Instability of Defection in the Prisoner's Dilemma Under Best Experienced Payoff Dynamics [*]

Srinivas Arigapudi[†], Yuval Heller[‡], Igal Milchtaich[§]

August 20, 2021

Abstract

We study population dynamics under which each revising agent tests each action k times, with each trial being against a newly drawn opponent, and chooses the action whose mean payoff was highest during the testing phase. When $k = 1$, defection is globally stable in the prisoner's dilemma. By contrast, when $k > 1$ we show that, if the gains from defection are not too large, there exists a globally stable state in which agents cooperate with probability between 28% and 50%. Next, we characterize stability of strict equilibria in general games. Our results demonstrate that the empirically plausible case of $k > 1$ can yield qualitatively different predictions than the case $k = 1$ commonly studied in the literature.

Keywords: learning, cooperation, best experienced payoff dynamics, sampling equilibrium, evolutionary stability. **JEL codes:** C72, C73.

[†]Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, Haifa 32000003, Israel. e-mail: arigapudi@campus.technion.ac.il; website: sites.google.com/view/srinivasarigapudi/home.

[‡]Department of Economics, Bar-Ilan University, Ramat Gan 5290002, Israel. e-mail: yuval.heller@biu.ac.il; website: https://sites.google.com/site/yuval26/.

[§]Department of Economics, Bar-Ilan University, Ramat Gan 5290002, Israel. e-mail: igal.milchtaich@biu.ac.il; website: https://sites.google.com/view/milchtaich/.

1. Introduction

The standard approach in game theory assumes that players form beliefs about the various uncertainties they face and then best respond to these beliefs. In equilibrium, the beliefs will be correct and the players will play a Nash equilibrium. However, in some economic environments where the players have limited information about the strategic situation, Nash equilibrium prediction is hard to justify. Consider the following example from Osborne and Rubinstein (1998). You are new to town and are planning your route to work. How do you decide which road to take? You know that other people use the roads, but have no idea which road is most congested. One plausible procedure is to try each route several times and then permanently adopt the one that was (on average) best. The outcome of this procedure is stochastic: you may sample the route that is in fact the best on a day when a baseball game congests it. Once you select your route, you become part of the environment that determines other drivers' choices.

This procedure is formalized as follows. Consider agents in a large population who are randomly matched to play a symmetric n-player game with a finite set of actions. Agents occasionally revise their action (which can also be interpreted as agents occasionally leaving the population and being replaced by new agents who base their behavior on the sampling procedure, as in the motivating example above). Each revising agent samples each feasible action k times and chooses the action that yields the highest average payoff (applying some tie-breaking rule).

This procedure induces a dynamic process according to which the distribution of actions in the population evolves (best experienced payoff dynamics: Sethi, 2000; Sandholm et al., 2019). An $S(k)$-equilibrium α^* is a rest point of the above dynamics. The equilibrium is locally stable if any distribution of actions in the population that is sufficiently close to α^* converges to α^*, and globally stable if any distribution of actions in the population with support that includes all actions converges to α^*.

The existing literature on payoff sampling equilibria (as surveyed below) has mainly focused on $S(1)$-equilibria, due to their tractability. It seems plausible that real-life behavior would rely on sampling each action more than once. A key insight of our analysis is that sampling actions several times might lead to qualitatively different results than sampling each action only once. In particular, in the prisoner's dilemma game, 1-payoff sampling dynamics yield the Nash equilibrium behavior, while k-payoff sampling dynamics (for $k > 1$) may induce substantial cooperation.

Recall that each player in the prisoner's dilemma game has two actions, cooperation c

and defection d, and the payoffs are as in Table 1, where $g, l > 0$. Sethi (2000) has shown that defection is the unique $S(1)$ globally stable $S(1)$-equilibrium. By contrast, our first main result (Theorem 1) shows that for any $k \geq 2$, a game for which the gains from defection are not too large (specifically, $g, l < \frac{1}{k-1}$) admits a globally stable state in which the rate of cooperation is between 28% and 50%.

	c	d
c	1 , 1	-l , 1+g
d	1+g , -l	0 , 0

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

Our remaining results characterize the local stability of strict equilibria for $k \geq 2$. Proposition 1 shows that defection in the prisoner's dilemma game is locally stable iff $l > \frac{1}{k-1}$. Theorem 2 extends the analysis to general symmetric games. It presents a simple necessary and sufficient condition for a strict symmetric equilibrium action a^* to be $S(k)$ locally stable (improving on the conditions presented in Sethi, 2000; Sandholm et al., 2020). Roughly speaking, the condition is that in any set of actions A' that does not include a^* there is an action that never yields the highest payoff when the corresponding sample includes a single occurrence of an action in A' and all the other sampled actions are a^*. Theorem 3 extends the characterization of local stability of strict equilibria to general asymmetric games.

Outline: In the remaining parts of the Introduction we review the related literature, and compare our predictions with the experimental findings. In Section 2, we introduce our model and the solution concept. We analyze the prisoner's dilemma in Section 3, and characterize the stability of strict equilibria in general symmetric games in Section 4. An extension of the analysis to asymmetric games is presented in Section 5.

1.1 Related Experimental Literature and Testable Predictions

The typical length of a lab experiment, as well as the subjects' cognitive costs, are likely to limit the sample sizes used by subjects to test the various actions to small values such as $k = 2$ or $k = 3$ (because larger samples induce too-high costs of non-optimal play during the sampling stage, and require larger cognitive effort to analyze). Proposition 2 shows that for these small sample sizes of $k = 2$ or 3, the k-payoff sampling dynamics admit a unique globally stable equilibrium, which depends on the parameter l. Specifically, everyone defecting is the globally stable equilibrium if $l > \frac{1}{k-1}$, while there is a substantial rate of cooperation between 24% and 33% if $l < \frac{1}{k-1}$.

The predictions of our model match quite well the empirical findings of the meta-study of Mengel (2018) concerning the behavior of subjects playing the one-shot prisoner's dilemma. Mengel (2018, Tables A.3, B.5) summarizes 29 sessions of lab experiments of that game in a "stranger" (one-shot) setting from 16 papers (with various values of g and l, both with median 1; the distribution of values is presented in Appendix B). The average rate of cooperation in these experiments is 37%. Our predictions are also broadly consistent with the experimentally observed comparative statics with respect to g and l, which is that the rate of cooperation is decreasing in l but is independent of g (see Mengel, 2018, Table B.5, where l is called $RISK^{Norm}$ and g is called $TEMPT^{Norm}$).[1]

The empirically observed average cooperation rate of 37% can also be explained by other theories. Specifically, it can be explained by agents making errors when the payoff differences are small (quantal response equilibrium, McKelvey and Palfrey, 1995), or by agents caring about the payoffs of cooperative opponents (inequality aversion à la Fehr and Schmidt, 1999, and reciprocity à la Rabin, 1993). Our model has two advantages in comparison with these alternative models. First, our model is parameter free (for a fixed k), while the existing models may require tuning their parameters to fit the experimental data (such as the parameter describing the agents' error rates in a quantal response equilibrium).

Second, the predictions of the existing models are arguably less compatible with the above-mentioned experimentally observed comparative statics. Quantal response equilibrium predicts that the cooperation rate decreases in both parameters. The other models predict that the cooperation rate decreases in g (because an increasing g increases the material payoff from defecting, while it does not change the payoff of a cooperative opponent), and their prediction with respect to l is ambiguous, because increasing l has two opposing effects: increasing the material gain from defection against a defecting opponent but decreasing the payoff of a cooperating opponent.

Our predictions might have an even better fit with experiments in which subjects have only partial information about the payoff matrix (a setting that might be relevant to many real-life interactions), such as a "black box" setting in which players do not know the game's structure and observe only their realized payoffs (see, e.g., Nax and Perc, 2015; Nax et al., 2016; Burton-Chellew et al., 2017).

---

[1]Theorem 1 shows that for g that is not too large (specifically, $g < \frac{1}{k-1}$), the minimal rate of cooperation in the globally stable state is 28% (when $l < \frac{1}{k-1}$). Proposition 2 allows arbitrary large g, which has the modest impact of slightly decreasing the minimal globally stable rate of cooperation to 24%.

1.2 Related Theoretical Literature

The payoff sampling dynamics approach employed in this paper was pioneered by Osborne and Rubinstein (1998) and Sethi (2000). The approach has been used in a variety of applications, including bounded-rationality models in industrial organization (Spiegler, 2006a,b), coordination games (Ramsza, 2005), trust and delegation of control (Rowthorn and Sethi, 2008), market entry (Chmura and Güth, 2011), ultimatum games (Miękisz and Ramsza, 2013), common-pool resources (Cárdenas et al., 2015), contributions to public goods (Mantilla et al., 2018), and finitely repeated games (Sethi, 2019).

Most of these papers mainly focus on 1-payoff sampling dynamic, in which each action is only sampled once.[2] One exception is Sandholm et al. (2019), which analyzes the stable $S(k)$-equilibrium in a centipede game and shows that it involves cooperative behavior even when the number of trials k of each action is large. Another is Sandholm et al. (2020), which presents general stability and instability criteria of $S(k)$-equilibria in general classes of games, thus providing a unified way of deriving many of the specific results the above papers derive, as well as several new results.

A related, alternative approach is *action sampling dynamics* (or sample best-response dynamics), according to which each revising agent obtains a small random sample of other players' actions, and chooses the action that is a best reply to that sample (see, e.g., Sandholm, 2001; Kosfeld et al., 2002; Kreindler and Young, 2013; Oyama et al., 2015; Heller and Mohlin, 2018; Salant and Cherry, 2020). The action sampling approach is a plausible heuristic when the players know the payoff matrix and are capable of strategic thinking but do not know the exact distribution of actions in the population.

2. Model

We consider a unit-mass continuum of agents who are randomly matched to play a symmetric n-player game $G = \{A, u\}$, where $A = \{a_1, a_2, \ldots, a_m\}$ is the (finite) set of actions and $u : A^n \to \mathbb{R}$ is the payoff function, which is invariant to permutations of its second through n-th arguments. An agent taking action a^1 against opponents playing $a^2, \ldots, a^n$, in any order, receives payoff $u(a^1, a^2, \ldots, a^n)$.

