Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xxxxxxx

pp. **X–XX**

"TEST TWO, CHOOSE THE BETTER" LEADS TO HIGH COOPERATION IN THE CENTIPEDE GAME

Segismundo S. Izquierdo*

BioEcoUva and Department of Industrial Organization
Universidad de Valladolid, Paseo del Cauce 59
47011 Valladolid, Spain

Luis R. Izquierdo

Department of Management Engineering
Universidad de Burgos, Avda. Cantabria s/n
09006, Burgos, Spain

(Communicated by the associate editor name)

Abstract. Explaining cooperative experimental evidence in the Centipede game constitutes a challenge for rational game theory. Traditional analyses of Centipede based on backward induction predict uncooperative behavior. Furthermore, analyses based on learning or adaptation under the assumption that those strategies that are more successful in a population tend to spread at a higher rate usually make the same prediction. In this paper we consider an adaptation model in which agents in a finite population do adopt those strategies that turn out to be most successful, according to their own experience. However, this behavior leads to an equilibrium with high levels of cooperation and whose qualitative features are consistent with experimental evidence.

1. Introduction.

The Centipede game [30] is a paradigm for modeling sequential interactions in which the temptation to secure short-term benefits can hinder the realization of much larger long-term gains. In the Centipede game, two players sequentially choose whether to stop or continue their interaction. Choosing to continue the interaction yields an immediate cost to the chooser, but a greater benefit to his opponent. Thus each decision to continue increases the joint total payoffs the players obtain. Play proceeds until one player decides to stop the interaction, or until a choice to continue is made in the final period. This final period is prespecified as part of the definition of the game. A Centipede game with four decision nodes is presented in fig. 1.

Centipede games can be understood as a stylized model of sequential contributions to the social good for settings with a commonly known terminal date. Applications can be found in a variety of domains of human decision making including sequential disarmament by two countries, sequential effort choice by two contributors to a project, and negotiations with sequential concessions by politicians nearing the

2020 *Mathematics Subject Classification.* Primary: 9110, 91A22; Secondary: 60J10.

Key words and phrases. evolutionary game dynamics, Centipede game, backwards induction, cooperation, simulation, best experienced payoff dynamics, finite population.

Financial support from grants PRX18-00182 and PRX19/00113 by the Spanish Ministry of Science, Innovation and Universities and the Fulbright Program, from MINECO/AEI/FEDER, UE (project ECO2017-83147-C2-2-P) and from "Junta de Castilla y León - Consejería de Educación" through BDNS 425389 is gratefully acknowledged.

* Corresponding author: Segismundo S. Izquierdo.

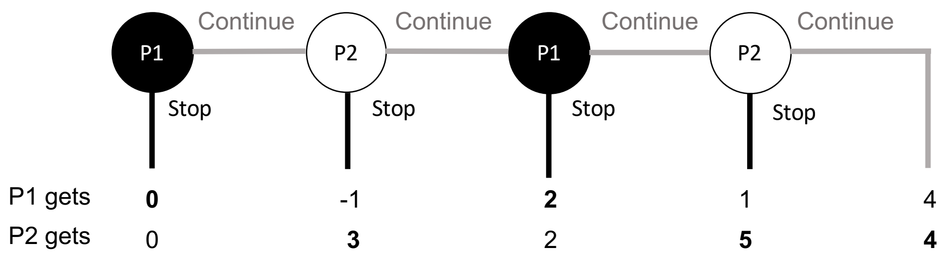


FIGURE 1. A centipede game with four decision nodes, each labeled with the deciding player. Payoffs for player 1 (P1) appear above those for player 2 (P2).

end of their terms. Examples can even be found in biology, among hermaphroditic sea bass, which take turns in laying small batches of eggs and fertilizing their mate's eggs [2].

Centipede games exhibit a tension between social optimality and individual incentives. Each additional round of play increases the players' joint total payoffs, but in the final period of the game choosing to continue can only hurt the player who chooses. In earlier periods, a player benefits from continuing if and only if he expects his opponent to continue in the next period. The logic of backward induction—formalized in the notion of subgame perfect equilibrium [35, 36]—predicts that the first player will stop the game at the very first decision node. The reason is the following: the player who chooses in the final node will surely stop (because continuing will only hurt him). Knowing this, the owner of the previous decision node will also stop, since he anticipates that the other player will stop the game in the following period. This backward induction logic unravels all the way back to the initial node. Thus backward induction predicts completely uncooperative behavior, with players obtaining the lowest possible joint total payoff.

The prediction that players stop immediately can be criticized from multiple points of view. Experimental evidence in Centipede shows that cooperative behavior, i.e., choosing to continue, often persists, with very few matches ending at the initial decision node, most reaching the last decision nodes, and a non-negligible fraction proceeding to the very last node [19, 21]. Experience generally, but not always, tends to make players stop at earlier stages [23, 7, 28, 9].

The backward induction prediction can also be criticized on theoretical grounds. Backward induction is founded on the assumption that each player always expects his opponent to behave rationally in the future, regardless of how he behaved in the past, and this assumption can be criticized not only from a descriptive point of view, but also from a normative perspective [3, 26, 40, 1, 29, 11, 24]. A decision not to stop at the initial node in Centipede could be taken as a signal of the intention not to stop at future nodes. Furthermore, choosing to stop at any decision node ensures that a player gets the second-worst payoff of those he can still obtain; the strong assumptions underlying backward induction are thus particularly pessimistic.

Traditional approaches to explain cooperative play in Centipede have relied on introducing the possibility that one's opponent may be altruistic [18, 19]. This complicates the model considerably, as traditional solution concepts require even the nature of this uncertainty about opponents' preferences to be common knowledge among the players. Most models based on population dynamics, including the replicator dynamics and the best response dynamics, also lead to the stop-immediately prediction in Centipede [8, 42]. However, adding noise to agents' decisions has been shown to lead to cyclical behavior under the replicator dynamic [27], and provides a possible explanation for cooperative behavior in noisy best response models [20].

Like traditional models in game theory, most population dynamic models assume that players have access to precise information about opponents' behavior—here, information about the population shares of each strategy, or about the expected payoffs that each strategy currently earns over all possible random matches. Such assumptions demand a lot from the players, particularly in games like Centipede, where a play of the game need not reveal the opponent's intended strategy (this is always the case if a play of the game does not reach the decision node at which the opponent intended to stop).

In this paper we take a different approach to defining population dynamics for games. Instead of assuming that players have precise information about aggregate behavior, we assume instead that they base their decisions entirely on their experiences playing the game, choosing the strategy that performed best during the most recent test of alternatives. Here we study the simplest form of this best-experienced-payoff process [33]: In each period, some agents are selected at random and given the opportunity to switch strategies. Upon such opportunities, each revising agent randomly selects a single alternate strategy. He tests his current strategy and the alternate strategy κ times each by playing them against randomly matched opponents. He then switches to the alternate strategy if the total payoff of the alternate strategy in the test is higher than the total payoff of his original strategy.

Using simulations and a deterministic mean dynamic approximation, we show in this paper that when individual agents follow this "test two, choose the better" rule in the Centipede, the distribution of choices in the population becomes concentrated on the last few nodes of the game. This prediction differs both from the traditional one based on subgame perfection, and also from those in the canonical framework of evolutionary game dynamics, but largely accords with experimental evidence and with the widespread intuitive impression that cooperative play in Centipede can be sustained.

Among the closest antecedents in the literature, Osborne and Rubinstein [22] define a static equilibrium concept for "procedurally rational players" who choose optimally given the information they possess. Specifically, they consider equilibrium behavior among players who include all of their strategies as candidates, test each κ times, and then choose the one whose total payoff is highest, breaking ties randomly. They show that if such players play the Centipede game, the equilibrium probability that player 1 stops immediately vanishes as the number of decision nodes grows to infinity. Sethi [37] introduces large population deterministic dynamics derived from the decision rule in [22]. Sandholm et al. [33] (see also [34]) prove that when players test all their strategies once before updating their strategy, deterministic dynamics in the Centipede game present an equilibrium in which almost every player continues up to their last three decision nodes.

We deviate from [33] in two main aspects. First, the assumption that a player will test all his available strategies before adopting a new one can be too stringent in some settings, especially if there is a large number of available strategies. A natural variation is to consider that a new strategy can be adopted after testing a limited number of alternative strategies; in the simplest setting, this leads to the *test-two* rule on which we focus. Second, real populations are necessarily finite, and deterministic dynamics results (which assume that the population is infinite) are not guaranteed to be good approximations for realistic population sizes. In this paper, besides considering deterministic dynamics, we focus on relatively small populations (mainly population sizes 10 and 100), so we can assess the practical relevance of the deterministic approximation.[1]

We start our analysis deriving the deterministic approximation (or mean-dynamic equations [32]) for the *test-two* rule with one trial ($\kappa = 1$), and presenting various results that characterize it. The backward-induction state, at which all players use their stop-immediately strategy, is always a rest point of the mean dynamics, but we prove that it is a repellor, i.e. solution trajectories starting close to the backward-induction state move away from it. We also find an interior rest point that we can compute exactly for centipedes of length $d \leq 8$, and through numerical analysis in longer games. At this interior rest point, most of the matches reach one of the last decision nodes. Specifically, for any length of the game, more than 94% of the matches reach one of the last five decision nodes, and if the number of decision nodes is greater than five the strategy distribution over the last nodes is basically the same. Furthermore, a numerical exploration of the mean dynamics suggests that the interior rest point attracts all trajectories except the stationary one at the backward induction state. We also derive the mean-dynamic equations for any number of trials κ and explore their behavior numerically for Centipedes with four nodes. In this case, we find that substantial levels of cooperation persist even for large numbers of trials (e.g. $\kappa = 100$) and the dynamics are cyclical, rather than contractive.

In finite populations, we study a short centipede ($d = 4$), which admits supporting graphical representations, and a longer centipede ($d = 10$) as an illustrative case of the results obtained for $d \geq 8$. Given that for large populations one can expect the deterministic approximation results to become more relevant, we focus here on relatively small groups, considering population sizes N between 10 and 100 agents.

For the one-trial case we find –in accordance with the deterministic approximation– an interior attractor[2] such that most matches end at one of the last nodes of the game. In finite populations where agents test each strategy a larger number of trials, we find –in accordance with the deterministic approximation– cyclical behavior. This cyclical behavior persists even if agents test their strategies against nearly all individuals in the other population. This is striking since when κ is large, the distribution of opponents' choices that revising agents face is similar to the actual current distribution in the population; this suggests that simulations should move towards Nash equilibria (all of which imply no cooperation at all). In fact, when agents test their strategies against all the agents in the opposing population (i.e. $\kappa = N$), a no-cooperation state is quickly reached, since in that case the *test-two* rule is effectively a pairwise version of the best-response protocol [10, 12, 43]. However, it is striking that the small variability introduced by making agents play against the whole population except for just one agent (i.e. $\kappa = N - 1$) can change

[1] Smead [39] also studies finite populations of agents who play the Centipede game, but they evolve according to a frequency-dependent Moran process. He focuses on the asymptotic behavior of the model and finds substantial levels of cooperation for some mutation rates.

[2] We use the term attractor in this context for states around which the process spends long consecutive periods of time.

the dynamics completely, leading to the cooperative cycles predicted by the mean dynamics.

To prove our results about the deterministic approximation, we use techniques from dynamical systems theory and we also employ algorithms from computational algebra and perturbation bounds from linear algebra. We complement this approach with numerical analyses of cases in which exact results cannot be obtained. All the analytical proofs are included in appendix A.1 and all the computational proofs are discussed in appendix A.2. The procedures followed to obtain the numerical results on the deterministic approximations are explained in appendix A.3.

The *Mathematica* notebook used to conduct the computational proofs and to obtain the numerical approximations is freely available at https://github.com/luis-r-izquierdo/bep-centipede. Details about all the functions implemented in the notebook are provided in the notebook itself.[3]

The results about the original dynamics on finite populations have been obtained running agent-based simulations, following the procedure detailed in appendix A.4. The agent-based model, which has been implemented in the open-source platform *NetLogo* [41], is freely available at https://luis-r-izquierdo.github.io/centipede-test-two.[4]

2. **The test-two dynamics.**

2.1. **Definition.** In this section we formally describe the stochastic process under study, henceforth the *test-two dynamics*. The two-player normal form game $G = \{(S^1, S^2), (A, B)\}$ is defined by pairs of strategy sets $S^p = \{1, \dots, s^p\}, p \in \{1, 2\}$ and payoff matrices $A, B \in \mathbb{R}^{s^1 \times s^2}$. A_{ij} and B_{ij} represent the two players' payoffs when strategy profile $(i, j) \in S^1 \times S^2$ is played. Our analysis of Centipede 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. If the number of decision nodes d in Centipede is even, each individual has an associated strategy $i \in \{1, ..., \frac{d}{2} + 1\}$, where strategy $i \leq \frac{d}{2}$ corresponds to "stop at your i-th decision node, and not before", and strategy $i = \frac{d}{2} + 1$ corresponds to "do not stop". The adaptation to an odd number of decision nodes is immediate, with $s^1 = \frac{d+3}{2}$ and $s^2 = s^1 - 1$. 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 $k \in \{0, \dots, s^p\}$, so that $[0]$ denotes continuing at all nodes, and $[k]$ with $k \geq 1$ denotes stopping at player p's kth-to-last node.

The payoff matrices (A, B) of Centipede's reduced normal form can be expressed concisely as

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

We consider two populations of N individuals each who play a Centipede game with d decision nodes. Individuals from one population take the role of player 1, and individuals from the other population take the role of player 2.

