Strategy sets closed under payoff sampling

Segismundo S. Izquierdo[a,*] and Luis R. Izquierdo[b]

[a] *Department of Industrial Organization and BioEcoUva, Universidad de Valladolid, Spain.*
[b] *Department of Management Engineering, Universidad de Burgos, Spain.*

Abstract

We consider population games played by procedurally rational players who, when revising their current strategy, test each of their available strategies independently in a series of random matches –i.e., a battery of tests–, and then choose the strategy that performed best in this battery of tests. This revision protocol leads to the so-called *payoff-sampling* dynamics (aka *test-all Best Experienced Payoff* dynamics).

In this paper we characterize the support of all the rest points of these dynamics in any game and analyze the asymptotic stability of the faces to which they belong. We do this by defining strategy sets *closed under payoff sampling*, and by proving that the identification of these sets can be made in terms of simple comparisons between some of the payoffs of the game.

JEL classification numbers: C72, C73.

Keywords: Payoff sampling; Best Experienced Payoff; Procedural rationality; Asymptotically stable sets; Closed Under Payoff Sampling

1 Introduction

In the context of players with bounded rationality or limited information, Osborne and Rubinstein (1998) introduced a decision rule for *procedurally rational* players. In a population game setting where players use pure strategies only (Sandholm, 2010), these procedurally rational players revise their pure strategy as follows. They first associate one payoff to each of their possible strategies by –literally or virtually– conducting a *battery of tests*. This battery of tests consists on testing each of their available strategies independently in κ matches or *trials*, with each trial involving a new set of randomly drawn co-players from the population. Revising players then choose the strategy that obtained the best average experienced payoff in the battery of tests.

An equilibrium under this procedure is a population state x such that the proportion of players using each strategy a_i (i.e., x_i) equals the probability of strategy a_i being

*Correspondence to: Department of Industrial Organization, Universidad de Valladolid, Dr. Mergelina s/n, 47011 Valladolid, Spain. *e-mail*: segis@eii.uva.es.

Abbreviations. BEP: Best Experienced Payoff; CUPS: Closed Under Payoff Sampling; CURB: Closed Under Rational Behavior.

selected as the best-experienced-payoff strategy in a κ-trial battery of tests conducted at state x. Using $w_i^\kappa(x)$ to denote this probability, a population state x is an equilibrium under procedural rationality if $x_i = w_i^\kappa(x)$ for every strategy a_i.

This equilibrium has been named *payoff-sampling equilibrium* in the literature (see e.g. Selten and Chmura (2008), Chmura and Güth (2011), Cárdenas et al. (2015), Sethi (2021) and Arigapudi et al. (2021, 2022)). When defining this equilibrium, most authors (e.g. Osborne and Rubinstein (1998); Selten and Chmura (2008); Cárdenas et al. (2015)) assume that ties are broken uniformly at random and use $S(\kappa)$ to denote the resulting (*uniform*) payoff-sampling equilibrium.[1]

Procedurally rational agents and their associated payoff-sampling equilibria have been used in a variety of applications, including consumer choice procedures and product pricing strategies (Spiegler, 2006a), markets with asymmetric information (Spiegler, 2006b), trust and delegation of control (Rowthorn and Sethi, 2008), the Traveler's Dilemma (Berkemer, 2008), market entry (Chmura and Güth, 2011), ultimatum bargaining (Miękisz and Ramsza, 2013), use of common-pool resources (Cárdenas et al., 2015), contributions to public goods (Mantilla et al., 2020), the Centipede game (Sandholm et al., 2019; Izquierdo and Izquierdo, 2021), the Prisoner's Dilemma (Arigapudi et al., 2021), coordination problems (Izquierdo et al., 2022), and finitely repeated games (Sethi, 2021).

Sethi (2000) introduced population dynamics based on the considered procedurally rational agents. These dynamics, which have been called *sampling dynamics* (see e.g. Miękisz and Ramsza (2013); Mantilla et al. (2020)) and *payoff-sampling dynamics* (see e.g. Sethi (2021); Arigapudi et al. (2021, 2022)), take the form

$$\dot{x}_i = w_i^\kappa(x) - x_i, \tag{1}$$

corresponding to a setting where agents occasionally and independently revise their current strategy and, when revising, they adopt strategy a_i with probability $w_i^\kappa(x)$. The process assumes a common rate of revision for every agent, so the "outflow" term for a_i-strategists in (1) is proportional to their presence in the population, x_i.

Sandholm et al. (2019) generalized payoff-sampling dynamics, allowing revising agents to consider *subsets* of their available strategies (rather than *all* available strategies) and to use different tie-breaking rules. This generalization led to the so-called family of *Best Experienced Payoff* (BEP) protocols and their associated dynamics.

The procedurally rational players and $S(\kappa)$ equilibria introduced by Osborne and Rubinstein (1998) correspond to $\text{BEP}_{all}(\kappa, \beta^{unif})$ protocols, where the subscript *all* indicates that all strategies are tested, κ is the number of trials that each strategy is tested, and β^{unif} indicates that ties are resolved uniformly at random. We refer to these $\text{BEP}_{all}(\kappa, \beta^{unif})$ protocols as *uniform payoff-sampling* protocols.

Sethi (2021) considers the family of *regular* tie-breaking rules β^r, which are those that place positive probability on choosing each of the strategies that achieve the best-

[1] Sandholm et al. (2019, 2020), Arigapudi et al. (2021), and Sethi (2021) also consider other tie-breaking rules.

experienced-payoff in a battery of tests. We refer to the corresponding $\text{BEP}_{all}(\kappa, \beta^r)$ protocols as *regular payoff-sampling* protocols.[2]

The protocols considered in this paper are $\text{BEP}_{all}(\kappa, \beta)$, i.e., we assume that revising agents test *all* their available strategies, and we allow for different number of trials κ and different tie-breaking rules β. Henceforth we refer to these $\text{BEP}_{all}(\kappa, \beta)$ protocols simply as *payoff-sampling* protocols. For brevity, we sometimes use BEP_{all} for $\text{BEP}_{all}(\kappa, \beta)$.

The relationship between Nash equilibria and payoff-sampling equilibria is well understood in the literature. Payoff-sampling equilibria are not necessarily Nash equilibria, and vice versa. Nonetheless, Osborne and Rubinstein (1998) showed that every two-player game has an $S(\kappa)$ equilibrium, for any number of trials $\kappa \in \mathbb{N}$, and that the limit of convergent sequences of $S(\kappa)$ equilibria as $\kappa \to \infty$ is a Nash equilibrium. Sandholm et al. (2020) extended these results to p-player games under any BEP protocol. We also know that, under payoff-sampling dynamics, asymptotically stable states are not necessarily Nash.[3] This contrasts with most other evolutionary dynamics.[4]

The connection between *strict* Nash equilibria and payoff-sampling equilibria –and their stability– is also well understood in the literature. Osborne and Rubinstein (1998) showed that strict Nash profiles correspond to $S(\kappa)$ equilibria, and that those are the only monomorphic $S(\kappa)$ equilibria, i.e., the only $S(\kappa)$ equilibria in which all players in each population use the same strategy. Sethi (2000) showed that strict Nash equilibria may be dynamically stable or unstable under payoff-sampling dynamics with $\kappa = 1$; and Sandholm et al. (2020) proved that, for a sufficiently large number of trials κ, strict Nash equilibria are asymptotically stable under any BEP protocol. The stability of strict Nash equilibria under payoff-sampling dynamics is also analyzed by Arigapudi et al. (2021) and Izquierdo and Izquierdo (2022b).

To our knowledge, besides their relationship with Nash and strict Nash equilibria, there are no general results about the structure and stability of payoff-sampling equilibria in the literature. In this paper, we shed some light on this issue. In particular, we provide necessary and sufficient conditions to characterize the support of all payoff-sampling equilibria and the dynamic stability of the faces to which they belong, in any game.

To illustrate this characterization, consider the following two-player symmetric game, with strategy set $\{a, b, c\}$ and payoff matrix (to the row player) E_1:

[2] Note that β^{unif} is a regular tie-breaking rule, so all uniform payoff-sampling protocols are regular.

[3] Sandholm et al. (2019) show that in the centipede game, for low number of trials κ, there is an interior asymptotically stable state at which most players cooperate up until the last five stages of the game.

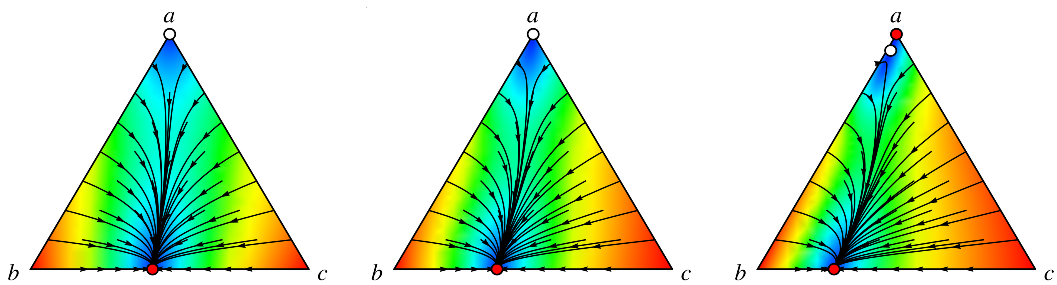
[4] In weakly payoff positive selection dynamics (i.e. dynamics where at least one of the pure strategies that obtains an expected payoff above average –assuming at least one such a strategy exists– has a positive growth rate), only Nash states can be Lyapunov stable (Weibull, 1995, p. 151). In the standard multi-population replicator dynamics, and in many other evolutionary dynamics, only strict Nash states can be asymptotically stable (Eshel and Akin, 1983; Ritzberger and Vogelsberger, 1990; Ritzberger and Weibull, 1995; Balkenborg and Schlag, 2007; Samuelson and Zhang, 1992).

$E_1 =$

	a	b	c
a	1	0	0
b	0	1	8
c	0	2	1

Strategy a is a strict Nash strategy (i.e., (a, a) is a strict Nash profile) and, consequently, for any κ, the population state[5] $(x_a, x_b, x_c) = (1, 0, 0)$ is an $S(\kappa)$ equilibrium. Given that it is strict Nash, this equilibrium is asymptotically stable for sufficiently large κ (Sandholm et al., 2020). Profiles (b, b) and (c, c) are not strict Nash, so the monomorphic population states $(0, 1, 0)$ and $(0, 0, 1)$ are not $S(\kappa)$ equilibria. An analysis based on strict Nash equilibria would finish here.

Nonetheless, note that the subset of strategies $H = \{b, c\}$ has an interesting property: if your co-player is using any strategy in H, you are better off choosing a strategy in H, rather than choosing any strategy not in H. Specifically, when testing each strategy (κ times) against co-players using strategies in H, strategy a will obtain a total payoff of 0, while strategies b and c will each obtain a total payoff of at least κ. Consequently, at any state whose support is contained in H, the strategy selected by a revising agent (i.e., the strategy that performs best in a battery of tests) will necessarily be in H. As a consequence, we will prove later that, for any $\kappa \in \mathbb{N}$, there is an $S(\kappa)$ equilibrium whose support is H (see figure 1). Furthermore, the face spanned by H, i.e. $\Delta_H = \{(0, x_b, x_c) \in \mathbb{R}^3_+ : x_b + x_c = 1\}$, is asymptotically stable under any payoff-sampling dynamic BEP_{all}.[6] As $\kappa \to \infty$, any convergent series of equilibria in Δ_H converges to the unique Nash equilibrium in Δ_H, which is $(0, \frac{7}{8}, \frac{1}{8})$.



(i) $\kappa = 1$ (ii) $\kappa = 2$ (iii) $\kappa = 4$

Figure 1: Uniform payoff sampling dynamics $\text{BEP}_{all}(\kappa, \beta^{unif})$ for different values of κ, for the game with payoff matrix E_1. In the figures, colors represent speed of motion: red is fastest, blue is slowest. Isolated rest points are represented with circles: red if the rest point is asymptotically stable, and white if it is unstable. Connected components of rest points are represented with lines: red if asymptotically stable, and white if unstable.

