Best Experienced Payoff Dynamics and Cooperation in the Centipede Game*

William H. Sandholm†, Segismundo S. Izquierdo‡, and Luis R. Izquierdo§

February 8, 2018

Abstract

We study population game dynamics under which each revising agent randomly selects a set of strategies according to a given test-set rule, plays each strategy in this set a fixed number of times, with each play of each strategy being against a newly drawn opponent, and chooses the strategy whose total payoff was highest, breaking ties according to a given tie-breaking rule. In the Centipede game, these *best experienced payoff dynamics* lead to cooperative play. Play at the almost globally stable state is concentrated on the last few nodes of the game, with the proportions of agents playing each strategy being dependent on the specification of the dynamics, but largely independent of the length of the game. The emergence of cooperative play is robust to allowing agents to test candidate strategies many times, and to introducing substantial proportions of agents who always stop immediately. Since best experienced payoff dynamics are defined by random sampling procedures, they are represented by systems of polynomial differential equations with rational coefficients, allowing us to establish key properties of the dynamics using tools from computational algebra.

---

*We thank Ken Judd, Panayotis Mertikopoulos, Erik Mohlin, Ignacio Monzón, Marzena Rostek, Ariel Rubinstein, Larry Samuelson, Ryoji Sawa, Lones Smith, Mark Voorneveld, Marek Weretka, and especially Antonio Penta for helpful discussions and comments. Financial support from the U.S. National Science Foundation (SES-1458992 and SES-1728853), the U.S. Army Research Office (MSN201957), project ECO2017-83147-C2-2-P (MINECO/AEI/FEDER, UE), and the Spanish Ministerio de Educación, Cultura, y Deporte (PRX15/00362 and PRX16/00048) is gratefully acknowledged.

†Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706, USA. e-mail: `whs@ssc.wisc.edu`; website: `www.ssc.wisc.edu/~whs`.

‡Department of Industrial Organization, Universidad de Valladolid, Paseo del Cauce 59, 47011 Valladolid, Spain. e-mail: `segis@eis.uva.es`; website: `www.segis.izqui.org`.

§Department of Civil Engineering, Universidad de Burgos, Edificio la Milanera, Calle de Villadiego, 09001 Burgos, Spain. e-mail: `lrizquierdo@ubu.es`; website: `www.luis.izqui.org`.

1. Introduction

The discrepancy between the conclusions of backward induction reasoning and observed behavior in certain canonical extensive form games is a basic puzzle of game theory. The Centipede game (Rosenthal (1981)), the finitely repeated Prisoner's Dilemma, and related examples can be viewed as models of relationships in which each participant has repeated opportunities to take costly actions that benefit his partner, and in which there is a commonly known date at which the interaction will end. Experimental and anecdotal evidence suggests that cooperative behavior may persist until close to the exogenous terminal date (McKelvey and Palfrey (1992)). But the logic of backward induction leads to the conclusion that there will be no cooperation at all.

Work on epistemic foundations provides room for wariness about unflinching appeals to backward induction. To support this prediction, one must assume that there is always common belief that all players will act as payoff maximizers at all points in the future, even when many rounds of previous choices argue against such beliefs.[1] Thus the simplicity of backward induction belies the strength of the assumptions needed to justify it, and this strength may help explain why backward induction does not yield descriptively accurate predictions in some classes of games.[2]

This paper studies a dynamic model of behavior in games that maintains the assumption that agents respond optimally to the information they possess. But rather than imposing strong assumptions about agents' knowledge of opponents' intentions, we suppose instead that agents' information comes from direct but incomplete experience with playing the strategies available to them. As with earlier work of Osborne and Rubinstein (1998) and Sethi (2000), our model is best viewed not as one that incorporates irrational choices, but rather as one of rational choice under particular restrictions on what agents know.

Following the standard approach of evolutionary game theory, we suppose that two populations of agents are recurrently randomly matched to play a two-player game. This framework accords with some experimental protocols, and can be understood more broadly as a model of the formation of social norms (Young (1998)). At random times, each agent receives opportunities to switch strategies. At these moments the agent decides whether to continue to play his current strategy or to switch to an alternative in a test set, which is determined using a prespecified stochastic rule τ. The agent then plays

[1] For formal analyses, see Binmore (1987), Reny (1992), Stalnaker (1996), Ben-Porath (1997), Halpern (2001), and Perea (2014).

[2] As an alternative, one could apply Nash equilibrium, which also predicts noncooperative behavior in the games mentioned above, but doing so replaces assumptions about future rationality with the assumption of equilibrium knowledge, which may not be particularly more appealing—see Dekel and Gul (1997).

each strategy he is considering against κ opponents drawn at random from the opposing population, with each play of each strategy being against a newly drawn opponent. He then switches to the strategy that achieved the highest total payoff, breaking ties according to a prespecified rule β.[3]

Standard results imply that when the populations are large, the agents' aggregate behavior evolves in an essentially deterministic fashion, obeying a differential equation that describes the expected motion of the stochastic process described above (Benaïm and Weibull (2003)). The differential equations generated by the protocol above, and our object of study here, we call *best experienced payoff dynamics*, or *BEP dynamics* for short.

Our model builds on earlier work on games played by "procedurally rational players". In particular, when the test-set rule τ specifies that agents always test all strategies, and when the tie-breaking rule β is uniform randomization, then the rest points of our process (with $\kappa = k$) are the $S(k)$ equilibria of Osborne and Rubinstein (1998). The corresponding dynamics were studied by Sethi (2000). The analyses in these papers differ substantially from ours, as we explain carefully below.

Our analysis of best experienced payoff dynamics in the Centipede game uses techniques from dynamical systems theory. What is more novel is our reliance on algorithms from computational algebra and perturbation bounds from linear algebra, which allow us to solve exactly for the rest points of our differential equations and to perform rigorous stability analyses. We complement this approach with numerical analyses of cases in which analytical results cannot be obtained.

Most of our results focus on dynamics under which each tested strategy is tested exactly once ($\kappa = 1$), so that agents' choices only depend on ordinal properties of payoffs. In Centipede games, under BEP dynamics with tie-breaking rules that do not abandon optimal strategies, or that break ties in favor of less cooperative strategies, the backward induction state—the state at which all agents in both populations stop at their first opportunity—is a rest point. However, we prove that this rest point is always *repelling*: the appearance of agents in either population who cooperate to any degree is self-reinforcing, and eventually causes the backward induction solution to break down completely.

For all choices of test-set and tie-breaking rules we consider, the dynamics have exactly one other rest point.[4] Except in games with few decision nodes, the form of this rest point is essentially independent of the length of the game. In all cases, the rest point has virtually

[3] Tie-breaking rules are important in the context of extensive form games, where different strategies that agree on the path of play earn the same payoffs.

[4] While traditional equilibrium notions in economics require stasis of choice, interior rest points of population dynamics represent situations in which individuals' choices fluctuate even as the expected change in *aggregate* behavior is null—see Section 2.3.

all players choosing to continue until the last few nodes of the game, with the proportions playing each strategy depending on the test-set and tie-breaking rules. Moreover, this rest point is dynamically stable, attracting solutions from all initial conditions other than the backward induction state. Thus under a range of specifications, if agents make choices based on experienced payoffs, testing each strategy in their test sets once and choosing the one that performed best, then play converges to a stable rest point that exhibits high levels of cooperation.

To explain why, we first observe that cooperative strategies are most disadvantaged when they are most rare—specifically, in the vicinity of the backward induction state. Near this state, the most cooperative agents would obtain higher *expected payoffs* by stopping earlier. However, when an agent considers switching strategies, he tests each strategy in his test set against new, independently drawn opponents. He may thus test a cooperative strategy against a cooperative opponent, and less cooperative strategies against less cooperative opponents, in which case his best *experienced payoff* will come from the cooperative strategy. Our analysis confirms that this possibility indeed leads to instability (Sections 5 and 6).[5] After this initial entry, the high payoffs generated by cooperative strategies when matched against one another spurs their continued growth. This growth is only abated when virtually all agents are choosing among the most cooperative strategies. The exact nature of the final mixture depends on the specification of the dynamics, and can be understood by focusing on the effects of a small fraction of possible test results (Section 5.1.3).

To evaluate the robustness of these results, we alter our model by replacing a fraction of each population with "backward induction agents". Such agents never consider any behavior other than stopping at their initial decision node. Despite interactions with invariably uncooperative agents, the remaining agents persist in behaving cooperatively, with the exact degree depending on the specification of the dynamics and the length of the game. For longer games (specifically, games with $d = 20$ decision nodes), which offer more opportunities for successfully testing cooperative strategies, cooperative behavior is markedly robust, persisting even when two-thirds of the population always stops immediately (Section 7).

Our final results consider the effects of the number of trials κ of each strategy in the test set on predictions of play. It seems clear that if the number of trials is made sufficiently

[5]Specifically, linearizing any given specification of the dynamics at the backward induction state identifies a single eigenvector with a positive eigenvalue (Appendix A). This eigenvector describes the mixture of strategies in the two populations whose entry is self-reinforcing, and identifies the direction toward which all other disturbances of the backward induction state are drawn. Direct examination of the dynamics provides a straightforward explanation why the given mixture of entrants is successful (Example 5.3).

large, so that the agents' information about opponents' behavior is quite accurate, then the population's behavior should come to resemble a Nash equilibrium. Indeed, when agents possess exact information, so that aggregate behavior evolves according to the best response dynamic (Gilboa and Matsui (1991), Hofbauer (1995b)), results of Xu (2016) imply that every solution trajectory converges to the set of Nash equilibria, all of which entail stopping at the initial node.

Our analysis shows, however, that stable cooperative behavior can persist even for substantial numbers of trials. In the Centipede games of length $d = 4$ that we consider in this analysis, a unique, attracting interior rest point with substantial cooperation persists for moderate numbers of trials. For larger numbers of trials, the attractor is always a cycle, and includes significant amounts of cooperation for numbers of trials as large as 200. We explain in Section 8 how the robustness of cooperation to fairly large numbers of trials can be explained using simple central limit theorem arguments.

Our main technical contribution lies in the use of methods from computational algebra and perturbation theorems from linear algebra to obtain analytical results about the properties of our dynamics. The starting point for this analysis, one that suggests a broader scope for our approach, is that decision procedures based on sampling from a population are described by multivariate polynomials with rational coefficients. In particular, BEP dynamics are described by systems of such equations, so finding their rest points amounts to finding the zeroes of these polynomial systems. To accomplish this, we compute a *Gröbner basis* for the set of polynomials that defines each instance of our dynamics; this new set of polynomials has the same zeroes as the original set, but its zeroes can be computed by finding the roots of a single (possibly high-degree) univariate polynomial.[6] Exact representations of these roots, known as *algebraic numbers*, can then be obtained by factoring the polynomial into irreducible components, and then using algorithms based on classical results to isolate each component's real roots.[7] These methods are explained in detail in Section 4. With these exact solutions in hand, we can rigorously assess the rest points' local stability through a linearization analysis. In order to obviate certain intractable exact calculations, this analysis takes advantage of both an eigenvalue perturbation theorem and a bound on the condition number of a matrix that does not require the computation of its inverse (Appendix B).

The code used to obtain the exact and numerical results is available as a *Mathematica* notebook posted on the authors' websites. An online appendix provides background and details about both the exact and the numerical analyses, reports certain numerical results

[6] See Buchberger (1965) and Cox et al. (2015). For applications of Gröbner bases in economics, see Kubler et al. (2014).

[7] See von zur Gathen and Gerhard (2013), McNamee (2007), and Akritas (2010).

in full detail, and presents a few proofs omitted here.

Traditionally, work in evolutionary game theory has focused on population dynamics under which Nash equilibria correspond to rest points, and sometimes conversely.[8] To the extent that observed behavior in populations differs from Nash equilibrium, models that aim to describe this behavior will need to take novel forms. The Gröbner basis techniques we use to compute rest points and the perturbation methods we develop to evaluate their local stability are not limited in scope to best experienced payoff dynamics, but can be applied to any game dynamic described by polynomials. Thus a basic contribution of this paper is to introduce techniques that will bring new models of population dynamics, especially ones motivated by descriptive considerations, under the purview of exact analysis.

Related literature

Previous work relating backward induction and deterministic evolutionary dynamics has focused on the replicator dynamic of Taylor and Jonker (1978) and the best response dynamic of Gilboa and Matsui (1991) and Hofbauer (1995b). Cressman and Schlag (1998) (see also Cressman (1996, 2003)) show that in generic perfect information games, every interior solution trajectory of the replicator dynamic converges to a Nash equilibrium. Likewise, Xu (2016) (see also Cressman (2003)) shows that in such games, every solution trajectory of the best response dynamic converges to a component of Nash equilibria. In both cases, the Nash equilibria approached need not be subgame perfect, and the Nash equilibrium components generally are not locally stable. Focusing on the Centipede game with three decision nodes, Ponti (2000) shows numerically that perturbed versions of the replicator dynamic exhibit cyclical behavior, with trajectories approaching and then moving away from the Nash component. In contrast, we show that best experienced payoff dynamics lead to a stable distribution of cooperative strategies far from the Nash component.

There are other deterministic dynamics from evolutionary game theory based on information obtained from samples.[9] Oyama et al. (2015) (see also Young (1993), Sandholm (2001), Kosfeld et al. (2002), and Kreindler and Young (2013)) study *sampling best response dynamics*, under which revising agents observe the choices made by a random sample of opponents, and play best responses to the distribution of choices in the sample. Droste et al. (2003) show that with a sample size of 1 and uniform tie-breaking, the unique rest

[8]For overviews of the relevant literature, see Weibull (1995) and Sandholm (2010b, 2015).

[9]In a related model of stochastic stability, Robson and Vega-Redondo (1996) assume that agents make choices based on outcomes from a single matching of population members.

point of this dynamic is interior with linearly decreasing strategy weights. Uniform tie-breaking is essential for this conclusion: if the other tie-breaking rules we consider here are used instead, the unique rest point of the resulting dynamic is the backward induction state.

While both are based on random samples, sampling best response dynamics and best experienced payoff dynamics differ in important respects. Sampling best response dynamics not only require agents to understand the game's payoff structure, but also require them to observe the *strategies* that opponents play. But in extensive form games like Centipede, playing the game against an opponent need not reveal the opponent's strategy; for instance, an agent who stops play at the initial node learns nothing at all about his opponent's intended play. In extensive form games, a basic advantage of dynamics based on experienced payoffs is that they only rely on information about opponents' intentions that agents can observe during play. As importantly for our results, agents modeled using best experienced payoff dynamics evaluate different strategies based on experiences with different samples of opponents. As we noted earlier, this property facilitates the establishment of cooperative behavior.

Osborne and Rubinstein's (1998) notion of $S(k)$ equilibrium corresponds to the rest points of the BEP dynamic under which agents test all strategies, subject each to k trials, and break ties via uniform randomization. While most of their analysis focuses on simultaneous move games, they show that in Centipede games, the probability with which player 1 stops immediately in any $S(1)$ equilibrium must vanish as the length of the game grows large. As we will soon see (Observation 3.1), this conclusion may fail if uniform tie-breaking is not assumed, with the backward induction state being an equilibrium. Nevertheless, more detailed analyses below will show that this equilibrium state is unstable under BEP dynamics.

Building on Osborne and Rubinstein (1998), Sethi (2000) introduces BEP dynamics under which all strategies are tested and ties are broken uniformly.[10] He shows that both dominant strategy equilibria and strict equilibria can be unstable under these dynamics, while dominated strategies can be played in stable equilibria. The latter fact is a basic component of our analysis of cooperative behavior. Here we introduce a general model of multiple-sample dynamics, and we develop techniques for analyzing these dynamics that not only allow us to make tight predictions in the Centipede game, but also provide analytical machinery for understanding other dynamic models.

Earlier efforts to explain cooperative behavior in Centipede and related games have

[10] Cárdenas et al. (2015) and Mantilla et al. (2017) use these dynamics to explain stable non-Nash behavior in public good games.

followed a different approach, applying equilibrium analyses to augmented versions of the game. The best known example of this approach is the work of Kreps et al. (1982). These authors modify the finitely repeated Prisoner's Dilemma by assuming that one player attaches some probability to his opponent having a fixed preference for cooperative play. They show that in all sequential equilibria of long enough versions of the resulting Bayesian game, both players act cooperatively for a large number of initial rounds.[11] To justify this approach, one must assume that the augmentation of the original game is commonly understood by the players, that the players act in accordance with a rather complicated equilibrium construction, and that the equilibrium knowledge assumptions required to justify sequential equilibrium apply. In contrast, our model makes no changes to the original game other than placing it in a population setting, and it is built upon the assumption that agents' choices are optimal given their experiences during play.

Further discussion of the literature, including models of behavior in games based on the logit choice rule, is offered in Section 9.

2. Best experienced payoff dynamics

In this section we consider evolutionary dynamics for populations of agents who are matched to play a two-player game. All of the definitions in this section are easily extended to single-population symmetric settings and to many-population settings.

2.1 Normal form games and population games

A two-player normal form game $G = \{(S^1, S^2), (U^1, U^2)\}$ is defined by pairs of strategy sets $S^p = \{1, \ldots, s^p\}$ and payoff matrices $U^p \in \mathbb{R}^{s^p \times s^q}$, $p, q \in \{1, 2\}$, $p \neq q$. U^p_{ij} represents the payoff obtained by player p when he plays strategy $i \in S^p$ against an opponent playing strategy $j \in S^q$. When considering extensive form games, our analysis focuses on the reduced normal form, whose strategies specify an agent's "plan of action" for the game, but not his choices at decision nodes that are ruled out by his own previous choices.

In our population model, members of two unit-mass populations are matched to play a two-player game. A *population state* ξ^p for population $p \in \{1, 2\}$ is an element of the simplex $\Xi^p = \{\xi^p \in \mathbb{R}^{s^p}_+ \colon \sum_{i \in S^p} \xi^p_i = 1\}$, with ξ^p_i representing the fraction of agents in population p

[11] A different augmentation is considered by Jehiel (2005), who assumes that agents bundle decision nodes from contiguous stages into analogy classes, and view the choices at all nodes in a class interchangeably. Alternatively, one can consider versions of Centipede in which the stakes of each move are small, and analyze these games using ε-equilibrium (Radner (1980)). But as Binmore (1998) observes, the existence of non-Nash ε-equilibrium depends on the relative sizes of the stakes and of ε, and the backward induction solution always persists as a Nash equilibrium, and hence as an ε-equilibrium.

using strategy $i \in S^p$. Thus ξ^p is formally equivalent to a mixed strategy for player p, and elements of the set $\Xi = \Xi^1 \times \Xi^2$ are formally equivalent to mixed strategy profiles. In a slight abuse of terminology, we also refer to elements of Ξ as population states.

2.2 Revision protocols and evolutionary dynamics

To define evolutionary game dynamics, we follow the now standard approach of specifying microfoundations in terms of revision protocols.[12] We suppose that at all times $t \in [0, \infty)$, each agent has a strategy he uses when matched to play game G. The distributions of these strategies are described by the population state $\xi(t) = (\xi^1(t), \xi^2(t))$. Agents occasionally receive opportunities to switch strategies according to independent rate 1 Poisson processes. An agent who receives an opportunity considers switching to a new strategy, making his decision by applying a revision protocol.

Formally, a *revision protocol* for population p is a map $\sigma^p \colon \mathbb{R}^{s^p \times s^q} \times \Xi \to (\Xi^p)^{s^p}$, where the Ξ before the arrow represents the set of population states, and the Ξ^p after the arrow represents the set of mixed strategies for population p. For any own payoff matrix $U^p \in \mathbb{R}^{s^p \times s^q}$ and opposing population state $\xi^q \in \Xi^q$, a revision protocol returns a matrix $\sigma^p(U^p, \xi^q)$ of *conditional switch probabilities*, where $\sigma^p_{ij}(U^p, \xi^q)$ is the probability that an agent playing strategy $i \in S^p$ who receives a revision opportunity switches to strategy $j \in S^p$.

If population sizes are large but finite, then a game and a pair of revision protocols defines a Markov process on the set Ξ of population states. During each short interval of time, a large number of agents in each population receive opportunities to switch strategies, applying the protocols σ^p to decide which strategy to choose next. But since the fraction of each population receiving revision opportunities during the interval is small, so is the change in the state over the interval. Intuition from the law of large numbers then suggests that over this interval, and over concatenations of such intervals, the state should evolve in an almost deterministic fashion, following the trajectory defined by the expected motion of the process. This claim is made rigorous by Benaïm and Weibull (2003), who show that over finite time spans, the stochastic evolution of play is very likely to hew closely to the solution of a differential equation. This differential equation, called the *mean dynamic*, describes the expected motion of the populations from each state:

$$(1) \qquad \dot{\xi}^p_i = \sum_{j \in S^p} \xi^p_j \sigma^p_{ji}(U^p, \xi^q) - \xi^p_i \quad \text{for all } i \in S^p \text{ and } p \in \{1, 2\}.$$

Equation (1) is easy to interpret. Since revision opportunities are assigned to agents

[12] See Björnerstedt and Weibull (1996), Weibull (1995), and Sandholm (2010a,b, 2015).

randomly, there is an outflow from each strategy i proportional to its current level of use. To generate inflow into i, an agent playing some strategy j must receive a revision opportunity, and applying his revision protocol must lead him to play strategy i.

2.3 Best experienced payoff protocols and dynamics

We now introduce the class of revision protocols and dynamics that we study in this paper. A best experienced payoff protocol is defined by a triple (τ, κ, β) consisting of a *test set rule* $\tau = (\tau_i^p)_{i \in S^p}^{p \in \{1,2\}}$, a *number of trials* κ, and a *tie-breaking rule* $\beta = (\beta_i^p)_{i \in S^p}^{p \in \{1,2\}}$. The triple (τ, κ, β) defines a revision protocol in the following way. When an agent currently using strategy $i \in S^p$ receives an opportunity to switch strategies, he draws a set of strategies $R^p \subseteq S^p$ to test according to the distribution τ_i^p on the power set of S^p. He then plays each strategy in R^p in κ random matches against members of the opposing population. He thus engages in $\#R^p \times \kappa$ random matches in total, facing distinct sets of opponents when testing different strategies. The agent then selects the strategy in R^p that earned him the highest total payoff, breaking ties according to rule β^p.