At every time period, an individual may revise his strategy with some probability $\gamma \in (0, 1)$; this probability is the same for every individual and independent between

[3] See also [33, section II, supplementary appendix] for an overview of the implemented functions.

[4] Other "best experienced payoff" dynamics –with different ways of choosing candidate strategies and breaking ties– in other games can be simulated with ABED [15], a more general software designed to simulate a wide range of evolutionary dynamics in finite populations.

individuals. Revising individuals choose a (uniformly) random alternate strategy, different from their current one, and play 2κ Centipede games: κ games (trials) using their current strategy and κ trials using the alternate one. Each of those κ trials is played with a newly picked random partner from the other population, without replacement while testing the same strategy. After all the revising players have tested their alternate strategies, they simultaneously decide whether they adopt their alternate strategy: if the total payoff obtained by the alternate strategy is greater than the total payoff obtained by their current strategy, they adopt their alternate strategy. Otherwise, they keep their current one.[5]

2.2. **Asymptotic or ultralong-run behavior.** Defining a state by the number of agents that are choosing each strategy in each population, it is not difficult to see that the *test-two dynamics* are Markov chains. In any game, for the test-two dynamics, states that correspond to pure Nash equilibria are absorbing states of the dynamic, but there could be more. In particular, in the Centipede game, for some parameterizations, there are more absorbing states besides the backwards induction state (which corresponds to the only pure Nash equilibrium).[6] Nonetheless, all the absorbing states share an important feature: every agent in population 1 is choosing the stop-immediately strategy, so at every absorbing state all games end at the first node. This result is gathered in proposition 1 and the proof is presented in appendix A.1.

Proposition 1. *In Centipede games of all lengths ($d \geq 2$), the* test-two dynamics *with any population size N, any number of trials κ, and any probability of revision $\gamma \in (0,1)$, eventually reach an absorbing state where all games end at the first node, regardless of initial conditions.*

Following the terminology of [4, 5], the stop-immediately situation where every game ends at the first period is the unique ultralong-run attractor of the process. Eventually, every realization of the process will end up there. However, for any given finite time window of analysis, this ultralong-run prediction can be, in practical terms, unattainable, since it may take an astronomically long time to reach any of the absorbing states.[7]

The attractive regimes that we focus on in this paper are what [4, 5] call the long-run attractors of the process (in contrast with the ultralong-run attractors). These attractive regimes are states or sets of states where long but finite realizations of the process are expected to spend a large fraction of time. To characterize these regimes, the deterministic approximation that we develop in the next section is particularly useful, especially when populations are large.

[5] The mean dynamics of this stochastic process in 1-population 3-strategy games can be easily analyzed using EvoDyn-3s [14]. This software generates phase portraits of evolutionary dynamics, as well as data for the analysis of their equilibria, using exact arithmetic.

[6] For instance, for $2 < \kappa \leq N$, the state where every agent chooses strategy 1 except for one agent in population 2 who uses strategy 2, i.e. $(N, 0, \dots, 0|N-1, 1, 0, \dots, 0)$, is also absorbing. At that state, a player 1 testing a strategy other than 1 would obtain a payoff of at most $2 + (\kappa - 1)(-1) = 3 - \kappa \leq 0$.

[7] In our computational experiments, reaching a stop-immediately absorbing state from random initial conditions was never observed for long centipedes ($d \geq 10$) with any number of trials $\kappa < N$ (not even for $\kappa = N - 1$), or for short centipedes with low number of trials κ. Only for the short centipede case ($d = 4$) and a number of trials that constitutes a large fraction of the population, it is not unlikely for the system to get to an absorbing stop-immediately state quickly.

3. **Deterministic approximation.**

In this section we present the deterministic approximation or mean-dynamic equations [32] for the *test-two dynamics* and derive several results for this approximation.

3.1. **Notation.**

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 ξ of Ξ as population states.

3.2. **Mean dynamics of the test-two dynamics with one trial ($\kappa = 1$).**

3.2.1. *Equations.*

The mean dynamic of the *test-two dynamics* for Centipede with $\kappa = 1$ can be written as

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

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

In the summations in (1a), the notation $L_-|L_+$ should be read as "if $h < i$ use L_- as the limit; if $h > i$ use L_+ as the limit"; likewise for those in (1b), but with h now compared to j.[8] Each term in the brackets in (1a) represents a comparison between the performances of strategies i and h. Terms that include $x_i + x_h$ represent cases in which the realized payoff to i is larger than that to h, so that it does not matter whether the revising agent is an i player who tests h or vice versa. The terms with x_i alone represent cases of payoff ties, which arise when i and h are both played against opponents choosing the same strategy $k < \min(i, h)$ that stops before either i or h; in this case, the agent will play i only if he was already doing so.

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

The following subsections present various results about the mean dynamic. Some results have been obtained analytically, others using (exact) symbolic computation, and some numerically.

---

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

3.2.2. *Analytical results.* It is easy to check that the backward induction state $\xi^\dagger$ (which corresponds to $x_1 = y_1 = 1$) is a rest point of the mean dynamic (1). The following proposition states that this rest point is always repelling.

Proposition 2. *In Centipede games of all lengths ($d \geq 2$), the backward induction state $\xi^\dagger$ is repelling under dynamic* (1).

Since the backward induction state is unstable, we next try to determine where 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. *In Centipede games of all lengths ($d \geq 2$), solutions to dynamic* (1) *from every initial condition $\xi \in \mathrm{bd}(\Xi) \setminus \{\xi^\dagger\}$ immediately enter $\mathrm{int}(\Xi)$.*

Together, propositions 2 and 3 imply that dynamic (1) has at least one interior rest point for any length $d \geq 2$. To gain a more precise understanding of the form and stability of such rest point(s), we turn to exact computations.

3.2.3. *Results based on exact computations.* Because the dynamic (1) is a system of polynomials with rational coefficients, its zeros can be found –at least in principle– by computing a Gröbner basis for the system (see [33] for details). In our case, it is computationally feasible to calculate all the rest points for $d \leq 8$.[9] Proposition 4 below states that there is just one other rest point besides the backward induction state, and it is asymptotically stable.

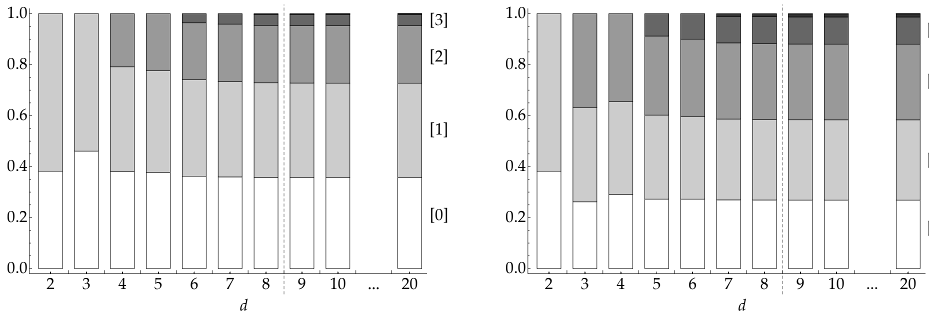
Proposition 4. *In Centipede games of lengths $2 \leq d \leq 8$,*

i Dynamic (1) *has exactly two rest points: $\xi^\dagger$, and $\xi^* = \xi^*(d) \in \mathrm{int}(\Xi)$.*
ii The rest point ξ^* is asymptotically stable.

The unique interior stable rest point is represented in fig. 2 (which also includes values for lengths d greater than 8 that have been calculated numerically). Note that the stable rest point is mostly cooperative, with play always concentrated at the last nodes of the game. Even the weakly dominated strategy for the last movers [0], i.e. the most cooperative strategy which plays "always continue", is played by approximately 25% of the population at the stable rest point, in any Centipede of any length d. The fact that testing of different strategies may occur against different opponents means that relations like dominance, which is intimately connected to backward induction, are not so compelling under the testing procedure considered here.[10]

[9]For lengths $2 \leq d \leq 3$, the system of polynomials can be solved analytically and has solutions that can be expressed in radicals. For lengths $4 \leq d \leq 8$, the solutions are algebraic numbers that cannot be expressed in radicals. For larger values of d, the computational burden of solving large systems of polynomial equations and working with algebraic numbers makes solving the system unfeasible in practical terms. As an example, when $d = 8$, the leading (univariate) polynomial from the Gröbner basis is of degree 128, and a coefficient of one of the polynomials in the basis has 775 digits.

[10]To understand why the weakly dominated strategy is present at the stable rest point, suppose that a population 2 agent tests strategies 2 and 3. It may happen that when she tests strategy 3, the opponent against whom she is matched plays strategy 3 (so the revising player obtains a payoff of 4 when testing strategy 3), and that when she tests strategy 2, the (new) opponent against whom she is matched plays strategy 1 or 2 (so the revising player obtains 0 or 2). Then the revising agent's best experienced payoff comes from her test of the weakly dominated strategy 3.



(i) penultimate mover (ii) last mover

Figure 2. The stable rest point of Centipede under dynamic (1) for game lengths $d = 2, \ldots, 10$ 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 ($d \leq 8$) and numerical ($d \geq 9$) results.

Propositions 2 to 4 together suggest that the unique interior rest point may attract all trajectories except the stationary one at the backward induction state. Using an algorithm from real algebraic geometry called *cylindrical algebraic decomposition* [6], we can prove this statement for Centipedes of lengths $d \leq 3$;[11] moreover, in the following subsection we provide numerical evidence that suggests that this statement is valid for Centipedes of any length.

Proposition 5. *In Centipede games of lengths $2 \leq d \leq 3$, solutions to dynamic (1) from every initial condition $\xi \in \Xi \setminus \{\xi^\dagger\}$ converge to the unique interior state $\xi^*(d)$.*

3.2.4. *Numerical results.* Because exact methods only allow us to determine the rest points of dynamic (1) in Centipede games of lengths $d \leq 8$, we use numerical methods to study games of lengths 8 through 20. Our numerical analysis –detailed in Appendix A.3– suggests that for game lengths $8 \leq d \leq 20$ there are exactly two rest points, the backward induction state $\xi^\dagger$, and an interior rest point $\xi^* = \xi^*(d)$. As fig. 2 illustrates, the form of the interior rest point follows the same pattern for any length $d \geq 7$, with nearly all of the mass placed on each population's four most cooperative strategies, and the weights on these strategies are essentially independent of the length of the game. Thus, for any length of the Centipede, play is always concentrated at the last nodes of the game, with more than 94% of the matches reaching one of the last 5 decision nodes.

In Appendix A.2 we provide precise numerical estimates of the interior rest points $\xi^*(d)$ (table 2) and of the eigenvalues of the Jacobian of the dynamic at $\xi^*(d)$ (table 3). In Appendix A.3 we provide numerical evidence that suggests that proposition 5 (which proves that the interior point is an almost globally asymptotically stable state for lengths $2 \leq d \leq 3$) extends to much longer Centipede games.

To summarize, the numerical results suggest that the conclusions about rest points established exactly for games of lengths $d \leq 8$ continue to hold for longer games: there are always exactly two rest points, the backward induction state $\xi^\dagger$,

[11]Exact implementations of this algorithm fail to terminate in longer games.

and a stable interior rest point ξ^* whose form barely varies with the length of the game. The result that the interior rest point is almost globally asymptotically stable also seems to extend for longer games.

3.3. **Mean Dynamics of the test-two dynamics with several trials.**

3.3.1. *Equations.* The approach followed to derive (1) is based on considering each of the strategies played by an agent's opponent when an agent tests a strategy i in his test set. For large κ we can obtain a formula with far fewer terms than looking at all the possible permutations by instead considering the possible total payoffs that an agent may obtain when testing a strategy, with their associated probability (i.e. working with the *distribution* of opponents' strategies when the agent tests strategy i). Using such formulas is essential for numerical computations when κ is not small.

To do this we introduce a number of definitions. For $p, q \in \{1, 2\}$, $p \neq q$, let

$$\mathbb{Z}_+^{s^q,\kappa} = \left\{ z \in \mathbb{Z}_+^{s^q} : \sum_{j \in S^q} z_j = \kappa \right\}$$

denote the set of possible (unnormalized) empirical distributions of opponents' strategies when a population p agent tests one of his own strategies κ times. When the state of population q is $\xi^q \in \Xi^q$, the probability that empirical distribution z occurs is the multinomial probability

$$M^{p,\kappa}(z, \xi^q) = \binom{\kappa}{z_1 \ \cdots \ z_{s^q}} (\xi_1^q)^{z_1} \ \cdots \ (\xi_{s^q}^q)^{z_{s^q}}.$$

And if a population p agent faces empirical distribution z when testing strategy $i \in S^p$, his total payoff is

$$\pi_i^p(z) = \sum_{j \in S^q} U_{ij}^p \, z_j$$

where U_{ij}^p is the corresponding payoff matrix, i.e., $U_{ij}^1 = A_{ij}$ and $U_{ij}^2 = B_{ji}$.