[5] In the examples, for clarity, we will use x_a to indicate the fraction of players in the population using strategy a (equivalently for the other strategies), instead of numbering them and using the notation $x_1, x_2, \ldots$.

[6] For $\kappa > 1$, this can be shown using the bound in the proof of proposition 3.6.

With the same motivation as Ritzberger and Weibull (1995), or Balkenborg et al. (2013), but considering different dynamics, in this paper we focus on the asymptotic stability of faces of the space of population states: subsets of states where some strategies are not used. Faces are associated with subsets of pure strategies; one subset for each player position. At one extreme of this spectrum we have monomorphic states (where all players in each population use the same strategy); if a monomorphic state is a rest point under a regular payoff-sampling dynamic, it must correspond to a strict Nash equilibrium. The opposite extreme is the whole set of population states, i.e. the maximal face, which includes all the strategies in the game. Informally, a face is asymptotically stable if it attracts all trajectories starting from (sufficiently close) nearby states. Ritzberger and Weibull (1995) consider regular[7] selection dynamics –such as the replicator dynamics and other sign-preserving selection dynamics–, and Balkenborg et al. (2013) analyze generalized best reply dynamics –which assume highly rational and highly informed players. Here we consider procedurally rational players and their associated payoff-sampling dynamics BEP_{all}.

Our main result (proposition 3.1) is a necessary and sufficient condition for a face to be invariant under every payoff-sampling protocol, leading to the definition of sets of strategies Closed Under Payoff Sampling (CUPS). If the number of trials is above a certain value, our condition is also necessary and sufficient for a face to be asymptotically stable under every payoff-sampling dynamic. Importantly, the characterization of CUPS sets (proposition 3.1) is made in terms of simple comparisons between some of the payoffs of the game.

We also prove that a) every CUPS face contains at least one payoff-sampling equilibrium, b) the support of every regular payoff-sampling equilibrium is a CUPS set, and c) every minimal CUPS set H contains at least one regular payoff-sampling equilibrium with support H, and no regular payoff-sampling equilibrium with support properly contained in H.

All these results combined can provide useful insights on the dynamics of payoff-sampling protocols in many games, using only a few comparisons between some of the payoffs of the game. This is illustrated for the Centipede game and for the Traveler's Dilemma in section 5.

As for the relation of CUPS sets to other setwise solution concepts, we show that CUPS sets are Closed Under Rational Behavior –CURB (Basu and Weibull, 1991)– and, consequently, every CUPS face contains an essential component of Nash equilibria (Ritzberger and Weibull, 1995). However, not all CURB sets are CUPS: CUPS sets are a refinement of CURB sets. Finally, there is no direct connection between CUPS sets and asymptotic stability under the (multi-population) replicator dynamics: CUPS faces may not be closed under the better-reply correspondence (Ritzberger and Weibull, 1995) –so they may not be asymptotically stable under the replicator dynamics– and vice versa.

[7] The term "regular", frequently used in evolutionary game theory to qualify certain dynamics (see e.g. Ritzberger and Weibull (1995)), has a different meaning in this sentence from the one used by Sethi (2021) to qualify certain tie-breakers. BEP_{all} dynamics are not regular dynamics in the first sense, since a strategy that is absent from the population can be introduced, if it performs well in a battery of tests.

The rest of the paper is structured as follows. Section 2 contains the notation and main definitions for symmetric p-player population games played in one population, and for payoff-sampling dynamics BEP$_{all}$, as well as some background on dynamical systems. Section 3 defines strategy sets Closed Under Payoff Sampling (CUPS), provides a simple way to characterize them from the game payoffs, and presents several results relating CUPS sets to equilibria and to asymptotic stability under payoff-sampling dynamics. In section 4 we show that CUPS sets are a refinement of CURB sets, and we also show that CUPS sets may or may not be closed under weakly better replies. In section 5 we include some applications, and section 6 presents several conclusions. The proofs, and the extension of our results to (symmetric or asymmetric) p-player games played in p populations, are presented in an appendix.

All figures in this paper can be easily replicated with open-source freely available software which also performs exact computations of rest points and exact linearization analyses (*BEP-3s-sp* (Izquierdo and Izquierdo, 2022a) for figures 2 and 3, and *EvoDyn-3s* (Izquierdo et al., 2018) for the other figures).

2 Payoff-sampling protocols and dynamics

2.1 Population games

For notational simplicity, we focus first on p-player symmetric games played in one population. The extension to multi-population (symmetric or asymmetric) games is presented in appendix A. Following Sandholm et al. (2020), we consider a unit-mass population of agents who are matched to play a symmetric p-player normal form game $G = \{A, U\}$. This game is defined by a strategy set $A = \{a_i\}_{i=1}^n$ containing n pure strategies, and a payoff function $U\colon A^p \to \mathbb{R}$, where $U(a_i; a_{j_1}, \ldots, a_{j_{p-1}})$ represents the payoff obtained by a strategy a_i player whose opponents play strategies $a_{j_1}, \ldots, a_{j_{p-1}}$. Our symmetry assumption requires that the value of U not depend on the ordering of the last $p-1$ arguments. For a tuple of $(p-1)$ strategies $\bar{a} \equiv (a_{j_1}, \ldots, a_{j_{p-1}})$, we write $U_{a_i,\bar{a}} \equiv U(a_i; a_{j_1}, \ldots, a_{j_{p-1}})$.

Aggregate behavior in the population is described by a *population state* $x = (x_i) \in \Delta_A \equiv \{x \in \mathbb{R}^n_+ \colon \sum_{i=1}^n x_i = 1\}$, with component x_i representing the fraction of agents in the population using strategy $a_i \in A$. The standard basis vector $e_i \in \Delta_A$ represents the pure (monomorphic) state at which all agents play strategy a_i.

The (expected) payoff function to strategy a_i at state x is the usual extension of the payoff function U to the simplex Δ_A, i.e., $U_i(x) = \sum_{\bar{a}\in A^{(p-1)}} \left(\prod_{j=1}^n (x_j)^{\mathbb{I}_j(\bar{a})}\right) U_{a_i,\bar{a}}$, where $\bar{a}$ is a $(p-1)$-tuple of strategies (one for each co-player), and the exponent $\mathbb{I}_j(\bar{a})$ is the number of occurrences of strategy a_j in $\bar{a}$.

2.2 Payoff-sampling dynamics BEP$_{all}$

Under a payoff-sampling protocol BEP$_{all}(\kappa, \beta)$, agents occasionally revise their current strategy by conducting a battery of tests involving all their strategies. The first parameter in a payoff-sampling protocol $\text{BEP}_{all}(\kappa,\beta)$, called the *number of trials* $\kappa \in \mathbb{N}$, specifies the number of times that each strategy will be played in the battery of tests. Thus, each strategy will be played by the revising agent over κ matches, with each match requiring a new independent sampling of $(p-1)$ co-players. The second parameter, namely the *tie-breaking rule* β, indicates the rule used to decide which strategy is selected when the best result (i.e. the greatest total –or, equivalently, average– experienced payoff) is obtained by more than one strategy.[8]

Under a payoff-sampling protocol $\text{BEP}_{all}(\kappa,\beta)$, let $w_i^{\kappa,\beta}(x)$ be the probability with which strategy a_i is selected by a revising agent at state x, i.e., the probability that strategy a_i obtains the best total payoff in κ trials, and, if there are ties, it is selected by the tie-breaking rule β. This probability is a continuous function of the population state x.

The calculation of the term $w_i^{\kappa,\beta}(x)$ for payoff-sampling processes $\text{BEP}_{all}(\kappa,\beta)$ is presented next, adapted from Sandholm et al. (2020). Let a *battery of tests* conducted by a revising agent be the process of testing κ times each of her n strategies, for which a total of $n\,\kappa\,(p-1)$ co-players need to be sampled. To represent the strategies used by the sampled co-players in a battery of tests, consider the indexes $i \in \{1,...,n\}, k \in \{1,...,\kappa\}$ and $o \in \{1,...,p-1\}$; let $a_{i,k,o}^{bat} \in A$ be the strategy of the o-th co-player sampled when conducting the k-th trial of strategy a_i; and let $a^{bat} \equiv (a_{i,k,o}^{bat})$ be the correspondingly indexed sequence of $n\,\kappa\,(p-1)$ strategies. Let $\Phi_{A,\kappa,p}$ be the set of all such indexed sequences of $n\,\kappa\,(p-1)$ strategies taken from A. For a battery of tests with sampled strategies a^{bat}, let $\pi^U(a^{bat})$ be the n-tuple of total experienced payoffs $(\pi_i^U(a^{bat}))$ obtained by each strategy $a_i \in A$, i.e.,

$$\pi_i^U(a^{bat}) = \sum_{k=1}^{\kappa} U(a_i; \bar{a}_{i,k}^{bat})$$

where $\bar{a}_{i,k}^{bat} \equiv (a_{i,k,1}^{bat}, a_{i,k,2}^{bat}, ..., a_{i,k,p-1}^{bat})$ is the $(p-1)$-tuple of strategies used by the sampled co-players of a revising agent when she conducts her k-th trial of strategy a_i.

Note that the probability of obtaining the sequence of strategies a^{bat}, in a battery of tests conducted at state x, is $\prod_{l=1}^{n} x_l^{\mathbb{I}_l(a^{bat})}$, where the exponent $\mathbb{I}_l(a^{bat})$ is the number of occurrences of strategy a_l in the sequence a^{bat}. Considering this, under a $\text{BEP}_{all}(\kappa,\beta)$ protocol, the probability that a revising agent chooses strategy a_i at population state x is given by

$$w_i^{\kappa,\beta}(x) = \sum_{j=1}^{n} x_j \sum_{a^{bat} \in \Phi_{A,\kappa,p}} \beta_{ji}\left(\pi^U(a^{bat})\right) \prod_{l=1}^{n} x_l^{\mathbb{I}_l(a^{bat})} \tag{2}$$

where the functions $\beta_{ji} : \mathbb{R}^n \to [0,1]$ define the tie-breaking rule. Considering an n-tuple of total payoffs $\pi \equiv (\pi_i)$, the functions $\beta_{ji}(\pi)$ are such that:

[8] We could allow each strategy to be tested a possibly different number of trials, and the selection be based on the greatest average payoff. Our results are qualitatively robust to this variation, considering that each strategy is tested at least κ times.

- $\beta_{ji}(\pi) = 1$ if $\pi_i > \pi_l$ for all $l \neq i$. I.e., $\beta_{ji}(\pi) = 1$ if strategy a_i is the only one to obtain the maximum total payoff.

- $\beta_{ji}(\pi) = 0$ if $\pi_i < \max_{l \in \{1,\ldots,n\}} \pi_l$. I.e., $\beta_{ji}(\pi) = 0$ if strategy a_i does not obtain the maximum total payoff.

- Otherwise, i.e., if strategy a_i obtains the maximum total payoff but it is not the only one to do so, the rule $\beta_{ji}(\pi)$ establishes the probability with which strategy a_i is chosen, depending on the total payoffs obtained by each strategy and on which strategy (a_j) is being currently used by the revising agent. We assume that one of the strategies with the highest payoff is chosen, so we have $\sum_{i:\pi_i = max_l(\pi_l)} \beta_{ji}(\pi) = 1$.

Regular tie-breaking rules β^r are such that $\beta^r_{ji}(\pi) > 0$ whenever strategy a_i obtains the maximum total payoff, i.e., whenever $\pi_i = \max_l \pi_l$.

