Direct reciprocity in structured populations

Matthijs van Veelen[a,b,1,2], **Julián García**[c,1], **David G. Rand**[a,d], and **Martin A. Nowak**[a,e,f]

[a]Program for Evolutionary Dynamics, [d]Department of Psychology, [e]Department of Mathematics, and [f]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; [b]Center for Research in Experimental Economics and Political Decision Making, University of Amsterdam, 1018 WB Amsterdam, The Netherlands; and [c]Research Group for Evolutionary Theory, Max Planck Institute for Evolutionary Biology, 23406 Plön, Germany

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved May 3, 2012 (received for review April 20, 2012)

Reciprocity and repeated games have been at the center of attention when studying the evolution of human cooperation. Direct reciprocity is considered to be a powerful mechanism for the evolution of cooperation, and it is generally assumed that it can lead to high levels of cooperation. Here we explore an open-ended, infinite strategy space, where every strategy that can be encoded by a finite state automaton is a possible mutant. Surprisingly, we find that direct reciprocity alone does not lead to high levels of cooperation. Instead we observe perpetual oscillations between cooperation and defection, with defection being substantially more frequent than cooperation. The reason for this is that "indirect invasions" remove equilibrium strategies: every strategy has neutral mutants, which in turn can be invaded by other strategies. However, reciprocity is not the only way to promote cooperation. Another mechanism for the evolution of cooperation, which has received as much attention, is assortment because of population structure. Here we develop a theory that allows us to study the synergistic interaction between direct reciprocity and assortment. This framework is particularly well suited for understanding human interactions, which are typically repeated and occur in relatively fluid but not unstructured populations. We show that if repeated games are combined with only a small amount of assortment, then natural selection favors the behavior typically observed among humans: high levels of cooperation implemented using conditional strategies.

repeated prisoner's dilemma | game theory

The problem of cooperation in its simplest and most challenging form is captured by the Prisoners' Dilemma. Two people can choose between cooperation and defection. If both cooperate, they get more than if both defect, but if one defects and the other cooperates, the defector gets the highest payoff and the cooperator gets the lowest. In the one-shot Prisoners' Dilemma, it is in each person's interest to defect, even though both would be better off had they cooperated. This game illustrates the tension between private and common interest.

However, people often cooperate in social dilemmas. Explaining this apparent paradox has been a major focus of research across fields for decades. Two important explanations for the evolution of cooperation that have emerged are reciprocity (1–19) and population structure (20–32). If individuals find themselves in a repeated Prisoner's Dilemma—rather than a one-shot version—then there are Nash equilibria where both players cooperate under the threat of retaliation in future rounds (1–19). The existence of such equilibria is a cornerstone result in economics (1–3), and the evolution of cooperation in repeated games is of shared interest for biology (4–10), economics (11–14), psychology (15), and sociology (16), with applications that range from antitrust laws (17) to sticklebacks (18), although it has been argued that firm empirical support in nonhuman animal societies is rare (19).

Population structure is equally important. If individuals are more likely to interact with others playing the same strategy, then cooperation can evolve even in one-shot Prisoner's Dilemmas, because then cooperators not only give, but also receive more cooperation than defectors (20–32). There are a host of different population structures and update rules that can cause the necessary assortment (28, 31). Whether thought of in terms of kin selection (20, 25, 26), group selection (24, 27, 32), both (29), or neither (30, 31), population structure can allow for the evolution of cooperative behavior that would not evolve in a well-mixed population. Assortment can, but does not have to be genetic, as for example in coevolutionary models based on cultural group selection (32).

In this article, we consider the interaction of these two mechanisms: direct reciprocity and population structure. We begin by re-examining the ability of direct reciprocity to promote cooperation in unstructured populations. Previous studies tend to consider strategies that only condition on the previous period (8–10) or use static equilibrium concepts that focus on infinitely many repetitions (11, 12). Although useful for analytical tractability, both of these approaches could potentially bias the results. Thus, we explore evolutionary dynamics that allow for an open-ended, infinite strategy space, and look at games where subsequent repetitions occur with a fixed probability δ.

To do so, we perform computer simulations where strategies are implemented using finite state automata (see Fig. 1 for examples), and compliment these simulations with analytical results, which are completely general and apply to all possible deterministic strategies. Our simulations contain a mutation procedure that guarantees that every finite state automaton can be reached from every other finite state automaton through a sequence of mutations. Thus, every strategy that can be encoded by a finite state automaton is a possible mutant. The mutants that emerge at a given time depend on the current state of the population: close-by mutants, requiring only one or two mutations, are more likely to arise than far away mutants, requiring many mutations.

Our computer program can, in principle, explore the whole space of deterministic strategies encoded by finite state automata. Fig. 1 shows that the population regularly transitions in and out of cooperation and gives a sample of the equilibrium strategies, with different degrees of cooperation, that surface temporarily. The variety of equilibria shows that evolution does explore a host of different possibilities for equilibrium behavior, and that it is as creative in constructing equilibria as it is in undermining them.

Based on previous analyses of repeated games, one might expect evolution to lead to high levels of cooperation for relatively small b/c ratios in our simulations, provided the continuation probability δ is reasonably large. However, this is not what we find. To understand why, we have to consider indirect invasions (33).

In a well-mixed population, the strategy Tit-for-Tat (TFT) can easily resist a direct invasion of ALLD (always choosing to defect, regardless) because ALLD performs badly in a population of TFT players, provided that the continuation probability δ is large enough. The strategy ALLC (unconditional cooperation), however,

EVOLUTION

SOCIAL SCIENCES

Author contributions: M.v.V. and J.G. designed research; M.v.V. and J.G. performed research; M.v.V., J.G., D.G.R., and M.A.N. analyzed data; and M.v.V., J.G., D.G.R., and M.A.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1] M.v.V. and J.G. contributed equally to this work.

[2] To whom correspondence should be addressed. E-mail: C.M.vanVeelen@uva.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1206694109/-/DCSupplemental.

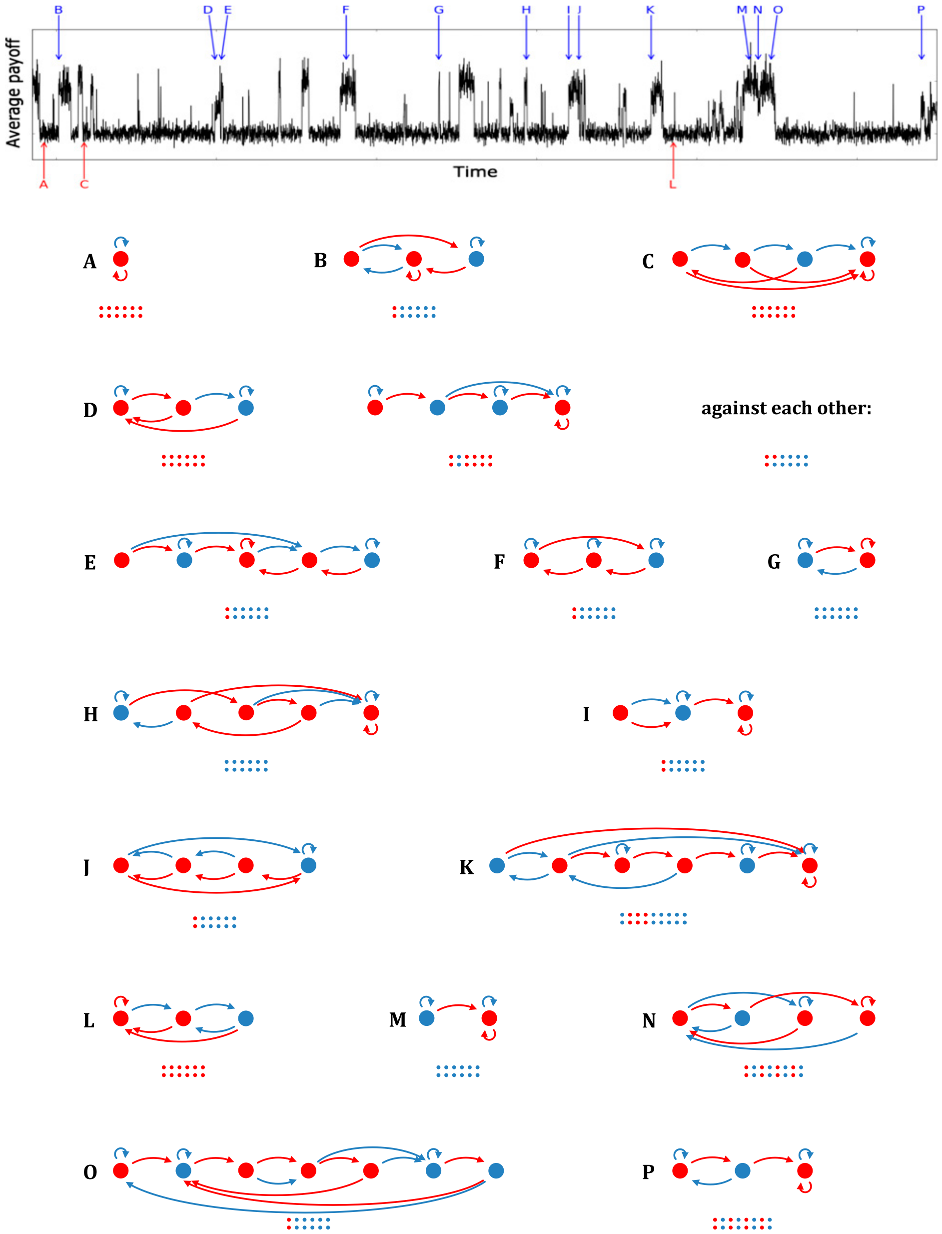


Fig. 1. Examples of equilibrium strategies observed during a simulation run. The top part of the figure depicts the average payoff during a part of a simulation run with continuation probability 0.85. The payoffs in the stage game are 2 for both players if they both cooperate, 1 for both if they both defect, and 3 for the defector and 0 for the cooperator if one defects and the other cooperates. These payoffs imply a benefit-to-cost ratio of $b/c = 2$. Because the game lengths are stochastic, there is variation in average payoff, even when the population makeup is constant. The different payoff plateaus indicate the population visiting different equilibria with different levels of cooperation and hence different expected average payoffs. Examples of equilibrium strategies, indicated by letters A through P, are also shown in the following way. The circles are the states and their colors indicate what this strategy plays when in that state; a blue circle means that the strategy will cooperate and a red one means that it will defect. The arrows reflect to which state this strategy goes, depending on the action played by its opponent; the blue arrows indicate where it goes if its opponent cooperates, the red ones where it goes if its opponent defects. The strategies all start in the leftmost state. The small colored dots indicate what the first few moves are if the strategy plays against itself, and in a mixture (case D) also what the two strategies play when they meet each other. The strategies vary widely in the ways in which they are reciprocal, the extent to which they are forgiving, the presence or absence of handshakes they use before they start cooperating, and the level of cooperation. The strategies shown here are discussed in greater detail in the *SI Appendix*.

can serve as a springboard for ALLD and thereby disrupt cooperation. ALLC is a neutral mutant of TFT: when meeting themselves and each other, both strategies always cooperate, and hence they earn identical payoffs. Thus, ALLC can become more abundant in a population of TFT players through neutral drift. When this process occurs, ALLD can then invade by exploiting ALLC. Unconditional cooperation is therefore cooperation's worst enemy (10).

Indirect invasions do not only destroy cooperation; they can also establish cooperation (for examples see the *SI Appendix*). If the population size is not too small, those indirect invasions in and out of cooperation come to dominate the dynamics in the population (34).

We can show that no strategy is ever robust against indirect invasions; there are always indirect paths out of equilibrium (34). Our simulations also suggest that in a well-mixed population, paths out of cooperation are more likely than paths into cooperation. Even for relatively high continuation probabilities, evolution consistently leads only to moderate levels of cooperation when averaged over time (unless the benefit-to-cost ratio of cooperation is very high) (see also *SI Appendix*, Fig. S5).

How, then, can we achieve high levels of cooperation? We find that adding a small amount of population structure increases the average level of cooperation substantially if games are repeated. The interaction between repetition and population structure is of primary importance, especially for humans, who tend to play games with many repetitions and who live in fluid, but not totally unstructured populations (35–40).