Therefore, if we let $\Pi_i^{p,\kappa}(\xi^q)$ be a random variable representing the total payoff obtained if strategy $i \in S^p$ is tested κ times when the state of the opposing population is ξ^q, then the distribution of $\Pi_i^{p,\kappa}(\xi^q)$ is

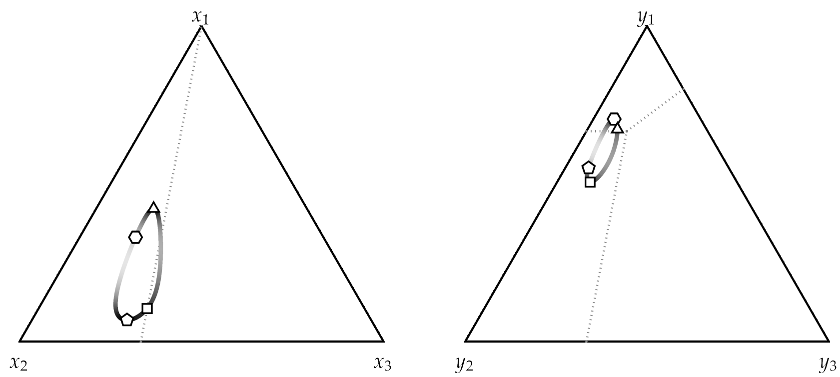
$$\Pr\left(\Pi_i^{p,\kappa}(\xi^q) = w_i^p\right) \ = \sum_{z \in \mathbb{Z}_+^{s^q,\kappa} \, : \, \pi_i^p(z) = w_i^p} M^{p,\kappa}(z, \xi^q).$$

We can then obtain the following equations for the considered test-two dynamic with κ trials:

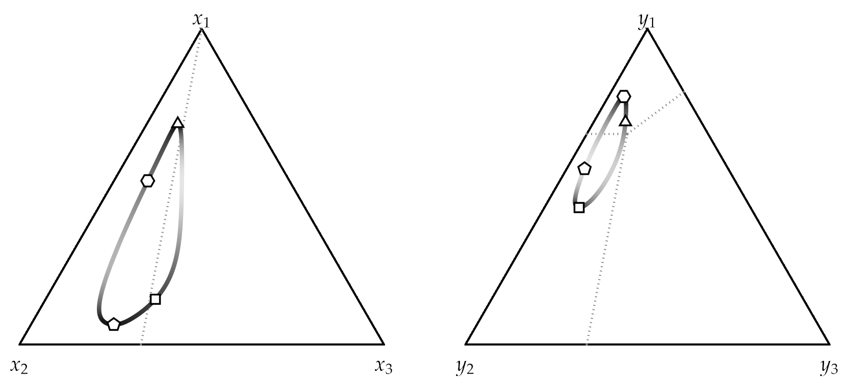
$$\dot{x}_i = \frac{1}{s^1 - 1} \sum_{w_i^1 \in W_i^{1,\kappa}} \Pr\big(\Pi_i^{1,\kappa}(y) = w_i^1\big) \sum_{h=1, h \neq i}^{s^1} \big(x_i \Pr\big(\Pi_h^{1,\kappa}(y) \leq w_i^1\big) + x_h \Pr\big(\Pi_h^{1,\kappa}(y) < w_i^1\big)\big) - x_i \tag{2a}$$

$$\dot{y}_j = \frac{1}{s^2 - 1} \sum_{w_j^2 \in W_j^{2,\kappa}} \Pr\big(\Pi_j^{2,\kappa}(x) = w_j^2\big) \sum_{h=1, h \neq j}^{s^2} \big(y_j \Pr\big(\Pi_h^{2,\kappa}(x) \leq w_j^2\big) + y_h \Pr\big(\Pi_h^{2,\kappa}(x) < w_j^2\big)\big) - y_j \tag{2b}$$

To interpret this formula, note that each term of the form



(I) $\kappa = 50$



(II) $\kappa = 100$

FIGURE 3. Stable cycles in Centipede of length $d = 4$ under dynamics (2) for $\kappa = 50$ and 100. Lighter shading represents faster motion. Shapes synchronize positions along the cycle.

$$x_i \frac{1}{s^1 - 1} \Pr\Big(\Pi_i^{1,\kappa}(y) = w_i^1\Big) \Pr\Big(\Pi_h^{1,\kappa}(y) \leq w_i^1\Big)$$

is the probability that a revising player 1 agent uses strategy i, selects strategy $h \neq i$ to test, obtains payoff w_i^1 when testing strategy i, and obtains some lower or equal payoff when testing strategy h, so the revising player keeps using strategy i.

3.3.2. *Numerical exploration.* Like in the one-trial case, the backward induction state $\xi^\dagger$ is also a rest point of the mean dynamic (2) for any κ. Nonetheless, a numerical exploration of the dynamic reveals that almost all solution trajectories converge to an interior rest point if the number of trials is low, and to a stable cycle if the number of trials is higher. As an example, in a Centipede of length $d = 4$, we can observe that when $\kappa \geq 40$, the interior rest point is not a global attractor anymore but the center of a stable cycle. The amplitude of this stable cycle seems to increase with the number of trials (fig. 3).

To assess the impact of the number of trials κ on the level of cooperation, we can compute the expected duration of play (i.e. the number of the terminal node at which a match ends) in the matches. Note that the duration of play ranges from

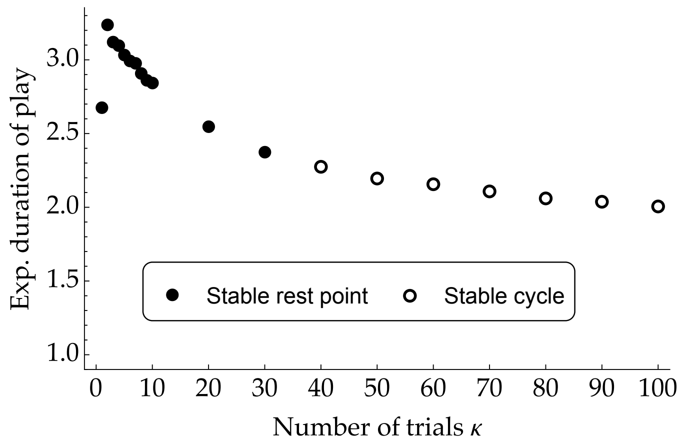


FIGURE 4. Expected duration of play at the stable rest point (for $\kappa \leq 30$) and integrated over the stable cycle (for $\kappa \geq 40$) in Centipede of length $d = 4$, for various numbers of trials κ.

1, if the match stops at the first decision node, to $d + 1$, if the match reaches the last terminal node. Figure 4 shows that going from one trial to two increases the expected duration of play (see appendix A.6 for an explanation of this effect), but from two trials onwards, the expected duration of play decreases as the number of trial increases. This makes intuitive sense –since the test-two rule approaches a pairwise version of best-response as κ increases–,[12] but it is striking to see how slowly the level of cooperation drops as the number of trials increases.

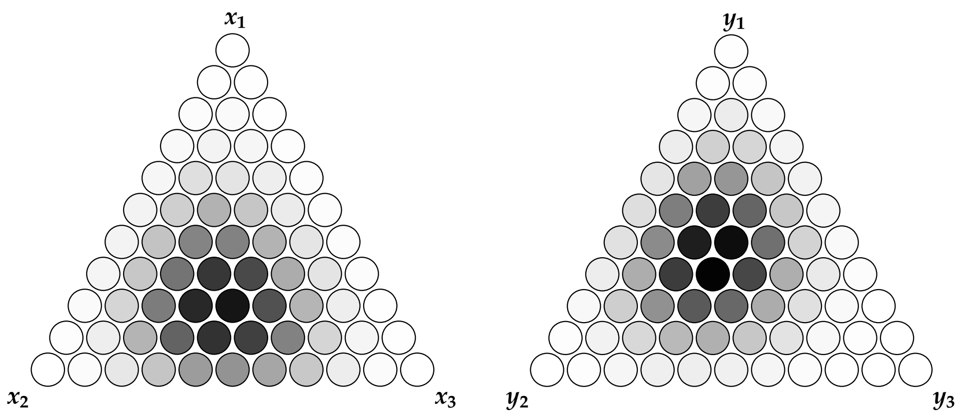
4. **The test-two dynamics on finite populations.** In this section we simulate and analyze the *test-two dynamics* on finite populations. We study the evolution of behavior in a short four-node Centipede and in a 10-node Centipede; the analysis of longer games is essentially the same as the case of length $d = 10$. In each game, we consider various choices of the number of agents in each population N and of the number of trials κ. We always use $\gamma = .1$ as probability of revision.[13]

4.1. **A short Centipede.** We first analyze the evolution of behavior in the four-node Centipede game pictured in fig. 1. In this case, we can represent aggregate behavior in each population by representing the population states x and y as points in the two-dimensional simplex –or, more precisely, as cells in discretized versions of the simplex.

4.1.1. *Four-node Centipede, one trial.* Figure 5 shows a plot describing the empirical distribution over population states in each population when strategies are tested once ($\kappa = 1$), in a population of size $N = 10$, over a long simulation time

[12]This intuition is formalized in [34, Proposition 3.2], where it is proved that any convergent sequence of rest points of the *test-two dynamics* approaches a Nash equilibrium as the number of trials κ increases.

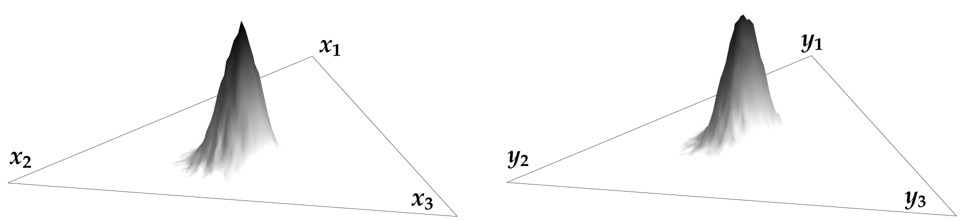
[13]This value was chosen because it allows for a rather quick evolution of the process towards its attractor, and at the same time it shows that the attractor is robust to simultaneous strategy revisions by more than one player in the same time period. (Simultaneous revisions are likely if the population size is large.) Our results present a small sensitivity to reductions or moderate increases in this parameter.



(i) first movers population (ii) second movers population

FIGURE 5. Plot of the empirical distribution on population states, in each population, in a four-node Centipede with $N = 10$ and $\kappa = 1$.

(see details in Appendix A.4). Only a minority of agents in each population stop at their initial node, and the empirical distributions exhibit considerable dispersion, as one would expect from dynamics in small populations. The means of the empirical distributions in the two populations are $\bar{x} = (0.229, 0.407, 0.364)$ and $\bar{y} = (0.361, 0.364, 0.275)$.



(i) first movers population (ii) second movers population

FIGURE 6. Smoothed 3D histogram of the empirical distribution on population states, in each population, for $N = 100$ and $\kappa = 1$.

Figure 6 presents the corresponding results for a population size of $N = 100$. Once again, only a minority of agents stop immediately, with the mean over states now equal to $\bar{x} = (0.210, 0.410, 0.379)$ and $\bar{y} = (0.346, 0.364, 0.290)$, which agrees well with the prediction of the mean dynamic equations (see fig. 2 and table 2). The dispersion of the population about its mean is much smaller, and the distributions over population states appear to be approximately multivariate normal (cf. [31, 15]).

These results and those to follow show that the logic of backward induction is not realized under the *test-two dynamics*. Indeed, the proportion of population 2 players choosing to continue at the last node—a weakly dominated strategy—is not vanishingly small, just like in the mean dynamic.

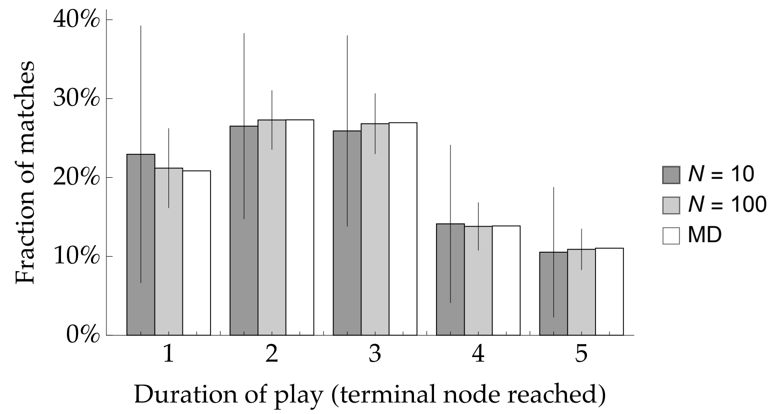


FIGURE 7. Expected fraction of matches that reach each terminal node $i \in \{1, ..., 5\}$ in a Centipede of length $d = 4$ with $\kappa = 1$. For $N = 10$ and $N = 100$, the height of each column corresponds to the average value over the empirical distribution on population states. The vertical lines correspond to the average $\pm$ one standard deviation. MD: mean dynamics.

Using the empirical distribution on population states, we can compute the expected distribution of the duration of play. In fig. 7 we present this information, corresponding to the empirical distributions from figs. 5 and 6. The pattern illustrated in fig. 7 agrees broadly with the qualitative results of experimental studies: most matches continue beyond the first decision node; (conditional) probabilities of stopping are higher at later decision nodes; and iii) there is a significant fraction of matches that get to the last terminal node.