To proceed more formally, let $\mathcal{P}(S^p)$ denote the power set of S^p. A *test-set distribution* τ_i^p used by a strategy $i \in S^p$ player is a probability distribution on $\mathcal{P}(S^p)$ that places all of its mass on sets $R^p \subseteq S^p$ that include strategy i and at least one other strategy. One basic instance is *test-all*, defined by

$$\tau_i^p(S^p) = 1. \tag{2}$$

A less demanding example, *test-two*, has the agent test his current strategy and another strategy chosen at random:

$$\tau_i^p(\{i, j\}) = \frac{1}{s^p - 1} \text{ for all } j \in S^p \setminus \{i\}. \tag{3}$$

In games like Centipede, in which we will number strategies according to when they stop play (see Section 3), another natural example is *test-adjacent*, under which an agent tests his current strategy and an adjacent alternative, chosen at random when more than one is available:

$$\begin{aligned} &\tau_i^p(\{i, i-1\}) = \tau_i^p(\{i, i+1\}) = \tfrac{1}{2} \text{ for } i \in S^p \setminus \{1, s^p\}, \\ &\tau_1^p(\{1, 2\}) = 1, \ \tau_{s^p}^p(\{s^p, s^p - 1\}) = 1. \end{aligned} \tag{4}$$

In what follows, we will denote the test-all, test-two, and test-adjacent rules by τ^{all}, τ^{two}, and τ^{adj}, respectively.

The *tie-breaking rule for strategy $i \in S^p$*, denoted β^p_i, is a function that for each vector π^p of realized payoffs and each set of tested strategies $R^p \subseteq S^p$ specifies the probability $\beta^p_{ij}(\pi^p, R^p)$ of playing strategy $j \in S^p$.[13] Since agents are payoff maximizers, $\beta^p_{ij}(\pi^p, R^p)$ may only be positive if $j \in \operatorname{argmax}_{k \in R^p} \pi^p_k$. If there is a unique optimal strategy in R^p, it is chosen with probability one; in general, $\beta^p_{i\cdot}(\pi^p, R^p)$ is a probability distribution on S^p whose support is contained in $\operatorname{argmax}_{k \in R^p} \pi^p_k$. In normal form games, tie-breaking rules only matter in nongeneric cases. But in extensive form games, it is common for different strategies to earn the same payoffs, giving tie-breaking rules greater importance.

In Centipede, a conservative tie-breaking rule is *min-if-tie*, defined by

$$\beta^p_{ij}(\pi^p, R^p) = 1 \ \text{ if } j = \min\left[\operatorname*{argmax}_{k \in R^p} \pi^p_k\right]. \tag{5}$$

The tie-breaking rules we find most natural are *stick-if-tie* rules, which always select the agent's current strategy if it is among the optimal tested strategies:

$$\beta^p_{ii}(\pi^p, R^p) = 1 \ \text{ whenever } i \in \operatorname*{argmax}_{k \in R^p} \pi^p_k. \tag{6}$$

Condition (6) completely determines β_i when only two strategies are tested at a time, as under test-set rules (3) and (4). One full specification for games with many strategies uses the smallest optimal strategy when the agent's current strategy is not optimal:

$$\beta^p_{ij}(\pi^p, R^p) = 1 \ \text{ if } i = j \in \operatorname*{argmax}_{k \in R^p} \pi^p_k, \text{ or if } i \notin \operatorname*{argmax}_{k \in R^p} \pi^p_k \text{ and } j = \min\left[\operatorname*{argmax}_{k \in R^p} \pi^p_k\right]. \tag{7}$$

For purposes of comparison, we also define the *uniform-if-tie* rule:

$$\beta^p_{ij}(\pi^p, R^p) = \frac{1}{\#(\operatorname{argmax}_{k \in R^p} \pi^p_k)} \quad \text{if } j \in \operatorname*{argmax}_{k \in R^p} \pi^p_k. \tag{8}$$

In what follows, we denote the min-if-tie rule, the stick-if-tie rule (7), and uniform-if-tie rule by $\beta^{\min}$, β^{stick}, and β^{unif}, respectively.

At last, given a collection (τ, κ, β), we define the corresponding *best experienced payoff protocol* as follows:

[13] For now the notation π^p is a placeholder; it will be defined in equation (9b). The values assigned to components of π^p corresponding to strategies outside of R^p are irrelevant, and $\beta^p_{ij}(\pi^p, R^p) = 0$ whenever $j \notin R^p$.

$$\sigma^p_{ij}(U^p,\xi^q) = \sum_{R^p\subseteq S^p} \tau^p_i(R^p)\left[\sum_r \left(\prod_{k\in R^p}\prod_{m=1}^{\kappa} \xi^q_{r_{km}}\right)\beta^p_{ij}(\pi^p(r),R^p)\right], \tag{9a}$$

$$\text{where } \pi^p_k(r) = \sum_{m=1}^{\kappa} U^p_{kr_{km}} \text{ for all } k\in R^p, \tag{9b}$$

where $U^p_{k\ell}$ denotes the payoff of a population p agent who plays k against an opponent playing ℓ, and where the interior sum in (9a) is taken over all lists $r\colon R^p\times\{1,\dots,\kappa\}\to S^q$ of the opponents' strategies during testing. One can verify that (9) is the formal expression for the procedure described at the start of this section. That every test of a strategy occurs against an independently drawn opponent is captured by the products appearing in the parentheses in expression (9a). As we noted in the introduction, this feature of the revision protocol plays a basic role in our analysis.

Inserting (9) into the mean dynamic equation (1) yields the *best experienced payoff dynamic* defined by τ, κ, and β, called the BEP(τ,κ,β) *dynamic* for short:

$$\dot\xi^p_i = \sum_{j\in S^p}\xi^p_j\left(\sum_{R^p\subseteq S^p}\tau^p_j(R^p)\left[\sum_r\left(\prod_{k\in R^p}\prod_{m=1}^{\kappa}\xi^q_{r_{km}}\right)\beta^p_{ji}(\pi^p(r),R^p)\right]\right) - \xi^p_i, \tag{B}$$

with $\pi^p(r)$ defined in (9b).

Outside of monomorphic (i.e. pure) cases, the rest points of the dynamic (B) should not be understood as equilibria in the traditional game-theoretic sense. Rather, they represent situations in which agents perpetually switch among strategies, but with the *expected* change in the use of each strategy equalling zero.[14] At states that are locally stable under the dynamic (B), fluctuations in any direction are generally undone by the action of (B) itself. Contrariwise, fluctuations away from unstable equilibria are reinforced, so we should not expect such states to be observed.

When each tested strategy is played exactly once, best experienced payoff dynamics only depend on ordinal properties of payoffs, and the formulas for specific instances of the dynamics are relatively simple. Using $\mathbf{1}[\cdot]$ to denote the indicator function, the BEP$(\tau^{\text{all}},1,\beta^{\min})$ dynamic is expressed as

$$\dot\xi^p_i = \sum_{r\colon S^p\to S^q}\left(\prod_{\ell\in S^p}\xi^q_{r_\ell}\right)\mathbf{1}\Big[i=\min\Big(\operatorname{argmax}_k U^p_{kr_k}\Big)\Big] - \xi^p_i. \tag{10}$$

Here the fact that an agent's choice probabilities do not depend on his current strategy

[14] Thus in the finite-population version of the model, variations in the use of each strategy would be observed. For a formal analysis, see Sandholm (2003).

lead to a particularly compact expression.

For its part, the BEP($\tau^{\text{two}}, 1, \beta^{\text{stick}}$) dynamic is written as

$$\dot{\xi}_i^p = \frac{1}{s^p - 1} \sum_{h \neq i} \sum_{(k,\ell) \in S^q \times S^q} \xi_k^q \, \xi_\ell^q \left(\xi_i^p \, \mathbf{1}[U_{ik}^p \geq U_{h\ell}^p] + \xi_h^p \, \mathbf{1}[U_{ik}^p > U_{h\ell}^p] \right) - \xi_i^p. \tag{11}$$

The first term in parentheses captures the case in which the revising agent is a strategy i player who continues to play strategy i, while the second term captures the case in which the revising agent is a strategy $h \neq i$ player who switches to strategy i. The stick-if-tie rule requires that different inequalities be applied in these two cases.[15]

We conclude this section with two comments. First, the rest points of the BEP($\tau^{\text{all}}, \kappa, \beta^{\text{unif}}$) dynamic are the $S(k)$ equilibria of Osborne and Rubinstein (1998), and the dynamic itself is the one studied by Sethi (2000). Second, we have the following immediate observation about pure and strict Nash equilibria under BEP dynamics.

Observation 2.1. *Under any BEP(τ, κ, β) dynamic for which β is a stick-if-tie rule* (6), *all pure Nash equilibria are rest points. Strict equilibria are rest points under all BEP dynamics.*

3. Centipede

We now focus our attention on the Centipede game of Rosenthal (1981). In doing so it will be useful to make some changes in notation. We write population states as $\xi = (x, y)$, and the set of population states as $\Xi = X \times Y$, where X and Y are the simplices in $\mathbb{R}^{s^1}$ and $\mathbb{R}^{s^2}$. Also, we write the payoff matrices for players 1 and 2 as $A \in \mathbb{R}^{s^1 \times s^2}$ and $B \in \mathbb{R}^{s^1 \times s^2}$, so that the players' payoff profile at strategy profile $(i, j) \in S^1 \times S^2$ is (A_{ij}, B_{ij}).

3.1 The Centipede game

Centipede is a two-player extensive form game with $d \geq 2$ decision nodes (Figure 1). Each node presents two actions, 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 stops ends the game. A player who continues suffers a cost of 1 but benefits his opponent 3, and sends the game to the next decision node if one exists.

In a Centipede game of length d, players 1 and 2 have $d^1 = \lfloor \frac{d+1}{2} \rfloor$ and $d^2 = \lfloor \frac{d}{2} \rfloor$ decision nodes, respectively. Thus player p has $s^p = d^p + 1$ strategies, where strategy $i < s^p$ is the

[15]The formula for the BEP($\tau^{\text{adj}}, 1, \beta^{\text{stick}}$) dynamic is quite similar to (11), but with the initial factor of $1/(s^p - 1)$ being replaced by $\frac{1}{2}$, with the initial sum being over $\{i-1, i+1\} \cap S^p$, and with the instances of ξ_1^p and $\xi_{s^p}^p$ appearing before the indicator functions being replaced by $2\xi_1^p$ and $2\xi_{s^p}^p$, respectively.

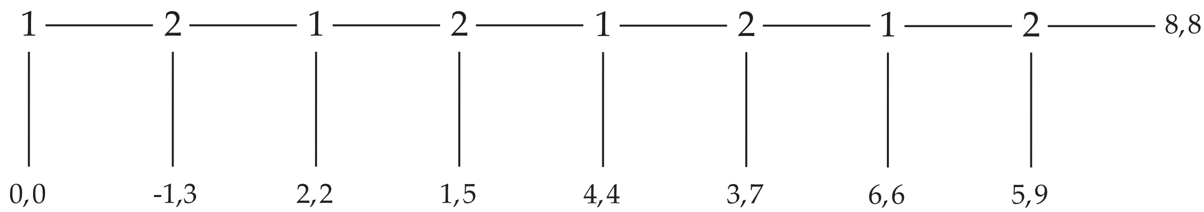


Figure 1: The Centipede game of length $d = 8$.

plan to continue at his first $i-1$ decision nodes and to stop at his ith decision node, and strategy s^p is the plan to continue at all d^p of his decision nodes. Of course, the portion of a player's plan that is actually carried out depends on the plan of his opponent. The payoff matrices (A, B) of Centipede's reduced normal form can be expressed concisely as

$$(12) \qquad (A_{ij}, B_{ij}) = \begin{cases} (2i-2, 2i-2) & \text{if } i \leq j, \\ (2j-3, 2j+1) & \text{if } j < i. \end{cases}$$

It will sometimes be convenient to number strategies starting from the end of the game. To do so, we write $[k] \equiv s^p - k$ for $k \in \{0, \ldots, d^p\}$, so that $[0]$ denotes continuing at all nodes, and $[k]$ with $k \geq 1$ denotes stopping at player p's kth-to-last node.

We noted above that best experienced payoff dynamics with $\kappa = 1$ only depend on ordinal properties of payoffs. In this case, what matters in (12) is that a player is better off continuing at a given decision node if and only if his opponent will continue at the subsequent decision node. If the cost of continuing is 1, this property holds as long as the benefit obtained when one's opponent continues exceeds 2. This ordering of payoffs also holds for typical specifications in which total payoffs grow exponentially over time.[16] When there are multiple trials of each tested strategy, then cardinal properties of payoffs matter; in this case, Rosenthal's (1981) specification (12) keeps the potential benefits from continuing relatively modest.

The backward induction solution to Centipede has both players stop at each of their decision nodes. We will thus call the population state $\xi^\dagger = (x^\dagger, y^\dagger) \in \Xi$ with $x_1^\dagger = y_1^\dagger = 1$ the (*reduced*) *backward induction state*. It is well known that all Nash equilibria of Centipede have player 1 stop at his initial node. This makes player 2 indifferent among all of her strategies, so Nash equilibrium requires that she choose a mixed strategy that makes stopping immediately optimal for player 1.

Of course, these predictions require assumptions about what the players know. In

[16]Suppose initial total payoffs are positive, that total payoffs are multiplied by $b > 1$ in each period, and that the player who was not the last to continue receives fraction $p \in (\frac{1}{2}, 1)$ of this total. Then continuing is not dominant when $\frac{p}{1-p} > b$, and the ordinal structure described above is obtained when $\frac{p}{1-p} < b^3$.

the traditional justification of Nash equilibrium, players are assumed to correctly anticipate opponents' play. Likewise, traditional justifications of the backward induction solution require agents to maintain common belief in rational future play, even if behavior contradicting this belief has been observed in the past.[17]

3.2 Best experienced payoff dynamics for the Centipede game

To obtain explicit formulas for best experienced payoff dynamics in the Centipede game, we substitute the payoff functions (12) into the general formulas from Section 2 and simplify the results. Here we present the two examples from the end of Section 2.3. Explicit formulas for the remaining BEP($\tau, 1, \beta$) dynamics with $\tau \in \{\tau^{\text{all}}, \tau^{\text{two}}, \tau^{\text{adj}}\}$ and $\beta \in \{\beta^{\min}, \beta^{\text{stick}}, \beta^{\text{unif}}\}$ are presented in Online Appendix VI.

The BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic for Centipede is given by[18]

$$\text{(13a)} \qquad \dot{x}_i = \left(\sum_{k=i}^{s^2} y_k\right)\left(\sum_{m=1}^{i} y_m\right)^{s^1-i} + \sum_{k=2}^{i-1} y_k \left(\sum_{\ell=1}^{k-1} y_\ell\right)^{i-k} \left(\sum_{m=1}^{k} y_m\right)^{s^1-i} - x_i,$$

$$\text{(13b)} \qquad \dot{y}_j = \begin{cases} \left(\sum_{k=2}^{s^1} x_k\right)(x_1 + x_2)^{s^2-1} + (x_1)^{s^2} - y_1 & \text{if } j = 1, \\ \left(\sum_{k=j+1}^{s^1} x_k\right)\left(\sum_{m=1}^{j+1} x_m\right)^{s^2-j} + \sum_{k=2}^{j} x_k \left(\sum_{\ell=1}^{k-1} x_\ell\right)^{j-k+1} \left(\sum_{m=1}^{k} x_m\right)^{s^2-j} - y_j & \text{otherwise.} \end{cases}$$

Under test-all with min-if-tie, the choice made by a revising agent does not depend on his original strategy. The first two terms of (13a) describe the two types of matchings that lead a revising agent in the role of player 1 to choose strategy i. First, it could be that when the agent tests i, his opponent plays i or higher (so that the agent is the one to stop the game), and that when the agent tests higher strategies, his opponents play strategies i or lower. In this case, only strategy i yields the agent his highest payoff. Second, it could be that when the agent tests i, his opponent plays strategy $k < i$; when he tests strategies between k and $i-1$, his opponents play strategies less than k; and when he tests strategies above i, his opponents play strategies less than or equal to k. In this case, strategy i is the lowest strategy that achieves the optimal payoff, and so is chosen by the revising agent under the min-if-tie rule. Similar logic, and accounting for the fact that player 2's jth node is followed by player 1's $(j+1)$st node, leads to equation (13b).

[17] See Binmore (1987), Reny (1992), Stalnaker (1996), Ben-Porath (1997), Dekel and Gul (1997), Halpern (2001), and Perea (2014).

[18] We follow the convention that a sum whose lower limit exceeds its upper limit evaluates to 0.

The BEP($\tau^{\text{two}}, 1, \beta^{\text{stick}}$) dynamic for Centipede is written as

$$\text{(14a)} \qquad \dot{x}_i = \frac{1}{s^1 - 1} \sum_{h \neq i} \left[\left(\sum_{k=h+1|i}^{s^2} \sum_{\ell=1}^{s^2|i} y_k y_\ell + \sum_{k=2}^{h|i-1} \sum_{\ell=1}^{k-1} y_k y_\ell \right) (x_i + x_h) + \sum_{k=1}^{h-1|i-1} (y_k)^2 x_i \right] - x_i,$$

$$\text{(14b)} \qquad \dot{y}_j = \frac{1}{s^2 - 1} \sum_{h \neq j} \left[\left(\sum_{k=h+2|j+1}^{s^1} \sum_{\ell=1}^{s^1|j+1} x_k x_\ell + \sum_{k=2}^{h+1|j} \sum_{\ell=1}^{k-1} x_k x_\ell \right) \left(y_j + y_h \right) + \sum_{k=1}^{h|j} (x_k)^2 y_j \right] - y_j.$$

In the summations in (14a), the notation $L_-|L_+$ should be read as "if $h < i$ use L_- as the limit; if $h > i$ use L_+ as the limit"; likewise for those in (14b), but with h now compared to j.[19] Each term in the brackets in (14a) represents a comparison between the performances of strategies i and h. Terms that include $x_i + x_h$ represent cases in which the realized payoff to i is larger than that to h, so that it does not matter whether the revising agent is an i player who tests h or vice versa. The terms with x_i alone represent cases of payoff ties, which arise when i and h are both played against opponents choosing the same strategy $k < i \wedge h$ that stops before either i or h; in this case, the stick-if-tie rule says that the agent should play i only if he had previously been doing so.[20]

We conclude this section with a simple observation about the backward induction solution of Centipede under best experienced payoff dynamics.

Observation 3.1. *Under any BEP(τ, κ, β) dynamic for which β is a stick-if-tie rule* (6) *or the min-if-tie rule* (5)*, the backward induction state $\xi^\dagger$ is a rest point.*

Osborne and Rubinstein (1998) show that if all strategies are tested once and ties are broken uniformly, then in a long Centipede game, stationarity requires that play is almost never is stopped at the initial node. Observation 3.1 shows that this conclusion depends on the assumption that ties are broken uniformly. If instead ties are broken in favor of an agents' current strategy or the lowest-numbered strategy, then the backward induction state is a rest point. Even so, more refined analyses in Sections 5 and 6 will explain why the backward induction state is not a compelling prediction of play even under these tie-breaking rules.

[19] We could replace | by $\wedge$ (the min operator) in all cases other than the two that include s^2 or s^1.

[20] To understand the functional form of (14a), consider a revising agent with test set $\{i, h\}$. If $i > h$, the initial double sum represents matchings in which i is played against an opponent choosing a strategy above h; if $i < h$, it represents matchings in which i is played against an opponent choosing strategy i or higher, while h is played against an opponent choosing strategy i or lower. The second double sum represents matchings in which i is played against an opponent choosing strategy $h \wedge (i-1)$ or lower, while h is played against an opponent choosing a still lower strategy. In all of these cases, i yields a larger payoff than h, so the revising agent selects i regardless of what he was initially playing. The final sum represents matchings in which i and h are both played against opponents who choose the same strategy $k < h \wedge i$, leading the agent to stick with his original strategy i.

4. Exact solutions of systems of polynomial equations

An important first step in analyzing a system of differential equations is to identify its rest points. From this point of view, a key feature of best experienced payoff dynamics is that they are defined by systems of polynomial equations with rational coefficients. In what follows, we explain the algebraic tools that we use to determine the exact values of the components of these rest points. The contents of this section are not needed to understand the presentation of our results, which begins in Section 5. Some readers may prefer to return to this section after reading those that follow.

4.1 Gröbner bases

Let $\mathbb{Q}[z_1, \ldots, z_n]$, or $\mathbb{Q}[z]$ for short, denote the collection (more formally, the *ring*) of polynomials in the variables $z_1, \ldots, z_n$ with rational coefficients. Let $F = \{f_1, \ldots f_m\} \subset \mathbb{Q}[z]$ be a set of such polynomials. Let Z be a subset of $\mathbb{R}^n$, and consider the problem of finding the set of points $z^* \in Z$ that are zeroes of all polynomials in F.

To do so, it is convenient to first consider finding all zeroes in $\mathbb{C}^n$ of the polynomials in F. In this case, the set of interest,

$$\mathbf{V}(f_1, \ldots, f_m) = \{z^* \in \mathbb{C}^n : f_j(z^*) = 0 \text{ for all } 1 \leq j \leq m\} \tag{15}$$

is called the *variety* (or *algebraic set*) *generated by* $f_1, \ldots, f_m$. To characterize (15), it is useful to introduce the *ideal generated by* $f_1, \ldots, f_m$:

$$\langle f_1, \ldots, f_m \rangle = \left\{ \sum_{j=1}^{m} h_j f_j : h_j \in \mathbb{C}[z] \text{ for all } 1 \leq j \leq m \right\}. \tag{16}$$

Thus the ideal (16) is the set of linear combinations of the polynomials $f_1, \ldots, f_m$, where the coefficients on each are themselves polynomials in $\mathbb{C}[z]$. It is easy to verify that any other collection of polynomials in $\mathbb{C}[z]$ whose linear combinations generate the ideal (16)—that is, any other *basis* for the ideal—also generates the variety (15).