To explore the effect of introducing population structure, we no longer have individuals meet entirely at random. Instead, a player with strategy S is matched with an opponent that also uses strategy S with probability $\alpha + (1 - \alpha)x_S$, where x_S is the frequency of strategy S in the population, and α is a parameter that can vary continuously between 0 and 1. With this matching process, the assortment parameter α is the probability for a rare mutant to meet a player that has the same strategy (21, 29, 41, 42).

If $\delta = 0$, we are back in the one-shot version of the game. If $\alpha = 0$, we study evolution in the unstructured (well-mixed) population. Thus, the settings where only one of the two mechanisms is present are included as special cases in our framework.

Fig. 2 shows how assortment and repetition together affect the average level of cooperation in our simulations. If there is no repetition, $\delta = 0$, we find a sharp threshold for the evolution of cooperation. If there is some repetition, $\delta > 0$, we find a more gradual rise in the average level of cooperation as assortment α increases. In the lower right region of Fig. 2, where repetition is high and assortment is low (but nonzero), we find behavior similar to what is observed in humans: the average level of cooperation is high, and the strategies are based on conditional cooperation (14, 43–46). In contrast, when assortment is high and repetition is low we observe the evolution of ALLC, which is rare among humans.

To gain a deeper understanding of these simulation results, we now turn to analytical calculations. For the stage game of the repeated game we consider the following payoff matrix.

	Cooperate	Defect
Cooperate	$b - c$	$-c$
Defect	b	0

Our analytical results are derived without restricting the strategy space and by considering both direct and indirect invasions (see the *SI Appendix* for calculation details). The analytical results are therefore even more general than the simulation results. The simulations allow for all strategies that can be represented by finite automata (a large, countably infinite set of strategies where the mutation procedure specifies which mutations are more or less likely to occur). The analytical treatment, on the other hand, considers and allows for all possible strategies. Thus, the analytical results are completely general and independent of any assumptions about specific mutation procedures.

We find that the parameter space, which is given by the unit square spanned by α and δ, can be divided into five main regions (Fig. 2*B*). We refer to the lower left corner, containing $\delta = 0, \alpha = 0$, as region 1, and number the remaining regions 2–5, proceeding counter clockwise around the pivot $(\delta, \alpha) = (0, c/b)$. To discuss these regions we introduce the cooperativity of a strategy, defined as the average frequency of cooperation a strategy achieves when playing against another individual using the same strategy. Cooperativity is a number between 0 and 1: fully defecting strategies, such as ALLD, have cooperativity 0, and fully cooperative strategies, such as ALLC or TFT (without noise), have cooperativity 1. In Fig. 3, we show representative simulation outcomes for parameter values in the center of each region.

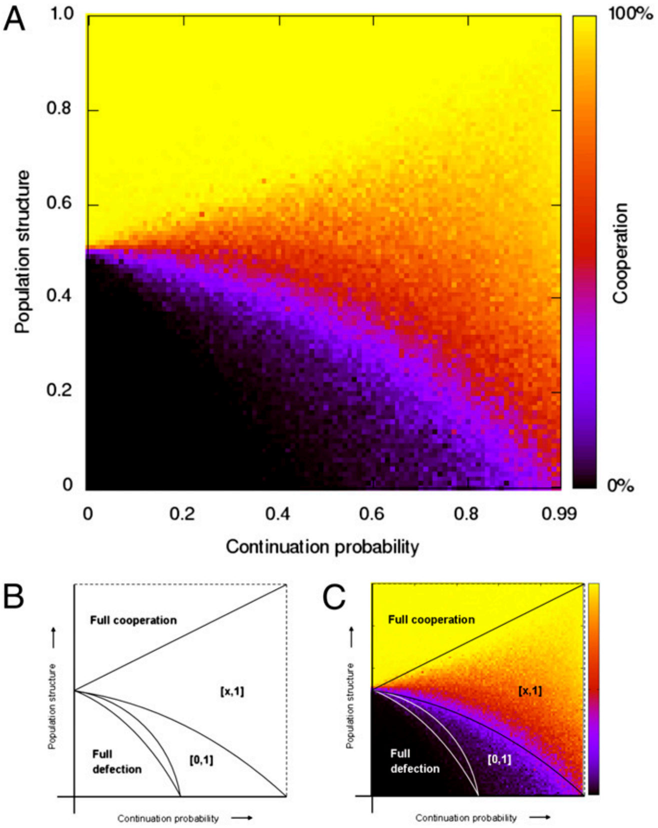


Fig. 2. Simulation results and theoretical prediction with repetition as well as assortment. (*A*) Every pixel represents a run of 500,000 generations, where every individual in every generation plays a repeated game once. The population size is 200 and the benefit-to-cost ratio is $b/c = 2$. The continuation probability δ (horizontal axis) indicates the probability with which a subsequent repetition of the stage game between the two players occurs. Therefore, a high continuation probability means that in expectation the game is repeated a large number of times and a continuation probability of 0 implies that the game is played exactly once. On the vertical axis we have a parameter α for the assortment introduced by population structure, which equals the probability with which a rare mutant meets another individual playing the same strategy, and that can also be interpreted as relatedness (21, 29, 41, 42). This parameter being 0 would reflect random matching. If it is 1, then every individual always interacts with another individual playing the same strategy. Both parameters—continuation probability δ and assortment α—are varied in steps of 0.01, which makes 10,100 runs in total. (*B* and *C*) A theoretical analysis with an unrestricted strategy space explains what we find in the simulations. This analysis divides the parameter space into five regions, as described in the main text (see the *SI Appendix* for a detailed analysis and a further subdivision). The border between regions 3 and 4 is an especially important phase transition, because above that line, fully defecting strategies no longer are equilibria. In the lower-right corner, where continuation probability is close to 1, adding only a little bit of population structure moves us across that border.

In regions 1 and 2, where both α and δ are less than c/b, all equilibrium strategies have cooperativity 0. In region 1, ALLD can directly invade every nonequilibrium strategy. In region 2 ALLD no longer directly invades all nonequilibrium strategies, but for every strategy with cooperativity larger than 0, there is at least one strategy that can directly invade. The fact that these direct invaders exist, however, does not prevent the population from spending

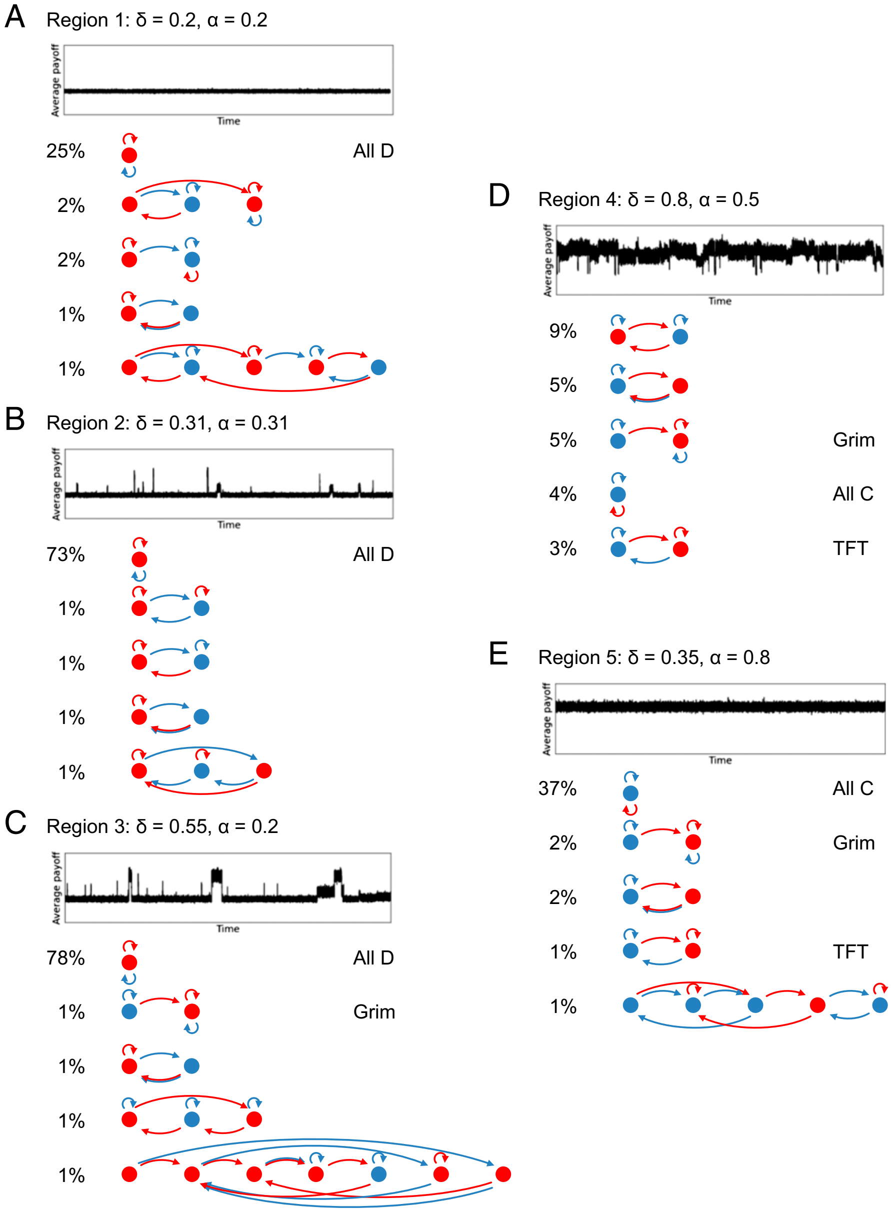


Fig. 3. Runs and top five strategies from the five regions. Each panel (*A*–*E*) shows simulation results using a (δ, α) pair taken from the center of the corresponding region of Fig. 2. In each panel, the average payoff over time is shown, as well the five most frequently observed strategies. If any of the strategies have an established name, it is also given. The simulation runs confirm the dynamic behavior the theoretical analysis suggests. In region 1 (*A*) we only see fully defecting equilibria; all strategies in the top five—and actually almost all strategies observed—always defect when playing against themselves. In region 2 (*B*) we observe that indirect invasions get the population away from full defection sometimes, but direct or indirect invasions bring it back to full defection relatively quickly. In region 3 (*C*) we observe different equilibria, ranging from fully defecting to fully cooperative. In region 4 (*D*) we observe high levels of cooperation, and although cooperative equilibria are regularly invaded indirectly, cooperation is always re-established swiftly. Furthermore, most of the cooperative strategies we observe are conditional. In region 5 (*E*) we observe full cooperation, and strategies that always cooperate against themselves. By far the most common strategy is ALLC, which cooperates unconditionally. Note that the top five also contain disequilibrium strategies. Strategies 2, 3 and 4 in the top five of region 2 are neutral mutants to the most frequent strategy there (ALLD), and ALLC, which came in fourth in region 4, is a neutral mutant to all fully cooperative equilibrium strategies.

some time in cooperative states. In our simulations, direct invasions out of cooperation are relatively rare, possibly because direct invaders are hard to find by mutation. Instead, cooperative states are more often left by indirect invasions.

In region 3 there exist equilibrium strategies for levels of cooperativity ranging from 0 to 1. In the simulations, we observe the population going from equilibrium to equilibrium via indirect invasions. The population spends time in states ranging from fully cooperative to completely uncooperative. As α and δ increase, indirect invasions that increase cooperation become more likely and indirect invasions that decrease cooperation become less likely; therefore the average level of cooperativity increases.

