc}$. As $\sigma_{dc} > \sigma_{cc}$ has never been observed in experiments, I therefore conclude that logit equilibrium implies $\sigma_{dc} = \sigma_{cd}$.

Proposition 4.1. *Let* (σ,σ) *be an MLE of a repeated PD with* $\max\{\sigma_{cc},\sigma_{cd},\sigma_{dc},\sigma_{dd}\} > 0.5$. *Then,*

1. $\sigma_{cd} \neq \sigma_{dc}$ *implies* $\sigma_{cd} < \sigma_{cc} = \sigma_{dd} < \sigma_{dc}$,
2. $\sigma_{cc} \neq \sigma_{dd}$ *implies* $\sigma_{cd} = \sigma_{dc} < \sigma_{cc}$.

With the additional restriction $\sigma_{cd} = \sigma_{dc}$, the two-dimensional manifold of belief-free MPEs reduces to a one-dimensional manifold of *belief-free Semi-Grim* MPEs satisfying $0 < \sigma_{dd} < \sigma_{dc} = \sigma_{cd} < \sigma_{cc} < 1$. The next proposition establishes that belief-free Semi-Grim MPEs exist if and only if the BOS axioms are satisfied (for comparability with BOS, the result is established for repeated "general" PDs).

Proposition 4.2. *In any repeated "general" PD, a one-dimensional manifold of belief-free Semi-Grim MPEs exists iff* $\delta > \delta_{BOS}$. *It consists of all strategy profiles* (σ,σ) *satisfying*

$$\sigma_{dd} = \frac{(p_{dc}-p_{cd})\,\delta\,\sigma_{cc} - p_{dd} - p_{dc} + p_{cd} + p_{cc}}{\delta\,(p_{dc}-p_{cd})}, \tag{12}$$

$$\sigma_{dc} = \sigma_{cd} = \frac{(p_{dc}-p_{cd})\,\delta\,\sigma_{cc} - p_{dc} + p_{cc}}{\delta\,(p_{dc}-p_{cd})}. \tag{13}$$

Eqs. (12), (13) yield a strategy profile (i.e. probabilities) if $\sigma_{cc} \geq \delta_{\mathrm{BOS}}/\delta$. At the threshold $\delta = \delta_{\mathrm{BOS}}$, a mixed Semi-Grim MPE satisfying $\sigma_{cc} = 1$, $\sigma_{cd} = \sigma_{dc} \in (0,1)$, and $\sigma_{dd} = 0$ appears, and considering the average behavior reviewed in Table 1, this seems to relate closely to the Markov strategy played by the cooperating players. This will be verified in detail in the next section.

Bhaskar et al. (2008) argue that robustness to imperfect monitoring (i.e. being belief-free) may not be the only plausible criterion for equilibrium selection. If a mixed equilibrium does not admit purification, then there is no reason why players should randomize in the specific way $\sigma_{cc} > \sigma_{dc} = \sigma_{cd} > \sigma_{dd}$ prescribed by the equilibrium. After all, the players are indifferent in all states. The standard argument justifying the mixed equilibrium is based on purification, but Bhaskar et al. show that the belief-free equilibria constructed above do not admit purification (in Markov strategies with one-period memory). In turn, Doraszelski and Escobar (2010) show that locally isolated MPEs satisfy a *regularity* condition that implies purifiability in Markov strategies.

Next, I show that (and when) such a *regular Semi-Grim MPE* (σ,σ) exists, i.e. an MPE satisfying $\sigma_{cc} = 1$, $\sigma_{cd} = \sigma_{dc} \in (0,1)$, $\sigma_{dd} = 0$ and inducing strict cooperation incentives $\tilde{\pi}_{cc} > 0$ and $\tilde{\pi}_{dd} < 0$ in the states (c,c) and (d,d). The strictness of these constraints implies local isolation, and thus regularity and purifiability (Doraszelski and Escobar, 2010). Regular Semi-Grim exists under the same conditions as belief-free Semi-Grim equilibrium if $p_{dc} + p_{cd} \geq p_{cc} + p_{dd}$ and under slightly weaker conditions otherwise.

Proposition 4.3. *A regular Semi-Grim MPE exists for all $\delta > \delta_{BOS}$ in general, and if $p_{cc} + p_{dd} > p_{dc} + p_{cd}$, then also for all*

$$\delta > 1 - \frac{\sqrt{2\sqrt{p_{cc}-p_{cd}}\sqrt{p_{dc}-p_{cc}}\sqrt{p_{dd}-p_{cd}}\sqrt{p_{dc}-p_{dd}}+(p_{dc}+p_{cd}-2p_{cc})p_{dd}+(p_{cc}-2p_{cd})p_{dc}+p_{cc}p_{cd}}}{p_{dc}-p_{cd}}. \tag{14}$$

The case $p_{dc} + p_{cd} \geq p_{cc} + p_{dd}$ is particularly illustrative. If $\delta = \delta_{\text{BOS}}$, the aforementioned equilibrium $\sigma_{cc} = 1$, $\sigma_{cd} = \sigma_{dc} \in (0,1)$, and $\sigma_{dd} = 0$ exists, then with $\tilde{\pi}_{cc} = \tilde{\pi}_{dd} = 0$. As δ increases, the belief-free MPE inducing $\tilde{\pi}_{cc} = \tilde{\pi}_{dd} = 0$ moves into the interior of the strategy space, while the corner solution turns into the strict, regular Semi-Grim MPE. If $p_{dc} + p_{cd} < p_{cc} + p_{dd}$, in turn, the gains from short-term defections are comparably low, and in this case, regular Semi-Grim MPEs exist under weaker conditions than belief-free Semi-Grim MPEs. The corresponding existence condition (14) clearly violates axiom $A4$, additive separability of the incentives, but interestingly, it also violates axiom $A5$, weight equality, despite the symmetry condition $\sigma_{s',s''} = \sigma_{s'',s'}$ for all s', s''. The reason is the same as with Grim. Since players are strictly best off cooperating in state (c,c), the payoff p_{cd} is irrelevant there, and hence $p_{dc} - p_{cc}$ and $p_{dd} - p_{cd}$ do not have equal weight overall.

Finally, note that the set of belief-free Semi-Grim MPEs characterized in Prop. 4.2 and the regular Semi-Grim MPE characterized in Prop. 4.3 are the only Semi-Grim MPEs, i.e. the only MPEs with the structure $\sigma_{cc} > \sigma_{cd} = \sigma_{dc} > \sigma_{dd}$. For, any other MPE with this structure would require randomization in state (c,c) or (d,d), with $\mu \neq 0$ in Eqs. (7), (8) to separate it from belief-free MPEs. For example, assume there exists an MPE with $\sigma_{cc} = 1 > \sigma_{cd} = \sigma_{dc} > \sigma_{dd} > 0$. This requires $\tilde{\pi}_{cd} = \tilde{\pi}_{dc} = \tilde{\pi}_{dd}$, and by $\mu \neq 0$ in Eqs. (7), (8), this implies $\sigma_{cd} = \sigma_{dc} = \sigma_{dd}$, contradicting $\sigma_{dc} > \sigma_{dd}$. Similar contradictions obtain in the other two cases, and thus, the set of Semi-Grim MPEs has been characterized completely.

5 Estimation of individual strategies

Table 1 shows that the average strategy is Semi-Grim in most treatments, and the previous section has shown that existence of the (belief-free) Semi-Grim MPE is equivalent to the predictive BOS criterion. Next, I show that the majority of individual subjects indeed uses Semi-Grim strategies when the respective equilibria exist. This establishes that the average behavior in Table 1 is not a weighted sum of entirely unrelated strategies, but of Semi-Grim strategies as claimed.

