



Evolutionary dynamics in finite populations can explain the full range of cooperative behaviors observed in the centipede game

David G. Rand^{a,b,*}, Martin A. Nowak^{a,c,d}

^a Program for Evolutionary Dynamics, Harvard University, Cambridge MA, USA

^b Department of Psychology, Harvard University, Cambridge, MA, USA

^c Department of Mathematics, Harvard University, Cambridge, MA, USA

^d Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

ARTICLE INFO

Article history:

Received 31 May 2011

Received in revised form

13 December 2011

Accepted 5 January 2012

Available online 14 January 2012

Keywords:

Fixed length games

Backwards induction

Weak selection

Evolutionary game theory

Evolution of cooperation

ABSTRACT

Classical economic models make behavioral predictions based on the assumption that people are fully rational and care only about maximizing their own payoffs. Although this approach successfully explains human behavior in many situations, there is a wealth of experimental evidence demonstrating conditions where people deviate from the predictions of these models. One setting that has received particular attention is fixed length repeated games. Iterating a social dilemma can promote cooperation through direct reciprocity, even if it is common knowledge that all players are rational and self-interested. However, this is not the case if the length of the game is known to the players. In the final round, a rational player will defect, because there is no future to be concerned with. But if you know the other player will defect in the last round, then you should defect in the second to last round, and so on. This logic of backwards induction leads to immediate defection as the only rational (sub-game perfect Nash equilibrium) strategy. When people actually play such games, however, immediate defection is rare. Here we use evolutionary dynamics in finite populations to study the centipede game, which is designed to explore this issue of backwards induction. We make the following observation: since full cooperation can risk-dominate immediate defection in the centipede game, stochastic evolutionary dynamics can favor both delayed defection and even full cooperation. Furthermore, our evolutionary model can quantitatively reproduce human behavior from two experiments by fitting a single free parameter, which is the product of population size and selection intensity. Thus we provide evidence that people's cooperative behavior in fixed length games, which is often called 'irrational', may in fact be the favored outcome of natural selection.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Evolutionary dynamics provide a powerful set of tools for investigating a range of issues in biology and the social sciences (Colman, 1995; Cressman, 2003; Dasgupta, 2009; Enquist and Leimar, 1993; Helbing and Yu, 2009; Hofbauer and Sigmund, 1998; Maynard Smith, 1982; McNamara et al., 1999; Nowak, 2006; Nowak and Sigmund, 2004). Formalizing evolutionary processes using mathematics gives insight into which genotypes, phenotypes and behaviors can be favored by natural selection. Fitter individuals become more common, while less fit individuals die out. Evolutionary dynamics can describe genetic or cultural evolution. In the latter context, individuals adopt the strategies or preferences of others who are more successful.

The study of cooperation through reciprocal interactions has been a particularly active area of research in both biology and economics (Axelrod and Hamilton, 1981; Fudenberg and Maskin, 1986, 1990; Lindgren, 1991; Nowak and Sigmund, 1992, 1993, 1998, 2005; Ohtsuki and Iwasa, 2006; Pacheco et al., 2006; Rand et al., 2009a; Sigmund, 2010; Trivers, 1971). When individuals interact repeatedly (direct reciprocity) or carry reputations across interactions with different partners (indirect reciprocity), both evolution and rational prospective reasoning can favor helping others at a cost to oneself. Results from many experimental studies are broadly consistent with the implications of these theoretical models (Bolton et al., 2005; Dal Bó, 2005; Dal Bó and Fréchet, 2011; Dreber et al., 2008; Frank et al., 1993; Fudenberg et al., in press; Milinski and Wedekind, 1998; Milinski et al., 2001, 2002a, 2002b; Rand et al., 2009b; Rockenbach and Milinski, 2006; Semmann et al., 2003; Wedekind and Milinski, 2000). The predictions about human behavior in repeated games generated by evolutionary game dynamics have thus far been largely qualitative as opposed to quantitative (an exception is Rand et al. (2009a)), and have typically

* Corresponding author at: Program for Evolutionary Dynamics, Harvard University, Cambridge, MA, USA. Tel.: +1 6075920218.

E-mail address: drand@fas.harvard.edu (D.G. Rand).

been in agreement with the predictions of economic models. In the present paper, we demonstrate that an evolutionary analysis can explain behavior in a context which has traditionally been very challenging for economic models, and furthermore that the evolutionary approach can quantitatively reproduce observed human behavior in the laboratory.

We focus on the ‘centipede game,’ in which players have a fixed number of potentially cooperative interactions (Rosenthal, 1982; Selten, 1978). Such fixed length repeated games have generated much interest and discussion in the economic, psychological and biological literature (Basu, 1990; Binmore, 1987; Cressman and Schlag, 1998; Kreps and Wilson, 1982a, 1982b; Kreps et al., 1982; Ladret and Lessard, 2008; Lessard and Ladret, 2007; McKelvey and Palfrey, 1992; Pettit and Sugden, 1989). Some scholars have argued that these games reveal a fundamental mismatch between the predictions of rational self-interest models and actual human behavior. While direct reciprocity can stabilize cooperation in repeated games among fully informed, rationally self-interested agents (Fudenberg and Maskin, 1986; Nowak and Sigmund, 1992), this is only true when there is uncertainty about when the game will end. If the total number of rounds is known to both players, the logic of backwards induction dooms cooperation. In the final round, defection is always the optimal move because there is no ‘shadow of the future’ to incentivize cooperation. But if both players will certainly defect in the last round, then the actions in the second-to-last round have no repercussions, and so on back to round 1. Therefore, cooperation collapses if rationality of both players is common knowledge (Kreps and Wilson, 1982a, 1982b; Kreps et al., 1982). The only sub-game perfect Nash equilibrium is immediate defection, and this equilibrium is selected by deterministic evolutionary dynamics in the two-player centipede game (Cressman and Schlag, 1998; Giovanni, 2000). Thus both players forgo a large potential payoff. Contrary to such predictions, however, behavioral experiments robustly demonstrate that most human subjects do not defect in the first round of fixed length games (Bornstein et al., 2004; Ho and Weigel, 2005; McKelvey and Palfrey, 1992; Nagel and Tang, 1988).

The centipede game (Rosenthal, 1982; Selten, 1978) was devised to explore this issue of backwards induction. It is similar to an alternating prisoners’ dilemma (Frean, 1994; Nowak and Sigmund, 1994), except that the game ends following the first defection. A particular form of the centipede game which we focus on here is defined such that payoffs grow exponentially with each round of cooperation, as opposed to linearly as in the standard prisoners’ dilemma. At the outset of the centipede game, there are two possible payoffs: a large payoff of a_1 and small payoff of a_0 ($a_1 > a_0$). One player is in control of the game at any given time. In each round, the controlling player has two options: (i) cooperate by passing control to the other player, in which case both payoffs are multiplied by a factor b ; or (ii) defect by taking the larger payoff and giving the smaller payoff to the other player. The game continues until either one player defects, or a fixed number of decisions occur. If the game reaches the final round and the controlling player cooperates, the payoffs are multiplied by b and then the cooperating player receives the smaller payoff while the other player receives the larger payoff. Sample payoff structures for four and six round games are shown in Fig. 1.