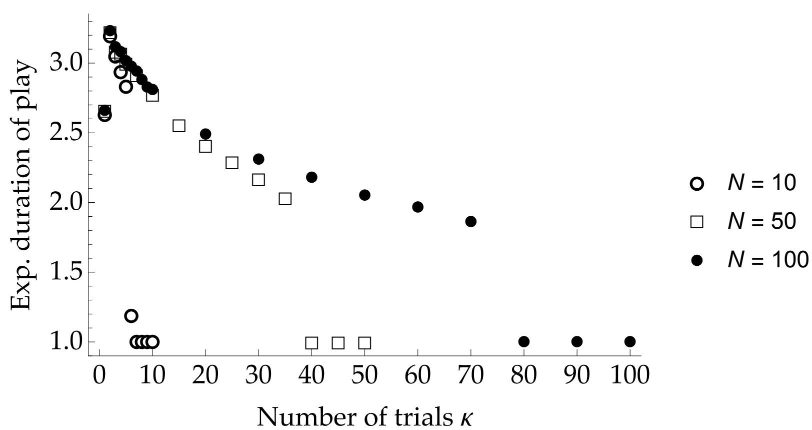


FIGURE 8. Expected duration of play, averaged over the empirical distribution on population states, in Centipede with $d = 4$, for various numbers of trials κ and population sizes N.

4.1.2. *Four-node Centipede, several trials.* To illustrate the effects of increasing the number of trials κ that agents perform, fig. 8 presents simulation data on expected match duration for various choices of κ and of the population size N, where again the expectations are taken over the empirical distributions from simulation runs. Depending on the number of trials, this process can present one or several absorbing stop-immediately states such that all players in the player 1 population stop at their first decision node, and most players in the player 2 population do the same. Figure 8 shows that increasing κ leads to a gradual decline of the (expected) duration of play. An exception occurs when κ increases from 1 to 2, just like we saw in the mean dynamic (see appendix A.6 for a discussion of this effect). As the number of trials κ increases, a value of κ is reached such that getting to an absorbing (stop-immediately) state is not a rare event in the considered time span anymore, and, if the number of trials is increased further, most simulations get to an absorbing state quickly. The expected duration of play in simulations with $N = 100$ that do not reach an absorbing state is well approximated by the mean dynamic (compare figs. 4 and 8).

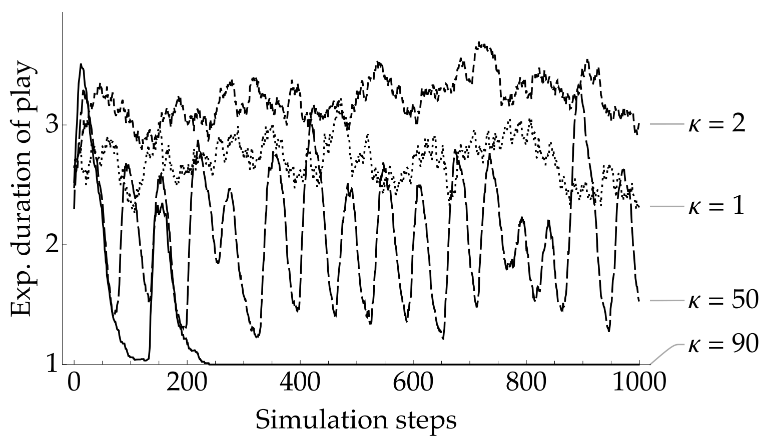


FIGURE 9. Sample paths of the expected duration of play in a 4-node Centipede played in populations of size $N = 100$ for various choices of κ.

Not apparent from fig. 8 is the fact that as the number of trials κ grows, simulation runs begin to exhibit a markedly cyclical behavior, just like we saw in the deterministic approximation. Using a population size of $N = 100$, fig. 9 shows sample paths of the expected duration of play for different number of trials κ during testing. The qualitative behavior of simulations that do not reach the absorbing state is reasonably well approximated by the mean dynamic (2), as can be appreciated comparing figs. 9 and 10. Finally, fig. 11 presents a histogram of the empirical distribution of play when $N = 100$ and $\kappa = 50$; the cyclical nature of the populations' behavior is manifest. As one would expect, the amplitude of the cycles in the finite-population simulation is greater than in the mean dynamic (compare figs. 3(i) and 11).

4.2. **Longer centipedes.** We now study dynamics in Centipede games with $d = 10$ nodes, as a representative case of long games.

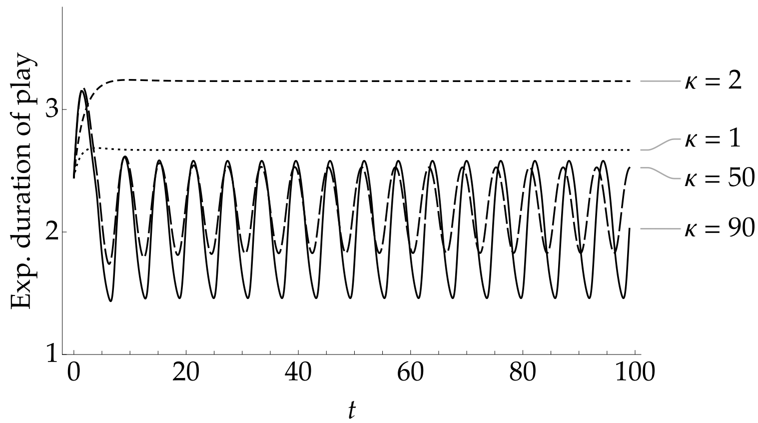
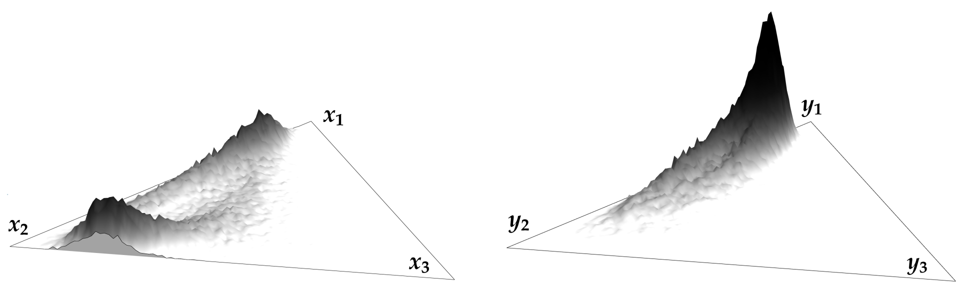


FIGURE 10. Expected duration of play over solution trajectories of the mean dynamic (2) in a 4-node Centipede for various choices of κ.



(I) first movers population (II) second movers population

FIGURE 11. Smoothed 3D histogram of the number of visits to states in each population for a 4-node centipede played in populations of size $N = 100$ with $\kappa = 50$ trials. Cyclical behavior leads the empirical distribution to take a crater-like form.

4.2.1. *Ten-node Centipede, one trial.* For one trial, fig. 12 shows the mean and standard deviation of match durations in Centipede games for different population sizes, again computed from the empirical distribution of states in simulations. In this case, we see that in virtually all matches, the game reaches one of the last 8 terminal nodes. As it happens with the interior rest point of the mean dynamics, it also turns out that this distribution of play remains essentially unchanged for Centipede games with larger numbers of decision nodes. In other words, the distribution of behavior under *test-two dynamics* in every long Centipede game is essentially the same, if we look backwards from the last decision node, with the large majority of games (i.e. more than 90%) reaching one of the last 5 decision nodes.

4.2.2. *Ten-node Centipede, several trials.* Figure 13 shows the expected duration of play, averaged over the empirical distribution on population states, for different

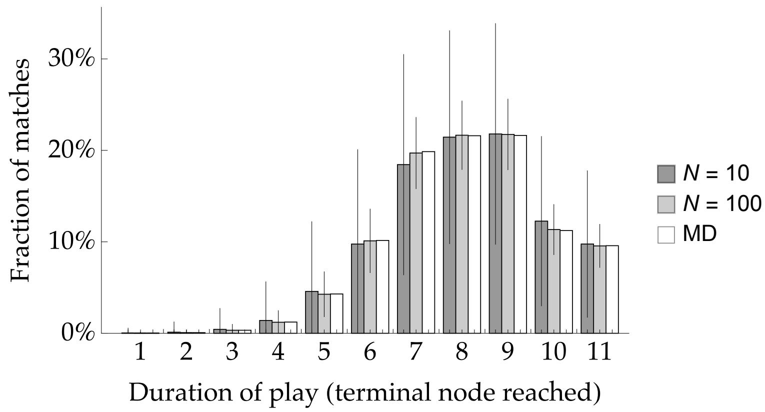


FIGURE 12. Expected fraction of matches that reach each terminal node $i \in \{1, ..., 11\}$. Centipede with 10 decision nodes ($d = 10$) and $\kappa = 1$. For $N = 10$ and $N = 100$, the height of each column corresponds to the average value over the empirical distribution on population states. The vertical lines correspond to the average $\pm$ one standard deviation.

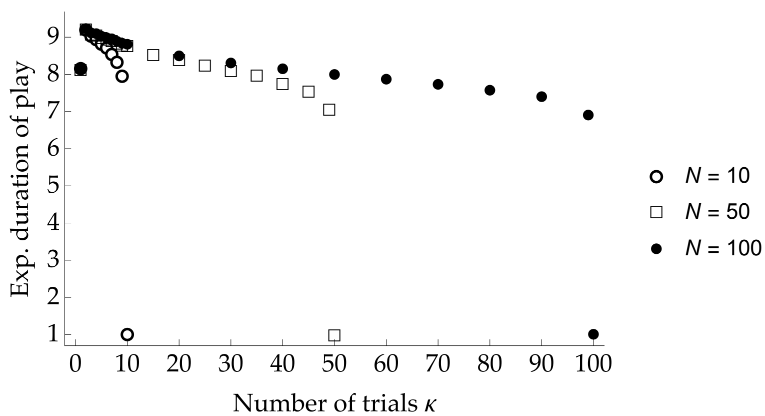


FIGURE 13. Expected duration of play in a Centipede game with 10 decision nodes, averaged over the empirical distribution on population states, for different number of trials κ.

values of κ. After an initial small jump upwards (see appendix A.6 for a discussion of this effect), there is a slow gradual decrease of the duration of play with κ, up until $\kappa = N - 1$, corresponding to the situation in which, while testing each strategy, a player faces all but one member of the opposing population. If κ is increased to N, tests are run against *all* members of the opposing population, so the *test-two dynamics* are a pairwise version of exact best response dynamics [12, 43], and a sharp discontinuity occurs: for this specification, but only for this specification, the logic of backward induction holds force, and cooperation unravels in short order until the backward induction state is reached. It is remarkable that even the small reduction in information arising from facing all but one member of the opposing population leads to a much more cooperative cyclical outcome.

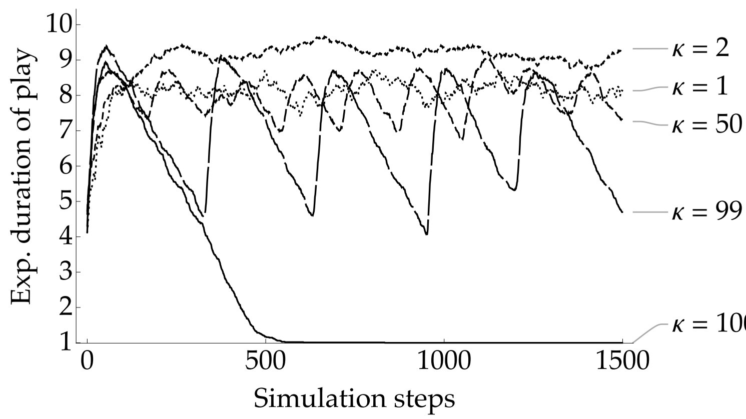


FIGURE 14. Evolution of the expected duration of play in a Centipede game with 10 decision nodes played in populations of size $N = 100$, for different number of trials κ.

Figure 14 shows individual sample paths of the expected duration of play for different numbers of trials κ in a Centipede of length $d = 10$ with $N = 100$. As in the shorter Centipede, we see here that the overall level of cooperation decreases as the number of trials grows, and that cycling through varying levels of cooperation is apparent when the number of trials is large (e.g. $\kappa \geq 40$). This cycling persists until $\kappa = 99$, and then vanishes entirely when $\kappa = 100$ (a parameterization that leads immediately to the backward induction state).

We now discuss the cycling behavior observed for large number of trials, taking as a reference a Centipede with 10 decision nodes, 100 players in each population, and $\kappa = 99$ (see fig. 14). Recall that in 10-node Centipedes, for $i \in \{1, ..., 5\}$ strategy i means "stop at your i-th decision node, and not before", and strategy 6 corresponds to "always continue". A population state is characterized by the number of players using strategies $(1, ..., 6)$ in each population.

In fig. 14 with $\kappa = 99$ it is apparent that the process cycles between state regions with large and much shorter duration of play. Picking up a state from the low-duration-of-play zone, we observe a population state such that the number of players is (0, 0, 76, 22, 2, 0) in the player 1 population, and (0, 32, 66, 2, 0, 0) in the player 2 population. The observed concentration on two or three strategies in each population (with progressive moves towards the strategies that stop earlier) is a characteristic effect of the best-response-like unraveling process in the Centipede. At the considered state, we can calculate the total payoffs that each revising agent can obtain when testing each possible strategy, as well as the probabilities of obtaining those payoffs, leading to Table 1. To distinguish between populations, we denote strategy i for player 1 as i_x and for player 2 as i_y.