Aggregate behavior in the population is described by a *population state* α lying in the unit simplex $\Delta \equiv \{\alpha = (\alpha_{a_i})_{i=1}^m \in \mathbb{R}_+^m \mid \sum_{i=1}^m \alpha_{a_i} = 1\}$, with α_{a_i} representing the fraction of

---

[2] See also the variant of the 1-payoff sampling dynamic presented in Rustichini (2003), according to which after an initial phase of sampling each action once, each player in each round chooses the action that has yielded the highest average payoff so far.

agents in the population using action a_i. The standard basis vector $e_a \in \Delta$ represents the pure, or monomorphic, state in which all agents play action a. Where no confusion is likely, we identify the action with the monomorphic state, denoting them both by a. The set of *interior population states*, in which all actions are used by a positive mass of agents, is $Int(\Delta) \equiv \Delta \cap \mathbb{R}^m_{++}$.

A sampling procedure involves the testing of the different actions against randomly drawn opponents, as explained next. Agents occasionally receive opportunities to switch actions (equivalently, this can be thought of as agents dying and being replaced by new agents). These opportunities do not depend on the currently used actions. That is, when the population state is $\alpha(t)$, the proportion of agents originally using an action a out of the agents who revise between time t and $t + dt$ is equal to their proportion in the population $\alpha_a(t)$.

When an agent receives a revision opportunity, he tries each of the feasible actions k times, using it each time against a newly drawn set of opponents from the population. Thus, the probability that the opponent's action is any $a \in A$ is $\alpha_a(t)$. The agent then chooses the action that yielded the highest mean payoff in these trials, employing some tie-breaking rule if more than one action yields the highest mean payoff. All of our results hold for any tie-breaking rule. Denote the probability that the chosen action is a by $w_{a,k}(\alpha(t))$.

As a result of the revision procedure described above, the expected change in the number of agents using an action a during an infinitesimal time interval of duration dt is

$$(2.1) \qquad w_{a,k}(\alpha(t))dt - \alpha_a(t)dt.$$

The first term in (2.1) is an inflow term, representing the expected number of revising agents who switch to action a, while the second term is an outflow term, representing the expected number of revising agents who are currently playing that action. In the limit $dt \to 0$, the rate of change of the fraction of agents using each action is given in vector notation by

$$(2.2) \qquad \dot{\alpha} = w_k(\alpha(t)) - \alpha(t),$$

where w_k is a vector whose a-th component is $w_{a,k}$. The system of differential equations (2.2) is called the *k-payoff sampling dynamic*. Its rest points are called *S(k)-equilibria*.

Definition 1 (Osborne and Rubinstein, 1998)**.** A population state $\alpha^* \in \Delta$ is an *S(k)-equilibrium* if $w_k(\alpha^*) = \alpha^*$.

An equilibrium is (locally) asymptotically stable if a population beginning near it remains close and eventually converges to the equilibrium, and it is (almost) globally asymptotically stable if the population converges to it from any initial interior state.

Definition 2. An $S(k)$-equilibrium α^* is $S(k)$ *asymptotically stable* if:

1. (*Lyapunov stability*) for every neighborhood U of α^* in Δ there is a neighborhood $V \subset U$ of α^* such that if $\alpha(0) \in V$, then $\alpha(t) \in U$ for all $t > 0$; and

2. there is some neighborhood U of α^* in Δ such that all trajectories initially in U converge to α^*; that is, $\alpha(0) \in U$ implies $\lim_{t\to\infty} \alpha(t) = \alpha^*$.

Definition 3. An $S(k)$-equilibrium α^* is $S(k)$ *globally asymptotically stable* if all interior trajectories converge to α^*; that is, $\alpha(0) \in Int(\Delta)$ implies $\lim_{t\to\infty} \alpha(t) = \alpha^*$.

3. The Prisoner's Dilemma

This section focuses on the (two-player) prisoner's dilemma game. The set of actions is given by $A = \{c, d\}$, where c is interpreted as cooperation and d as defection. The payoffs are as described in Table 1, with $g, l > 0$: when both players cooperate they get payoff 1, when they both defect they get 0, and when one player defects and the other cooperates, the defector gets $1+g$ and the cooperator gets $-l$.

3.1 Stability of Defection

Sethi (2000, Example 5) analyzes the 1-payoff sampling dynamic and shows that everyone defecting is globally stable.

Claim 1 (Sethi, 2000)**.** *Defection is $S(1)$ globally asymptotically stable.*

The argument behind Claim 1 is as follows. When an agent samples the action c (henceforth, the *c-sample*) her payoff is higher than when sampling the action d (henceforth, the *d-sample*) iff the opponent has cooperated in the c-sample and defected in the d-sample (which happens with probability $\alpha_c \cdot \alpha_d$). Therefore, the 1-payoff sampling dynamic is given by

$$\dot{\alpha}_c = w_{c,1}(\alpha) - \alpha_c = \alpha_c \cdot \alpha_d - \alpha_c = \alpha_c(1 - \alpha_c) - \alpha_c = -\alpha_c^2 \quad (< 0 \text{ if } \alpha_c > 0).$$

The unique rest point $\alpha_c^* = 0$ is the unique $S(1)$-equilibrium, and it is easy to see that it is globally asymptotically stable.

Our next result shows that for $k \geq 2$ everyone defecting is no longer globally asymptotically stable (indeed, it is not even locally asymptotically stable) if $l < \frac{1}{k-1}$.

Proposition 1. *Let $k \geq 2$ and assume that[3] $l \neq \frac{1}{k-1}$. Defection is $S(k)$ asymptotically stable if and only if $l > \frac{1}{k-1}$.*

Proposition 1 is implied by the results of Sandholm et al. (2020), and as we show in Section 4.3, it also follows from Theorem 2 below. For completeness, we provide a direct sketch of proof.

Sketch of Proof. Consider a population state in which a small fraction ϵ of the agents co-operate and the remaining $1 - \epsilon$ agents defect. A revising agent most likely sees all the opponents defecting, both in the c-sample and in the d-sample. With a probability of approximately $k\epsilon$, the agent sees a single cooperation in the c-sample and no cooperation in the d-sample, and so c yields a mean payoff of $\frac{1-(k-1)\cdot l}{k}$ and d yields 0. The former is higher iff $l < \frac{1}{k-1}$. Thus, if the last inequality holds, then the prevalence of cooperation gradually increases, and the population drifts away from the state where everyone defects. By contrast, if $l > \frac{1}{k-1}$, then cooperation yields the higher mean payoff only if the c-sample includes at least two cooperators, which happens with a negligible probability of order ϵ^2. Therefore, in this case, cooperation gradually dies out, and the population converges to the state where everyone defects. □

3.2 Stability of (Partial) Cooperation

Next we show that for any $k \geq 2$, if g and l are sufficiently small, then the prisoner's dilemma game admits an $S(k)$ globally asymptotically stable $S(k)$-equilibrium in which the frequency of cooperation is between 28% and 50% and is increasing in k.

Theorem 1. *For $k \geq 2$ and $g, l < \frac{1}{k-1}$, the unique $S(k)$ globally asymptotically stable equilibrium α^k in the prisoner's dilemma game satisfies $0.28 < \alpha_c^k < 0.5$. Moreover, $\alpha_c^{k'} < \alpha_c^k$ for all $2 \leq k' < k$.*

The intuition for Theorem 1 is as follows. The condition $g, l < \frac{1}{k-1}$ implies that cooperation yields a higher average payoff than defection iff the opponent has cooperated more

---

[3]The stability of defection in the borderline case of $l = \frac{1}{k-1}$ depends on the tie-breaking rule, because observing a single c in the c-sample and no c's in the d-sample produces a tie between the two samples. If one assumes a uniform tie-breaking rule, then action c wins with a probability of $\frac{k}{2}\epsilon - O(\epsilon^2)$, which is greater than ϵ if and only if $k > 2$. Thus, with this rule, defection is stable if $k = 2$ and unstable if $k > 2$.

times in the d-sample than in the c-sample. If α_c is close to zero, then the probability of this event is roughly $k \cdot \alpha_c > \alpha_c$. The symmetry between the two samples implies that the probability that the opponent has cooperated more times in the d-sample is less than 0.5. Thus, there exists $\alpha_c^k < 0.5$, which is not close to zero, for which the probability of cooperation yielding a higher average payoff is equal to α_c^k. The proof shows that this equality holds for $\alpha_c^k > 0.28$, and that α_c is globally stable.

Proof. The proof of the theorem uses a number of claims, whose formal proofs are given in Appendix A. In what follows, we state each claim and present a sketch of proof.

Notation For $j \le k$, let $f_{k,p}(j) \equiv \binom{k}{j}p^j(1-p)^{k-j}$ be the probability mass function of a binomial random variable with parameters k and p. Let $Tie(k,p) = \sum_{j=0}^{k}(f_{k,p}(j))^2$ be the probability of having a tie between two independent binomial random variables with parameters k, p, and let $Win(k,p) = 0.5 \cdot (1 - Tie(k,p))$ be the probability that the first random variable has a larger value than the second. Let $p \equiv \alpha_c$ denote the proportion of cooperating agents in the population.