The econometric approach and the considered strategies closely follows Dal Bo and Fréchette (2011). The considered strategies are always-defect, always-cooperate, Grim, TFT, and a reciprocal strategy with prolonged punishment $(\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}) = (1, 0, 0, 0.5)$. The latter strategy is a Markov variant of the T2 strategy considered by Dal Bo and Fréchette (2011), and it it is a weakened version of perfect TFT $(1, 0, 0, 1)$, which is theoretically plausible[3] but in its strict form not identified in the data (see Dal Bo and Fréchette, 2011, and Fudenberg et al., 2012). In addition, I consider regular Semi-Grim and belief-free Semi-Grim (of the linear continuum of belief-free Semi-Grim MPEs, the median one is taken).

Econometric model

Similarly to Dal Bo and Fréchette (2011), we have to account for noise in the analysis. I assume that subjects of a given type play their equilibrium action with probability $1-\gamma$, $\gamma \in (0, 1)$, and that they randomize uniformly with probability γ (in each

[3]PTFT is theoretically more effective than TFT if most subjects play cooperative strategies (Nowak et al., 1993) and (Imhof et al., 2007), as original TFT struggles to restore cooperation after unilateral defection. PTFT is also called “win-stay, lose-shift”.

round, taking independent draws). This approach estimates the same population weights as Dal Bo and Fréchette's approach if only pure equilibria are considered, and it generalizes their approach straightforwardly to mixed MPEs. Econometrically, the population is described as a mixture of a finite set K of components, where for all $k \in K$, members of component k cooperate with (perturbed) probability $\sigma_\omega(k)$ in state ω and they have weight $\rho(k)$ in the population. Using $o_{s,t} \in \{0,1\}$ and $\omega_{s,t} = \{\emptyset, cc, cd, dc, dd\}$ to denote choice and state of the decision number t of subject $s \in S$, where $o_{s,t} = 1$ denotes cooperation and $o_{s,t} = 0$ denotes defection, the log-likelihood of the model is

$$LL = \sum_{s \in S} \log \sum_{k \in K} \prod_t \left(\sigma_{\omega_{s,t}}(k)\right)^{o_{s,t}} \cdot \left(1 - \sigma_{\omega_{s,t}}(k)\right)^{1-o_{s,t}}. \tag{15}$$

The likelihood is maximized jointly over all parameters using the robust, gradient-free NEWUOA algorithm (Powell, 2006), and convergence has been verified using a Newton-Raphson algorithm. Thus, standard errors can be taken from the information matrix (McLachlan and Peel, 2000). I consider all four experiments reviewed in Table 1, but in the case of Fudenberg et al. (2012), I focus on the "standard" treatment without exogenous noise (as equilibrium predictions are not available otherwise). In all cases, I use the observations from the second halves of the experiments, i.e. when individual behavior has largely stabilized, which follows Dal Bo and Fréchette (2011) and Fudenberg et al. (2012).

The set of relevant strategies (i.e. the model dimension) is estimated as follows. First, I eliminate all components with equilibria that do not exist, and similarly, I eliminate the non-equilibrium strategies Always-Coop, TFT, T05 if Grim is not an equilibrium. Second, I eliminate components with insignificant weights. Here I use the conservative "BIC criterion" and eliminate a component if its elimination does not increase $BIC = -LL + \log n \cdot D/2$ (with n as number of subjects and D as number of parameters). In our case, this seems sufficient, as it leaves us with just 2-4 components in most cases. In general, alternative approaches may suggest the elimination of further components (Biernacki et al., 2000).

Finally, actions in round 1 are not prescribed uniquely in MPEs and there usually are three possible equilibria for round-1 behavior. The standard assumption is that players treat the initial state equivalently to the way they treat (c,c), but I intend to estimate the weights of the various MPEs without testing the joint assumption on MPE and round-1 behavior. For this reason, I allow that all players may either be "cautious", and cooperate with a maximal probability of $\sigma_\emptyset$ in round 1, or not cautious, and cooperate with probability σ_{cc} in round 1. Regardless of whether they are cautious in round 1 or not, all players are assumed to play their (perturbed) equilibrium strategies in all subsequent rounds. Both $\sigma_\emptyset$ and the weight ρ_{Caut} of cautious players are estimated from the data.

Results

Table 2 presents the estimated strategy weights for the four experiments discussed before, and supplementing it, Table 3 shows which equilibria exist in the various treatments and how the Semi-Grim predictions relate to the respective choices of the average "cooperating" subject, i.e. to the subjects with $\hat{\sigma}_{cc} > 0.25$. The remaining subjects (i.e. those with $\hat{\sigma}_{cc} \leq 0.25$) usually play always-defect, i.e. their relation to Semi-Grim is irrelevant. The equilibrium predictions for the remaining states σ_{cc} and σ_{dd} are equal or close to 1 and 0 (respectively) in all cases, which is as observed. The main observations can be summarized first; their discussion follows.

Result 5.1. *If Semi-Grim equilibria exist, then the majority of subjects plays Semi-Grim. Otherwise, the majority plays Always-defect or Grim. The reciprocal strategies TFT and T05 are assigned to minorities in either case: to 10%–20% of the subjects if Semi-Grim equilibria exist, and to 20%–40% if not.*

Thus, the results clearly support the hypothesis that the majority of subjects plays Semi-Grim, which followed from Table 1. Most importantly, the inclination to play Semi-Grim actually increases as the overall inclination to cooperate increases. To see this, look at the four treatments where subjects cooperate most consistently in state (c,c), i.e. at the treatments where the share of "cooperating subjects" ($\hat{\sigma}_{cc} > 0.25$) is at least 80%. By Table 3, these are the treatments DF4, DF6, DO, and FRD. In three of these four treatments, more than 80% of the subjects play Semi-Grim, i.e. regular or belief-free Semi-Grim. In the remaining treatments, between 47% and 73% of the subjects play Semi-Grim if it exists, i.e. the majority

Table 2: Weights of the various strategies in the four experiments

(a) Duffy and Ochs (2009)

A-Def	A-Coop	Grim	BF-SG	Reg-SG	TFT	T05	γ	$\sigma_\emptyset$	ρ_{Caut}	LL
–	–	–	0.5 (0.07)	0.39 (0.08)	0.11 (–)	–	0.01 (0)	0.1 (0.04)	0.29 (0.06)	−946.3

(b) Fudenberg et al. (2012), treatment 6 (“no noise”)

A-Def	A-Coop	Grim	BF-SG	Reg-SG	TFT	T05	γ	$\sigma_\emptyset$	ρ_{Caut}	LL
–	0.2 (0.07)	–	0.24 (0.06)	0.37 (0.08)	0.19 (–)	–	0.01 (0)	0.06 (0.05)	0.17 (0.06)	−429.1

(c) Blonski et al. (2011)