In the centipede game, players are presented with a social dilemma. Consider the case shown in Fig. 1, where $a_1 = 0.4$, $a_0 = 0.1$, and $b = 2$. If you cooperate and your co-player also cooperates, you can earn more than if you defect and keep the larger amount now (1.6 vs. 0.4 if starting in the first round, for example). However, if you cooperate and the co-player defects, you earn less than if you defect now (0.2 vs. 0.4 if starting in the first round). Both players can potentially earn high payoffs if they cooperate repeatedly. Yet in each round, there is a temptation to

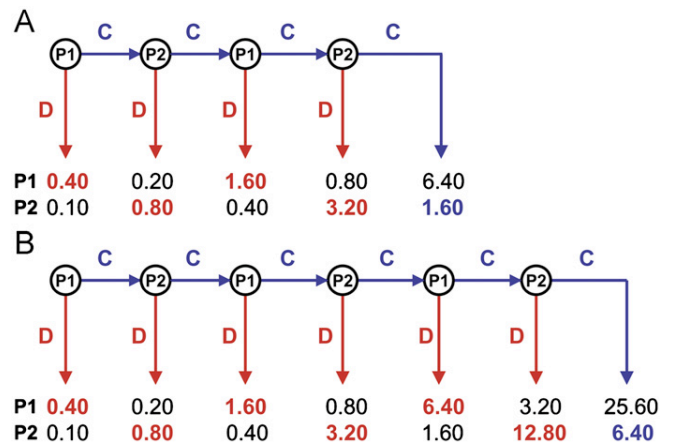


Fig. 1. Payoff structure for a four (A) and six (B) round centipede game, using payoff values $a_1 = 0.4$, $a_0 = 0.1$ and $b = 2$. The game alternates between Player 1 and Player 2. At each node one player chooses whether to cooperate or defect. If he defects, he gets the larger payoff (initially valued at a_1) and the other player gets the lower payoff (initially valued at a_0). If he cooperates, both payoffs are multiplied by b and play passes to the other player. If a player cooperates in the final round, the payoffs are multiplied by b and then the cooperator automatically receives the smaller payoff. The deciding player is shown inside the circle for each decision node. Defection is represented by a red down arrow, while cooperation is a blue arrow to the right. At the base of each downward arrow, the payoffs of both players are shown. The payoff of the player choosing to end the game at that node is shown in bold.

defect and ensure a larger payoff for yourself. Because the game ends immediately when one player defects, there is no possibility of retaliation; and because the game has a maximum length, backward induction unravels cooperation.

Numerous behavioral experiments have been conducted to see how people play the centipede game (Bornstein et al., 2004; Ho and Weigel, 2005; McKelvey and Palfrey, 1992; Nagel and Tang, 1988). Contrary to the logic of backwards induction, most players cooperate initially, and start to defect only as the end of the game approaches. One explanation of this behavior assumes that there is incomplete information about the other player (Fudenberg and Levine, 1997; Kreps and Wilson, 1982a, 1982b; Kreps et al., 1982; McKelvey and Palfrey, 1992). If your opponent might not be rational, or might be altruistic and care about your payoff as well as her own, then it can be rational to cooperate initially. However, it is never rational to cooperate in the final round, regardless of your beliefs about the other's strategy. Another explanation, ‘bounded rationality,’ asserts that subjects can only backward induct several steps due to cognitive limitations (Ellingsen and Östling, 2007; Neyman, 1985; Simon, 1972). It has also been shown that extrinsic sources of strategy variation (such as local mutation) can promote full cooperation in fixed length games, although not the intermediate levels of cooperation observed empirically among humans (McNamara et al., 2004).

Here we study the evolution of cooperation in the centipede game using finite-sized stochastic population dynamics (Imhof and Nowak, 2006; Nowak et al., 2004; Taylor et al., 2004; Traulsen et al., 2007). We show that natural selection can favor both full cooperation as well as partial cooperation, without assumptions about other-regarding preferences or cognitive limitations. Furthermore, our evolutionary model quantitatively reproduces the behavior of humans from two experiments.

2. Model

We consider a well-mixed population of size N playing a centipede game of maximum length L . The strategy of player i ,

Table 1

Payoff matrix for a centipede game with a large initial payoff of $a_1=0.4$, a small initial payoff of $a_0=0.1$, and a cooperative multiplier of $b=2$. Shown is the row player's payoff. In a four round game, the available strategies are $s=0$ (immediate defection) through $s=4$ (full cooperation); in a six round game all seven strategies shown are possible.

	$s=0$	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$
$s=0$	0.25	0.6	0.6	0.6	0.6	0.6	0.6
$s=1$	0.15	0.5	1.2	1.2	1.2	1.2	1.2
$s=2$	0.15	0.3	1	2.4	2.4	2.4	2.4
$s=3$	0.15	0.3	0.6	2	4.8	4.8	4.8
$s=4$	0.15	0.3	0.6	1.2	4	9.6	9.6
$s=5$	0.15	0.3	0.6	1.2	2.4	8	19.2
$s=6$	0.15	0.3	0.6	1.2	2.4	4.8	16

s_i is defined as the last round in which player i will cooperate ($s_i \leq L$). Thus immediate defection is $s=0$, full cooperation is $s=L$, and intermediate cooperation is $0 < s < L$. When two players interact, each is equally likely to be either Player 1 or 2. A player's expected payoff in a given interaction is the average of the payoffs received in both roles. We begin by assuming that players choose the same strategy in both roles; we then show that our results are robust to allowing players to choose different strategies in the role of Player 1 versus Player 2. As the population is well mixed, each player's average payoff π is the average of the expected payoffs when playing with each other member of the population. The resulting payoff matrix using $a_1=0.4$, $a_0=0.1$, and $b=2$ is shown in Table 1.

We model the change of strategies in the population using a probabilistic process in which higher payoff strategies are more likely to spread while lower payoff strategies are more likely to die out. If game payoffs represent reproductive success, this corresponds to the process of (haploid) genetic evolution. These dynamics can also describe a process of social learning (Aoki et al., 2011; Boyd and Richerson, 1988; Sigmund et al., 2010), in which people engage in payoff-based imitation. There are a wide range of contexts in which people may choose strategies by imitating successful others. Payoff-based imitation is a useful heuristic which can often allow people to find successful strategies much more quickly (and with lower cognitive costs) than deductive, prospective reasoning (Rendell et al., 2010). In our model, a strategy specifies a particular set of behaviors, and evolution occurs based on the material payoff earned by that behavior. In the context of humans, therefore, each strategy can be thought of as corresponding to a set of preferences that cause agents with strategy s_i to rationally choose to cooperate through period i , and agents then adopt each other's preferences with adoption being biased by the material payoff: you are more likely to adopt the norms and preferences of a person you see as being more successful.

We use the frequency dependent Moran process (Hauert et al., 2007; Moran, 1962; Nowak et al., 2004). A player's fitness is defined as $f=1-w+w\pi$, where w is the intensity of selection and π_j is the expected payoff of agent j . In each generation, one agent i is randomly chosen to change strategy (i.e., die). With probability $1-u$, agent i picks another agent j to imitate, proportional to fitness f . The selection intensity, w , determines how much the payoff contributes to the fitness of a player. In the limit $w \rightarrow 0$, we have 'weak selection': the game contributes virtually nothing to fitness and reproduction is almost entirely at random with respect to game strategy. Or put differently, learning is extremely error prone and learners pick teachers to imitate almost at random (with respect to game payoff); this could represent a situation in which many factors other than behavior in the game in question contribute to a person's success. As w increases towards 1, learning occurs with ever greater fidelity, such that

learners are increasingly likely to imitate players with higher game payoffs. With probability u , on the other hand, a 'mutation' occurs instead and agent i randomly picks a strategy between 0 and L . In the context of learning, mutation corresponds to either innovation/experimentation or confusion about the teacher's strategy.

We begin by considering the low mutation limit. When u is sufficiently small, there are rarely more than two strategies present in the population at once. Starting from a homogeneous population of players all of whom use the same strategy, a mutant strategy will either become extinct or go to fixation before another mutant arises. The frequency of each strategy in the population averaged over time (i.e., frequency in the steady state distribution) can be directly calculated from the fixation probabilities of each possible mutant arising in a population of each possible resident (Fudenberg and Imhof, 2006; Imhof et al., 2005; Nowak et al., 2004). We use the calculation method introduced by Antal et al. (2009a), which is reproduced in Appendix A. Later, we will also consider higher mutation rates using agent based simulations.