Well-known results of Benaïm and Weibull (2003) show that the behavior of a large but finite population following the procedure presented above is closely approximated by the solution of the associated *mean dynamic*, a differential equation which describes the expected motion of the population from each state. This mean dynamic for $\text{BEP}_{all}(\kappa, \beta)$ processes is:

$$\dot{x}_i = w_i^{\kappa,\beta}(x) - x_i \tag{3}$$

An equilibrium $S(\kappa, \beta)$ under a payoff-sampling protocol is a population state x satisfying

$$w^{\kappa,\beta}(x) = x$$

where $w^{\kappa,\beta}(x)$ is the vector whose components are $w_i^{\kappa,\beta}(x)$.

The $S(\kappa)$ equilibria of Osborne and Rubinstein (1998) are the $S(\kappa, \beta^{unif})$ equilibria, which correspond to the uniform payoff-sampling protocol $\text{BEP}_{all}(\kappa, \beta^{unif})$.

2.3 Background on dynamical systems: invariant and asymptotically stable faces

Consider a C^1 differential equation $\dot{x} = V(x)$ defined on Δ_A whose forward solutions do not leave Δ_A. A set $Y \subseteq \Delta_A$ is called *forward invariant* if any solution path starting in Y remains in Y for the entire future: $x(t, x^0) \in Y$ for all $x^0 \in Y$ and $t \in \mathbb{R}_+$. It is called invariant if, moreover, any solution path that at some time is in Y has also been in Y for the entire past. A point $x^* \in \Delta_A$ is called a stationary point or a rest point if $\{x^*\}$ is an invariant set, satisfying $V(x^*) = 0$.

For any nonempty subset of strategies $H \subseteq A$, let Δ_H be the face (or subsimplex) of Δ_A spanned by the strategies in H, i.e., $\Delta_H = \{x \in \Delta_A : x_i = 0 \text{ if } a_i \notin H\}$. BEP dynamics are C^1 and satisfy $\dot{x}_i \geq -x_i$, which implies that if some strategy is initially present in the population, it will remain present forever (it can only vanish asymptotically). Consequently, if $x(t, x^0)$ is in some face Δ_H at some time t, the path $x(t, x^0)$ has been in that face Δ_H for the entire past, and if Δ_H is forward invariant, then it is invariant.

A closed invariant set $Y \subseteq \Delta_A$ is (Lyapunov) stable if for every neighborhood O of Y there exists a neighborhood O' of Y such that $x(t, x^0) \in O$ for all $x^0 \in O' \cap \Delta_A$ and all $t \geq 0$. If a set Y is not Lyapunov stable it is unstable, and it is repelling if there is a neighborhood O of Y such that solutions from all initial conditions in $(O \setminus Y) \cap \Delta_A$ leave O.

A closed invariant set $Y \subseteq \Delta_A$ is asymptotically stable if it is stable and there is some neighborhood O of Y such that $x(t, x^0) \to Y$ as $t \to \infty$ for all $x^0 \in O \cap \Delta_A$.

3 Sets closed under payoff sampling: characterization and properties

As a preparation for the definition of CUPS sets, we first define sets closed under a specific payoff-sampling protocol $\mathrm{BEP}_{all}(\kappa, \beta)$. Informally, a set of strategies H is closed under a payoff-sampling protocol if any revising agent using that protocol will necessarily choose some strategy in H when players use strategies in H. Note that the set of strategies that a revising agent using a $\mathrm{BEP}_{all}(\kappa, \beta)$ protocol may select with positive probability at state x is the set of strategies a_i such that $w_i^{\kappa,\beta}(x) > 0$.

Definition 1. *Closed under a payoff-sampling protocol. A nonempty subset of strategies $H \subseteq A$ is closed under a $BEP_{all}(\kappa, \beta)$ protocol if for all $x \in \Delta_H$, $w^{\kappa,\beta}(x) \in \Delta_H$.*

If H is closed under a $\mathrm{BEP}_{all}(\kappa, \beta)$ protocol, any player using such a revision protocol at any state $x \in \Delta_H$ will choose some strategy in H. This makes the face Δ_H invariant under the corresponding payoff-sampling dynamics. It can be easily seen that being closed under a protocol is also a necessary condition for Δ_H to be invariant under the protocol, so it is a sufficient and necessary condition for invariance of Δ_H.

Definition 2. *Closed under payoff sampling. A nonempty subset of strategies $H \subseteq A$ is closed under payoff sampling (CUPS) if H is closed under every payoff-sampling protocol $BEP_{all}(\kappa, \beta)$. If H is a CUPS set, we say that Δ_H is a CUPS face.*

A CUPS face is invariant under $\mathrm{BEP}_{all}(\kappa, \beta)$ dynamics for every $\kappa \in \mathbb{N}$ and for every tie-breaking rule β. The fact that H is a CUPS set implies that, under payoff sampling, if a revising agent using some strategy in H meets co-players who use strategies in H, then the selected strategy is necessarily in H, regardless of the number of trials κ and of the tie-breaking rule. However, we will show later that a sufficient (and necessary) condition to be a CUPS set is to be closed under (any) one regular payoff-sampling protocol $\mathrm{BEP}_{all}(\kappa, \beta^r)$, i.e., being closed for some κ and some β^r implies being closed for every κ and every β.

Clearly, the whole strategy set A is CUPS. Our next proposition shows that the other CUPS sets can be easily characterized from the game payoffs.

Proposition 3.1. *A nonempty subset of strategies $H \subset A$ is closed under payoff sampling (CUPS) if and only if:*

$$\max_{a_i \in (A \setminus H),\, \bar{a} \in H^{(p-1)}} U_{a_i, \bar{a}} < \max_{a_j \in H} \min_{\bar{a} \in H^{(p-1)}} U_{a_j, \bar{a}} \tag{4}$$

The term $(\max_{a_j \in H} \min_{\bar{a} \in H^{(p-1)}} U_{a_j,\bar{a}})$ in proposition 3.1 is the maxmin payoff in H: the maximum payoff that can be guaranteed to be obtained or exceeded by some strategy in H when meeting co-players using strategies in H. For a strategy a_i that is not in H, the term $(\max_{\bar{a} \in H^{(p-1)}} U_{a_i,\bar{a}})$ is the maximum payoff that an a_i-player may obtain when meeting co-players using strategies in H. It can be easily seen that the CUPS sets of size 1, when they exist, are precisely the strict Nash strategies of the game (i.e., strategies a_j such that the strategy profile $(a_j, a_j, ..., a_j)$ is a strict Nash equilibrium of the game). To see this, note that the condition for a set with a single strategy $\{a_j\}$ to be a CUPS set is $U(a_i; a_j, a_j, ..., a_j) < U(a_j; a_j, a_j, ..., a_j)$ for all $a_i \neq a_j$, which is the definition of a strict Nash strategy. From this point of view, a CUPS set is a setwise generalization of the strict Nash property.

Proposition 3.2, which can be seen as a consequence of proposition 3.1, shows that if a set is closed under some regular $\text{BEP}_{all}(\kappa, \beta^r)$ protocol, then it is CUPS.

Proposition 3.2. *A subset of strategies is closed under a regular payoff-sampling protocol $BEP_{all}(\kappa, \beta^r)$ if and only if it is CUPS.*

Propositions 3.1 and 3.2 together show that being CUPS (i.e. satisfying (4)) is a necessary and sufficient condition for a face to be invariant under any given regular $\text{BEP}_{all}(\kappa, \beta^r)$ dynamics, and a sufficient condition for invariance of a face under any given $\text{BEP}_{all}(\kappa, \beta)$ dynamics.

Considering that $w^{\kappa,\beta}$ is a continuous function that maps CUPS faces onto themselves, we could expect the following existence result for $S(\kappa, \beta)$ equilibria in every CUPS face.

Proposition 3.3. *Every CUPS face contains at least one $S(\kappa, \beta)$ equilibrium, for every number of trials κ and any tie-breaking rule β.*

In turn, for regular tie-breaking rules, the support of an $S(\kappa, \beta^r)$ equilibrium must be a CUPS set. This is equivalent to stating that regular payoff-sampling equilibria must belong to the relative interior of some CUPS face.

Proposition 3.4. *The support of every regular $S(\kappa, \beta^r)$ equilibrium is a CUPS set.*

Proposition 3.4 shows that the support of any regular $S(\kappa, \beta^r)$ equilibrium is a CUPS set, and Proposition 3.3 shows that, for any κ and β, a CUPS set contains the support of at least one $S(\kappa, \beta)$ equilibrium. In other words, H being a CUPS set is a necessary condition to have a regular $S(\kappa, \beta^r)$ equilibrium whose support is H; and it is a sufficient condition to have an $S(\kappa, \beta)$ equilibrium whose support is contained in H.

Definition 3. *Minimal CUPS set. A minimal CUPS set is a CUPS set that does not contain any proper CUPS subset.*

Given that the whole strategy set A is a CUPS set, the existence of at least one minimal CUPS set is guaranteed. If H is a minimal CUPS set, we say that Δ_H is a *minimal CUPS face.*

Proposition 3.5. *Consider any number of trials $\kappa \in \mathbb{N}$ and any regular tie-breaking rule β^r. A set H is a minimal CUPS set if and only if there is a regular $S(\kappa, \beta^r)$ equilibrium with support H and there is no regular $S(\kappa, \beta^r)$ equilibrium with support properly contained in H.*

Proposition 3.5 implies that, if H is a minimal CUPS set, then, for any $\kappa \in \mathbb{N}$ and any regular tie-breaking rule β^r, there is some regular $S(\kappa, \beta^r)$ equilibrium in the (relative) interior of Δ_H, and there are no $S(\kappa, \beta^r)$ equilibria in the boundary of Δ_H. Besides, minimal CUPS faces are the smallest faces that can be asymptotically stable under regular BEP_{all} dynamics (recall that being CUPS is a necessary and sufficient condition for a face to be invariant under regular BEP_{all} dynamics.)

Our last results in this section are about stability of CUPS faces. First, proposition 3.6 shows that, for a sufficiently large number of trials, CUPS faces are asymptotically stable under BEP_{all} dynamics.

Proposition 3.6. *If H is CUPS, then there is a finite k_0 such that, for $\kappa > k_0$, face Δ_H is asymptotically stable under every $BEP_{all}(\kappa, \beta)$ dynamics.*

So, under any regular BEP_{all} dynamics, being CUPS is a necessary condition for asymptotic stability of a face (since it is a necessary condition for invariance). And, for a sufficiently large number of trials, being CUPS is also a sufficient condition for the face to be asymptotically stable, in this case under every payoff-sampling dynamics BEP_{all} (regular or not). The proof of proposition 3.6 provides a finite value κ_0 (not necessarily the smallest one) that guarantees asymptotic stability for any $\kappa > \kappa_0$, regardless of the tie-breaking rule.

Last, considering $\text{BEP}_{all}(1, \beta)$ dynamics, we extend (to CUPS faces) a sufficient condition by Sethi (2000) for a strict Nash equilibrium to be a repellor, as well as more general conditions (by Sandholm et al. (2020)) for such equilibria to be unstable.

Considering a CUPS set H and its complement $H^c \equiv A \setminus H$, we say that strategy $a_j \in H^c$ *supports* invasion (of H) by strategy $a_k \in H^c$ if

$$\min_{\bar{a} \in H^{(p-2)}} U(a_k; a_j, \bar{a}) > \max_{a_i \in H, \bar{a} \in H^{(p-1)}} U_{a_i, \bar{a}}$$

Or, for the two-player case, if

$$U_{a_k, a_j} > \max_{a_i, a_l \in H} U_{a_i, a_l}$$

I.e., a_j supports invasion (of H) by a_k if the presence of one a_k-strategist and one a_j-strategist in a strategy profile where the other players use strategies in H ensures that strategy a_k obtains a greater payoff than the best payoff that the strategies in H can obtain when playing among themselves. In the dynamics (3), the presence of strategy a_j in proportion x_j at states close to Δ_H guarantees a minimum inflow for $\dot{x}_k$ that is proportional to x_j and to $(p-1)$. Note that strategy $a_j \in H^c$ may support invasion by a_j itself.