Treat	A-Def	A-Coop	Grim	BF-SG	Reg-SG	TFT	T05	γ	$\sigma_\emptyset$	ρ_{Caut}	LL
Agg	0.29 (0.03)	0 (0)	0.15 (0.03)	0.21 (0.04)	0.23 (0.04)	0.11 (0.02)	0.01 (–)	0.03 (0)	0.33 (0.02)	0.6 (0.04)	−3115.9
1	0.31 (0.08)	0.03 (0.02)	0.14 (0.07)	0.52 (–)	–	–	–	0.01 (0)	0.17 (0.04)	0.55 (0.1)	−412.2
2	–	–	0.7 (0.13)	–	–	0.3 (–)	–	0.01 (0)	0.02 (0.01)	0.97 (0.02)	−115
3	1 (–)	–	–	–	–	–	–	0.14 (0.02)	1.21 (*NaN*)	0.51 (*NaN*)	−168.6
4	0.54 (0.07)	–	0.24 (0.07)	–	–	0.18 (0.06)	0.03 (–)	0.04 (0.01)	0.37 (0.04)	0.81 (0.08)	−542.1
5	0.29 (0.09)	–	–	0.29 (0.1)	0.3 (0.12)	0.12 (–)	–	0.01 (0.01)	0.27 (0.06)	0.59 (0.1)	−368.9
6	0.8 (0.09)	–	0.2 (–)	–	–	–	–	0.02 (0.01)	0.57 (0.17)	0.63 (0.33)	−53.3
7	–	–	0.46 (0.12)	–	0.54 (–)	–	–	0.01 (0)	0.04 (0.01)	0.88 (0.05)	−161.1
8	0.4 (0.09)	–	–	0.04 (0.04)	0.56 (–)	–	–	0.02 (0)	0.38 (0.07)	0.66 (0.11)	−278.7
9	0.22 (0.07)	–	0.31 (0.11)	0.11 (0.08)	0.36 (–)	–	–	0.02 (0.01)	0.4 (0.05)	0.49 (0.09)	−359
10	0.17 (0.06)	–	–	0.25 (0.1)	0.48 (0.1)	0.1 (–)	–	0.05 (0.01)	0.5 (0.05)	0.52 (0.1)	−516.6

(d) Dal Bo and Fréchette (2011)

Treat	A-Def	A-Coop	Grim	BF-SG	Reg-SG	TFT	T05	γ	$\sigma_\emptyset$	ρ_{Caut}	LL
Agg	0.14 (0.03)	0.03 (0.02)	0.14 (0.04)	0.2 (0.04)	0.17 (0.06)	0.24 (0.04)	0.07 (–)	0.08 (0)	0.33 (0.01)	0.49 (0.04)	−9435.8
1	1 (–)	–	–	–	–	–	–	0.1 (0)	0.35 (*NaN*)	0.5 (*NaN*)	−1412.7
2	0.47 (0.08)	–	0.15 (0.07)	–	–	0.38 (–)	–	0.09 (0)	0.32 (0.02)	0.83 (0.08)	−1530.4
3	0.45 (0.07)	–	0.17 (0.06)	–	–	0.25 (0.07)	0.12 (–)	0.08 (0.01)	0.29 (0.01)	0.96 (0.04)	−2216.8
4	0.08 (0.04)	–	–	0.87 (0.06)	–	–	0.06 (–)	0.01 (0)	0.18 (0.02)	0.2 (0.07)	−1112.8
5	0.35 (0.07)	–	–	0.41 (0.08)	0.11 (0.06)	0.13 (–)	–	0.03 (0)	0.37 (0.02)	0.58 (0.09)	−1953.4
6	–	–	–	0.1 (0.05)	0.75 (0.09)	–	0.15 (–)	0.03 (0)	0.7 (0.03)	0.39 (0.09)	−1008.2

Legend: “A-Def” is Always-Defect $(\sigma_{cc},\sigma_{cd},\sigma_{dc},\sigma_{dd}) = (0,0,0,0)$, “A-Coop” is Always-Cooperate $(1,1,1,1)$, Grim is $(1,0,0,0)$, “BF-SG” is belief-free Semi-Grim $\approx (0.9,0.5,0.5,0.1)$ (depending on treatment parameters, see also Prop. 4.2), “Reg-SG” is regular Semi-Grim $\approx (1,0.3,0.3,0)$ (depending on treatment parameters, see also Prop. 4.3), “TFT” is $(1,0,1,0)$, and “T05” is a TFT strategy with prolonged punishment $(1,0,0,0.5)$. Subjects of a given type are assumed to play always-defect if their strategy is not a best response to itself (i.e. if it is not an equilibrium).

γ is the noise parameter (probability of randomizing uniformly instead of playing the strategy), ρ_{Caut} is the share of cautious players, and $\sigma_\emptyset$ is the probability of cooperation of cautious players in round 1.

Table 3: Behavior of the "cooperating" subjects ($\hat{\sigma}_{cc} > 0.25$) in relation to the δ-thresholds in the various treatments

		δ-thresholds			Threshold met			Behavior of "cooperators"				Semi-Grim pred. $\sigma_{cd,dc}$	
Treatment	δ	δ_{Grim}	$\delta_{\text{BF-SG}}$	$\delta_{\text{Reg-SG}}$	δ_{Grim}	$\delta_{\text{BF-G}}$	$\delta_{\text{Reg-SG}}$	Share	$\hat{\sigma}_{cc}$	$\hat{\sigma}_{dd}$	$\hat{\sigma}_{cd,dc}$	Regular	Belief-free
Blonski et al. (2011)													
1	0.75	0.5	0.667	0.667	×	×	×	0.475	0.92	0.043	0.44	0.256	0.5
2	0.75	0.52	0.904	0.865	×			0.075	1	0.014	0.375	-	-
3	0.5	0.667	0.8	0.8				0.3	0.917	0.024	0	-	-
4	0.75	0.667	0.8	0.8	×			0.333	0.982	0.063	0.326	-	-
5	0.875	0.667	0.8	0.8	×	×	×	0.525	0.993	0.046	0.449	0.228	0.5
6	0.5	0.326	0.797	0.741	×			0	0	0	0	-	-
7	0.75	0.326	0.797	0.741	×		×	0.2	0.977	0	0.424	0.667	-
8	0.875	0.326	0.797	0.741	×	×	×	0.4	0.98	0.017	0.455	0.329	0.843
9	0.75	0.429	0.667	0.661	×	×	×	0.7	0.956	0.017	0.201	0.299	0.611
10	0.75	0.455	0.572	0.572	×	×	×	0.725	0.903	0.04	0.201	0.132	0.405
Dal Bo and Fréchette (2011)													
1	0.5	0.719	0.815	0.815				0.364	0.701	0.108	0.457	-	-
2	0.75	0.719	0.815	0.815	×			0.682	0.755	0.082	0.48	-	-
3	0.5	0.401	0.606	0.605	×			0.5	0.625	0.151	0.432	-	-
4	0.75	0.401	0.606	0.605	×	×	×	0.921	0.927	0.161	0.492	0.214	0.553
5	0.5	0.078	0.394	0.343	×	×	×	0.696	0.832	0.129	0.429	0.333	0.789
6	0.75	0.078	0.394	0.343	×	×	×	1	0.943	0.146	0.34	0.137	0.693
Duffy and Ochs (2009)													
	0.9	0.5	0.667	0.667	×	×	×	0.929	0.969	0.127	0.351	0.094	0.5
Fudenberg et al. (2012), "no-noise" treatment													
6	0.875	0.25	0.4	0.4	×	×	×	0.917	0.973	0.098	0.555	0.041	0.5

Note: The "δ-thresholds" refer to the minimal δ such that the respective equilibria exist (Grim, Belief-free Semi-Grim, Regular Semi-Grim), see Eqs. (5), (6), (14). In addition, shares and average strategies of "cooperators" (subjects with $\hat{\sigma}_{cc} > 0.25$) are provided, and the cooperation rates $\sigma_{dc} = \sigma_{cd}$ predicted by the two Semi-Grim equilibria.

of subjects across all experiments. The strategies of the subjects not playing Semi-Grim in these cases (i.e. when Semi-Grim exists) depend on the experiment. In Fudenberg et al. (2012), they play heuristics such as Always-Cooperate or TFT, in Blonski et al. (2011) they play Always-Defect or Grim, and in Dal Bo and Fréchette (2011) they play Always-Defect or TFT/T05. Thus, the observation that most subjects play Semi-Grim (when it is an equilibrium) is the only observation common to all experiments, and it shows that the majority of subjects adapt Semi-Grim strategies regardless of how the minority of "other players" behaves.