Claim 2. *Assume that $g, l \in (0, \frac{1}{k-1})$. The k-payoff sampling dynamic is given by*

$$(3.1) \qquad \dot{p} = Win(k,p) - p.$$

Sketch of Proof. The condition $g, l < \frac{1}{k-1}$ implies that action c has a higher mean payoff iff the c-sample includes more cooperating opponents than the d-sample does. The number of cooperators in each sample has a binomial distribution with parameters k and p, and so the probability of c having a higher mean payoff is $Win(k,p)$ (which we substitute in (2.2)). □

For $k \ge 2$ and $0 \le p \le 1$, denote the expression on the right-hand side of Eq. (3.1) by

$$(3.2) \qquad h_k(p) = Win(k,p) - p.$$

Claim 3. *For $k \ge 2$, the function h_k satisfies $h_k(0) = 0, h_k(1) = -1$ and $h_k'(0) > 0$.*

Sketch of Proof. When $p = 0$ (resp., $= 1$), in both samples all the opponents are defectors (resp., cooperators). The conclusion implies that $Win(k,p) = 0$ for $p \in \{0,1\}$, which, in turn, implies that $h_k(0) = 0$ and $h_k(1) = -1$. Next, observe that for $p = \epsilon << 1$, $Win(k,p) \approx k\epsilon$, which is approximately the probability of having at least one cooperator in the c-sample. Thus, $h_k(\epsilon) \approx k\epsilon - \epsilon$, which implies that $h_k'(0) = k - 1 > 0$. □

Claim 4. *For $k \geq 2$, the expression $h_k(p)$ is concave in p, and satisfies $h_k(p) < h_{k+1}(p)$ for $p \in (0,1)$, $h_k\left(\frac{1}{2}\right) < 0$, and $\lim_{k \to \infty} h_k\left(\frac{1}{2}\right) = 0$.*

Sketch of Proof. Observe that $Tie(k,p)$ is close to 1 when p is close to either zero or one, and is smaller for intermediate p's. The formal proof shows (by analyzing the characteristic function) that $Tie(k,p)$ is (1) convex in p, (2) decreasing in k (i.e., the larger the number of actions in each sample, the smaller the probability of having exactly the same number of cooperators in both samples), and (3) converges to zero as k tends to ∞. These findings imply that $h_k(p) = (0.5 \cdot (1 - Tie(k,p)) - p$ is concave in p and increasing in k, and that

$$h_k\left(\frac{1}{2}\right) = \frac{1}{2} \cdot \left(1 - Tie\left(k, \frac{1}{2}\right)\right) - \frac{1}{2} < \frac{1}{2} - \frac{1}{2} = 0, \quad \text{and}$$

$$\lim_{k \to \infty} h_k\left(\frac{1}{2}\right) = \lim_{k \to \infty}\left(\frac{1}{2} \cdot \left(1 - Tie\left(k, \frac{1}{2}\right)\right) - \frac{1}{2}\right) = \left(\frac{1}{2} - 0\right) - \frac{1}{2} = 0.$$

$\square$

It follows from Claims 3 and 4 that for $k \geq 2$ the equation $h_k(p) = 0$ has a unique solution in the interval $(0,1)$, that this solution $p(k)$ corresponds to an $S(k)$ globally asymptotically stable state, that it satisfies $p(k) < 0.5$, and that it is increasing in k.

To complete the proof of Theorem 1, it remains to show that $p(2) > 0.28$. This inequality is an immediate corollary of the fact that for $p = 0.28$,

$$h_2(p = 0.28) = 2p(1 - p)^3 + p^2(1 - p^2) - p \approx 0.001 > 0.$$

$\square$

Figure 1 shows the k-payoff sampling dynamic and the $S(k)$ globally stable $S(k)$-equilibria for various values of k.

Remark 1. Theorem 1 shows that partial cooperation is globally stable for any fixed k if the parameters g and l are sufficiently small (with the upper bound depending on k). At the same time, it is well known that, for fixed g and l, defection is globally stable if k is sufficiently large (Sandholm et al., 2020, Prop. 4.3; see also Osborne and Rubinstein, 1998, Prop. 4).

Theorem 1 leaves open the question of the stability of partial cooperation when either g or l is larger than $\frac{1}{k-1}$. The next proposition answers this question for small sample sizes of $k = 2$ or 3, leaving full characterization of larger sample sizes for future research. As the proposition states, everyone defecting is globally asymptotically stable if $l > \frac{1}{k-1}$, while if the reverse inequality holds, the globally asymptotically stable state has a substantial rate of cooperation of between 24% and 33%.
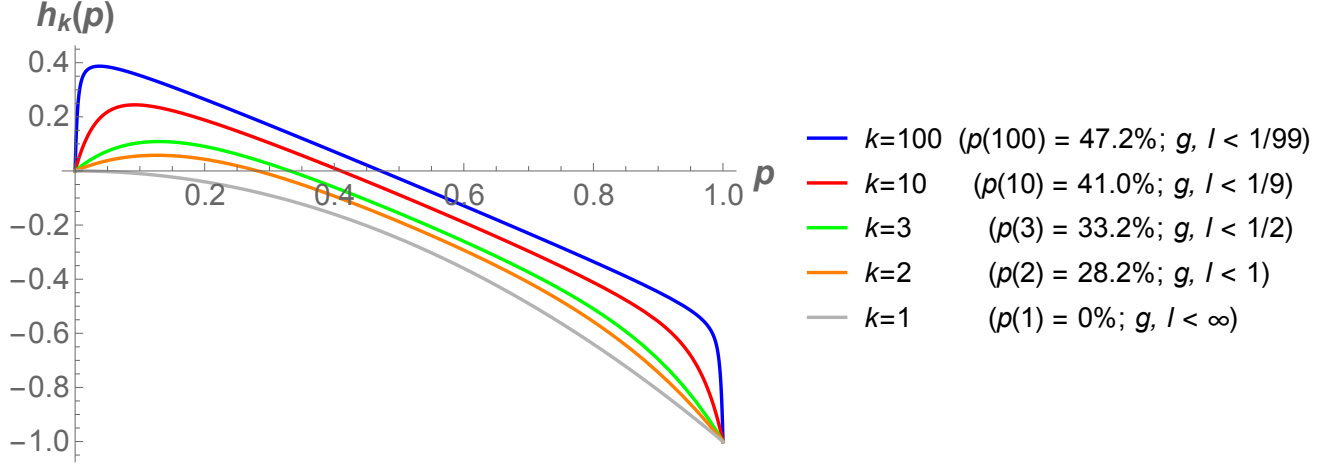
Figure 1: The function h_k and its zero $p(k)$ for various values of k.

Proposition 2. *For $k \in \{2, 3\}$, the unique $S(k)$ globally asymptotically stable $S(k)$-equilibrium α^k satisfies $\alpha^k_c = 0$ if $l > \frac{1}{k-1}$, and $\alpha^k_c \in (0.24, 0.33)$ if $l < \frac{1}{k-1}$.*

The proof of the proposition is presented in Appendix A.6.

4. General Symmetric Games

In this section we extend Proposition 1 by presenting a necessary and sufficient condition for $S(k)$ asymptotic stability of actions in general symmetric games. Our characterization uses the following two definitions.

4.1 Definitions

Definition 4. *A symmetric n-player game with payoff function $u : A^n \to \mathbb{R}$ is generic if for any two sequences of action profiles $\left(\left(a_j^1, a_j^2, \ldots, a_j^n\right)\right)_{j=1}^L$ and $\left(\left(\tilde{a}_j^1, \tilde{a}_j^2, \ldots, \tilde{a}_j^n\right)\right)_{j=1}^L$ of equal length L, the equality*

$$\sum_{j=1}^L u\left(a_j^1, a_j^2, \ldots, a_j^n\right) = \sum_{j=1}^L u\left(\tilde{a}_j^1, \tilde{a}_j^2, \ldots, \tilde{a}_j^n\right)$$

implies $\left\{a_j^1\right\}_{j=1}^L = \left\{\tilde{a}_j^1\right\}_{j=1}^L$.

Thus, a symmetric game is generic if the sums of payoffs of two sequences of action profiles are equal only if every action that appears as the first action in one of the profiles

−11−

in one sequence also appears as the first action in the other sequence. Note that this definition of genericity is rather weak. A stronger definition could be used instead; the use of a weak definition only strengthens our results below. Observe also that if each entry in the payoff matrix is independently drawn from a continuous (atomless) distribution, then the resulting random symmetric game is generic with probability one. Clearly, in a generic game, every pure Nash equilibrium is a strict equilibrium.

Definition 5. For an action a^* in a symmetric n-player game, and for $a, a' \in A \setminus \{a^*\}$:

1. action a *directly* $S(k)$ *supports* a' *against* a^* if

$$u(a', a, a^*, \ldots, a^*) + (k-1) \cdot u(a', a^*, \ldots, a^*) > k \cdot u(a^*, \ldots, a^*); \text{ and}$$

2. action a $S(k)$ *supports* a' *by spoiling* a^* if

$$k \cdot u(a', a^*, \ldots, a^*) > u(a^*, a, a^*, \ldots, a^*) + (k-1) \cdot u(a^*, \ldots, a^*)$$

$$\text{and } u(a', a^*, \ldots, a^*) > u(b, a^*, \ldots, a^*) \ \forall b \notin \{a^*, a'\}.$$

Action a $S(k)$ *single supports*, *double supports*, or just *supports* action a' against a^* if exactly one of conditions 1 and 2, both conditions, or at least one condition, respectively, holds. The notion of *weak* $S(k)$ *support* and the related terms (weak direct support, weak support by spoiling, weak single support, and weak double support) are defined similarly, except that the strict inequalities in 1 and 2 are replaced by weak inequalities.

Less formally, action a supports action a' against action a^* if a single appearance of a in a population in which almost everyone plays a^* can make the mean payoff in the a'-sample the highest one. For an action a that directly supports a' (a supporter of a', in the terminology of Sandholm et al., 2020, p.12), a single appearance of a in the a'-sample (with all other actions being a^*) is sufficient to make the mean payoff larger than that yielded by a^*, and thus to make it the largest payoff. For a that supports a' by spoiling a^* (a benefiting spoiler of a^*, in that terminology), a single appearance of a in the a^*-sample is sufficient to make the mean payoff smaller than that yielded by a'. This makes the latter the largest mean payoff if a' is the second-best reply to a^* (while if another action a'' is second best, then a'' yields a higher payoff, assuming that the a''-sample includes only a^*'s). Note that, in a generic game, the second-best reply is unique: the set $\text{argmax}_{b \neq a^*} u(b, a^*, \ldots, a^*)$ is a singleton.

Also, in a generic game, the notions of support and weak support coincide: action a $S(k)$ supports a' against a^* if and only if it weakly $S(k)$ supports a' against a^*.

Observe that if a^* is a symmetric equilibrium action, then $S(k)$ support against it is "easier" the smaller k is: if action a (weakly) $S(k)$ supports a' against a^*, then it (respectively, weakly) $S(k')$ supports a' against a^* for all $k' < k$.

4.2 Result

It is well known (Osborne and Rubinstein, 1998) that an action a^* can be an $S(k)$-equilibrium only if a^* is a symmetric Nash equilibrium (otherwise, $w_{a^*,k}(e_{a^*}) = 0$, which contradicts e_{a^*} being a rest point). Moreover, if the tie-breaking rule assigns positive probability to all co-winning actions, then a^* must be a strict equilibrium (otherwise, $w_{a^*,k}(e_{a^*}) < 1$, which again contradicts e_{a^*} being a rest point). Thus, being a strict symmetric equilibrium is essentially a necessary condition for an action to be $S(k)$ asymptotically stable. Our next result characterizes the conditions for a strict symmetric equilibrium action to be $S(k)$ asymptotically stable when $k \geq 2$ or $n \geq 3$.