Note that, when those individuals who are the last ones to stop (in the considered state, those playing 5_x) try another strategy that stops at a later stage or does not stop, in this case 6_x, 6_x is at no disadvantage with respect to 5_x, and the test using 6_x will provide a higher payoff than the test using 5_x if the individual that is left

TABLE 1. Possible payoffs obtained by each strategy, with their probabilities, at a state next to the bottom of a cycle

Str	#	**0.32**	**0.66**	**0.02**	Str	#	**0.76**	**0.22**	**0.02**
1_x		0	0	0	1_y		297	297	297
2_x		198	198	198	2_y	32	495	495	495
3_x	76	303	300	300	3_y	66	468	465	465
4_x	22	241	239	236	4_y	2	450	448	445
5_x	2	239	237	235	5_y		448	446	444
6_x		239	237	235	6_y		448	446	444

Str: strategy; #: number of players using the corresponding strategy. For each strategy, the three values on the right of the # column are the three possible total payoffs obtainable at the considered state by the corresponding strategy in 99 trials. Top numbers in bold: probability of obtaining the payoffs on that column, at the considered state.

out when playing 6_x with 99 partners (out of 100) stops at an earlier node than the individual left out when testing 5_x. Specifically, at the considered population state, the probability that a revising player playing 5_x who tests 6_x decides to switch to 6_x is $(0.66 \times 0.32 + 0.02 \times 0.98) \approx 0.23$.

Since the best-response-like unraveling process leads to stopping at earlier nodes, those players who are currently the last ones to stop have more stop-later strategies that can be tested, and those stop-later strategies can provide (with relatively high probability) better results than the current strategies, due to the other population's strategy variability (which will usually include at least two strategies, as players adapt their strategies gradually). Once there is one or a few players in the population who continue until the last stages, it becomes advantageous for any player who meets them to stop at later stages of the game rather than stopping early, so the process quickly moves towards stopping at the last stages, and, once there, the slower-motion best-response-like process begins to unravel again, creating the cycle-like behavior observed before (a more detailed discussion is provided in appendix A.5).

5. Concluding remarks.

The backward-induction solution in the Centipede game has been shown to be very sensitive to the assumptions that players make about the other players in the game. If a player thinks that her partner, with some small probability, may be an "altruist" whose preferences do not correspond to the payoffs of the game, it can be rational to keep cooperation for many periods [19, 18]. Here we provide a different potential explanation for cooperation in the Centipede: we assume that players preferences do correspond to the payoffs of the game and that players follow an adaptive approach to change their behavior (i.e. strategy), according to the results they obtain when testing alternative strategies, choosing the strategy that performs best in the test. This experience-based adaptive mechanism leads to population equilibria that are consistent with the characteristic qualitative features observed in experimental studies. A key factor for this result is the variability on the comparative results that two strategies can obtain when tested, which in turn is due to the fact that each strategy is effectively tested against potentially different partners. As a result, strategies that often perform well in the game, such as "always continue" in the Centipede, can be present in equilibria, even if there is a weakly dominating alternative, or an alternative with better average performance. The equilibrium we study is made up by the strategies that are most often the most successful at that equilibrium (as one could expect, being an evolutionary equilibrium), and it constitutes an interesting example of how adopting better-performing strategies can lead to completely different equilibria when agents base their decisions on –potentially variable– experienced payoffs (even if averaged over many trials), versus when strategies' performance is measured by a precise deterministic payoff.

Acknowledgments. This paper is the result of a wonderful and inspiring collaboration with the late Bill Sandholm. If you find something unclear or mistaken, it is certainly our fault; if you find something beautiful or inspiring in this work, it is certainly thanks to him. We deeply miss his mind-blowing intelligence, his generosity, his playfulness and his scientific rigor. But above all, we feel profound gratefulness. We are extremely grateful for having had the opportunity to learn so much from such a great, exceptional, distinguished and amazing friend.

Appendix.

A.1. **Proofs of analytical results.**

Proof of Proposition 1. The proof consists in showing that a necessary condition for a state to be absorbing is that all agents in population 1 are choosing strategy 1, and that it is possible to go from any non-absorbing state to an absorbing state in a finite number of steps. Note that the Markov chain is aperiodic since $\gamma \in (0, 1)$.

Consider a state $(n_1, \ldots, n_{s^1} | m_1, \ldots, m_{s^2})$, where n_i denotes the number of agents in the first movers population with strategy i and m_j the number of agents in the second-movers population with strategy j. The backwards induction state $(N, 0, \ldots, 0|N, 0, \ldots, 0)$ is absorbing for any d, κ and $\gamma \in (0, 1)$, but there can be other absorbing states nearby (see footnote 6).

Assume that we are at a state where $n_1 \neq N$, i.e. not every first mover stops at their first decision node. The following argument shows that such a state is not absorbing. Let i_{max} be the greatest strategy number used in the first population, i.e. $i_{max} = \max(i; n_i > 0)$, and let j_{max} be the greatest strategy number used in the second population, i.e. $j_{max} = \max(j; m_j > 0)$. Naturally, $i_{max} \geq 2$ and $j_{max} \geq 1$. Consider the following two possibilities:

- $i_{max} > j_{max}$. In this case, there is at least one agent in population 1 who is using strategy $i_{max} > j_{max} \geq 1$ and would never stop the game, because all her opponents in population 2 choose to stop at an earlier node. At this state, there is a positive probability that only this agent revises her strategy, she considers strategy $j_{max} < i_{max}$ as an alternate strategy, and she obtains a greater payoff when testing j_{max} than when testing i_{max}.[14] This means that the considered state is not absorbing, since there is a positive probability of moving to another state where one (and only one) agent in population 1 has switched to a lower strategy. This move will not change the value of j_{max} but will potentially bring i_{max} down to j_{max}.

[14] This is the case because at that state, j_{max} obtains at least the same payoff as i_{max} against any given opponent, and a strictly greater payoff than i_{max} against opponents who are using j_{max}.

- $i_{max} \leq j_{max}$. In this case, there is at least one agent in population 2 who is using strategy $j_{max} \geq i_{max} \geq 2$ and would never stop the game, because all his opponents in population 1 choose to stop at an earlier node. At this state, there is a positive probability that only this agent revises his strategy, he considers strategy $(i_{max} - 1) < j_{max}$ as an alternate strategy, and he obtains a greater payoff when testing $(i_{max} - 1)$ than when testing j_{max}.[15] This means that the considered state is not absorbing, since there is a positive probability of moving to another state where one (and only one) agent in population 2 has switched to a lower strategy. This move will not change the value of i_{max} but will potentially bring j_{max} down to $(i_{max} - 1)$.

By applying the logic above repeatedly, we are able to identify a path that can take the dynamics from any state where $n_1 \neq N$ to a state where $n_1 = N$. Thus, all absorbing states must have $n_1 = N$. The following argument shows that it is always possible to go from any non-absorbing state where $n_1 = N$ to the backwards induction (absorbing) state.

Consider a state where $n_1 = N$ and it is not absorbing. At this state, all strategies for the second movers yield the same payoff (i.e. 0), so no agent in population 2 would change her strategy if given the opportunity. Since the state is not absorbing by assumption, this means that at least one agent in population 1 could change her strategy if given the opportunity. Since all agents in population 1 are facing the same situation, this means that there is a positive probability that all of them change their strategy if given the opportunity (and the probability of switching to any given strategy is the same for all of them).

Let $\alpha > 1$ be the minimum strategy number (greater than 1) that could be selected by revising agents in population 1. There is a positive probability that all agents in population 1 revise their strategy simultaneously and they all adopt strategy α. After this, there is a positive probability that all agents in population 2 revise their strategy simultaneously and adopt the (unique) best response to strategy α, which –for the second movers– is strategy number $(\alpha - 1)$. The key is that now all agents in each population are choosing the same strategy, so there is a positive probability that the usual best-response-like unraveling process develops. To be specific, there is now a positive probability that all agents in population 1 revise their strategy simultaneously and adopt strategy $(\alpha - 1)$, which is the best response to second movers' $(\alpha - 1)$. If α was equal to 2, we would be now at the backward induction (absorbing) state. If not, we can repeat this argument as many times as necessary to get to the backwards induction (absorbing) state.

In summary, we have proved that all absorbing states have $n_1 = N$, and that it is possible to go from any non-absorbing state to an absorbing state in a finite number of steps. Applying standard results in the theory of Markov chains we can conclude that the process will eventually get to one of the absorbing states (where $n_1 = N$) and stay there. Thus, eventually, a state where all games end at the first node will be reached. □

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

[15] This is the case because at that state, $(i_{max} - 1)$ obtains at least the same payoff as j_{max} against any given opponent, and a strictly greater payoff than j_{max} against opponents who are using i_{max}.

Writing our dynamics as

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

we have $V\colon \operatorname{aff}(\Xi) \to T\Xi$, and so $DV(\xi)z \in T\Xi$ for all $\xi \in \Xi$ and $z \in T\Xi$. We can thus view $DV(\xi)$ as a linear map from $T\Xi$ to itself, and the behavior of the dynamics in the neighborhood of a rest point is determined by the eigenvalues and eigenvectors of this linear map. The latter are obtained by computing the eigenvalues and eigenvectors of the product matrix $\mathbf{\Phi} DV(\xi)\mathbf{\Phi}$, where $V\colon \mathbb{R}^s \to \mathbb{R}^s$ 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 $I - \frac{1}{s^1}\mathbf{11'} \in \mathbb{R}^{s^1 \times s^1}$ and $I - \frac{1}{s^2}\mathbf{11'} \in \mathbb{R}^{s^2 \times s^2}$, where $\mathbf{1} = (1, \dots, 1)'$. Since V maps Ξ into $T\Xi$, the projection is only needed when there are eigenspaces of $DV(\xi)$ that intersect both the set $T\Xi$ and its complement.

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

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

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

on $T\Xi$.

Examining display (1), it is easy to verify that, for $d \geq 4$, all states in $\{(x, y) \in \operatorname{aff}(\Xi)\colon x = (1, 0 \dots, 0)', y_1 = 1\}$ are rest points. The existence of these sets of rest points implies that the derivative matrices $DV(\xi^\dagger)$ have eigenvalues equal to zero, so that the standard results from linearization theory cannot be applied. To show state $\xi^\dagger$ is nevertheless repelling, we appeal to results from center manifold theory ([16, 17], [25]) which describe the behavior of nonlinear dynamics near nonhyperbolic rest points. 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 *center subspace* $E^c \subseteq T\Xi$ is the span of the real and imaginary parts of eigenvectors corresponding to eigenvalues with zero real part.

Let $A^{cs} = E^c \oplus E^s + \xi^\dagger$ be the affine space that is parallel to $E^c \oplus E^s$ and that passes through $\xi^\dagger$. Below we show that under the considered test-two dynamic ((1)), the subspace $E^c \oplus E^s$ has dimension $d - 1$, and the affine space A^{cs} is a supporting hyperplane to Ξ at $\xi^\dagger$.

Linearization is much less simple for nonhyperbolic rest points than for hyperbolic ones—see [25]. However, for our purposes, it is enough that there exists a (local) *center-stable manifold* M^{cs} that is tangent to A^{cs}, and is invariant under (D) [17]. This manifold need not be unique; see [17, section 4], for an example. But for any choice of center-stable manifold M^{cs}, there is a neighborhood $O \subset \operatorname{aff}(\Xi)$ of $\xi^\dagger$ satisfying $O \cap \Xi \cap M^{cs} = \{\xi^\dagger\}$ such that solutions to (D) from initial conditions in $(O \cap \Xi) \setminus \{\xi^\dagger\}$ eventually move away from $\xi^\dagger$; see [16, p. 336], or see [25, Theorem 2.12.2], for a closely related and more explicitly presented result. This and the properties from the previous paragraph imply that $\xi^\dagger$ is a repellor of the considered test-two dynamics on Ξ.

Starting from (1), we compute that under the considered test-two dynamic,

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

For $d \geq 2$, the eigenvalues of $DV(\xi^\dagger)$ with respect to $T\Xi$ and the bases for their eigenspaces are:

$$0, \qquad \left\{\xi^2 - \xi^j : j \in \{3, \ldots, s^2\}\right\} \text{ if } d \geq 4; \tag{3}$$

$$-\tfrac{1}{m}, \qquad \left\{\delta^2 - \delta^i : i \in \{3, \ldots, s^1\}\right\} \text{ if } d \geq 3; \tag{4}$$

$$\lambda_- \equiv -\tfrac{1}{2m} - \sqrt{1 + (\tfrac{1}{2m})^2}, \qquad \left\{(-\lambda_-, \tfrac{\lambda_-}{m}, \ldots, \tfrac{\lambda_-}{m} \,\Big|\, -1, \tfrac{1}{n}, \ldots, \tfrac{1}{n})'\right\}; \text{ and} \tag{5}$$

$$\lambda_+ \equiv -\tfrac{1}{2m} + \sqrt{1 + (\tfrac{1}{2m})^2}, \qquad \left\{(-\lambda_+, \tfrac{\lambda_+}{m}, \ldots, \tfrac{\lambda_+}{m} \,\Big|\, -1, \tfrac{1}{n}, \ldots, \tfrac{1}{n})'\right\}. \tag{6}$$

The eigenvectors in (3) span the center subspace E^c of the linear equation $\dot{z} = DV(\xi^\dagger)\, z$, while the eigenvectors in (4) and (5) span the stable subspace E^s. The normal vector $z^\perp$ to the hyperplane $E^c \oplus E^s$ is the orthogonal projection onto $T\Xi$ of the auxiliary vector

$$z^\perp_{aux} = \tfrac{1}{\lambda_-}\delta^1 - \xi^1$$

which satisfies

$$(z^\perp)'(\delta^i - \delta^1) = -\tfrac{1}{\lambda_-} > 0 \;\text{ for } i \in S^1 \setminus \{1\}, \;\text{ and}$$
$$(z^\perp)'(\xi^j - \xi^1) = 1 > 0 \;\text{ for } j \in S^2 \setminus \{1\}.$$

The collection of vectors $\{\delta^i - \delta^1 : i \in S^1\} \cup \{\xi^j - \xi^1 : j \in S^2\}$ describes the motions along all edges of the convex set Ξ emanating from state $\xi^\dagger$. Thus the fact that their inner products with $z^\perp$ are all positive implies that the translation of $E^c \oplus E^s$ to $\xi^\dagger$ is a hyperplane that supports Ξ at $\xi^\dagger$.

