Best Experienced Payoff Dynamics and Cooperation in the Centipede Game*

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

March 7, 2019

Abstract

We study population game dynamics under which each revising agent tests each of his strategies 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. In the Centipede game, these *best experienced payoff dynamics* lead to cooperative play. When strategies are tested once, 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 largely independent of the length of the game. Testing strategies many times leads to cyclical play.

*We thank Dan Friedman, Drew Fudenberg, Ken Judd, Panayotis Mertikopoulos, Erik Mohlin, Ignacio Monzón, Arthur Robson, Marzena Rostek, Ariel Rubinstein, Larry Samuelson, Ryoji Sawa, Andy Skrzypacz, 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 (W911NF-17-1-0134 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`.

‡BioEcoUva Research Institute on Bioeconomy. Department of Industrial Organization, Universidad de Valladolid, Paseo del Cauce 59, 47011 Valladolid, Spain. e-mail: `segis@eii.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 plays each of his strategies 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

[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).

to the strategy that achieved the highest total payoff, breaking ties in favor of the lowest-numbered strategy. 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)). We study the properties of this differential equation when agents play the Centipede game.

Our model builds on earlier work on games played by "procedurally rational players". If we replaced our tie-breaking rule with uniform tie-breaking, then the rest points of the process (with $\kappa = k$) would correspond to the $S(k)$ equilibria of Osborne and Rubinstein (1998). The corresponding dynamics were studied by Sethi (2000). These and other dynamics are instances of the broader family of *best experienced payoff dynamics*, or *BEP dynamics* for short (Sandholm et al. (2019)), which allow for variation in how ties are resolved and in the selection of sets of candidate strategies considered by revising agents. The results we present here are robust to many different model specifications within the family of BEP dynamics.

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 in Centipede games with up to six decision nodes. We complement this approach with numerical analyses of cases in which analytical results cannot be obtained.

Our initial 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 the BEP dynamics studied here, 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.

We next obtain strong lower bounds on the total weight placed on cooperative strategies at any other rest points of the BEP dynamic. At any such rest point, the probability that play during a random match leads to one of the last five terminal nodes is above .96, and the probability that play leads to one of the last seven terminal nodes is virtually 1. We then use tools from computational algebra to perform an exact analysis of games with up to six decision nodes, and we perform numerical analyses of longer games. In all cases, we find that besides the unstable backward induction state, the dynamics have exactly one other rest point.[3] The form of this rest point is essentially independent of the length of the game. The rest point has virtually all players choosing to continue until the last few nodes of the game. Moreover, this rest point is dynamically stable, attracting solutions from all initial conditions other than the backward induction state. Thus if agents make choices based on experienced payoffs, testing each strategy 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 of his strategies 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.[4] 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.

Our final results consider the effects of the number of trials κ of each strategy during testing on predictions of play. It seems clear that if the number of trials is made sufficiently 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 (1995)), 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. To start, we prove that the backward induction state is unstable as long as the number of trials κ is less than the length of the game. For larger number of trials, the backward induction state becomes locally stable, but numerical ev-

[3]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 2.2.

[4]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 3.3).

idence suggests that its basin of attraction is very small. Examining Centipede games of length $d = 4$ in detail, we find that a unique, attracting interior rest point with substantial cooperation persists for moderate numbers of trials. With many trials, numerical analysis suggests the attractor is always a single cycle which includes significant amounts of cooperation for numbers of trials as large as 200. We discuss in Section 4 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 prove 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 zeros 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 zeros as the original set, but its zeros can be computed by finding the roots of a single (possibly high-degree) univariate polynomial.[5] 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.[6] 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.

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

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 (1995). Cressman and Schlag (1998)

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

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

(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 for small and moderate numbers of tests, best experienced payoff dynamics lead to a stable distribution of cooperative strategies far from the Nash component.

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.[7] 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 2.1), 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.[8] 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. Berkemer (2008) considers the local stability of the unique rationalizable strategy profile in the traveler's dilemma of Basu (1994) under Sethi's (2000) dynamics, obtaining a sufficient condition for the instability of the rationalizable state. He shows numerically that the stable $S(1)$ equilibrium becomes independent of the number of strategies in the game, and provides evidence from agent-based simulations that larger numbers of trials during testing can lead to cyclical behavior.[9]

Earlier efforts to explain cooperative behavior in Centipede and related games have followed a different approach, applying equilibrium analyses to augmented versions of

[7]For extensions of $S(k)$ equilibrium to more complex testing procedures, see Rustichini (2003).

[8]Cárdenas et al. (2015) and Mantilla et al. (2019) use these dynamics to explain stable non-Nash behavior in public good games.

[9]For complementary models of dynamics based on a single sample, see Sandholm (2001), Kosfeld et al. (2002), Droste et al. (2003), and Oyama et al. (2015).

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.[10] 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.

2. Best experienced payoff dynamics in the Centipede Game

2.1 Normal form games and population games

A two-player normal form game $G = \{(S^1, S^2), (A, B)\}$ is defined by pairs of strategy sets $S^p = \{1, \ldots, s^p\}$ and payoff matrices $A, B \in \mathbb{R}^{s^p \times s^q}$, $p, q \in \{1, 2\}$, $p \neq q$. A_{ij} and B_{ij} represent the two players' payoffs when strategy profile $(i, j) \in S^1 \times S^2$ is played. 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* for population 1 is an element of $X = \{x \in \mathbb{R}^{s^1}_+ \colon \sum_{i \in S^1} x_i = 1\}$, where x_i is the fraction of population 1 players choosing strategy i. Likewise $Y = \{y \in \mathbb{R}^{s^2}_+ \colon \sum_{i \in S^2} y_i = 1\}$ is the set of population states for population 2. Thus x and y are formally equivalent to mixed strategies for players 1 and 2, and elements of the set $\Xi = X \times Y$ are formally equivalent to mixed strategy profiles. In a slight abuse of terminology, we also refer to elements of Ξ as population states.

[10]McKelvey and Palfrey (1992) show that this analysis extends to the Centipede game. 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, following Radner (1980), one can consider versions of Centipede in which the stakes of each move are small, and analyze these games using ε-equilibrium; see Friedman and Oprea (2012) for a discussion. 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.

2.2 Revision protocols and evolutionary dynamics

To define evolutionary game dynamics, we follow the standard approach of specifying microfoundations in terms of revision protocols.[11] We suppose that at all times $t \in [0, \infty)$, each agent has a strategy he uses when matched to play game G. The empirical distributions of these strategies are described by the population state $\xi(t) = (x(t), y(t))$.

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 1 is described by a map $(A, y) \mapsto \sigma^1(A, y) \in X^{s^1}$ that assigns own payoff matrices and opposing population states to matrices of *conditional switch probabilities*, where $\sigma^1_{ij}(A, y)$ is the probability that an agent playing strategy $i \in S^1$ who receives a revision opportunity switches to strategy $j \in S^1$. Likewise, a revision protocol for population 2 is described by a map $(B, x) \mapsto \sigma^2(B, x) \in Y^{s^2}$ with an analogous interpretation.[12]

It is well known that if the population sizes are large, the Markov process implicitly defined by the above procedure is well approximated by solutions to a differential equation defined by the expected motion of the process (Benaïm and Weibull (2003)). Here this differential equation takes the form

$$
\begin{aligned}
\dot{x}_i &= \sum_{j \in S^1} x_j \sigma^1_{ji}(A, y) - x_i \quad \text{for all } i \in S^1, \\
\dot{y}_i &= \sum_{j \in S^2} y_j \sigma^2_{ji}(B, x) - y_i \quad \text{for all } i \in S^2.
\end{aligned}
\tag{1}
$$

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

Outside of monomorphic (i.e. pure) cases, the rest points of the dynamic (1) 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 equaling zero.[13] At states that are locally stable

[11] See Björnerstedt and Weibull (1996), Weibull (1995), Sandholm (2010a,b, 2015), and Izquierdo et al. (2019).