Theorem 2. *Suppose that $k \geq 2$ or $n \geq 3$. A necessary and sufficient condition for a strict symmetric equilibrium action a^* in a generic symmetric n-player game to be $S(k)$ asymptotically stable is that, for the set $A^* \equiv A \backslash \{a^*\}$,*

I. every nonempty subset $A' \subseteq A^*$ includes an action a' that is not $S(k)$ supported against a^* by any action in A'.

In a non-generic game, condition I is still necessary for $S(k)$ asymptotic stability, and a sufficient condition is

II. every nonempty subset $A' \subseteq A^*$ includes an action a' that is not weakly $S(k)$ supported against a^* by any action in A'.

In conditions I and II, "is not $S(k)$ supported by" and "is not weakly $S(k)$ supported by" can be replaced by "does not $S(k)$ support" and "does not weakly $S(k)$ support", respectively, as the conditions resulting from these replacements, I' and II', are equivalent to I and II.

Sketch of Proof. Suppose that there is a nonempty subset of actions $A' \subseteq A^*$, with cardinality $m \geq 1$, such that each action $a' \in A'$ is supported by some action in A'. Consider an initial population state α in which a fraction $1 - \epsilon$ of the agents play a^* and $\frac{\epsilon}{m}$ play each of the actions in A'. Since each $a' \in A'$ is supported by some action a in A', there is a probability of approximately $k(n-1)\frac{\epsilon}{m}$ of having action a appear in the sample and thus making

the mean payoff yielded by a' the highest one. It follows that $w_{a',k}(\alpha) = k(n-1)\frac{\epsilon}{m} > \frac{\epsilon}{m} = \alpha_{a'}$ for all $a' \in A'$. Thus the frequency of all actions in A' increases, which implies that a^* is not asymptotically stable. The formal proof (Appendix A.4) formalizes this intuition, by examining the Jacobian matrix at e_{a^*} and showing that it admits an eigenvalue greater than 1.

Next, suppose that every nonempty subset $A' \subseteq A^*$ includes an action a' that is not weakly $S(k)$ supported by any action in A'. Consider a state in which $1 - \epsilon$ of the agents play a^*. We know that there exists an action a' that is not weakly $S(k)$ supported by any action in A^*. This implies that the probability of action a' having the maximal mean payoff in an agent's sample is $O(\epsilon^2)$, and, thus, $w_{a',k}(\alpha) = O(\epsilon^2)$. As the frequency of a' becomes negligible, we can iterate the argument for $A' = A^*\backslash\{a'\}$, and find another action a'' for which $w_{a'',k}(\alpha) = O(\epsilon^2)$, etc. The formal proof (Appendix A.4) shows that (a) condition II implies that all the eigenvalues of the Jacobian matrix are negative, and (b) the phrase "is not weakly $S(k)$ supported by" can be replaced by "does not weakly $S(k)$ support." □

We remark that condition II is sufficient for asymptotic stability also when $k = 1$ and $n = 2$. However, condition I is not necessary in this case, as demonstrated by defection in the prisoner's dilemma.

The following corollary shows that any strict symmetric equilibrium is characterized by a threshold $k_0 < \infty$ that determines the equilibrium's asymptotic stability for any $k \geq 2$. (Note that the corollary does not give any information regarding $S(1)$ stability.)

Corollary 1. *Let a^* be a strict symmetric equilibrium action in a symmetric game. There exists an integer k_0 such that, for $k \geq 2$, action a^* is $S(k)$ asymptotically stable iff $k \geq k_0$.*

Proof. Let $\bar{k} \geq 2$ be a sufficiently large integer such that, for any action $a' \neq a^*$ and action profiles $(a^1, ..., a^n)$ and $(a'^1, ..., a'^n)$,

$$\bar{k} \cdot \left(u(a^*, ..., a^*) - u(a', a^*, ..., a^*) \right) > u(a^1, ..., a^n) - u(a'^1, ..., a'^n).$$

The inequality implies that there exists no pair of actions $a, a' \neq a^*$ such that action a weakly supports action a' against a^*, which in view of Theorem 2 implies that a^* is an $S(\bar{k})$ asymptotically stable equilibrium. Let k_0 be the smallest number for which a^* is an $S(k_0)$ asymptotically stable equilibrium. This implies that there is no pair of actions $a, a' \neq a^*$ such that action a $S(k_0)$ supports a' against a^*. The last observation at the end of Section 4.1 implies that for any $k > k_0$ there is no pair of actions $a, a' \neq a^*$ such that action a $S(k)$ supports a' against a^*, which implies that, for $k \geq 2$, action a^* is an $S(k)$ asymptotically stable equilibrium iff $k \geq k_0$. □

4.3 Applications

In this subsection we demonstrate the usefulness of the necessary and sufficient conditions identified in Theorem 2 by studying the $S(k)$ asymptotic stability of strict equilibria in a number of applications.

The first two applications deal with two-action games ($A = \{a^*, a'\}$). Observe that in such games condition I (resp., I') in Theorem 2 reduces to the requirement that the other action a' does not $S(k)$ (resp., weakly) support itself against a^*.

Prisoner's dilemma   Cooperation cannot support itself by spoiling, because it always increases the payoff of a defecting opponent. For $k \geq 2$, cooperation directly (resp., weakly) $S(k)$ supports itself against defection iff $l < \frac{1}{k-1}$ (resp., $\leq \frac{1}{k-1}$), because

$$u(c,c) + (k-1) \cdot u(c,d) > k \cdot u(d,d) \Leftrightarrow 1 + (k-1) \cdot (-l) > 0 \Leftrightarrow l < \frac{1}{k-1}.$$

By Theorem 2, the last finding implies that defection is asymptotically stable if $l > \frac{1}{k-1}$ and is not asymptotically stable if $l < \frac{1}{k-1}$, which proves Proposition 1.

Public good games   Consider a symmetric n-player game, with $n \geq 2$, where the set of actions is $A = \{c, nc\}$, with c and nc interpreted as contributing or not contributing, respectively, a fixed amount of some private good. The amount of public good produced is $\varphi(l)$, where l is the number of contributions and φ is a nondecreasing production function with $\varphi(0) = 0$ and $\varphi(1) < 1$. Each of the contributors gets a payoff of $\varphi(l) - 1$, while for a non-contributor the payoff is $\varphi(l)$. The assumption $\varphi(1) < 1$ implies that no one contributing is a strict equilibrium (and it is the unique equilibrium if $\varphi(l+1) - \varphi(l) < 1$ for all l). Observe that contributing cannot support itself by spoiling, and it directly $S(k)$ supports itself against non-contributing iff

$$u(c,c,nc,...,nc) + (k-1) \cdot u(c,nc,...,nc) > k \cdot u(nc,...,nc) \Leftrightarrow$$

$$\varphi(2) - 1 + (k-1) \cdot (\varphi(1) - 1) > 0 \Leftrightarrow k < 1 - \frac{1 - \varphi(2)}{1 - \varphi(1)}.$$

By Theorem 2, the above finding implies that if $n \geq 3$ or $k \geq 2$, then a sufficient condition for not contributing to be $S(k)$ asymptotically stable is that

$$k > 1 - \frac{1 - \varphi(2)}{1 - \varphi(1)}.$$

–15–

The corresponding weak inequality is a necessary condition.

Coordination games   Consider a symmetric n-player game, with $n \geq 2$, where the set of actions is $A = \{a_1, ..., a_m\}$, with $m \geq 2$. If all players choose the same action a_j, everyone gets payoff u_j, where $u_1 \geq u_2 \geq ... \geq u_m > 0$. If all players do not choose the same action, then everyone gets zero. The game admits m strict equilibria: everyone playing a_1, everyone playing a_2, ..., everyone playing a_m. Next we characterize the stability of the strict symmetric equilibrium action a_l, for any $1 \leq m$. No action $a_i \neq a_l$ $S(k)$ supports an action $a_j \neq a_l$ against a_l by spoiling, because

$$u(a_j, a_l, ..., a_l) = 0 < u(a_l, a_i, a_l, .., a_l) + (k - 1) \cdot u(a_l, ..., a_l) = (k - 1) \cdot u_l.$$

The only action that might directly $S(k)$ support $a_j \neq a_l$ against a_l is a_j itself, and this can happen only if $n = 2$. This is so because if $a_i \neq a_j$ or $n > 2$, then the following inequality holds:

$$u(a_j, a_i, a_l, ..., a_l) + (k - 1) \cdot u(a_j, a_l, ..., a_l) = 0 < k \cdot u(a_l, ..., a_l) = k \cdot u_l.$$

If $n = 2$, then a_j directly supports itself against a_l iff

$$u(a_j, a_j) + (k - 1) \cdot u(a_j, a_l) > k \cdot u(a_l, a_l) \Leftrightarrow u_j + 0 > k \cdot u_l.$$

By Theorem 2, this implies that all the strict equilibria are $S(k)$ asymptotically stable for any k if there are at least three players. In the two-player case, and for $k \geq 2$, the strict symmetric equilibrium action a_l is $S(k)$ asymptotically stable if $u_l > \frac{u_1}{k}$ and it is not $S(k)$ asymptotically stable if $u_l < \frac{u_1}{k}$.

4.4 Comparison with Sandholm et al. (2020)

We conclude this section by comparing Theorem 2 with the conditions for stability of strict equilibria presented in Sandholm et al. (2020, Section 5) (which, in turn, improve on the conditions presented in Sethi, 2000). For simplicity, the comparison focuses on the case of generic symmetric games.[4] As in Theorem 2, we use the notation $A^* \equiv A \backslash \{a^*\}$.

Sandholm et al. (2020) identify two necessary conditions for stability of a strict equilibrium. They are the negations of conditions 1 and 2 in the following proposition.

---

[4] In addition, our setting concerns only revising agents who test all feasible actions. The more general setting studied by Sandholm et al. allows dynamics in which revising agents test only some of the actions.

Adaptation of Proposition 5.4 (Sandholm et al., 2020)   Let a^* be a strict symmetric equilibrium action in a generic symmetric game. For $k \geq 2$, action a^* is *not* $S(k)$ asymptotically stable if either:

1. $\exists A' \subseteq A^*$ such that every $a' \in A'$ is directly $S(k)$ supported by some action in A'; or

2. $\exists A' \subseteq A^*$ such that every action $a' \in A'$ $S(k)$ supports some action in A'.