For $d \geq 4$, the affine set through $\xi^\dagger$ defined by the eigenvectors with zero eigenvalue consists entirely of rest points. This fact and Corollary 3.3 of [38] imply that this affine set is the unique center manifold through $\xi^\dagger$. □

Proof of Proposition 3. The proof follows closely [33]. For completeness we reproduce here the required results. 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 A.1. [33] *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*

$$v(t) \geq \mathrm{e}^{-t}\left(v(0) + \int_0^t \mathrm{e}^s\, a(s)\,\mathrm{d}s\right) \quad \textit{for all } t \in [0,T]. \tag{7}$$

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 A.1 to prove inward motion from the boundary under test-two 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). Let $\{\xi(t)\}_{t\geq 0}$ be the solution to the dynamic with initial condition $\xi(0)$. Let $\mathcal{S}_0 = \mathrm{supp}(\xi(0))$ and $Q = \frac{1}{2}\min\{\xi_h(0) \colon h \in \mathcal{S}_0\}$, and, finally, let $\mathcal{S}_1 = \{i \in \mathcal{S} \setminus \mathcal{S}_0 \colon r_i(\xi(0)) > 0\}$ 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 $r_i(\chi) \geq R$ 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 A.1 shows that

$$\xi_i(t) \geq R(1 - \mathrm{e}^{-t}) \text{ for all } t \in [0, T] \text{ and } i \in \mathcal{S}_1. \tag{8}$$

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 A.1 imply that

$$\xi_j(t) \geq c\, Q^a\, \mathrm{e}^{-t} \int_0^t \mathrm{e}^s (R(1 - \mathrm{e}^{-s}))^b \,\mathrm{d}s \text{ for all } t \in [0, T]. \tag{9}$$

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 test-two dynamic from states $\xi(0) \neq \xi^\dagger$ immediately enter $\mathrm{int}(\Xi)$ 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.

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; a revising agent tests strategy i with probability $c \geq \frac{1}{max(s^1,s^2)-1}$; then, if both of the tested strategies (i and another strategy) are tested against opponents playing h, strategy i is selected, as it is the unique best response to opponents playing h. Let $\mathcal{S}_{\mathrm{I}}$ denote the set of strategies added during this stage. The assumption that $\xi(0) \neq \xi^\dagger$ implies that $\mathcal{S}_0 \cup \mathcal{S}_{\mathrm{I}}$ contains $1^{[1]}$, $1^{[2]}$, and $2^{[1]}$.

(II) Second, we add the strategies $j \in S^2 \setminus (\mathcal{S}_0 \cup \mathcal{S}_{\mathrm{I}})$. We can do so because j is tested by a revising agent with probability $c \geq \frac{1}{max(s^1,s^2)-1}$ and because j provides the highest payoff when it is tested against $2^{[1]}$ and the alternative strategy is tested against $1^{[1]}$.

(III) Third, we add the strategies $k \in S^1 \setminus (\mathcal{S}_0 \cup \mathcal{S}_{\mathrm{I}})$. We can do so because k is tested by a revising agent with probability $c \geq \frac{1}{max(s^1,s^2)-1}$ and because k provides the highest payoff when it is tested against $2^{[2]}$ and the alternative strategy is tested against $1^{[2]}$. $\square$

A.2. **Proofs of computational results.**

Proof of Proposition 4. The exact rest points have been obtained by computing a Gröbner basis for the system of polynomials with rational coefficients (1), following the methods described in detail in [33, appendix C and section I in the supplementary appendix]. The only difference with [33] is that here we have to impose the additional constraint $x_1 \neq 1$ to get a polynomial system with a finite solution set. This was done by adding the auxiliary variable z and the polynomial $(x_1-1)*z = 1$.

The code used to compute these rest points has been implemented in the open-source *Mathematica* notebook which is freely available at https://github.com/luis-r-izquierdo/bep-centipede. In particular, the function implemented for this purpose is `ExactRestPoints`. To obtain all the rest points of the mean dynamic (1) in a Centipede game with d decision nodes, it is sufficient to run the following code in the *Mathematica* notebook:

```
ExactRestPoints["two", "stick", d]
```

Using this function the program computes the exact rest points up to $d = 8$. Table 2 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 2.1. The numbers reported in the table are decimal approximations, since the exact values are algebraic numbers that do not admit an exact decimal representation.

To prove that the interior rest point of the mean dynamic (1) is asymptotically stable, in principle we could use linearization. However, since the components of ξ^* are algebraic numbers, computing the eigenvalues of the Jacobian of the dynamic at the interior rest point requires finding the exact roots of a polynomial with algebraic coefficients, a computationally intensive problem. Fortunately, we can prove local stability without doing so using an eigenvalue perturbation theorem [13, Observation 6.3.1], as detailed in [33, Appendix C]. The approximate eigenvalues are reported in Table 3.

The computational function implemented to prove stability, whose code is open for inspection, is `LocalStabilityOfInteriorRestPoint`. To replicate this computational proof for a Centipede game with d decision nodes, it is sufficient to run the following code in the *Mathematica* notebook:

```
LocalStabilityOfInteriorRestPoint["two", "stick", d]
```
□

Proof of Proposition 5. To argue that the interior rest point $\xi^* = (x^*, y^*)$ is almost globally stable we introduce the candidate Lyapunov function

$$L(x, y) = \sum_{i=2}^{s_1} (x_i - x_i^*)^2 + \sum_{j=2}^{s_2} (y_j - y_j^*)^2. \tag{10}$$

In words, $L(x, y)$ is the squared Euclidean distance from (x, y) to (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. For the Centipede game of lengths 2 and 3, we are able to verify that L is a Lyapunov function using an algorithm from real algebraic geometry called *cylindrical algebraic decomposition* [6]. However, exact implementations of this algorithm fail to terminate in longer games.

The computational function implemented to prove almost global stability, whose code is open for inspection, is `GlobalStabilityOfInteriorRestPoint`. To replicate this computational proof for a Centipede game with d decision nodes, it is sufficient to run the following code in the *Mathematica* notebook:

```
GlobalStabilityOfInteriorRestPoint["two", "stick", d]
```
□

Table 2. The interior rest point $\xi^* = \xi^*(d)$ of the dynamic for Centipede of lengths $d \in \{2, \ldots, 20\}$. p denotes the penultimate player, q the last player. The dashed lines separated exact ($d \leq 8$) from numerical ($d \geq 9$) results. The numbers shown are approximations, since the exact values are algebraic numbers that do not admit an exact decimal representation.

p	[7]	[6]	[5]	[4]	[3]	[2]	[1]	[0]
2	-	-	-	-	-	-	.618034	.381966
3	-	-	-	-	-	-	.539189	.460811
4	-	-	-	-	-	.208426	.411450	.380124
5	-	-	-	-	-	.223867	.398692	.377441
6	-	-	-	-	.035722	.223253	.378763	.362262
7	-	-	-	-	.040882	.225279	.374384	.359455
8	-	-	-	.002980	.042792	.225384	.371574	.357271
9	-	-	-	.003239	.043396	.225559	.370966	.356839
10	-	-	.000138	.003311	.043558	.225576	.370747	.356670
11	-	-	.000145	.003327	.043595	.225585	.370707	.356641
12	-	4.19×10^{-6}	.000147	.003330	.043603	.225586	.370697	.356633
13	-	4.32×10^{-6}	.000147	.003331	.043604	.225586	.370695	.356632
14	9.04×10^{-8}	4.34×10^{-6}	.000147	.003331	.043604	.225586	.370695	.356632
15	9.24×10^{-8}	4.34×10^{-6}	.000147	.003331	.043604	.225586	.370695	.356632
16	9.27×10^{-8}	4.34×10^{-6}	.000147	.003331	.043604	.225586	.370695	.356632
17	9.28×10^{-8}	4.34×10^{-6}	.000147	.003331	.043604	.225586	.370695	.356632
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
20	9.28×10^{-8}	4.34×10^{-6}	.000147	.003331	.043604	.225586	.370695	.356632

q	[7]	[6]	[5]	[4]	[3]	[2]	[1]	[0]
2	-	-	-	-	-	-	.618034	.381966
3	-	-	-	-	-	.369102	.369102	.261795
4	-	-	-	-	-	.344955	.364555	.290490
5	-	-	-	-	.087713	.310211	.329668	.272409
6	-	-	-	-	.100021	.304394	.323241	.272345
7	-	-	-	.010544	.104027	.298920	.317193	.269316
8	-	-	-	.011813	.105888	.297664	.315745	.268891
9	-	-	.000650	.012191	.106378	.297094	.315103	.268585
10	-	-	.000692	.012297	.106528	.296977	.314969	.268537
11	-	2.42×10^{-5}	.000701	.012321	.106559	.296944	.314931	.268520
12	-	2.51×10^{-5}	.000703	.012326	.106566	.296938	.314925	.268518
13	6.17×10^{-7}	2.53×10^{-5}	.000703	.012327	.106567	.296937	.314923	.268517
14	6.33×10^{-7}	2.53×10^{-5}	.000703	.012327	.106567	.296936	.314923	.268517
15	6.35×10^{-7}	2.53×10^{-5}	.000703	.012327	.106567	.296936	.314923	.268517
16	6.36×10^{-7}	2.53×10^{-5}	.000703	.012327	.106567	.296936	.314923	.268517
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
20	6.36×10^{-7}	2.53×10^{-5}	.000703	.012327	.106567	.296936	.314923	.268517

TABLE 3. Approximate eigenvalues of $DV(\xi^*)$ for the dynamic.

$d=2$	$-.8090 \pm .4468\,i$											
$d=3$	$-.9071$	$-.6556 \pm .3376\,i$										
$d=4$	$-.8715 \pm .1608\,i$	$-.5238 \pm .2572\,i$										
$d=5$	$-.6851 \pm .1242\,i$	$-.4412 \pm .1803\,i$	-1.0845									
$d=6$	$-.5778 \pm .1093\,i$	$-.3687 \pm .1286\,i$	-1.1369	$-.8289$								
$d=7$	$-.4762 \pm .0829\,i$	$-.3145 \pm .0920\,i$	-1.2171	$-.8220$	$-.7632$							
$d=8$	$-.4262 \pm .0657\,i$	$-.2781 \pm .0620\,i$	-1.2194	$-.9406$	$-.7827$	$-.5425$						
$d=9$	$-.3647 \pm .0549\,i$	$-.2471 \pm .0516\,i$	-1.2203	$-.9703$	$-.8741$	$-.6137$	$-.5069$					
$d=10$	$-.3394 \pm .0404\,i$	$-.2248 \pm .0383\,i$	-1.1978	$-.9917$	$-.9230$	$-.6928$	$-.5786$	$-.4027$				
$d=11$	$-.2982 \pm .0404\,i$	$-.2040 \pm .0378\,i$	-1.1856	$-.9975$	$-.9427$	$-.7738$	$-.6353$	$-.4986$	$-.3874$			
$d=12$	$-.2835 \pm .0306\,i$	$-.1878 \pm .0305\,i$	-1.1666	$-.9995$	$-.9562$	$-.8129$	$-.6930$	$-.5637$	$-.4666$	$-.3280$		
$d=13$	$-.2531 \pm .0331\,i$	$-.1731 \pm .0312\,i$	-1.1568	$-.9999$	$-.9583$	$-.8390$	$-.7301$	$-.6389$	$-.5184$	$-.4212$	$-.3214$	
$d=14$	$-.2431 \pm .0260\,i$	$-.1610 \pm .0261\,i$	-1.1429	-1.0000	$-.9637$	$-.8555$	$-.7466$	$-.6921$	$-.5724$	$-.4800$	$-.3974$	$-.2804$

A.3. **Procedures to obtain numerical results on the deterministic approximations.**

A.3.1. *Numerical results for the dynamics with one trial.* The numerical estimates of rest points for $d > 8$ have been computed following the methods described in detail in [33, section II.3 in the supplementary appendix]. The code used to compute these rest points has been implemented in the open-source *Mathematica* notebook, in the function `RationalApproximateRestPoint`. This function computes a rational approximation of the stable interior rest point using exact arithmetic. To replicate this computation in a Centipede game with d decision nodes, it is sufficient to run the following code in the *Mathematica* notebook:

```
RationalApproximateRestPoint["two", "stick", d]
```

Table 2 reports the approximate values of the interior rest points $\xi^* = \xi^*(d)$ up to $d = 20$, but the pattern continues for longer games.

The numerical estimates of the eigenvalues of the Jacobian of the dynamic at the interior rest point have been computed using the function `EigenvaluesAtRationalApproximateRestPoint`. This function computes the exact eigenvalues of the Jacobian of the dynamic at the rational approximation to the interior rest point obtained from a call to `RationalApproximateRestPoint`. To replicate this computation in a Centipede game with d decision nodes, it is sufficient to run the following code in the *Mathematica* notebook:

```
EigenvaluesAtRationalApproximateRestPoint["two", "stick", d]
```

Table 3 reports the approximate values of the eigenvalues up to $d = 20$.