Definition 4. *(Partially) inferior CUPS set. A CUPS set H is partially inferior if there is a non-empty subset of strategies $H_0^c \subseteq (A \setminus H)$ such that every strategy $a_j \in H_0^c$ supports invasion by at least one strategy $a_k \in H_0^c$. If every strategy $a_j \in H_0^c$ supports invasion by at least two strategies $a_k \neq a_{k'} \in H_0^c$, we say that H is partially twice inferior. If the condition holds for $H_0^c = (A \setminus H)$, we call the set inferior (or twice inferior, respectively).*

Proposition 3.7. *In symmetric games with more than two players:*

- *If a CUPS set H is inferior, then face Δ_H is repelling under $BEP_{all}(1, \beta)$ dynamics.*
- *If a CUPS set H is partially inferior, then face Δ_H is unstable under $BEP_{all}(1, \beta)$ dynamics.*

In symmetric two-player games, the previous result applies to twice inferior (respectively, partially twice inferior) CUPS sets.

To illustrate the usefulness of the propositions derived in this section, consider the following p-player symmetric game, with strategy set $A = \{a, b, c\}$ and payoff matrix:

$E_2 =$

	number of other players using strategy c: 0	1	≥ 2
a	1	1	0
b	1	1	1
c	0	2	2

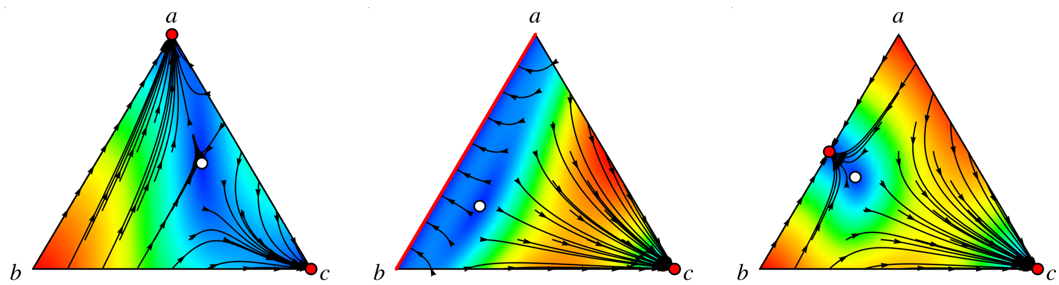
In this example, the sets of strategies $H_1 = \{a, b\}$ and $H_2 = \{c\}$ are both CUPS, since

$$\max_{a_i \in (A \setminus H_1),\, \bar{a} \in H_1^{(p-1)}} U_{a_i, \bar{a}} = 0 < 1 = \max_{a_j \in H_1} \min_{\bar{a} \in H_1^{(p-1)}} U_{a_j, \bar{a}}$$

and $(c, ..., c)$ is a strict Nash profile.

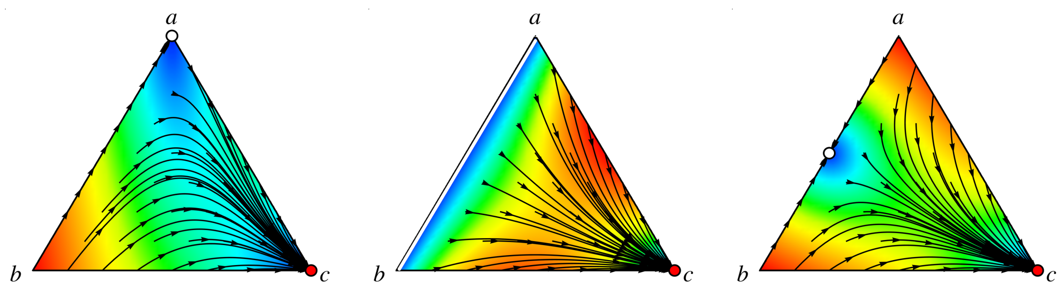
Thus, proposition 3.1 tells us that face $\Delta_{\{a,b\}}$ is invariant under any BEP$_{all}(\kappa, \beta)$ protocol and that e_c is an $S(\kappa, \beta)$ equilibrium for any $\kappa \in \mathbb{N}$ and any tie-breaking rule β. Proposition 3.3 implies that, besides e_c, there is at least another $S(\kappa, \beta)$ equilibrium in the face $\Delta_{\{a,b\}}$. See figures 2 and 3.

It is easy to see that CUPS sets $H_1 = \{a, b\}$ and $H_2 = \{c\}$ are both minimal, and that there is no other minimal CUPS set. Naturally, the whole strategy set A is CUPS. Thus, applying propositions 3.4 and 3.5, we can assert that for any $\kappa \in \mathbb{N}$ and any regular tie-breaking rule β^r, every regular $S(\kappa, \beta^r)$ equilibrium besides e_c either lies in the relative interior of face $\Delta_{\{a,b\}}$ (and there is at least one such equilibrium, since $H_1 = \{a, b\}$ is minimal –see proposition 3.5) or has full support (and such an equilibrium may or may not exist, since CUPS set A is not minimal). See figures 2(iii) and 3(iii).



(i) $\mathrm{BEP}_{all}\left(\kappa = 2, \beta^{min}\right)$ (ii) $\mathrm{BEP}_{all}\left(\kappa = 2, \beta^{stick-unif}\right)$ (iii) $\mathrm{BEP}_{all}\left(\kappa = 4, \beta^{unif}\right)$

Figure 2: Different payoff-sampling dynamics for the game with payoff matrix E_2 and $p = 3$ players. The tie-breaker β^{min} in (i) chooses in alphabetical order in case of tie. The tie-breaker $\beta^{stick-unif}$ in (ii) selects the strategy that the revising agent is currently using if it is among the best in the test; otherwise, a uniformly random choice is conducted among the best strategies in the test. Finally, the tie-breaker β^{unif} in (iii) makes a uniformly random choice among the best strategies in the test. The red line in (ii) denotes an asymptotically stable connected component of rest points.



(i) $\mathrm{BEP}_{all}\left(\kappa = 1, \beta^{min}\right)$ (ii) $\mathrm{BEP}_{all}\left(\kappa = 1, \beta^{stick-unif}\right)$ (iii) $\mathrm{BEP}_{all}\left(\kappa = 1, \beta^{unif}\right)$

Figure 3: Different payoff-sampling dynamics for the game with payoff matrix E_2 and $p = 3$ players. For the definition of the different tie-breakers, see figure 2. The white line in (ii) denotes an unstable connected component of rest points.

In terms of asymptotic stability, proposition 3.6 implies that face $\Delta_{\{a,b\}}$ and strict Nash state e_c are both asymptotically stable under every $\text{BEP}_{all}(\kappa, \beta)$ dynamics for a sufficiently large number of trials κ (see figure 2).

Finally, as for instability, note that CUPS set $H_1 = \{a, b\}$ is inferior, since strategy $c \in H_1^c$ supports invasion (of H_1) by itself:[9]

$$\min_{\bar{a} \in H_1^{(p-2)}} U(c; c, \bar{a}) = 2 > 1 = \max_{a_i \in H_1, \bar{a} \in H_1^{(p-1)}} U_{a_i, \bar{a}}.$$

Thus, proposition 3.7 implies that in game E_2 with more than 2 players, face $\Delta_{\{a,b\}}$ is repelling under all $\text{BEP}_{all}(1, \beta)$ dynamics (see figure 3).

4 CUPS and CURB sets: payoff sampling versus rational behavior

Following Basu and Weibull (1991), we call a nonempty set of strategies $H \subseteq A$ Closed Under Rational Behavior (CURB) if it contains all its best replies, i.e., if $\text{BR}(x) \subseteq H$ for every $x \in \Delta_H$, where BR is the pure best-reply correspondence which maps population states to their pure best-reply strategies (Ritzberger and Weibull, 1995).

Proposition 4.1. *CUPS sets are closed under rational behavior.*

The reason why CUPS sets are CURB is that strategies that do not belong to a CUPS set cannot be best reply to any strategy profile (for co-players) made up by strategies in the CUPS set (or to any state in the CUPS face), as they obtain a lower payoff than the maxmin payoff in the CUPS set (recall proposition 3.1).

Since CUPS sets are CURB, CUPS faces contain an essential connected component of Nash equilibria, which satisfies strong setwise refinement criteria (Ritzberger and Weibull, 1995).[10] While all CUPS sets are CURB, not all CURB sets are CUPS, as the following example illustrates. Consider a two-player symmetric game with strategy set $A = \{a, b, c\}$ and payoff matrix E_3:

$E_3 =$

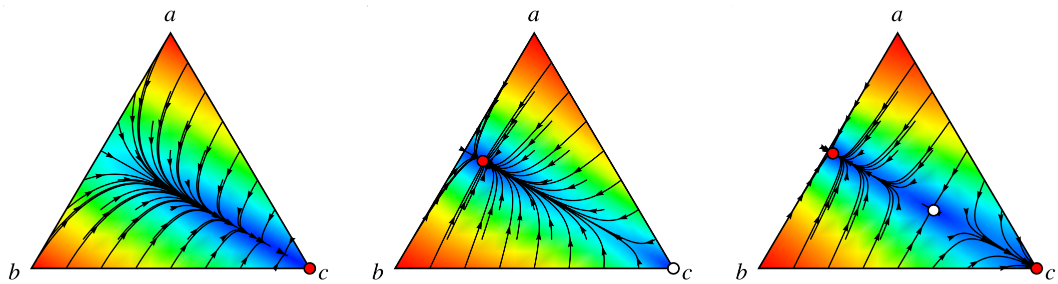
	a	b	c
a	1	5	0
b	5	1	0
c	2	2	2

In this example, the set of strategies $H_1 = \{a, b\}$ is not CUPS, given that the maxmin payoff in H_1 is $1 < \max(U_{c,a}, U_{c,b}) = 2$. Sets $\{a\}$ and $\{b\}$ are not CUPS either –as the strategies are not strict Nash–, so there is no regular $S(\kappa, \beta^r)$ equilibrium in Δ_{H_1}, for

[9] In the two-player case, consider $U(c, c)$ instead of $\min_{\bar{a} \in H_1^{(p-2)}} U(c; c, \bar{a})$.

[10] Ritzberger and Weibull (1995) show that if a set of strategies is closed under some behavior correspondence –a family that includes the best-response correspondence–, then the associated face contains an essential connected component of Nash equilibria (van Damme, 1991), which is consequently hyperstable and strategically stable (Kohlberg and Mertens, 1986).

any κ (see proposition 3.4). However, H_1 is a minimal CURB set, Δ_{H_1} contains the Nash equilibrium $(x_a, x_b, x_c) = (\frac{1}{2}, \frac{1}{2}, 0)$, and a sequence of $S(\kappa, \beta^r)$ equilibria can converge to this Nash equilibrium state from the interior of the simplex (see figure 4). In contrast, $H_2 = \{c\}$ is a CUPS set (profile (c, c) is strict Nash), so $e_c = (0, 0, 1)$ is an $S(\kappa, \beta)$ equilibrium for every κ and β.



(i) $\kappa = 1$ (ii) $\kappa = 2$ (iii) $\kappa = 3$

Figure 4: Uniform payoff sampling dynamics $\text{BEP}_{all}(\kappa, \beta^{unif})$ for different values of κ, for the game with payoff matrix E_3. $H_1 = \{a, b\}$ is minimal CURB but it is not CUPS. Face Δ_{H_1} contains no $S(\kappa)$ equilibrium. Profile (c, c) is strict Nash, so $H_2 = \{c\}$ is minimal CURB and CUPS.