For our purposes, the most useful basis for the ideal (16) is the *reduced lex-order Gröbner basis*, which we denote by $G \subset \mathbb{Q}[z]$. This basis, which contains no superfluous polynomials and is uniquely determined by its ideal and the ordering of the variables, has this convenient property: it consists of polynomials in z_n only, polynomials in z_n and z_{n-1} only, polynomials in z_n, z_{n-1}, and z_{n-2} only, and so forth. Thus if the variety (15) has cardinality $|\mathbf{V}| < \infty$, then it can be computed sequentially by solving univariate polynomials and substituting backward.[21]

In many cases, including all that arise in this paper, the basis G is of the simple form

$$G = \{g_n(z_n),\, z_{n-1} - g_{n-1}(z_n),\, \ldots\, ,\, z_1 - g_1(z_n)\} \tag{17}$$

for some univariate polynomials $g_n, \ldots, g_1$, where g_n has degree $\deg(g_n) = |\mathbf{V}|$ and where $\deg(g_k) < |\mathbf{V}|$ for $k < n$.[22] In such cases, one computes the variety (15) by finding the $|\mathbf{V}|$ complex roots of g_n, and then substituting each into the other $n-1$ polynomials to obtain the $|\mathbf{V}|$ elements of (15).[23]

4.2 Algebraic numbers

The first step in finding the zeros of the polynomials in $G \subset \mathbb{Q}[z]$ is to find the roots of the univariate polynomial g_n. There are well-known limits to what can be accomplished here: Abel's theorem states that there is no solution in radicals to general univariate polynomial equations of degree five or higher. Nevertheless, tools from computational algebra allow us to represent such solutions exactly.

Let $\bar{\mathbb{Q}} \subset \mathbb{C}$ denote the set of *algebraic numbers*: the complex numbers that are roots of polynomials with rational coefficients. $\bar{\mathbb{Q}}$ is a subfield of $\mathbb{C}$, and this fact and the definition of algebraic numbers are summarized by saying that $\bar{\mathbb{Q}}$ is the *algebraic closure* of $\mathbb{Q}$.[24]

Every univariate polynomial $g \in \mathbb{Q}[x]$ can be factored as a product of *irreducible polynomials* in $\mathbb{Q}[x]$, which cannot themselves be further factored into products of nonconstant elements of $\mathbb{Q}[x]$.[25] If an irreducible polynomial $h \in \mathbb{Q}[x]$ is of degree k, it has k distinct roots $a_1, \ldots, a_k \in \bar{\mathbb{Q}}$. The multiple of h whose leading term has coefficient 1 is called the

[21] The notion of Gröbner bases and the basic algorithm for computing them are due to Buchberger (1965). Cox et al. (2015) provide an excellent current account of Gröbner basis algorithms, as well as a thorough introduction to the ideas summarized above.

[22] According to the *shape lemma*, a sufficient condition for the reduced lex-order basis to be of form (17) is that each point in (15) have a distinct z_n component and that (16) be a *radical ideal*, meaning that if it includes some integer power of a polynomial, then it includes the polynomial itself. See Becker et al. (1994) and Kubler et al. (2014).

[23] Although we are only interested in elements of the variety (15) that lie in the state space Ξ, the solution methods described above only work if (15) has a finite number of solutions in $\mathbb{C}^n$. For some specifications of the BEP dynamic, the latter property fails, but we are able to circumvent this problem by introducing additional variables and equations—see Section 6.2.

[24] Like $\mathbb{C}$, $\bar{\mathbb{Q}}$ is *algebraically closed*, in that every univariate polynomial with coefficients in $\bar{\mathbb{Q}}$ has a root in $\bar{\mathbb{Q}}$. It follows from this and the existence of lex-order Gröbner bases that when the variety (15) has a finite number of elements, the components of its elements are algebraic numbers.

[25] "Typical" polynomials in $\mathbb{Q}[x]$ are irreducible: for instance, the quadratic $ax^2 + bx + c$ with $a, b, c \in \mathbb{Q}$ is only reducible if $\sqrt{b^2 - 4ac} \in \mathbb{Q}$. By *Gauss's lemma*, polynomial factorization in $\mathbb{Q}[x]$ is effectively equivalent to polynomial factorization in $\mathbb{Z}[x]$. For an excellent presentation of polynomial factorization algorithms, see von zur Gathen and Gerhard (2013, Ch. 14–16).

minimal polynomial of these roots. One often works instead with the multiple of h that is *primitive* in $\mathbb{Z}[x]$, meaning that its coefficients are integers with greatest common divisor 1.

Each algebraic number is uniquely identified by its minimal polynomial h and a label that distinguishes the roots of h from one another. For instance, one can label each root $a_j \in \bar{\mathbb{Q}}$ with a numerical approximation that is sufficiently accurate to distinguish a_j from the other roots. In computer algebra systems, the algebraic numbers with minimal polynomial h are represented by pairs consisting of h and an integer in $\{1, \ldots, k\}$ which ranks the roots of h with respect to some ordering; for instance, the lowest integers are commonly assigned to the real roots of h in increasing order. Just as the symbol $\sqrt{2}$ is a label for the positive solution to $x^2 - 2 = 0$, the approach above provides labels for every algebraic number.[26]

If the Gröbner basis G is of form (17), then we need only look for the roots of the irreducible factors h of the polynomial g_n, which are the possible values of $x_n \in \bar{\mathbb{Q}}$; then substitution into the univariate polynomials $g_{n-1}, \ldots, g_1$ determines the corresponding values of the other variables. The fact that these latter values are generated from a fixed algebraic number allows us to work in subfields of $\bar{\mathbb{Q}}$ in which arithmetic operations are easy to perform. If the minimal polynomial h of $\alpha \in \bar{\mathbb{Q}}$ has degree $\deg(h)$, then for any polynomial f, one can find a polynomial f^* of degree $\deg(f^*) < \deg(h)$ such that $f(\alpha) = f^*(\alpha)$. It follows that the values of $g_{n-1}(\alpha), \ldots, g_1(\alpha)$ are all elements of

$$\mathbb{Q}(\alpha) = \left\{ \sum_{k=0}^{\deg(h)-1} a_k \, \alpha^k : \; a_0, \ldots, a_{\deg(h)-1} \in \mathbb{Q} \right\} \subset \bar{\mathbb{Q}},$$

called the *field extension* of $\mathbb{Q}$ generated by α. Straightforward arguments show that the representation of elements of $\mathbb{Q}(\alpha)$ by sequences of coefficients $(a_0, \ldots, a_d)$ makes addition and multiplication in $\mathbb{Q}(\alpha)$ simple to perform. For further details on algebraic numbers and field extensions, we refer the reader to Dummit and Foote (2004, Chapter 13) and Cohen (1993, Chapter 4).

4.3 Examples

To illustrate the techniques above, we use them to compute the rest points of the $\mathrm{BEP}(\tau, 1, \beta^{\min})$ dynamic in the Centipede game with $d = 3$ nodes, where τ is either τ^{two} (equation (3)) or τ^{adj} (equation (4)). Although $s^1 = d^1 + 1 = 3$ and $s^2 = d^2 + 1 = 2$, we need

[26] There are exact methods based on classic theorems of Sturm and Vincent for isolating the real roots of a polynomial with rational coefficients; see McNamee (2007, Ch. 2 and 3) and Akritas (2010).

only write down the laws of motion for three components, say x_1, x_2, and y_1, since the remaining components are then given by $x_3 = 1 - x_1 - x_2$ and $y_2 = 1 - y_1$.[27]

Example 4.1. The BEP($\tau^{\text{two}}, 1, \beta^{\min}$) dynamic in Centipede of length $d = 3$ is

$$\begin{aligned} \dot{x}_1 &= \tfrac{1}{2}\left(y_1(x_1 + x_2) + y_1(x_1 + x_3)\right) - x_1, \\ \dot{x}_2 &= \tfrac{1}{2}\Big(y_2(x_1 + x_2) + (y_2 + (y_1)^2)(x_2 + x_3)\Big) - x_2, \\ \dot{y}_1 &= ((x_2 + x_3)(x_1 + x_2) + (x_1)^2)(y_1 + y_2) - y_1 \end{aligned} \tag{18}$$

(see Online Appendix VI). To find the rest points of this system, we substitute $x_3 = 1 - x_1 - x_2$ and $y_2 = 1 - y_1$ in the right-hand sides of (18) to obtain a system of three equations and three unknowns. We then compute a Gröbner basis of form (17) for the right-hand sides of (18):

$$\begin{aligned} \Big\{3(y_1)^4 - 8(y_1)^3 + 13(y_1)^2 - 12y_1 + 4,\ 4x_2 + 3(y_1)^3 - 5(y_1)^2 + 6y_1 - 4, \\ 8x_1 - 3(y_1)^3 + 2(y_1)^2 - 9y_1 + 2\Big\}. \end{aligned} \tag{19}$$

The initial quartic in (19) has roots 1, $\frac{2}{3}$, and $(1 \pm \sqrt{7}\,\mathrm{i})/2$. Of course, only the first two roots could be components of states in Ξ. Substituting $y_1 = 1$ in the remaining polynomials in (19) and equating them to 0 yields $x_1 = 1$ and $x_2 = 0$, which with the simplex constraints gives us the backward induction state $\xi^\dagger$. Substituting $y_1 = \frac{2}{3}$ instead yields the interior state $\xi^* = (x^*, y^*) = ((\frac{1}{2}, \frac{1}{3}, \frac{1}{6}), (\frac{2}{3}, \frac{1}{3}))$. This is the complete set of rest points of the dynamic (18). ◆

Example 4.2. The BEP($\tau^{\text{adj}}, 1, \beta^{\min}$) dynamic in Centipede of length $d = 3$ is

$$\begin{aligned} \dot{x}_1 &= y_1(x_1 + \tfrac{1}{2}x_2) - x_1, \\ \dot{x}_2 &= \Big(y_2(x_1 + \tfrac{1}{2}x_2) + (y_2 + (y_1)^2)(\tfrac{1}{2}x_2 + x_3)\Big) - x_2, \\ \dot{y}_1 &= ((x_2 + x_3)(x_1 + x_2) + (x_1)^2)(y_1 + y_2) - y_1 \end{aligned} \tag{20}$$

We again compute a Gröbner basis of form (17):

$$\begin{aligned} \Big\{2(y_1)^6 - 8(y_1)^5 + 19(y_1)^4 - 29(y_1)^3 + 28(y_1)^2 - 16y_1 + 4,\ 2x_2 + 2(y_1)^5 - 4(y_1)^4 \\ + 7(y_1)^3 - 7(y_1)^2 + 4y_1 - 2,\ 2x_1 - 2(y_1)^4 + 4(y_1)^3 - 7(y_1)^2 + 5y_1 - 2\Big\}. \end{aligned} \tag{21}$$

[27]The Gröbner basis algorithm sometimes runs faster if all components are retained and the left-hand sides of the constraints $\sum_{i=1}^{s^1} x_i - 1 = 0$ and $\sum_{j=1}^{s^2} y_j - 1 = 0$ are included in the initial set of polynomials. Our *Mathematica* notebook includes both implementations.

The initial polynomial in (21) has root 1, which again generates the backward induction state $\xi^\dagger$. Dividing this polynomial by $y_1 - 1$ yields

$$2(y_1)^5 - 6(y_1)^4 + 13(y_1)^3 - 16(y_1)^2 + 12y_1 - 4,$$

which is an irreducible quintic. Using the algorithms mentioned above, one can show that this quintic has one real root, which we designate by $y_1^* = \mathrm{Root}[2\alpha^5 - 6\alpha^4 + 13\alpha^3 - 16\alpha^2 + 12\alpha - 4, 1] \approx .7295$, and four complex roots. Substituting y_1^* into the remaining expressions in (21) and using the simplex constraints, we obtain an exact expression for the interior rest point ξ^*, whose components are elements of the field extension $\mathbb{Q}(y_1^*)$; their approximate values are $\xi^* = (x^*, y^*) \approx ((.5456, .4047, .0497), (.7295, .2705))$. ◆

5. Analysis of test-all, min-if-tie dynamics

In this section and Section 6.1, we analyze BEP($\tau, 1, \beta^{\min}$) dynamics in Centipede games, focusing on the tie-breaking rule that is most favorable to backward induction. Before proceeding, we review some standard definitions and results from dynamical systems theory, and follow this with a simple example.

Consider a C^1 differential equation $\dot{\xi} = V(\xi)$ defined on Ξ whose forward solutions $\{x(t)\}_{t\geq 0}$ do not leave Ξ. State ξ^* is a *rest point* if $V(\xi^*) = 0$, so that the unique solution starting from ξ^* is stationary. Rest point ξ^* is *Lyapunov stable* if for every neighborhood $O \subset \Xi$ of ξ^*, there exists a neighborhood $O' \subset \Xi$ of ξ^* such that every forward solution that starts in O' is contained in O. If ξ^* is not Lyapunov stable it is *unstable*, and it is *repelling* if there is a neighborhood $O \subset \Xi$ of ξ^* such that solutions from all initial conditions in $O \smallsetminus \{\xi^*\}$ leave O.

Rest point ξ^* is *attracting* if there is a neighborhood $O \subset \Xi$ of ξ^* such that all solutions that start in O converge to ξ^*. A state that is Lyapunov stable and attracting is *asymptotically stable*. In this case, the maximal (relatively) open set of states from which solutions converge to ξ^* is called the *basin* of ξ^*. If the basin of ξ^* contains int(Ξ), we call ξ^* *almost globally asymptotically stable*; if it is Ξ itself, we call ξ^* *globally asymptotically stable*.

The C^1 function $L\colon O \to \mathbb{R}_+$ is a *strict Lyapunov function* for rest point $\xi^* \in O$ if $L^{-1}(0) = \{\xi^*\}$, and if its time derivative $\dot{L}(\xi) \equiv \nabla L(\xi)' V(\xi)$ is negative on $O \smallsetminus \{\xi^*\}$. Standard results imply that if such a function exists, then ξ^* is asymptotically stable.[28] If L is a strict Lyapunov function for ξ^* with domain $O = \Xi \smallsetminus \{\xi^\dagger\}$ and $\xi^\dagger$ is repelling, then ξ^* is almost globally asymptotically stable; if the domain is Ξ, then ξ^* is globally asymptotically stable.

[28]See, e.g., Sandholm (2010b, Appendix 7.B).

Example 5.1. As a preliminary, we consider BEP($\tau, 1, \beta^{\min}$) dynamics for the Centipede game of length 2. Since each player has two strategies, all test-set rules τ have revising agents test both of them. Focusing on the fractions of agents choosing to continue, we can express the dynamics as

$$\begin{aligned}\dot{x}_2 &= y_2 - x_2,\\ \dot{y}_2 &= x_2 x_1 - y_2.\end{aligned} \tag{22}$$

By way of interpretation, a revising agent in population 1 chooses to continue if his opponent when he tests continue also continues. A revising agent in population 2 chooses to continue if her opponent continues when she tests continue, and her opponent stops when she tests stop.[29]

Writing $1 - x_2$ for x_1 in (22) and then solving for the zeroes, we find that the unique rest point of (22) is the backward induction state: $x_2^\dagger = y_2^\dagger = 0$. Moreover, defining the function $L\colon [0,1]^2 \to \mathbb{R}_+$ by $L(x_2, y_2) = \frac{1}{2}((x_2)^2 + (y_2)^2)$, we see that $L^{-1}(0) = \{\xi^\dagger\}$ and that

$$\dot{L}(x_2, y_2) = x_2\dot{x}_2 + y_2\dot{y}_2 = x_2 y_2 - (x_2)^2 + y_2 x_2 - y_2(x_2)^2 - (y_2)^2 = -(x_2 - y_2)^2 - y_2(x_2)^2,$$

which is nonpositive on $[0,1]^2$ and equals zero only at the backward induction state. Since L is a strict Lyapunov function for $\xi^\dagger$ on Ξ, state $\xi^\dagger$ is globally asymptotically stable. ◆

In light of this example, our analyses of dynamics using the min-if-tie rule $\beta^{\min}$ will focus on Centipede games of lengths $d \geq 3$. In the remainder of this section we suppose that τ is the test-all rule τ^{all}. Section 6.1 considers the test-set rules τ^{two} and τ^{adj}, and Section 6.2 considers the tie-breaking rules β^{stick} and β^{unif}.

5.1 Rest points and local stability

5.1.1 Analytical results

As we know from Observation 3.1, the backward induction state $\xi^\dagger$ of the Centipede game is a rest point of the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic. Our first result shows that this rest point is always repelling.

Proposition 5.2. *In Centipede games of lengths $d \geq 3$, the backward induction state $\xi^\dagger$ is repelling under the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic.*

The proof of Proposition 5.2, which is presented in Appendix A, is based on a somewhat nonstandard linearization argument. While we are directly concerned with the behavior

[29] Compare the discussion after equation (13) and Example 5.3 below.

of the BEP dynamics on the state space Ξ, it is useful to view equation (B) as defining dynamics throughout the affine hull $\text{aff}(\Xi) = \{(x, y) \in \mathbb{R}^{s^1+s^2} \colon \sum_{i\in S^1} x_i = \sum_{j\in S^2} y_j = 1\}$, which is then invariant under (B). Vectors of motion through aff(Ξ) are elements of the tangent space $T\Xi = \{(z^1, z^2) \in \mathbb{R}^{s^1+s^2} \colon \sum_{i\in S^1} z^1_i = \sum_{j\in S^2} z^2_j = 0\}$. Note that $T\Xi$ is a subspace of $\mathbb{R}^{s^1+s^2}$, and that aff(Ξ) is obtained from $T\Xi$ via translation: $\text{aff}(\Xi) = T\Xi + \xi^\dagger$.

A standard linearization argument is enough to prove that $\xi^\dagger$ is unstable. Let the vector field $V \colon \text{aff}(\Xi) \to T\Xi$ be defined by the right-hand side of (B). To start the proof, we obtain an expression for the derivative matrix $DV(\xi^\dagger)$ that holds for any game length d. We then derive formulas for the d linearly independent eigenvectors of $DV(\xi^\dagger)$ in the subspace $T\Xi$ and for their corresponding eigenvalues. We find that $d-1$ of the eigenvalues are negative, and one is positive. The existence of the latter implies that $\xi^\dagger$ is unstable.

To prove that $\xi^\dagger$ is repelling, we show that the hyperplane through $\xi^\dagger$ defined by the span of the set of $d-1$ eigenvectors with negative eigenvalues supports the convex state space Ξ at state $\xi^\dagger$. Results from dynamical systems theory—specifically, the Hartman-Grobman and stable manifold theorems (Perko (2001, Sec. 2.7–2.8))—then imply that in some neighborhood $O \subset \text{aff}(\Xi)$ of $\xi^\dagger$, the set of initial conditions from which solutions converge to $\xi^\dagger$ is disjoint from $\Xi \setminus \{\xi^\dagger\}$, and that solutions from the remaining initial conditions eventually move away from $\xi^\dagger$.

The following example provides intuition for the instability of the backward induction state; the logic is similar in longer games and for other specifications of the dynamics.

Example 5.3. In a Centipede game of length $d = 4$, writing out display (13) shows that the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic is described by

$$\begin{aligned} \dot{x}_1 &= (y_1)^2 - x_1, & \dot{y}_1 &= (x_2 + x_3)(x_1 + x_2)^2 + (x_1)^3 - y_1, \\ \dot{x}_2 &= (y_2 + y_3)(y_1 + y_2) - x_2, & \dot{y}_2 &= x_3 + x_2 x_1(x_1 + x_2) - y_2, \\ \dot{x}_3 &= y_3 + y_2 y_1 - x_3, & \dot{y}_3 &= x_2(x_1)^2 + x_3(x_1 + x_2) - y_3. \end{aligned} \tag{23}$$

The linearization of this system at $(x^\dagger, y^\dagger)$ has the positive eigenvalue 1 corresponding to eigenvector $(z^1, z^2) = ((-2, 1, 1), (-2, 1, 1))$ (see equation (31)). Thus at state $(x, y) = ((1 - 2\varepsilon, \varepsilon, \varepsilon), (1 - 2\varepsilon, \varepsilon, \varepsilon))$ with $\varepsilon > 0$ small, we have $(\dot{x}, \dot{y}) \approx ((-2\varepsilon, \varepsilon, \varepsilon), (-2\varepsilon, \varepsilon, \varepsilon))$.

To understand why the addition of agents in both populations playing the cooperative strategies 2 and 3 is self-reinforcing, we build on the discussion following equation (13). Consider, for instance, component y_3, which represents the fraction of agents in population 2 who continue at both decision nodes. The last expression in (23) says that a revising population 2 agent switches to strategy 3 if (i) when testing strategy 3 she meets an opponent playing strategy 2, and when testing strategies 1 and 2 she meets opponents playing strategy 1, or (ii) when testing strategy 3 she meets an opponent playing strategy 3, and when testing strategy 2 she meets an opponent playing strategy 1 or 2. These events have total probability $\varepsilon(1-\varepsilon)+\varepsilon(1-2\varepsilon)^2 \approx 2\varepsilon$. Since there are $y_3=\varepsilon$ agents currently playing strategy 3, outflow from this strategy occurs at rate ε. Combining the inflow and outflow terms shows that $\dot{y}_3 \approx 2\varepsilon-\varepsilon=\varepsilon$. Analogous arguments explain the changes in the values of the other components of the state. ◆

It may seem surprising that the play of a weakly dominated strategy—continuing by the last mover at the last decision node—is positively reinforced at an interior population state. This is possible because revising agents test each of their strategies against newly drawn opponents: as just described, a revising population 2 agent will choose to continue at both of her decision nodes if her opponent's strategy when she tests strategy 3 is more cooperative than her opponents' strategies when she tests her own less cooperative strategies.

If population 2 agents understand that they are playing Centipede, one might appeal to the principle of admissibility to argue that they should not choose to continue at their last decision node. One can capture this principle in our context by modeling play in the length d Centipede game using the dynamics we have defined for the length $d-1$ game.[30]

Since the backward induction state is unstable, we next look for attractors to which the dynamics may converge.

Proposition 5.4. *For Centipede games of lengths* $3 \leq d \leq 6$*, the BEP*$(\tau^{\mathrm{all}}, 1, \beta^{\min})$ *dynamic has exactly two rest points,* $\xi^\dagger$ *and* $\xi^* \in \mathrm{int}(\Xi)$*. The rest point* ξ^**, whose exact components are known, is asymptotically stable.*

Our proof that the dynamics considered in the proposition have exactly two rest points uses the Gröbner basis and algebraic number algorithms presented in Section 4. Exact solutions can only be obtained for games of length at most 6 because of the computational demands of computing the Gröbner bases. One indication of these demands is that when $d=6$, the univariate polynomial from the Gröbner basis is of degree 221.[31]

Table 1 reports the approximate values of the interior rest points ξ^*, referring to strategies using the last-to-first notation $[k]$ introduced in Section 3.1. Evidently, the masses on each strategy are nearly identical for games of lengths 4, 5, and 6, with nearly all of the weight in both populations being placed on continuing to the end, stopping at the last node, or stopping at the penultimate node.

[30]Because of the logical complications inherent in the foundations for iterated admissibility (Samuelson (1992), Brandenburger et al. (2008)), we hesitate to consider iterative truncation of our game dynamics.