We also conducted an extensive numerical computation that suggests that proposition 5 (which proves that the interior point is an almost globally asymptotically stable state for lengths $2 \leq d \leq 3$) extends to much longer Centipede games. Specifically, we tried to verify numerically that the squared Euclidean distance to the interior rest point (x^*, y^*), i.e. function $W(x, y)$ (11), is a Lyapunov function for longer games.

$$W(x,y) = \sum_{i=1}^{s_1}(x_i - x_i^*)^2 + \sum_{j=1}^{s_2}(y_j - y_j^*)^2 \tag{11}$$

In particular, 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 $\dot{W}$ at each point. In all instances, the approximate version of $\dot{W}$ evaluated to a negative number. This numerical approach is much stronger (and much faster) than one based on the computation of numerical solutions to the differential equation (1): it not only avoids the numerical errors inherent in obtaining approximate solutions to (1), but also provides evidence about the global structure of the dynamics.

The evaluation of the floating-point approximation of $\dot{L}$ at various points can be replicated using the function `NumericalGlobalStabilityOfInteriorRestPointLyapunov`.

A.3.2. *Numerical results for the dynamics with several trials.* The numerical exploration of the mean dynamic with several trials has been conducted using the function `NDSolveMeanDynamicsManyTrials`. This function uses *Mathematica*'s `NDSolve` function to compute a numerical solution of the mean dynamic (2), where the number of trials κ and the initial condition of the solution can be specified by the user. The solution is computed until the time at which the norm of the law of motion is sufficiently small (where what constitutes sufficiently small can be chosen by the user) or until a maximum time chosen by the user is reached. The function also graphs the components of the state as a function of time, reports the terminal point and the time at which this point is reached, and integrates the expected duration of play over the solution trajectory. This function was used in producing figs. 3, 4 and 10. As an example, the solution trajectories shown in fig. 10 can be obtained by running the following code:

```
NDSolveMeanDynamicsManyTrials["two", "stick", 4, k],
```

where `k` must be replaced with the desired number of trials.

A.4. **Procedure to obtain numerical results from simulations with finite populations.** The agent-based model used to simulate the *test-two dynamics* on finite populations has been implemented in the open-source platform *NetLogo* [41] and is freely available at https://luis-r-izquierdo.github.io/centipede-test-two.

The reported values for the estimated cooperative regime of the *test-two dynamic* are calculated using one long run in every case, starting from random initial conditions. We make sure that the obtained sample corresponds to a single regime by checking that various time-averaged statistics present stability along the sample.

Concretely, to report average values for the test-two dynamic, in each run, we let the process evolve for 10^4 time steps before measuring values, to let the effect of the random initial condition fade away.[16] We then measure values for 10^5 time steps. To check that the considered time framework captures a stable regime in every case (stable on average along time, as the population state is always changing and, for large κ, the dynamic often presents a cyclical character), and that the process does not present a change of regime during the simulated run, we calculate at every time step, according to the population state, the expected fraction f_j of matches with duration of play (reached terminal node) $j \in \{1, ..., d+1\}$. We then compare the average values of f_j corresponding to the last 10% periods of the sample with the average ones corresponding to the whole sample. In most of the simulated runs, the differences between the two average values for f_j obtained this way were below 1% for every j. In fact, only for the short centipede ($d = 4$), smallest population size ($N = 10$) and number of trials $\kappa = 6$, in fig. 8, we represent a value from a run that changed regime and got absorbed at an stop-immediately state during the reporting periods (between periods 10^4 and 11×10^4).

A.5. **Analysis of the cycling behavior for large number of trials.** In this section we discuss with more detail the cycling behavior observed for large number of trials, taking as a reference a Centipede with 10 decision nodes, 100 players in each population, and number of trials $\kappa = 99$ (see fig. 14). Recall that, for $i \in \{1, ..., 5\}$, strategy i means "stop at your i-th decision node, and not before", and strategy 6 corresponds to "always continue".

Looking at a random simulation when approaching the low-duration-of-play zone of the typical cycle-like patterns that are observed for high κ, we observe a population state such that the number of players using strategies $(1, ..., 6)$ is (0, 66, 34, 0, 0, 0) in the player 1 population, and (4, 89, 7, 0, 0, 0) in the player 2 population. At the considered state, we can calculate the total payoffs that each revising agent can obtain when testing each possible strategy, as well as the probabilities of obtaining

[16] The evolution of the process towards the attractive cooperative regime is usually much quicker than that, as can be seen in figs. 9 and 14.

those payoffs, leading to table 4, where we denote strategy i for player 1 as i_x and for player 2 as i_y.

Table 4. Possible payoffs obtained by each strategy, with their probabilities, at a state near the bottom of a cycle

Str	#	**0.04**	**0.89**	**0.07**	Str	#	**0.66**	**0.34**
1_x		0	0	0	1_y	4	297	297
2_x	66	189	186	186	2_y	89	300	297
3_x	34	114	112	109	3_y	7	266	264
4_x		107	105	103	4_y		266	264
5_x		107	105	103	5_y		266	264
6_x		107	105	103	6_y		266	264

Str: strategy; #: number of players using the corresponding strategy. For each strategy, the values on the right of the # column are the possible total payoffs obtainable at the considered state by the corresponding strategy in 99 trials. Top numbers in bold: probability of obtaining the payoffs on that column, at the considered state.

When those individuals who are the last ones to stop (in the considered state, those playing 3_y) try another strategy that stops at a later stage or does not stop, in this case 4_y, 5_y or 6_y, those strategies are at no disadvantage with respect to 3_y, and the test using them can provide a higher payoff than the test using 3_y with a considerably high probability ($0.34 \cdot 0.66 = 0.22$). This shows that it is very likely that some of those individuals using 3_y will change to using 4_y, 5_y or 6_y, and, more generally, it illustrates how likely it is that cooperative strategies appear in the population as the unraveling process develops towards the backward induction state. Once we have some strategies stopping at the last decision nodes in one population, the other population follows easily. In our example, after some simulation steps we observe a population state such that the number of players using strategies $(1, ..., 6)$ is (0, 78, 15, 2, 2, 3) in the player 1 population, and (20, 75, 2, 1, 1, 1) in the player 2 population. At this state we obtain table 5

Table 5. Possible payoffs obtained by each strategy, with their probabilities, at a state leaving the bottom of a cycle

Str	#	**0.2**	**0.75**	**0.02**	**0.01**	**0.01**	**0.01**	Str	#	**0.78**	**0.15**	**0.02**	**0.02**	**0.03**
1_x		0	0	0	0	0	0	1_y	20	297	297	297	297	297
2_x	78	141	138	138	138	138	138	2_y	75	264	261	261	261	261
3_x	15	76	74	71	71	71	71	3_y	2	263	261	258	258	258
4_x	2	80	78	76	73	73	73	4_y	1	271	269	267	264	264
5_x	2	83	81	79	77	74	74	5_y	1	275	273	271	269	266
6_x	3	84	82	80	78	76	73	6_y	1	272	270	268	266	264

Str: strategy; #: number of players using the corresponding strategy. For each strategy, the values on the right of the # column are the possible total payoffs obtainable at the considered state by the corresponding strategy in 99 trials. Top numbers in bold: probability of obtaining the payoffs on that column, at the considered state.

The first two columns of possible payoff values for the player 1 population show that revising players using 3_x will likely adopt a strategy that stops later, and, looking at the first two columns of possible payoff values for the player 2 population, it can be seen that the same happens for the majority of the player 2 population, who are using strategy 2_y: if they test 4_y, 5_y or 6_y there is a probability of at least 0.95 that the corresponding strategy will be adopted. After a few moves in that direction, it soon becomes advantageous in population one to use the last strategies too, instead of 2_x, which quickly leads the process towards the last strategies in both populations (see table 6).

Table 6. Possible payoffs obtained by each strategy, with their probabilities, at a state leaving the bottom of a cycle

Str	#	**0.24**	**0.55**	**0.03**	**0.04**	**0.08**	**0.06**	Str	#	**0.73**	**0.10**	**0.04**	**0.04**	**0.09**
1_x		0	0	0	0	0	0	1_y	24	297	297	297	297	297
2_x	73	129	126	126	126	126	126	2_y	55	279	276	276	276	276
3_x	10	116	114	111	111	111	111	3_y	3	303	301	298	298	298
4_x	4	149	147	145	142	142	142	4_y	4	325	323	321	318	318
5_x	4	173	171	169	167	164	164	5_y	8	339	337	335	333	330
6_x	9	177	175	173	171	169	166	6_y	6	330	328	326	324	322

Str: strategy; #: number of players using the corresponding strategy. For each strategy, the values on the right of the # column are the possible total payoffs obtainable at the considered state by the corresponding strategy in 99 trials. Top numbers in bold: probability of obtaining the payoffs on that column, at the considered state.

A.6. **Initial effect of increasing the number of trials.** In this section we discuss the initial effect that increasing the number of trials κ has on the expected duration of play, averaged over the empirical distribution on population states at the cooperative regime (see figs. 8 and 13), and the corresponding effect in the deterministic approximation (see fig. 4).

To explain this effect, i.e., the increase on the expected game duration when moving from $\kappa = 1$ to $\kappa = 2$, consider a centipede with d decision nodes and $\kappa = 1$. For this example, let us refer to strategies using the last-to-first notation $[k]$ introduced in section 2.1. The payoffs depending on the stopping strategy are shown in table 7. Suppose (as it will be often the case near the attractor) that all players are stopping at one of their last 5 decision nodes: the strategies being played are $[0]_x, [1]_x, \ldots, [4]_x, [0]_y, [1]_y, \ldots, [4]_y$ and the duration of play is between $d-7$ and $d+1$. Then (see table 7),

- All the strategies being played in population 1 obtain a payoff $\pi \geq d-9$, while any alternative strategy for player 1 obtains a payoff $\pi \leq d-10$ and consequently cannot be adopted.
- All the strategies being played in population 2 obtain a payoff $\pi \geq d-8$, with $[5]_y$ being the only other strategy that, if tested, could be adopted by a revising player 2, as it obtains a payoff $\pi = d-7$.

But for $[5]_y$ to be adopted, the revising player 2 when testing his current strategy must obtain a payoff $\pi = d-8$, i.e., it must meet a $[4]_x$ player. The probability of that event is $x_{[4]}$, the fraction of $[4]_x$ players in population 1.

TABLE 7. Payoff obtained by each player depending on the stopping strategy

Duration of play	$\cdots$	$d-10$	$d-9$	$d-8$	$d-7$	$d-6$	$d-5$	$d-4$	$d-3$	$d-2$	$d-1$	d	$d+1$
Payoff player 1	$\cdots$	$d-13$	$d-10$	$d-11$	$d-8$	$d-9$	$d-6$	$d-7$	$d-4$	$d-5$	$d-2$	$d-3$	d
Payoff player 2	$\cdots$	$d-9$	$d-10$	$d-7$	$d-8$	$d-5$	$d-6$	$d-3$	$d-4$	$d-1$	$d-2$	$d+1$	d
Stopping strategy	$\cdots$	$[6]_y$	$[5]_x$	$[5]_y$	$[4]_x$	$[4]_y$	$[3]_x$	$[3]_y$	$[2]_x$	$[2]_y$	$[1]_x$	$[1]_y$	none

TABLE 8. Payoff obtained by player 2 depending on the stopping strategies, two trials

Stopping strategy	1^{st} trial:	$[4]_x$	$[4]_y$	$[3]_x$	$[3]_y$	$[2]_x$	$[2]_y$	$[1]_x$	$[1]_y$	none
2^{nd} trial:										
$[4]_x$		$2d-16$	$2d-13$	$2d-14$	$2d-11$	$2d-12$	$2d-9$	$2d-10$	$2d-7$	$2d-8$
$[4]_y$		$2d-13$	$2d-10$	$2d-11$	$2d-8$	$2d-9$	$2d-6$	$2d-7$	$2d-4$	$2d-5$
$[3]_x$		$2d-14$	$2d-11$	$2d-12$	$2d-9$	$2d-10$	$2d-7$	$2d-8$	$2d-5$	$2d-6$
$[3]_y$		$2d-11$	$2d-8$	$2d-9$	$2d-6$	$2d-7$	$2d-4$	$2d-5$	$2d-2$	$2d-3$
$[2]_x$		$2d-12$	$2d-9$	$2d-10$	$2d-7$	$2d-8$	$2d-5$	$2d-6$	$2d-3$	$2d-4$
$[2]_y$		$2d-9$	$2d-6$	$2d-7$	$2d-4$	$2d-5$	$2d-2$	$2d-3$	$2d$	$2d-1$
$[1]_x$		$2d-10$	$2d-7$	$2d-8$	$2d-5$	$2d-6$	$2d-3$	$2d-4$	$2d-1$	$2d-2$
$[1]_y$		$2d-7$	$2d-4$	$2d-5$	$2d-2$	$2d-3$	$2d$	$2d-1$	$2d+2$	$2d+1$
$none$		$2d-8$	$2d-5$	$2d-6$	$2d-3$	$2d-4$	$2d-1$	$2d-2$	$2d+1$	$2d$

First row: stopping strategy at the first trial. First column: stopping strategy at the second trial.