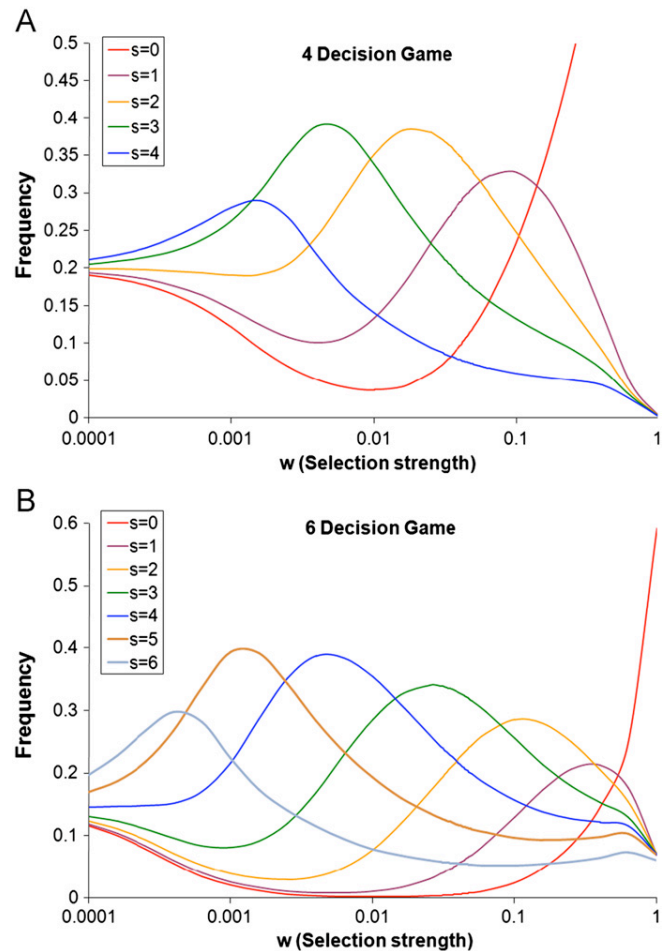


Fig. 2. Decreasing the selection strength favors increasingly cooperative strategies. The frequency of each strategy in the four round (A) and six round (B) games is shown for variable selection strength, using $a_1=0.4$, $a_0=0.1$, $b=2$ and $N=1000$. At strong selection $w=1$, immediate defection $s=0$ dominates the population. As w decreases, $s=1$ (cooperating once and then defecting) becomes the most common strategy, followed by $s=2$, and so on until finally full cooperation becomes most frequent. Nash equilibrium analysis finds that full cooperation is irrational under all circumstances. Nonetheless, full cooperation can be favored by evolution in finite populations if the intensity of selection is sufficiently weak.

3. Results

We begin by examining the four round centipede game. Using payoffs $a_1=0.4$, $a_0=0.1$, and $b=2$ and population size $N=1000$, we vary selection strength from weaker selection $w=0.0001$ to stronger selection $w=1$. We then ask which of the five possible strategies $s=[0,1,2,3,4]$ can be favored by natural selection. Sub-game perfect Nash equilibrium analysis predicts that all players should defect immediately ($s=0$). Our evolutionary analysis, however, finds that for certain values of w , each of the five strategies can be most common (Fig. 2(A)).

When selection is strong, $w=1$, the evolutionary model agrees with the sub-game perfect Nash equilibrium analysis: selection opposes any cooperation, and immediate defection, $s=0$, is the most frequent strategy. But as the intensity of selection w decreases, each increasingly cooperative strategy in turn becomes

most frequent. Finally, as we approach the weak selection limit, ($w \rightarrow 0$), full cooperation, $s=4$, is the most frequent strategy. The same pattern occurs in the six round game (Fig. 2(B)). This pattern is also not unique to the particular set of payoff values used in Fig. 2, as shown in Fig. 3. Thus, in a stochastic evolutionary framework with rare mutations, either immediate defection, delayed defection or full cooperation can be the favored outcomes depending on the intensity of selection. Reducing intensity of selection promotes the evolution of cooperation in fixed length games.

We now present an intuition underlying the success of cooperation in our model. When selection is strong, what matters for a strategy's success is its ability to resist invasion. If a single mutant is at a disadvantage in a given resident population, that mutant will almost certainly die out. As a result, the resident strategy will persist. A strategy which cooperates for i rounds is

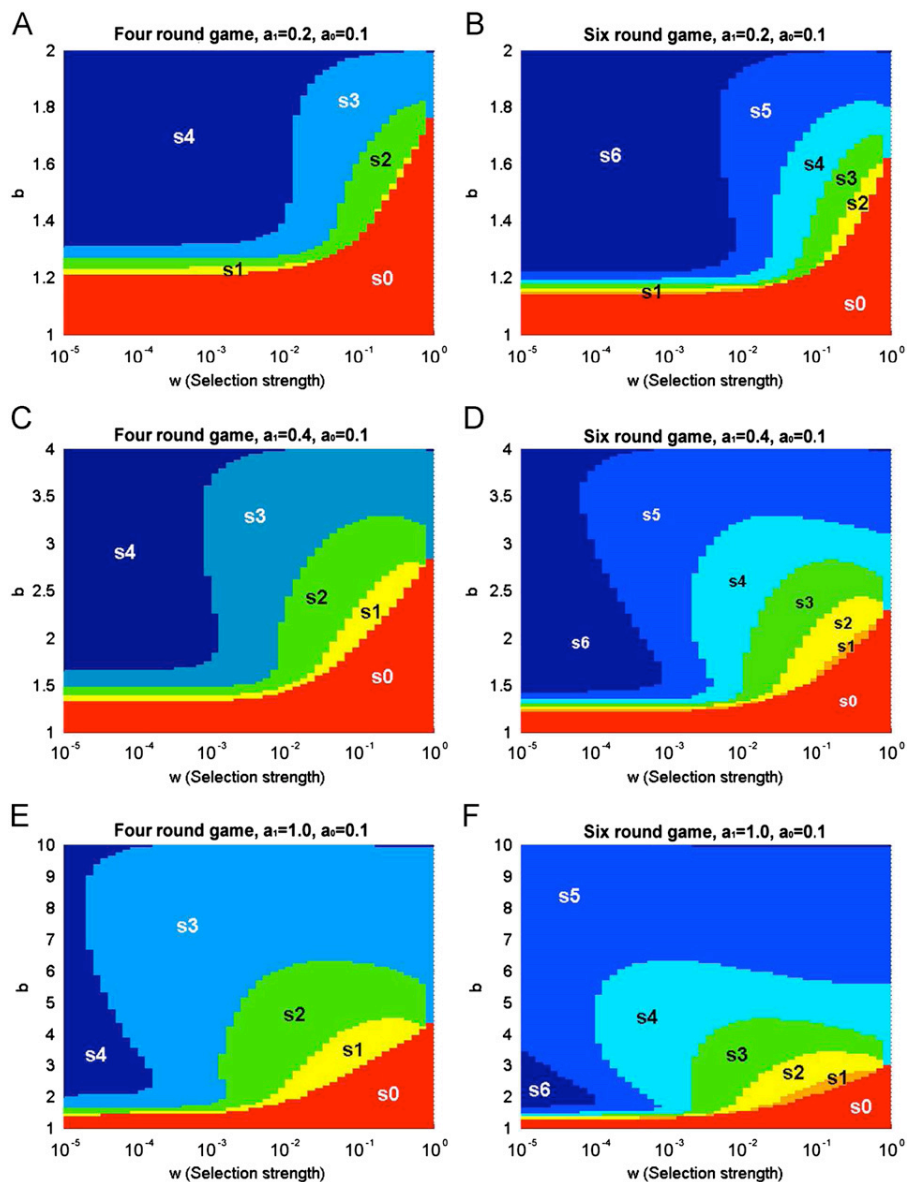


Fig. 3. As the selection strength w decreases from $w=1$, increasingly cooperative strategies are favored by natural selection over a range of payoff values. The smaller initial payoff is fixed at $a_0=0.1$. In each panel, the most common strategy is indicated over a range of b and w values, and the value of the larger initial payoff a_1 is varied across panels, taking on the values of $a_1=0.2$ (A), (B), $a_1=0.4$ (C), (D), or $a_1=1$ (E), (F). Panels in the left column show results for the four round centipede game (A), (C), (E), while panels in the right column show results for the size round game (B), (D), (F). All calculations use $N=1000$.