[12] When σ^1_{ij} and σ^2_{ij} are independent of the current strategy i, as is true for the dynamic (3) we focus on here, it is equivalent to interpret the process as one in which agents play a fixed strategy until leaving the population, when they are replaced by new agents whose strategies are determined by applying σ^1 and σ^2.

[13] 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).

under the dynamic (1), fluctuations in any direction are generally undone by the action of (1) itself. Contrariwise, fluctuations away from unstable equilibria are reinforced, so we should not expect such states to be observed.

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* τ, a *number of trials* κ, and a *tie-breaking rule* β. 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^p on the collection of subsets of S^p with at least two elements. 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. The triple (τ, κ, β) thus defines a revision protocol σ^p for each population p. Inserting these revision protocols into equation (1) defines a *best experienced payoff dynamic*.

Our analysis here focuses on the test-set rule *test-all*, τ^{all}, under which a revising agent tests all of his strategies, and on the tie-breaking rule *min-if-tie*, $\beta^{\min}$, which chooses the lowest-numbered optimal strategy. We refer to the resulting dynamics (1) as *BEP*$(\tau^{\text{all}}, \kappa, \beta^{\min})$ *dynamics*.

BEP dynamics based on other specifications of test-set and tie-breaking rules are studied in a companion paper, Sandholm et al. (2019); they are also discussed briefly in Section 3.4. Importantly, if we retain τ^{all}, but replace $\beta^{\min}$ with uniform tie-breaking, then the rest points of the dynamic (1) are the $S(k)$ equilibria of Osborne and Rubinstein (1998) (with $k = \kappa$), and the dynamic itself is the one studied by Sethi (2000). In extensive form games like Centipede, different strategies often earn the same payoffs, so the choice of tie-breaking rule matters. Because of our convention for numbering strategies in the Centipede game, the min-if-tie rule will be the one that is least conducive to cooperative play.

Choice probabilities under best experienced payoff dynamics depend only on the payoffs strategies earn during testing; they do not require agents to track the choices made by their opponents. This property makes the dynamics appealing as a simple model of play for extensive form games. Typically, a single play of an extensive form game does not reveal the strategy chosen by one's opponent, but only the portion of that strategy required to determine the path of play. Consequently, it is not straightforward

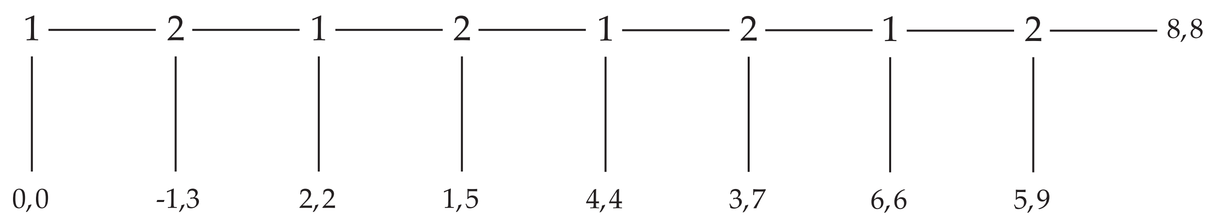


Figure 1: The Centipede game of length $d = 8$.

to specify how agents should use their experience of play to assess opponents' choices of strategies. Because they focus on the performances of own strategies, best experienced payoff dynamics avoid such ambiguities.

2.4 The Centipede game

Centipede (Rosenthal (1981)) 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 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

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

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 (2) 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. When there are multiple trials of each tested strategy ($\kappa = 1$), then cardinal properties of payoffs matter; in this case, Rosenthal's (1981) specification (2) 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 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.

2.5 Best experienced payoff dynamics for the Centipede game

We can now introduce the explicit formula for the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic in the Centipede game.[14]

$$\text{(3a)} \qquad \dot{x}_i = \left(\sum_{k=i}^{s^2} y_k\right)\left(\sum_{m=1}^{i} y_m\right)^{s^1-i} + \sum_{k=2}^{i-1} y_k \left(\sum_{\ell=1}^{k-1} y_\ell\right)^{i-k} \left(\sum_{m=1}^{k} y_m\right)^{s^1-i} - x_i,$$

$$\text{(3b)} \qquad \dot{y}_j = \begin{cases} \left(\displaystyle\sum_{k=2}^{s^1} x_k\right)(x_1 + x_2)^{s^2-1} + (x_1)^{s^2} - y_1 & \text{if } j = 1, \\ \left(\displaystyle\sum_{k=j+1}^{s^1} x_k\right)\left(\displaystyle\sum_{m=1}^{j+1} x_m\right)^{s^2-j} + \displaystyle\sum_{k=2}^{j} x_k \left(\displaystyle\sum_{\ell=1}^{k-1} x_\ell\right)^{j-k+1} \left(\displaystyle\sum_{m=1}^{k} x_m\right)^{s^2-j} - y_j & \text{otherwise.} \end{cases}$$

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 (3a) 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

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

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 (3b).

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

Observation 2.1. *Under the BEP($\tau^{\text{all}}, \kappa, \beta^{\min}$) dynamic, 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 stopped at the initial node. Observation 2.1 shows that this conclusion depends on the assumption that ties are broken uniformly. If instead ties are broken in favor of the lowest-numbered strategy (or, alternatively, an agent's current strategy), then the backward induction state is a rest point. Even so, the analyses to follow will explain why the backward induction state is not a compelling prediction of play even under these tie-breaking rules.

3. Analysis of dynamics with one trial of each strategy

In this section we analyze the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic in Centipede. Since tie-breaking rule $\beta^{\min}$ selects the optimal strategy that stops soonest, it is the tie-breaking rule that is most favorable toward 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 $\operatorname{int}(\Xi)$, 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.[15] 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.

Example 3.1. 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

$$\begin{aligned} \dot{x}_2 &= y_2 - x_2, \\ \dot{y}_2 &= x_2 x_1 - y_2. \end{aligned} \tag{4}$$

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.[16]

Writing $1 - x_2$ for x_1 in (4) and then solving for the zeros, we find that the unique rest point of (4) 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 to come will focus on Centipede games of lengths $d \geq 3$.

3.1 Analytical results

As we know from Observation 2.1, 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

[15] See, e.g., Sandholm (2010b, Appendix 7.B).
[16] Compare the discussion after equation (3) and Example 3.3 below.

point is always repelling.

Proposition 3.2. *In Centipede games of lengths $d \geq 3$, the backward induction state $\xi^\dagger$ is repelling under the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic.*

The proof of Proposition 3.2, which is presented in Appendix A, is based on a somewhat nonstandard linearization argument. While we are directly concerned with the behavior of the BEP dynamics on the state space Ξ, it is useful to view equation (1) 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 (1). Vectors of motion through aff(Ξ) are elements of the tangent space $T\Xi = \{(z^1, z^2) \in \mathbb{R}^{s^1+s^2} \colon \sum_{i \in S^1} z_i^1 = \sum_{j \in S^2} z_j^2 = 0\}$. 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 (1). 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 $\Xi \smallsetminus \{\xi^\dagger\}$, 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 BEP dynamics.

Example 3.3. In a Centipede game of length $d = 4$, writing out display (3) shows that the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic is described by

$$
\begin{aligned}
\dot{x}_1 &= (y_1)^2 - x_1, & \dot{y}_1 &= (x_2 + x_3)(x_1 + x_2)^2 + (x_1)^3 - y_1, \\
\dot{x}_2 &= (y_2 + y_3)(y_1 + y_2) - x_2, & \dot{y}_2 &= x_3 + x_2 x_1 (x_1 + x_2) - y_2, \\
\dot{x}_3 &= y_3 + y_2 y_1 - x_3, & \dot{y}_3 &= x_2 (x_1)^2 + x_3 (x_1 + x_2) - y_3.
\end{aligned} \tag{5}
$$

The linearization of this system at $(x^\dagger, y^\dagger)$ has the positive eigenvalue 1 corresponding to eigenvector $(z^1, z^2) = ((-2, 1, 1), (-2, 1, 1))$ (this is equation (15) with $m \equiv d^1 = 2$ and $n \equiv d^2 = 2$). 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 (3). 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 (5) 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.