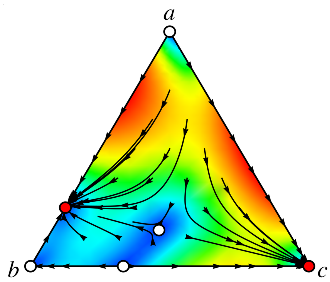
Considering stability under the replicator dynamics, CUPS sets need not be closed under weakly better replies (Ritzberger and Weibull, 1995), which is a sufficient and necessary condition for the corresponding face to be asymptotically stable under the multi-population replicator dynamics, and, more generally, under any sign-preserving dynamics. Consider for instance a symmetric two-player game with strategy set $A = \{a, b, c\}$ and payoff matrix E_4:

$E_4 =$

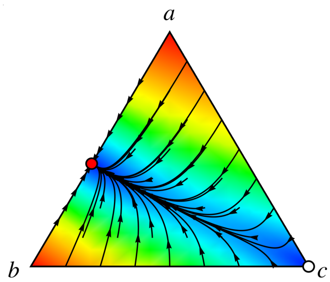
	a	b	c
a	1	4	0
b	4	3	0
c	2	2	2

In this example, the set of strategies $H = \{a, b\}$ is CUPS, because strategy c obtains a payoff of 2, less than the maxmin payoff in H, which is 3. But H is not closed under weakly better replies, because at state $e_a = (1, 0, 0) \in \Delta_H$, strategy $c \notin H$ is a better reply to e_a than e_a itself. As can be seen in figure 5, face Δ_H is not asymptotically stable in the Replicator Dynamics, given that from any neighborhood of e_a there is a trajectory towards e_c. By contrast, face Δ_H is asymptotically stable under any $\text{BEP}_{all}(\kappa, \beta)$ dynamics at least for every $\kappa > 3$ (this can be proved using the bound provided in the proof of proposition 3.6).

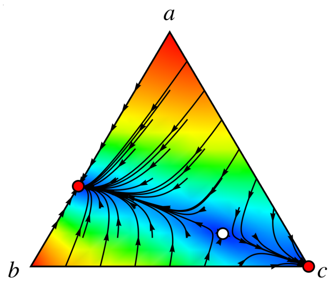
Sets closed under weakly better replies need not be CUPS either, as our next example shows. Consider a symmetric two-player game with strategy set $A = \{a, b, c\}$ and payoff matrix E_5:



(i) Replicator



(ii) Uniform payoff sampling, $\kappa = 1$



(iii) Uniform payoff sampling, $\kappa = 3$

Figure 5: (i) Replicator, (ii) Uniform payoff sampling $\text{BEP}_{all}(\kappa = 1, \beta^{unif})$, and (iii) Uniform payoff sampling $\text{BEP}_{all}(\kappa = 3, \beta^{unif})$ dynamics for the game with payoff matrix E_4. Strategy set $H = \{a, b\}$ is CUPS, but it is not closed under weakly better replies. Face Δ_H is not asymptotically stable in the Replicator Dynamics but, for every κ, it contains some $S(\kappa)$ equilibria, and face Δ_H is asymptotically stable under $\text{BEP}_{all}(\kappa, \beta)$ dynamics at least for every $\kappa > 3$.

$E_5 =$

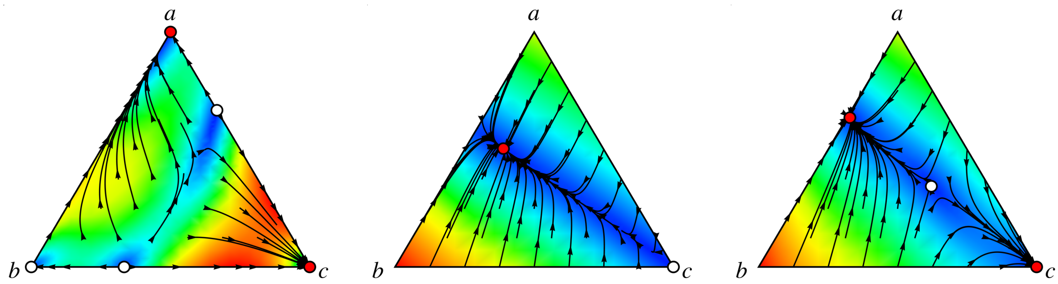
	a	b	c
a	1	3	0
b	1	2	0
c	0	1	2

In this example, the set of strategies $H = \{a, b\}$ is closed under weakly better replies (and also CURB), and Δ_H is asymptotically stable in the Replicator Dynamics (see figure 6), but H is not CUPS, because the maxmin payoff in H is 1, which is also achieved by strategy $c \notin H$ when playing with $b \in H$. Consequently, Δ_H is not invariant under regular payoff sampling, and, given that $\{a\}$ and $\{b\}$ are not CUPS (the strategies are not strict Nash), Δ_H contains no $S(\kappa)$ equilibria, although there can be $S(\kappa)$ equilibria arbitrarily close to Δ_H for large enough κ.

5 Applications

5.1 The Centipede Game

Centipede (Rosenthal, 1981) is a two-player extensive form game with $d \geq 2$ decision nodes, and $d + 1$ final nodes. Consider, for instance, a Centipede game with 8 decision nodes, as shown in Figure 7. Each decision node presents two strategies: stop and continue. The nodes are arranged linearly, with the first one assigned to player 1 and subsequent ones assigned in an alternating fashion. A player who decides to stop ends the game. A player who decides to continue suffers a cost of 1 but benefits his opponent with 3, and sends the game to the next decision node, if one exists. For player $p \in \{1, 2\}$, let strategy $i_p \in A_p$ be the plan to stop at his ith decision node and not before, with an additional strategy for the plan to continue at all his decision nodes. Of course,



(i) Replicator

(ii) Uniform payoff sampling, $\kappa = 1$

(iii) Uniform payoff sampling, $\kappa = 2$

Figure 6: (i) Replicator, (ii) Uniform payoff sampling $\text{BEP}_{all}(\kappa = 1, \beta^{unif})$, and (iii) Uniform payoff sampling $\text{BEP}_{all}(\kappa = 2, \beta^{unif})$ dynamics for the game with payoff matrix E_5. Strategy set $H = \{a, b\}$ is closed under weakly better replies, but not CUPS. Face Δ_H is asymptotically stable in the Replicator Dynamics, but it contains no $S(\kappa)$ equilibria.

the portion of a player's plan that is actually carried out depends on the plan of his opponent.

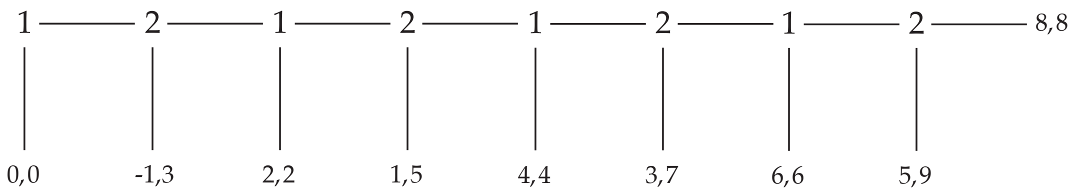


Figure 7: The Centipede game with $d = 8$ decision nodes. Each decision node is labeled with a single number (1 or 2) denoting the deciding player. Each of the $d + 1 = 9$ final nodes is labeled with a pair of payoffs (π_1, π_2), where π_p denotes the payoff obtained by player p.

In a centipede game, the best reply to a player that stops at some node other than the first one, is to stop at the previous node, and all strategies for player 2 are a best reply to strategy 1 for player 1 (i.e. the stop-at-first-node strategy). Given that CUPS sets[11] are CURB, it follows that any CUPS product set $(H_1 \times H_2)$ has to include the stop-at-first-node strategy for player 1 in H_1, and has to include all strategies for the second player in H_2. Considering that the payoff to the stop-at-first-node strategy for the first player is 0 and that all his other strategies may provide a higher payoff (when meeting some strategy in $H_2 = A_2$), it follows that the only CUPS set is the product set of all strategies: $(A_1 \times A_2)$. Consequently, every regular $S(\kappa, \beta^r)$ equilibrium (and, in particular, every $S(\kappa)$ equilibrium) in the Centipede game must have full support. Sandholm et al. (2019) show that, in any $S(1)$ equilibrium, most players in a centipede game continue until their last three decision nodes.

[11] See in appendix A the extension of the definition of CUPS sets to games played in several populations, which is quite straightforward.

5.2 The Traveler's Dilemma

The Traveler's Dilemma (Basu, 1994) is a normal form analogue of the Centipede game in which the unique rationalizable strategy earns the players far less than many other strategy profiles. The payoff matrix for the n-strategy Traveler's Dilemma is E_6.

$E_6 =$

	1	2	3	4	…	n
1	2	4	4	4	$\cdots$	4
2	0	3	5	5	$\cdots$	5
3	0	1	4	6	$\cdots$	6
4	0	1	2	5	$\ddots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\ddots$	$n+2$
n	0	1	2	$\dots$	$n-2$	$n+1$

Strategy 1 is the unique rationalizable strategy, and profile (1,1) is the unique Nash equilibrium.

Let us analyze this game when played in one population. Considering that the best reply to strategy $i > 1$ is strategy $i-1$, and that CUPS sets are CURB, if a CUPS set contains strategy i, then it has to contain strategies $i-1, i-2, ..., 1$. It can also be seen from the payoff matrix and proposition 3.1 that if H is CUPS and $\{1,2,3\} \subseteq H$ then $H = A$. Consequently, the only CUPS sets are $\{1\}, \{1,2\}$ and A. The first strategy is strict Nash, so e_1 is an $S(\kappa, \beta)$ equilibrium, whose stability is analyzed in Sandholm et al. (2020). For $\kappa = 1$, it is easy to check that the only equilibrium in the face spanned by $\{1,2\}$ is e_1, because the dynamics on that face satisfy $\dot{x}_2 = -x_2^2$. Proposition 3.7 also tells us that this equilibrium e_1 is unstable under any $\text{BEP}_{all}(1, \beta)$ for $n > 4$.[12] As the only other CUPS set is A, any other regular $S(\kappa, \beta^r)$ equilibrium must have full support.

6 Conclusions

We have defined strategy sets Closed Under Payoff Sampling (CUPS) and shown that a necessary and sufficient condition to be CUPS is to be closed under some regular payoff-sampling dynamics $\text{BEP}_{all}(\kappa, \beta^r)$. This means that the property of being closed under a regular payoff-sampling protocol $\text{BEP}_{all}(\kappa, \beta^r)$ is independent of the number of trials κ and of the (regular) tie-breaking rule. We have also provided a simple rule to identify CUPS sets from the payoffs of the game .

The identification of CUPS sets in a game yields useful insights on its dynamics under payoff sampling. Being CUPS is a sufficient condition for a face to be invariant under every payoff-sampling dynamics, and it is a necessary and sufficient condition for a face to be invariant under any regular payoff-sampling dynamics (proposition 3.1). Also, for a sufficiently large number of trials, CUPS faces are asymptotically stable under

[12] Let $H = \{1\}$. If $n > 4$, every strategy in $H_0^c = \{n-1, n\} \subseteq (A \setminus H)$ supports invasion by both strategies in H_0^c, so CUPS set $H = \{1\}$ is partially twice inferior.

any payoff-sampling dynamics: even if some (sufficiently few) players adopt a strategy outside the support of the face, the population will tend to move back to the face. For payoff-sampling dynamics with $\kappa = 1$ trial, we have provided sufficient conditions for CUPS faces to be repelling and to be unstable.

CUPS sets are also useful to characterize the support of payoff-sampling equilibria. For a start, every CUPS face contains at least one payoff-sampling equilibrium. We have also proved that the support of every regular payoff-sampling equilibrium is a CUPS set, and that every minimal CUPS set H contains at least one regular payoff-sampling equilibrium with support H, and no regular payoff-sampling equilibrium with support properly contained in H.