easily invaded by a strategy which cooperates for $i-1$ rounds, and the strategy which never cooperates cannot be invaded by any other strategies. Thus full defection is the only evolutionarily stable strategy, and is favored when selection is strong. When selection is weak, however, the dynamics are different. Even if a single mutant is at a disadvantage in a given resident population, that strategy will sometimes become more common through random drift. Thus successful strategies must not only resist invasion, but must also perform well in heterogeneous populations where both strategies have non-negligible frequencies. In the limit of weak selection, success in these pairwise competitions depends on risk-dominance (Antal et al., 2009a, 2009b). One strategy risk-dominates another by earning a higher payoff when both strategies are equally common, which is the case on average in the weak selection limit. As can be seen from the payoff matrix of the centipede game (Table 1) full cooperation can risk-dominate immediate defection. Even though full cooperation performs poorly when somewhat less cooperative strategies are present, it can perform well in pairwise play against immediate defection. When selection is weak, full cooperation can frequently invade immediate defection, given the right payoffs: random drift can help a full cooperator mutant overcome the initial disadvantage faced when arising in a population of immediate defectors. As selection becomes stronger, however, it becomes harder and harder for full cooperation to overcome this initial disadvantage. Full cooperation invading immediate defection is replaced by bistability between the two strategies in the deterministic limit, and we revert to the ESS/Nash case.

Next we compare the predictions of our evolutionary model to experimental data. McKelvey and Palfrey (1992) performed behavioral experiments using four and six round centipede games (payoffs as shown in Fig. 1(A) and (B)); we do not consider the experimental data from a four round game with higher payoffs run by McKelvey and Palfrey (1992), as there was only one session of data and the behavior was not significantly different from that observed in the four round game shown in Fig. 1(A)). For each game, they present the average number of cooperative moves (i.e., average game lengths) from three different

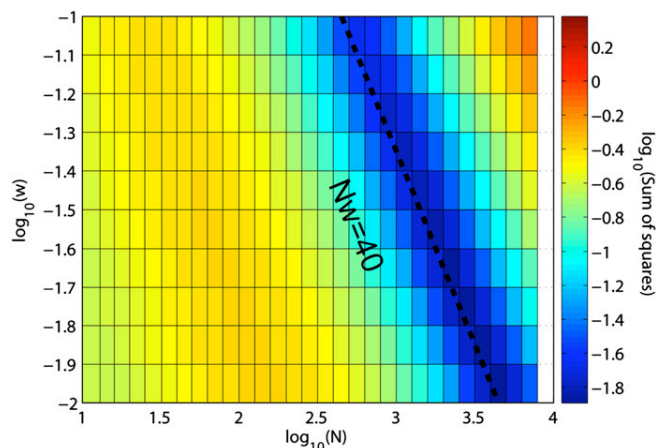


Fig. 4. We ask whether the model can reproduce the frequency of each game length observed in the behavioral experiments. To fit the model to experimental data, we minimize the sum of squared differences between the model predictions and the observed values. We calculate the sum of squared differences over both the four and six round games simultaneously, to produce a good fit to all data. The colormap indicates this sum of squares (a measure of badness of fit), such that lower z-axis values (shown in blue) indicate a better fit to the experimental data. As can be seen, the best fits (i.e., lowest z-axis values) lie along a line of constant $Nw=40$. Thus we effectively fit the data using only a single free parameter quantity Nw .

experimental sessions. They observed only few encounters that ended with immediate defection: 7% of four round games and <1% of six round games. Full cooperation was also rare: 5% of four round games and 1% of six round games. Most subjects showed an intermediate level of cooperation: defection was most likely to occur on round two or three of the four round game, and round three or four of the six round game.

To fit our evolutionary model to the experimental data, we can vary the population size N and the selection intensity w . However, the model's behavior, and the resulting quality of the fit, is largely constant along lines of constant Nw (see Fig. 4). The quantity Nw can be seen as a measure of selection versus random drift in the evolutionary process. Smaller values of Nw augment random drift. We seek the value of Nw which minimizes the sum of squared

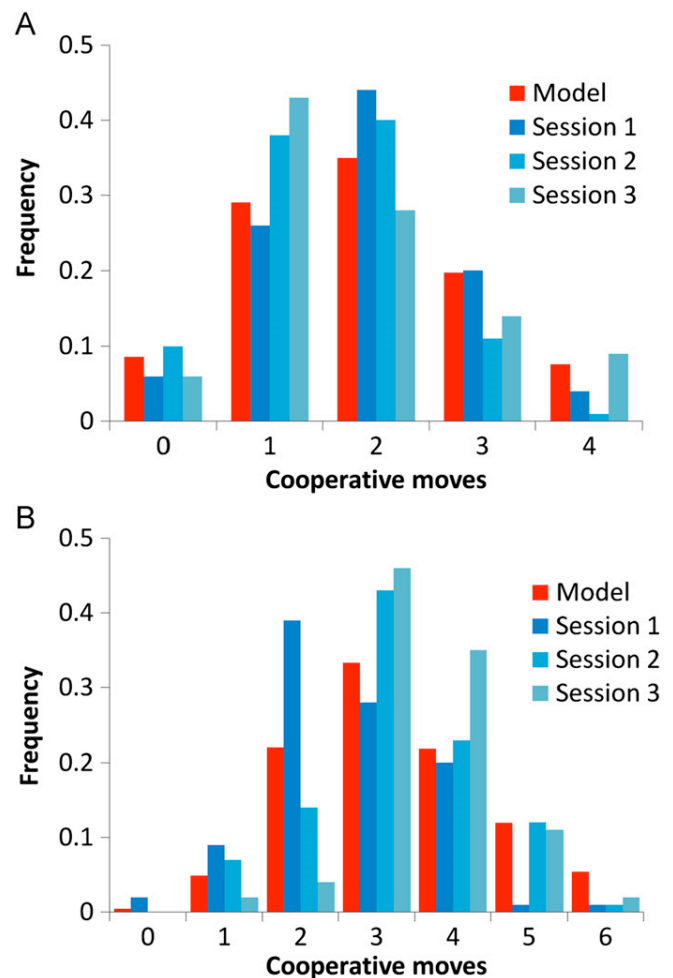


Fig. 5. Finite population size analysis quantitatively reproduces human behavior from two laboratory experiments. Model predictions (Red) are generated using payoffs $a_1=0.4$, $a_0=0.1$ and $b=2$ and model parameters $N=1000$ and $w=0.04$. The mean game length is shown for each of three separate experimental sessions (blue) of McKelvey and Palfrey, 1992 for the four round game (A) and six round game (B). For all but one of the 12 game length frequency data points, the model prediction lies within the range of the mean values from the three experimental sessions, and is not significantly different from the average frequency (Logistic regression with robust standard errors clustered on subjects and session, $p > 0.10$ for all). The one prediction which lies outside of the range of experimental data (full cooperation in the six round game, $p < 0.001$) exceeds the closest experimental session value by only 4%. Furthermore, this deviation may be explained by in-game learning across the 10 interactions played by each subject in the experiment: the model prediction is not significantly different from the data when controlling for interaction number ($p=0.534$).

differences between the model predictions and the experimental data. Predictions and data for both the four and six round games are compared simultaneously, to generate the best overall fit. We find that the evolutionary model quantitatively reproduces the behavioral data from both experiments for $Nw=40$ (Fig. 5). Our fit is also robust to including strategies which behave differently as Player 1 and Player 2 (see Appendix B). We note that in the experiments, subjects played in only one role (unlike in our model where players are equally like to be Player 1 or Player 2) and that the number of players per session was much smaller than the $N=1000$ population size we consider (although the results vary little when using a smaller population size of $N=100$, see Appendix C). As we are seeking to model the intuitions that subjects bring into the lab, rather than the in-game learning that occurs during the experiment, we chose these model features to better represent the environment outside the lab in which subjects' intuitions seem likely to have developed. See below for further discussion.