In region 4 there still exist equilibrium strategies for different levels of cooperativity, but fully defecting strategies are no longer equilibria. All equilibria are at least somewhat cooperative. Indirect invasions by fully defecting strategies are possible, and they do occur, but they result in relatively short-lived excursions into fully defecting disequilibrium states, which can be directly invaded by

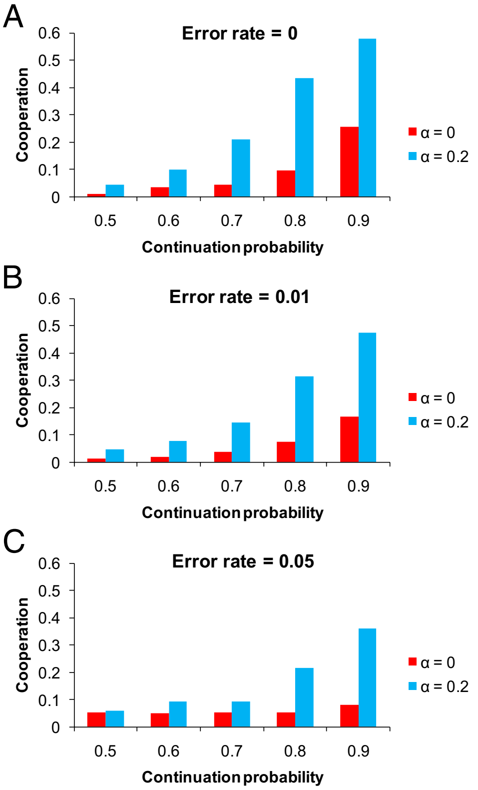


Fig. 4. Simulation results with and without noise. (*A–C*) The simulations shown in Figs. 1–3 have no errors; individuals observe the actions of their opponent with perfect accuracy, and make no mistakes in executing their own actions. That is of course a stylized setting, and it is more reasonable to assume that in reality errors do occur (11, 44, 45, 48–50). For five values of δ and two values of α we therefore repeat our simulations, but now with errors; once with an execution error of 1% per move, and once with an execution error of 5% per move. With errors, even ALLD against itself sometimes plays C, so the benchmark of no cooperation becomes the error rate, rather than 0. Errors decrease the evolution of cooperation somewhat, but the results do not change qualitatively. If anything, the effect of the combination of repetition and population structure is more pronounced; at an error rate of 5% both mechanisms have only a very small effect by themselves, but together make a big difference at sizable continuation probabilities.

strategies with positive cooperativity. As a result, cooperativity is high across much of region 4. In particular, reciprocal cooperative strategies, which condition their cooperation on past play, are common, for example TFT and Grim.

Finally, in region 5 all equilibrium strategies have maximum cooperativity, and ALLC can directly invade every strategy that is not fully cooperative. It is disadvantageous to defect regardless of the other's behavior. Therefore, not only is cooperativity high, but specifically unconditional cooperation (ALLC) is the most common cooperative strategy by a wide margin.

The five regions are separated by four curves, which are calculated in the *SI Appendix*. One remarkable finding is the following. If assortment is sufficiently high, $\alpha > c/b$, then introducing repetition (choosing $\delta > 0$) can be bad for cooperation. The intuition behind this finding is that reciprocity not only protects cooperative strategies from direct invasions by defecting ones, but also shields somewhat cooperative strategies from direct invasion by more cooperative strategies. (Details are in the *SI Appendix*, along with an explanation how repetition can facilitate indirect invasions into ALLC for large enough δ. See also ref. 47 and references therein for games other than the Prisoner's Dilemma in which repetition can be bad, even without population structure). If we move horizontally through the parameter space, starting in region 5 and increasing the continuation probability, average cooperativity in the simulations therefore first decreases (Fig. 2). Later, cooperativity goes back up again. The reason for this result is that equilibrium strategies can start with "handshakes," which require one or more mutual defections before they begin to cooperate with themselves (see strategies B, E, F, I, J, K, and O in Fig. 1). The loss of cooperativity for any given handshake decreases if the expected length of the repeated game increases, because the handshake then becomes a relatively small fraction of total play. That effect is only partly offset by the fact that an increase in continuation probability also allows for equilibria with longer handshakes.

In our simulations, individuals do not make mistakes; they both perceive the other's action and execute their own strategy with perfect accuracy. Mistakes, however, are very relevant for repeated games (11, 44, 45, 48–50). Therefore, we also ran simulations with errors for a selection of parameter combinations to check the robustness of our results. In those runs, every time a player chooses an action, C or D, there is some chance that the opposite move occurs. Fig. 4 compares error-free simulations with those that have a 1% and 5% error rate. We find that our conclusions are robust with respect to errors: it is the interaction between repetition and structure that yields high cooperation. Another classic extension is to include complexity costs (12, 48). In addition, here one can reasonably expect that the simulation results will be very similar as long as complexity costs are sufficiently small.

In summary, we have shown that repetition alone is not enough to support high levels of cooperation, but that repetition together with a small amount of population structure can lead to the evolution of cooperation. In particular, in the parameter region where repetition is common and assortment is small but nonzero, we find a high prevalence of conditionally cooperative strategies. These findings are noteworthy because human interactions are typically repeated and occur in the context of population structure. Moreover, experimental studies show that humans are highly cooperative in repeated games and use conditional strategies (14, 43–46). Thus, our results seem to paint an accurate picture of cooperation among humans. Summarizing, one can say that one possible recipe for human cooperation may have been "a strong dose of repetition and a pinch of population structure."

ACKNOWLEDGMENTS. We thank Tore Ellingsen, Drew Fudenberg, Corina Tarnita, and Jörgen Weibull for comments. This study was supported by the Netherlands Science Foundation, the National Institutes of Health, and the Research Priority Area Behavioral Economics at the University of Amsterdam; D.G.R. is supported by a grant from the John Templeton Foundation.

1. Friedman J (1971) A noncooperative equilibrium for supergames. *Rev Econ Stud* 38:1–12.
2. Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54:533–554.
3. Abreu D (1988) On the theory of infinitely repeated games with discounting. *Econometrica* 56:383–396.
4. Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211:1390–1396.
5. Selten R, Hammerstein P (1984) Gaps in Harley's argument on evolutionarily stable learning rules and in the logic of "tit for tat." *Behav Brain Sci* 7:115–116.
6. Boyd R, Lorberbaum JP (1987) No pure strategy is stable in the repeated prisoner's dilemma game. *Nature* 327:58–59.
7. May R (1987) More evolution of cooperation. *Nature* 327:15–17.
8. Nowak MA, Sigmund K (1992) Tit for tat in heterogeneous populations. *Nature* 355:250–253.
9. Nowak MA, 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.
10. Imhof LA, Fudenberg D, Nowak MA (2005) Evolutionary cycles of cooperation and defection. *Proc Natl Acad Sci USA* 102:10797–10800.

11. Fudenberg D, Maskin E (1990) Evolution and cooperation in noisy repeated games. *Am Econ Rev* 80:274–279.
12. Binmore KG, Samuelson L (1992) Evolutionary stability in repeated games played by finite automata. *J Econ Theory* 57:278–305.
13. Kim Y-G (1994) Evolutionarily stable strategies in the repeated prisoner's dilemma. *Math Soc Sci* 28:167–197.
14. Dal Bó P, Fréchette GR (2011) The evolution of cooperation in infinitely repeated games: Experimental evidence. *Am Econ Rev* 101:411–429.
15. Liberman V, Samuels SM, Ross L (2004) The name of the game: predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Pers Soc Psychol Bull* 30:1175–1185.
16. Bendor J, Swistak P (1995) Types of evolutionary stability and the problem of cooperation. *Proc Natl Acad Sci USA* 92:3596–3600.
17. Abreu D, Pearce D, Stacchetti E (1990) Optimal cartel equilibrium with imperfect monitoring. *J Econ Theory* 39:251–269.
18. Milinski M (1990) No alternative to Tit for Tat in sticklebacks. *Anim Behav* 39: 989–991.
19. Clutton-Brock T (2009) Cooperation between non-kin in animal societies. *Nature* 462: 51–57.
20. Hamilton WD (1964) The genetical evolution of social behaviour. I. *J Theor Biol* 7:1–16.
21. Eshel I, Cavalli-Sforza LL (1982) Assortment of encounters and evolution of cooperativeness. *Proc Natl Acad Sci USA* 79:1331–1335.
22. Nowak MA, May RM (1992) Evolutionary games and spatial chaos. *Nature* 359:826–829.
23. Durrett R, Levin S (1994) The importance of being discrete (and spatial). *Theor Popul Biol* 46:363–394.
24. Wilson DS, Dugatkin LA (1997) Group selection and assortative interactions. *Am Nat* 149:336–351.
25. Rousset F, Billiard S (2000) A theoretical basis for measures of kin selection in subdivided populations: Finite populations and localized dispersal. *J Evol Biol* 13:814–825.
26. Rousset F (2004) *Genetic Structure and Selection in Subdivided Populations* (Princeton Univ Press, Princeton, NJ).
27. Traulsen A, Nowak MA (2006) Evolution of cooperation by multilevel selection. *Proc Natl Acad Sci USA* 103:10952–10955.
28. Fletcher JA, Doebeli M (2009) A simple and general explanation for the evolution of altruism. *Proc Biol Sci* 276:13–19.
29. van Veelen M (2009) Group selection, kin selection, altruism and cooperation: When inclusive fitness is right and when it can be wrong. *J Theor Biol* 259:589–600.
30. Tarnita CE, Ohtsuki H, Antal T, Fu F, Nowak MA (2009) Strategy selection in structured populations. *J Theor Biol* 259:570–581.
31. Nowak MA, Tarnita CE, Antal T (2010) Evolutionary dynamics in structured populations. *Philos Trans R Soc Lond B Biol Sci* 365:19–30.
32. Richerson P, Boyd R (2005) *Not by Genes Alone. How Culture Transformed Human Evolution* (Univ of Chicago Press, Chicago).
33. van Veelen M (2012) Robustness against indirect invasions. *Games Econ Behav* 74:382–393.
34. van Veelen M, García J (2010) In and out of equilibrium: Evolution of strategies in repeated games with discounting, TI discussion paper 10-037/1. Available at http://www.tinbergen.nl/ti-publications/discussion-papers.php?paper=1587.
35. Marlowe FW (2005) Hunter-gatherers and human evolution. *Evol Anthropol* 14: 54–67.
36. Palla G, Barabási A-L, Vicsek T (2007) Quantifying social group evolution. *Nature* 446: 664–667.
37. Boyd R, Richerson PJ (1988) The evolution of reciprocity in sizable groups. *J Theor Biol* 132:337–356.
38. Ohtsuki H, Nowak MA (2007) Direct reciprocity on graphs. *J Theor Biol* 247: 462–470.
39. Tarnita CE, Wage N, Nowak MA (2011) Multiple strategies in structured populations. *Proc Natl Acad Sci USA* 108:2334–2337.
40. Rand DG, Arbesman S, Christakis NA (2011) Dynamic networks promote cooperation in experiments with humans. *Proc Natl Acad Sci USA* 108:19193–19198.
41. Grafen A (1985) A geometric view of relatedness. *Oxford Surveys in Evolutionary Biology* 2:28–90.
42. Bergstrom T (2003) The algebra of assortative encounters and the evolution of cooperation. *Int Game Theory Rev* 5:211–228.
43. Wedekind C, Milinski M (1996) Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus generous Tit-for-Tat. *Proc Natl Acad Sci USA* 93:2686–2689.
44. Aoyagi M, Frechette G (2009) Collusion as public monitoring becomes noisy: Experimental evidence. *J Econ Theory* 144:1135–1165.
45. Fudenberg D, Rand DG, Dreber A (2012) Slow to anger and fast to forgive: Cooperation in an uncertain world. *Am Econ Rev* 102:720–749.
46. Rand DG, Ohtsuki H, Nowak MA (2009) Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *J Theor Biol* 256:45–57.
47. Dasgupta P (2009) Trust and cooperation among economic agents. *Philos Trans R Soc Lond B Biol Sci* 364:3301–3309.
48. Hirshleifer J, Martinez Coll JC (1988) What strategies can support the evolutionary emergence of cooperation? *J Confl. Res.* 32:367–398.
49. Boyd R (1989) Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *J Theor Biol* 136:47–56.
50. Wu J, Axelrod R (1995) Coping with noise in the iterated prisoner's dilemma. *J Confl Res* 39:183–189.

Supporting Information for:

Direct reciprocity in structured populations

Matthijs van Veelen, Julián García, David G. Rand and Martin A. Nowak

1 How the combination of population structure and repetition can help the evolution of cooperation.

We begin with a simple illustration of how the two ingredients of the model together can allow cooperation to evolve. Therefore we consider a population structure where the population is divided into groups of equal size. If this division is made in an assortative way, then this implies that players that play strategy A face other players that play strategy A with a probability that is larger than the probability that random group formation would give. Figure S1 below depicts the extreme case, where assortment is complete.

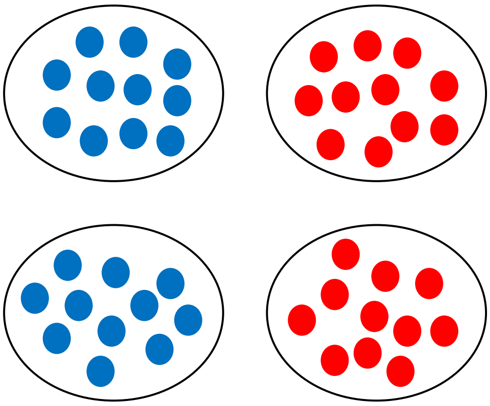


Fig. S1. A population divided in groups with complete assortment.