Regarding its relation with other setwise solution concepts, CUPS sets are a refinement of strategy sets Closed Under Rational Behavior (CURB). Given that as the number of trials κ goes to infinity, any payoff-sampling dynamics BEP_{all} becomes a version of best response dynamics, one can consider payoff-sampling dynamics BEP_{all} as noisy best response dynamics, with more noise (i.e. greater variance in the information obtained by sampling) for lower values of κ. While all CURB faces are asymptotically stable under best response dynamics (Balkenborg et al., 2013), only those that are also CUPS can be asymptotically stable under regular payoff sampling. For large enough number of trials, CUPS faces are indeed asymptotically stable under any payoff-sampling dynamics. In contrast, CURB faces that are not CUPS cannot contain any regular payoff-sampling equilibrium (for any number of trials); moreover, as illustrated in figure 4(i), such faces may be far away from any payoff-sampling equilibrium for low number of trials.

Acknowledgements

We are very grateful to Tilman Börgers for helpful comments on earlier versions of this manuscript. Financial support from the Spanish State Research Agency (PID2020-118906GB-I00/AEI/10.13039/501100011033), from the Regional Government of Castilla y León and the EU-FEDER program (CLU-2019-04), from the Spanish Ministry of Science, Innovation and Universities, and from the Fulbright Program (PRX19/00113, PRX21/00295), is gratefully acknowledged.

A Appendix: Games played in p populations. Asymmetric games.

In this section we adapt the definitions and results for CUPS sets and faces to the context of p-player games (either symmetric or asymmetric) played in p populations. The propositions in this section are straightforward adaptations of the propositions for the single-population case, and so are their proofs, which we omit.

As argued by Ritzberger and Weibull (1995), many economic applications call for multi-population, rather than single-population dynamics: the player roles may be different and the game may not be symmetric. This leads to the study of evolutionary dynamics (for p-player games) played in p populations, with each player role corresponding to a distinct population (Sandholm, 2010).

Let $G = (\mathcal{P}, A, U)$ be a finite p-player population game, where $\mathcal{P} = \{1, ..., p\}$ is the set of populations or player positions, $A = \times_{\mathfrak{p}\in\mathcal{P}} A^{\mathfrak{p}}$ is the set of pure strategy profiles, with $A^{\mathfrak{p}}$ being the set of $n_{\mathfrak{p}}$ pure strategies for population $\mathfrak{p}$, and $U : A \to \mathbb{R}^p$ is the payoff function, extended to the set of population states $[\Delta]_A = \times_{\mathfrak{p}\in\mathcal{P}} \Delta_{A^{\mathfrak{p}}}$ in the usual way. A (global) population state $x \equiv (x^1, x^2, ..., x^p)$ is a point in the polyhedron $[\Delta]_A$, with $x^{\mathfrak{p}} \in \Delta_{A^{\mathfrak{p}}}$ corresponding to the state in population $\mathfrak{p}$. We represent by $U^{\mathfrak{p}}(a_i^{\mathfrak{p}}; \bar{a})$ the payoff to a player from population $\mathfrak{p}$ using strategy $a_i^{\mathfrak{p}} \in A^{\mathfrak{p}}$ when the other players use the strategies indicated in the (partial) strategy profile $\bar{a} \in \times_{(o\in\mathcal{P}, o\neq\mathfrak{p})} A^o$.

Let a *battery of tests* conducted by a revising agent from population $\mathfrak{p}$ be the process of testing κ times each of her $n_{\mathfrak{p}}$ strategies, for which a total of $n_{\mathfrak{p}}\, \kappa\, (p-1)$ co-players need to be sampled. To represent the strategies used by the sampled co-players in a battery of tests, let $a^{-\mathfrak{p}} \equiv (a_{i,k,o}^{-\mathfrak{p}})$ be an indexed sequence of $n_{\mathfrak{p}}\, \kappa\, (p-1)$ strategies, considering three indexes. The first index $i \in \{1, ..., n_{\mathfrak{p}}\}$ corresponds to the strategy being tested; the second index $k \in \{1, ..., \kappa\}$ corresponds to the trial number; the third index $o \in \mathcal{P} \setminus \mathfrak{p}$ corresponds to the population from which a co-player is sampled, so $a_{i,k,o}^{-\mathfrak{p}} \in A^o$ is the strategy of the co-player from population $o \neq \mathfrak{p}$ sampled when conducting the k-th trial of strategy $a_i^{\mathfrak{p}} \in A^{\mathfrak{p}}$. Let $\Phi_{A,\kappa}^{-\mathfrak{p}}$ be the set of all such indexed sequences of $n_{\mathfrak{p}}\, \kappa\, (p-1)$ strategies.

For $a^{-\mathfrak{p}} \in \Phi_{A,\kappa}^{-\mathfrak{p}}$, let $\pi^{\mathfrak{p}}(a^{-\mathfrak{p}})$ be the $n_{\mathfrak{p}}$-tuple of total payoffs $(\pi_i^{\mathfrak{p}}(a^{-\mathfrak{p}}))$ obtained by each strategy $a_i^{\mathfrak{p}} \in A^{\mathfrak{p}}$, i.e.,

$$\pi_i^{\mathfrak{p}}(a^{-\mathfrak{p}}) = \sum_{k=1}^{\kappa} U^{\mathfrak{p}}(a_i^{\mathfrak{p}}; \bar{a}_{i,k}^{-\mathfrak{p}})$$

where $\bar{a}_{i,k}^{-\mathfrak{p}} \equiv (a_{i,k,o}^{-\mathfrak{p}})_{o\in\mathcal{P}\setminus\mathfrak{p}}$ is the $(p-1)$-tuple of strategies used by co-players of a revising agent from population $\mathfrak{p}$ when conducting the k-th trial of strategy $a_i^{\mathfrak{p}} \in A^{\mathfrak{p}}$.

Under a BEP$_{all}$ protocol, the probability that a revising agent from population $\mathfrak{p}$ chooses strategy $a_i^{\mathfrak{p}} \in A^{\mathfrak{p}}$ at population state x is given by

$$w_i^{\mathfrak{p},\kappa,\beta}(x) = \sum_{j=1}^{n^{\mathfrak{p}}} x_j^{\mathfrak{p}} \sum_{a^{-\mathfrak{p}}\in\Phi_{A,\kappa}^{-\mathfrak{p}}} \beta_{ji}^{\mathfrak{p}}(\pi^{\mathfrak{p}}(a^{-\mathfrak{p}})) \prod_{o\in\mathcal{P}\setminus\mathfrak{p}} \prod_{l=1}^{n_o} (x_l^o)^{\mathbb{I}_l^o(a^{-\mathfrak{p}})} \tag{A.1}$$

where the exponent $\mathbb{I}_l^o(a^{-\mathfrak{p}})$ indicates the number of occurrences of strategy a_l^o (whose prevalence is x_l^o, in population o) in the tuple $a^{-\mathfrak{p}}$, and where the functions $\beta_{ji}^{\mathfrak{p}} : \mathbb{R}^{n_{\mathfrak{p}}} \to [0,1]$ define the tie-breaking rule. And the payoff-sampling dynamics BEP$_{all}(\kappa, \beta)$ is given by

$$\dot{x}_i^{\mathfrak{p}} = w_i^{\mathfrak{p},\kappa,\beta}(x) - x_i^{\mathfrak{p}} \tag{A.2}$$

for each $\mathfrak{p} \in \mathcal{P}$ and $i \in \{1, .., n^{\mathfrak{p}}\}$.

Let $\mathcal{H}$ be the set of all nonempty product sets $H \subseteq A$, i.e., $H = \times_{\mathfrak{p}\in\mathcal{P}} H^{\mathfrak{p}}$, where $\emptyset \neq H^{\mathfrak{p}} \subseteq A^{\mathfrak{p}}$, for all $\mathfrak{p} \in \mathcal{P}$. For any $H \in \mathcal{H}$, let $[\Delta]_H = \times_{\mathfrak{p}\in\mathcal{P}} \Delta_{H^{\mathfrak{p}}}$ be the face of the polyhedron $[\Delta]_A$ spanned by H. $[\Delta]_H$ is itself a polyhedron of (global) population states associated with the reduced game in which the pure strategy set in population $\mathfrak{p}$ is $H^{\mathfrak{p}}$ (Ritzberger and Weibull, 1995).

In the next definition, $x = (x^1, x^2, ..., x^p)$ is a global population state, with $x^{\mathfrak{p}}$ being the state in population $\mathfrak{p}$, and $w^{\kappa,\beta}(x)$ is a global "inflow" vector made up by the inflow vectors $w^{\mathfrak{p},\kappa,\beta}(x)$ for each population.

Definition 5. *A product set $H \in \mathcal{H}$ is closed under a BEP$_{all}(\kappa, \beta)$ protocol if for all $x \in [\Delta]_H$, $w^{\kappa,\beta}(x) \in [\Delta]_H$.*

Definition 6. *Closed under payoff sampling. A product set $H \in \mathcal{H}$ is closed under payoff sampling (CUPS) if H is closed under every BEP$_{all}(\kappa, \beta)$ protocol.*

If H is a CUPS product set we say that $[\Delta]_H$ is a CUPS face.

Given a product set $H \in \mathcal{H}$, and considering a particular population $\mathfrak{p} \in \mathcal{P}$, let $H^{-\mathfrak{p}} = \times_{(o \in \mathcal{P}, o \neq \mathfrak{p})} H^o$ be the product set of the subsets of strategies H^o in populations o other than $\mathfrak{p}$. $H^{-\mathfrak{p}}$ contains all the strategy profiles of co-players that a player from population $\mathfrak{p}$ may face when revising at a state $x \in [\Delta]_H$.

Proposition A.1. *A product set $H \in \mathcal{H}$ is closed under payoff sampling if and only if for every population $\mathfrak{p} \in \mathcal{P}$:*

$$\max_{a_i \in (A^{\mathfrak{p}} \setminus H^{\mathfrak{p}}),\, \bar{a} \in H^{-\mathfrak{p}}} U^{\mathfrak{p}}(a_i; \bar{a}) < \max_{a_j \in H^{\mathfrak{p}}} \min_{\bar{a} \in H^{-\mathfrak{p}}} U^{\mathfrak{p}}(a_j; \bar{a})$$

Propositions 3.2 and 3.3 can be adapted directly to the multi-population case, replacing "a nonempty subset of strategies" with "a nonempty product set $H \in \mathcal{H}$". For proposition 3.4, adapted below, we need to consider, instead of the support of an equilibrium (single population case), the product set of the supports of the equilibrium in each population.

Proposition A.2. *If x is a regular $S(\kappa, \beta^r)$ equilibrium, then the product set of the supports of x in each population $\mathfrak{p} \in \mathcal{P}$ is a CUPS set.*

For the adaptation of proposition 3.5 to the multi-population setting, let the (relative) interior of $[\Delta]_H$, represented as int($[\Delta]_H$), be the set of population states $x \in [\Delta]_H$ such that $x_i^{\mathfrak{p}} > 0$ for every $\mathfrak{p} \in \mathcal{P}$ and for every i such that $a_i^{\mathfrak{p}} \in H^{\mathfrak{p}}$. And let the boundary of $[\Delta]_H$, represented as bd($[\Delta]_H$), be the set of population states $x \in [\Delta]_H$ such that $x_i^{\mathfrak{p}} = 0$ for some $\mathfrak{p} \in \mathcal{P}$ and some i such that $a_i^{\mathfrak{p}} \in H^{\mathfrak{p}}$.

Proposition A.3. *Consider any number of trials $\kappa \in \mathbb{N}$ and any regular tie-breaking rule β^r. H is a minimal CUPS set if and only if there is a regular $S(\kappa, \beta^r)$ equilibrium in int($[\Delta]_H$) and there is no regular $S(\kappa, \beta^r)$ equilibrium in bd($[\Delta]_H$).*

Our next two propositions can be adapted directly from the single-population case.