Our calculations thus far have assumed a vanishingly small mutation rate. This assumption simplifies the analysis, allowing direct numerical calculation of steady state frequencies. Here we show that the results are robust to considering higher mutation rates. To do so, agent based simulations are required. We simulate a Moran process with each agent having a strategy chosen from the finite set of possible strategies (s_0 thru s_4 for the four round game, s_0 thru s_6 for the six round game). The population is initialized randomly, the Moran process is simulated, and the frequency of each strategy is then averaged over 10^7 simulated generations. We find that our results are unchanged when the low mutation limit assumption is relaxed. Agent based simulations using higher mutation rates are also able to quantitatively reproduce the observed human behavior (Fig. 6). Interestingly, we observe a positive linear relationship between mutation rate u and the best-fit value of Nw (Fig. 7). We can tune the randomness of the evolutionary process by either increasing u or reducing Nw .

4. Discussion

We have explored the ability of stochastic evolutionary dynamics to predict human behavior in the fixed-length centipede game. In contrast to the predictions of classical economic game theory, subjects in experimental centipede games exhibit substantial levels of cooperation (McKelvey and Palfrey, 1992). We show that such behavior can be favored by natural selection when accounting for the stochastic nature of evolutionary dynamics. When the intensity of selection is strong and evolution is largely deterministic, selection favors the sub-game perfect Nash equilibrium of immediate defection. But at lower intensities of selection, both delayed defection and full cooperation can be favored. The strategy which does best in a perfectly precise world does not necessarily triumph in the presence of stochastic effects. (See also Smead (2008), which explores extinction leading to homogeneous populations of non-equilibrium strategies in finite populations.)

Our analysis focuses on weak selection as the source of randomness in the evolutionary process. Weak selection corresponds to uncertainty regarding payoff: sometimes people imitate the strategies of lower performing others (or in the genetic context, sometimes lower payoff strategies out-reproduce higher payoff ones). This stands in contrast to previous approaches for understanding cooperation in the centipede game. Both incomplete information in economics (Kreps and Wilson, 1982a, 1982b; Kreps et al., 1982) and mutation in biology (McNamara et al., 2004) involve uncertainty about strategy rather than payoff. In the context of incomplete information, you are unsure of the

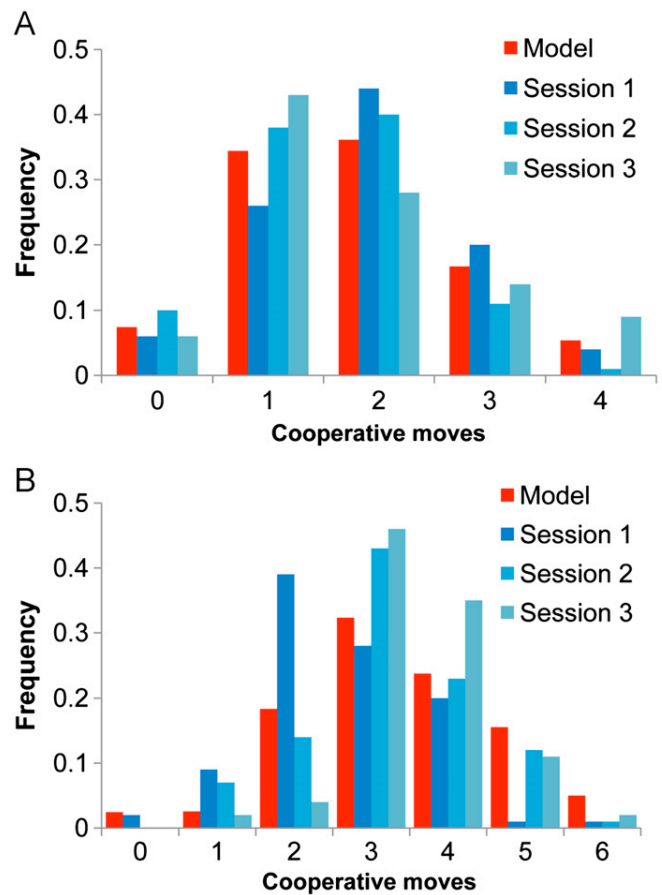


Fig. 6. Agent based simulations ($N=250$ due to computational constraints) show that our model can fit the experimental data outside of the low mutation limit. Shown is the model (red) and data (blue) for the four round (A) and six round (B) games at $u=0.01$, using the best-fit value of $Nw=100$. The quality of the fit is quite similar to that obtained using the low mutation limit.

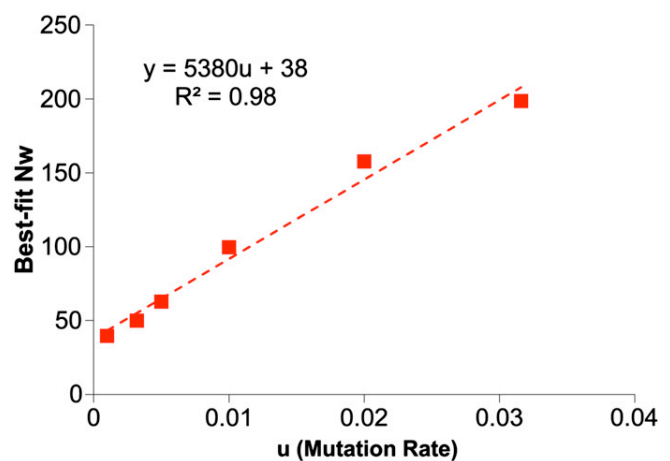


Fig. 7. A linear relationship exists between the mutation rate u and the best-fit value of Nw . Agent based simulations with $N=250$ are used to explore evolutionary dynamics with larger mutation rates. For various mutation rates, the selection strength resulting in the best fit to the experimental data is shown. The resulting relationship between u and Nw is well fit by a linear regression, $R^2=0.98$. Moreover, the regression predicts that when $u=0$, the best fit is given by $Nw=38$. This is very close to the value $Nw=40$ determined by the low mutation limit calculation. Such consistency is indicative of the robustness of the low mutation limit analysis.

preferences (and therefore the behavioral strategy) of your partner when reasoning about what strategy to adopt yourself. In the context of mutation, higher payoff agents are always more likely to reproduce, but sometimes their strategy is passed on incorrectly (or put differently, players sometimes get confused when trying to imitate a higher payoff other and adopt the wrong strategy).

Incomplete information can make any strategy rational except for cooperation in the final round (Kreps and Wilson, 1982a, 1982b; Kreps et al., 1982); and local mutation in a fixed length prisoner's dilemma game can result in a sharp transition from no cooperation to full cooperation when mutations are sufficiently common (McNamara et al., 2004). Here we show that weak selection in finite populations can favor both intermediate and extreme strategies, combining the strengths of previous approaches. Thus our framework can potentially help explain a variety of social dilemma behaviors in biological contexts involving finite time horizons; for example, interactions occurring within the course of a single breeding season such as competitive egg ejection in the Greater Ani, a communally nesting bird species (Riehl, 2011; Riehl and Jara, 2009). Furthermore, with only one free parameter, and assuming only self-regarding preferences, our evolutionary model is able to quantitatively reproduce human cooperative behavior from two experiments. Therefore these results also suggest the potential power of finite population analysis for understanding the origins of human behavior.

We are not seeking to model the evolution of peoples' strategies over the course of an experimental session in the lab (Goeree and Holt, 1999). Rather, we suggest that subjects enter the lab with intuitions which have been shaped by natural selection in the context of daily life experiences. In this view, laboratory experiments provide a snapshot of the end-product of an evolutionary process, which can be compared to steady state distributions from evolutionary models (Rand, 2011; Rand and Nowak, 2011; Rand et al., 2009a). Our model does more than restate the experimental observations in terms of a psychological profile or a utility function. Instead, it suggests a specific evolutionary mechanism underlying the results. The strategies observed in these behavioral experiments, as opposed to being altruistic or irrational, may in fact have been favored by the competitive dynamics of an evolutionary process. Further exploration of the possible link between stochastic evolutionary models and human behavior is an exciting topic for both theoretical and experimental research.