Theorem 2 strengthens this result by omitting "directly" from condition 1, thus weakening the condition. Moreover, it shows that this weaker condition (call it 1′) is actually equivalent to condition 2, and that both 1′ and 2 are in fact necessary *and sufficient* conditions for asymptotic stability.

Sandholm et al. (2020) present the following sufficient condition for stability.

Definition 6. Action a *tentatively $S(k)$ supports action a' by spoiling a^** if

$$k \cdot u(a', a^*, \ldots, a^*) > u(a^*, a, a^*, \ldots, a^*) + (k-1) \cdot u(a^*, \ldots, a^*).$$

Adaptation of Prop. 5.9 (Sandholm et al., 2020)   Let a^* be a strict symmetric equilibrium action in a generic symmetric game. For $k \geq 2$, action a^* is $S(k)$ asymptotically stable if

3. there exists an ordering of A^* such that no action a in this set directly $S(k)$ supports or tentatively $S(k)$ supports by spoiling a^* any weakly higher action a'.

Theorem 2 strengthens this result by omitting "tentatively" from condition 3, thus weakening the condition and making it *necessary* and sufficient for stability.[5] Sufficiency still holds because the weaker condition (call it 3′) implies that the lowest action in every subset $A' \subseteq A^*$ does not $S(k)$ support any action in A'. Necessity holds because it is not very difficult to see that 3′ is implied by condition I in Theorem 2. (Order A^* by recursively removing an element that is not $S(k)$ supported against a^* by any of the current elements.)

5. Asymmetric Games

In what follows we adapt our model and the characterization of $S(k)$ asymptotic stability to asymmetric games.

---

[5] Condition 3 is not necessary for asymptotic stability. In the symmetric two-player game defined by the following payoff matrix, action a^* is $S(2)$ asymptotically stable as it satisfies the condition in Theorem 2, yet it does not satisfy condition 3 due to action a'' tentatively supporting itself by spoiling.

a^*	8	9	3
a'	7	5	2
a''	6	4	1

Each player i has a finite set of actions A_i and a payoff function $u_i : \prod_{j=1}^{n} A_j \to \mathbb{R}$. The player is represented by a distinct population of agents, the i-population, whose state α^i is an element of the unit simplex in $\mathbb{R}^{|A_i|}$. The state of all n populations is given by $\alpha = (\alpha^1, \alpha^2, \ldots, \alpha^n) \in \Delta$, where Δ is the Cartesian product of the players' unit simplices.

The population state α determines for each player i the probability vector $w_k^i(\alpha(t))$ specifying the probability that each of the player's actions yields the highest mean payoff in k trials, employing some tie-breaking rule. For any $k \geq 1$, the k-payoff sampling dynamic is given by

(5.1) $\dot{\alpha}^i = w_k^i(\alpha(t)) - \alpha^i(t)$.

A population state α^* is an $S(k)$-equilibrium if $w_k^i(\alpha^*) = (\alpha^*)^i$ for each player i. Asymptotic stability and global asymptotic stability are defined as in the symmetric case.

The notion of supporting an action against an action profile $a^* = (a_1^*, a_2^*, \ldots, a_n^*)$ is conceptually similar to that in symmetric games. To present it in a formally similar way, consider the disjoint union $A^* = \bigcup_{i=1}^{n} (A_i \setminus \{a_i^*\})$. Each element of A^* is of the form a_i: a specific action of a specified player i such that $a_i \neq a_i^*$. For such an element, (a_i, a_{-i}^*) denotes the action profile in which player i plays a_i and all the other players play according to a^*. For $a_i, a_j \in A^*$ with $i \neq j$, (a_i, a_j, a_{-ij}^*) denotes the action profile in which player i plays a_i, player j plays a_j, and all the other players play according to a^*.

Definition 7. For an action profile a^* in an n-player game, and for $a_i, a_j \in A^*$:

1. action a_i *directly $S(k)$ supports* a_j *against* a^* if $i \neq j$ and

$$u_j(a_i, a_j, a_{-ij}^*) + (k-1) \cdot u_j(a_j, a_{-j}^*) > k \cdot u_j(a^*); \text{ and}$$

2. action a_i *$S(k)$ supports* a_j *by spoiling* a^* if $i \neq j$ and

$$k \cdot u_j(a_j, a_{-j}^*) > u_j(a_i, a_{-i}^*) + (k-1) \cdot u_j(a^*) \quad \text{and} \quad u_j(a_j, a_{-j}^*) > u_j(b_j, a_{-j}^*) \ \forall b_j \neq a_j \in A^*.$$

Action a_i *$S(k)$ single supports, double supports,* or just *supports* action a_j against a^* if exactly one of conditions 1 and 2, both conditions, or at least one condition, respectively, holds. *Weak $S(k)$ support,* and the related terms, are defined similarly, except that the strict inequalities in 1 and 2 are replaced by weak inequalities.

Next we adapt the characterization of $S(k)$ asymptotic stability to asymmetric games.

Theorem 3. *For $k \geq 2$, a necessary condition for a strict equilibrium a^* in an n-player game to be $S(k)$ asymptotically stable is that condition I (equivalently, I') in Theorem 2 holds, and a sufficient condition is that II (equivalently, II') holds.*

Obviously, the necessary conditions coincide with the sufficient ones if the game is generic, in the standard sense. The proof of the theorem, which is very similar to that of Theorem 2, is presented in Appendix A.5.

When the underlying game is symmetric, there are two different best experienced payoff dynamics that are applicable to it. The baseline, *one-population* dynamics presented in Section 2 (specifically, (2.2)) assumes a single population from which the players are sampled. Moreover, players are not assigned roles in the game; there is no player 1, player 2, etc. An alternative dynamics that can be applied to the game are the *n-population* dynamics 5.1. Although meant for asymmetric games, they can be used to study symmetric games in which the players are arbitrarily numbered, with the *i*-population representing player[6] *i*.

In other evolutionary dynamics it is often the case that stability under the one-population dynamics is not equivalent to stability under the *n*-population dynamics. (For example, it is well known that the mixed equilibrium of a hawk-dove game is stable under the one-population replicator dynamics but is not stable under the two-population replicator dynamics.) Our next result shows that this is not the case here.

Corollary 2. *For $k \geq 2$, a strict symmetric equilibrium action a^* in a symmetric n-player game is $S(k)$ asymptotically stable under the one-population dynamics (2.2) if and only if everyone playing a^* is $S(k)$ asymptotically stable under the n-population dynamics (5.1).*

The simple proof, which is given in Appendix A.7, relies on the fact that our two definitions of an action supporting another action against a strict symmetric equilibrium (action) essentially coincide when the underlying game is symmetric.

5.1 Applications

We conclude this section with demonstrating the usefulness of Theorem 3 by applying it to the study of $S(k)$ asymptotic stability of strict equilibria in asymmetric prisoner's dilemma and hawk-dove games.

---

[6]The *n*-population dynamics can also capture environments in which players from a single population play in all roles, the roles are observable, and a player conditions her action on her role.

Asymmetric Prisoner's Dilemma		
	c_2	d_2
c_1	1 , 1	$-l_1$, $1+g_2$
d_1	$1+g_1$, $-l_2$	0 , 0

Asymmetric Hawk-Dove		
	D_2	H_2
D_1	1 , 1	l_1 , $1+g_2$
H_1	$1+g_1$, l_2	0 , 0

Table 2: Payoff Matrices of Asymmetric Games ($g_1, g_2, l_1, l_2 > 0$; in hawk-dove, also $l_1, l_2 < 1$)

Asymmetric prisoner's dilemma   The left-hand side of Table 2 presents the payoff matrix of an asymmetric prisoner's dilemma, in which the unique equilibrium is $d = (d_1, d_2)$, mutual defection. Action c_1 cannot support c_2 by spoiling, because cooperation increases the payoff of a defecting opponent. It directly (weakly) supports c_2 against d iff $l_2 < \frac{1}{k-1}$ (respectively, $\leq \frac{1}{k-1}$), because

$$u_2(c_1, c_2) + (k-1) \cdot u_2(d_1, c_2) > k \cdot u_2(d) \Leftrightarrow 1 + (k-1) \cdot (-l_2) > 0.$$

The same holds with 1 and 2 interchanged. This implies, by Theorem 3, that for any $k \geq 2$ mutual defection is $S(k)$ asymptotically stable if $\max(l_1, l_2) > \frac{1}{k-1}$ and is not $S(k)$ asymptotically stable if $\max(l_1, l_2) < \frac{1}{k-1}$. Note that, if $l_1 = l_2$, then in accordance with Corollary 2 the condition for stability coincides with that of the one-population model in the symmetric prisoner's dilemma.

Asymmetric hawk-dove   The right-hand side of Table 2 presents the payoff matrix of the asymmetric hawk-dove game. The game admits two strict equilibria, (D_1, H_2) and (H_1, D_2), in which one player plays hawk and the other plays dove. Observe that action D_2 can $S(k)$ support action H_1 against (D_1, H_2) only by direct support, which is obtained iff

$$u_1(H_1, D_2) + (k-1) \cdot u_1(H_1, H_2) > k \cdot u_1(D_1, H_2) \Leftrightarrow 1 + g_1 > k \cdot l_1.$$

Action H_1 can $S(k)$ support action D_2 against (D_1, H_2) only by spoiling, which is obtained iff

$$k \cdot u_2(D_1, D_2) > u_2(H_1, H_2) + (k-1) \cdot u_2(D_1, H_2) \Leftrightarrow k > (k-1) \cdot (1+g_2) \Leftrightarrow g_2 < \frac{1}{k-1}.$$

By Theorem 3, this implies that (D_1, H_2) is asymptotically stable if $g_1 < k \cdot l_1 - 1$ or $g_2 > \frac{1}{k-1}$ and is not asymptotically stable if $g_1 > k \cdot l_1 - 1$ and $g_2 < \frac{1}{k-1}$. The conditions for asymptotic stability of the other strict equilibrium are obtained by interchanging 1 and 2.