Proposition A.4. *If $[\Delta]_H$ is a CUPS face, then it is asymptotically stable under every payoff-sampling dynamic BEP$_{all}(\kappa, \beta)$ with $\kappa > k_0$, for some finite k_0.*

Proposition A.5. *If $H \in \mathcal{H}$ is CUPS, then H is Closed Under Rational Behavior (CURB).*

For this last proposition, and similarly to the single-population case, we call a product set H closed under rational behavior (CURB) if it contains all its best replies, i.e., if $\mathrm{BR}(x) \subseteq H$ for every $x \in [\Delta_H]$, where BR is the pure best-reply correspondence which maps populations states to their pure best-reply strategy combinations (Basu and Weibull, 1991; Ritzberger and Weibull, 1995).

Last, for the instability results, considering a CUPS product set $H = \times_{\mathfrak{p}\in\mathcal{P}} H^{\mathfrak{p}}$ and defining $H^{-\mathfrak{p}_1,\mathfrak{p}_2} = \times_{(o\in\mathcal{P}, o\neq\mathfrak{p}_1, o\neq\mathfrak{p}_2)} H^{o}$, we say that strategy $a_j^{\mathfrak{p}_1} \in (H^{\mathfrak{p}_1})^c$ *supports* invasion (of H) by strategy $a_k^{\mathfrak{p}_2} \in (H^{\mathfrak{p}_2})^c$ if

$$\min_{\bar{a}\in H^{-\mathfrak{p}_1,\mathfrak{p}_2}} U^{\mathfrak{p}_2}(a_k^{\mathfrak{p}_2}; a_j^{\mathfrak{p}_1}, \bar{a}) > \max_{a_i\in H^{\mathfrak{p}_2}} \max_{\bar{a}\in H^{-\mathfrak{p}_2}} U^{\mathfrak{p}_2}(a_i; \bar{a})$$

Or, for the two-population case, if

$$U^{\mathfrak{p}_2}(a_k^{\mathfrak{p}_2}; a_j^{\mathfrak{p}_1}) > \max_{a_i\in H^{\mathfrak{p}_2}} \max_{a_m\in H^{\mathfrak{p}_1}} U^{\mathfrak{p}_2}(a_i; a_m)$$

A CUPS product set H is *partially twice inferior* if there is a product subset of strategies $H_0^c \subseteq \times_{\mathfrak{p}}(H^{\mathfrak{p}})^c$, with elements for at least two populations, such that every strategy in every population with elements in H_0^c supports invasion by at least two other strategies in H_0^c. If the condition holds for $H_0^c = \times_{\mathfrak{p}}(H^{\mathfrak{p}})^c$, we call the set twice inferior.

Proposition A.6. *For p-player games played in p populations:*

- *If a CUPS set H is twice inferior, then face Δ_H is repelling under $BEP_{all}(1,\beta)$ dynamics.*
- *If a CUPS set H is partially twice inferior, then face Δ_H is unstable under $BEP_{all}(1,\beta)$ dynamics.*

B Appendix: Proofs

Proof of proposition 3.1. Let $H^c \equiv (A \setminus H)$. The condition in proposition 3.1 is equivalent to

$$\max_{a_i\in H^c, \bar{a}\in H^{(p-1)}} U_{a_i,\bar{a}} < \max_{a_j\in H} \min_{\bar{a}\in H^{(p-1)}} U_{a_j,\bar{a}}. \tag{B.1}$$

Let $M_{(H^c,H)} \equiv \max_{a_i\in H^c,\bar{a}\in H^{(p-1)}} U_{a_i,\bar{a}}$ and let $Maxmin_H \equiv \max_{a_j\in H} \min_{\bar{a}\in H^{(p-1)}} U_{a_j,\bar{a}}$. The best payoff that the strategies in H^c can obtain at a state $x \in \Delta_H$ is lower or equal than $\kappa\, M_{(H^c,H)}$, and this upper bound for the maximum payoff of the strategies in H^c is obtained with positive probability at any $x \in \mathrm{int}(\Delta_H)$ (i.e., the relative interior of Δ_H). The best payoff obtained by the strategies in H at a state $x \in \Delta_H$ is greater or equal than $\kappa\, Maxmin_H$, and this lower bound for the best payoff of the strategies in H is obtained with positive probability at any $x \in \mathrm{int}(\Delta_H)$. Consequently, if (B.1) holds, then $w_i^{\kappa,\beta}(x) = 0$ for all $x \in \Delta_H$ and all i such that $a_i \in H^c$, and also for any κ and any tie-breaking rule β, proving that H is CUPS. Similarly, considering any $x \in \text{int}(\Delta_H)$, we find that (B.1) is a necessary condition for H to be CUPS: if (B.1) does not hold, there is some $a_i \in H^c$ such that $w_i^{\kappa,\beta^r}(x) > 0$ for any $x \in \text{int}(\Delta_H)$, any $\kappa \in \mathbb{N}$ and any regular tie-breaking rule β^r. $\square$

Proof of proposition 3.2. From the proof of proposition 3.1 we know that (B.1) is a necessary condition for a set H to be closed under some (any) regular BEP$_{all}(\kappa, \beta^r)$ protocol, and (B.1) is also a sufficient condition to be closed under every BEP$_{all}(\kappa, \beta)$ protocol. $\square$

Proof of proposition 3.3. If H is CUPS, then $w_i^{\kappa,\beta}(x) = 0$ for every i such that $a_i \in H^c$, $x \in \Delta_H$, $\kappa \in \mathbb{N}$ and tie-breaking rule β, so $w^{\kappa,\beta}(x)$ is a continuous function that maps the compact Δ_H onto itself. By Brouwer's fixed point theorem, Δ_H contains at least one fixed point x such that $w^{\kappa,\beta}(x) = x$, which is an $S(\kappa, \beta)$ equilibrium. $\square$

Proof of proposition 3.4. The proof is conducted by contradiction. Let $\underline{x}$ be a regular $S(\kappa, \beta^r)$ equilibrium and let H be the support of $\underline{x}$. Note that there is a positive probability that all strategies in H, when tested κ times at state $\underline{x}$, obtain an average payoff no greater than the maxmin value $Maxmin_H \equiv \max_{a_j \in H} \min_{\bar{a} \in H^{(p-1)}} U_{a_j,\bar{a}}$. Additionally, if H is not CUPS, there is a positive probability that some strategy $a_i \in H^c$ with $\underline{x}_i = 0$ obtains an average payoff greater than or equal to $Maxmin_H$ and is selected (because the tie-breaking rule is regular), so $w_i^{\kappa,\beta^r}(\underline{x}) > 0$, contradicting the fact that $\underline{x}$ is an equilibrium with support H. $\square$

Proof of proposition 3.5. Let H be a minimal CUPS set. By proposition 3.3, for every κ and β^r, H contains the support of at least one $S(\kappa, \beta^r)$ equilibrium. By proposition 3.4, no proper subset of H can be the support of a regular payoff-sampling equilibrium $S(\kappa, \beta^r)$, so there is no regular $S(\kappa, \beta^r)$ equilibrium $\underline{x}$ with $supp(\underline{x}) \subset H$. Now, fix κ and β^r and let y be an $S(\kappa, \beta^r)$ equilibrium such that there is no $S(\kappa, \beta^r)$ equilibrium y' with $supp(y') \subset supp(y) \equiv J$. By proposition 3.4, J is CUPS. If J is not a minimal CUPS set, then, by proposition 3.3, there is some $S(\kappa, \beta^r)$ equilibrium y' with $supp(y') \subset J$ (a contradiction), so J is a minimal CUPS set. $\square$

Proof of proposition 3.6. If $H = A$, the result is immediate. Otherwise, the proof considers that, in a neighborhood of Δ_H, a revising agent conducting a battery of tests will meet mainly co-players using strategies in H, and, for states sufficiently close to Δ_H, the probability that a revising agent meets more than one co-player using some strategy in $H^c = A \backslash H$ (i.e., not in H) becomes negligible, compared with the probability of meeting either none or just one of such "deviating" co-players. Defining $s_i^{\kappa,\beta}(x)$ as the probability of selecting strategy a_i under those most-likely events, and letting $\epsilon \equiv \sum_{i:a_i \in H^c} x_i$, we will show that:

i) $w_i^{\kappa,\beta}(x) = s_i^{\kappa,\beta}(x) + O(\epsilon^2)$ as $\epsilon \to 0$.

ii) There is a finite bound k_0 such that, for $\kappa > k_0$ and every i such that $a_i \in H^c$, $s_i^{\kappa,\beta}(x) = 0$.

Consequently, for $\kappa > k_0$ the dynamics (3) are such that $\sum_{i:a_i \in H^c} \dot{x}_i = \dot{\epsilon} = -\epsilon + O(\epsilon^2)$, which guarantees asymptotic stability of Δ_H.

Let $s_i^{\kappa,\beta}(x)$ be the probability of selecting strategy a_i at state x in a battery of tests such that either none or exactly one of the $n\,(p-1)\,\kappa$ sampled co-players (the "deviating co-player") uses some strategy a_j in H^c, while all the other co-players use strategies in H. Note in (2) that the probability of meeting more than one deviating co-player in a battery of tests at state x – corresponding to all sequences of strategies a^{bat} with at least two strategies belonging to H^c – involves a sum of monomials $\prod_{l=1}^{n} x_l^{\mathbb{I}_l(a^{bat})}$ such that the sum of the exponents $\mathbb{I}_l(a^{bat})$ corresponding to strategies $a_l \in H^c$ is at least two. Considering $\epsilon \equiv \sum_{i:a_i \in H^c} x_i$, which implies $x_i \leq \epsilon$ for all i such that $a_i \in H^c$, we consequently have that $w_i^{\kappa,\beta}(x) = s_i^{\kappa,\beta}(x) + O(\epsilon^2)$ as $\epsilon \to 0$.

There are three possible cases that we need to consider to calculate $s_i^{\kappa,\beta}(x)$:

a) There is no deviating co-player met in a battery of tests. In this case, the best payoff achieved by the strategies in H is bounded below by $LB_H^a \equiv \kappa\, Maxmin_H$, where $Maxmin_H \equiv \max_{a_j \in H} \min_{\bar{a} \in H^{(p-1)}} U_{a_j,\bar{a}}$. And the best payoff achieved by the strategies in H^c is bounded above by $UB_{H^c}^a \equiv \kappa \max_{a_i \in H^c, \bar{a} \in H^{p-1}} U_{a_i,\bar{a}}$. Consequently, considering that H is CUPS, the best payoff in this case is obtained exclusively by strategies in H.

b) The deviating co-player is met when testing some strategy $a_i \in H$. In this case, the best payoff achieved by the strategies in H is bounded below by $LB_H^b \equiv (\kappa - 1)\, Maxmin_H + \max_{a_m \in M(H)} \min_{\bar{a} \in H^{p-1}, a_j \in H^c} U_{a_m,\bar{a}^j}$, where:

   - $M(H)$ is the set of maxmin strategies $M(H) \equiv \{a_i \in H : \min_{\bar{a} \in H^{(p-1)}} U_{a_i,\bar{a}} = Maxmin_H\}$ and
   - $\bar{a}^j$ is a modification of a $(p-1)$-tuple of strategies $\bar{a} \in H^{(p-1)}$, in which one of the strategies has been replaced by strategy $a_j \in H^c$.