Acknowledgments

We thank Megan Frederickson, Drew Fudenberg, Michael Manapat, Erik Mohlin, Hisashi Ohtsuki, Richard Rand, Christina Reihl, Corina Tarnita and Matthijs van Veelen for helpful comments and feedback. Support from the John Templeton Foundation is gratefully acknowledged.

Appendix A. Low mutation limit calculation method

Here we reproduce the details of the calculation method used in our low mutation limit calculations, first introduced in Antal et al. (2009a). If viable mutants are very rare, the population spends almost all of its time in a homogeneous state. When a mutant arises, it either goes to fixation or dies out before the next mutant arises, returning the system to a homogeneous population. Let s_i be the frequency of strategy i , with a total of M strategies. We can then assemble a transition matrix between homogeneous states of the system. The transition probability

from state i to state j is the product of the probability of a mutant of type j arising, $1/M$, and the fixation probability of a single mutant j in a population of i players, $\rho_{j,i}$. The probability of staying in state i is thus $1 - \sum_{j \neq i} \rho_{j,i} / M$, where $\rho_{i,i} = 1$. This transition matrix can then be used to calculate the steady state frequency distribution s^n of strategies:

$$\begin{matrix}
 & \text{O} & \text{1} \\
 \text{O} & s_1^n & s_2^n \\
 \text{1} & s_2^n & s_M^n
 \end{matrix}
 \begin{matrix}
 & \text{O} & \text{1} \\
 \text{O} & \frac{\rho_{1,1}}{M-1} & \frac{\rho_{1,2}}{M-1} \\
 \text{1} & \frac{\rho_{2,1}}{M-1} & \frac{\rho_{2,2}}{M-1} \\
 & \vdots & \vdots \\
 \text{1} & \frac{\rho_{M,1}}{M-1} & \frac{\rho_{M,2}}{M-1}
 \end{matrix}
 \begin{matrix}
 & \text{O} & \text{1} \\
 \text{O} & s_1^n & s_2^n \\
 \text{1} & s_2^n & s_M^n
 \end{matrix}$$

As this is a stochastic matrix, the largest eigenvalue of this matrix will equal 1. The eigenvector corresponding to the eigenvalue 1 will give the steady state distribution of the stochastic process.

Using the Moran process, the fixation probability $\rho_{B,A}$ (the probability that a single A mutant introduced into a population of B-players will take over) can be calculated as follows. Let the pairwise payoffs of A and B be

$$\begin{matrix}
 & \text{A} & \text{B} \\
 \text{A} & a & b \\
 \text{B} & c & d
 \end{matrix}$$

where the value indicates the row player's payoff. In a population of x A-players and $N-x$ B-players, the fitness of an A-player f_x and B-player g_x are

$$\begin{matrix}
 f_x = \frac{1}{N} [w a x + (1-w) b d N - x b] \\
 g_x = \frac{1}{N} [w b c x + (1-w) d N - x c]
 \end{matrix}$$

where w is the intensity of selection.

The fixation probability of a single A-player in a population of B-players can then be calculated (Nowak, 2006) as follows:

$$\rho_{B,A} = \frac{1}{1 - \sum_{k \neq 1} \frac{g_k}{f_k}}$$

The calculations presented in the main text numerically evaluate Eq. (4) for each strategy pair, and then numerically solve Eq. (1) to determine the steady state frequency of each strategy.

Appendix B. Strategies which play differently as Player 1 and Player 2

In the main text, we assume that players use the same strategy regardless of their role. Here, we consider the case when players can choose different strategies as Player 1 and Player 2. This increases the size of the strategy set. In Table B.1, we present the payoff matrix when using this expanded strategy set. A strategy is now defined as $s_{i,j}$, where i is the number of times the player cooperates before defecting when Player 1, and j is the number of times the player cooperates before defecting when Player 2. Using this expanded strategy set, we can again nicely reproduce the experimental data, with only a small change in the optimal Nw , at $Nw \approx 45$. The best fit to the experimental data using the expanded strategy set is shown in Fig. B.1.

Table B.1

Payoff matrix allowing for different strategies as Player 1 and Player 2 for a centipede game with a large initial payoff of $a_1=0.4$, a small initial payoff of $a_0=0.1$, and a cooperative multiplier of $b=2$. Shown is the row player's payoff. For strategy s_{ij} , i is the number of times the player cooperates when Player 1, and j is the number of times the player cooperates when Player 2.

	$s_{0,0}$	$s_{0,1}$	$s_{0,2}$	$s_{0,3}$	$s_{1,0}$	$s_{1,1}$	$s_{1,2}$	$s_{1,3}$	$s_{2,0}$	$s_{2,1}$	$s_{2,2}$	$s_{2,3}$	$s_{3,0}$	$s_{3,1}$	$s_{3,2}$	$s_{3,3}$
$s_{0,0}$	0.25	0.25	0.25	0.25	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
$s_{0,1}$	0.25	0.25	0.25	0.25	0.4	0.4	0.4	0.4	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
$s_{0,2}$	0.25	0.25	0.25	0.25	0.4	0.4	0.4	0.4	1	1	1	1	6.6	6.6	6.6	6.6
$s_{0,3}$	0.25	0.25	0.25	0.25	0.4	0.4	0.4	0.4	1	1	1	1	3.4	3.4	3.4	3.4
$s_{1,0}$	0.15	0.85	0.85	0.85	0.5	1.2	1.2	1.2	0.5	1.2	1.2	1.2	0.5	1.2	1.2	1.2
$s_{1,1}$	0.15	0.85	0.85	0.85	0.3	1	1	1	1.7	2.4	2.4	2.4	1.7	2.4	2.4	2.4
$s_{1,2}$	0.15	0.85	0.85	0.85	0.3	1	1	1	0.9	1.6	1.6	1.6	6.5	7.2	7.2	7.2
$s_{1,3}$	0.15	0.85	0.85	0.85	0.3	1	1	1	0.9	1.6	1.6	1.6	3.3	4	4	4
$s_{2,0}$	0.15	0.45	3.25	3.25	0.5	0.8	3.6	3.6	0.5	0.8	3.6	3.6	0.5	0.8	3.6	3.6
$s_{2,1}$	0.15	0.45	3.25	3.25	0.3	0.6	3.4	3.4	1.7	2	4.8	4.8	1.7	2	4.8	4.8
$s_{2,2}$	0.15	0.45	3.25	3.25	0.3	0.6	3.4	3.4	0.9	1.2	4	4	6.5	6.8	9.6	9.6
$s_{2,3}$	0.15	0.45	3.25	3.25	0.3	0.6	3.4	3.4	0.9	1.2	4	4	3.3	3.6	6.4	6.4
$s_{3,0}$	0.15	0.45	1.65	12.85	0.5	0.8	2	13.2	0.5	0.8	2	13.2	0.5	0.8	2	13.2
$s_{3,1}$	0.15	0.45	1.65	12.85	0.3	0.6	1.8	13	1.7	2	3.2	14.4	1.7	2	3.2	14.4
$s_{3,2}$	0.15	0.45	1.65	12.85	0.3	0.6	1.8	13	0.9	1.2	2.4	13.6	6.5	6.8	8	19.2
$s_{3,3}$	0.15	0.45	1.65	12.85	0.3	0.6	1.8	13	0.9	1.2	2.4	13.6	3.3	3.6	4.8	16