Finally, look at the strategies of subjects when Semi-Grim equilibria do not exist. These are the strategies in the treatments BOS 2–4,6 and DF 1–3. In these cases, the majority of subjects is classified as Always-Defect or Grim (note that the weight of Semi-Grim is set to zero in the cases, as the mixed equilibrium strategies cannot be computed if the equilibria do not exist). As Table 3 shows, in treatments BOS 2,4 and DF 2,3, Grim equilibria exist, and the discount factor δ is about 0.1 below the threshold for existence of Semi-Grim. In these cases, the average cooperating subject already plays Semi-Grim $\sigma_{cc} > \sigma_{cd,dc} > \sigma_{dd}$, with $\hat{\sigma}_{cd,dc} \in (0.3, 0.5)$ as reported in Table 3. These cooperation probabilities are similar to those in treatments where Semi-Grim equilibria exist, which in turn are equal to the Semi-Grim predictions in these cases, but due to non-existence of Semi-Grim equilibria, the cooperating subjects are to be classified as playing Grim or TFT in these cases. Arguably, these subjects actually play Semi-Grim ε-equilibria, which suggests that the weights of Grim and TFT are overestimated in these cases. An evaluation of such mixed ε-equilibria is left as further research, however.

6 Conclusion

This paper proposed a novel explanation of behavior in repeated Prisoner's Dilemmas that fits both subjects' strategies and the BOS criterion. Accordingly, subjects play a mixed, non-reciprocal Semi-Grim strategy $\sigma_{cc} > \sigma_{dc} \approx \sigma_{cd} > \sigma_{dd}$. This strategy closely fits choices in four recent experiments, i.e. both average and individual behavior in these experiments, and it relates closely to several recent developments in related literature: axiomatic equilibrium selection in repeated PDs (Blonski et al., 2011), robustness to imperfect monitoring (Ely and Välimäki, 2002) and purifiability (Doraszelski and Escobar, 2010), and Markov logit equilibrium (McKelvey and Palfrey, 1995; Breitmoser et al., 2010). The results appear to be very robust, as the majority of subjects plays Semi-Grim strategies whenever a Semi-Grim equilibrium exists, and the weights of Semi-Grim strategies is actually the largest (above 80%) in treatments where most subjects cooperate. These results further strengthen the observation that subjects' behavior in repeated PDs seems to be predictable, following Blonski et al. (2011) and Dal Bo and Fréchette (2011) who showed that the emergence of cooperation is predictable. This positive result appears to be very promising in relation to the embarrassment of riches implied by Folk theorems, suggesting that substantial equilibrium selection actually takes place in repeated games.

At the same time, the analysis departs from the literature following Axelrod (1980a,b), which focused on reciprocal strategies, i.e. Markov strategies satisfying either $\sigma_{dc} > \sigma_{cd}$ or $\sigma_{dd} > \sigma_{cd}$. The support for this department is fairly strong, as perfect TFT is not played by subjects (see Dal Bo and Fréchette, 2011, Fudenberg et al., 2012, and the weight of the T05 strategy in 2), and as TFT is played by few, if any, subjects. For, if TFT would have substantial weight, then $\sigma_{dc} \neq \sigma_{cd}$ should be significant, since *all* otherwise discussed strategies (including Semi-Grim) imply $\sigma_{dc} = \sigma_{cd}$.

Finally, this paper has been the first to consider usage of mixed strategies such as Semi-Grim in econometric analyses of repeated PDs, and thus also the first to show that subjects play belief-free equilibria. Hence, there is ample opportunity to extend this research. In particular, it appears to be most interesting to see how predictive non-reciprocal, mixed equilibria are with respect other repeated games. While the concept of belief-free equilibria generalizes straightforwardly to other constituent games, based on the results of Ely et al. (2005), a generalization of the no-reciprocity condition $\sigma_{dc} = \sigma_{cd}$ does not seem to be available immediately. This may be an obstacle in such generalizations, but the observed relation to logit equilibrium may be of help here. In light of the above results, however, further research along these lines seems warranted.

References

Axelrod, R. (1980a). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(1):3–25.

Axelrod, R. (1980b). More effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(3):379–403.

Barlo, M., Carmona, G., and Sabourian, H. (2009). Repeated games with one-memory. *Journal of Economic Theory*, 144(1):312–336.

Bhaskar, V., Mailath, G., and Morris, S. (2008). Purification in the infinitely-repeated prisoners' dilemma. *Review of Economic Dynamics*, 11(3):515–528.

Biernacki, C., Celeux, G., and Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(7):719–725.

Blonski, M., Ockenfels, P., and Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, 3(3):164–192.

Breitmoser, Y. (2012). Strategic reasoning in p-beauty contests. *Games and Economic Behavior*, 75(2):555–569.

Breitmoser, Y., Tan, J., and Zizzo, D. (2010). Understanding perpetual r&d races. *Economic Theory*, 44(3):445–467.

Bruttel, L. and Kamecke, U. (2012). Infinity in the lab. how do people play repeated games? *Theory and Decision*, 72(2):205–219.

Capra, C., Goeree, J., Gomez, R., and Holt, C. (1999). Anomalous behavior in a traveler's dilemma? *American Economic Review*, 89(3):678–690.

Choi, S., Gale, D., and Kariv, S. (2008). Sequential equilibrium in monotone games: A theory-based analysis of experimental data. *Journal of Economic Theory*, 143(1):302–330.

Dal Bo, P. (2005). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review*, 95(5):1591–1604.

Dal Bo, P. and Fréchette, G. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101(1):411–429.

Doraszelski, U. and Escobar, J. F. (2010). A theory of regular markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification. *Theoretical Economics*, 5:369–402.

Duffy, J. and Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, 66(2):785–812.

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

Ely, J. and Välimäki, J. (2002). A robust folk theorem for the prisoner's dilemma. *Journal of Economic Theory*, 102(1):84–105.

Engle-Warnick, J. and Slonim, R. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior and Organization*, 55(4):553–573.

Engle-Warnick, J. and Slonim, R. (2006). Inferring repeated-game strategies from actions: evidence from trust game experiments. *Economic Theory*, 28(3):603–632.

Fey, M., McKelvey, R., and Palfrey, T. (1996). An experimental study of constant-sum centipede games. *International Journal of Game Theory*, 25(3):269–287.

Fudenberg, D., Rand, D., and Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*.

Fudenberg, D. and Yamamoto, Y. (2010). Repeated games where the payoffs and monitoring structure are unknown. *Econometrica*, 78(5):1673–1710.

Goeree, J., Holt, C., and Laury, S. (2002a). Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics*, 83(2):255–276.

Goeree, J., Holt, C., and Palfrey, T. (2002b). Quantal response equilibrium and overbidding in private-value auctions. *Journal of Economic Theory*, 104(1):247–272.

Govindan, S. and Wilson, R. (2006). Sufficient conditions for stable equilibria. *Theoretical Economics*, 1(2):167–206.

Govindan, S. and Wilson, R. (2012). Axiomatic equilibrium selection for generic two-player games. *Econometrica*, 80(4):1639–1699.

Hörner, J. and Lovo, S. (2009). Belief-free equilibria in games with incomplete information. *Econometrica*, 77(2):453–487.

Hörner, J. and Olszewski, W. (2006). The folk theorem for games with private almost-perfect monitoring. *Econometrica*, 74(6):1499–1544.

Imhof, L., Fudenberg, D., and Nowak, M. (2007). Tit-for-tat or win-stay, lose-shift? *Journal of Theoretical Biology*, 247(3):574–580.