Since the backward induction state is unstable, we next try to determine where in the state space the dynamics may converge. As a start, we prove that except at the rest point $\xi^\dagger$, motion from states on the boundary of the state space proceeds immediately into the interior of the state space.

Proposition 3.4. *In Centipede games of all lengths $d \geq 3$, solutions to the BEP$(\tau^{\text{all}}, 1, \beta^{\min})$ dynamic from every initial condition $\xi \in \text{bd}(\Xi) \setminus \{\xi^\dagger\}$ immediately enter $\text{int}(\Xi)$.*

The proof of Proposition 3.4, which is presented in Appendix B, starts with a simple differential inequality (Lemma B.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 (20)). 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.

Together, Propositions 3.2 and 3.4 imply that in Centipede games of length $d \geq 3$, the BEP$(\tau^{\text{all}}, 1, \beta^{\min})$ dynamic has an interior rest point. Our next result places strong bounds on the weights on the most cooperative strategies in any such rest point. For any population state $\xi = (x, y)$ and for $k \in \{0, \ldots, d^1\}$, let $\bar{x}_{[k]} = x_{[0]} + \cdots + x_{[k]}$ be the mass of population 1 agents who stop at their kth-to-last decision node or later. Likewise, for $\ell \in \{0, \ldots, d^2\}$, let $\bar{y}_{[\ell]} = y_{[0]} + \cdots + y_{[\ell]}$.

Proposition 3.5. *Let $\xi = (x, y)$ be a rest point of the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic in Centipede other than the backward induction state $\xi^\dagger$.*

(*i*) *If d is even, so that player 2 moves last, then $\bar{y}_{[2]} > .9657$, $\bar{x}_{[2]} > 1-10^{-4}$, and $\bar{y}_{[3]} > 1-10^{-16}$.*

(*ii*) *If d is odd, so that player 1 moves last, then $\bar{x}_{[2]} > .9657$, $\bar{y}_{[2]} > 1-10^{-4}$, and $\bar{x}_{[3]} > 1-10^{-16}$.*

The proposition implies that the probability that play in a random match leads to the last five terminal nodes, namely $\bar{y}_{[2]}\bar{x}_{[2]}$, is at least .9656, and play is virtually guaranteed to reach the last eight terminal nodes.

Proof. We consider the case in which d is even. We start by introducing three inequalities that a rest point (x, y) must satisfy:

$$1 - \bar{x}_{[k]} \leq (1 - \bar{y}_{[k]})^{k+1}, \tag{6}$$

$$1 - \bar{y}_{[k+1]} \leq (1 - \bar{x}_{[k]})^{k+2}, \tag{7}$$

$$\bar{x}_{[k]} \geq \left(1 - (1 - \bar{y}_{[k+1]})^{k+1}\right)(1 - \bar{y}_{[k+1]}). \tag{8}$$

A necessary condition for a population 1 agent to choose a strategy outside of $\{[k], \ldots, [0]\}$ is that when testing strategies in this set, he is never matched with a population 2 agent playing a strategy in $\{[k], \ldots, [0]\}$. This fact gives us inequality (6). Likewise, for a population 2 agent not to choose a strategy in $\{[k+1], \ldots, [0]\}$, it is necessary that when testing strategies in this set, she is never matched with a population 1 opponent playing a strategy in $\{[k], \ldots, [0]\}$; this gives us inequality (7). Finally, inequality (8) follows from this observation: for a revising population 1 agent to choose a strategy in $\{[k], \ldots, [0]\}$, it is enough that both (i) when playing at least one such strategy, her match opponent chooses a strategy in $\{[k+1], \ldots, [0]\}$ and (ii) when playing strategy $[k+1]$, her opponent does not choose a strategy in $\{[k+1], \ldots, [0]\}$. Inequality (8) can be also be written as

$$1 - \bar{x}_{[k]} \leq (1 - \bar{y}_{[k+1]})^{k+2} + \bar{y}_{[k+1]}. \tag{9}$$

Combining (7) and (9) and rearranging yields the inequality

$$\left((1 - \bar{y}_{[k+1]})^{k+2} + y_{[k+1]}\right)^{k+2} + y_{[k+1]} - 1 \geq 0. \tag{10}$$

When $k = 0$, the left-hand side of (10) is a polynomial whose real roots are 0 and r, with r

close to but greater than .4301, and which is negative on $(0, r)$ and positive on (r, ∞). (This r is the lone real root of the polynomial $z^3 - 2z^2 + 3z - 1$.) Thus any rest point (x, y) with $\bar{y}_{[1]} \neq 0$ satisfies $\bar{y}_{[1]} \geq r > .4301$. By Proposition 3.4, this requirement holds for every rest point besides $\xi^\dagger$.

Now, applying (6) and (7) sequentially yields the inequalities

$$\begin{aligned} \bar{x}_{[1]} &\geq 1 - (1 - y_{[1]})^2 \geq 1 - (1 - r)^2 > .6752, \\ \bar{y}_{[2]} &\geq 1 - (1 - x_{[1]})^3 \geq 1 - (1 - r)^6 > .9657, \\ \bar{x}_{[2]} &\geq 1 - (1 - y_{[2]})^3 \geq 1 - (1 - r)^{18} > 1 - 10^{-4}, \\ \bar{y}_{[3]} &\geq 1 - (1 - x_{[2]})^4 \geq 1 - (1 - r)^{72} > 1 - 10^{-16}. \quad \blacksquare \end{aligned} \tag{11}$$

Part of the intuition behind the proof of Proposition 3.5 is straightforward. Consider an interior rest point $\xi = (x, y)$ of the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic when d is even. Inequality (6) says that if $1 - \bar{y}_{[k]}$ is small (i.e., if few population 2 players stop before $[k]$), then $1 - \bar{x}_{[k]}$ is even smaller (i.e., few population 1 players stop before $[k]$), since some test of a more cooperative strategy will very likely lead to a high payoff. Likewise, by (7), if $1 - \bar{x}_{[k]}$ is small then $1 - \bar{y}_{[k+1]}$ is smaller still. These inequalities imply that any lower bound on $\bar{y}_{[k]}$ will quickly propagate into much stronger lower bounds on $\bar{x}_{[k]}$, $\bar{y}_{[k+1]}$, $\bar{x}_{[k+1]}$, and so on (see (11)). Initiating this chain of reasoning requires a less obvious step: we combine inequality (7) with a weak bound (9) on $1 - \bar{x}_{[k]}$ that depends on $\bar{y}_{[k+1]}$ alone. This combination gives us the initial inequality $\bar{y}_{[1]} \geq .4301$, which in turn leads to the strong bounds stated in the proposition.

3.2 Results based on exact computations

Proposition 3.5 places strong lower bounds on the degree of cooperation arising at any rest point other than the unstable backward induction state. To gain a more precise understanding of the form and stability of such rest points, we turn to exact computations. Because the dynamic (3) is a system of polynomials with rational coefficients, its zeros can in principle be found by computing a Gröbner basis for the system. The Gröbner basis is a new system of equations that has the same zeros as the original system, but can be solved by backward substitution. Once the Gröbner basis has been obtained, polynomial factoring and root finding algorithms can be used to identify its zeros, and hence the zeros of the original system. Applying these techniques, which are described in detail in the Online Appendix, leads to part (i) of the following result.

Proposition 3.6. *In Centipede games of lengths* $3 \leq d \leq 6$,

	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 $\xi^* = \xi^*(d)$ of the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic. p denotes the owner of the penultimate decision node, q the owner of the last decision node.

(i) *The BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic has exactly two rest points: $\xi^\dagger$, and $\xi^* = \xi^*(d) \in \text{int}(\Xi)$.*

(ii) *The rest point ξ^* is asymptotically stable.*

Exact solutions can only be obtained for games of length at most 6 because of the computational demands of computing the Gröbner bases. Two indications of these demands are that when $d = 6$, the leading (univariate) polynomial from the Gröbner basis is of degree 221, and a coefficient of one of the polynomials in the basis has 13,278 digits.