[31]We will see that games of lengths 7 and 8 can be handled for some other choices of τ and β. See Table 7 for a summary.

	population p				population q			
	[3]	[2]	[1]	[0]	[3]	[2]	[1]	[0]
$d = 3$			.618034	.381966		.381966	.381966	.236068
$d = 4$		.113625	.501712	.384663		.337084	.419741	.243175
$d = 5$		.113493	.501849	.384658	.001462	.335672	.419706	.243160
$d = 6$	3.12×10^{-9}	.113493	.501849	.384658	.001462	.335672	.419706	.243160

Table 1: "Exact" interior rest points ξ^* of the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic. p denotes the owner of the penultimate decision node, q the owner of the last decision node.

$d = 3$	$-1 \pm .3820$	-1	
$d = 4$	$-1.1411 \pm .3277\,\mathrm{i}$	$-.8589 \pm .3277\,\mathrm{i}$	
$d = 5$	$-1.1355 \pm .3284\,\mathrm{i}$	$-.8645 \pm .3284\,\mathrm{i}$	-1
$d = 6$	$-1.1355 \pm .3284\,\mathrm{i}$	$-.8645 \pm .3284\,\mathrm{i}$	$-1 \pm 9.74 \times 10^{-5}\,\mathrm{i}$

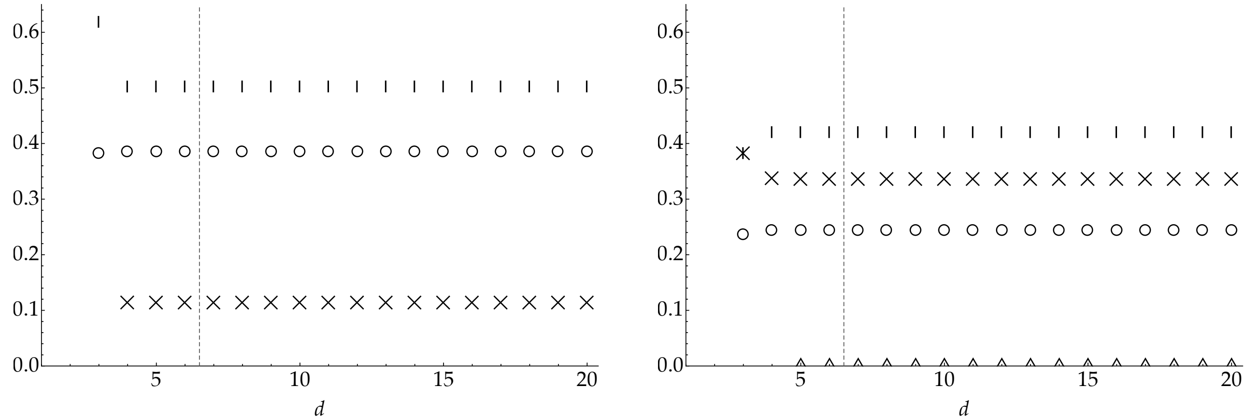
Table 2: Eigenvalues of the derivative matrices $DV(\xi^*)$ of the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic.

In principle, it is possible to prove the local stability of the rest points ξ^* using linearization. But since the components of ξ^* are algebraic numbers, computing the eigenvalues of $DV(\xi^*)$ requires finding the exact roots of a polynomial with algebraic coefficients, a computationally intensive problem. Fortunately, we can prove local stability without doing so. Instead, we compute the eigenvalues of the matrix $DV(\xi)$, where ξ is a rational point that is very close to ξ^*, showing that these eigenvalues all have negative real part. Proposition B.1 in Appendix B establishes an upper bound on the distances between the eigenvalues of $DV(\xi)$ and $DV(\xi^*)$. Importantly, this bound can be evaluated without having to compute the roots of a polynomial with algebraic coefficients or to invert a matrix with algebraic components, as both of these operations quickly become computationally infeasible. Combining these steps allows us to conclude that the eigenvalues of $DV(\xi^*)$ also have negative real part. For a detailed presentation of this argument, see Appendix B.

The approximate eigenvalues of $DV(\xi^*)$ are reported in Table 2. Note that the eigenvalues for games of length 5 and 6 are nearly identical, with the replacement of an eigenvalue of -1 by a pair of complex eigenvalues that are very close to -1.

5.1.2 Numerical results

Because exact methods only allow us to determine the rest points of the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic in Centipede games of lengths $d \leq 6$, we use numerical methods to study games of lengths 7 through 20. We find that in all cases there are two rest points, the backward



(i) penultimate mover

(ii) last mover

Figure 2: The stable rest point of Centipede under the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic for game lengths $d = 3, \ldots, 20$. Markers ○, |, ×, and △, represent weights on strategies [0], [1], [2], and [3] (continue at all decision nodes; stop at the last, second-to-last, or third-to-last decision node). Other weights are less than 10^{-8}. The dashed line separates exact ($d \leq 6$) and numerical ($d \geq 7$) results.

induction state $\xi^\dagger$, and an interior rest point ξ^*. As Figure 2 illustrates, the form of the interior rest point follows the pattern from Table 1: regardless of the length of the game, nearly all of the mass is placed on each population's three most cooperative strategies, and the weights on these strategies are essentially independent of the length of the game. Precise numerical estimates of these rest points are provided in Online Appendix VIII, and numerical estimates of the eigenvalues of the derivative matrices $DV(\xi^*)$ are presented in Online Appendix IX. The latter are essentially identical to those presented in Table 2 for $d = 6$, with the addition of an eigenvalue of ≈ -1 for each additional decision node.

These numerical results strongly suggest that the conclusions about rest points established analytically for games of lengths $d \leq 6$ continue to hold for longer games: there are always exactly two rest points, the backward induction state $\xi^\dagger$, and an stable interior rest point ξ^* whose form barely varies with the length of the game.

5.1.3 Intuition

We now explain why the stable rest point $\xi^* = (x^*, y^*)$ of the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic concentrates its mass on each population's three most cooperative strategies, and consider the effects of other test-set rules on which strategies will be played.

Revising agents are least likely to choose cooperative strategies when few opponents are already cooperating. Thus cooperative strategies should have the most difficulty gaining ground near the backward induction state $\xi^\dagger$. But our arguments above (cf. Example 5.3) explain why a small invasion of cooperative strategies at $\xi^\dagger$ is self-reinforcing. Once such strategies have established a toehold, it is easy for them to become more prevalent: when a player tests a cooperative strategy against a cooperative opponent, this strategy is likely to have performed best.

To understand why each player's three most cooperative strategies predominate, we examine the test results that lead to each strategy being chosen. We assume for concreteness that d is even, so that we can call agents in the player 1 and player 2 roles "first movers" and "last movers" respectively.

The first two terms of equation (13b) give the total probability of test results that would lead a last mover to play strategy j. The terms differ in how play ended when this agent tested strategy j: in the first term, this test ended with the agent stopping the game at his jth decision node; in the second term, the opponent stopped the game at an earlier node.

Now suppose that nearly all first movers play one of their three most cooperative strategies. Then the choice probabilities of last movers are determined by the terms from (13b) whose multiplicands all contain at least one of $x_{[0]}$, $x_{[1]}$, and $x_{[2]}$. Such terms only exist for the last mover's four most cooperative strategies:

$$\Pr(j = [0]) \approx x_{[0]}(x_{[d^1]} + \cdots + x_{[1]}) + x_{[1]}(x_{[d^1]} + \cdots + x_{[2]})^2,$$

$$\Pr(j = [1]) \approx x_{[0]} + x_{[1]}(x_{[d^1]} + \cdots + x_{[2]})\,(x_{[d^1]} + \cdots + x_{[1]}),$$

$$\Pr(j = [2]) \approx (x_{[1]} + x_{[0]})\,(x_{[d^1]} + \cdots + x_{[1]})^2, \tag{24}$$

$$\Pr(j = [3]) \approx (x_{[2]} + x_{[1]} + x_{[0]})\,(x_{[d^1]} + \cdots + x_{[2]})^3. \tag{25}$$

When a last mover tests strategy [0], his opponent necessarily stops before him, so the probability that he chooses [0] comes from the second term in (13b). Examining this term reveals that there are two likely ways in which strategy [0] can have the optimal test result: if it is tested against a first mover playing [0] and strategy [1] is not, or if it is tested against a first mover playing strategy [1] and strategies [1] and [2] are tested against first movers playing less cooperative strategies. When population 1's state is x^*, then the contributions of $x^*_{[0]}$, $x^*_{[1]}$, and $x^*_{[2]}$ determine $y^*_{[0]}$ precisely:

$$x^*_{[0]}(x^*_{[2]} + x^*_{[1]}) + x^*_{[1]}(x^*_{[2]})^2 \approx (.384658)(.615342) + (.501849)(.113493)^2 \approx .243160 \approx y^*_{[0]}.$$

There are two likely ways in which a last mover's strategy [1] can obtain the optimal test result, one for each term from (13b): if it is tested against [0], so that it ends the game with the maximal payoff, or if is tested against [1] (and so does not end the game), strategy [2] is tested against a less cooperative strategy, and strategy [0] is not tested against the most cooperative strategy. At x^* the probability of a last mover choosing [1] is

$$x^*_{[0]} + x^*_{[1]}x^*_{[2]}(x^*_{[2]} + x^*_{[1]}) \approx .384658 + (.501849)(.113493)(.615342) \approx .419706 \approx y^*_{[1]}.$$

The remaining last-mover strategies are only likely to have the optimal test result when they end the game and more cooperative strategies do not earn a higher payoff. At x^* the probabilities that strategies [2] or [3] are chosen are

$$(26) \quad \begin{aligned} &(x^*_{[1]} + x^*_{[0]})\,(x^*_{[2]} + x^*_{[1]})^2 \approx (.886507)(.615342)^2 \approx .335672 \approx y^*_{[2]} \text{ and} \\ &1 \cdot (x^*_{[2]})^3 \approx (.113493)^3 \approx .001462 \approx y^*_{[3]}. \end{aligned}$$

Equations (24) and (26) show that even if first movers only played strategies [0] and [1], last movers would still play strategy [2] frequently, suggesting why both populations must use at least their three most cooperative strategies at the stable state.

Contrariwise, the cubic term in equation (25) impedes the sustained use of more than three strategies. Indeed, the less cooperative is strategy j, the more strategies there are that are more cooperative than j, and for each such strategy k, there are fewer test results under which k fails to outperform j. Because of these two factors, the masses on less-cooperative strategies at the stable rest point of the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic drop extremely rapidly. Under the test-set rules τ^{two} and τ^{adj}, a less-cooperative strategy can be chosen so long as it performs as well as (at most) one more-cooperative strategy. Because of this, the masses on less-cooperative strategies at the stable rest point do not decay at as dramatic a pace. See Section 6.1 and Online Appendix VIII for details.

5.2 Global convergence

The facts that the vertex $\xi^\dagger$ is repelling, the interior rest point ξ^* is attracting, and these are the only two rest points give us a strong reason to suspect that state ξ^* attracts all solutions of the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic other than the stationary solution at $\xi^\dagger$.[32] In this section, we argue that this is indeed the case.

To start, we prove that excepting the rest point $\xi^\dagger$, the boundary of the state space is repelling.

Proposition 5.5. *In Centipede games of all lengths $d \geq 3$, solutions to the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic from every initial condition $\xi \in \text{bd}(\Xi) \setminus \{\xi^\dagger\}$ immediately enter $\text{int}(\Xi)$.*

[32] For there to be other solutions that did not converge to ξ^* without the dynamics having another rest point, the flow of the dynamics would need to have very special topological properties. For instance, in a two-dimensional setting, this could occur if ξ^* were contained in a pair of concentric closed orbits, the inner repelling and the outer attracting.

Of course, since the rest point ξ^* is very close to the boundary for even moderate game lengths d, the distance of motion from the boundary from some initial conditions must be very small.

The proof of Proposition 5.5, which is presented in Appendix C, starts with a simple differential inequality (Lemma C.1) that lets us obtain explicit positive lower bounds on the use of any initially unused strategy i at times $t \in (0, T]$. The bounds are given in terms of the probabilities of test results that lead i to be chosen, and thus, backing up one step, in terms of the usage levels of the opponents' strategies occurring in those tests (equation (63)). With this preliminary result in hand, we prove inward motion from $\xi \neq \xi^\dagger$ by constructing a sequence that contains all unused strategies, and whose kth strategy could be chosen by a revising agent after a test result that only includes strategies that were initially in use or that appeared earlier in the sequence. Variations on this argument establish inward motion for all other BEP dynamics we consider (Remark C.2).

To argue that ξ^* is almost globally stable we introduce the candidate Lyapunov function

$$L(x, y) = \sum_{i=2}^{s_1}(x_i - x_i^*)^2 + \sum_{j=2}^{s_2}(y_j - y_j^*)^2. \tag{27}$$

In words, $L(x, y)$ is the squared Euclidean distance of (x, y) from (x^*, y^*) if the points in the state space Ξ are represented in $\mathbb{R}^d$ by omitting the first components of x and y.

The Gröbner basis techniques from Section 4 do not allow for inequality constraints—here, the nonnegativity constraints on the components of the state—and so are not suitable for establishing that L is a Lyapunov function.[33] We therefore evaluate this claim numerically. For games of lengths 4 through 20, we chose one billion (10^9) points from the state space Ξ uniformly at random, and evaluated a floating-point approximation of $\dot{L}$ at each point. In all instances, the approximate version of $\dot{L}$ evaluated to a negative number. This numerical procedure covers the state space fairly thoroughly for the game lengths we consider,[34] and so provides strong numerical evidence that the interior rest point ξ^* is an almost global attractor, and also about the global structure of the dynamics.

[33]In principle, we could verify that L is a Lyapunov function using an algorithm from real algebraic geometry called *cylindrical algebraic decomposition* (Collins (1975)), but exact implementations of this algorithm fail to terminate once the dimension of the state space exceeds 2.

[34]By a standard combinatoric formula, the number of states in a grid in $\Xi = X \times Y$ with mesh $\frac{1}{m}$ is $\binom{m+s^1-1}{m}\binom{m+s^2-1}{m}$. Applying this formula shows for a game of length 10, 10^9 is between the numbers of states in grids in Ξ of meshes $\frac{1}{17}$ (since $\binom{22}{17}^2 = 693{,}479{,}556$) and $\frac{1}{18}$ (since $\binom{23}{18}^2 = 1{,}132{,}255{,}201$). For a game of length 15, the comparable meshes are $\frac{1}{10}$ and $\frac{1}{11}$, and for length 20, $\frac{1}{7}$ and $\frac{1}{8}$.

6. Analyses of other test-set and tie-breaking rules

6.1 Other test-set rules

We have focused on tie-breaking rule $\beta^{\min}$ because it is the most favorable to the backward induction states. On the contrary, the test-set rule τ^{all} works in favor of the emergence of cooperative behavior. A revising agent who tests all of his strategies has many opportunities to test a cooperative strategy against a cooperative opponent, and so to obtain a high payoff. By reducing the number of such chances, testing fewer strategies makes switches to cooperative strategies considerably less likely, especially starting from initial states with little cooperative play. This leads us to consider BEP dynamics based on the test-two and test-adjacent rules. In summary, we find that the qualitative predictions described in the previous section go largely unchanged when these less ambitious test-set rules are used. The conclusion for the test-adjacent rule is particularly striking, as it shows that cooperative play will emerge even if it only may do so one step at a time.

6.1.1 Test-two

Our analytical results for the BEP($\tau^{\text{two}}, 1, \beta^{\min}$) dynamic are as follows

Proposition 6.1. *Consider the BEP($\tau^{\text{two}}, 1, \beta^{\min}$) dynamic in Centipede games.*

(*i*) *For games of all lengths d, the rest point $\xi^\dagger$ is a repellor.*

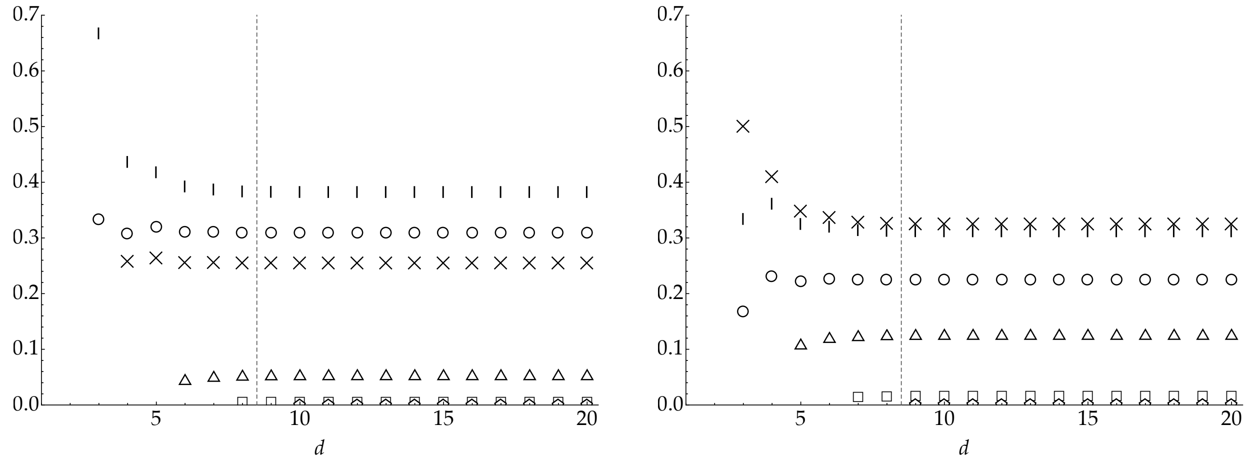
(*ii*) *For games of lengths $3 \leq d \leq 8$, the dynamic has exactly two rest points, $\xi^\dagger$ and $\xi^* \in \text{int}(\Xi)$. The rest point ξ^* is asymptotically stable.*

(*iii*) *For games of all lengths d, solutions from every initial condition $\xi \in \text{bd}(\Xi) \smallsetminus \{\xi^\dagger\}$ immediately enter* $\text{int}(\Xi)$.

The only differences between the conclusions stated in Proposition 6.1 and those from the previous section appears in part (ii): because testing only two strategies leads the polynomials that define the dynamics to have fewer terms, we are able to analytically characterize the rest points for games of lengths up to $d = 8$.

The proof that $\xi^\dagger$ is a repellor is presented in Appendix A.2. The approach is that described in Section 5.1.1, but in this case the characterization of the eigenvectors of $DV(\xi^\dagger)$ is more difficult. The characterization of the rest points is again proved using the algebraic methods from Section 4.

Analogues of the numerical results from Section 5 also go through. As shown in Figure 3, the form of the stable rest point settles down once $d = 8$, with the interior rest point ξ^* here concentrating its mass on the penultimate player's four most cooperative strategies and the last player's five most cooperative strategies. Online Appendices VIII and IX



(i) penultimate mover (ii) last mover

Figure 3: The stable rest point of Centipede under the BEP($\tau^{\text{two}}, 1, \beta^{\text{min}}$) dynamic for game lengths $d = 3, \ldots, 20$. Markers ○, |, ×, △, □, and ⬠ represent weights on strategies [0], [1], [2], [3], [4], and [5]. Other weights are less than 10^{-4}. The dashed line separates exact ($d \leq 8$) and numerical ($d \geq 9$) results.

present precise values of the rest points for game lengths up to 20 and of the eigenvalues of the derivative matrices $DV(\xi^*)$. In all cases, all eigenvalues have negative real part, providing strong numerical evidence of the local stability of ξ^*. Finally, for games of lengths 3 through 20, we computed the time derivative of the squared distance function

$$L(x, y) = \sum_{i=1}^{s_1} (x_i - x_i^*)^2 + \sum_{j=1}^{s_2} (y_j - y_j^*)^2. \tag{28}$$

at one billion randomly chosen points in Ξ, finding that it evaluates to a negative number in all cases.[35] This again provides strong numerical evidence that the interior rest point ξ^* is an almost global attractor.

6.1.2 Test-adjacent

Under the test-adjacent rule, revising agents only consider their current strategy and one adjacent strategy: jumps from uncooperative to cooperative strategies are forbidden. Despite this impediment to cooperative play, Proposition 6.2 shows that our analytical results for the BEP($\tau^{\text{adj}}, 1, \beta^{\text{min}}$) dynamic agree qualitatively with those stated previously.

Proposition 6.2. *The conclusions of Proposition 6.1 also hold for the BEP*($\tau^{\text{adj}}, 1, \beta^{\text{min}}$) *dynamic, except that part* (*ii*) *holds for games of lengths* $3 \leq d \leq 7$.

[35] This is the only instance in which we use the Lyapunov function (28) instead of (27).

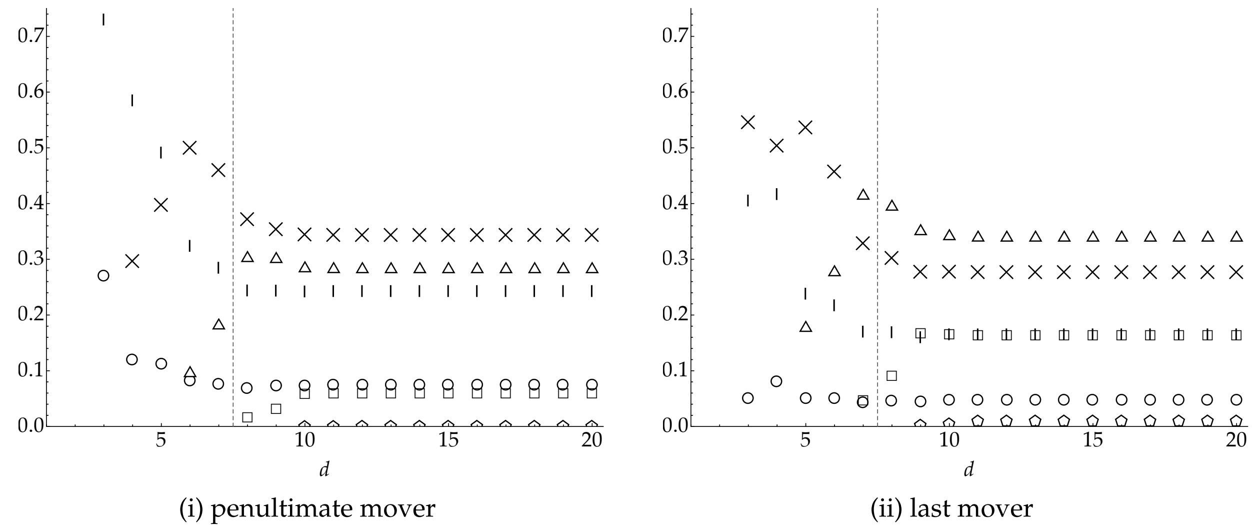


Figure 4: The stable rest point of Centipede under the BEP($\tau^{\text{adj}}, 1, \beta^{\text{min}}$) dynamic for game lengths $d = 3, \ldots, 20$. Markers ○, |, ×, △, □, and ⬠ represent weights on strategies [0], [1], [2], [3], [4], and [5]. Other weights are less than 10^{-5}. The dashed line separates exact ($d \leq 7$) and numerical ($d \geq 8$) results.