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

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

McKelvey, R. and Palfrey, T. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38.

McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. Wiley series in probability and statistics.

Murnighan, J. and Roth, A. (1983). Expecting continued play in prisoner's dilemma games. *Journal of Conflict Resolution*, 27(2):279–300.

Nowak, M., Sigmund, K., et al. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58.

Powell, M. (2006). The newuoa software for unconstrained optimization without derivatives. *Large-Scale Nonlinear Optimization*, pages 255–297.

Rapoport, A. and Mowshowitz, A. (1966). Experimental studies of stochastic models for the prisoner's dilemma. *Behavioral Science*, 11(6):444–458.

Roth, A. and Murnighan, J. (1978). Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psychology*, 17(2):189–198.

Sabourian, H. (1998). Repeated games with *m*-period bounded memory (pure strategies). *Journal of Mathematical Economics*, 30(1):1–35.

Stahl, D. et al. (1991). The graph of prisoners' dilemma supergame payoffs as a function of the discount factor. *Games and Economic Behavior*, 3(3):368–384.

A Proofs

Proof of Proposition 3.1. In standardized PDs, it is cumbersome but straightforward to verify that $\tilde{\pi}_{s',s''} := \pi_{s',s''}(c) - \pi_{s',s''}(d)$ for all $s', s'' \in S$ satisfy Eqs. (7)–(8) with

$$\mu = \frac{(1-\delta)\,(\delta\,(b\sigma_{dd} - \sigma_{dc} - \sigma_{cd} - b\sigma_{cc} + 2\sigma_{cc}) - a\delta\,(2\sigma_{dd} - \sigma_{dc} - \sigma_{cd}) + b - a - 1)}{\substack{1-\delta\left(\sigma_{dd}^2 - 2\sigma_{dd} - 2\sigma_{cd}\sigma_{dc} + \sigma_{dc} + \sigma_{cd} + \sigma_{cc}^2\right) \\ -\delta^2\,(\sigma_{dd} - \sigma_{cc})\,(\sigma_{dc}\sigma_{dd} + \sigma_{cd}\sigma_{dd} - 2\sigma_{cc}\sigma_{dd} - 2\sigma_{cd}\sigma_{dc} + \sigma_{cc}\sigma_{dc} + \sigma_{cc}\sigma_{cd})}}.$$

Thus, $\mu = 0$ implies $\tilde{\pi}_{cc} = \tilde{\pi}_{cd} = \tilde{\pi}_{dc} = \tilde{\pi}_{dd}$. Solving its enumerator for σ_{dc} yields

$$\sigma_{dc} = -\frac{(b-2a)\,\delta\sigma_{dd} + (a-1)\,\delta\sigma_{cd} + (2-b)\,\delta\sigma_{cc} + b - a - 1}{(a-1)\,\delta} \tag{16}$$

and substituting this for σ_{cd} in $\tilde{\pi}_{cc} = 0$, and straightforward but tedious algebraic manipulations, yields σ_{cd} as claimed above. Substituting this for σ_{cd} in Eq. (16), finally, this yields σ_{dc} as claimed above. Finally, the set of strategy profiles satisfying these constraints is not empty if $\sigma_{cd} > 0$ and $\sigma_{dc} < 1$, while $\sigma_{cd} < 1$ and $\sigma_{dc} > 0$

hold true in any standardized PD. At the critical point $\sigma_{cc} = 1, \sigma_{dd} = 0$, $\sigma_{dc} < 1$ is equivalent to $a > 1/\delta$ and $\sigma_{cd} > 0$ is equivalent to $b < (a-\delta)/(1-\delta)$. $\square$

Proof of Proposition 4.1. Recall $\tilde{\pi}_{s',s''}$ as defined in Eqs. (7)–(8) and that $\log\left[(1-\sigma_{s',s''})/\sigma_{s',s''}\right] = -\lambda\tilde{\pi}_{s',s''}$ in all (s',s'') if (σ,σ) is an MLE.

1. First, by $\tilde{\pi}_{cd} - \tilde{\pi}_{dc} = (\sigma_{cd} - \sigma_{dc}) \cdot \mu$, $\sigma_{dc} \neq \sigma_{cd}$ implies $\mu > 0$. For, $\mu = 0$ implies $\tilde{\pi}_{cd} = \tilde{\pi}_{dc}$ and thus $\sigma_{dc} = \sigma_{cd}$ in MLE, while in case $\mu < 0$, $\sigma_{cd} \gtrless \sigma_{dc}$ implies $\tilde{\pi}_{cd} \lesseqgtr \tilde{\pi}_{dc}$ and thus $\sigma_{cd} \lesseqgtr \sigma_{dc}$, a contradiction. Next, by $\mu > 0$, $\sigma_{dc} \lesseqgtr \sigma_{cc}$ implies $\tilde{\pi}_{cc} \lesseqgtr \tilde{\pi}_{cd}$ and thus $\sigma_{cc} \lesseqgtr \sigma_{cd}$, and vice versa. Thus, either $\sigma_{cd} < \sigma_{cc} = \sigma_{dd} < \sigma_{dc}$ or $\sigma_{dc} < \sigma_{cc} = \sigma_{dd} < \sigma_{cd}$. It remains to show $\sigma_{cd} < \sigma_{dc}$. For contradiction, assume $\sigma_{dc} < \sigma_{cc} = \sigma_{dd} < \sigma_{cd}$. If $\sigma_{cc} = \sigma_{dd}$, $\tilde{\pi}_{cd}$ is falling in σ_{cd} and increasing in σ_{dc}, and in the limiting case $\sigma_{dc} = \sigma_{cc} = \sigma_{dd} = \sigma_{cd}$, it evaluates to

   $$\tilde{\pi}_{cd} = -(d-1)\left(-b\sigma_{cc} + a\sigma_{cc} + \sigma_{cc} - 1\right).$$

   Thus, $\tilde{\pi}_{cd} < 0$, which implies $\sigma_{cd} < 0.5$ and thus contradicts the assumption that $\max\{\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}\} > 0.5$.

2. First, by $\tilde{\pi}_{cc} - \tilde{\pi}_{dd} = (\sigma_{dd} - \sigma_{cc}) \cdot \mu$, $\sigma_{cc} \neq \sigma_{dd}$ implies $\mu < 0$. For, $\mu = 0$ implies $\sigma_{cc} = \sigma_{dd}$, and in case $\mu > 0$, $\sigma_{cc} \gtrless \sigma_{dd}$ implies $\tilde{\pi}_{cc} \lesseqgtr \tilde{\pi}_{dd}$ and thus $\sigma_{cc} \lesseqgtr \sigma_{dd}$ (contradiction). In turn, $\mu < 0$ implies $\sigma_{cd} = \sigma_{dc}$, as $\sigma_{cd} \neq \sigma_{dc}$ implies $\mu > 0$ by the argument made in point 1. It remains to show that $\sigma_{dc} < \sigma_{cc}$, or equivalently $\tilde{\pi}_{cc} < \tilde{\pi}_{dc}$. Using $\sigma_{cd} = \sigma_{dc}$, $\tilde{\pi}_{cc} - \tilde{\pi}_{dc}$ simplifies toward

   $$\tilde{\pi}_{cc} - \tilde{\pi}_{dc} = \frac{(d-1)(\sigma_{dc}-\sigma_{cc})(d(b\sigma_{dd}-2a\sigma_{dd}+2a\sigma_{dc}-2\sigma_{dc}-b\sigma_{cc}+2\sigma_{cc})+b-a-1)}{d\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{dc}^2+2\sigma_{dc}+\sigma_{cc}^2\right)+2d^2(\sigma_{dc}-\sigma_{cc})(\sigma_{dd}-\sigma_{cc})(\sigma_{dd}-\sigma_{dc})-1}.$$