Obviously, with complete assortment the probability of a cooperator interacting with a cooperator is 1, while the probability of a defector meeting a cooperator is 0.

Complete assortment implies that already in a regular, unrepeated prisoners dilemma cooperation is always selected for, because cooperators playing against cooperators earn more than defectors playing against defectors.

A bit less extreme is the case in Figure S2 below, where the population frequency of cooperators is 0.5, but the probability of a cooperator interacting with a cooperator is

$38/55$, while the probability of a defector meeting a cooperator is $17/55$. (Consequently, the probability of a defector interacting with a defector is $38/55$, and the probability of a cooperator meeting a defector is $17/55$).

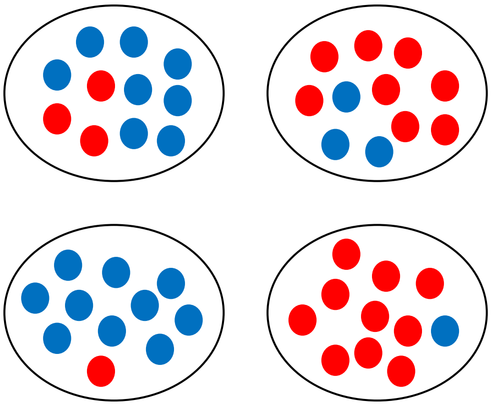


Fig. S2. A population divided in groups with less than complete assortment.

In a regular, unrepeated prisoners dilemma with $T = 3$, $R = 2$, $P = 1$ and $S = 0$ this implies that the expected payoff for a cooperator is $(38/55) \times 2 + (17/55) \times 0 \approx 1.38$, while the expected payoff for a defector is $(38/55) \times 1 + (17/55) \times 3 \approx 1.62$. Cooperative behaviour is therefore not selected with this population structure and without repetition.

If we replace the unrepeated, one shot prisoners' dilemma with the repeated version with a continuation probability δ and if we replace the strategy Cooperate by Tit-for-Tat (*TFT*) and Defect by Always Defect (*All D*), then the payoffs change. For the repeated game, we compute normalised discounted payoffs as $(1 - \delta) \sum_{i=0}^{\infty} \delta^i \pi_i$, where π_i is the payoff in round *i* and the term $(1 - \delta)$ normalises the discounted sum of the payoffs. With $\delta = 0.5$ the payoff matrix changes from the first to the second payoff matrix below.

	Cooperate	*Defect*
Cooperate	2	0
Defect	3	1

	TFT	*All D*
TFT	2	0.5
All D	2	1

With repetition, reciprocity therefore will reduce how much a cooperator loses when it meets a defector and how much a defector gains from meeting a cooperator. This allows cooperation to win where it could not win before; with the population structure we just had, the expected payoff for Tit-for-Tat is $(38/55) \times 2 + (17/55) \times 0.5 \approx 1.55$ while the expected payoff for a defector is $(38/55) \times 1 + (17/55) \times 2 \approx 1.31$. From the second payoff matrix it is clear that cooperation would also not be selected with repetition only and no population structure. Therefore this example illustrates that the combination of population structure and repetition can work in cases where none of the two would lead to the evolution of cooperation on its own.

For a population structure that is subdivided into groups, assortment can be measured as $P(T|T) - P(T|N)$ (see 1–5). This measure is the probability that a player with strategy T interacts with another player that also plays strategy T minus the probability that a player with a strategy other than strategy T interacts with a player that plays strategy T. This equals the probability with which a rare mutant meets a copy of itself, because $P(T|N) = 0$ if T is rare, and it can be interpreted as relatedness.

This is only an illustration that considers a few strategies and suggest why population structure and repetition (which allows for reciprocity to evolve) *might* reinforce each other. Below we will consider a setting that allows for all possible strategies, and where population structure and repetition are both parametrized by continuous parameters. For this very general setting it turns out that one can in fact make predictions, but we have to account for the complexities that come with the generality of the strategy set.

Other population structures that deviate from random matching are for instance models with local interaction or graph structured populations. Given some very mild conditions one can also characterize the population structure in, amongst others, graph structured populations, with a single measure σ, which together with the payoff matrix suffices to determine the direction of selection (6). Also here, the reason why the combination of population structure and repetition can work is that with repetition cooperators that interact with defectors will transfer less payoff to them, if the cooperators are reciprocal.

We are not the first to point to the importance of the interaction between assortment and repetition. The seminal paper of Axelrod and Hamilton (7) discussed both elements: their stability analysis considered repeated games in well-mixed populations, while the role of assortment was pointed out verbally. In the current paper, we present a theoretical analysis of both aspects, describing exactly how the two ingredients interact and demonstrating the need for both in order to explain human conditional cooperation. More recent examples of papers with models that combine conditional strategies and (moderate levels of) assortment are (5) and (8–10), which are discussed in more depth in section 5 of the SI.

Another interesting way to look at these two ingredients is that on top of the assortment that is induced by the population structure, reciprocity can also be seen as generating assortment. This assortment is then not one in strategies, but in cooperative actions, and is created, not instantaneously, but with a (small) delay. By matching the previous move of its opponent, Tit-for-Tat for instance increases assortment in cooperative behaviours in any population. Both ingredients can therefore be seen as different sources of assortment (11). We should stress, however, that the analysis below shows that it is also very much worth treating them as separate ingredients, and not as interchangeable sources of assortment. The picture that the equilibrium analysis paints for the case in which there is only assortment in the more standard sense is very different from what it finds for the case in which there is only repetition. While the analysis predicts only one equilibrium level of cooperation for the case of assortment without repetition, it predicts a multiplicity of equilibrium levels of cooperation in the case of repetition without (spatial) assortment, with movement between them. The finding that repetition and reciprocity can actually also harm cooperation – which is the case in a part of the parameter space where assortment is high – also shows that there is not a straightforward way to substitute one for the other and get similar results, even when only considering the *average* level of cooperation.

2 The simulations

The population consists of an even number of individuals, which are matched in pairs. Each individual is endowed with a finite state automaton, which codes for a strategy for the repeated game. Not all possible strategies in the repeated prisoners dilemma are finite state automata, but one can show that in a very natural sense, all strategies can be approximated arbitrarily closely by a finite state automaton (12). Using finite state automata therefore is only a very mild restriction on the strategy space.

The two individuals in every pair play the repeated game with each other, where the number of repetitions is a random variable; the first round is played for sure, and then every next repetition occurs with probability δ. This implies that the number of rounds follows a geometric distribution, where the probability that the game lasts at least i rounds is δ^{i-1}, and the probability that it lasts exactly i rounds is $\delta^{i-1}(1-\delta)$. The payoffs gathered by an individual in the different rounds are added up to give its total payoff. This follows the standard interpretation of δ as a continuation probability in repeated games.

The total payoffs they gather from their interaction are used in the update step in which the next generation is drawn. Every pair in the next generation is drawn following the same procedure. First the first individual in that pair is drawn, where each individual in the parent population has a probability proportional to payoff to be the parent (the probability that individual j in the parent population is drawn as a parent is its payoff divided by the sum of the payoffs of all individuals in the parent population). For the second individual, a nested procedure applies. With probability α the parent of the first individual is chosen to also become the parent of the second individual. With probability $1-\alpha$ the second individual is drawn in the same way as the first one, which means that each individual in the parent population has a probability proportional to payoff to become the parent. The parent of the first individual is not excluded there, so in that sense one could say that it is with replacement.

Two things are worth noting here. The first is that with $\alpha = 0$ we are back in the random matching procedure. The second is that the expected number of offspring of every member of the parent population is not affected by the value of α in this procedure; for every α the expected number of offspring of an individual in the parent population is its own payoff divided by the average payoff in the parent population.

Then every individual has a probability with which it mutates. The mutation procedure is the same as in (12); if an individual gets to mutate, then either a state is added, deleted, the action that is played when in a state is changed, or which state to go to, depending on the action of the other, is changed. If a state is added, one arrow is randomly selected to point to this state. If a state is deleted, then all arrows pointing towards that state are one by one randomly reassigned to another state, where each state is equally likely to be chosen.

This ensures that every finite state automaton can be reached from every other finite state automaton in a finite number of mutation steps. The mutation step completes the cycle, which is then repeated a large number of times.

One advantage of the matching procedure is that it allows us to vary the population structure continuously, in a way that is consistent with the classical setting in (1), where the parameter m has the same role as α here. Note that for an alternative wide set of population structures and update rules, one could compute a characteristic sigma (6) for each combination of a population structure and update rule, which could also play a role similar to the assortment parameter, but to inverse engineer population structures that give a continuum of such sigma’s is much harder. What we did here is basically form groups of size 2 – where the examples in Figures S1 and S2 have groups of size 11 – in a way that can be done for every α between 0 and 1, and that is consistent with (1–5).

The simulation program, including more detailed explanation, can be found on www.evolutionandgames.com and the code is available on request.

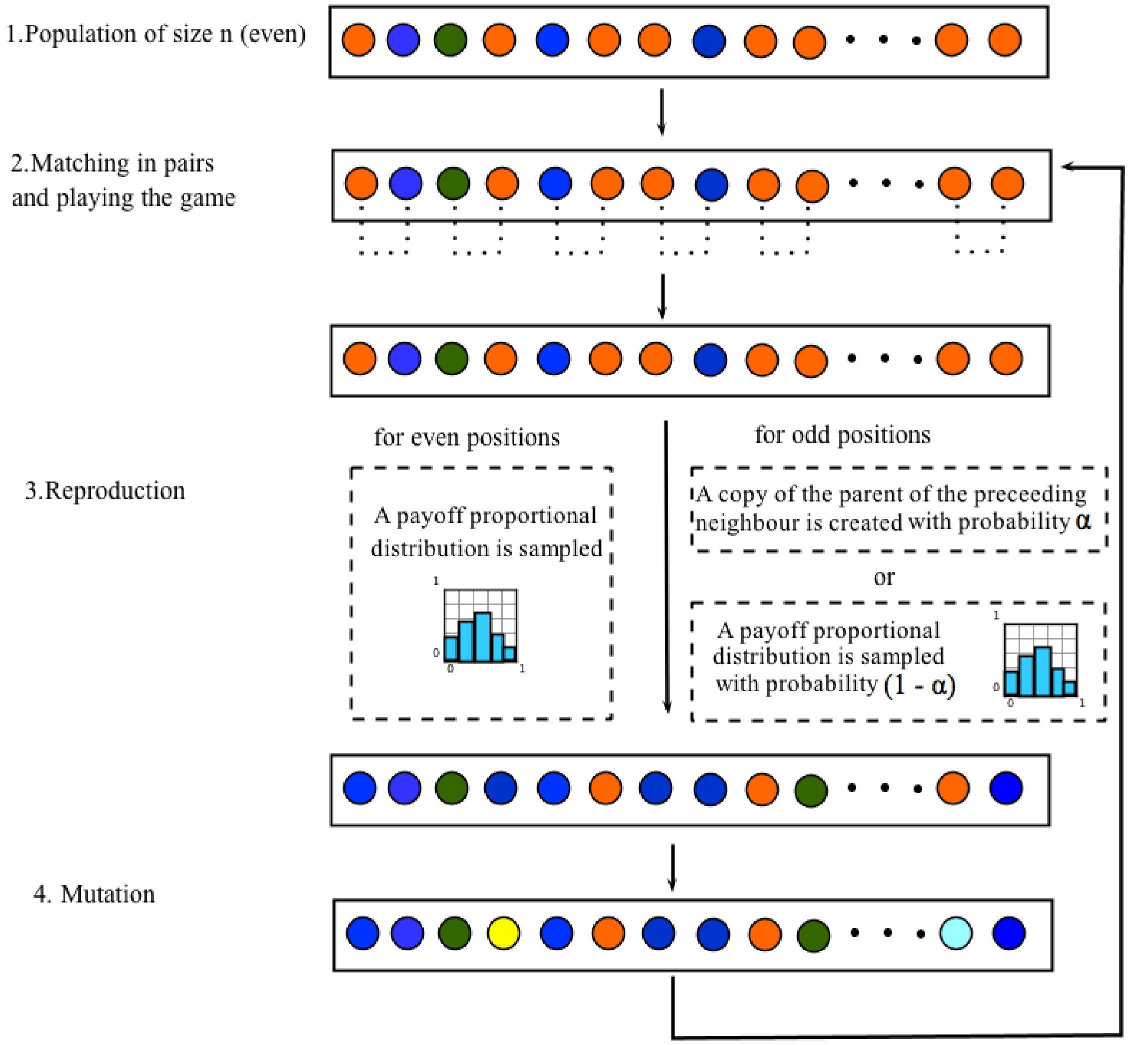


Fig. S3. The life cycle in the simulations.