Appendix

A. Proofs

A.1 Proof of Claim 2

Recall that $p \equiv \alpha_c$ and $1 - p \equiv \alpha_d$ denote the proportion of agents in the population playing actions c and d, respectively. An agent's c-sample includes k actions of the opponents. If $j \in \{0, ..., k\}$ of these are c, then the agent's mean payoff (when playing c) is $j - (k - j)l$. Similarly, if $j' \in \{0, ..., k\}$ of the sampled actions in the agent's d-sample are c, then the mean payoff (when playing action d) is $j'(1 + g)$. The difference between the two payoffs is

$$(j - (k - j)l) - j'(1 + g) = j - j' - ((k - j)l + j'g).$$

This expression is clearly negative if $j \leq j'$. If $j \geq j' + 1$, it is positive, since $(k - j)l + j'g \leq \frac{k-j+j'}{k-1} \leq 1$ and the first inequality becomes an equality only if $j = k$ and $j' = 0$ while the second one does so only if $j = j' + 1$.

These conclusions prove that the c-sample yields a superior payoff iff it includes more cooperations than the d-sample does. As the number of cooperators in each sample has a binomial distribution with parameters k and p, we conclude that $w_{c,k} = Win(k, p)$.

A.2 Proof of Claim 3

A binomial random variable with parameters k, p has a degenerate distribution if $p \in \{0, 1\}$, and so $Win(k, 0) = Win(k, 1) = 0$. From Eq. (3.2) we get $h_k(0) = 0 - 0 = 0, h_k(1) = 0 - 1 = -1$. To show that $h'_k(0) > 0$ for $k = 2, 3, 4, \ldots$, we use the fact that

$$h_k(p) = Win(k, p) - p = 0.5(1 - Tie(k, p)) - p = 0.5\left(1 - \sum_{j=0}^{k} (f_{k,p}(j))^2\right) - p$$

$$= 0.5(1 - ((1 - p)^{2k} + O(p^2)) - p \quad (O(p^2) \text{ denotes the terms with degree} \geq 2)$$

$$= 0.5(1 - (1 - 2pk + O(p^2)) - p + O(p^2) = kp - p + O(p^2) = (k - 1)p + O(p^2).$$

From the above expression, it follows that $h'_k(0) = k - 1 > 0$ for $k > 1$.

A.3 Proof of Claim 4

Let $\{A_{j,p}\}_{j=1}^k$, $\{B_{j,p}\}_{j=1}^k$ be $2k$ independent Bernoulli random variables with parameter p. Then $X_{k,p} = \sum_{i=1}^k A_{j,p}$ and $Y_{k,p} = \sum_{j=1}^k B_{j,p}$ are i.i.d. binomial random variables with parameters k and p. Eq. (3.2) can be expressed as

$$
\begin{aligned}
h_k(p) &= P(X_{k,p} > Y_{k,p}) - p \\
&= \frac{1}{2}\left(P(X_{k,p} > Y_{k,p}) + P(X_{k,p} < Y_{k,p})\right) - p \quad \text{(since } X_{k,p} \text{ and } Y_{k,p} \text{ are } i.i.d.\text{)} \\
&= \frac{1}{2}\left(1 - P(X_{k,p} = Y_{k,p})\right) - p.
\end{aligned}
$$

(A.1)

For $j = 1, 2, \ldots, k$, let $Z_{j,p} = A_{j,p} - B_{j,p}$. Clearly, $\{Z_{j,p}\}_{j=1}^n$ are i.i.d., with distribution given by

$$
P(Z_{j,p} = -1) = P(Z_{j,p} = 1) = pq \text{ and } P(Z_{j,p} = 0) = 1 - 2pq,
$$

where $q = 1 - p$.

Consider the characteristic function $\varphi_k(\cdot; p)$ of the random variable $Z_p^k = X_{k,p} - Y_{k,p} = \sum_{j=1}^k Z_{j,p}$:

$$
\begin{aligned}
\varphi_k(t; p) &= \mathbb{E}\left[e^{itZ_p^k}\right] = \mathbb{E}\left[e^{it\sum_{j=1}^k Z_{j,p}}\right] = \left(\mathbb{E}[e^{itZ_{1,p}}]\right)^k \\
&= \left(e^{it(-1)}pq + e^{it(1)}pq + e^{it(0)}(1 - 2pq)\right)^k = \left(1 + pq(e^{-it} + e^{it} - 2)\right)^k \\
&= (1 + 2pq(\cos(t) - 1))^k \qquad (\text{since } e^{-it} + e^{it} = 2\cos(t)) \\
&= \left(1 - 4p(1 - p)\sin^2\left(\frac{t}{2}\right)\right)^k \qquad \left(\text{since } q = 1 - p \text{ and } \cos(t) = 1 - 2\sin^2\left(\frac{t}{2}\right)\right).
\end{aligned}
$$

The base of the last exponent, with power k, is an expression that is convex as a function of p, lies between 0 and 1 and, if $0 < p < 1$ and t is not a whole multiple of π, is different from 0 and 1. Therefore, the same is true for $\varphi_k(t; p)$ and, if $0 < p < 1$ and t is not a whole multiple of π, $\varphi_k(t; p) > \varphi_{k+1}(t; p)$ and $\lim_{k \to \infty} \varphi_k(t; p) = 0$. It follows, in view of Eq. (A.1) and the fact that (see Fact 1)

$$
P(X_{k,p} = Y_{k,p}) = P(Z_p^k = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_k(t; p)dt,
$$

that the function h_k is concave and, for $0 < p < 1$, the sequence $(h_k(p))_{k=1}^\infty$ is strictly increasing and converges to $\frac{1}{2} - p$. This completes the proof.

Fact 1. $P(Z_p^k = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_k(t; p) dt.$

Proof of Fact 1. From the definition of $\varphi_k(t; p)$, we have the following:

$$\int_{-\pi}^{\pi} \varphi_k(t; p) dt = \int_{-\pi}^{\pi} \mathbb{E}\left[e^{itZ_p^k}\right] dt = \mathbb{E}\left[\int_{-\pi}^{\pi} e^{itZ_p^k} dt\right] = \mathbb{E}\left[\int_{-\pi}^{\pi} e^{itZ_p^k}\left(\mathbb{1}_{\{Z_p^k=0\}} + \mathbb{1}_{\{Z_p^k \neq 0\}}\right) dt\right]$$

$$= \mathbb{E}\left[\int_{-\pi}^{\pi} e^{itZ_p^k} \mathbb{1}_{\{Z_p^k=0\}} dt\right] + \mathbb{E}\left[\int_{-\pi}^{\pi} e^{itZ_p^k} \mathbb{1}_{\{Z_p^k \neq 0\}} dt\right]$$

$$= \mathbb{E}\left[\int_{-\pi}^{\pi} 1 \cdot \mathbb{1}_{\{Z_p^k=0\}} dt\right] + \mathbb{E}[0] = \mathbb{E}\left[2\pi \cdot \mathbb{1}_{\{Z_p^k=0\}}\right] = 2\pi \cdot P(Z_p^k = 0).$$

From the above series of equalities, it follows that $P(Z_p^k = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_k(t; p) dt.$

A.4 Proof of Theorem 2

For completeness, we present all details of the proof, although various steps are analogous to arguments presented in the proofs of Sandholm et al. (2020, Section 5).

Suppose, first, that the game is generic. Let T be the nonnegative $|A^*| \times |A^*|$ matrix whose element in row a' and column a is

$$T_{a'a} = \begin{cases} 2 \text{ if action } a \text{ double } S(k) \text{ supports action } a' \text{ against } a^* \\ 1 \text{ if action } a \text{ single } S(k) \text{ supports action } a' \text{ against } a^* \\ 0 \text{ if action } a \text{ does not } S(k) \text{ support action } a' \text{ against } a^* \end{cases}$$

We will now compute the Jacobian of the k-payoff sampling dynamic (2.2) in the (monomorphic population state corresponding to the) strict symmetric equilibrium action a^*. Suppose that the frequency α_{a^*} of action a^* in the population is $1 - \epsilon$, where $\epsilon > 0$ is a small number. Denote by $\alpha^* \equiv \alpha|_{A^*}$ the frequencies of the actions in the set A^*, that is, $\alpha_a^* = \alpha_a$ for all $a \in A^*$. Clearly, $\sum_{a \in A^*} \alpha_a^* = \epsilon$, which implies that the Euclidean norm of α^* is of order ϵ, that is, $|\alpha^*| = (\sum_{a \in A^*} \alpha_a^{*2})^{\frac{1}{2}} = O(\epsilon)$.

The probability that a non-equilibrium action $a' \in A^*$ yields the best payoff is roughly equal to the probability that it yields a higher payoff than the strict symmetric equilibrium action a^* does when both actions are tested k times in the population state α. When $a' \in A^*$ is tested k times, with a very high probability (of $(1 - \epsilon)^{k(n-1)}$) it encounters the equilibrium action a^* each time. The probability that a^* is encountered $k - 2$ or fewer times is of order ϵ^2 or higher, and these higher-order terms can be neglected for stability analysis. Action a' yields a higher payoff than a^* does in the following two cases:

–23–

Case 1: When a' is tested, one of the $k(n-1)$ opponents plays a non-equilibrium action a and the remaining opponents play a^*. Action a' obtains the highest mean payoff if a directly $S(k)$ supports it against a^*, that is, $u(a',a,a^*,\ldots,a^*) + (k-1) \cdot u(a',a^*,\ldots,a^*) > k \cdot u(a^*,a^*,\ldots,a^*)$.

Case 2: When a^* is tested, one of the $k(n-1)$ opponents plays a' and the others play a^*. Action a' obtains the maximal mean payoff if a $S(k)$ supports it by spoiling a^*, that is, $k \cdot u(a',a^*,\ldots,a^*) > u(a^*,a,a^*,\ldots,a^*) + (k-1) \cdot u(a^*,\ldots,a^*)$ and $u(a',a^*,\ldots,a^*) > u(b,a^*,\ldots,a^*)$ for all $b \notin \{a',a^*\}$.