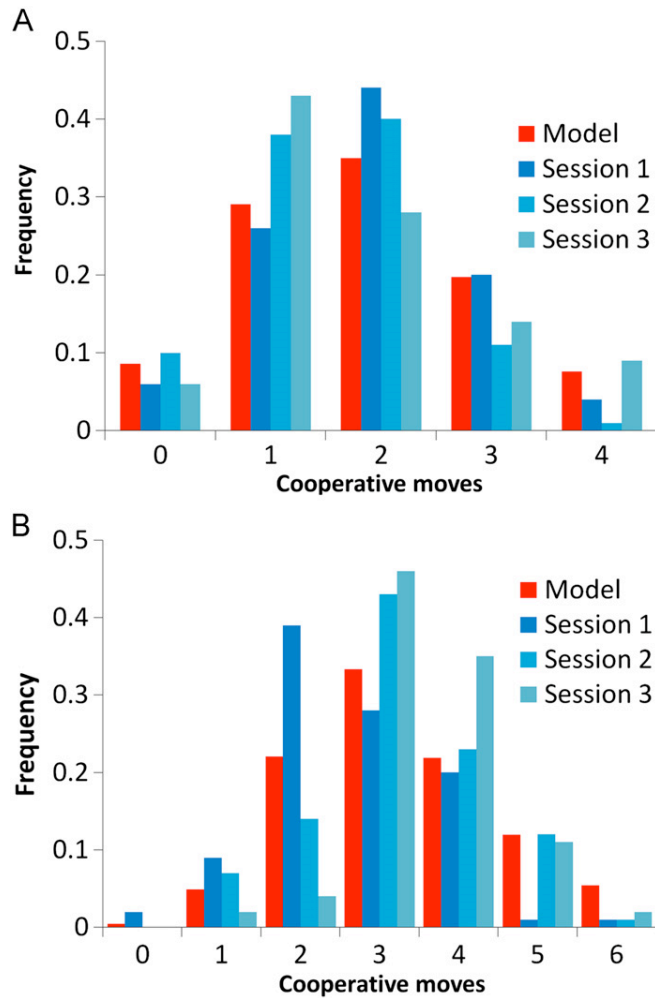


Fig. B.1. Best fit to the experimental data using an expanded strategy set. Here we include strategies which play differently as Player 1 and Player 2. For example, a strategy might defect on round 1 as Player 1, but cooperate throughout as Player 2. This increases the number of strategies in the four round game from 5 to 9, and in the six round game from 7 to 16. Interestingly, these additional strategies perform very poorly, and are present at < 1% frequency in the relevant Nw range. We find that the model using this full strategy set fits the data well for $Nw=45$.

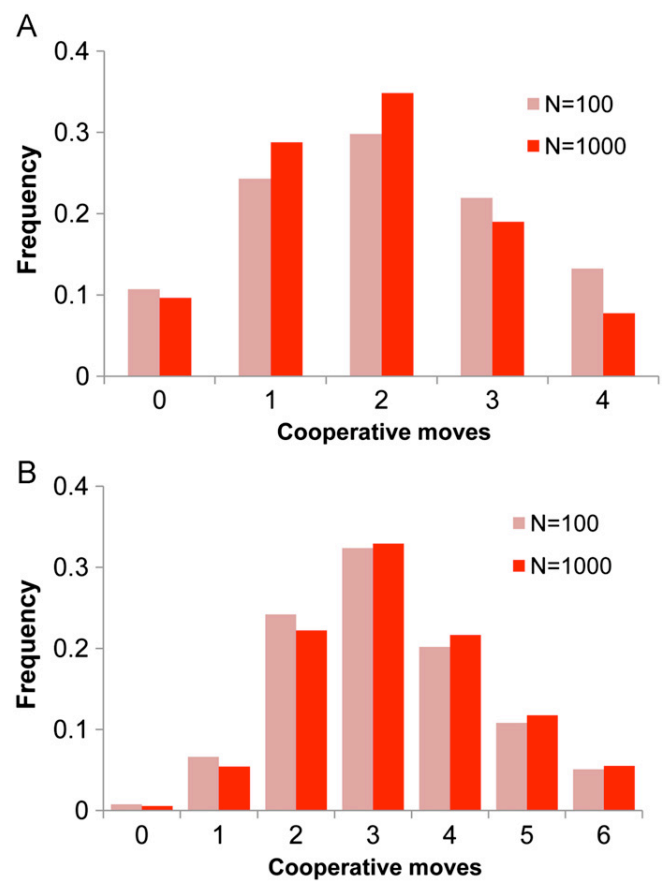


Fig. C.1. Strategy frequencies in the four (A) and six (B) round games using $a_1=0.4$, $a_0=0.1$ and $b=2$, holding constant $Nw=40$ and comparing $N=100$ (with $w=0.4$) and $N=1000$ (with $w=0.04$).

Appendix C. Fitted outcomes for four and six round games using $N=100$

In the main text Fig. 5, we fit the experimental data using the population size $N=1000$ and selection strength $w=0.04$. Here we show, in Fig. C.1, that the result is qualitatively similar using the

smaller population size of $N=100$ (and correspondingly larger selection strength $w=0.4$).