   For contradiction assume $\sigma_{dc} \geq \sigma_{cc}$. The denominator of the fraction is generally decreasing in σ_{dd}, and in the limiting case $\sigma_{dd} = 0$ it is

   $$2d\left(d\sigma_{cc} - 1\right)\sigma_{dc}^2 - 2d\left(d\sigma_{cc}^2 - 1\right)\sigma_{dc} + d\sigma_{cc}^2 - 1 < 0.$$

   Thus, it is generally negative. The numerator of the right-hand side is negative if

$$\sigma_{dd} < \frac{(2a-2)\,d\,\sigma_{dc} + (2-b)\,d\,\sigma_{cc} + b - a - 1}{(2a-b)\,d} =: \tilde{\sigma}_{dd}.$$

Thus, in case $\sigma_{dd} < \tilde{\sigma}_{dd}$, $\tilde{\pi}_{cc} - \tilde{\pi}_{dc}$ is positive, contradicting the initial assumption $\sigma_{cc} < \sigma_{dc}$. Alternatively, in case $\sigma_{dd} \geq \tilde{\sigma}_{dd}$, the cooperation incentive $\tilde{\pi}_{dc}$ is decreasing in σ_{dd}, and in the limiting case $\sigma_{dd} = \tilde{\sigma}_{dd}$,

$$\tilde{\pi}_{dc} = \frac{(d-1)\,(b\,d\,\sigma_{dc} - b\,d\,\sigma_{cc} + b - a)}{2d\,\sigma_{dc} - 2d\,\sigma_{cc} + 1}.$$

Thus, $\tilde{\pi}_{dc} < 0$ follows if $\sigma_{dc} > \sigma_{cc}$. Since $\tilde{\pi}_{dc} < 0$ also implies $\sigma_{dc} < 0.5$, this contradicts $\max\{\sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}\} > 0.5$. □

Proof of Proposition 4.2. Eqs. (7)–(8) hold equivalently here, now with $\mu = r_1/r_2$ where

$$\begin{aligned} r_1 = \delta\,(p_{dc} + p_{cd})\,(\sigma_{dd} - \sigma_{cc}) - 2\delta\,p_{cc}\,(\sigma_{dd} - \sigma_{dc}) \\ - 2\delta\,p_{dd}\,(\sigma_{dc} - \sigma_{cc}) - p_{dd} + p_{dc} + p_{cd} - p_{cc} \end{aligned}$$

and $r_2 \neq 0$. Thus, $r_1 = 0$ again yields $\tilde{\pi}_{cc} = \tilde{\pi}_{cd} = \tilde{\pi}_{dc} = \tilde{\pi}_{dd}$. Solving $r_1 = 0$ for σ_{dc},

$$\sigma_{dc} = \frac{2\delta\,(p_{cc}\,\sigma_{dd} - p_{dd}\,\sigma_{cc}) - \delta\,(p_{dc} + p_{cd})\,(\sigma_{dd} - \sigma_{cc}) + p_{dd} - p_{dc} - p_{cd} + p_{cc}}{2\delta\,(p_{cc} - p_{dd})},$$

and substituting this into $\tilde{\pi}_{cc} = 0$ yields

$$\begin{aligned} \delta^2\,(p_{dc} - p_{cd})\,(\sigma_{dd} - \sigma_{cc}) - \delta\,(p_{dc} - p_{cd})\,(\sigma_{dd} - \sigma_{cc} - 1) \\ + \delta\,(p_{dd} - p_{cc}) - p_{dd} - p_{dc} + p_{cd} + p_{cc} = 0. \end{aligned}$$

Solving these two conditions for $(\sigma_{dd}, \sigma_{dc})$ yields Eqs. (12), (13). As for existence of these MPEs, $\sigma_{dd} \geq 0$ holds true (at $\sigma_{cc} = 1$) iff $\delta \geq (p_{dc} + p_{dd} - p_{cd} - p_{cc})/(p_{dc} - p_{cd})$, while $\sigma_{dd} < \sigma_{dc} \leq 1$ is satisfied for all $\sigma_{cc} \in [0, 1]$. □

Proof of Proposition 4.3. If $\sigma_{cc} = 1$, $\sigma_{cd} = \sigma_{dc}$, and $\sigma_{dd} = 0$, the cooperation incentive in state (d, c), $\tilde{\pi}_{dc} := \pi_{dc}(c) - \tilde{\pi}_{dc}(d)$, is

$$\tilde{\pi}_{dc} = \frac{\delta\,(p_{dd}+\delta\,(p_{dc}-p_{cd})-p_{dc}+p_{cd}-p_{cc})\,\sigma_{dc}^2-\left(\delta^2\,(p_{dc}-p_{cd})+2\,\delta\,(p_{dd}-p_{dc})+p_{dc}+p_{cd}-p_{dd}-p_{cc}\right)\sigma_{dc}-(1-\delta)\,(p_{dd}-p_{cd})}{2\,\delta\,(\sigma_{dc}-1)\,\sigma_{dc}+1}. \tag{17}$$

First, I show that the two conditions $\tilde{\pi}_{cc} > \tilde{\pi}_{dc}$ and $\tilde{\pi}_{dc} = 0$ imply that σ is a mixed MPE. By $\sigma_{dc} = \sigma_{cd}$ and Eq. (8), $\tilde{\pi}_{dc} = \tilde{\pi}_{cd}$, i.e. $\tilde{\pi}_{dc} = 0$ implies $\tilde{\pi}_{cd} = 0$. Further, by $\tilde{\pi}_{cc} > \tilde{\pi}_{dc}$ and Eqs. (7)–(8), $\sigma_{cc} = 1 > \sigma_{dc}$ implies $\mu < 0$, and by $\sigma_{dd} = 0 < \sigma_{dc}$ this implies $\tilde{\pi}_{dd} < \tilde{\pi}_{dc} = 0$. Hence, any strategy profile satisfying $\tilde{\pi}_{cc} > \tilde{\pi}_{dc} = 0$ (besides $\sigma_{cc} = 1, \sigma_{dd} = 0$) is mixed MPE with the claimed incentive structure.

Second, I derive the existence condition. $\tilde{\pi}_{dc} = 0$ obtains if

$$\sigma_{dc} = \frac{(2\,\delta-1)\,p_{dd}+(1-\delta)^2\,p_{dc}+\left(1-\delta^2\right)p_{cd}-p_{cc}\pm\sqrt{r}}{(2\,\delta^2-2\,\delta)\,(p_{dc}-p_{cd})-2\,\delta\,(p_{cc}-p_{dd})} \tag{18}$$

with

$$\begin{aligned}
r = {} & (p_{dd}-p_{dc}-p_{cd}+p_{cc})^2+4\,\delta\,\left((p_{dc}+p_{cd})\,p_{dd}+p_{cc}\,(p_{dc}+p_{cd}-2\,p_{dd})-p_{dc}^2-p_{cd}^2\right)\\
& -2\,\delta^2\,\left((p_{dc}+p_{cd})\,p_{dd}+p_{cc}\,(p_{dc}+p_{cd}-2\,p_{dd})-3\,p_{dc}^2+4\,p_{cd}\,p_{dc}-3\,p_{cd}^2\right)\\
& +\delta^4\,(p_{dc}-p_{cd})^2-4\,\delta^3\,(p_{dc}-p_{cd})^2
\end{aligned}$$