The proofs of the results in Proposition 6.2 follow the same lines as before. However, the proof that $x^\dagger$ is repelling under test-adjacent is more involved because the derivative matrices $DV(\xi^\dagger)$ are not diagonalizable, requiring us to consider generalized eigenvectors, and because the pattern followed by these generalized eigenvectors is not simple. See Appendix A.3 for details.

Figure 4 presents exact ($d \leq 7$) and numerical ($d \geq 8$) coordinates of the stable rest point ξ^*. Here the components of the rest point do not settle down until the length of the game is 10, and the mass is concentrated on the penultimate player's five most cooperative strategies and the last player's six most cooperative strategies.

To assess global stability for games of length $d \geq 3$, we again evaluated the time derivative of the function L from (27) at one billion randomly chosen points in Ξ. For games of lengths $d \leq 6$, the time derivative always evaluates to a negative number, indicating that L is a strict Lyapunov function, and ξ^* an almost global attractor. For games of length $d \geq 7$, the time derivative of L is positive at some states, and so L is not a Lyapunov function. This is not entirely surprising, since the test-adjacent rule forbids agents playing uncooperative strategies to switch directly to the strategies most commonly used at the interior rest point. To obtain weaker evidence that ξ^* is an almost global attractor, we computed numerical solutions to the BEP($\tau^{\text{adj}}, 1, \beta^{\text{min}}$) dynamic from one million randomly chosen initial conditions, and then checked whether each solution entered a ball of radius 10^{-4} around a floating-point approximation of ξ^*. For games of lengths 7 through 20, we found that this is indeed the case.

population p	[5]	[4]	[3]	[2]	[1]	[0]
test-all	.0000	.0000	.0000	.1135	.5018	.3847
test-two	.0002	.0040	.0514	.2545	.3817	.3083
test-adjacent	.0006	.0581	.2820	.3432	.2423	.0739

population q	[5]	[4]	[3]	[2]	[1]	[0]
test-all	.0000	.0000	.0015	.3357	.4197	.2432
test-two	.0008	.0146	.1239	.3246	.3111	.2249
test-adjacent	.0101	.1623	.3386	.2766	.1646	.0478

Table 3: The interior attractor of BEP($\tau, 1, \beta^{\min}$) dynamics under different test-set rules τ ($d \geq 12$).

The requirement of step-by-step adjustments also suggests that the BEP($\tau^{\mathrm{adj}}, 1, \beta^{\min}$) dynamic should require more time to travel from the vicinity of the backward induction state to the stable rest point than dynamics based on other test-set rules, particularly in longer games. To assess this claim, we computed numerical solutions to the BEP($\tau, 1, \beta^{\min}$) dynamics with $\tau \in \{\tau^{\mathrm{all}}, \tau^{\mathrm{two}}, \tau^{\mathrm{adj}}\}$ in Centipede games of lengths $d = 10$ and $d = 20$, with initial conditions near the backward induction state. For both game lengths, the time to reach a neighborhood of ξ^* increases substantially if one changes the test-set rule from τ^{all} to τ^{two}, and then from τ^{two} to τ^{adj}. See Online Appendix IV for details.

Table 3 presents the components of the stable rest point of the three BEP($\tau, 1, \beta^{\min}$) dynamics for Centipede games of lengths $d \geq 12$, long enough that the first four digits of the components are fixed. One fact worth noting is that the fully cooperative strategies receive far less weight under test-adjacent than under the other tie-breaking rules. To understand why, note that (i) a revising player may choose strategy [0] only if he was already playing this strategy, or if he was playing strategy [1] and tests strategies [1] and [0] (rather than [1] and [2]); and (ii) the min-if-tie rule implies that a player testing [0] and [1] will only choose [0] if his opponent when testing [0] plays a more cooperative strategy than his opponent when testing [1].

6.2 Other tie-breaking rules

The min-if-tie rule $\beta^{\min}$ is the tie-breaking rule that is least supportive of cooperative play. In this section we briefly describe the analogues of the forgoing results when other tie-breaking rules are used. Specifically, we consider the stick-if-tie rule β^{stick} (equation (7)), we which we feel is the most natural specification, and the uniform-if-tie rule β^{unif} (equation (8)), which when combined with the test-all rule τ^{all} defines the $S(k)$ equilibrium of Osborne and Rubinstein (1998) and the corresponding dynamics of Sethi (2000).

While a variety of details of the analyses depend on the choice of tie-breaking rule, our main conclusions do not: in all cases, play under BEP dynamics in Centipede from almost all initial conditions converges to an interior rest point at which choices are concentrated on the most cooperative strategies, with weights that become essentially independent of the length of the game. Tables 4–6 report the weights on the most cooperative strategies at the interior rest point for each combination of test-set and tie-breaking rules. Table 7 presents the maximal lengths of Centipede for which rest points can be characterized analytically, along with the degree of the leading polynomial of the Gröbner basis in each case. Full descriptions of the rest points and the eigenvalues of the derivative matrices $DV(\xi^*)$ for game lengths $d \in \{2, \ldots, 20\}$ are provided in Online Appendices VIII and IX.[36]

Tables 4 and 5 show that under test-set rules τ^{all} and τ^{two}, switching the tie-breaking rule from β^{min} to β^{stick} or β^{unif} leads to a stable rest point at which the fully cooperative strategy [0] gains weight at the expense of less cooperative strategies. As anticipated in the previous section, these switches in the tie-breaking rule lead to more substantial increases in cooperation under the τ^{adj} rule, as Table 6 shows.

Under uniform tie-breaking β^{unif}, the backward induction state is $\xi^\dagger$ not a rest point, as population 2 agents randomize among whatever strategies are in their test sets. In other respects, the properties and analysis of these dynamics are similar to those under β^{min}.

State $\xi^\dagger$ is a rest point under the β^{stick} rule, and the behavior of the dynamics on $\mathrm{aff}(\Xi)$ in the vicinity of this state complicates the analysis. For each choice of test-set rule τ and for most game lengths, there are continuous sets of rest points of the $\mathrm{BEP}(\tau, 1, \beta^{\mathrm{stick}})$ dynamic in $\mathrm{aff}(\Xi)$. These sets consist of points satisfying $x_1 = 1$, and the only state in Ξ they contain is the backward induction state $\xi^\dagger$.[37] The existence of these sets of rest points implies that the derivative matrices $DV(\xi^\dagger)$ have eigenvalues equal to zero, so that the standard results from linearization theory used earlier (see Section 5.1.1) cannot be applied. To show state $\xi^\dagger$ is nevertheless repelling, we appeal to results from center manifold theory (Kelley (1967a,b), Perko (2001)) which describe the behavior of nonlinear dynamics near nonhyperbolic rest points. See Appendix A.4 and Online Appendix V for details.[38]

[36]Example 5.1 shows that in Centipede of length $d = 2$, the $\mathrm{BEP}(\tau, 1, \beta^{\mathrm{min}})$ dynamic converges to the backward induction state $\xi^\dagger$ from all initial conditions. Switching the tie-breaking rule to β^{stick} or β^{unif} introduces new positive terms to the expression for $\dot{y}_2$ in (22), corresponding to cases in which population 2 agents test both strategies against opponents who stop immediately. Because of these terms, the dynamics possess an almost globally stable interior rest point. As discussed after Example 5.3, one could rule out this behavior by fiat by imposing the principle of admissibility.

[37]Examining display (14), it is easy to verify that under the $\mathrm{BEP}(\tau^{\mathrm{two}}, 1, \beta^{\mathrm{stick}})$ dynamic, all states in $\{(x, y) \in \mathrm{aff}(\Xi) \colon x = (1, 0 \ldots, 0)', y_1 = 1\}$ are rest points. Under τ^{all} and τ^{adj}, the sets of rest points in $\mathrm{aff}(\Xi)$ with $x_1 = 1$ vary with the length of the game in more complicated ways.

[38]Another difficulty arises because the Gröbner basis techniques from Section 4 require the set of polynomials under consideration to have a finite number of zeros. To circumvent this problem, we record the rest

	population p				population q			
	[3]	[2]	[1]	[0]	[3]	[2]	[1]	[0]
min	.0000	.1135	.5018	.3847	.0015	.3357	.4197	.2432
stick/min	.0000	.1106	.4999	.3895	.0014	.3315	.4163	.2508
uniform	.0000	.1080	.4980	.3939	.0013	.3277	.4131	.2580

Table 4: The interior attractor of BEP($\tau^{\text{all}}, 1, \beta$) under different tie-breaking rules β ($d \geq 6$). p denotes the owner of the penultimate decision node, q the owner of the last decision node.

population p	[5]	[4]	[3]	[2]	[1]	[0]
min	.0002	.0040	.0514	.2545	.3817	.3083
stick	.0001	.0033	.0436	.2256	.3707	.3566
uniform	.0001	.0034	.0439	.2261	.3690	.3575

population q	[5]	[4]	[3]	[2]	[1]	[0]
min	.0008	.0146	.1239	.3246	.3111	.2249
stick	.0007	.0123	.1066	.2969	.3149	.2685
uniform	.0007	.0124	.1071	.2960	.3124	.2713

Table 5: The interior attractor of BEP($\tau^{\text{two}}, 1, \beta$) under different tie-breaking rules β ($d \geq 12$).

population p	[5]	[4]	[3]	[2]	[1]	[0]
min	.0006	.0581	.2820	.3432	.2423	.0739
stick	.0000	.0047	.1253	.3417	.3672	.1611
uniform	.0000	.0042	.1200	.3393	.3708	.1658

population q	[5]	[4]	[3]	[2]	[1]	[0]
min	.0101	.1623	.3386	.2766	.1646	.0478
stick	.0002	.0360	.2419	.3375	.2623	.1222
uniform	.0001	.0333	.2362	.3348	.2685	.1271

Table 6: The interior attractor of BEP($\tau^{\text{adj}}, 1, \beta$) under different tie-breaking rules β ($d \geq 12$).

	test-all	test-two	test-adjacent
min	6 (221)	8 (97)	7 (202)
stick	5 (65)	8 (128)	6 (47)
uniform	6 (168)	8 (128)	7 (230)

Table 7: The maximal lengths d of Centipede for which the rest points of the BEP(τ, κ, β) dynamic have been computed analytically, and the degrees of the leading polynomial of the Gröbner basis (in parentheses).

7. Robustness to including backward induction agents

To this point we have assumed that each population is homogeneous, with all members employing the same best experienced payoff protocol. Given that the predictions generated by this protocol differ dramatically from traditional ones, it is natural to ask what happens if populations are heterogeneous, with some agents playing in accordance with backward induction—that is, always stopping immediately. Because including a small proportion of agents who stop immediately only slightly alters the differential equations that describe the behavior of the majority of the agents, doing so has little effect on predictions.[39] What happens with larger proportions?

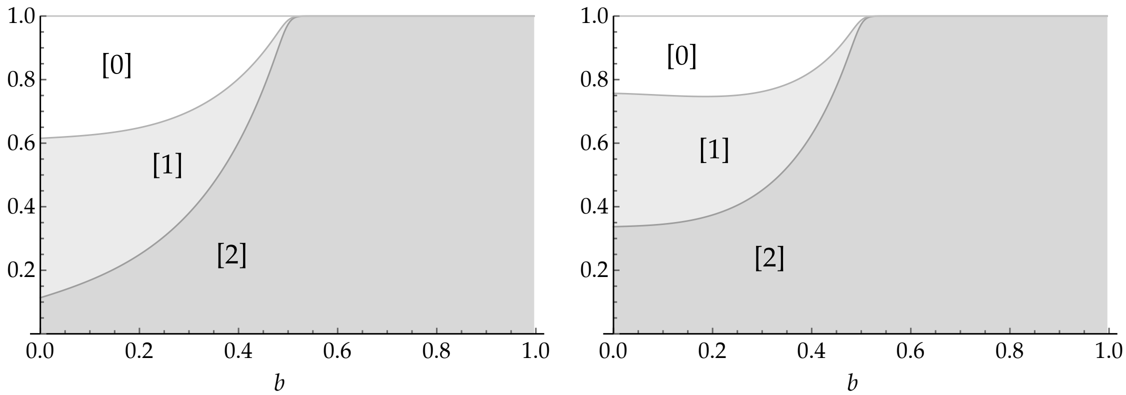
Figures 5 and 6 present the distributions of strategies chosen by BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) agents at the stable rest point as the fraction b of backward induction agents in each population is varied from 0 to 1. In a Centipede game of length $d = 4$ (Figure 5), having 20% backward induction agents has a limited effect on the behavior of the BEP agents, and cooperation is not eliminated completely until the population is about evenly split. In a game of length $d = 20$ (Figure 6), the robustness of cooperation is much more substantial. When 40% of each population consists of backward induction agents, the behavior of the BEP agents differs little from their behavior when there are no backward induction agents at all. Even if 70% of agents are assumed to stop immediately, the behavior of the remaining agents is strikingly cooperative. Thus in longer games, cooperative play by a minority of agents can be sustained, even if the majority is fully uncooperative.

One reason for the persistence of cooperation illustrated in Figure 6 is the nature of the test-all rule: when agents test all available strategies, a longer game provides more opportunities for a successful test of a cooperative strategy. To assess the extent to which this effect drives cooperation, we can instead suppose that the BEP agents use the test-two rule. Figure 7 illustrates the results. While cooperation drops off more quickly than under the test-all rule, it still persists when there are large fractions of backward induction agents. When this fraction is $b = \frac{1}{3}$, the distribution of choices among the BEP agents is nearly uniform, and cooperative choices by the BEP agents is still quite common when $b = \frac{2}{3}$. Thus cooperative play is robust to the introduction of agents who always stop immediately even when agents' test sets are small.

---

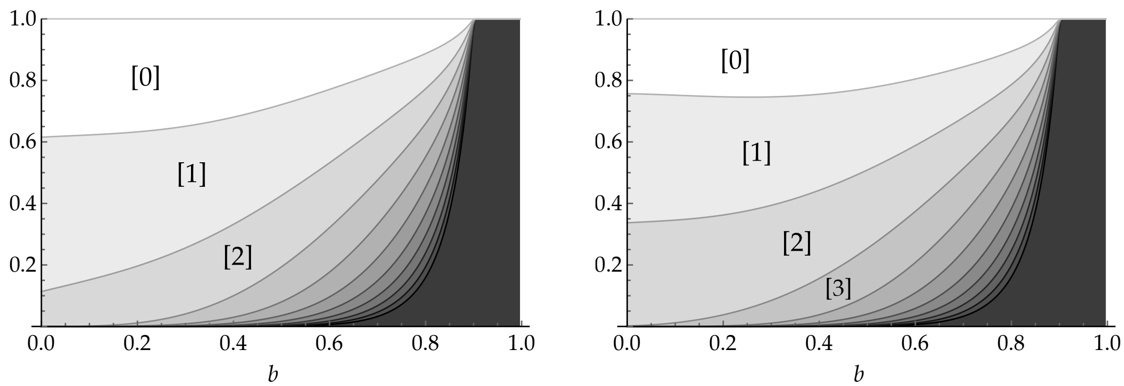
point $\xi^\dagger$ and then eliminate the extraneous rest points using a standard trick, representing the nonequality $x_1 \neq 1$ by introducing a new variable z and new equality $(x_1 - 1)z = 1$ as inputs of the Gröbner basis algorithm.

[39]For formal results along these lines, see Hofbauer and Sandholm (2011).



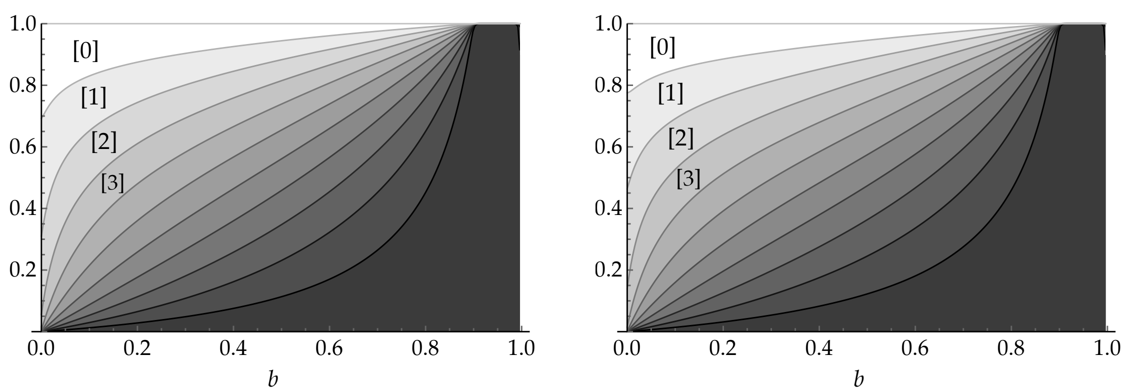
(i) population 1 (ii) population 2

Figure 5: Behavior of BEP($\tau^{\text{all}}, 1, \beta^{\min}$) agents at the stable rest point in Centipede of length $d = 4$ when proportion $b \in [0, 1]$ of each population always stops immediately.



(i) population 1 (ii) population 2

Figure 6: Behavior of BEP($\tau^{\text{all}}, 1, \beta^{\min}$) agents at the stable rest point in Centipede of length $d = 20$ when proportion $b \in [0, 1]$ of each population always stops immediately.



(i) population 1 (ii) population 2

Figure 7: Behavior of BEP($\tau^{\text{two}}, 1, \beta^{\min}$) agents at the stable rest point in Centipede of length $d = 20$ when proportion $b \in [0, 1]$ of each population always stops immediately.

8. Larger numbers of trials

The analysis thus far has focused on cases in which agents test each strategy in their test sets exactly once. We now examine aggregate behavior when each strategy is subject to larger numbers of trials.

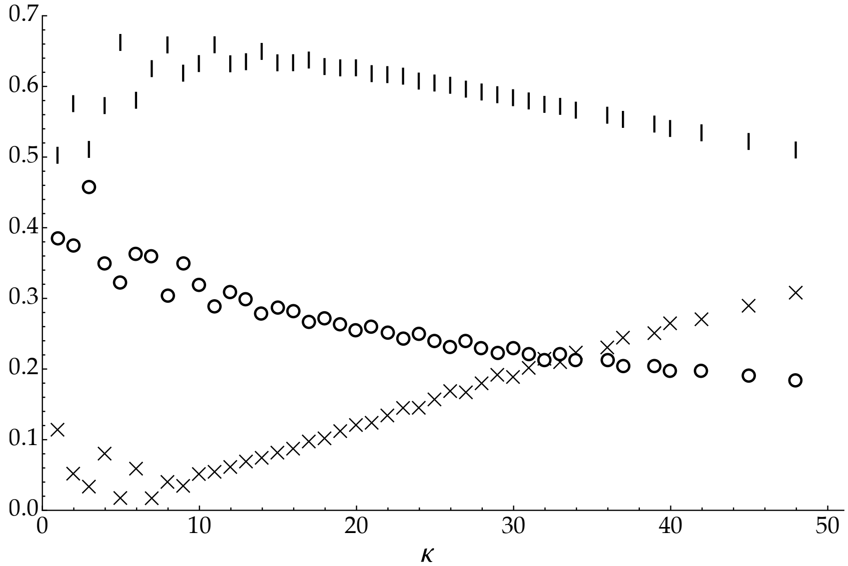
8.1 Persistence of cooperative behavior

When agents test their strategies more thoroughly, the distributions of opponents' choices they face when testing each strategy will come to resemble the current distribution of play in the opposing population. Since agents choose the strategy whose total payoff during testing was highest, this suggests that the rest points of the resulting dynamics should approximate Nash equilibria. Indeed, when agents possess exact information, so that play adjusts according to the exact best response dynamic (Gilboa and Matsui (1991), Hofbauer (1995b)), results of Xu (2016) imply that every solution trajectory converges to the set of Nash equilibria; in Centipede, all Nash equilibria entail all population 1 agents stopping immediately.

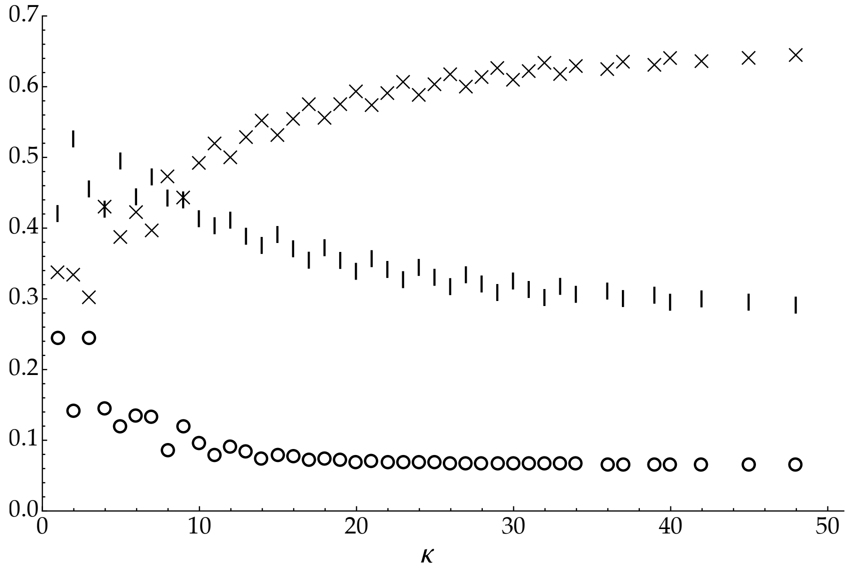
While the intuition suggested above is correct for large enough numbers of trials, it is nevertheless the case that stable cooperative behavior can persist when the number of trials of each strategy is substantial. To illustrate this, we consider play in the Centipede game of length $d = 4$ under the BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamic. Figures 8 and 9 present the stable rest points of this dynamic for numbers of trials κ up to 50, which we computed using numerical methods.[40] While increasing the number of trials shifts mass toward uncooperative strategies, it is clear from the figures that this shifting takes place gradually: even with rather thorough testing, significant levels of cooperation are still maintained. We note as well that the fraction of population 2 agents who play the weakly dominated strategy 3 (always continue) becomes fixed between 7% and 6.5% once $\kappa \geq 15$, even as the fraction of population 1 agents who play strategy 3 remains far from 0 (specifically, between 28% and 18%).