Table 1 reports the approximate values of the interior rest points $\xi^* = \xi^*(d)$, referring to strategies using the last-to-first notation $[k]$ introduced in Section 2.4. 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.

In principle, it is possible to prove the local stability of the rest points $\xi^* = \xi^*(d)$ 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 C.1 in Appendix C 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 C.

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.

$d = 3$	$-1 \pm .3820$	-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($\tau^{\mathrm{all}}, 1, \beta^{\mathrm{min}}$) dynamic.

3.3 Numerical results

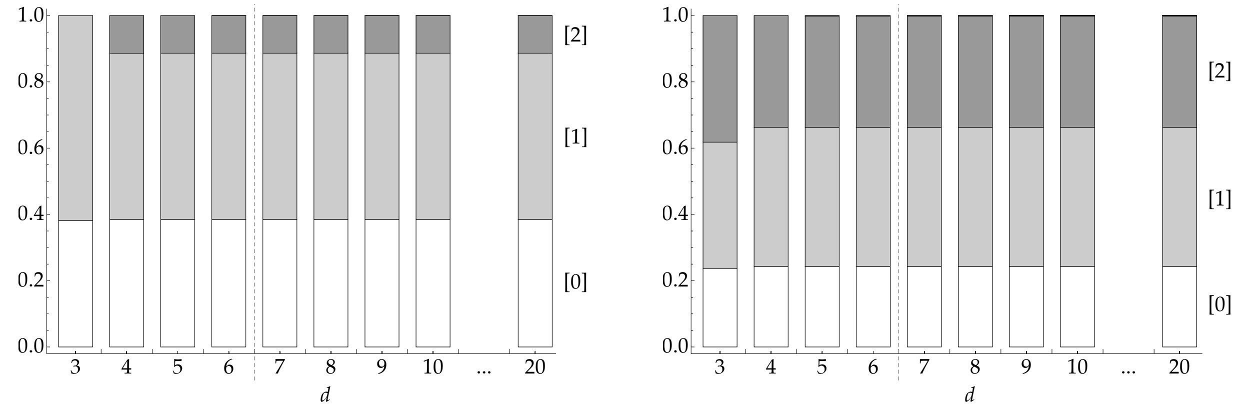
Because exact methods only allow us to determine the rest points of the BEP($\tau^{\mathrm{all}}, 1, \beta^{\mathrm{min}}$) dynamic in Centipede games of lengths $d \leq 6$, we use numerical methods to study games of lengths 7 through 20. We know from Proposition 3.5 that at any rest point besides the backward induction state $\xi^\dagger$, the weight on strategies that stop before either player's third-to-last node is very small. This suggests that the presence of earlier nodes should have little bearing on how the game is played.

Our numerical analysis shows that for game lengths $7 \leq d \leq 20$ there are exactly two rest points, the backward induction state $\xi^\dagger$, and an interior rest point $\xi^* = \xi^*(d)$. 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 IV, as are numerical estimates of the eigenvalues of the derivative matrices $DV(\xi^*)$. 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 a stable interior rest point ξ^* whose form barely varies with the length of the game.

The facts that the vertex $\xi^\dagger$ is repelling, the interior rest point $\xi^* = \xi^*(d)$ 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($\tau^{\mathrm{all}}, 1, \beta^{\mathrm{min}}$) dynamic other than the stationary solution at $\xi^\dagger$.[17] To argue that ξ^* is almost globally stable we introduce the candidate Lyapunov function

[17]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.



(i) penultimate mover (ii) last mover

Figure 2: The stable rest point $\xi^* = \xi^*(d)$ of Centipede under the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic for game lengths $d = 3, \ldots, 10$ and $d = 20$. Stacked bars, from the bottom to the top, represent weights on strategy [0] (continue at all decision nodes), [1] (stop at the last node), [2] (stop at the second-to-last node), etc. The dashed line separates exact ($d \leq 6$) and numerical ($d \geq 7$) results.

$$L(x, y) = \sum_{i=2}^{s_1} (x_i - x_i^*)^2 + \sum_{j=2}^{s_2} (y_j - y_j^*)^2. \tag{12}$$

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 used in Section 3.2 are not suitable for establishing that L is a Lyapunov function. For the Centipede game of length $d = 3$, we are able to verify that L is a Lyapunov function using an algorithm from real algebraic geometry called *cylindrical algebraic decomposition* (Collins (1975)). However, exact implementations of this algorithm fail to terminate in longer games.

We therefore verify numerically that L is a Lyapunov function. 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,[18] and so provides strong numerical evidence that the interior rest point ξ^* is an almost global attractor.

[18] 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 $\frac{1}{8}$.

3.4 Other specifications of the dynamics

To test the robustness of the results above, we repeat the analyses for other specifications of BEP($\tau, 1, \beta$) dynamics. In addition to the test-all rule τ^{all}, we also studied a test-set rule under which the revising agent only considers his current strategy and one other strategy (τ^{two}), as well as a rule under which the revising agent only considers his current strategy and one adjacent strategy (τ^{adj}). The qualitative behavior under these test-set rules is similar to that under τ^{all}. The differences worth mentioning are that stable play is concentrated on a larger number of strategies (e.g., nine strategies in total have mass of at least .01 under BEP($\tau^{\text{two}}, 1, \beta^{\text{min}}$), and that the rate of decay of the weights on strategies that stop earlier is not as severe as under τ^{all}. The intuition here is simple. Under test-all, a revising agent will try out all of his most cooperative strategies, providing many opportunities for some such strategy to perform best; if instead only two strategies are tested at a time, the selective pressure against less cooperative strategies is weaker.

In addition, we also considered alternative tiebreaking rules: *stick/min-if-tie*, which chooses the agent's current strategy if it is optimal, and chooses the lowest-numbered strategy otherwise, and *uniform-if-tie*, which randomizes uniformly among the optimal strategies (as in Osborne and Rubinstein (1998) and Sethi (2000)). As we noted in Section 2.5, uniform tie-breaking implies that the backward induction state $\xi^{\dagger}$ is not a rest point, rendering a stability analysis of this rest point unnecessary. In other respects, alternate choices of tiebreaking rule have little qualititative impact on behavior.

In summary, the results presented in previous sections are highly robust to alternative specifications of the dynamics.

4. 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 κ, focusing on BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics.

4.1 Instability and stability of the backward induction state

Proposition 3.2 shows that the backward induction state $\xi^{\dagger}$ is a repellor under the BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamic with $\kappa = 1$. The following proposition shows that $\xi^{\dagger}$ remains unstable as long as the number of trials κ is less than the length of the game d, and then becomes stable for larger numbers of trials. The statement is complicated slightly by the dependence of the crossover point on whether d is even or odd.

Proposition 4.1. *Under the BEP($\tau^{\text{all}}, \kappa, \beta^{\min}$) dynamic in the Centipede game of length d, the backward induction state $\xi^\dagger$ is unstable if $\kappa \le 2\lfloor \frac{d}{2} \rfloor$ and is asymptotically stable otherwise.*

Like those of the earlier stability analyses, the proof of Proposition 4.1, which is presented in Appendix D, is based on linearization. The key observation is that linearizations of BEP dynamics around pure rest points are driven by match results in which exactly one out of all κs^p match partners plays a strategy different from the equilibrium strategy. We show that if $\kappa \le d - 1$ (for d even), or $\kappa \le d$ (for d odd), there is enough sensitivity to perturbations of the state to ensure the existence of an unstable manifold through $\xi^\dagger$; however, unlike in the $\kappa = 1$ case, $\xi^\dagger$ need not be a repellor. Conversely, if these inequalities are violated, strategy $1 \in S^1$ earns the highest total payoff after any matching with exactly one discrepant opponent. This insensitivity of population 1's behavior to small changes in population 2's behavior ensures local stability.