   And the best payoff achieved by the strategies in H^c is bounded above by $UB_{H^c}^b \equiv \kappa \max_{a_i \in H^c, \bar{a} \in H^{p-1}} U_{a_i,\bar{a}}$. Consequently, for

$$\kappa > b_1 \equiv \frac{Maxmin_H - \max_{a_m \in M(H)} \min_{\bar{a} \in H^{p-1}, a_j \in H^c} U_{a_m,\bar{a}^j}}{Maxmin_H - \max_{a_i \in H^c, \bar{a} \in H^{p-1}} U_{a_i,\bar{a}}} < \infty$$

   the best payoff in this case is obtained exclusively by strategies in H.

c) The deviating co-player is met when testing some strategy $a_i \in H^c$. In this case, the best payoff achieved by the strategies in H is bounded below by $LB_H^c \equiv \kappa\, Maxmin_H$, and the best payoff achieved by the strategies in H^c is bounded above by $UB_{H^c}^c \equiv (\kappa - 1) \max_{a_i \in H^c, \bar{a} \in H^{p-1}} U_{a_i,\bar{a}} + \max_{a_i \in H^c, \bar{a} \in H^{p-1}, a_j \in H^c} U_{a_i,\bar{a}^j}$, with $\bar{a}^j$ defined as before. Consequently, for

$$\kappa > b_2 \equiv \frac{\max_{a_i \in H^c, \bar{a} \in H^{p-1}, a_j \in H^c} U_{a_i,\bar{a}^j} - \max_{a_i \in H^c, \bar{a} \in H^{p-1}} U_{a_i,\bar{a}}}{Maxmin_H - \max_{a_i \in H^c, \bar{a} \in H^{p-1}} U_{a_i,\bar{a}}} < \infty$$

   the best payoff in this case is obtained exclusively by strategies in H.

Looking at the three previous cases, we have that, for $\kappa > k_0 \equiv \max(b_1, b_2)$, the best payoff in the three cases above is obtained exclusively by strategies in H, which implies that, for i such that $a_i \in H^c$, we have $s_i^{\kappa,\beta}(x) = 0$ and $w_i^{\kappa,\beta}(x) = O(\epsilon^2)$, which, considering (3), leads to $\sum_{i:a_i \in H^c} \dot{x}_i = \dot{\epsilon} = -\epsilon + O(\epsilon^2)$. Consequently, there is some positive $\epsilon_0 > 0$ and some positive constant $\lambda > 0$ such that, for every x with $\sum_{i:a_i \in H^c} x_i = \epsilon < \epsilon_0$, we have $\dot{\epsilon} \leq -\lambda\,\epsilon$, proving asymptotic stability of the face Δ_H. □

Proof of proposition 3.7. Let us first prove the repelling result, based on Sethi (2000). As in the previous proof, let $H^c = (A \backslash H)$ and let $\epsilon = \sum_{j:a_j \in H^c} x_j$. We will show that, if H is an inferior CUPS set, then $\text{BEP}_{all}(1, \beta)$ dynamics satisfy the equation $\dot{\epsilon} \geq (p-2)\epsilon + O(\epsilon^2)$. This implies that, for $p > 2$, face Δ_H is repelling: there is a neighborhood of Δ_H in which, from any initial state $x^0 \notin \Delta_H$ (so $\epsilon(t=0) = \sum_{j:a_j \in H^c} x_j^0 > 0$) in the neighborhood, the value of $\epsilon(t)$ grows exponentially until the state leaves the neighborhood.

Suppose that strategy $a_j \in H^c$ supports invasion (of H) by strategy $a_k \in H^c$. This implies that

$$\dot{x}_k \geq (p-1)x_j + O(\epsilon^2) - x_k \tag{B.2}$$

which follows from considering the event (cases) in which, when a revising agent tests strategy a_k, one of the $(p-1)$ sampled co-players uses strategy a_j, while all other co-players in the battery of tests use some strategy in H: if a_j supports a_k, this event leads to the revising agent choosing strategy a_k, and it happens with probability $(p-1)x_j(1-\epsilon)^{n(p-1)-1}$, a function which, considering that $x_j \leq \epsilon$ and the binomial expansion of $(1-\epsilon)^{n(p-1)-1}$, is $(p-1)x_j + O(\epsilon^2)$.

Now, let B be the subset of strategies in H^c that support some strategy in H^c. If H is inferior, then $B = H^c$, and from (B.2), we have:

$$\dot{\epsilon} = \sum_{k:a_k \in H^c} \dot{x}_k \geq (p-1) \sum_{j:a_j \in B} x_j - \sum_{k:a_k \in H^c} x_k + O(\epsilon^2) = (p-2)\epsilon + O(\epsilon^2)$$

proving that Δ_H is repelling for $p > 2$.

If there are several strategies supporting a_k, the term x_j in (B.2) can be substituted by the sum of the proportions of those supporting strategies. It is then easy to adapt the last step, considering the subset of strategies in H^c that support at least two strategies in H^c, to show that, if H is twice inferior, then

$$\dot{\epsilon} \geq (2p-3)\epsilon + O(\epsilon^2)$$

proving that Δ_H is repelling for $p \geq 2$.

Let us now consider the instability result. Consider the partially inferior set H, its complement $H^c \equiv (A \setminus H)$, a subset $H_1^c \subset H^c$ satisfying the condition required for H to be partially inferior, and the set $H_2^c \equiv (H^c \setminus H_1^c)$. Let $\epsilon_1 \equiv \sum_{j:a_j \in H_1^c} x_j$ and $\epsilon_2 \equiv \sum_{j:a_j \in H_2^c} x_j$. We will show that there are positive constants $\epsilon_{1,0}$ and $\epsilon_{2,0}$ such that ϵ_1 grows exponentially in the neighborhood O^0 of Δ_H defined by ($\epsilon_1 < \epsilon_{1,0}$ and $\epsilon_2 < \epsilon_{2,0}$), so any state trajectory with initial value $x^0 \in O^0$ such that $\epsilon_1 > 0$ (i.e., such that $\epsilon_1(t=0) = \sum_{j:a_j \in H_1^c} x_j^0 > 0$) leaves O^0, proving instability.

Considering the same event as before for a revising agent, if strategy $a_j \in H_1^c$ supports invasion (of H) by strategy $a_k \in H_1^c$ we have:

$$\dot{x}_k \geq (p-1)x_j(1-\epsilon_1-\epsilon_2)^{n(p-1)-1} - x_k$$

So, for any positive constant value $\epsilon_{2,0}$, if $\epsilon_2 < \epsilon_{2,0}$, then

$$\dot{x}_k \geq (p-1)x_j(1-\epsilon_1-\epsilon_{2,0})^{n(p-1)-1} - x_k \geq (p-1)x_j[(1-\epsilon_{2,0})^{n(p-1)-1} - f(\epsilon_1)] - x_k \quad \text{(B.3)}$$

where $f(\epsilon_1)$ is $O(\epsilon_1)$: this can be checked easily from the binomial expansion of $((1-\epsilon_{2,0})-\epsilon_1)^{n(p-1)-1}$, substituting 1 (if the coefficient in the expansion is negative) or 0 (if the coefficient is positive) for the $(1-\epsilon_{2,0})$ terms of the expansion. Adding (B.3) for $k : a_k \in H_1^c$, and considering that $\epsilon_1 = \sum_{k:a_k \in H_1^c} x_k$, we obtain

$$\dot{\epsilon_1} \geq [(p-1)(1-\epsilon_{2,0})^{n(p-1)-1} - 1]\epsilon_1 - O(\epsilon_1^2) \quad \text{(B.4)}$$

By taking $\epsilon_{2,0}$ sufficiently small, we can make $(1-\epsilon_{2,0})^{n(p-1)-1}$ as close to 1 as we want. This, combined with (B.4) implies that for any positive values $\lambda_1 < 1$ and $\lambda_2 < 1$ (take e.g. $\lambda_1 = \lambda_2 = 0.9$), we can find positive values $\epsilon_{2,0}$ and $\epsilon_{1,0}$ such that if $\epsilon_2 < \epsilon_{2,0}$ and $\epsilon_1 < \epsilon_{1,0}$ then

$$\dot{\epsilon_1} \geq [(p-1)\lambda_2 - 1]\lambda_1\epsilon_1$$

which (taking e.g. $\lambda_1 = \lambda_2 = 0.9$) leads to the exponential growth of ϵ_1 in O^0 for $p > 2$. The adaptation of the proof to partially twice inferior sets is done as before, considering that, if there are several strategies supporting a_k, the term x_j in (B.3) can be substituted by the sum of the proportions of those supporting strategies. □

Proof of proposition 4.1. The proof coincides with the first part of the proof of proposition 3.1. □

Bibliography

Arigapudi, S., Heller, Y., and Milchtaich, I. (2021). Instability of defection in the prisoner's dilemma under best experienced payoff dynamics. *Journal of Economic Theory*, 197:105174.

Arigapudi, S., Heller, Y., and Schreiber, A. (2022). Sampling dynamics and stable mixing in hawk-dove games. Working paper available at https://arxiv.org/abs/2107.08423.

Balkenborg, D. and Schlag, K. H. (2007). On the evolutionary selection of sets of Nash equilibria. *Journal of Economic Theory*, 133(1):295–315.

Balkenborg, D. G., Hofbauer, J., and Kuzmics, C. (2013). Refined best-response correspondence and dynamics. *Theoretical Economics*, 8(1):165–192.

Basu, K. (1994). The traveler's dilemma: Paradoxes of rationality in game theory. *American Economic Review Papers and Proceedings*, 84:391–395.

Basu, K. and Weibull, J. W. (1991). Strategy subsets closed under rational behavior. *Economics Letters*, 36(2):141–146.

Benaïm, M. and Weibull, J. W. (2003). Deterministic approximation of stochastic evolution in games. *Econometrica*, 71(3):873–903.

Berkemer, R. (2008). Disputable advantage of experience in the traveler's dilemma. Unpublished manuscript, Technical University of Denmark. Abstract in *International Conference on Economic Science with Heterogeneous Interacting Agents, Warsaw, 2008*.

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

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

Eshel, I. and Akin, E. (1983). Coevolutionary instability of mixed nash solutions. *J. Math. Biol*, 18(2):123–133.

Izquierdo, L. R. and Izquierdo, S. S. (2022a). BEP-3s-sp. Software available at https://doi.org/10.5281/zenodo.7156265.

Izquierdo, L. R., Izquierdo, S. S., and Rodriguez, J. (2022). Fast and scalable global convergence in single-optimum decentralized coordination problems. *IEEE Transactions on Control of Network Systems*.

Izquierdo, L. R., Izquierdo, S. S., and Sandholm, W. H. (2018). *EvoDyn-3s*: A Mathematica computable document to analyze evolutionary dynamics in 3-strategy games. *SoftwareX*, 7:226–233.

Izquierdo, S. S. and Izquierdo, L. R. (2021). "Test two, choose the better" leads to high cooperation in the centipede game. *Journal of Dynamics and Games*.

Izquierdo, S. S. and Izquierdo, L. R. (2022b). Stability of strict equilibria in best experienced payoff dynamics: Simple formulas and applications. *Journal of Economic Theory*, 105553.

Kohlberg, E. and Mertens, J.-F. (1986). On the Strategic Stability of Equilibria. *Econometrica*, 54(5):1003.

Mantilla, C., Sethi, R., and Cárdenas, J. C. (2020). Efficiency and stability of sampling equilibrium in public goods games. *Journal of Public Economic Theory*, 22(2):355–370.

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

Osborne, M. and Rubinstein, A. (1998). Games with Procedurally Rational Players. *American Economic Review*, 88(4):834–847.

Ritzberger, K. and Vogelsberger, K. (1990). The Nash Field. *IAS (IHS) Research Report 263*.

Ritzberger, K. and Weibull, J. W. (1995). Evolutionary Selection in Normal-Form Games. *Econometrica*, 63(6):1371.

Rosenthal, R. W. (1981). Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, 25:92–100.

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

Samuelson, L. and Zhang, J. (1992). Evolutionary stability in asymmetric games. *Journal of Economic Theory*, 57(2):363–391.

Sandholm, W. H. (2010). *Population games and evolutionary dynamics*. The 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.

Selten, R. and Chmura, T. (2008). Stationary concepts for experimental 2x2-games. *American Economic Review*, 98(3):938–66.

Sethi, R. (2000). Stability of Equilibria in Games with Procedurally Rational Players. *Games and Economic Behavior*, 32(1):85–104.

Sethi, R. (2021). Stable sampling in repeated games. *Journal of Economic Theory*, 197:105343.

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.

van Damme, E. (1991). *Stability and Perfection of Nash Equilibria*. Springer Berlin Heidelberg, Berlin, Heidelberg.

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