While surprising at first glance, these facts can be explained by considering both the expectations and the *dispersions* in the payoffs obtained through repeated trials of each strategy. As an illustration, consider the stable rest point when $\kappa = 32$, namely $\xi^* = (x^*, y^*) \approx ((.2140, .5738, .2122), (.6333, .3010, .0657))$. Let Π_j be a random variable representing the payoff obtained by a population 2 agent who plays strategy j in a single

[40]To perform computations efficiently, we use a presentation of BEP dynamics stated in terms of multinomial distributions. When the number of trials is not small, this presentation has far fewer terms than definition (B). See Online Appendix VII for details.



(i) population 1



(ii) population 2

Figure 8: The stable rest point of the BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamic in the Centipede game of length $d = 4$, $\kappa = 1, \ldots, 50$. Markers $\bigcirc$, $|$, and $\times$, represent weights on strategies 3 (= [0]), 2 (= [1]), and 1 (= [2]).

random match at this state. By equation (12) (or Figure 1), the expected payoffs to this agent's three strategies are

$$\mathbb{E}(\Pi_1) = (0,3,3)\cdot x^* = 2.3580, \;\; \mathbb{E}(\Pi_2) = (0,2,5)\cdot x^* = 2.2086, \;\; \mathbb{E}(\Pi_3) = (0,2,4)\cdot x^* = 1.9964.$$

From this we anticipate that the strategy weights in population 2 satisfy $y_1^* > y_2^* > y_3^*$.

To explain why these weights take the values they do, we also need to know how dispersed the payoffs from testing each strategy are. We thus compute the variances of the single-test payoffs Π_j:

$$\mathrm{Var}(\Pi_1) = 1.5138, \;\; \mathrm{Var}(\Pi_2) = 2.7223, \;\; \mathrm{Var}(\Pi_3) = 1.7048.$$

Using these calculations and the central limit theorem, we find that the difference between the average payoffs from 32 tests of strategy 3 and 32 tests of strategy 2 is approximately normally distributed with mean $\mathbb{E}(\Pi_3) - \mathbb{E}(\Pi_2) = -.2122$ and standard deviation $\sqrt{(\mathrm{Var}(\Pi_3) + \mathrm{Var}(\Pi_2))/32} \approx .3720$. The latter statistic is commensurate with the former. Thus the weakly dominated strategy 3 yields a higher total payoff than the dominating strategy 2 with approximate probability $\mathbb{P}(Z \geq .57) \approx .28$, and so is not a rare event. Likewise, evaluating the appropriate multivariate normal integrals shows that the probabilities of strategies 1, 2, and 3 yielding the highest total payoff are approximately .61, .32, and .07, figures which accord fairly well with the components of y^*.

As the number of trials κ becomes larger, greater averaging reduces the variation in each strategy's payoffs per trial. At the same time, increasing κ increases the weight x_1^* on stopping immediately at the expense of population 1's other two strategies, reducing the differences in the expected payoffs of population 2's strategies. This explains why the strategy weights in population 2 do not vary very much as κ increases, and why the weight on the weakly dominated strategy hardly varies at all.

8.2 Convergence to cycles

Figure 8 does not record rest points for certain numbers of trials above 34. For these values of κ, the population state does not converge to a rest point. Instead, our numerical analyses indicate that for all κ with empty entries in Figure 8 and all κ between 51 and 100, the BEP($\tau^{\text{all}}, \kappa, \beta^{\min}$) dynamic converges to a periodic orbit. Figure 10 presents the cycles under the BEP($\tau^{\text{all}}, \kappa, \beta^{\min}$) dynamics for κ = 50, 100, and 200. In all three cases, we observe substantial levels of cooperative play in population 1 over the course of the cycle, with the fraction of the population choosing to continue at the initial node varying between .50 and .83 for $\kappa = 50$, between .28 and .70 for $\kappa = 100$, and between .16 and .45

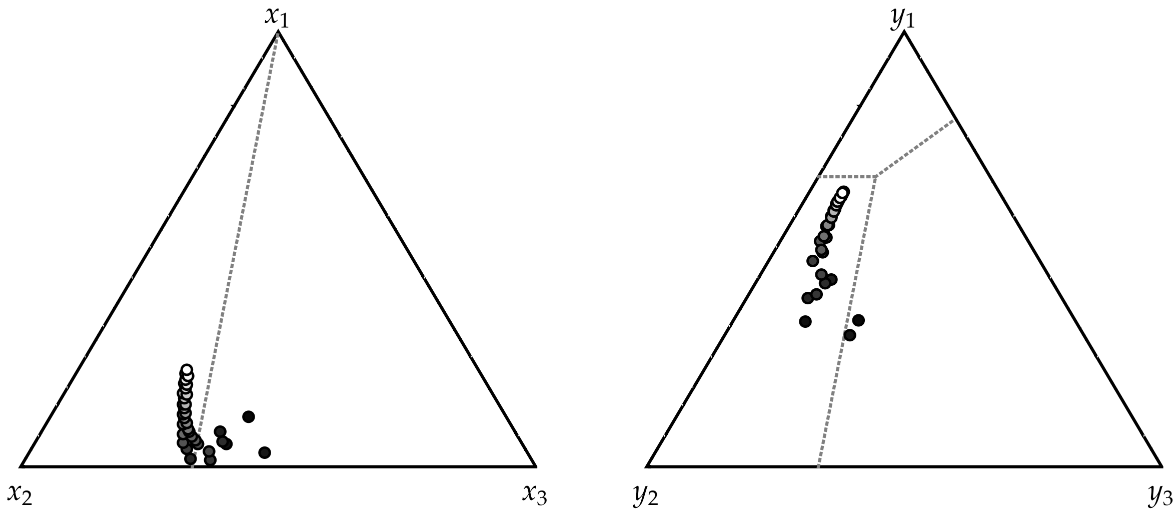


Figure 9: The stable rest point in Centipede of length $d = 4$ under BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 1, \ldots, 34$ trials of each tested strategy. Lighter shading corresponds to larger numbers of trials. Dashed lines represent boundaries of best response regions.

for $\kappa = 200$. These examples illustrate that cooperative behavior can persist even when agents have substantial amounts of information about opponents' play.

From a methodological point of view, the existence of almost globally attracting limit cycles under BEP dynamics suggests that solution concepts like $S(k)$ equilibrium and logit equilibrium that are motivated as steady states of dynamic disequilibrium processes should be applied with some caution. Existence results for such solution concepts can generally be proved by appeals to suitable fixed point theorems. But the fact that static solutions exist need not imply that any are stable, and it may happen that no static solution provides a good prediction of the behavior of the underlying dynamic process.

9. Discussion

In this paper, we have introduced a class of game dynamics built on natural assumptions about the information agents obtain when revising, and have shown that these dynamics lead to cooperative behavior in the Centipede game. One key feature of the agents' revision process is that conditional on the current population state, the experienced payoffs to each strategy are independent of one another. This allows cooperative strategies with suboptimal expected payoffs to be played with nonnegligible probabilities, even when the testing of each strategy involves substantial numbers of trials. The use of any such strategy increases the expected payoffs of other cooperative strategies, creating a virtuous circle that sustains cooperative play.

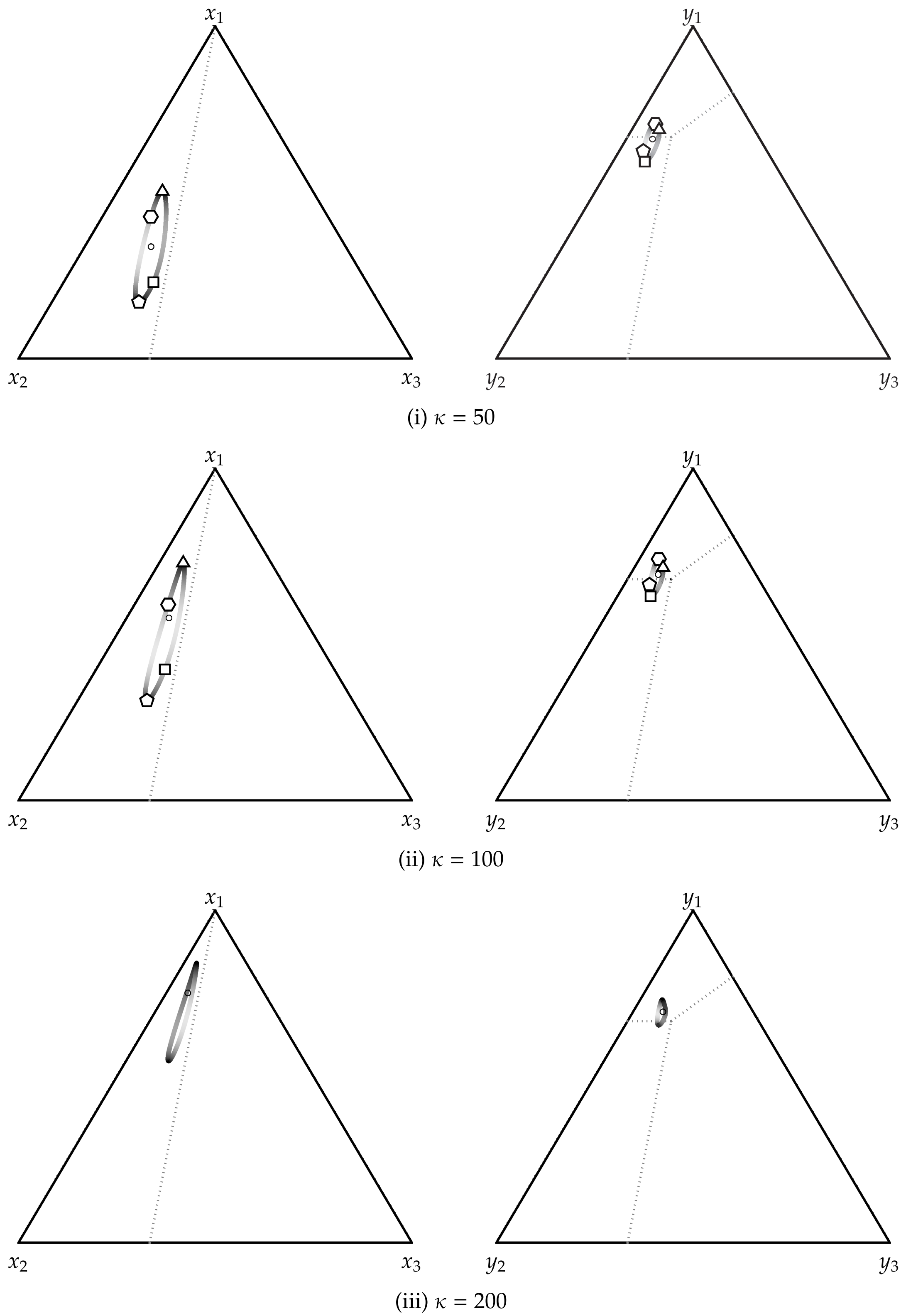


Figure 10: Stable cycles in Centipede of length $d = 4$ under BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 50$, 100, and 200. Lighter shading represents faster motion. The small circles represent the unstable interior rest points. For $\kappa = 50$ and 100, shapes synchronize positions along the cycle.

9.1 Comparisons to discrete choice models

A related and thoroughly-studied approach to modeling choice under population game dynamics is to use constructions from discrete choice theory.[41] The most common specification employed is the logit choice model, in large part because of its simple functional form.[42] Equilibria of the logit model are commonly used in the analysis of experimental data, as they offer a convenient way of incorporating suboptimal choices into equilibrium predictions.[43]

Logit dynamics and BEP dynamics share a key feature: both incorporate independent realizations of the payoffs to different strategies. These dynamics also differ in basic ways. While specifying a high noise level for the logit dynamic leads to near-uniform randomization, choosing small numbers of trials under BEP dynamics yields interesting predictions of play. In addition, while the payoff perturbations used to define logit dynamics take the same form regardless of the current population state, the variations in payoffs under BEP dynamics depend on the state in a fundamental way. This allows strict equilibria to be rest points of all BEP dynamics (cf. Observation 2.1), while under logit dynamics the boundary of the state space is always repelling.

From a computational standpoint, logit choice probabilities can be expressed in a closed form, but one involving transcendental functions. This makes exact computation of logit equilibrium impossible. The fact that BEP dynamics are described by polynomial equations makes them susceptible to analysis using computational algebra, allowing us to obtain exact expressions for rest points, as well as rigorous proofs of dynamic stability.

As important as these operational differences are the conceptual differences between discrete choice and BEP models for games. The former proceed by augmenting a game with a payoff noise structure that is completely distinct from the game itself. By contrast, BEP models do not introduce new sources of payoffs. Rather, they drop standard but strong assumptions concerning knowledge about opponents' behavior, assuming that agents make choices using only the information gained from limited experiences of play. This step in the direction of realism leads to a model whose predictions can differ from traditional ones and, at least in the case of Centipede, may better reflect observed behavior.

When the number of trials of each tested strategy is not small, the central limit theorem allows us to approximate choice probabilities under BEP dynamics by probit choice

[41] See, e.g., Fudenberg and Levine (1998) and Hofbauer and Sandholm (2002, 2007).

[42] See the previous references, as well as Blume (1997).

[43] McKelvey and Palfrey (1995) introduce equilibrium notions defined in terms of discrete choice models under the name *quantal response equilibrium*, and place emphasis on logit equilibrium. McKelvey and Palfrey (1998) extend this notion to extensive form games, focusing on a definition of equilibrium based on the agent normal form, and use this concept to fit experimental data from the Centipede game.

probabilities—that is, by probabilities obtained by perturbing strategies' payoffs with independent, normally distributed disturbance terms—but with the variances in these disturbances depending upon the current state (see Section 8). This fact suggests an approach to studying the behavior of BEP dynamics with large numbers of trials, for which exact expressions are too cumbersome for use even in numerical analysis. We are developing this approach in our current research.

9.2 Analyses of other game dynamics

Our ability to analyze BEP dynamics using tools from computational algebra depended only on the fact that these dynamics are described by polynomial equations with rational coefficients. There are many other dynamics that share this property, including other dynamics based on random sampling and various sorts of imitative dynamics.[44] The powerful algebraic tools used here have the potential to generate new qualitative insights even about well-studied game dynamics, offering a fertile domain for future work.

9.3 Analyses of general extensive form games

A basic message of our analysis is that natural weakenings of assumptions about agents' knowledge, ones that do not in turn require common knowledge of a more complicated interaction, can lead to predictions of play that differ markedly from ones based on backward induction and that accord in broad terms with experimental and anecdotal evidence. Our analysis has focused on the Centipede game because it provides the simplest model of ongoing relationships with a fixed terminal date. It is important to confirm that the ideas underlying the stability of cooperative behavior here remain relevant in other games in which violations of backward induction can be mutually beneficial, and to identify other properties of games that work for or against cooperative play. We leave these questions for future research.

[44]For the former, see Oyama et al. (2015); for the latter, see Taylor and Jonker (1978), Weibull (1995), Hofbauer (1995a), and Sandholm (2010b).

Appendix

A. Repulsion from the backward induction state

Letting $s = s^1 + s^2$, we denote the tangent space of the state space $\Xi = X \times Y$ by $T\Xi = TX \times TY = \{(z^1, z^2)' \in \mathbb{R}^s \colon \sum_{i \in S^1} z_i^1 = 0 \text{ and } \sum_{j \in S^2} z_j^2 = 0\}$, and we denote the affine hull of Ξ by $\text{aff}(\Xi) = T\Xi + \xi^\dagger$. Writing our dynamics as

(D) $\qquad \dot{\xi} = V(\xi),$

we have $V \colon \text{aff}(\Xi) \to T\Xi$, and so $DV(\xi)z \in T\Xi$ for all $\xi \in \Xi$ and $z \in T\Xi$. We can thus view $DV(\xi)$ as a linear map from $T\Xi$ to itself, and the behavior of the dynamics in the neighborhood of a rest point is determined by the eigenvalues and eigenvectors of this linear map. The latter are obtained by computing the eigenvalues and eigenvectors of the product matrix $\mathbf{\Phi} D\mathcal{V}(\xi)\mathbf{\Phi}$, where $\mathcal{V} \colon \mathbb{R}^s \to \mathbb{R}^s$ is the natural extension of V to $\mathbb{R}^s$, and $\mathbf{\Phi}$ is the orthogonal projection of $\mathbb{R}^s$ onto $T\Xi$, i.e., the block diagonal matrix with diagonal blocks $I - \frac{1}{s^1}\mathbf{1}\mathbf{1}' \in \mathbb{R}^{s^1 \times s^1}$ and $I - \frac{1}{s^2}\mathbf{1}\mathbf{1}' \in \mathbb{R}^{s^2 \times s^2}$, where $\mathbf{1} = (1, \ldots, 1)'$. Since V maps Ξ into $T\Xi$, the projection is only needed when there are eigenspaces of $DV(\xi)$ that intersect both the set $T\Xi$ and its complement.

In what follows we write $\delta^i \in \mathbb{R}^s$ and $\varepsilon^j \in \mathbb{R}^s$ for the standard basis vectors corresponding to strategies $i \in S^1$ and $j \in S^2$, respectively. We also write all expressions in terms of the numbers of decision nodes rather than the numbers of strategies, as doing so usually generates more compact expressions. To eliminate superscripts we use the notations $m \equiv d^1 = s^1 - 1$ and $n \equiv d^2 = s^2 - 1$ for the numbers of decision nodes.

For BEP dynamics with tie-breaking rule $\beta^{\min}$, we prove that the backward induction state $\xi^\dagger$ is a repellor using the following argument. Computing the eigenvalues and eigenvectors of $DV(\xi^\dagger)$ as described above, we find that under each of the three test-set rules we consider, $\xi^\dagger$ is a *hyperbolic* rest point, meaning that all of the eigenvalues have nonzero real part.

The linearization of the dynamic (D) at rest point $\xi^\dagger$ is the linear differential equation

(L) $\qquad \dot{z} = DV(\xi^\dagger)z$

on $T\Xi$. The *stable subspace* $E^s \subseteq T\Xi$ of (L) is the span of the real and imaginary parts of the eigenvectors and generalized eigenvectors of $DV(\xi^\dagger)$ corresponding to eigenvalues with negative real part. The *unstable subspace* $E^u \subseteq T\Xi$ of (L) is defined analogously. The basic theory of linear differential equations implies that solutions to (L) on E^s converge to the origin at an exponential rate, that solutions to (L) on E^u diverge from the origin at an exponential rate, and that the remaining solutions approach E^u and then diverge from the origin at an exponential rate.

Let $A^s = E^s + \xi^\dagger$ and $A^u = E^u + \xi^\dagger$ denote the affine spaces that are parallel to E^s and E^u and that pass through $\xi^\dagger$. In Sections A.1–A.3, we prove that under the BEP$(\tau, 1, \beta^{\min})$ dynamics with test-set tules τ^{all}, τ^{two}, and τ^{adj}, the dimensions of E^s and E^u are $d-1$ and 1, and that A^s is a supporting hyperplane to Ξ at $\xi^\dagger$.

Combining these facts with fundamental results from dynamical systems theory lets us complete the proof that $\xi^\dagger$ is a repellor under each dynamic. By the *Hartman-Grobman theorem* (Perko (2001, Section 2.8)), there is a homeomorphism h between a neighborhood of $\xi^\dagger$ in aff(Ξ) and a neighborhood of $\mathbf{0}$ in $T\Xi$ that maps solutions of (D) to solutions of (L). By the the *stable manifold theorem* (Perko (2001, Section 2.7)), there is an invariant *stable manifold* $M^s \subset$ aff(Ξ) of dimension $\dim(E^s) = d-1$ that is tangent to A^s at $\xi^\dagger$ such that solutions to (D) in M^s converge to $\xi^\dagger$ at an exponential rate. Combining these results shows that there is a neighborhood $O \subset$ aff(Ξ) of $\xi^\dagger$ with these properties: $O \cap \Xi \cap M^s = \{\xi^\dagger\}$; the initial conditions in O from which solutions converge exponentially quickly to $\xi^\dagger$ are those in $O \cap M^s$; and solutions from initial conditions in $(O \cap \Xi) \setminus \{\xi^\dagger\}$ eventually move away from $\xi^\dagger$. Thus the properties stated in the previous paragraph imply that state $\xi^\dagger$ is a repellor of the dynamic (D) on Ξ.

For dynamics with tie-breaking rule β^{stick}, the rest point $\xi^\dagger$ is not hyperbolic, so results from center manifold theory are needed for the local stability analysis. These are explained in Section A.4.

A.1 Test-all, min-if-tie

It is easy to verify that under the BEP$(\tau^{\text{all}}, 1, \beta^{\min})$ dynamic,

$$D\mathcal{V}(\xi^\dagger) = \left(\begin{array}{cccc|cccc}
-1 & 0 & \cdots & 0 & m+1 & 1 & \cdots & 1 \\
0 & -1 & \ddots & \vdots & 0 & 1 & \cdots & 1 \\
\vdots & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & -1 & 0 & 1 & \cdots & 1 \\
\hline
n+1 & 1 & \cdots & 1 & -1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 1 & 0 & -1 & \ddots & \vdots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 \\
0 & 1 & \cdots & 1 & 0 & \cdots & 0 & -1
\end{array}\right)$$

For $d \geq 3$, the eigenvalues of $DV(\xi^\dagger)$ with respect to $T\Xi$ and the bases for their eigenspaces are as follows:

$$\text{(29)} \qquad -1, \qquad \left\{\delta^2 - \delta^i : i \in \{3, \ldots, m+1\}\right\} \cup \left\{\varepsilon^2 - \varepsilon^j : j \in \{3, \ldots, n+1\}\right\};$$

$$\text{(30)} \qquad -1 - \sqrt{mn}, \quad \left\{(\sqrt{mn}, -\sqrt{n/m}, \ldots, -\sqrt{n/m} \,|\, -n, 1, \ldots, 1)'\right\}; \text{ and}$$

$$\text{(31)} \qquad -1 + \sqrt{mn}, \quad \left\{(-\sqrt{mn}, \sqrt{n/m}, \ldots, \sqrt{n/m} \,|\, -n, 1, \ldots, 1)'\right\}.$$

The eigenvectors in (29) and (30) span the stable subspace E^s of the linear equation (L). The normal vector to E^s is

$$z^\perp = \left(-\frac{n+1}{m+1}\sqrt{\frac{m}{n}}, \frac{n+1}{(m+1)\sqrt{mn}}, \ldots, \frac{n+1}{(m+1)\sqrt{mn}} \,\middle|\, -1, \frac{1}{n}, \ldots, \frac{1}{n}\right)'.$$

This vector satisfies

$$\text{(32)} \qquad (z^\perp)'(\delta^i - \delta^1) = \frac{n+1}{\sqrt{mn}} > 0 \text{ for } i \in S^1 \smallsetminus \{1\} \text{ and}$$

$$\text{(33)} \qquad (z^\perp)'(\varepsilon^j - \varepsilon^1) = \frac{n+1}{n} > 0 \text{ for } j \in S^2 \smallsetminus \{1\}.$$

The collection of vectors $\{\delta^i - \delta^1 : i \in S^1\} \cup \{\varepsilon^j - \varepsilon^1 : j \in S^2\}$ describes the motions along all edges of the convex set Ξ emanating from state $\xi^\dagger$. Thus the fact that their inner products with $z^\perp$ are all positive implies that the translation of E^s to $\xi^\dagger$ is a hyperplane that supports Ξ at $\xi^\dagger$. Since the remaining eigenvalue, from (31), is positive, the arguments from the start of the section allow us to conclude that $\xi^\dagger$ is a repellor.