In order to assess the practical relevance of the stability of the backward induction state for larger numbers of trials, we use numerical analysis to estimate the basin of attraction of $\xi^\dagger$, and to determine the position of the interior saddle point of the dynamics, which lies on the manifold separating the basin of $\xi^\dagger$ from the basin of the main attractor. We focus for tractability on games of length $d = 4$ and numbers of trials $\kappa \le 100$. Details of these analyses are presented in Online Appendices V and VI.

We have two main findings from this numerical analysis. First, the basin of attraction is always minuscule, with volumes always smaller than 0.01% of the total volume of the state space Ξ. Second, $\xi^\dagger$ is almost completely nonrobust to changes in behavior in population 1. Evidence for this lies in the position of the saddle points, which have more than 99.8% of population 1 agents choosing strategy 1, indicating that changes in the behavior of 0.2% of population 1 agents are enough to disrupt the stability of $\xi^\dagger$. Thus the exact stability analysis of the backward induction state for larger numbers of trials is undercut by a thorough numerical analysis of the dynamics in the vicinity of that state.

4.2 Persistence of the stable interior rest point

When agents test their strategies 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 (1995)), 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($\tau^{\text{all}}, \kappa, \beta^{\min}$) dynamic. Figures 3 and 4 present the stable rest points of this dynamic for numbers of trials κ up to 50, which we computed using numerical methods. 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 [0] (always continue) becomes fixed between 7% and 6.5% once $\kappa \geq 15$, even as the fraction of population 1 agents who play strategy [0] 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 random match at this state. By equation (2) (or Figure 1), the expected payoffs to this agent's three strategies are

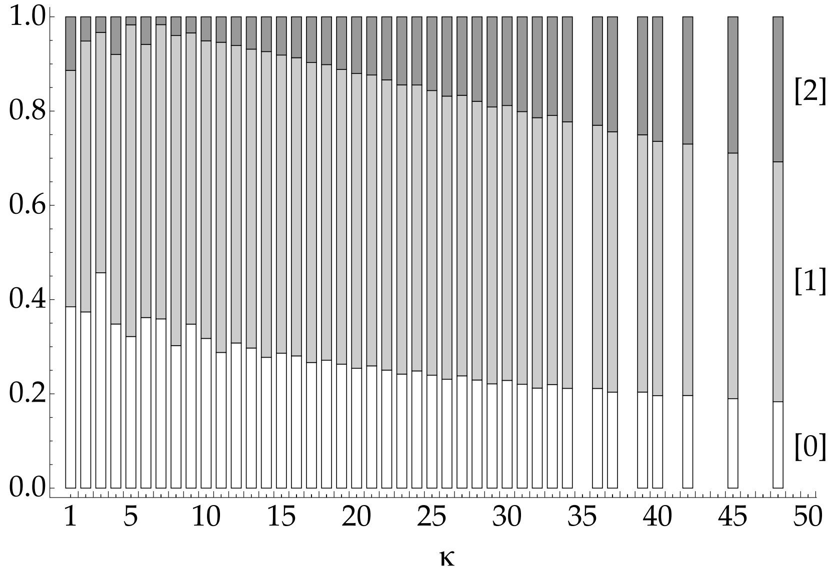
$$\mathbb{E}(\Pi_1) = (0,3,3)\cdot x^* = 2.3580,\;\; \mathbb{E}(\Pi_2) = (0,2,5)\cdot x^* = 2.2086,\;\; \mathbb{E}(\Pi_3) = (0,2,4)\cdot x^* = 1.9964.$$

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

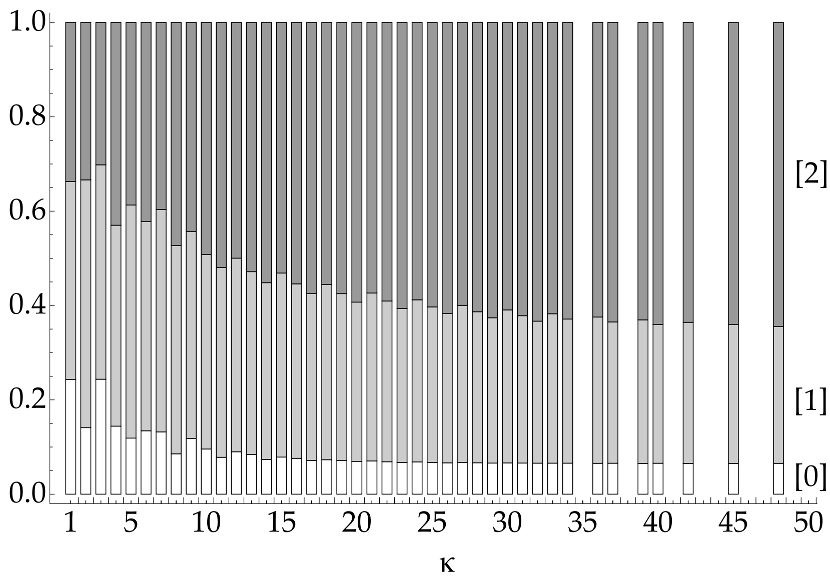
To explain why these weights take the values they do, we also need to know how dispersed the payoffs from testing each strategy are. We thus compute the variances of the single-test payoffs Π_j:

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

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 $\sqrt{(\text{Var}(\Pi_3) + \text{Var}(\Pi_2))/32} \approx .3720$. 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.



(i) population 1



(ii) population 2

Figure 3: The stable interior rest point of the BEP($\tau^{\mathrm{all}}, \kappa, \beta^{\mathrm{min}}$) dynamic in the Centipede game of length $d = 4$, $\kappa = 1, \ldots, 50$. Stacked bars, from the bottom to the top, represent weights on strategies [0], [1] and [2].

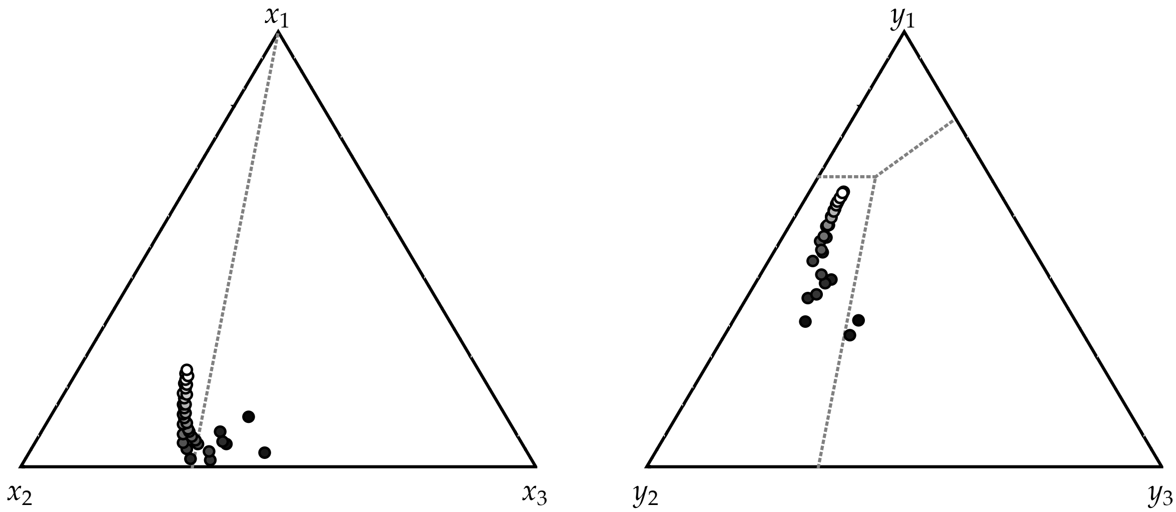


Figure 4: The stable interior rest point in Centipede of length $d = 4$ under BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 1, \ldots, 34$ trials of each tested strategy. Lighter shading corresponds to larger numbers of trials. Dashed lines represent boundaries of best response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

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.