These strategy profiles exist if $r \geq 0$, and solving $r = 0$ for δ, this yields the lower bound claimed in Eq. (14). Now, evaluating $\tilde{\pi}_{cc} - \tilde{\pi}_{dc}$ at $\sigma_{cc} = 1$, $\sigma_{cd} = \sigma_{dc}$, $\sigma_{dd} = 0$ yields

$$\tilde{\pi}_{cc} - \tilde{\pi}_{dc} = \frac{(1-\sigma_{dc})\,(\delta\,(2\,p_{dd}\,\sigma_{dc}-2\,p_{cc}\,\sigma_{dc}-2\,p_{dd}+p_{dc}+p_{cd})+p_{dd}-p_{dc}-p_{cd}+p_{cc})}{2\,\delta\,(\sigma_{dc}-1)\,\sigma_{dc}+1} \tag{19}$$

and at the limiting strategy $\sigma_{dc}\big|_{r=0}$, it is positive if and only if

$$\frac{(d-1)^2\,(p_{dc}-p_{cd})\,(p_{dd}-p_{dc}-p_{cd}+p_{cc})}{p_{cc}-p_{dd}+(1-\delta)\,(p_{dc}-p_{cd})} > 0. \tag{20}$$

This is satisfied if and only if $p_{cc} + p_{dd} > p_{dc} + p_{cd}$. Otherwise, the limiting strategy

σ_{dc} does not solve $r = 0$. Instead, it solves $\tilde{\pi}_{cc} = \tilde{\pi}_{dc}$, which yields

$$\sigma_{dc} = \frac{(2\delta - 1)\, p_{dd} + (1-\delta)\, p_{dc} + (1-\delta)\, p_{cd} - p_{cc}}{2\delta(p_{dd} - p_{cc})}. \tag{21}$$

Substituting it into $\tilde{\pi}_{dc} = 0$, and solving for δ yields $\delta > \delta_{\text{BOS}}$. $\square$

Lemma A.1. *If both players play strategy* σ *in a "simple" repeated PD, then the expected payoffs in the four states are*

$$\pi_{cc}=\frac{\delta\left(a\sigma_{cc}\sigma_{dd}^2-\sigma_{cc}\sigma_{dd}^2+\sigma_{dd}^2-a\sigma_{cc}^2\sigma_{dd}+\sigma_{cc}^2\sigma_{dd}-a\sigma_{dd}-\sigma_{dd}-2a\sigma_{cc}\sigma_{cd}\sigma_{dc}+2\sigma_{cc}\sigma_{cd}\sigma_{dc}-2\sigma_{cd}\sigma_{dc}+a\sigma_{cc}^2\sigma_{dc}-\sigma_{cc}^2\sigma_{dc}+\sigma_{dc}+a\sigma_{cc}^2\sigma_{cd}-\sigma_{cc}^2\sigma_{cd}+\sigma_{cd}+\sigma_{cc}^2+a\sigma_{cc}-\sigma_{cc}\right)+a\delta^2\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)-a\sigma_{cc}+\sigma_{cc}-1}{\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{dc}+\sigma_{cd}+\sigma_{cc}^2\right)+\delta^2\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)-1},$$

$$\pi_{cd}=\frac{\begin{array}{c}\delta^2\left(a\sigma_{dc}^2\sigma_{dd}^2-a\sigma_{cc}\sigma_{dc}\sigma_{dd}^2-\sigma_{cc}\sigma_{dc}\sigma_{dd}^2+\sigma_{dc}\sigma_{dd}^2-a\sigma_{cd}^2\sigma_{dd}^2+a\sigma_{cc}\sigma_{cd}\sigma_{dd}^2+\sigma_{cc}\sigma_{cd}\sigma_{dd}^2+a\sigma_{cd}\sigma_{dd}^2-a\sigma_{cc}\sigma_{dd}^2-\sigma_{cc}\sigma_{dd}^2-2a\sigma_{cd}\sigma_{dc}^2\sigma_{dd}+2a\sigma_{cd}^2\sigma_{dc}\sigma_{dd}-a\sigma_{cd}\sigma_{dc}\sigma_{dd}-\sigma_{cd}\sigma_{dc}\sigma_{dd}+a\sigma_{cc}^2\sigma_{dc}\sigma_{dd}+\sigma_{cc}^2\sigma_{dc}\sigma_{dd}-a\sigma_{cc}^2\sigma_{cd}\sigma_{dd}-\sigma_{cc}^2\sigma_{cd}\sigma_{dd}+\\ a\sigma_{cc}^2\sigma_{dd}+\sigma_{cc}^2\sigma_{dd}+2\sigma_{cc}\sigma_{cd}\sigma_{dc}^2+a\sigma_{cd}\sigma_{dc}^2-\sigma_{cd}\sigma_{dc}^2-\sigma_{cc}^2\sigma_{dc}^2-2\sigma_{cc}\sigma_{cd}^2\sigma_{dc}-a\sigma_{cd}^2\sigma_{dc}+\sigma_{cd}^2\sigma_{dc}+a\sigma_{cc}\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\sigma_{dc}-a\sigma_{cc}^2\sigma_{dc}+\sigma_{cc}^2\sigma_{cd}^2-\sigma_{cc}^2\sigma_{cd}\right)+\delta\left(a\sigma_{dc}\sigma_{dd}^2-\sigma_{cd}\sigma_{dd}^2+\sigma_{dd}^2-a\sigma_{cd}\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dc}\sigma_{dd}-a\sigma_{dc}\sigma_{dd}-\sigma_{dc}\sigma_{dd}+a\sigma_{cd}\sigma_{dd}+\sigma_{cd}\sigma_{dd}\right.\\ \left.-a\sigma_{dd}-\sigma_{dd}-a\sigma_{cd}\sigma_{dc}^2-\sigma_{cd}\sigma_{dc}^2+\sigma_{dc}^2+a\sigma_{cd}^2\sigma_{dc}+\sigma_{cd}^2\sigma_{dc}-a\sigma_{cc}\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\sigma_{dc}-2\sigma_{cd}\sigma_{dc}+a\sigma_{cc}^2\sigma_{dc}+a\sigma_{dc}-\sigma_{cd}^2-\sigma_{cc}^2\sigma_{cd}+\sigma_{cd}+\sigma_{cc}^2\right)-a\sigma_{dc}+\sigma_{cd}-1\end{array}}{\left(\delta\left(\sigma_{dc}-\sigma_{cd}\right)+1\right)\left(\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{dc}+\sigma_{cd}+\sigma_{cc}^2\right)+\delta^2\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)-1\right)},$$