In Figures 1 and 3 of the main text we represent automata by their smallest equivalent. All strategies for which the output in every state is D, for instance, are reduced to the simplest *All D*: a single state automaton with output D. That also implies that in Figure 3 of the main text we lump every implementation of *All D* together. There is good reason to do that – they are all the same strategy – but it should be noted that mutation probabilities to other strategies of course depend on the actual implementation.

Note that the results will naturally be sensitive to the choice of the mutation procedure. If we change to a strange mutation procedure that never adds a state with output C and never changes the output in a state to a C, then cooperation will obviously never evolve from *All D*. Also the requirement that it should be possible to go from one automaton to any other automaton by a sequence of mutation steps with positive probability does not guarantee anything, as one can always make some mutations extremely unlikely compared to others. What we do have is that mutations are not biased against D or C, so without selection pressure, automata evolve that on average cooperate half the time. But even that does not imply we should get a picture like Figure 2 in the main text; if adding a state is sufficiently less likely than deleting one, Figure 2 will consist of a black block everywhere below $\alpha = 0.5$ and a yellow block everywhere above. General statements about the robustness to all changes in mutation probabilities therefore are hard to construct. It seems, however, that simulation results are relatively robust to changes in the probabilities of adding, deleting and changing states and arrows, as long as they remain of the same order of magnitude.

3 Simulation results

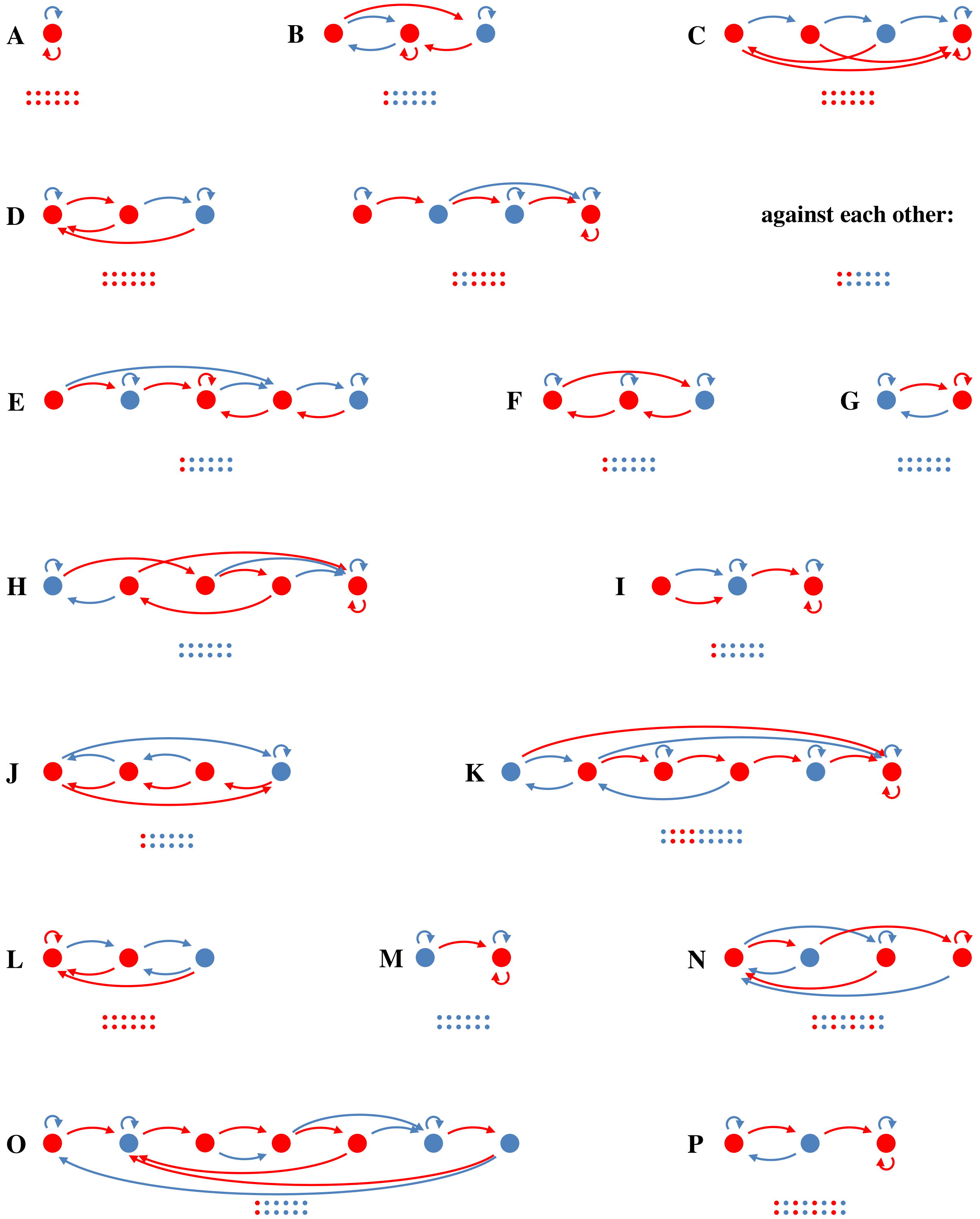


Fig. S4. The strategies from Fig. 1 in the main text.

3.1 Description of the strategies in Figure 1 in the main text; $\delta = 0.85$, $\alpha = 0$.

The population spends a lot of time in *All D* (A) or similar, fully uncooperative equilibria (C, L), but also visits more cooperative equilibria. Some of those are classic strategies, like *TFT* (G) or *Grim Trigger* (M), but more often strategies will be combinations of different components. They vary in their reciprocity, their use of codes before they start cooperating, their ability to exploit naïve cooperators, and their level of cooperation if it is established.

All equilibria in which there is cooperation must be reciprocal, but while some strategies are completely unforgiving (I, K, M), others punish defections with a minimum of 1 (G, N), 2 (B, F) or 3 (H, J) defections, where in some of those a return to cooperation does (B, G, H, N) and in others does not (F, J) require its opponent to "apologize" by playing cooperate at the right time. More complex forms of reciprocity, with different cooperative states, are the strategies at E and O.

Some strategies use a code before starting to cooperate, where the codes we see in this part of the run are to defect once (B, E, F, I, J, O), or to play a sequence of one cooperation and three defections (K). If the opponent does not provide this code, then in some cases cooperation is started anyway (I, J), while in other cases the strategy defects until the right code is provided after all (B, F, O), or never gives it another chance and simply defects forever (K). These strategies thereby also differ in how well they exploit some more naïve cooperators. The strategy at E has alternative codes leading to different cooperative states.

The level of cooperation that is achieved – once past the initial code, if there is one – can also differ; in these examples they then cooperate every period (B, E-K, M) or every other period (N, P) when playing against themselves.

The population at D is a mix of two strategies, where each strategy performs poorly against itself, but both trigger cooperation in each other. They thereby do well against each other. While the "less nice" one of the two (in this case the second one) gets the better end of the deal, it is nonetheless prevented from outperforming the other overall by its bad performance against itself.

3.2 Levels of cooperation for $\alpha = 0$ and varying continuation probability

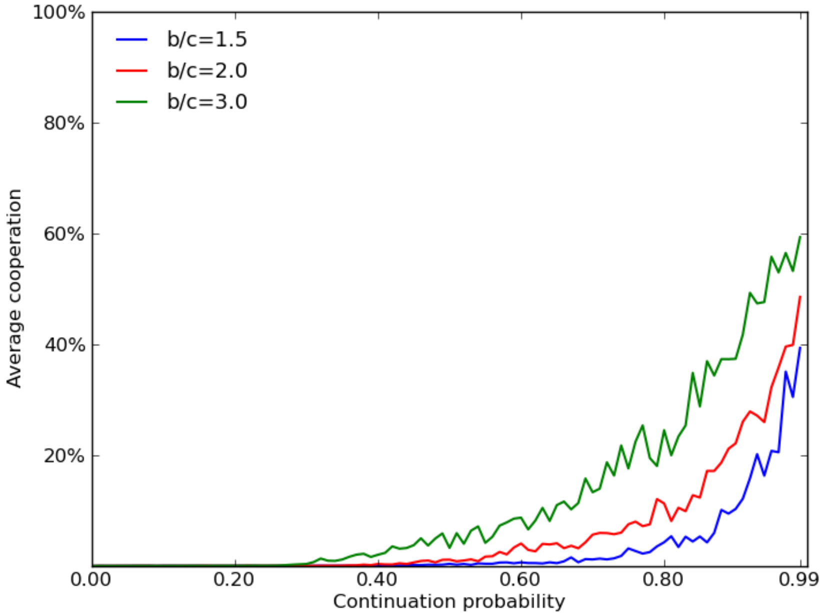


Fig. S5. Levels of cooperation for different b/c **ratios and continuation probabilities**

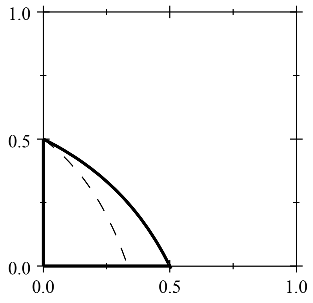
Every data point is the average level of cooperation during four runs of 500 000 generations each. The game is a repeated prisoners' dilemma with varying *b*/*c*-ratios. The red line is for the stage game where a co-operator gets a payoff of 2 when it interacts with another co-operator, 0 if it interacts with a defector, and where a defector gets a payoff of 3, when it interacts with a co-operator, and a payoff of 1 if it interacts with another defector. This gives a *b*/*c*-ratio of 2. The green and the blue line are for stage games with *b*/*c* equal to 3 and 1.5, respectively. Continuation probability is varied in steps of 0.01 (and only goes up to .99, because 1.0 would imply that games would never actually end). Even for $\delta = 0.99$ defection was still quite common.

4 Theoretical analysis

4.1 Description of the regions

All figures are drawn for $\frac{b}{c} = 2$.

Region 1



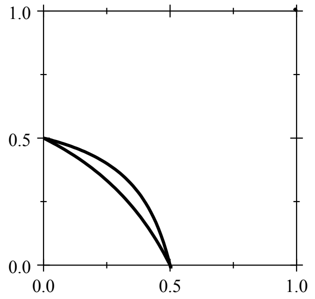
Region 1 is the area below the line $\alpha = \frac{c-\delta b}{b-\delta b}$. The region is split in two by the line $\alpha = \frac{c-\delta(c+b)}{b-\delta(c+b)}$.

In this region, *All D* is the most prominent equilibrium. It is not the only one, but what all other equilibria in this region share with *All D,* is that they always play D when playing against themselves. *All D* can also directly invade any other strategy which does cooperate when it plays against itself, even if only a little. Therefore, in this region, there is no equilibrium that is even somewhat cooperative.

If r is below the threshold depicted by the dashed line, then all equilibria are even robust against indirect invasions (RAII). An indirect invasion occurs when a neutral mutant opens the door for another mutant, where the second mutant only has a selective advantage after the neutral mutant has gained a certain share in the population, but not before.

Above this threshold indirect invasions into *All D* and other equilibria are possible, but because within region 1 there are no equilibria that are even somewhat cooperative, the population cannot settle on a (somewhat) cooperative equilibrium after the indirect invasion. Moreover, whatever gets established with the indirect invasion can in turn be invaded directly by *All D*. We therefore expect the population to quickly revert to a fully uncooperative equilibrium after any indirect invasion in the part of region 1 in which indirect invasions are possible.

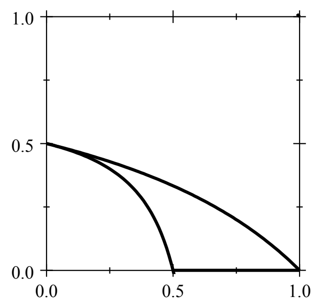
Region 2



Region 2 is the area above the line $\alpha = \frac{c-\delta b}{b-\delta b}$, and below $\alpha = \frac{c-\delta b}{b-2\delta b+\delta c}$.