4.3 Convergence to cycles

Figure 3 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 3 and all κ between 51 and 100, the BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamic converges to a periodic orbit. Figure 5 presents the cycles under the BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 50, 100$, 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 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 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.

5. Conclusion

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.

Appendix

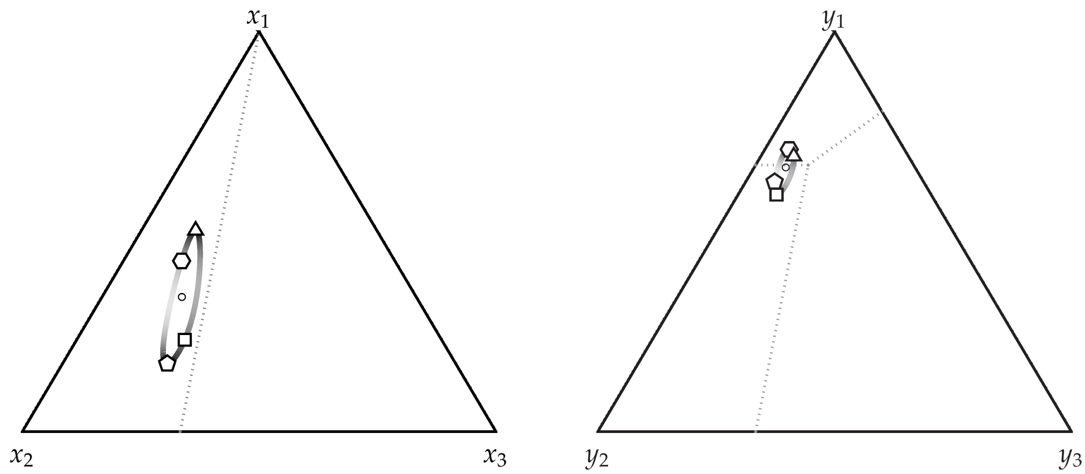
A. Proof of Proposition 3.2

A.1 Generalities

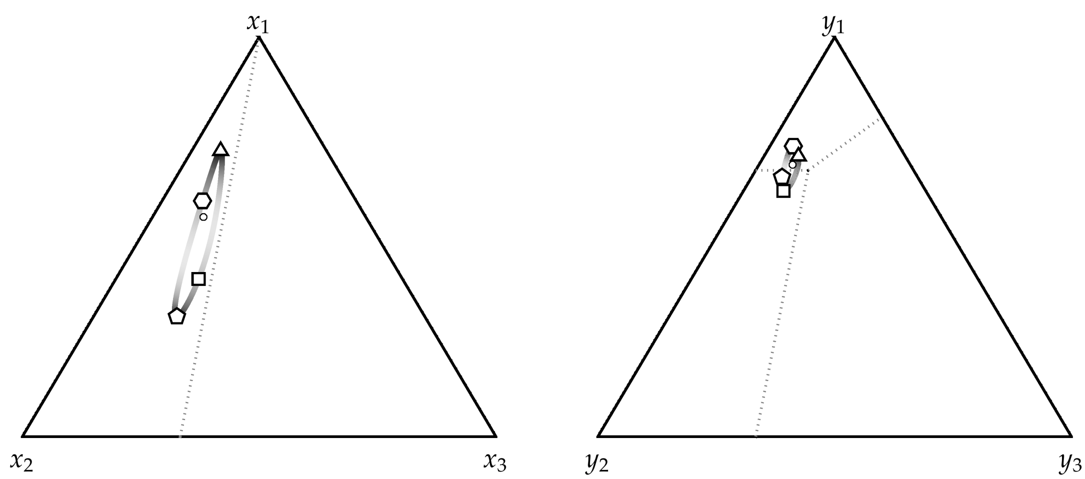
Letting $s = s^1 + s^2$, we denote the tangent space of the state space $\Xi = X \times Y$ by $T\Xi = TX \times TY = \{(z^1, z^2)' \in \mathbb{R}^s \colon \sum_{i \in S^1} z^1_i = 0 \text{ and } \sum_{j \in S^2} z^2_j = 0\}$, and we denote the affine hull of Ξ by $\text{aff}(\Xi) = T\Xi + \xi^\dagger$.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

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

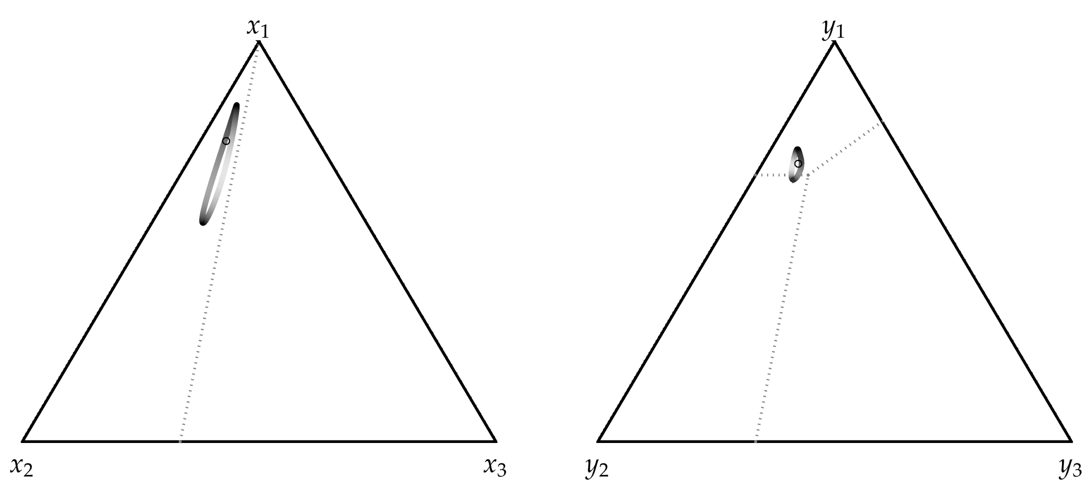
we have $V \colon \text{aff}(\Xi) \to T\Xi$, and so $DV(\xi)z \in T\Xi$ for all $\xi \in \Xi$ and $z \in T\Xi$. We can thus view $DV(\xi)$ as a linear map from $T\Xi$ to itself, and the behavior of the dynamics in the neighborhood of a rest point is determined by the eigenvalues and eigenvectors of this



(i) $\kappa = 50$



(ii) $\kappa = 100$



(iii) $\kappa = 200$

Figure 5: 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.

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.

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

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 Section A.2, we prove that under the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic, 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. 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 *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.

A.2 Computation of eigenvalues and eigenvectors

Starting from formula (3) and using the notations $m \equiv d^1 = s^1 - 1$ and $n \equiv d^2 = s^2 - 1$, it is easy to verify that under the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic,

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

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

$$\text{(13)} \qquad -1, \qquad \left\{\delta^2 - \delta^i \colon i \in \{3, \ldots, m+1\}\right\} \cup \left\{\varepsilon^2 - \varepsilon^j \colon j \in \{3, \ldots, n+1\}\right\};$$

$$\text{(14)} \qquad -1 - \sqrt{mn}, \quad \left\{(\sqrt{mn}, -\sqrt{n/m}, \ldots, -\sqrt{n/m} \mid -n, 1, \ldots, 1)'\right\}; \text{ and}$$

$$\text{(15)} \qquad -1 + \sqrt{mn}, \quad \left\{(-\sqrt{mn}, \sqrt{n/m}, \ldots, \sqrt{n/m} \mid -n, 1, \ldots, 1)'\right\}.$$

The eigenvectors in (13) and (14) span the stable subspace E^s of the linear equation (L). The normal vector to E^s is

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

This vector satisfies

$$\text{(16)} \qquad (z^\perp)'(\delta^i - \delta^1) = \frac{n+1}{\sqrt{mn}} > 0 \text{ for } i \in S^1 \setminus \{1\} \text{ and}$$

$$\text{(17)} \qquad (z^\perp)'(\varepsilon^j - \varepsilon^1) = \frac{n+1}{n} > 0 \text{ for } j \in S^2 \setminus \{1\}.$$

The collection of vectors $\{\delta^i - \delta^1 \colon i \in S^1\} \cup \{\varepsilon^j - \varepsilon^1 \colon j \in S^2\}$ 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 (15), is positive, the arguments from the start of the section allow us to conclude that $\xi^\dagger$ is a repellor.