$$\pi_{dc}=\frac{\begin{array}{c}-\delta^2\left(\sigma_{dc}^2\sigma_{dd}^2-a\sigma_{cc}\sigma_{dc}\sigma_{dd}^2-\sigma_{cc}\sigma_{dc}\sigma_{dd}^2-a\sigma_{dc}\sigma_{dd}^2-2\sigma_{dc}\sigma_{dd}^2-\sigma_{cd}^2\sigma_{dd}^2+a\sigma_{cc}\sigma_{cd}\sigma_{dd}^2+\sigma_{cc}\sigma_{cd}\sigma_{dd}^2+\sigma_{cd}\sigma_{dd}^2+a\sigma_{cc}\sigma_{dd}^2+\sigma_{cc}\sigma_{dd}^2-2\sigma_{cd}\sigma_{dc}^2\sigma_{dd}+2\sigma_{cd}^2\sigma_{dc}\sigma_{dd}+a\sigma_{cd}\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dc}\sigma_{dd}+a\sigma_{cc}^2\sigma_{dc}\sigma_{dd}+\sigma_{cc}^2\sigma_{dc}\sigma_{dd}+2a\sigma_{dc}\sigma_{dd}+2\sigma_{dc}\sigma_{dd}\right.\\ \left.-a\sigma_{cc}^2\sigma_{cd}\sigma_{dd}-\sigma_{cc}^2\sigma_{cd}\sigma_{dd}-2a\sigma_{cd}\sigma_{dd}-2\sigma_{cd}\sigma_{dd}-a\sigma_{cc}^2\sigma_{dd}-\sigma_{cc}^2\sigma_{dd}+2a\sigma_{cc}\sigma_{cd}\sigma_{dc}^2+a\sigma_{cd}\sigma_{dc}^2+3\sigma_{cd}\sigma_{dc}^2-a\sigma_{cc}^2\sigma_{dc}^2-a\sigma_{dc}^2-\sigma_{dc}^2-2a\sigma_{cc}\sigma_{cd}^2\sigma_{dc}-a\sigma_{cd}^2\sigma_{dc}-3\sigma_{cd}^2\sigma_{dc}-a\sigma_{cc}\sigma_{cd}\sigma_{dc}-\sigma_{cc}\sigma_{cd}\sigma_{dc}-\sigma_{cc}^2\sigma_{dc}+a\sigma_{cc}^2\sigma_{cd}^2+a\sigma_{cd}^2+\sigma_{cd}^2+a\sigma_{cc}^2\sigma_{cd}+2\sigma_{cc}^2\sigma_{cd}\right)\\ -\delta\left(\sigma_{dc}\sigma_{dd}^2-a\sigma_{cd}\sigma_{dd}^2-\sigma_{dd}^2+a\sigma_{cd}\sigma_{dc}\sigma_{dd}-\sigma_{cd}\sigma_{dc}\sigma_{dd}-a\sigma_{dc}\sigma_{dd}-\sigma_{dc}\sigma_{dd}+a\sigma_{cd}\sigma_{dd}+\sigma_{cd}\sigma_{dd}+a\sigma_{dd}+\sigma_{dd}-a\sigma_{cd}\sigma_{dc}^2-\sigma_{cd}\sigma_{dc}^2+a\sigma_{dc}^2+a\sigma_{cd}^2\sigma_{dc}+\sigma_{cd}^2\sigma_{dc}+a\sigma_{cc}\sigma_{cd}\sigma_{dc}-\sigma_{cc}\sigma_{cd}\sigma_{dc}+2\sigma_{cd}\sigma_{dc}+\sigma_{cc}^2\sigma_{dc}+\sigma_{dc}-a\sigma_{cd}^2-a\sigma_{cc}^2\sigma_{cd}-a\sigma_{cd}-2\sigma_{cd}-\sigma_{cc}^2\right)\\ +\left(a+1\right)\delta^3\left(\sigma_{dc}-\sigma_{cd}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)+\sigma_{dc}-a\sigma_{cd}-1\end{array}}{\left(\delta\left(\sigma_{dc}-\sigma_{cd}\right)+1\right)\left(\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{dc}+\sigma_{cd}+\sigma_{cc}^2\right)+\delta^2\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)-1\right)},$$

$$\pi_{dd}=\frac{\delta\left(a\sigma_{dc}\sigma_{dd}^2-\sigma_{dc}\sigma_{dd}^2+a\sigma_{cd}\sigma_{dd}^2-\sigma_{cd}\sigma_{dd}^2-a\sigma_{cc}\sigma_{dd}^2+\sigma_{cc}\sigma_{dd}^2+\sigma_{dd}^2-2a\sigma_{cd}\sigma_{dc}\sigma_{dd}+2\sigma_{cd}\sigma_{dc}\sigma_{dd}+a\sigma_{cc}^2\sigma_{dd}-\sigma_{cc}^2\sigma_{dd}-2\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{dc}+\sigma_{cd}+\sigma_{cc}^2\right)+\delta^2\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)-a\sigma_{dd}+\sigma_{dd}-1}{\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{dc}+\sigma_{cd}+\sigma_{cc}^2\right)+\delta^2\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dc}\sigma_{dd}+\sigma_{cd}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-2\sigma_{cd}\sigma_{dc}+\sigma_{cc}\sigma_{dc}+\sigma_{cc}\sigma_{cd}\right)-1}$$

Proof. These results follow straightforwardly from algebraic manipulations of Eqs. (1)-(3). □

Lemma A.2. (σ,σ) *is an MLE of a "simple" repeated PD if and only if* $\log(1-\sigma_{s',s''})/\sigma_{s',s''}=-\lambda\tilde{\pi}_{s',s''}$ *for all* $s',s''\in S$*, where* $\tilde{\pi}_{s',s''}:=\pi_{s',s''}(c)-\tilde{\pi}_{s',s''}(d)$ *are*

$$\tilde{\pi}_{cc}=\frac{(d-1)\left(-\delta\left(b\sigma_{dd}^2-a\sigma_{dd}^2-2b\sigma_{cc}\sigma_{dd}+a\sigma_{cc}\sigma_{dd}+a\sigma_{dd}-b\sigma_{dc}^2+a\sigma_{dc}^2+\sigma_{dc}^2+2b\sigma_{cc}\sigma_{dc}-2\sigma_{cc}\sigma_{dc}-a\sigma_{dc}-a\sigma_{cc}^2+\sigma_{cc}^2\right)+a\delta^2\left(\sigma_{dc}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{dc}\right)+b\sigma_{cc}-a\sigma_{cc}+\sigma_{cc}-1\right)}{\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{dc}^2+2\sigma_{dc}+\sigma_{cc}^2\right)+2\delta^2\left(\sigma_{dc}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{dc}\right)-1}$$

$$\tilde{\pi}_{dc}=\frac{(d-1)\left(-\delta\left(b\sigma_{dd}^2-a\sigma_{dd}^2-2b\sigma_{dc}\sigma_{dd}+a\sigma_{dc}\sigma_{dd}+a\sigma_{dd}+b\sigma_{dc}^2+a\sigma_{dc}^2-\sigma_{dc}^2-a\sigma_{cc}\sigma_{dc}+2\sigma_{cc}\sigma_{dc}-a\sigma_{dc}-\sigma_{cc}^2\right)+a\delta^2\left(\sigma_{dc}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{dc}\right)+b\sigma_{dc}-a\sigma_{dc}+\sigma_{dc}-1\right)}{\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{dc}^2+2\sigma_{dc}+\sigma_{cc}^2\right)+2\delta^2\left(\sigma_{dc}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{dc}\right)-1}$$

$$\tilde{\pi}_{dd}=\frac{(d-1)\left(\delta\left(b\sigma_{dd}^2-2b\sigma_{dc}\sigma_{dd}+2\sigma_{dc}\sigma_{dd}+a\sigma_{cc}\sigma_{dd}-2\sigma_{cc}\sigma_{dd}-a\sigma_{dd}+b\sigma_{dc}^2-a\sigma_{dc}^2-\sigma_{dc}^2+a\sigma_{dc}+\sigma_{cc}^2\right)+a\delta^2\left(\sigma_{dc}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{dc}\right)+b\sigma_{dd}-a\sigma_{dd}+\sigma_{dd}-1\right)}{\delta\left(\sigma_{dd}^2-2\sigma_{dd}-2\sigma_{dc}^2+2\sigma_{dc}+\sigma_{cc}^2\right)+2\delta^2\left(\sigma_{dc}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{cc}\right)\left(\sigma_{dd}-\sigma_{dc}\right)-1}.$$

Proof. $\sigma_{cd}=\sigma_{dc}$ implies $\tilde{\pi}_{cd}=\tilde{\pi}_{dc}$ and the remainder follows from the definition Eq. (11), using Eq. (1) and Lemma A.1. □