Region 2 is similar to region 1 in the sense that all cooperative strategies can directly be invaded. Apart from being similar, it is also different from region 1, in the sense that it is not *All D* that can directly invade cooperative strategies here. It can be invaded directly, but by a strategy that is not fully defecting when it plays against itself. This invader, in turn, is also vulnerable to direct invasion by less cooperative strategies. These mutations however are relatively unlikely to appear in our simulations, because it takes a bit of tinkering to construct them (see Theorem 3). The population can therefore be stuck for a while in such a cooperative disequilibrium state. It turns out that such states are regularly left through indirect invasion after all (first a neutral mutant, then one with an actual advantage) which suggests that even though the possibility of a direct invasion is there, it may be relatively hard to find. Arriving at these (somewhat) cooperative strategies also happens by indirect invasions. The fact that both getting there and away regularly happens by indirect invasions, in spite of the presence of a (hard to find) direct way out, makes the dynamics in region 2 a bit more like region 3 – where indirect invasions in and out of equilibrium are all that is there – than like region 1 – where all there is, is neutral movement between fully defecting states (see Figure 3 in the main text).

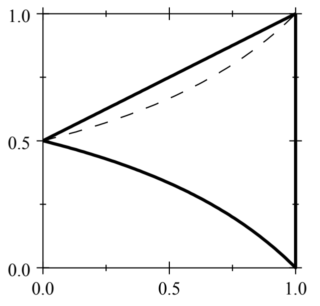
Region 3



Region 3 is the area above the line $\alpha = \frac{c-\delta b}{b-2\delta b+\delta c}$, and below $\alpha = \frac{c-\delta c}{b-\delta c}$.

In this region there is a host of equilibria, ranging from fully defecting (for instance *All D*) to fully cooperative (for instance *TFT* or *Grim Trigger*). The population moves between these equilibria by means of indirect invasions.

Region 4



Region 4 is the area above the line $\alpha = \frac{c-\delta c}{b-\delta c}$ and below $\alpha = \frac{\delta(b-c)+c}{b}$. The dotted line is $\alpha = \frac{c}{b-\delta(b-c)}$.

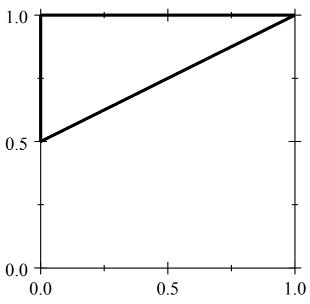
In region 4 there are also multiple equilibria. The difference between region 3 and region 4 is that in the latter none of them are fully defecting. In other words: all equilibria now are at least somewhat cooperative. Indirect invasions by fully defecting strategies are possible, and do occur, but then result in relatively short-lived stays in a defecting disequilibrium state.

Going from region 3 to region 4 is an important shift, because in region 4 full defection no longer is an equilibrium. The reason why the crossing of the border between those regions is especially interesting for humans is that the amount of population structure needed to reach the threshold is decreasing as the continuation probability increases. With continuation probabilities close to 1, even a very small positive α is enough to move into region 4 and lose the fully defecting equilibria.

A remarkable thing is that while repetition helps cooperation in the lower half of region 4, it actually harms cooperation in the upper half. The reason is that reciprocity can also safeguard a moderate level of defection from direct invasions of a higher level of cooperation. In the results section we will also see how not only in the lower half, but also in the upper half, close to region 5, indirect invasions with decreasing levels of cooperation are possible.

Above the dotted line, *All C* and all other strategies that always play *C* against themselves are RAII.

Region 5



Region 5 is the area above the line $\alpha = \frac{\delta(b-c)+c}{b}$.

In Region 5, *All C* is the most prominent equilibrium. It is not the only one, but what all other equilibria in this region share with *All C* is that they always play *C* in equilibrium. In fact, any strategy that always plays C against itself is an equilibrium in region 5, and moreover even RAII. *All C* itself can also directly invade any other strategy which plays defect against itself, even if it does so only once. Therefore, in this region, there is no equilibrium that falls short of full cooperation.

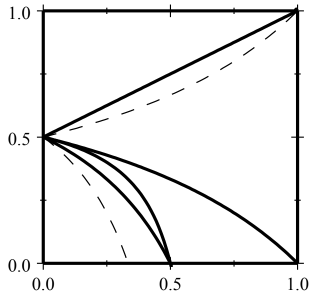


Fig. S6. All regions together in one figure.

4.2 Definitions

The characteristics of and boundaries between these regions are given by the results presented below. The results are general in the sense that they look at equilibrium strategies in the unrestricted strategy space of the repeated game (see 12 for the case without population structure). We will apply a few well known concepts; an equilibrium, an evolutionarily stable strategy (ESS) and a neutrally stable strategy (NSS), which is also known as *weak* ESS. We also use the concept of a strategy that is robust against indirect invasions (RAII). In this setting their definitions all have to be extended – or applied in an unusually broad sense – in order to encompass population structure. With random matching ($\alpha = 0$) they revert to their normal definitions.

Definition 1: A strategy S – or a strategy profile (S, S) – is a (symmetric) equilibrium if for every strategy S' the following holds:

$$\alpha\pi(S', S') + (1 - \alpha)\pi(S', S) \leq \pi(S, S)$$

Without structure – that is, with $\alpha = 0$ – this is the definition of a symmetric Nash equilibrium (13, 14 see also 15).

Definition 2: A strategy S is an *evolutionarily stable strategy* if for every strategy $S' \neq S$ the following holds:

1) $\alpha\pi(S',S') + (1-\alpha)\pi(S',S) \leq \pi(S,S)$

2) If *1)* holds with equality, then

$$\pi(S',S') - \pi(S',S) < \pi(S,S') - \pi(S,S)$$

Without structure – that is, with $\alpha = 0$ – this is the standard definition of an evolutionarily stable strategy (16, see also 15) because the equality in condition *1)* for $\alpha = 0$ implies that condition *2)* simplifies to $\pi(S',S') < \pi(S,S')$. In order to see why also for $\alpha > 0$ this should be the condition, let p_ϵ be a mixture of a $1-\epsilon$ share of incumbents S and an ϵ share of mutants S'. With assortment α, the probability with which an individual playing S is facing copy of itself, if it finds itself in population p_ϵ, is $\alpha + (1-\alpha)(1-\epsilon)$, while the probability with which S' is facing copy of itself is $\alpha + (1-\alpha)\epsilon$. The payoffs of both strategies against the whole population is therefore:

$$\pi(S,p_\epsilon) = \alpha\pi(S,S) + (1-\alpha)\{\epsilon\pi(S,S') + (1-\epsilon)\pi(S,S)\}$$

$$= \pi(S,S) + \epsilon(1-\alpha)\{\pi(S,S') - \pi(S,S)\}$$

$$\pi(S',p_\epsilon) = \alpha\pi(S',S') + (1-\alpha)\{\epsilon\pi(S',S') + (1-\epsilon)\pi(S',S)\}$$

$$= \alpha\pi(S',S') + (1-\alpha)\pi(S',S)$$

$$+\epsilon(1-\alpha)\{\pi(S',S') - \pi(S',S)\}$$

It is clear that there is an $\bar{\epsilon} \in (0,1)$ such that $\pi(S,p_\epsilon) > \pi(S',p_\epsilon)$ for all $\epsilon \in (0,\bar{\epsilon})$ if and only if conditions *1)* and *2)* apply. The more general definitions in (1) and (17) can therefore encompass this simple population structure already (see 18).

Definition 3: A strategy S is a *neutrally stable strategy* if for every strategy $S' \neq S$ the following holds:

1) $\alpha\pi(S',S') + (1-\alpha)\pi(S',S) \leq \pi(S,S)$

2) If *1)* holds with equality, then

$$\pi(S',S') - \pi(S',S) \leq \pi(S,S') - \pi(S,S)$$

Without structure – that is, with $\alpha = 0$ – this is the standard definition of a neutrally stable strategy (19, see also 15). A neutrally stable strategy is sometimes also referred to as a *weakly* evolutionarily stable strategy

If S is not a neutrally stable strategy, there is apparently a strategy S' for which either

$$\alpha\pi(S',S') + (1-\alpha)\pi(S',S) > \pi(S,S),$$

or

$$\alpha\pi(S',S') + (1-\alpha)\pi(S',S) = \pi(S,S) \text{ and}$$

$$\pi(S',S') - \pi(S',S) > \pi(S,S') - \pi(S,S),$$

In that case we say that S' can *directly* invade S.

We say that a strategy $S = S_0$ can be *indirectly invaded* if there is a sequence of strategies $S_1, \dots, S_{m+1}$ such that $S_1, \dots, S_m$ is a sequence of neutral mutants, that is, for $i = 1, \dots, m-1$

$$\pi(S_{i+1},S_{i+1}) = \pi(S_{i+1},S_i) = \pi(S_i,S_{i+1}) = \pi(S_i,S_i)$$

and S_{m+1} can directly invade S_m, that is

$$\alpha\pi(S_{m+1}, S_{m+1}) + (1-\alpha)\pi(S_{m+1}, S_m) > \pi(S_m, S_m)$$

or

$$\alpha\pi(S_{m+1}, S_{m+1}) + (1-\alpha)\pi(S_{m+1}, S_m) = \pi(S_m, S_m) \text{ and}$$

$$\pi(S_{m+1}, S_{m+1}) - \pi(S_{m+1}, S_m) > \pi(S_m, S_{m+1}) - \pi(S_m, S_m),$$

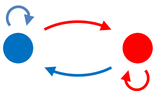
With repeated games, where neutral mutants abound, it is important to account for indirect invasions too, and hence we use the concept of robustness against them:

Definition 4: A strategy S is a *robust against indirect invasions* if it cannot be indirectly invaded.

Without structure – that is, with $\alpha = 0$ – this is the normal definition of a strategy that is RAII (20). Every strategy that is ESS is also RAII, but not vice versa, and every strategy that RAII is NSS, but not vice versa.

These definitions allow the statement of all the results that follow.

Example of an indirect invasion in TFT for $\alpha = 0$



TFT

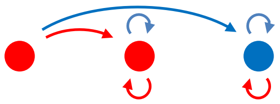
Neutral mutant

Advantageous mutant

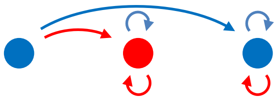
Examples of indirect invasions in All D for $\alpha = 0$

Into full cooperation (but not an equilibrium)

All D



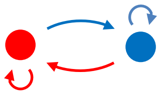
Neutral mutant



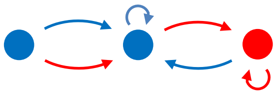
Advantageous mutant

Into a (mixed, but not completely cooperative) equilibrium

All D



Neutral mutant



Advantageous mutant

Note that we restrict ourselves to pure strategies; not only do the definitions restrict the possible equilibrium strategies to pure ones, also the possible invaders against which stability is to be checked are pure. Thereby we follow the literature; all classic papers on evolution in repeated games do that (7, 21–25). We are however working on a more technical follow-up paper that shows that the same theorems that we prove in this paper also hold if we do admit mixed strategies.

4.3 Results

For all results we will use the following payoff matrix.

	Cooperate	*Defect*
Cooperate	*b – c*	*- c*
Defect	*b*	*0*

As all definitions above are insensitive to adding a constant to all payoffs, this is only a convenient normalisation. Every result derived below therefore applies equally well to other possible payoff matrices with the same costs and benefits of switching between cooperation and defection, such as for instance

	Cooperate	*Defect*
Cooperate	*1 + b – c*	*1 – c*
Defect	*1 + b*	*1*

In region 1 All D can directly invade all cooperative strategies.

The boundary between region 1 and 2 is given by $\alpha = \frac{c-\delta b}{b-\delta b}$. If α is below this threshold, then any strategy that is not fully defecting, when playing against itself, can directly be invaded by *All D*. In order to verify that this is indeed the threshold, we first compute the payoffs of an at least somewhat cooperative strategy, and of *All D*, both against themselves and against each other.

When *All D* plays against itself, it plays defect every round.

$$\begin{array}{ll} \mathit{All\ D} & D\ D\ D\ D\ \ldots\ldots \\ \mathit{All\ D} & D\ D\ D\ D\ \ldots\ldots \end{array}$$

Therefore it gets payoff 0 every round, which thereby also is the normalized, discounted payoff of *All D* against itself. This is denoted as $\pi(AllD, AllD) = 0$.