The probability that action a' yields the best payoff is therefore given by

(A.2) $\qquad w_{a',k}(\alpha) = k(n-1) \sum_{a \in A^*} T_{a'a}\alpha_a + O(|\alpha^*|^2),$

and the k-payoff sampling dynamic (2.2) can be written as

$$\dot{\alpha}_{a'} = w_{a',k}(\alpha) - \alpha_{a'} = k(n-1) \sum_{a \in A^*} T_{a'a}\alpha_a - \alpha_{a'} + O(|\alpha^*|^2).$$

In matrix notation,

(A.3) $\qquad \dot{\alpha}^* = f(\alpha^*) \equiv (k(n-1)T - I)\alpha^* + O(|\alpha^*|^2),$

where I is the $|A^*| \times |A^*|$ identity matrix and $O(|\alpha^*|^2)$ here is an $|A^*|$-dimensional vector with elements of order $|\alpha^*|^2$ or higher. Let J denote the Jacobian matrix of f evaluated at the origin:

$$J = \left.\frac{\partial f(\alpha^*)}{\partial \alpha^*}\right|_{\alpha^* = \underbrace{(0,0,\ldots,0)}_{|A^*| \text{ zeros}}} = k(n-1)T - I.$$

The asymptotic stability of the system (A.3) can be analyzed by examining the eigenvalues of the Jacobian matrix J.

A sufficient condition for a^* to be $S(k)$ asymptotically stable is that all the eigenvalues of J have negative real parts (see, e.g., Sandholm, 2010, Corollary 8.C.2). A sufficient condition for it *not* to be $S(k)$ asymptotically stable is that at least one of the eigenvalues has a positive real part. The first condition holds, in particular, if the only eigenvalue of T is zero, in other words, if the spectral radius ρ of that matrix is 0, as this condition means that the only eigenvalue of J is -1. The second condition holds if $\rho \geq 1$. This is because the

spectral radius of a nonnegative matrix is an eigenvalue with a nonnegative eigenvector (Horn and Johnson, 1985, Theorem 8.3.1), and so $\rho \geq 1$ implies that J has the eigenvalue $k(n-1)\rho - 1 \geq 2 \cdot 1 - 1 > 0$, with a corresponding nonnegative eigenvector. It therefore suffices to show that $\rho = 0$ holds if condition I in the theorem holds and $\rho \geq 1$ holds if the condition does not hold. Observe that conditions I and I' can be rephrased as follows: every principal submatrix of T has a row or column, respectively, where all entries are zero. Therefore, to complete the proof of the theorem, it remains only to establish the following fact (which in particular proves the equivalence of I and I').

Fact 2. *Let M be a square matrix of nonnegative integers, and let ρ be its spectral radius. If every principal submatrix of M has a row of zeros, then $\rho = 0$. Otherwise, $\rho \geq 1$. The same is true when we replace "row" by "column."*

Proof of Fact 2. Suppose that $\rho > 0$. Let v be a corresponding nonnegative right eigenvector. Since $(Mv)_i = \rho v_i > 0$ for every index i with $v_i > 0$, the set α of all such indices defines a principal submatrix M' (obtained from M by deleting all rows and all columns with indices not in α) with the property that every row includes at least one nonzero entry.

Conversely, suppose that M has a principal submatrix M', defined by some set of indices α, with the above property. Let v be a column vector of 0's and 1's where an entry is 1 if and only if its index lies in α. It is easy to see that $Mv \geq v$. This vector inequality implies that $\rho \geq 1$ (Horn and Johnson, 1985, Theorem 8.3.2).

The proof for "column" is obtained by replacing M by its transpose.

We now drop the assumption that the game is generic. This change means that cases 1 and 2 need to be extended by replacing the respective strict inequalities with weak ones. When a weak inequality in case 1 or 2 holds as equality, the probability of moving to action a' depends on the tie-breaking rule. Adding these probabilities to $T_{a'a}$, for all $a', a \in A^*$, defines a new matrix $\bar{T} \geq T$, which replaces T in (A.3). If the probabilities were replaced by 1's, the result would be a matrix $\bar{\bar{T}} \geq \bar{T}$, where all entries are integers. In view of Fact 2, conditions II and II' are both equivalent to the condition that the spectral radius of $\bar{\bar{T}}$ is 0. That condition implies that the spectral radius of $\bar{T}$ is also 0, which implies that a^* is $S(k)$ asymptotically stable. If condition I or I' does not hold, then the spectral radius of T is 1 or greater. The same is then true for $\bar{T}$, which implies that a^* is not $S(k)$ asymptotically stable.

A.5 Proof of Theorem 3

Let T be the $|A^*| \times |A^*|$ matrix defined exactly as in the proof of Theorem 2, and denote by $\alpha^* \equiv \alpha|_{A^*}$ the vector of dimension $\sum_{i=1}^{n}(|A_i| - 1)$ whose components are the frequencies of the actions in the set A^* (that is, the players' non-equilibrium actions). By arguments similar to those employed in that proof, if the game is generic, then the Jacobian J of the k-payoff sampling dynamic (5.1) is given by $J = kT - I$, so that

$$\dot{\alpha}^* = (kT - I)\alpha^* + O(|\alpha^*|^2).$$

The rest of the proof, including the treatment of the non-generic case, is essentially the same as for Theorem 2.

A.6 Proof of Proposition 2: Global Stability for $k \in \{2, 3\}$

In this subsection, we characterize the $S(2)$ and $S(3)$ globally asymptotically stable equilibria explicitly for all parameter configurations g, l in the prisoner's dilemma.

$S(2)$ analysis: In Theorem 1 we solved the case $g, l < 1$. Here, we consider the remaining three cases. Recall that, when testing each action twice, we have:

$$\text{When a player samples } c \text{ she} \begin{cases} \text{gets 2 with probability } p^2 \\ \text{gets } 1 - l \text{ with probability } 2p(1 - p) \\ \text{gets } -2l \text{ with probability } (1 - p)^2 \end{cases}$$

$$\text{When a player samples } d \text{ she} \begin{cases} \text{gets } 2(1 + g) \text{ with probability } p^2 \\ \text{gets } 1 + g \text{ with probability } 2p(1 - p) \\ \text{gets 0 with probability } (1 - p)^2 \end{cases}$$

Case I: $l < 1 < g$. Action c has a higher mean payoff iff the c-sample includes at least one cooperation and the d-sample does not include any cooperation. Thus, the 2-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^2(1 - p)^2 + 2p(1 - p)(1 - p)^2 - p = (1 - p)^2(p^2 + 2p(1 - p)) - p$$
$$= (1 - p)^2(1 - (1 - p)^2) - p = (1 - p)^2 - (1 - p)^4 - p.$$

The rest points of the above dynamic are 0 and 0.245. It is straightforward to verify that 0

is unstable and that 0.245 is globally stable.

Case II: $g < 1 < l$. Action c has a higher mean payoff iff the c-sample includes two cooperations, while the d-sample includes at most one cooperation. The 2-*payoff sampling dynamic* in this case is given by $\dot{p} = p^2(1 - p^2) - p$. The unique rest point is 0, which is globally stable.

Case III: $g, l > 1$. Action c has a higher mean payoff iff the c-sample includes two cooperations, while the d-sample does not include any cooperation. The 2-*payoff sampling dynamic* in this case is given by $\dot{p} = p^2(1 - p)^2 - p$. As in the previous case, the unique rest point is 0, which is globally stable.

$S(3)$ analysis: In Theorem 1 we solved the case $g, l < \frac{1}{2}$. Here, we consider the remaining cases. Recall that, when testing each action thrice, we have:

When a player samples c she $\begin{cases} \text{gets } 3 \text{ with probability } p^3 \\ \text{gets } 2 - l \text{ with probability } 3p^2(1 - p) \\ \text{gets } 1 - 2l \text{ with probability } 3p(1 - p)^2 \\ \text{gets } - 2l \text{ with probability } (1 - p)^3 \end{cases}$

When a player samples d she $\begin{cases} \text{gets } 3(1 + g) \text{ with probability } p^3 \\ \text{gets } 2(1 + g) \text{ with probability } 3p^2(1 - p) \\ \text{gets } 1 + g \text{ with probability } 3p(1 - p)^2 \\ \text{gets } 0 \text{ with probability } (1 - p)^3 \end{cases}$

Case I: $l < \frac{1}{2}, \frac{1}{2} < g < 2$, and $g + l < 1$. Action c has a higher mean payoff iff, when the c-sample includes at least two cooperations, the d-sample includes at most one cooperation, or when the c-sample includes exactly one cooperation, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(3p(1-p)^2 + (1-p)^3) + 3p^2(1-p)(3p(1-p)^2 + (1-p)^3) + 3p(1-p)^2(1-p)^3 - p$$
$$= p^2(1-p)^2(3 - 2p)(1 + 2p) + 3p(1-p)^5 - p.$$

The rest points of the above dynamic are 0 and 0.323. It is straightforward to verify that 0 is unstable and that 0.323 is globally stable.

Case II: $l < \frac{1}{2}, \frac{1}{2} < g < 2$, and $g + l > 1$. Action c has a higher mean payoff iff, when the c-sample includes three cooperations, the d-sample includes at most one cooperation, or

when the c-sample includes either one or two cooperations, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(3p(1-p)^2 + (1-p)^3) + 3p^2(1-p)(1-p)^3 + 3p(1-p)^2(1-p)^3 - p$$
$$= p^3(1-p)^2(1+2p) + 3p(1-p)^4 - p.$$

The rest points of the above dynamic are 0 and 0.250. It is straightforward to verify that 0 is unstable and that 0.250 is globally stable.

Case III: $l < \frac{1}{2}, g > 2$. Action c has a higher mean payoff iff, when the c-sample includes at least one cooperation, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(1-p)^3 + 3p^2(1-p)(1-p)^3 + 3p(1-p)^2(1-p)^3 - p$$
$$= (1 - (1-p)^3)(1-p)^3 - p = (1-p)^3 - (1-p)^6 - p.$$

The rest points of the above dynamic are 0 and 0.245. It is straightforward to verify that 0 is unstable and that 0.245 is globally stable.

Case IV: $\frac{1}{2} < l < 2, g < \frac{1}{2}$ and $g + l < 1$. Action c has a higher mean payoff iff, when the c-sample includes three cooperations, the d-sample includes at most two cooperations, or when the c-sample includes exactly two cooperations, the d-sample includes at most one cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(1-p^3) + 3p^2(1-p)(3p(1-p)^2) + (1-p)^3) - p$$
$$= p^3(1-p^3) + 3p^2(1-p)^3(1+2p) - p.$$

The unique rest point is 0, which is globally stable.