Remark A.1. There is no loss in replacing the normal vector $z^\perp \in T\Xi$ with an auxiliary vector $z^\perp_{\text{aux}} \in \mathbb{R}^s$ whose orthogonal projection onto $T\Xi$ is a multiple of $z^\perp$. For instance, subtracting $\frac{n+1}{(m+1)\sqrt{mn}}(1, \ldots, 1 \,|\, 0, \ldots, 0)' + \frac{1}{n}(0, \ldots, 0 \,|\, 1, \ldots, 1)'$ from $z^\perp$ and multiplying the result by $\frac{n}{n+1}$ yields the simple auxiliary vector $z^\perp_{\text{aux}} = -\sqrt{\frac{n}{m}}\,\delta^1 - \varepsilon^1$. Replacing $z^\perp$ by $z^\perp_{\text{aux}}$ in (32) and (33) yields the inner products $\sqrt{\frac{n}{m}} > 0$ and $1 > 0$, respectively. To have less cumbersome expressions we will work with auxiliary vectors in what follows.

A.2 Test-two, min-if-tie

Under the BEP($\tau^{\text{two}}, 1, \beta^{\min}$) dynamic,

$$
D\mathcal{V}(\xi^{\dagger}) = \left(\begin{array}{ccccc|cccc}
0 & \frac{1}{m} & \cdots & \cdots & \frac{1}{m} & 2 & 1 & \cdots & 1 \\
0 & -\frac{1}{m} & \ddots & \ddots & \vdots & 0 & \frac{1}{m} & \cdots & \frac{1}{m} \\
0 & 0 & -\frac{2}{m} & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \frac{1}{m} & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & -\frac{m}{m} & 0 & \frac{1}{m} & \cdots & \frac{1}{m} \\
\hline
2 & 1 & \cdots & \cdots & 1 & 0 & \frac{1}{n} & \cdots & \frac{1}{n} \\
0 & \frac{1}{n} & \cdots & \cdots & \frac{1}{n} & 0 & -\frac{1}{n} & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \frac{1}{n} \\
0 & \frac{1}{n} & \cdots & \cdots & \frac{1}{n} & 0 & \cdots & 0 & -\frac{n}{n}
\end{array}\right)
$$

For $d \geq 3$, the eigenvalues and eigenvectors of $DV(\xi^{\dagger})$ with respect to $T\Xi$ are

$$(34)\qquad -\tfrac{i}{m}, \qquad \delta^{i} - \delta^{i+1} \ \ (i \in \{2, \ldots, m\});$$

$$(35)\qquad -\tfrac{j}{n}, \qquad \varepsilon^{j} - \varepsilon^{j+1} \ \ (j \in \{2, \ldots, n\}, \text{ so } n \geq 4);$$

$$(36)\qquad \lambda_{-} \equiv \tfrac{-(m+n)-\sqrt{m-n+(2mn)^{2}}}{2mn}, \qquad (a_{1}^{-}, a_{2}^{-}, \ldots, a_{m+1}^{-} \mid b_{1}^{-}, b_{2}^{-}, \ldots, b_{n+1}^{-})'; \text{ and}$$

$$(37)\qquad \lambda_{+} \equiv \tfrac{-(m+n)+\sqrt{m-n+(2mn)^{2}}}{2mn}, \qquad (a_{1}^{+}, a_{2}^{+}, \ldots, a_{m+1}^{+} \mid b_{1}^{+}, b_{2}^{+}, \ldots, b_{n+1}^{+})',$$

where letting $\star$ represent either $+$ or $-$, we have

$$(38)\qquad b_{n+1}^{\star} = 1;\ \ b_{j}^{\star} = \tfrac{n\lambda_{\star}+j+1}{n\lambda_{\star}+j-1}\, b_{j+1}^{\star}\ (j = n, \ldots, 2);\ \ b_{1}^{\star} = -n(\lambda_{\star}+1)(\lambda_{\star}+\tfrac{1}{m});$$

$$(39)\qquad a_{m+1}^{\star} = \tfrac{n}{m}(\lambda_{\star}+\tfrac{1}{m});\ \ a_{i}^{\star} = \tfrac{m\lambda_{\star}+i+1}{m\lambda_{\star}+i-1}\, a_{i+1}^{\star}\ (i = m, \ldots, 2);\ \ a_{1}^{\star} = -n(\lambda_{\star}+1).$$

When d is even, then the eigenvalues from (34) and (35) are the same, $\lambda_{-} = -\frac{m+1}{m}$, and $\lambda_{+} = \frac{m-1}{m}$. When d is odd, all eigenvalues are distinct, and λ_{-} and λ_{+} are irrational.

(The components (38) and (39) of the eigenvectors from (36) and (37) are obtained as follows: By construction $DV(\xi^{\dagger})$ maps $T\Xi$ to itself, and by definition

$$(40)\qquad \left(DV(\xi^{\dagger}) - \lambda^{\star} I\right)\begin{pmatrix} a^{\star} \\ b^{\star} \end{pmatrix} = \mathbf{0}.$$

Using the facts that $a_{1}^{\star} = -\sum_{i=2}^{m+1} a_{i}^{\star}$ and $b_{1}^{\star} = -\sum_{j=2}^{n+1} b_{j}^{\star}$, we can obtain simple expressions for the first components of each block of (40), the last components of each block, and the differences between the $(i+1)$st and ith components of the upper block ($i \in \{2, \ldots m\}$) and the $(j+1)$st and jth components of the lower block ($j \in \{2, \ldots n\}$). Each expresses a component of $b^{\star}$ as a multiple of a component of $a^{\star}$. Setting $b_{n+1}^{\star} = 1$ and applying these expressions yields (38) and (39), along with the fact that $(\lambda_\star + \frac{1}{m})(\lambda_\star + \frac{1}{n}) = 1$.)

The eigenvectors in (34)–(36) span the stable subspace E^s of the linear equation (L). The normal vector to E^s is the orthogonal projection onto $T\Xi$ of the auxiliary vector

$$z^{\perp}_{\text{aux}} = -\tfrac{m-n+\sqrt{m-n+(2mn)^2}}{2mn}\delta^1 - \varepsilon^1,$$

which satisfies

$$\begin{aligned}
&(z^{\perp})'(\delta^i - \delta^1) = (z^{\perp}_{\text{aux}})'(\delta^i - \delta^1) = \tfrac{m-n+\sqrt{m-n+(2mn)^2}}{2mn} > 0 \ \text{ for } i \in S^1 \smallsetminus \{1\}, \ \text{ and}\\
&(z^{\perp})'(\varepsilon^j - \varepsilon^1) = (z^{\perp}_{\text{aux}})'(\varepsilon^j - \varepsilon^1) = 1 > 0 \ \text{ for } j \in S^2 \smallsetminus \{1\}.
\end{aligned}$$

Thus the translation of E^s to $\xi^\dagger$ supports Ξ at $\xi^\dagger$. Since the remaining eigenvalue, from (37), is positive, the arguments from the start of the section imply that $\xi^\dagger$ is a repellor.

A.3 Test-adjacent, min-if-tie

Under the BEP($\tau^{\text{adj}}, 1, \beta^{\min}$) dynamic,

$$D\mathcal{V}(\xi^\dagger) = \left(\begin{array}{ccccc|ccccc}
0 & \frac{1}{2} & 0 & \cdots & 0 & 2 & 1 & \cdots & \cdots & 1\\
0 & -\frac{1}{2} & \ddots & \ddots & \vdots & 0 & 1 & \cdots & \cdots & 1\\
\vdots & \ddots & \ddots & \frac{1}{2} & 0 & 0 & \cdots & \cdots & \cdots & 0\\
\vdots & \ddots & \ddots & -\frac{1}{2} & 1 & \vdots & \ddots & \ddots & \ddots & \vdots\\
0 & \cdots & \cdots & 0 & -1 & 0 & \cdots & \cdots & \cdots & 0\\
\hline
2 & 1 & \cdots & \cdots & 1 & 0 & \frac{1}{2} & 0 & \cdots & 0\\
0 & 1 & \cdots & \cdots & 1 & 0 & -\frac{1}{2} & \ddots & \ddots & \vdots\\
0 & \cdots & \cdots & \cdots & 0 & \vdots & \ddots & \ddots & \frac{1}{2} & 0\\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & -\frac{1}{2} & 1\\
0 & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & -1
\end{array}\right)$$

We first consider the case in which $d = 3$, for which the eigenvalues and eigenvectors of $DV(x^\dagger, y^\dagger)$ with respect to $T\Xi$ are:

(41) $\quad -1, \quad (0, 2, -2 \,|\, -1, 1)'$;

(42) $\quad \lambda_- \equiv \frac{-3-\sqrt{17}}{4}, \quad (\lambda_- + 1, -\lambda_- - 1, 0 \,|\, 1, -1)'$; and

(43) $\quad \lambda_+ \equiv \frac{-3+\sqrt{17}}{4}, \quad (\lambda_+ + 1, -\lambda_+ - 1, 0 \,|\, 1, -1)'$.

The eigenvectors in (41) and (42) span the stable subspace E^s of the linear equation (L). The normal vector to E^s is

$$z^{\perp} = \left(-\tfrac{2+\sqrt{17}}{3}, \tfrac{\sqrt{17}-1}{6}, \tfrac{5+\sqrt{17}}{6} \mid -1, 1\right)', \text{ which satisfies}$$
$$(z^{\perp})'(\delta^2-\delta^1) = \tfrac{1+\sqrt{17}}{2},\ (z^{\perp})'(\delta^3-\delta^1) = \tfrac{3+\sqrt{17}}{2}, \text{ and } (z^{\perp})'(\varepsilon^2-\varepsilon^1) = 2,$$

so the same arguments as above imply that $\xi^\dagger$ is a repellor.

For $d \geq 4$, the eigenvalues of $DV(\xi^\dagger)$ with respect to $T\Xi$ are $\frac{1}{2}$, $-\frac{3}{2}$, -1 (with algebraic multiplicity 2), and $-\frac{1}{2}$ (with algebraic multiplicity $d-4$). In the first three cases, the bases for the corresponding eigenspaces are as follows:

$$\text{(44)} \qquad \tfrac{1}{2}, \quad \left\{\delta^2-\delta^1+\varepsilon^2-\varepsilon^1\right\};$$

$$\text{(45)} \qquad -\tfrac{3}{2}, \quad \left\{\delta^1-\delta^2+\varepsilon^2-\varepsilon^1\right\};$$

$$\text{(46)} \qquad -1, \quad \Big\{\delta^1+2\sum_{i=3}^{m}(-1)^i\delta^i+(-1)^{m+1}\delta^{m+1}-\varepsilon^1-2\sum_{j=3}^{n}(-1)^j\varepsilon^j-(-1)^{n+1}\varepsilon^{n+1};$$
$$\tfrac{1}{2}\delta^1+\delta^2+3\sum_{i=3}^{m}(-1)^i\delta^i+\tfrac{3}{2}(-1)^{m+1}\delta^{m+1}-\varepsilon^1+\varepsilon^2\Big\}.$$

The geometric multiplicity of the eigenvalue $-\frac{1}{2}$ is 1 if $d=5$ and is 2 if $d \geq 6$. Thus for $d \geq 7$, $DV(\xi^\dagger)$ is not diagonalizable. The generalized eigenvectors for the eigenvalue $-\frac{1}{2}$, which correspond to two blocks of the real Jordan matrix similar to $DV(\xi^\dagger)$, are the $m-2$ and $n-2$ columns of the following two matrices:

$$\text{(47)} \qquad \left(\begin{array}{cccccccc|ccccc}
0 & 2 & -2 & 0 & 0 & 0 & \dots & 0 & -1 & 1 & 0 & \dots & 0 \\
-1 & 1 & 4 & -4 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & -2 & 2 & 8 & -8 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & -2^{m-4} & 2^{m-4} & 2^{m-2} & -2^{m-2} & 0 & 0 & 0 & 0 & \dots & 0
\end{array}\right)' \text{ and}$$

$$\text{(48)} \qquad \left(\begin{array}{ccccccccc|cccccccccc}
\frac{1}{2} & \frac{3}{2} & -2 & 0 & 0 & 0 & 0 & \dots & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 5 & -1 & -4 & 0 & 0 & 0 & \dots & 0 & -2 & 0 & 0 & 2 & 0 & 0 & 0 & \dots & 0 \\
0 & 8 & 2 & -2 & -8 & 0 & 0 & \dots & \vdots & -4 & 0 & 0 & 0 & 4 & 0 & 0 & \dots & 0 \\
0 & 16 & 0 & 4 & -4 & -16 & 0 & \dots & 0 & -8 & 0 & 0 & 0 & 0 & 8 & 0 & \dots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 2^{n-2} & 0 & \dots & 0 & 2^{n-4} & -2^{m-4} & -2^{n-2} & 0 & -2^{n-3} & 0 & 0 & 0 & 0 & \dots & 0 & 2^{n-3} & 0
\end{array}\right)'.$$

The eigenvectors in (46)–(48) span the stable subspace E^s of the linear equation (L). The normal vector to E^s is the orthogonal projection onto $T\Xi$ of the auxiliary vector

$$z^{\perp}_{\text{aux}} = \sum_{i=1}^{m}(2^{2-m} - 3 \cdot 2^{1-i})\,\delta^i + \sum_{j=1}^{n}(2^{2-n} - 3 \cdot 2^{1-j})\,\varepsilon^j.$$

Thus $z^{\perp}$ satisfies

$$\begin{aligned}
&(z^{\perp})'(\delta^{m+1} - \delta^1) = (z^{\perp}_{\text{aux}})'(\delta^{m+1} - \delta^1) = 3 - 2^{2-m} > 0,\\
&(z^{\perp})'(\varepsilon^{n+1} - \varepsilon^1) = (z^{\perp}_{\text{aux}})'(\varepsilon^{n+1} - \varepsilon^1) = 3 - 2^{2-n} > 0,\\
&(z^{\perp})'(\delta^{i} - \delta^1) = (z^{\perp}_{\text{aux}})'(\delta^{i} - \delta^1) = 3 \cdot (1 - 2^{1-i}) > 0 \;\text{ for } i \in S^1 \smallsetminus \{1, m+1\}, \;\text{ and}\\
&(z^{\perp})'(\varepsilon^{j} - \varepsilon^1) = (z^{\perp}_{\text{aux}})'(\varepsilon^{j} - \varepsilon^1) = 3 \cdot (1 - 2^{1-j}) > 0 \;\text{ for } j \in S^2 \smallsetminus \{1, n+1\}.
\end{aligned}$$

Since the remaining eigenvalue, from (44), is positive, the arguments from the start of the section imply that $\xi^{\dagger}$ is a repellor.

A.4 Test-all, stick/min-if-tie

Under BEP dynamics based on the stick/min-if tie rule β^{stick}, the backward induction state $\xi^{\dagger}$ is not a hyperbolic rest point, and in particular has some zero eigenvalues. To show that it is nevertheless a repellor, we follow the same line of argument as above, but replace the arguments based on the Hartman-Grobman and stable manifold theorems with arguments from center manifold theory.

Consider again the dynamic (D) and its linearization at $\xi^{\dagger}$ (L) from the start of this section. When $\xi^{\dagger}$ is not a hyperbolic rest point of (D), we can define for (L) the stable subspace $E^s \subseteq T\Xi$, the unstable subspace $E^u \subseteq T\Xi$, and the *center subspace* $E^c \subseteq T\Xi$, where the last is the span of the real and imaginary parts of eigenvectors corresponding to eigenvalues with zero real part.

Let $A^{cs} = E^c \oplus E^s + \xi^{\dagger}$ be the affine space that is parallel to $E^c \oplus E^s$ and that passes through $\xi^{\dagger}$. Below we show that under the BEP($\tau^{\text{all}}, 1, \beta^{\text{stick}}$) dynamic, the subspace $E^c \oplus E^s$ has dimension $d - 1$, and the affine space A^{cs} is a supporting hyperplane to Ξ at $\xi^{\dagger}$. We prove the corresponding results for dynamics based on test-set rules τ^{two} and τ^{adj} in Online Appendix V.

Linearization is much less simple for nonhyperbolic rest points than for hyperbolic ones—see Perko (2001). However, for our purposes, it is enough that there exists a (local) *center-stable manifold* M^{cs} that is tangent to A^{cs}, and is invariant under (D) (Kelley (1967b)). This manifold need not be unique; see Kelley (1967b, Section 4) for an example.[45]

[45] Under the BEP($\tau^{\text{two}}, 1, \beta^{\text{stick}}$) dynamic, the affine set through $\xi^{\dagger}$ defined by the eigenvectors with zero eigenvalue consists entirely of rest points; see footnote 37 and Online Appendix V.2. This fact and Corollary 3.3 of Sijbrand (1985) imply that this affine set is the unique center manifold through $\xi^{\dagger}$.

But for any choice of center-stable manifold M^{cs}, there is a neighborhood $O \subset \text{aff}(\Xi)$ of $\xi^\dagger$ satisfying $O \cap \Xi \cap M^{cs} = \{\xi^\dagger\}$ such that solutions to (D) from initial conditions in $(O \cap \Xi) \setminus \{\xi^\dagger\}$ eventually move away from $\xi^\dagger$; see Kelley (1967a, p. 336), or see Perko (2001, Theorem 2.12.2) for a closely related and more explicitly presented result. This and the properties from the previous paragraph imply that $\xi^\dagger$ is a repellor of the BEP($\tau, 1, \beta^{\text{stick}}$) dynamics on Ξ.

Turning to the computation, we observe that under the BEP($\tau^{\text{all}}, 1, \beta^{\text{stick}}$) dynamic,

$$D\mathcal{V}(\xi^\dagger) = \left(\begin{array}{cccc|cccc} -1 & 0 & \cdots & 0 & m+1 & 1 & \cdots & 1 \\ 0 & -1 & \ddots & \vdots & 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 0 & 1 & \cdots & 1 \\ \hline n+1 & 1 & \cdots & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & 1 & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 1 & \cdots & 1 & 0 & \cdots & \cdots & 0 \end{array}\right)$$

The eigenvalues of $DV(\xi^\dagger)$ and bases for their eigenspaces are

(49) $\quad 0, \qquad \left\{\varepsilon^2 - \varepsilon^j \colon j \in \{3, \ldots, n+1\}\right\};$

(50) $\quad -1, \qquad \left\{\delta^2 - \delta^i \colon i \in \{3, \ldots, m+1\}\right\};$

(51) $\quad \lambda_- \equiv -\frac{1}{2} - \sqrt{mn + \frac{1}{4}}, \qquad \left\{(-\lambda_-, \frac{\lambda_-}{m}, \ldots, \frac{\lambda_-}{m} \,\middle|\, -n, 1, \ldots, 1)'\right\};$ and

(52) $\quad \lambda_+ \equiv -\frac{1}{2} + \sqrt{mn + \frac{1}{4}}, \qquad \left\{(-\lambda_+, \frac{\lambda_+}{m}, \ldots, \frac{\lambda_+}{m} \,\middle|\, -n, 1, \ldots, 1)'\right\}.$

The eigenvectors in (49) span the center subspace E^c of the linear equation (L), while the eigenvectors in (50) and (51) span the stable subspace E^s. The normal vector to the hyperplane $E^c \oplus E^s$ is the orthogonal projection onto $T\Xi$ of the auxiliary vector

$$z^\perp_{\text{aux}} = -\sqrt{\frac{n}{m}}\,\delta^1 - \varepsilon^1,$$

which satisfies

$$(z^\perp)'(\delta^i - \delta^1) = (z^\perp_{\text{aux}})'(\delta^i - \delta^1) = \sqrt{\tfrac{n}{m}} > 0 \;\text{ for } i \in S^1 \setminus \{1\} \;\text{ and}$$
$$(z^\perp)'(\varepsilon^j - \varepsilon^1) = (z^\perp_{\text{aux}})'(\varepsilon^i - \varepsilon^1) = 1 > 0 \;\text{ for } j \in S^2 \setminus \{1\}.$$

Since the remaining eigenvalue, from (52), is positive, the arguments above imply that $\xi^\dagger$ is a repellor.

B. Local stability analysis of interior rest points

The interior rest point ξ^* of the dynamic $\dot{x} = V(x)$ is locally stable if all eigenvalues of the derivative matrix $DV(\xi^*)$ have negative real part. Since each entry of the derivative matrix $D\mathcal{V}(\xi^*)$ is a polynomial with many terms that is evaluated at state whose components are algebraic numbers, it is not feasible to compute its eigenvalues exactly. We circumvent this problem by computing the eigenvalues of the derivative matrix at a nearby rational state ξ, and making use of a bound on the distances between the eigenvalues of the two matrices. This bound is established in Proposition B.1, and a more quickly computable alternative is provided in Proposition III.4 of the online appendix.