References

- Antal, T., Nowak, M.A., Traulsen, A., 2009a. Strategy abundance in 2×2 games for arbitrary mutation rates. *J. Theor. Biol.* 257, 340–344.
- Antal, T., Traulsen, A., Ohtsuki, H., Tarnita, C.E., Nowak, M.A., 2009b. Mutation–selection equilibrium in games with multiple strategies. *J. Theor. Biol.* 258, 614–622.
- Aoki, K., Lehmann, L., Feldman, M.W., 2011. Rates of cultural change and patterns of cultural accumulation in stochastic models of social transmission. *Theor. Populat. Biol.* 79, 192–202. doi:10.1016/j.tpb.2011.02.001.
- Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- Basu, K., 1990. On the non-existence of a rationality definition for extensive games. *Int. J. Game Theory* 19, 33–44.
- Binmore, K., 1987. Modeling rational players. *Econ. Philos.* 3, 179–214.
- Bolton, G.E., Katok, E., Ockenfels, A., 2005. Cooperation among strangers with limited information about reputation. *J. Publ. Econ.* 89, 1457–1468. doi:10.1016/j.jpubeco.2004.03.008.
- Bornstein, G., Kugler, T., Ziegelmeyer, A., 2004. Individual and group decisions in the centipede game: are groups more rational players? *J. Exp. Soc. Psychol.* 40, 599–605.
- Boyd, R., Richerson, P.J., 1988. An evolutionary model of social learning: the effects of spatial and temporal variation. In: Zentall, T.R., Galef, B.G.J. (Eds.), *Social Learning: Psychological and Biological Perspectives*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, pp. 29–48.
- Colman, A.M., 1995. *Game Theory and its Applications in the Social and Biological Sciences*. Routledge, New York.
- Cressman, R., 2003. *Evolutionary Dynamics and Extensive Form Games*. The MIT Press, Cambridge, MA.
- Cressman, R., Schlag, K.H., 1998. The dynamic (in)stability of backwards induction. *J. Econ. Theory* 83, 260–285.
- Dal Bó, P., 2005. Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *Am. Econ. Rev.* 95, 1591–1604.
- Dal Bó, P., Fréchette, G.R., 2011. The evolution of cooperation in infinitely repeated games: experimental evidence. *Am. Econ. Rev.* 101, 411–429.
- Dasgupta, P., 2009. Trust and cooperation among economic agents. *Philos. Trans. R. Soc. B: Biol. Sci.* 364, 3301–3309. doi:10.1098/rstb.2009.0123.
- Dreber, A., Rand, D.G., Fudenberg, D., Nowak, M.A., 2008. Winners don't punish. *Nature* 452, 348–351.
- Ellingsen, T., Östling, R., 2007. Communication and coordination: the case of boundedly rational players. *Mimeo*.
- Enquist, M., Leimar, O., 1993. The evolution of cooperation in mobile organisms. *Anim. Behav.* 45, 747–757. doi:10.1006/ambe.1993.1089.
- Frank, R.H., Gilovich, T., Regan, D.T., 1993. The evolution of one-shot cooperation: an experiment. *Ethol. Sociobiol.* 14, 247–256.
- Frean, M.R., 1994. The prisoner's dilemma without synchrony. *Proceedings: Biological Sciences* 257, pp. 75–79.
- Fudenberg, D., Maskin, E., 1986. The Folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54, 533–554.
- Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. *Am. Econ. Rev.* 80, 274–279.
- Fudenberg, D., Levine, D.K., 1997. Measuring players' losses in experimental games. *Q. J. Econ.* 112, 507–536.
- Fudenberg, D., Imhof, L.A., 2006. Imitation processes with small mutations. *J. Econ. Theory* 131, 251–262.
- Fudenberg, D., Rand, D.G., Dreber, A., in press. Slow to anger and fast to forgive: cooperation in an uncertain world. *Am. Econ. Rev.*
- Giovanni, P., 2000. Cycles of learning in the centipede game. *Games Econ. Behav.* 30, 115–141. doi:10.1006/game.1998.0707.
- Goeree, J.K., Holt, C.A., 1999. Stochastic game theory: For playing games, not just for doing theory. *Proceedings of the National Academy of Sciences* 96, pp. 10564–10567. doi:10.1073/pnas.96.19.10564.
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., Sigmund, K., 2007. Via freedom to coercion: the emergence of costly punishment. *Science* 316, 1905–1907.
- Helbing, D., Yu, W., 2009. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences* 106, pp. 3680–3685. doi:10.1073/pnas.0811503106.
- Ho, T., Weigel, K., 2005. Trust building among strangers. *Manage. Sci.*, 51.
- Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.
- Imhof, L.A., Nowak, M.A., 2006. Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* 52, 667–681.
- Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. *Proc. Nat. Acad. Sci. U.S.A.* 102, 10797–10800.
- Kreps, D.M., Wilson, R., 1982a. Sequential equilibria. *Econom.: J. Econom. Soc.*, 863–894.
- Kreps, D.M., Wilson, R., 1982b. Reputation and imperfect information. *J. Econ. Theory* 27, 253–279.
- Kreps, D.M., Milgrom, P., Roberts, J., 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *J. Econ. Theory* 27, 245–252.
- Ladret, V., Lessard, S., 2008. Evolutionary game dynamics in a finite asymmetric two-deme population and emergence of cooperation. *J. Theor. Biol.* 255, 137–151. doi:10.1016/j.jtbi.2008.07.025.
- Lessard, S., Ladret, V., 2007. The probability of fixation of a single mutant in an exchangeable selection model. *J. Math. Biol.* 54, 721–744. doi:10.1007/s00285-007-0069-7.
- Lindgren, K., 1991. Evolutionary phenomena in simple dynamics. *Artif. Life II* 10, 295–312.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK.
- McKelvey, R.D., Palfrey, T.R., 1992. An experimental study of the centipede game. *Econometrica* 60, 803–836.
- McNamara, J.M., Gasson, C.E., Houston, A.I., 1999. Incorporating rules for responding into evolutionary games. *Nature* 401, 368–371.
- McNamara, J.M., Barta, Z., Houston, A.I., 2004. Variation in behaviour promotes cooperation in the prisoner's dilemma game. *Nature* 428, 745–748.
- Milinski, M., Wedekind, C., 1998. Working memory constrains human cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences of the United States of America* 95, pp. 13755–13758.
- Milinski, M., Semmann, D., Krambeck, H., 2002a. Donors to charity gain in both indirect reciprocity and political reputation. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269, pp. 881–883. doi:10.1098/rspb.2002.1964.
- Milinski, M., Semmann, D., Krambeck, H.J., 2002b. Reputation helps solve the 'tragedy of the commons'. *Nature* 415, 424–426.
- Milinski, M., Semmann, D., Bakker, T.C.M., Krambeck, H.-J., 2001. Cooperation through indirect reciprocity: image scoring or standing strategy? *Proc. R. Soc. London, Ser. B: Biol. Sci.* 268, 2495–2501. doi:10.1098/rspb.2001.1809.
- Moran, P.A.P., 1962. *The Statistical Processes of Evolutionary Theory*. Clarendon, Oxford, UK.
- Nagel, R., Tang, F.F., 1988. Experimental results on the centipede game in normal form: an investigation on learning. *J. Math. Psychol.* 42, 356–384.
- Neyman, A., 1985. Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Econ. Lett.* 19, 227–229.
- Nowak, M., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature* 364, 56–58.
- Nowak, M.A., 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press of Harvard University Press, Cambridge, MA and London, England.
- Nowak, M.A., Sigmund, K., 1992. Tit-for-tat in heterogeneous populations. *Nature* 355, 250–253.
- Nowak, M.A., Sigmund, K., 1994. The alternating prisoner's dilemma. *J. Theor. Biol.* 168, 219–226. doi:10.1006/jtbi.1994.1101.
- Nowak, M.A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577. doi:10.1038/31225.
- Nowak, M.A., Sigmund, K., 2004. Evolutionary dynamics of biological games. *Science* 303, 793–799.
- Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. *Nature* 437, 1291–1298.
- Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability infinite populations. *Nature* 428, 646–650.
- Ohtsuki, H., Iwasa, Y., 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* 239, 435–444.
- Pacheco, J.M., Santos, F.C., Chalub, F.A.C.C., 2006. Stern-judging: a simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comput. Biol.* 2, e178.
- Pettit, P., Sugden, R., 1989. The backward induction paradox. *J. Philos.* 86, 169–182.
- Rand, D.G., 2011. The promise of mechanical turk: how online labor markets can help theorists run behavioral experiments. *J. Theor. Biol.*, 10.1016/j.jtbi.2011.03.004.
- Rand, D.G., Nowak, M.A., 2011. The evolution of antisocial punishment in optional public goods games. *Nat. Commun.* 2, 434. doi:10.1038/ncomms1442_S1.html.
- Rand, D.G., Ohtsuki, H., Nowak, M.A., 2009a. Direct reciprocity with costly punishment: generous tit-for-tat prevails. *J. Theor. Biol.* 256, 45–57.
- Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A., 2009b. Positive interactions promote public cooperation. *Science* 325, 1272–1275.
- Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Fogarty, L., Ghirlanda, S., Lillcrap, T., Laland, K.N., 2010. Why copy others? Insights from the social learning strategies tournament. *Science* 328, 208–213. doi:10.1126/science.1184719.
- Riehl, C., 2011. Living with strangers: direct benefits favour non-kin cooperation in a communally nesting bird. *Proc. R. Soc. B: Biol. Sci.* 278, 1728–1735. doi:10.1098/rspb.2010.1752.
- Riehl, C., Jara, L., 2009. Natural history and reproductive biology of the communally breeding Greater Ani (*Crotophaga major*) at Gatún Lake, Panama. *Wilson J. Ornithol.* 121, 679–687. doi:10.1676/09-017.1.
- Rockenbach, B., Milinski, M., 2006. The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444, 718–723.
- Rosenthal, R., 1982. Games of perfect information, predatory pricing, and the chain store paradox. *J. Econ. Theory* 25, 92–100.
- Selten, R., 1978. The chain-store paradox. *Theory Decis.* 9, 127–159.
- Semmann, D., Krambeck, H.-J., Milinski, M., 2003. Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature* 425, 390–393.
- Sigmund, K., 2010. *The Calculus of Selfishness*. Princeton University Press, Princeton.

- Sigmund, K., De Silva, H., Traulsen, A., Hauert, C., 2010. Social learning promotes institutions for governing the commons. *Nature* 466, 861–863.
- Simon, H.A., 1972. In: Radner, C.B.M. a.R. (Ed.), *Theories of Bounded Rationality. Decisions and Organisation*, North Holland, Amsterdam.
- Smead, R., 2008. The evolution of cooperation in the centipede game with finite populations. *Philosophy of Science* 75, 157–177.
- Taylor, C., Fudenberg, D., Sasaki, A., Nowak, M.A., 2004. Evolutionary game dynamics in finite populations. *Bull.Math. Biol.* 66, 1621–1644.
- Traulsen, A., Pacheco, J.M., Nowak, M.A., 2007. Pairwise comparison and selection temperature in evolutionary game dynamics. *J. Theor. Biol.* 246, 522–529, 10.1016/j.jtbi.2007.01.002.
- Trivers, R., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Wedekind, C., Milinski, M., 2000. Cooperation through image scoring in humans. *Science* 288, 850–852. doi:10.1126/science.288.5467.850.