TABLE 9. Payoff obtained by each player depending on the stopping strategy

Duration of play	$\cdots$	$d-10$	$d-9$	$d-8$	$d-7$	$d-6$	$d-5$	$d-4$	$d-3$	$d-2$	$d-1$	d	$d+1$
Payoff player 1	$\cdots$	-1	5	4	10	9	15	14	20	19	25	24	30
Payoff player 2	$\cdots$	6	5	11	10	16	15	21	20	26	25	31	30
Stopping strategy	$\cdots$	$[6]_y$	$[5]_x$	$[5]_y$	$[4]_x$	$[4]_y$	$[3]_x$	$[3]_y$	$[2]_x$	$[2]_y$	$[1]_x$	$[1]_y$	none

On the other hand, if the number of trials is 2, strategy $[5]_y$ would obtain a total payoff $\pi = 2d - 14$ and, given that the strategies being played $([0]_y, [1]_y, \ldots, [4]_y)$ would obtain one of the payoffs shown in table 8 (at the crossing cell of the duration of play in each trial), the only way strategy $[5]_y$ can be adopted is if player 2 when testing his current strategy meets a $[4]_x$ player in both trials, which is usually much more unlikely than meeting a $[4]_x$ player in one trial, as the first probability is $x_{[4]} \frac{N\, x_{[4]} - 1}{N-1} < x_{[4]}^2$ while the second probability is $x_{[4]}$.

This increased difficulty to be selected when $\kappa = 2$ (vs. $\kappa = 1$) for the strategies that are near the most frequent ones at the equilibrium but are less cooperative explains the quicker decay in the fraction of less-cooperative strategies and the associated difference between the expected game duration obtained for $\kappa = 1$ and $\kappa = 2$. This effect does not go on beyond $\kappa = 2$ because from $\kappa = 3$ there are additional events that can make strategy $[5]_y$ beat the strategies that stop later, besides the event that they meet a $[4]_x$ player in all their trials. For instance, for $\kappa = 3$, and taking, e.g., $d = 20$ (just to work with numbers for the payoffs instead of letters), $[5]_y$ would obtain a payoff $\pi = 3 \cdot 13 = 39$, and it would be selected if the current strategy being tested meets a $[4]_x$ player three times (obtaining $3 \cdot 12 = 36 < 39$), but also if the current strategy being tested meets $[4]_x$ twice and $[3]_x$ once, obtaining a payoff $\pi = 12 + 12 + 14 = 38 < 39$.

The effect just described depends on the cost/gain relation of the Centipede. Let us look at the change in payoffs after a player chooses to continue as the combination of a gain b for each of the players and a cost c incurred only by the player who continues. The standard payoffs that we have considered here (see section 2.1) correspond to a cost of continuing $c = 4$ and a gain for each partner from a continuation $b = 3$ (so the net cost of the decision to continue is $b - c = -1$). If the cost of continuing is $c = 7$ and the shared gain for each partner from a continuation is $b = 6$, then the payoff structure is as shown in table 9.

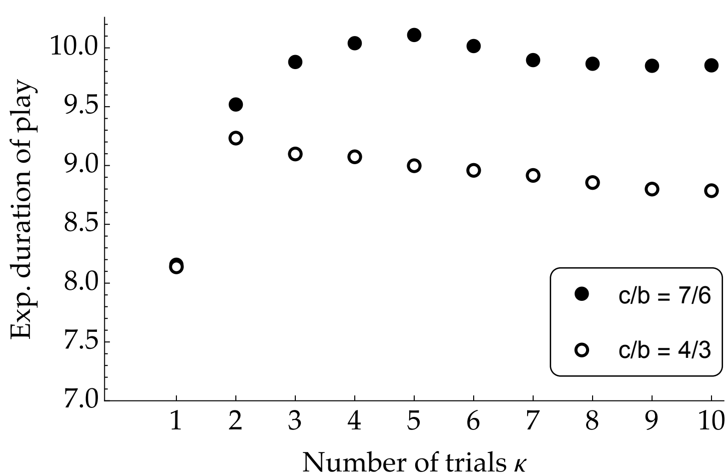


FIGURE 15. Expected duration of play in two 10-node Centipede games with different cost/gain ratios c/b, averaged over the empirical distribution on population states, for different number of trials κ. $N = 50$.

In this case, if the strategies being played are $[0]_x, [1]_x, \ldots, [4]_x, [0]_y, [1]_y, \ldots, [4]_y$, the payoff when testing $[5]_y$ would be 11 per match, making a total payoff $\pi = 11 \cdot \kappa$, while the payoff using the current strategy being played by player 2 would be 10 per match if meeting $[4]_x$ or at least 15 per match otherwise, leading to a total payoff $\pi = 10 \cdot \kappa$ if meeting $[4]_x$ at every trial, or at least $\pi = 10 \cdot \kappa + 5$ otherwise. This means that playing with an $[4]_x$ player gives a payoff advantage of 1 to $[5]_y$ compared to the player 2 strategies that stop later than $[5]_y$, but those other strategies get a payoff advantage of at least 5 over $[5]_y$ if they play with a player 1 that stops after $[4]_x$. Then, for $\kappa \leq 5$, the chances of strategy $[5]_y$ being chosen always correspond to just one possible event, i.e., that the current strategy meets a $[4]_x$ strategist on all its κ trials; and this event has a probability that decreases steeply with κ. As before, that decline in probabilities is not so steep for more than 5 trials, since in that case there are more events leading to $[5]_y$ being selected. We can then expect the average duration of play to be increasing with κ up to $\kappa = 5$. This is corroborated by the data shown in Fig 15.

REFERENCES

[1] E. Ben-Porath, Rationality, Nash Equilibrium and Backwards Induction in Perfect- Information Games, *The Review of Economic Studies*, **64** (1997), 23, doi:10.2307/2971739.

[2] K. G. Binmore, *Natural justice*, Oxford University Press, 2011.

[3] K. Binmore, Modeling rational players: Part I, *Economics and Philosophy*, **3** (1987), 179–214, doi:10.1017/S0266267100002893.

[4] K. Binmore and L. Samuelson, An Economist's Perspective on the Evolution of Norms, *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft*, **150** (1994), 45–63, URL `https://www.jstor.org/stable/40753015`.

[5] K. G. Binmore, L. Samuelson and R. Vaughan, Musical Chairs: Modeling Noisy Evolution, *Games and Economic Behavior*, **11** (1995), 1–35, doi:10.1006/GAME.1995.1039.

[6] G. E. Collins, Quantifier elimination for the theory of real closed fields by cylindrical algebraic decomposition, in *Second GI Conference on Automata Theory and Formal Languages*, vol. 33 of Lecture Notes in Computer Science, Springer, Berlin, 1975, 134–183, doi:10.1007/3-540-07407-4_17.

[7] J. C. Cox and D. James, On replication and perturbation of the McKelvey and Palfrey centipede game experiment, in *Replication in Experimental Economics (Research in Experimental Economics, Volume 18)* (eds. C. A. Deck, E. Fatas and T. Rosenblat), vol. 18, Emerald Group Publishing Ltd., 2015, 53–94, doi:10.1108/S0193-230620150000018003.

[8] R. Cressman and K. H. Schlag, The Dynamic (In)Stability of Backwards Induction, *Journal of Economic Theory*, **83** (1998), 260–285, doi:10.1006/jeth.1996.2465.

[9] M. Embrey, G. R. Fréchette and S. Yuksel, Cooperation in the Finitely Repeated Prisoner's Dilemma, *The Quarterly Journal of Economics*, **133** (2018), 509–551, doi:10.1093/qje/qjx033.

[10] I. Gilboa and A. Matsui, Social stability and equilibrium, *Econometrica*, **59** (1991), 859–867, doi:10.2307/2938230.

[11] J. Y. Halpern, Substantive Rationality and Backward Induction, *Games and Economic Behavior*, **37** (2001), 425–435, doi:10.1006/GAME.2000.0838.

[12] J. Hofbauer, Stability for the best response dynamics, 1995, Unpublished manuscript, University of Vienna.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd edition, Cambridge University Press, New York, 2013.

[14] L. R. Izquierdo, S. S. Izquierdo and W. H. Sandholm, EvoDyn-3s: A mathematica computable document to analyze evolutionary dynamics in 3-strategy games, *SoftwareX*, **7** (2018), 226 – 233, doi:10.1016/j.softx.2018.07.006.

[15] L. R. Izquierdo, S. S. Izquierdo and W. H. Sandholm, An introduction to ABED: Agent-based simulation of evolutionary game dynamics, *Games and Economic Behavior*, **118** (2019), 434–462, doi:10.1016/j.geb.2019.09.014.

[16] A. Kelley, Stability of the center-stable manifold, *Journal of Mathematical Analysis and Applications*, **18** (1967), 336–344, doi:10.1016/0022-247X(67)90061-3.

[17] A. Kelley, The stable, center-stable, center, center-unstable, and unstable manifolds, *Journal of Differential Equations*, **3** (1967), 546–570, doi:10.1016/0022-0396(67)90016-2.
[18] D. M. Kreps, P. Milgrom, J. Roberts and R. Wilson, Rational cooperation in the finitely repeated prisoners' dilemma, *Journal of Economic Theory*, **27** (1982), 245–252, doi:10.1016/0022-0531(82)90029-1.
[19] R. D. McKelvey and T. R. Palfrey, An Experimental Study of the Centipede Game, *Econometrica*, **60** (1992), 803–836, doi:10.2307/2951567.
[20] R. D. Mckelvey and T. R. Palfrey, Quantal Response Equilibria for Extensive Form Games, *Experimental Economics*, **1** (1998), 9–41, doi:10.1023/A:1009905800005.
[21] R. Nagel and F. F. Tang, Experimental Results on the Centipede Game in Normal Form: An Investigation on Learning, *Journal of Mathematical Psychology*, **42** (1998), 356–384, doi:10.1006/jmps.1998.1225.
[22] M. J. Osborne and A. Rubinstein, Games with Procedurally Rational Players, *The American Economic Review*, **88** (1998), 834–847, URL `https://www.jstor.org/stable/117008`.
[23] I. Palacios-Huerta and O. Volij, Field Centipedes, *American Economic Review*, **99** (2009), 1619–1635, doi:10.1257/aer.99.4.1619.
[24] A. Perea, Belief in the opponents' future rationality, *Games and Economic Behavior*, **83** (2014), 231–254, doi:10.1016/j.geb.2013.11.008.
[25] L. Perko, *Differential Equations and Dynamical Systems*, Third edition edition, Springer, New York, 2001.
[26] P. Pettit and R. Sugden, The Backward Induction Paradox, *The Journal of Philosophy*, **86** (1989), 169, doi:10.2307/2026960.
[27] G. Ponti, Cycles of Learning in the Centipede Game, *Games and Economic Behavior*, **30** (2000), 115–141, doi:10.1006/game.1998.0707.
[28] B. D. Pulford, E. M. Krockow, A. M. Colman and C. L. Lawrence, Social Value Induction and Cooperation in the Centipede Game, *PLOS ONE*, **11** (2016), e0152352, doi:10.1371/journal.pone.0152352.
[29] P. J. Reny, Backward Induction, Normal Form Perfection and Explicable Equilibria, *Econometrica*, **60** (1992), 627, doi:10.2307/2951586.
[30] R. W. Rosenthal, Games of perfect information, predatory pricing and the chain-store paradox, *Journal of Economic Theory*, **25** (1981), 92–100, doi:10.1016/0022-0531(81)90018-1.
[31] W. H. Sandholm, Evolution and equilibrium under inexact information, *Games and Economic Behavior*, **44** (2003), 343–378, doi:10.1016/S0899-8256(03)00026-5.
[32] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, MIT Press, Cambridge, 2010.
[33] W. H. Sandholm, S. S. Izquierdo and L. R. Izquierdo, Best experienced payoff dynamics and cooperation in the Centipede game, *Theoretical Economics*, **14** (2019), 1347–1386, doi:10.3982/TE3565.
[34] W. H. Sandholm, S. S. Izquierdo and L. R. Izquierdo, Stability for best experienced payoff dynamics, *Journal of Economic Theory*, **185** (2020), 104957, doi:10.1016/j.jet.2019.104957.
[35] R. Selten, Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragetragheit, *Zeitschrift für die Gesamte Staatswissenschaft*, **121** (1965), 301–324; 667–689, URL `http://www.jstor.org/stable/40748884`.
[36] R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory*, **4** (1975), 25–55, doi:10.1007/BF01766400.
[37] R. Sethi, Stability of Equilibria in Games with Procedurally Rational Players, *Games and Economic Behavior*, **32** (2000), 85–104, doi:10.1006/GAME.1999.0753.
[38] J. Sijbrand, Properties of center manifolds, *Transactions of the American Mathematical Society*, **289** (1985), 431–469, doi:10.2307/2000247.
[39] R. Smead, The evolution of cooperation in the centipede game with finite populations*, *Philosophy of Science*, **75** (2008), 157–177, doi:10.1086/590197.
[40] R. Stalnaker, Knowledge, Belief and Counterfactual Reasoning in Games, *Economics and Philosophy*, **12** (1996), 133, doi:10.1017/S0266267100004132.
[41] U. Wilensky, Netlogo, Software. `http://ccl.northwestern.edu/netlogo/`. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL., 1999.
[42] Z. Xu, Convergence of best-response dynamics in extensive-form games, *Journal of Economic Theory*, **162** (2016), 21–54, doi:10.1016/j.jet.2015.12.001.

[43] D. Zusai, Gains in evolutionary dynamics: A unifying and intuitive approach to linking static and dynamic stability, 2018, URL https://arxiv.org/abs/1805.04898, Unpublished manuscript, Temple University.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: segis@eii.uva.es
E-mail address: lrizquierdo@ubu.es