As in Appendix A, let $s = s^1 + s^2 = d + 2$, let $\dot{\xi} = V(\xi)$, $V\colon \operatorname{aff}(\Xi) \to T\Xi$ denote an instance of the BEP dynamics (B), and let $\mathcal{V}\colon \mathbb{R}^s \to \mathbb{R}^s$ denote the natural extension of V to $\mathbb{R}^s$. Observe that if $D\mathcal{V}(\xi)$ is diagonalizable, then so is $DV(\xi)$, and all eigenvalues of the latter are eigenvalues of the former. To state the proposition, we write $S = S^1 \cup S^2$ and omit population superscripts to define

$$(53) \qquad \Delta = \max_{i\in S} \max_{k\in S} \sum_{j\in S} \frac{\partial^2 \mathcal{V}_i}{\partial \xi_j \partial \xi_k}(1,\ldots,1\,|\,1,\ldots,1).$$

Proposition B.1. *Suppose that $D\mathcal{V}(\xi)$ is (complex) diagonalizable with $D\mathcal{V}(\xi) = Q\operatorname{diag}(\lambda)Q^{-1}$, and let λ^* be an eigenvalue of $D\mathcal{V}(\xi^*)$. Then there is an eigenvalue λ_i of $D\mathcal{V}(\xi)$ such that*

$$(54) \qquad \left|\lambda^* - \lambda_i\right| < \frac{2\Delta}{s^{s/2-1}} \frac{\operatorname{tr}(Q^*Q)^{s/2}}{|\det(Q)|} \sum_{k\in S} |\xi_k - \xi_k^*|.$$

The eigenvalue perturbation theorem (56) that begins the proof of the proposition bounds the distances between the eigenvalues of $D\mathcal{V}(\xi^*)$ and $D\mathcal{V}(\xi)$, but neither term on its right-hand side is feasible to compute. The second paragraph of the proof provides a bound on the condition number $\kappa_\infty(Q)$ that does not require the computation of the inverse of the (algebraic-valued) eigenvector matrix Q. The third paragraph provides a bound on the norm of $D\mathcal{V}(\xi) - D\mathcal{V}(\xi^*)$, which is needed because numerically evaluating of the entries of $D\mathcal{V}(\xi^*)$ with guaranteed precision is computationally infeasible. Two further devices that we employ to improve the bound and speed its computation are described after the proof of the proposition.

Proof. For $M \in \mathbb{R}^{s \times s}$, let

$$\vert\vert\vert M \vert\vert\vert_\infty = \max_{1 \le i \le s} \sum_{j=1}^{s} |M_{ij}|. \tag{55}$$

denote the maximum row sum norm of M. Let $\kappa_\infty(Q) = \vert\vert\vert Q\vert\vert\vert_\infty \, \vert\vert\vert Q^{-1}\vert\vert\vert_\infty$ be the condition number of Q with respect to norm (55). The following eigenvalue perturbation theorem (Horn and Johnson (2013, Observation 6.3.1)) follows from the Geršgorin disk theorem and the submultiplicativity of matrix norms:

$$\left|\lambda^* - \lambda_i\right| \le \kappa_\infty(Q) \, \vert\vert\vert D\mathcal{V}(\xi) - D\mathcal{V}(\xi^*)\vert\vert\vert_\infty. \tag{56}$$

To bound $\kappa_\infty(Q)$, let $\vert\vert\vert M\vert\vert\vert_2$ denote the spectral norm of M (i.e., the largest singular value of M), and let $\kappa_2(Q) = \vert\vert\vert Q\vert\vert\vert_2 \, \vert\vert\vert Q^{-1}\vert\vert\vert_2$ be the condition number of Q with respect to this norm. Since the maximum row sum and spectral norms differ by a factor of at most $\sqrt{s}$ (Horn and Johnson (2013, Problem 5.6.P23)), it follows that

$$\kappa_\infty(Q) \le s\kappa_2(Q). \tag{57}$$

Also, Guggenheimer et al. (1995) (see also Merikoski et al. (1997)) show that

$$\kappa_2(Q) < \frac{2}{|\det(Q)|} \left(\frac{\operatorname{tr}(Q^* Q)}{s}\right)^{s/2}. \tag{58}$$

To bound the final expression in (56), recall that each component of expression (B) for the BEP dynamics $\dot{x} = \mathcal{V}(x)$ is the difference between a sum of monomials in the components of ξ with positive coefficients and a linear term. Thus the second derivatives of $\mathcal{V}_i(\xi)$ are sums of monomials with positive coefficients. Since every component of every state $\xi \in \Xi$ is at most 1, we therefore have

$$\max_{\xi \in \Xi} \left| \frac{\partial^2 \mathcal{V}_i}{\partial \xi_j \partial \xi_k}(\xi) \right| \le \frac{\partial^2 \mathcal{V}_i}{\partial \xi_j \partial \xi_k}(1, \ldots, 1 \,|\, 1, \ldots, 1). \tag{59}$$

Thus the fundamental theorem of calculus, (59), and (53) imply that

$$\begin{aligned} \vert\vert\vert D\mathcal{V}(\xi) - D\mathcal{V}(\xi^*)\vert\vert\vert_\infty &\le \max_{i \in S} \sum_{j \in S} \sum_{k \in S} \frac{\partial^2 \mathcal{V}_i}{\partial \xi_j \partial \xi_k}(1, \ldots, 1 \,|\, 1, \ldots, 1) \times |\xi_k - \xi_k^*| \\ &\le \Delta \sum_{k \in S} |\xi_k - \xi_k^*|. \end{aligned} \tag{60}$$

Combining inequalities (56), (57), (58), and (60) yields inequality (54). ■

When applying Proposition B.1, one can choose Q to be any matrix of eigenvectors of $D\mathcal{V}(\xi)$. Guggenheimer et al. (1995) suggest that choosing the eigenvectors to have Euclidean norm 1 (which if done exactly makes the expression in parentheses in (58) equal 1) leads to the lowest bounds. We apply this normalization in the final step of our analysis. Finally, recalling that $s = d + 2$, we can speed up calculations substantially by replacing the $s \times s$ matrix $D\mathcal{V}(\xi^*)$ with a $d \times d$ matrix with the same eigenvalues as $DV(\xi)\colon T\Xi \to T\Xi$. We accomplish this by adapting a dimension-reduction method from Sandholm (2007). See Proposition III.4 in the online appendix for the bound's final form.

To use this bound to establish the stability of the interior rest point ξ^*, we choose a rational point ξ close to ξ^*, compute the eigenvalues of the derivative matrix $DV(\xi)$, and evaluate the bound from Proposition III.4. The eigenvalues of $DV(\xi)$ all have negative real part so long as ξ is reasonably close to ξ^*.[46] If ξ is close enough to ξ^* that the bound is smaller than the magnitude of the real part of any eigenvalue of $DV(\xi)$, we can conclude that the eigenvalues of $DV(\xi^*)$ all have negative real part, and hence that ξ^* is asymptotically stable.

Selecting state ξ involves a tradeoff: choosing ξ closer to ξ^* reduces the bound, but doing so also leads the components of ξ to have larger numerators and denominators, which slows the computation of the bound significantly. In all cases, we are able to choose ξ satisfactorily and to conclude that ξ^* is asymptotically stable. For further details about how the computations are implemented, see the online appendix.

C. Inward motion from the boundary

The following differential inequality will allow us to obtain simple lower bounds on the use of initially unused strategies. In all cases in which we apply the lemma, $v(0) = 0$.

Lemma C.1. *Let $v\colon [0,T] \to \mathbb{R}_+$ satisfy $\dot{v}(t) \geq a(t) - v(t)$ for some $a\colon [0,T] \to \mathbb{R}_+$. Then*

$$(61) \qquad v(t) \geq \mathrm{e}^{-t}\left(v(0) + \int_0^t \mathrm{e}^s\, a(s)\,\mathrm{d}s\right) \text{ for all } t \in [0,T].$$

Proof. Clearly $v(t) = v(0) + \int_0^t \dot{v}(s)\,\mathrm{d}s \geq v(0) + \int_0^t (a(s) - v(s))\,\mathrm{d}s$. The final expression is the time t value of the solution to the differential equation $\dot{v}(s) + v(s) = a(s)$ with initial

[46] Among the cases in which we solve for ξ^* exactly, the smallest magnitude of a real part of an eigenvalue occurs under the BEP($\tau^{\text{adj}}, 1, \beta^{\min}$) dynamic when $d = 7$, in which case there is an eigenvalue approximately equal to $-.1125$. See Online Appendix IX.

condition $v(0)$. Using the integrating factor e^s to solve this equation yields the right-hand side of (61). ■

For the analysis to come, it will be convenient to work with the set $\mathcal{S} = S^1 \cup S^2$ of all strategies from both populations, and to drop population superscripts from notation related to the state—for instance, writing ξ_i rather than ξ_i^p.

We use Lemma C.1 to prove inward motion from the boundary under BEP dynamics in the following way. Write $\dot{\xi}_i = r_i(\xi) - \xi_i$, where $r_i(\xi)$ is the polynomial appearing in the formula (B) for the BEP$(\tau, 1, \beta)$ dynamic. Let $\{\xi(t)\}_{t\geq 0}$ be the solution of (B) with initial condition $\xi(0)$. Let $\mathcal{S}_0 = \mathrm{supp}(\xi(0))$ and $Q = \min\{\xi_h(0)\colon h \in \mathcal{S}_0\}$, and, finally, let $\mathcal{S}_1 = \{i \in \mathcal{S} \smallsetminus \mathcal{S}_0 \colon r_i(\xi(0)) > 0\}$ and $R = \frac{1}{2}\min\{r_k(\xi(0))\colon r_k(\xi(0)) > 0\}$.

By the continuity of (B), there is a neighborhood $O \subset \Xi$ of $\xi(0)$ such that every $\chi \in O$ satisfies $\chi_h > Q$ for all $h \in \mathcal{S}_0$ and $\dot{\chi}_i \geq R$ for all $i \in \mathcal{S}_1$. And since (B) is smooth, there is a time $T > 0$ such that $\xi(t) \in O$ for all $t \in [0, T]$. Thus applying Lemma C.1 shows that

$$\xi_i(t) \geq R(1 - \mathrm{e}^{-t}) \text{ for all } t \in [0, T] \text{ and } i \in \mathcal{S}_1. \tag{62}$$

Now let $\mathcal{S}_2$ be the set of $j \notin \mathcal{S}_0 \cup \mathcal{S}_1$ for which there is a term of polynomial r_j whose factors all correspond to elements of $\mathcal{S}_0$ or $\mathcal{S}_1$. If this term has a factors in $\mathcal{S}_0$, b factors in $\mathcal{S}_1$, and coefficient c, then the foregoing claims and Lemma C.1 imply that

$$\xi_j(t) \geq c\, Q^a\, \mathrm{e}^{-t} \int_0^t \mathrm{e}^s (R(1 - \mathrm{e}^{-s}))^b\, \mathrm{d}s \text{ for all } t \in [0, T]. \tag{63}$$

Proceeding sequentially, we can obtain positive lower bounds on the use of any strategy for times $t \in (0, T]$ by considering as-yet-unconsidered strategies k whose polynomials r_k have a term whose factors all correspond to strategies for which lower bounds have already been obtained. Below, we prove that solutions to the BEP$(\tau^{\mathrm{all}}, 1, \beta^{\mathrm{min}})$ dynamic from states $\xi(0) \neq \xi^\dagger$ immediately enter $\mathrm{int}(\Xi)$ by showing that the strategies in $\mathcal{S} \smallsetminus \mathcal{S}_0$ can be considered in a sequence that satisfies the property just stated. The proofs for other BEP dynamics are described in the remarks that follow.

To proceed, we use the notations $i^{[1]}$ and $i^{[2]}$ to denote the ith strategies of players 1 and 2. We also introduce the linear order $\prec$ on $\mathcal{S}$ defined by $1^{[1]} \prec 1^{[2]} \prec 2^{[1]} \prec 2^{[2]} \prec 3^{[1]} \prec \ldots$, which arranges the strategies according to how early they stop play in Centipede.

Proof of Proposition 5.5. Fix an initial condition $\xi(0) \neq \xi^\dagger$. We can sequentially add all strategies in $\mathcal{S} \smallsetminus \mathcal{S}_0$ in accordance with the property above as follows:

(I) First, we add the strategies $\{i \in \mathcal{S} \smallsetminus \mathcal{S}_0 \colon i \prec \max \mathcal{S}_0\}$ in decreasing order. At the point that i has been added, i's successor h has already been added, and strategy i is the unique best response when the revising agent tests all strategies against opponents playing h. Let $\mathcal{S}_{\mathrm{I}}$ denote the set of strategies added during this stage. The assumption that $\xi(0) \neq \xi^{\dagger}$ implies that $\mathcal{S}_0 \cup \mathcal{S}_{\mathrm{I}}$ contains $1^{[1]}$, $1^{[2]}$, and $2^{[1]}$.

(II) Second, we add the strategies $j \in S^2 \smallsetminus (\mathcal{S}_0 \cup \mathcal{S}_{\mathrm{I}})$. We can do so because j is the unique best response when it is tested against $2^{[1]}$ and all other strategies are tested against $1^{[1]}$.

(III) Third, we add the strategies $k \in S^1 \smallsetminus (\mathcal{S}_0 \cup \mathcal{S}_{\mathrm{I}})$. We can do so because k is the unique best response when it is tested against $2^{[2]}$ and other strategies are tested against $1^{[2]}$. ■

Remark C.2. (i) Since the preceding argument only used unique best responses, it applies equally well under any tie-breaking rule. In the case of the uniform-if-tie rule β^{unif}, a simple argument along the lines above also establishes motion into the interior from the backward induction state $\xi^{\dagger}$.

(ii) Likewise, while Proposition 5.5 concerns the test-all rule τ^{all}, the analysis above works equally well for the test-two rule τ^{two}.

(iii) To establish entry into $\mathrm{int}(\Xi)$ under the test-adjacent rule τ^{adj}, we must start with the following preliminary stage: Let $m = \max \mathcal{S}_0$ be a population p strategy, let $q \neq p$, and let $\ell = \max \mathcal{S}_0 \cap S^q$. In the preliminary stage, we sequentially add strategies in the set $\{i \in S^p \smallsetminus \mathcal{S}_0 \colon i \prec m\}$ in decreasing order, stopping once we add the successor of ℓ. We can add each such strategy h because player p's succeeding strategy is either in $\mathcal{S}_0$ or has already been added, so that agents playing that strategy can have strategy h in their test set; and because the revising agent will prefer h to the succeeding strategy if both are tested against opponents playing $\ell \prec h$. After these strategies are added, stage (I) can proceed, but replacing $\max \mathcal{S}_0$ with ℓ, and stages (II) and (III) work as before provided that the strategies added in these stages are added in increasing order.

References

Akritas, A. G. (2010). Vincent's theorem of 1836: Overview and future research. *Journal of Mathematical Sciences*, 168:309–325.

Becker, E., Marinari, M. G., Mora, T., and Traverso, C. (1994). The shape of the Shape Lemma. In von zur Gathen, J. and Giesbrecht, M., editors, *ISSAC '94: Proceedings of the international symposium on symbolic and algebraic computation*, pages 129–133. ACM.

Ben-Porath, E. (1997). Rationality, Nash equilibrium and backwards induction in perfect-information games. *Review of Economic Studies*, 64:23–46.

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

Binmore, K. (1987). Modeling rational players. *Economics and Philosophy*, 3-4:179–214 and 9–55.

Binmore, K. (1998). *Game Theory and the Social Contract, Volume 2: Just Playing*. MIT Press, Cambridge.

Björnerstedt, J. and Weibull, J. W. (1996). Nash equilibrium and evolution by imitation. In Arrow, K. J. et al., editors, *The Rational Foundations of Economic Behavior*, pages 155–181. St. Martin's Press, New York.

Blume, L. E. (1997). Population games. In Arthur, W. B., Durlauf, S. N., and Lane, D. A., editors, *The Economy as an Evolving Complex System II*, pages 425–460. Addison-Wesley, Reading, MA.

Brandenburger, A., Friedenberg, A., and Keisler, H. J. (2008). Admissibility in games. *Econometrica*, 76:307–352.

Buchberger, B. (1965). *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal*. PhD thesis, University of Innsbruck. Translated by M.P. Abramson as "An algorithm for finding the basis elements of the residue class ring of a zero-dimensional polynomial ideal" in *Journal of Symbolic Computation* 41 (2006), 475–511.

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

Cohen, H. (1993). *A Course in Computational Algebraic Number Theory*. Springer, Berlin.

Collins, G. E. (1975). Quantifier elimination for the theory of real closed fields by cylindrical algebraic decomposition. In *Second GI Conference on Automata Theory and Formal Languages*, volume 33 of *Lecture Notes in Computer Science*, pages 134–183. Springer, Berlin.

Cox, D., Little, J., and O'Shea, D. (2015). *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer International, Cham, Switzerland, fourth edition.

Cressman, R. (1996). Evolutionary stability in the finitely repeated prisoner's dilemma game. *Journal of Economic Theory*, 68:234–248.

Cressman, R. (2003). *Evolutionary Dynamics and Extensive Form Games*. MIT Press, Cambridge.

Cressman, R. and Schlag, K. H. (1998). On the dynamic (in)stability of backwards induction. *Journal of Economic Theory*, 83:260–285.

Dekel, E. and Gul, F. (1997). Rationality and knowledge in game theory. In Kreps, D. M. and Wallis, K. F., editors, *Advances in Economics and Econometrics: Theory and Applications*, volume 1, pages 87–172. Cambridge University Press, Cambridge.

Droste, E., Kosfeld, M., and Voorneveld, M. (2003). Best-reply matching in games. *Mathematical Social Sciences*, 46:291–309.

Dummit, D. S. and Foote, R. M. (2004). *Abstract Algebra*. Wiley, Hoboken, NJ, third edition.

Fudenberg, D. and Levine, D. K. (1998). *The Theory of Learning in Games*. MIT Press, Cambridge.

Gilboa, I. and Matsui, A. (1991). Social stability and equilibrium. *Econometrica*, 59:859–867.

Guggenheimer, H. W., Edelman, A. S., and Johnson, C. R. (1995). A simple estimate of the condition number of a linear system. *College Mathematics Journal*, 26:2–5.

Halpern, J. Y. (2001). Substantive rationality and backward induction. *Games and Economic Behavior*, 37:425–435.

Hofbauer, J. (1995a). Imitation dynamics for games. Unpublished manuscript, University of Vienna.

Hofbauer, J. (1995b). Stability for the best response dynamics. Unpublished manuscript, University of Vienna.

Hofbauer, J. and Sandholm, W. H. (2002). On the global convergence of stochastic fictitious play. *Econometrica*, 70:2265–2294.

Hofbauer, J. and Sandholm, W. H. (2007). Evolution in games with randomly disturbed payoffs. *Journal of Economic Theory*, 132:47–69.

Hofbauer, J. and Sandholm, W. H. (2011). Survival of dominated strategies under evolutionary dynamics. *Theoretical Economics*, 6:341–377.

Horn, R. A. and Johnson, C. R. (2013). *Matrix Analysis*. Cambridge University Press, New York, second edition.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123:81–104.

Kelley, A. (1967a). Stability of the center-stable manifold. *Journal of Mathematical Analysis and Applications*, 18:336–344.

Kelley, A. (1967b). The stable, center-stable, center, center-unstable, and unstable manifolds. *Journal of Differential Equations*, 3:546–570.

Kosfeld, M., Droste, E., and Voorneveld, M. (2002). A myopic adjustment process leading to best reply matching. *Journal of Economic Theory*, 40:270–298.

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

Kreps, D. M., Milgrom, P., Roberts, J., and Wilson, R. (1982). Rational cooperation in the finitely repeated Prisoner's Dilemma. *Journal of Economic Theory*, 27:245–252.

Kubler, F., Renner, P., and Schmedders, K. (2014). Computing all solutions to polynomial equations in economics. In Schmedders, K. and Judd, K. L., editors, *Handbook of Computational Economics*, volume 3, pages 599–652. Elsevier, Amsterdam.

Mantilla, C., Sethi, R., and Cárdenas, J. C. (2017). Efficiency and stability of sampling equilibrium in public good games. Unpublished manuscript, Universidad del Rosario, Columbia University, and Universidad de Los Andes.

McKelvey, R. D. and Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica*, 60:803–836.

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

McKelvey, R. D. and Palfrey, T. R. (1998). Quantal response equilibria for extensive form games. *Experimental Economics*, 1:9–41.

McNamee, J. M. (2007). *Numerical Methods for Roots of Polynomials, Part I*. Elsevier, Amsterdam.

Merikoski, J. K., Urpala, U., Virtanen, A., Tam, T.-Y., and Uhlig, F. (1997). A best upper bound for the 2-norm condition number of a matrix. *Linear Algebra and its Applications*, 254:355–365.

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

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

Perea, A. (2014). Belief in the opponents' future rationality. *Games and Economic Behavior*, 83:231–254.

Perko, L. (2001). *Differential Equations and Dynamical Systems*. Springer, New York, third edition.

Ponti, G. (2000). Cycles of learning in the Centipede game. *Games and Economic Behavior*, 30:115–141.

Radner, R. (1980). Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. *Journal of Economic Theory*, 22:136–154.

Reny, P. J. (1992). Backward induction, normal form perfection, and explicable equilibria. *Econometrica*, 60:627–649.

Robson, A. and Vega-Redondo, F. (1996). Efficient equilibrium selection in evolutionary games with random matching. *Journal of Economic Theory*, 70:65–92.

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

Samuelson, L. (1992). Dominated strategies and common knowledge. *Games and Economic Behavior*, 4:284–313.

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

Sandholm, W. H. (2003). Evolution and equilibrium under inexact information. *Games and Economic Behavior*, 44:343–378.

Sandholm, W. H. (2007). Evolution in Bayesian games II: Stability of purified equilibria. *Journal of Economic Theory*, 136:641–667.

Sandholm, W. H. (2010a). Pairwise comparison dynamics and evolutionary foundations for Nash equilibrium. *Games*, 1:3–17.

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

Sandholm, W. H. (2015). Population games and deterministic evolutionary dynamics. In Young, H. P. and Zamir, S., editors, *Handbook of Game Theory*, volume 4, pages 703–778. Elsevier, Amsterdam.

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

Sijbrand, J. (1985). Properties of center manifolds. *Transactions of the American Mathematical Society*, 289:431–469.

Stalnaker, R. (1996). Knowledge, belief, and counterfactual reasoning in games. *Economics and Philosophy*, 12:133–163.

Taylor, P. D. and Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156.

von zur Gathen, J. and Gerhard, J. (2013). *Modern Computer Algebra*. Cambridge University Press, Cambridge, third edition.

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

Xu, Z. (2016). Convergence of best-response dynamics in extensive-form games. *Journal of Economic Theory*, 162:21–54.

Young, H. P. (1993). The evolution of conventions. *Econometrica*, 61:57–84.

Young, H. P. (1998). *Individual Strategy and Social Structure*. Princeton University Press, Princeton.