B. Proof of Proposition 3.4

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 B.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*

$$v(t) \geq \mathrm{e}^{-t}\left(v(0) + \int_0^t \mathrm{e}^s\, a(s)\, \mathrm{d}s\right) \text{ for all } t \in [0,T]. \tag{18}$$

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 condition $v(0)$. Using the integrating factor e^s to solve this equation yields the right-hand side of (18). ■

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 B.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 (1) for the BEP$(\tau, 1, \beta)$ dynamic. Let $\{\xi(t)\}_{t\geq 0}$ be the solution to the dynamic with initial condition $\xi(0)$. Let $\mathcal{S}_0 = \operatorname{supp}(\xi(0))$ 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 (1), 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 (1) is smooth, there is a time $T > 0$ such that $\xi(t) \in O$ for all $t \in [0,T]$. Thus applying Lemma B.1 shows that

$$\xi_i(t) \geq R(1 - \mathrm{e}^{-t}) \text{ for all } t \in [0,T] \text{ and } i \in \mathcal{S}_1. \tag{19}$$

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 B.1 imply that

$$\xi_j(t) \geq c\, Q^a\, \mathrm{e}^{-t} \int_0^t \mathrm{e}^s (R(1 - \mathrm{e}^{-s}))^b\, \mathrm{d}s \text{ for all } t \in [0,T]. \tag{20}$$

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($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic from states $\xi(0) \neq \xi^\dagger$ immediately enter int(Ξ) by showing that the strategies in $\mathcal{S} \setminus \mathcal{S}_0$ can be considered in a sequence that satisfies the property just stated.

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 $1^{[1]} \prec 1^{[2]} \prec 2^{[1]} \prec 2^{[2]} \prec 3^{[1]} \prec \dots$, which arranges the strategies according to how early they stop play in Centipede.

Proof of Proposition 3.4. Fix an initial condition $\xi(0) \neq \xi^\dagger$. We can sequentially add all strategies in $\mathcal{S} \setminus \mathcal{S}_0$ in accordance with the property above as follows:

(I) First, we add the strategies $\{i \in \mathcal{S} \setminus \mathcal{S}_0 \colon i \prec \max \mathcal{S}_0\}$ 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 $\mathcal{S}_{\text{I}}$ denote the set of strategies added during this stage. The assumption that $\xi(0) \neq \xi^\dagger$ implies that $\mathcal{S}_0 \cup \mathcal{S}_{\text{I}}$ contains $1^{[1]}$, $1^{[2]}$, and $2^{[1]}$.

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

C. Proof of Proposition 3.6(ii)

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 a 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 C.1.

As in Appendix A, let $s = s^1 + s^2 = d + 2$, let $\dot{\xi} = V(\xi)$, $V\colon \text{aff}(\Xi) \to T\Xi$ denote an instance of the BEP dynamics, 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

$$(21) \qquad \Delta = \max_{i\in S}\max_{k\in S}\sum_{j\in S}\frac{\partial^2\mathcal{V}_i}{\partial\xi_j\partial\xi_k}(1,\ldots,1\,|\,1,\ldots,1).$$

Proposition C.1. *Suppose that $D\mathcal{V}(\xi)$ is (complex) diagonalizable with $D\mathcal{V}(\xi) = Q\,\mathrm{diag}(\lambda)Q^{-1}$, 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*

$$(22) \qquad \left|\lambda^* - \lambda_i\right| < \frac{2\Delta}{s^{s/2-1}}\,\frac{\mathrm{tr}(Q^*Q)^{s/2}}{|\det(Q)|}\sum_{k\in S}|\xi_k - \xi_k^*|.$$

The eigenvalue perturbation theorem (24) 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

$$(23) \qquad |||M|||_\infty = \max_{1\le i\le s}\sum_{j=1}^{s}|M_{ij}|.$$

denote the maximum row sum norm of M. Let $\kappa_\infty(Q) = |||Q|||_\infty\,|||Q^{-1}|||_\infty$ be the condition number of Q with respect to norm (23). 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:

$$(24) \qquad \left|\lambda^* - \lambda_i\right| \le \kappa_\infty(Q)\,|||D\mathcal{V}(\xi) - D\mathcal{V}(\xi^*)|||_\infty.$$

To bound $\kappa_\infty(Q)$, let $|||M|||_2$ denote the spectral norm of M (i.e., the largest singular value of M), and let $\kappa_2(Q) = |||Q|||_2\,|||Q^{-1}|||_2$ 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

$$(25) \qquad \kappa_\infty(Q) \le s\kappa_2(Q).$$

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

$$\kappa_2(Q) < \frac{2}{|\det(Q)|}\left(\frac{\mathrm{tr}(Q^*Q)}{s}\right)^{s/2}. \tag{26}$$

To bound the final expression in (24), note that by construction, each component of 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

$$\max_{\xi\in\Xi}\left|\frac{\partial^2 \mathcal{V}_i}{\partial\xi_j\partial\xi_k}(\xi)\right| \le \frac{\partial^2 \mathcal{V}_i}{\partial\xi_j\partial\xi_k}(1,\ldots,1\,|\,1,\ldots,1). \tag{27}$$

Thus the fundamental theorem of calculus, (27), and (21) imply that

$$\begin{aligned}|||D\mathcal{V}(\xi) - D\mathcal{V}(\xi^*)|||_\infty &\le \max_{i\in S}\sum_{j\in S}\sum_{k\in S}\frac{\partial^2 \mathcal{V}_i}{\partial\xi_j\partial\xi_k}(1,\ldots,1\,|\,1,\ldots,1)\times|\xi_k - \xi_k^*| \\ &\le \Delta\sum_{k\in S}|\xi_k - \xi_k^*|. \end{aligned} \tag{28}$$

Combining inequalities (24), (25), (26), and (28) yields inequality (22). ■

When applying Proposition C.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 (26) equal 1) leads to the lowest bounds. We apply this normalization in the final step of our analysis.

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 C.1. The eigenvalues of $DV(\xi)$ all have negative real part so long as ξ is reasonably close to ξ^*.

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.

D. Proof of Proposition 4.1

Letting $K = \{1, \ldots, \kappa\}$, we can write the population 1 equations of the BEP($\tau^{\text{all}}, \kappa, \beta^{\min}$) dynamic as

$$\text{(29)} \qquad \dot{x}_i = \sum_{r\colon S^1 \times K \to S^2} \left(\prod_{\ell \in S^1,\, \lambda \in K} y_{r_{\ell\lambda}} \right) \mathbf{1}\left[i = \min\left(\operatorname*{argmax}_{k \in S^1} \pi_k^1(r) \right) \right] - x_i, \quad \text{where } \pi_k^1(r) = \sum_{m=1}^{\kappa} A_{k r_{km}}.$$

The result function $(\ell, \lambda) \mapsto r_{\ell\lambda}$ specifies the strategy in S^2 played by an agent's match partner during the λth test of strategy ℓ for all $\ell \in S^1$ and $\lambda \in K$. The second piece of (29) specifies the probability of a given result, and the third piece indicates whether strategy i is the minimal optimal strategy for this result.

If there are two or more occurrences of strategies from S^2 other than 1, then all partial derivatives of the product in (29) equal 0. Thus for the purposes of computing the Jacobian, we need only consider results in which there are 0 or 1 match partners playing strategies other than strategy $1 \in S^2$. These results comprise the following possibilities:

(i) If all match partners play strategy $1 \in S^2$, then strategy $1 \in S^1$ earns total payoff $\kappa \cdot 0$ and all other strategies earn total payoff $\kappa \cdot (-1)$, so strategy 1 has the best experienced payoff.