Now consider a strategy *X* that does not always play *D* when it meets itself. In that case there is obviously at least one moment in time where it plays *C* against itself. Assume that the first time it does so is at time $n + 1$, where $n \geq 0$. This implies that the first n times, *X* plays *D* against itself. Note that n can be 0, so a strategy that starts cooperating immediately is not excluded.

$$\begin{array}{llll} X & \overbrace{D\ldots D}^{n\geq 0} & C & ?\ ?\ ?\ \ldots\ldots \\ X & D\ldots D & C & ?\ ?\ ?\ \ldots\ldots \end{array}$$

If all the question marks were *C*'s – note that playing against itself, they come in pairs of *C*'s or pairs of *D*'s – then it would earn a discounted, normalized payoff of $\delta^n(b - c)$. That puts a maximum on the payoff of *X* playing against a copy of itself; $\pi(X, X) \leq \delta^n(b - c)$.

$$\begin{array}{llll} X & \overbrace{D\ldots D}^{n\geq 0} & C & ?\ ?\ ?\ \ldots\ldots \\ \mathit{All\ D} & D\ldots D & D & D\ D\ D\ \ldots\ldots \end{array}$$

When *X* and *All D* are playing each other, then the best case scenario for *X* – and the worst case scenario for *All D* – is if all the question marks are *D*'s. This way *X* only loses c once, and *All D* only gains b once, which then happens in round $n+1$. This gives the following upper and lower bound, respectively.

$$\pi(X, All\ D) \leq (1-\delta)\delta^n(-c)$$

$$\pi(All\ D, X) \geq (1-\delta)\delta^n b.$$

With these bounds, we can show the following.

Theorem 1 If $\alpha < \frac{c-\delta b}{b-\delta b}$, then no strategy *X* with $\pi(X,X) > 0$ is an equilibrium.

Proof If $\alpha < \frac{c-\delta b}{b-\delta b}$ and $\pi(X,X) > 0$, then

$$\alpha\pi(All\ D, All\ D) + (1-\alpha)\pi(All\ D, X) = (1-\alpha)\pi(All\ D, X)$$

$$> \left(1 - \frac{c-\delta b}{b-\delta b}\right)(1-\delta)\delta^n b = \delta^n(b-c) \geq \pi(X,X)$$

Hence *X* is not an equilibrium.

In Region 1, 2 and 3, All D is a neutrally stable strategy.

In regions 1 to 3, *All D* cannot be directly invaded, or, in other words, it is a neutrally stable strategy (NSS). In the proof, we first look at strategies *X* that do not always play *D* when they meet themselves, and for which $\pi(X,X) > 0$, and then at strategies *X* that do, and for which therefore $\pi(X,X) = 0$. If $\pi(X,X) > 0$, we again assume that it plays *C* for the first time in round $n+1$, with $n \geq 0$.

Theorem 2 If $\alpha < \frac{c-\delta c}{b-\delta c}$, then *All D* is a neutrally stable strategy.

Proof Assume that $\alpha < \frac{c-\delta c}{b-\delta c}$.

If $\pi(X,X) > 0$, then, since $\pi(X,X) > 0 > \pi(X, All\ D)$,

$$\alpha\pi(X,X) + (1-\alpha)\pi(X, All\ D) <$$

$$< \frac{c-\delta c}{b-\delta c}\pi(X,X) + \left(1 - \frac{c-\delta c}{b-\delta c}\right)\pi(X, All\ D)$$

$$\leq \frac{c-\delta c}{b-\delta c}\delta^n(b-c) + \frac{b-c}{b-\delta c}(1-\delta)\delta^n(-c) = 0$$

$$= \pi(All\ D, All\ D)$$

If $\pi(X,X) = 0$, then it must always play *D* against itself, and hence also against *All D*. Thereby

$$\alpha\pi(X,X) + (1-\alpha)\pi(X, All\ D) = 0 = \pi(All\ D, All\ D)$$

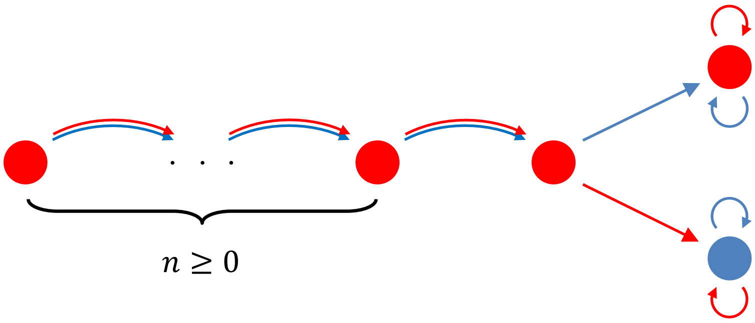
and also

$$\pi(X,X) - \pi(X, All\ D) = 0 \leq 0 = \pi(All\ D, X) - \pi(All\ D, All\ D).$$

Because $\pi(X,X) \geq 0$, this completes the proof, as *All D* is now shown to satisfy Definition 3 for all *X*.

Also in region 2 all equilibria are fully defecting

Consider a strategy X that does not always play defect against itself (i.e., $\pi(X, X) > 0$) and let $n \geq 0$ be a number of defections before the first time that C is played. We have already seen that $\pi(X, X) \leq \delta^n(b - c)$. Now consider strategy X_n which is defined by the following automaton.



When X_n plays itself, we get the following sequence of action profiles

$$\begin{array}{llll} X_n & \overbrace{D\ldots D}^{n \geq 0} & D & C\ \ C\ \ C\ \ldots\ldots \\ X_n & D\ldots D & D & C\ \ C\ \ C\ \ldots\ldots \end{array}$$

Thereby $\pi(X_n, X_n) = \delta^{n+1}(b - c)$. When X and X_n are playing each other, we get the following sequence of action profiles.

$$\begin{array}{llll} X & \overbrace{D\ldots D}^{n \geq 0} & C & ?\ \ ?\ \ ?\ \ldots\ldots \\ X_n & D\ldots D & D & D\ \ D\ \ D\ \ldots\ldots \end{array}$$

Thereby we find the following upper and lower bounds:

$$\pi(X, X_n) \leq (1-\delta)\delta^n(-c)$$

$$\pi(X_n, X) \geq (1-\delta)\delta^n b.$$

Theorem 3 If $0 \leq \alpha < \frac{c-\delta b}{b-2\delta b+\delta c}$ and $\delta < \frac{c}{b}$, then no strategy *X* with $\pi(X,X) > 0$ is an equilibrium.

Proof Assume that indeed $0 \leq \alpha < \frac{c-\delta b}{b-2\delta b+\delta c}$, $\delta < \frac{c}{b}$ and $\pi(X,X) > 0$. First note that with $b > c > 0$, $\delta < \frac{c}{b}$ implies that also $\delta < \frac{b}{2b-c}$ and thus $b - 2\delta b + \delta c > 0$. Therefore

$$\alpha\pi(X_n, X_n) + (1-\alpha)\pi(X_n, X) \geq$$

$$\alpha\delta^{n+1}(b-c) + (1-\alpha)(1-\delta)\delta^n b =$$

$$\delta^n\big((1-\delta)b + \alpha(-b + 2\delta b - \delta c)\big) >$$

$$\delta^n\left((1-\delta)b + \frac{c-\delta b}{b-2\delta b+\delta c}(-b+2\delta b-\delta c)\right) =$$

$$\delta^n(b-c) \geq \pi(X,X)$$

Thereby *X* is not an equilibrium.

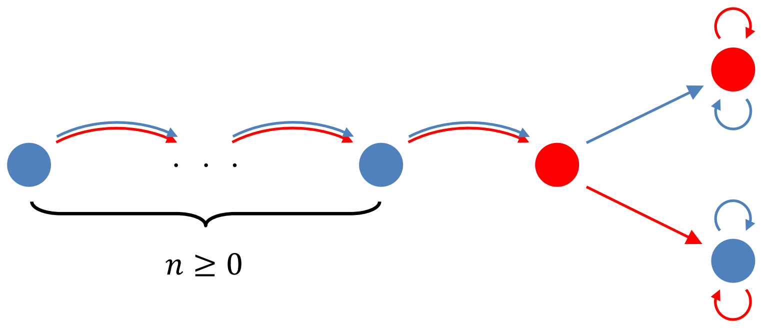
In region 3, 4 and 5, Grim Trigger (and TFT) are neutrally stable strategies

The strategy *Grim*, or *Grim Trigger*, never triggers itself, and therefore results in pairs of C's when playing against itself.

$$\begin{array}{lllllll} Grim & C & C & C & C & \ldots \\ Grim & C & C & C & C & \ldots \end{array}$$

Its payoff against itself therefore are $b - c$.

Consider strategy Y_n which is defined by the following automaton.



It is clear that, given a strategy Y that plays D for the first time in round $n + 1$, such a strategy cannot get a higher continuation payoff after round $n + 1$ against *Grim Trigger* than Y_n does, nor can it get higher continuation payoffs against itself, after round $n + 1$.

$$\begin{array}{llllll} Y_n & \overbrace{C \ldots C}^{n > 0} & D & C & C & C \ \ldots \\ Y_n & C \ldots C & D & C & C & C \ \ldots \end{array}$$

Thereby we have $\pi(Y_n, Y_n) = (1 - (1 - \delta)\delta^n)(b - c)$. Against *Grim Trigger* we see

$$\begin{array}{llll} Grim & \overbrace{C\ldots C}^{n\geq 0} & C & D\ D\ D\ .\ .\ .\ . \\ Y_n & C\ldots C & D & D\ D\ D\ .\ .\ .\ . \end{array}$$

Thereby we have $\pi(Grim, Y_n) = (1-\delta^n)(b-c) - (1-\delta)\delta^n c$ and $\pi(Y_n, Grim) = (1-\delta^n)(b-c) + (1-\delta)\delta^n b$.

Theorem 4 If $\alpha > \frac{c-\delta b}{b-2\delta b+\delta c}$ or $\delta > \frac{c}{b}$ then *Grim Trigger* is a neutrally stable strategy.

Proof Assume that *Y* is a strategy for which $\pi(Y,Y) < b-c$. Then there must be a number $n \geq 0$ such that against itself, *Y* first plays n times *C* and then *D* the $n+1$st time. If $\delta > \frac{c}{b}$, then not only $\pi(Y,Y) < b-c = \pi(Grim, Grim)$, but also $\pi(Y, Grim) \leq \pi(Y_n, Grim) = b-c+\delta^n c - \delta^{n+1} b < b-c = \pi(Grim, Grim)$. That implies that certainly for $0 \leq \alpha \leq 1$

$$\alpha\pi(Y,Y) + (1-\alpha)\pi(Y, Grim) \leq \pi(Grim, Grim)$$

If $\delta \leq \frac{c}{b}$ and $\alpha > \frac{c-\delta b}{b-2\delta b+\delta c}$, then

$$\alpha\pi(Y,Y) + (1-\alpha)\pi(Y, Grim) \leq$$

$$\alpha\pi(Y_n,Y_n) + (1-\alpha)\pi(Y_n, Grim) =$$

$$\alpha(1-(1-\delta)\delta^n)(b-c) + (1-\alpha)[(1-\delta^n)(b-c) + (1-\delta)\delta^n b] <$$

$$(b-c+\delta^n c - \delta^{n+1} b) + \alpha\delta^n(-b+2\delta b - \delta c) <$$

$$(b-c+\delta^n c - \delta^{n+1} b) + \frac{c-\delta b}{b-2\delta b+\delta c}\delta^n(-b+2\delta b-\delta c) =$$

$$(b - c + \delta^n c - \delta^{n+1} b) - \delta^n (c - \delta b) = b - c = \pi(Grim, Grim)$$

Now assume that $\pi(Y, Y) = b - c$. Then it must always play *C* against itself, and hence also against *Grim Trigger*. Thereby

$$\alpha\pi(Y, Y) + (1 - \alpha)\pi(Y, Grim) = b - c = \pi(Grim, Grim)$$

and also

$$(Y, Y) - \pi(Y, Grim) = 0 \leq 0 = \pi(Grim, Y) - \pi(Grim, Grim).$$

Because $(Y, Y) \leq b - c$, this completes the proof, as *Grim Trigger* is now shown to satisfy Definition 3 for all *Y*.

The same can be shown to hold for *TFT*.