Case V: $\frac{1}{2} < l < 2, g < \frac{1}{2}$, and $g + l > 1$. Action c has a higher mean payoff iff, when the c-sample includes three cooperations, the d-sample includes at most two cooperations, or when the c-sample includes exactly two cooperations, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(1-p^3) + 3p^2(1-p)(1-p)^3 - p = p^3(1-p^3) + 3p^2(1-p)^4 - p.$$

The unique rest point is 0, which is globally stable.

Case VI: $\frac{1}{2} < l < 2, \frac{1}{2} < g < 2$. Action c has a higher mean payoff if, when the c-sample includes three cooperations, the d-sample includes at most one cooperation, or when the c-

sample includes exactly two cooperations, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(3p(1-p)^2 + (1-p)^3) + 3p^2(1-p)(1-p)^3 - p$$
$$= p^3(1-p)^2(1+2p) + 3p^2(1-p)^4 - p.$$

The unique rest point is 0, which is globally stable.

Case VII: $\frac{1}{2} < l < 2, g > 2$. Action c has a higher mean payoff iff, when the c-sample includes at least two cooperations, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(1-p)^3 + 3p^2(1-p)(1-p)^3 - p = p^2(1-p)^3(3-2p) - p.$$

The unique rest point is 0, which is globally stable.

Case VIII: $l > 2, g < \frac{1}{2}$. Action c has a higher mean payoff iff, when the c-sample includes three cooperations, the d-sample includes at most two cooperations. Thus, the 3-*payoff sampling dynamic* in this case is given by $\dot{p} = p^3(1-p^3) - p$. The unique rest point is 0, which is globally stable.

Case IX: $l > 2, \frac{1}{2} < g < 2$. Action c has a higher mean payoff iff, when the c-sample includes three cooperations, the d-sample includes at most one cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by

$$\dot{p} = p^3(3p(1-p)^2 + (1-p)^3) - p = p^3(1-p)^2(1+2p) - p.$$

The unique rest point is 0, which is globally stable.

Case X: $l > 2, g > 2$. Action c has a higher mean payoff iff, when the c-sample includes three cooperations, the d-sample does not include any cooperation. Thus, the 3-*payoff sampling dynamic* in this case is given by $\dot{p} = p^3(1-p)^3 - p$. The unique rest point is 0, which is globally stable.

A.7 Proof of Corollary 2

Suppose that action a^* is not $S(k)$ asymptotically stable under the one-population dynamics, and so there is a subset $A' \subseteq A \backslash \{a^*\}$ such that all actions in A' are supported against a^* by actions in A'. Let $\bar{A}' = \dot{\bigcup}_{i=1}^{n} A'$ be the disjoint union of n copies of A'. It follows immediately from Definitions 5 and 7 and the symmetry of the game that all actions in $\bar{A}'$

Figure 2: Values of g and l in the 29 Experiments Summarized in Mengel (2018, Table A.3)
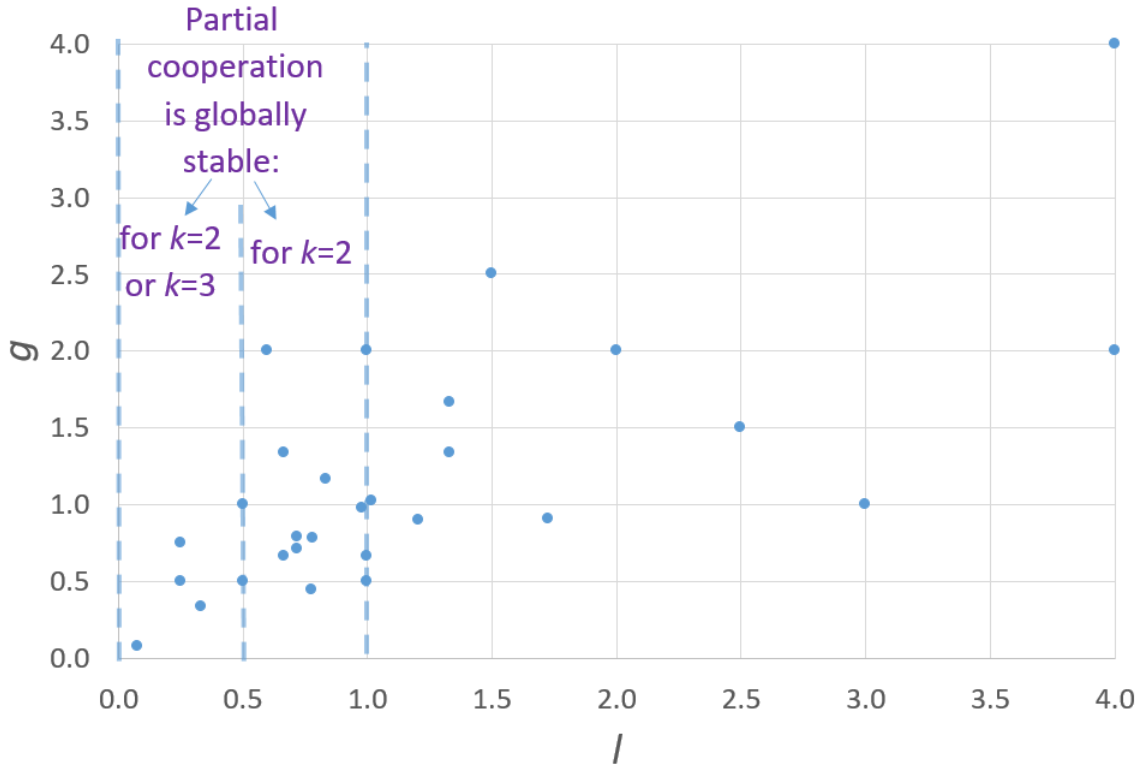
are supported by actions in $\bar{A}'$ against $\bar{a}^* \equiv (a^*, a^*, \ldots, a^*)$. Therefore, the strategy profile $\bar{a}^*$ is not $S(k)$ asymptotically stable under the n-population dynamics.

Conversely, suppose that the last conclusion holds, so that there is a subset $A' \subseteq A^* \equiv \dot{\bigcup}_{i=1}^{n}(A\backslash\{a^*\})$ such that all actions in A' are supported by actions in A' against $\bar{a}^*$. Let $\bar{A}' = \{a \in A \mid$ there is some player i and a corresponding action $a_i \in A'$ with $a = a_i\}$ be the set of all actions that are included in A' for at least one player. It is easy to see that all actions in $\bar{A}'$ are supported against a^* by actions in $\bar{A}'$, and so action a^* is not $S(k)$ asymptotically stable under the one-population dynamics.

B. Values of g and l in Prisoner's Dilemma Experiments

Figure 2 shows the values of g and l in the 29 experiments of the one-shot prisoner's dilemma (taken from 16 papers) as summarized in the meta-study of Mengel (2018, Table A.3). The figure shows that most of these experiments satisfy the condition for global stability of partial cooperation for $k = 2$ (namely, $l < 1$), and quite a few of them also satisfy the condition for $k = 3$ ($l < 0.5$).

References

Burton-Chellew, M. N., El Mouden, C., and West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society B: Biological Sciences*, 284:20170067.

Cárdenas, J., Mantilla, C., and Sethi, R. (2015). Stable sampling equilibrium in common pool resource games. *Games*, 6(3):299–317.

Chmura, T. and Güth, W. (2011). The minority of three-game: An experimental and theoretical analysis. *Games*, 2(3):333–354.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Heller, Y. and Mohlin, E. (2018). Social learning and the shadow of the past. *Journal of Economic Theory*, 177:426–460.

Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press.

Kosfeld, M., Droste, E., and Voorneveld, M. (2002). A myopic adjustment process leading to best-reply matching. *Games and Economic Behavior*, 40(2):270–298.

Kreindler, G. E. and Young, H. P. (2013). Fast convergence in evolutionary equilibrium selection. *Games and Economic Behavior*, 80:39–67.

Mantilla, C., Sethi, R., and Cárdenas, J. C. (2018). Efficiency and stability of sampling equilibrium in public goods games. *Journal of Public Economic Theory*, 22(2):355–370.

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

Mengel, F. (2018). Risk and temptation: A meta-study on prisoner's dilemma games. *The Economic Journal*, 128(616):3182–3209.

Miękisz, J. and Ramsza, M. (2013). Sampling dynamics of a symmetric ultimatum game. *Dynamic Games and Applications*, 3(3):374–386.

Nax, H. H., Burton-Chellew, M. N., West, S. A., and Young, H. P. (2016). Learning in a black box. *Journal of Economic Behavior & Organization*, 127:1–15.

Nax, H. H. and Perc, M. (2015). Directional learning and the provisioning of public goods. *Scientific Reports*, 5(1):1–6.

Osborne, M. J. and Rubinstein, A. (1998). Games with procedurally rational players. *American Economic Review*, 88(4):834–847.

Oyama, D., Sandholm, W. H., and Tercieux, O. (2015). Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics*, 10(1):243–281.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302.

Ramsza, M. (2005). Stability of pure strategy sampling equilibria. *International Journal of Game Theory*, 33(4):515–521.

Rowthorn, R. and Sethi, R. (2008). Procedural rationality and equilibrium trust. *The Economic Journal*, 118(530):889–905.

Rustichini, A. (2003). Equilibria in large games with continuous procedures. *Journal of Economic Theory*, 111(2):151–171.

Salant, Y. and Cherry, J. (2020). Statistical inference in games. *Econometrica*, 88(4):1725–1752.

Sandholm, W. H. (2001). Almost global convergence to p-dominant equilibrium. *International Journal of Game Theory*, 30(1):107–116.

Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. MIT Press.

Sandholm, W. H., Izquierdo, S. S., and Izquierdo, L. R. (2019). Best experienced payoff dynamics and cooperation in the centipede game. *Theoretical Economics*, 14(4):1347–1385.

Sandholm, W. H., Izquierdo, S. S., and Izquierdo, L. R. (2020). Stability for best experienced payoff dynamics. *Journal of Economic Theory*, 185:104957.

Sethi, R. (2000). Stability of equilibria in games with procedurally rational players. *Games and Economic Behavior*, 32(1):85–104.

Sethi, R. (2019). Procedural rationality in repeated games. Unpublished manuscript, Barnard College, Columbia University.

Spiegler, R. (2006a). Competition over agents with boundedly rational expectations. *Theoretical Economics*, 1(2):207–231.

Spiegler, R. (2006b). The market for quacks. *The Review of Economic Studies*, 73(4):1113–1131.