(ii) If the lone match against another strategy $j \in S^2 \setminus \{1\}$ occurs when the revising agent plays strategy $1 \in S^1$, then total payoffs are as above, and strategy 1 has the best experienced payoff.

(iii) If the lone match against another strategy occurs when the revising agent plays strategy $i \in S^1 \setminus \{1\}$, and if this match occurs against an opponent playing strategy $j \in S^2 \setminus \{1\}$, then (using the payoffs A_{ij} defined in (2)) strategy 1 is the minimal strategy earning the best experienced payoff if

$$\kappa \cdot 0 \geq (\kappa - 1) \cdot (-1) + \begin{cases} 2i - 2 & \text{if } i \leq j, \\ 2j - 3 & \text{if } i > j; \end{cases}$$

otherwise, strategy i uniquely obtains the best experienced payoff.

Accounting for all of these possibilities, including the fact that the matches in cases (ii) and (iii) can occur during any of the κ tests of the strategy in question, we have

$$\begin{aligned}
\dot{x}_1 &= (y_1)^{\kappa s^1} + \kappa (y_1)^{\kappa s^1 - 1}\left(\sum_{j=2}^{s^2} y_j + \sum_{i=2}^{s^1}\left(\sum_{j=2}^{i-1} y_j \mathbf{1}_{2j-2\le\kappa} + \sum_{j=i}^{s^2} y_j \mathbf{1}_{2i-1\le\kappa}\right)\right) - x_1 + O((y_{-1})^2),\\
\dot{x}_i &= \kappa (y_1)^{\kappa s^1 - 1}\left(\sum_{j=2}^{i-1} y_j \mathbf{1}_{2j-3\ge\kappa} + \sum_{j=i}^{s^2} y_j \mathbf{1}_{2i-2\ge\kappa}\right) - x_i + O((y_{-1})^2),
\end{aligned} \tag{30a}$$

where $y_{-1} = \sum_{j=2}^{s^2} y_j$ and $i \in S^1 \setminus \{1\}$.

Turning to population 2, the test results with 0 or 1 match against opponents playing strategies other than $1 \in S^1$ comprise these possibilities:

(i) If all match partners play strategy $1 \in S^1$, then all strategies earn total payoff $\kappa \cdot 0$, so strategy 1 is the minimal strategy earning the best experienced payoff.

(ii) If the lone match against another strategy occurs when the revising agent plays strategy $j \in S^2$, then strategy j earns a positive total payoff and other strategies earn total payoff 0, so strategy j has the best experienced payoff.

Accounting for both possibilities, we obtain

$$\begin{aligned}
\dot{y}_1 &= (x_1)^{\kappa s^2} + \kappa (x_1)^{\kappa s^2 - 1}\sum_{i=2}^{s^1} x_i - y_1 + O((x_{-1})^2),\\
\dot{y}_j &= \kappa (x_1)^{\kappa s^2 - 1}\sum_{i=2}^{s^1} x_i - y_j + O((x_{-1})^2),
\end{aligned} \tag{30b}$$

where $x_{-1} = \sum_{i=2}^{s^1} x_i$ and $j \in S^1 \setminus \{1\}$.

Taking the derivative of (30) at state $\xi^\dagger$, we obtain the following matrix. (We write this matrix for the case of d even, so that $s^1 = s^2$; roughly speaking, the case of d odd corresponds to removing the final column.)

$$
D\mathcal{V}(\xi^\dagger) = \left(\begin{array}{ccccc|ccccc}
-1 & 0 & \cdots & \cdots & 0 & \kappa s^1 & + & \cdots & \cdots & + \\
0 & -1 & \ddots & \ddots & \vdots & 0 & \kappa\mathbf{1}_{2i-2\geq\kappa} & \cdots & \cdots & \kappa\mathbf{1}_{2i-2\geq\kappa} \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \kappa\mathbf{1}_{2j-3\geq\kappa} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & 0 & -1 & 0 & \kappa\mathbf{1}_{2j-3\geq\kappa} & \cdots & \kappa\mathbf{1}_{2j-3\geq\kappa} & \kappa\mathbf{1}_{2i-2\geq\kappa} \\
\hline
\kappa s^2 & \kappa & \cdots & \cdots & \kappa & -1 & 0 & \cdots & \cdots & 0 \\
0 & \kappa & \cdots & \cdots & \kappa & 0 & -1 & \ddots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\
0 & \kappa & \cdots & \cdots & \kappa & 0 & \cdots & 0 & 0 & -1
\end{array}\right). \tag{31}
$$

Each + above represents the number that makes the column sum in the block equal κs^1.

If all indicator functions in the upper-right block of $D\mathcal{V}(\xi^\dagger)$ equal 0, then each + in (31) equals κs^1, implying that this block is the zero operator on TY. In this case (31) acts as a block triangular matrix on $T\Xi$, and so its lone eigenvalue with respect to $T\Xi$ is -1, implying that $\xi^\dagger$ is stable.

To check that the indicators are all 0 when d is even, it is enough to consider the indicator for entry $j = i = s^2 \equiv \frac{1}{2}d + 1$, which is 0 if and only if $\kappa \geq d + 1$. When d is odd, it is enough to check the indicator for entry $j = i = s^1 - 1 \equiv \frac{1}{2}(d + 1)$, which is 0 if and only if $\kappa \geq d$. We conclude that $\xi^\dagger$ is asymptotically stable in these cases.

To show that $\xi^\dagger$ is unstable in the remaining cases (when $\kappa \geq 2$ and $d \geq 3$), write

$$
\chi^\kappa_{ij} = \begin{cases} \mathbf{1}_{2i-2\geq\kappa} & \text{if } i \leq j, \\ \mathbf{1}_{2j-3\geq\kappa}, & \text{if } i > j, \end{cases} \qquad \chi^\kappa_{i\Sigma} = \sum_{j=2}^{s^2} \chi^\kappa_{ij}, \quad \text{and } \chi^\kappa_{\Sigma\Sigma} = \sum_{i=2}^{s^1}\sum_{j=2}^{s^2} \chi^\kappa_{ij}
$$

for $i, j \geq 2$. A straightforward calculation shows that $\lambda = \kappa\sqrt{\chi^\kappa_{\Sigma\Sigma}} - 1$ is an eigenvalue of $D\mathcal{V}(\xi^\dagger)$ corresponding to eigenvector

$$
z = (-\chi^\kappa_{\Sigma\Sigma}, \chi^\kappa_{2\Sigma}, \ldots, \chi^\kappa_{s^1\Sigma} \mid -(s^2 - 1)\sqrt{\chi^\kappa_{\Sigma\Sigma}}, \sqrt{\chi^\kappa_{\Sigma\Sigma}}, \ldots, \sqrt{\chi^\kappa_{\Sigma\Sigma}}).
$$

Since $\kappa \geq 2$, λ is positive whenever at least one of the indicators in $D\mathcal{V}(\xi^\dagger)$ equals 1. Combining this with the previous argument, we conclude that $\xi^\dagger$ is unstable whenever it is not asymptotically stable.

References

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

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

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.

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

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.

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.

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.

Friedman, D. and Oprea, R. (2012). A continuous dilemma. *American Economic Review*, 102:337–363.

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

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

Izquierdo, L. R., Izquierdo, S. S., and Sandholm, W. H. (2019). An introduction to ABED: Agent-based simulation of evolutionary game dynamics. Unpublished manuscript, Universidad de Burgos, Universidad de Valladolid, and University of Wisconsin.

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

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

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. (2019). Efficiency and stability of sampling equilibrium in public good games. *Journal of Public Economic Theory*, forthcoming.

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

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.

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

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

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. (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.

Sandholm, W. H., Izquierdo, S. S., and Izquierdo, L. R. (2019). Stability for best experienced payoff dynamics. Unpublished manuscript, University of Wisconsin, Universidad de Valladolid, and Universidad de Burgos.

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

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