In region 4 and 5, all equilibria are at least somewhat cooperative

The boundary between region 3 and 4 is given by $\alpha = \frac{c - \delta c}{b - \delta c}$. When we cross that boundary, from region 3 into region 4, *All D* ceases to be a neutrally stable strategy and even an equilibrium. The same is true for all fully defecting strategies; none of them is an equilibrium in regions 4 and 5. If *Z* is a strategy that always plays defect against itself, then there is at least one strategy that can directly invade it in region 4 and 5, and that strategy is *Grim Trigger*.

Z *D D D D* . . . .
Z *D D D D* . . . .

Grim *C C C C* . . . .
Grim *C C C C* . . . .

Grim	*C*	*D*	*D*	*D*	.	.
Z	*D*	?	?	?	.	.

The discounted, normalized payoff of *Z* playing against a copy of itself is $\pi(Z, Z) = 0$ and the discounted, normalized payoff of *Grim Trigger* playing against a copy of itself is $\pi(Grim, Grim\ ) = b - c$. The payoff of *Grim Trigger* against *Z* is at least $-(1 - \delta)c$. The payoff of *Z* against *Grim Trigger* is at most $(1 - \delta)b$.

The proof of the next theorem uses the fact that *Grim Trigger* can directly invade any *Z* with $\pi(Z, Z) = 0$ for $\alpha > \frac{c-\delta c}{b-\delta c}$.

Theorem 5 If $\alpha > \frac{c-\delta c}{b-\delta c}$, then no strategy *Z* with $\pi(Z, Z) = 0$ is an equilibrium.

Proof Assume that indeed $\alpha > \frac{c-\delta c}{b-\delta c}$ and $\pi(Z, Z) = 0$. If $\pi(Grim, Z) > 0$, then

$$\alpha\pi(Grim, Grim) + (1 - \alpha)\pi(Grim, Z) > 0 = \pi(Z, Z)$$

If $(Grim, Z) \leq 0$, then

$$\alpha\pi(Grim, Grim) + (1 - \alpha)\pi(Grim, Z) \geq$$

$$\alpha(b - c) - (1 - \alpha)(1 - \delta)c >$$

$$\frac{c - \delta c}{b - \delta c}(b - c) - \frac{b - c}{b - \delta c}(1 - \delta)c =$$

$$0 = \pi(Z, Z)$$

Thereby *Z* is not an equilibrium.

In region 5 all equilibria are fully cooperative

If a strategy is not fully cooperative, then *All C* can directly invade it. When *All C* plays against itself, it plays *C* every round.

$$\begin{array}{ll} \textit{All C} & C\,C\,C\,C\ \ .\ .\ .\ . \\ \textit{All C} & C\,C\,C\,C\ \ .\ .\ .\ . \end{array}$$

Therefore it gets payoff $b-c$ every round, which is also the normalized, discounted payoff of *All C* against itself.

When *Y* not always plays *C* against itself, and $\pi(Y,Y) < b-c$, then there is obviously at least one moment in time where it plays *D*. Assume that the first time it does play *D* is at time $n+1$, where $n \geq 0$, which implies that the first n times, Y plays *C* against itself. Note that n can be 0, so a strategy *Y* that starts defecting in round 1 is not excluded here.

$$\begin{array}{lllll} & \overbrace{}^{n\geq 0} & & & \\ Y & C\ldots C & D & ?\ \ ?\ \ ?\ \ .\ .\ .\ . \\ Y & C\ldots C & D & ?\ \ ?\ \ ?\ \ .\ .\ .\ . \end{array}$$

If all the question marks were *C*'s – when *Y* is playing against itself, they come in pairs of *C*'s or pairs of *D*'s – then it would earn a discounted, normalized payoff of $(1-\delta^n+\delta^{n+1})(b-c)$, so that puts a maximum on the payoff of *Y* playing against a copy of itself; $\pi(Y,Y) \leq (1-\delta^n+\delta^{n+1})(b-c)$.

$$\begin{array}{lllll} & \overbrace{}^{n\geq 0} & & & \\ Y & C\ldots C & D & ?\ \ ?\ \ ?\ \ .\ .\ .\ . \\ \textit{All C} & C\ldots C & C & C\ \ C\ \ C\ \ .\ .\ .\ . \end{array}$$

When *Y* and *All C* are playing each other, then the best case scenario for *Y*, and the worst case scenario for *All C*, is if all the question marks are *D*′s. This way *Y* also gets b in every period following period $n+1$, and *All C* loses c in all of them. This then gives the following upper and lower bound, respectively;

$$\pi(Y, All\ C) \leq (1-\delta^n)(b-c)+\delta^n b$$

$$\pi(All\ C, Y) \geq (1-\delta^n)(b-c)-\delta^n c.$$

With these bounds, we can show the following.

Theorem 6 If $\alpha > \frac{\delta(b-c)+c}{b}$, then no strategy *Y* with $\pi(Y,Y) < b-c$ is an equilibrium.

Proof Assume that indeed $\alpha > \frac{\delta(b-c)+c}{b}$ and $\pi(Y,Y) < b-c$. Then

$$\alpha\pi(AllC, AllC) + (1-\alpha)\pi(AllC, Y) \geq$$

$$\alpha(b-c) + (1-\alpha)[(1-\delta^n)(b-c)-\delta^n c] >$$

$$\frac{\delta(b-c)+c}{b}(b-c) + \frac{(1-\delta)(b-c)}{b}\big((1-\delta^n)(b-c)-\delta^n c\big) =$$

$$(b-c)(1-\delta^n+\delta^{n+1}) \geq \pi(Y,Y)$$

Hence *Y* is not an equilibrium.

In region 1a all fully defecting strategies are RAII. Everywhere else they are not.

To show that in the left / down part of region 1 all fully defecting strategies are RAII, we begin by computing bounds on the payoffs of all possible mutants after a sequence of neutral mutants. Starting from strategy $X = X_0$, which is fully defecting when it plays against itself, we assume a sequence of neutral mutants $X_1, \ldots, X_m$, for which it also must be the case that all of them always play *D* when they plays against themselves. Any non-neutral mutant X_{m+1} must however play *C* at least once. This gives us the following sequences and bounds on payoffs:

$$\begin{array}{lllllll} X_m & \overbrace{D\ldots D}^{n\geq 0} & D & ? & ? & ? & \ldots. \\ X_{m+1} & D\ldots D & C & ? & ? & ? & \ldots. \end{array}$$

The best case scenario for X_{m+1} would be if all the question marks for X_m were *C*'s and all question marks for X_{m+1} were *D*'s. Therefore it follows that $\pi(X_{m+1}, X_m) \leq \delta^n(-(1-\delta)c + \delta b)$.

When X_{m+1} plays against itself, we get the following:

$$\begin{array}{lllllll} X_{m+1} & \overbrace{D\ldots D}^{n>0} & C & ? & ? & ? & \ldots. \\ X_{m+1} & D\ldots D & C & ? & ? & ? & \ldots. \end{array}$$

The best case scenario for X_{m+1} would be if all the question marks were *C*'s, which implies that $\pi(X_{m+1}, X_{m+1}) \leq \delta^n(b - c)$.

Theorem 7 If $\alpha < \frac{c-\delta(c+b)}{b-\delta(c+b)}$ and $\delta < \frac{c}{c+b}$, then any strategy *X* with $\pi(X, X) = 0$ is robust against indirect invasions (RAII).

Proof Assume that indeed $\alpha < \frac{c-\delta(c+b)}{b-\delta(c+b)}$, $\delta < \frac{c}{c+b}$ and $\pi(X,X) = 0$.

Let $X_0 = X$ and let $X_1, \dots, X_m$ be a sequence of neutral mutants. This implies that $\pi(X_i, X_i) = 0$ also for $i = 1, \dots, m$. For any strategy X_{m+1} we now find that

$$\alpha\pi(X_{m+1}, X_{m+1}) + (1-\alpha)\pi(X_{m+1}, X_m) \leq$$

$$\alpha\delta^n(b-c) + (1-\alpha)\delta^n(-(1-\delta)c + \delta b) <$$

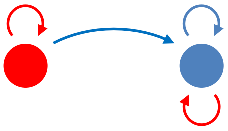
$$\frac{c-\delta(c+b)}{b-\delta(c+b)}(b-c) + \frac{b-c}{b-\delta(c+b)}(-(1-\delta)c + \delta b) =$$

$$0 = \pi(X_m, X_m)$$

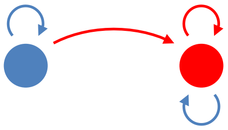
Hence an indirect invasion into X is not possible. The strict inequality is justified because $\delta < \frac{c}{c+b}$ implies that $-(1-\delta)c + \delta b < b - c$

The flip side of this result is that if $\alpha > \frac{c-\delta(c+b)}{b-\delta(c+b)}$, any strategy X with $\pi(X,X) = 0$ *can* in fact be invaded indirectly. This can be shown by constructing mutant strategies X_1 and X_2 for which those bounds are attained. Therefore we take the following two automata.

X_1:



X_2:



In region 4b and 5 all fully cooperative strategies are RAII. Everywhere else they are not.

To show that in region 5 and in the upper sliver of region 4 all fully cooperative strategies are indeed RAII, we begin by computing bounds on the payoffs of all possible mutants after a sequence of neutral mutants. Starting from strategy $Y = Y_0$, which is fully cooperative when it plays against itself, we assume a sequence of neutral mutants $Y_1, \ldots, Y_m$ for which it also must be the case that all of them always play *C* when they plays against themselves. Any non-neutral mutant Y_{m+1} must however play *D* at least once. This gives us the following sequences and bounds on payoffs:

$$\begin{array}{llll} Y_m & \overbrace{C\ldots C}^{n \geq 0} & C & ?\ \ ?\ \ ?\ \ldots\ldots \\ Y_{m+1} & C\ldots C & D & ?\ \ ?\ \ ?\ \ldots\ldots \end{array}$$

The best case scenario for Y_{m+1} would be if all the question marks for Y_m were *C*'s and all question marks for Y_{m+1} were *D*'s. Therefore it follows that $\pi(Y_{m+1}, Y_m) \leq b - (1 - \delta^n)c$

When Y_{m+1} plays against itself, we get the following:

$$\begin{array}{llll} Y_{m+1} & \overbrace{C\ldots C}^{n > 0} & D & ?\ \ ?\ \ ?\ \ldots\ldots \\ Y_{m+1} & C\ldots C & D & ?\ \ ?\ \ ?\ \ldots\ldots \end{array}$$

The best case scenario for Y_{m+1} would be if all the question marks were *C*'s, which implies that $\pi(Y_{m+1}, Y_{m+1}) \leq (1 - \delta^n + \delta^{n+1})(b - c)$.

Theorem 8 If $\alpha > \frac{c}{b-\delta(b-c)}$ then any strategy Y with $\pi(Y,Y) = b - c$ is robust against indirect invasions (RAII).

Proof Assume that indeed $\alpha > \frac{c}{b-\delta(b-c)}$ and $\pi(Y,Y) = b - c$. Let $Y_0 = Y$ and let $Y_1, \dots, Y_m$ be a sequence of neutral mutants. This implies that $\mu(Y_i, Y_i) = b - c$ also for $i = 1, \dots, m$. For any strategy Y_{m+1} we now find that

$$\alpha\pi(Y_{m+1}, Y_{m+1}) + (1-\alpha)\pi(Y_{m+1}, Y_m) \leq$$

$$\alpha(1 - \delta^n + \delta^{n+1})(b-c) + (1-\alpha)(b - (1-\delta^n)c) <$$

$$\frac{c}{b-\delta(b-c)}(1 - \delta^n + \delta^{n+1})(b-c) + \frac{(1-\delta)(b-c)}{b-\delta(b-c)}(b - (1-\delta^n)c) =$$

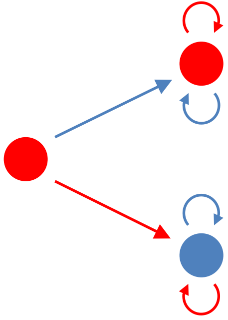
$$b - c = \pi(Y_m, Y_m)$$

Hence an indirect invasion into Y is not possible.

The flip side of this result is that if $< \frac{c}{b-\delta(b-c)}$, any strategy Y with $\pi(Y,Y) = b - c$ *can* in fact be invaded indirectly. This can be shown by constructing mutant strategies Y_1 and Y_2 for which those bounds are attained. Therefore we take the following two automata.

Y_1:

Y_2:



The figures below show the thresholds between the regions for b/c equal to 3, 2 and 1.5, respectively. Notice that at $b/c = 1.5$ a new region has appeared that does not even have equilibria.

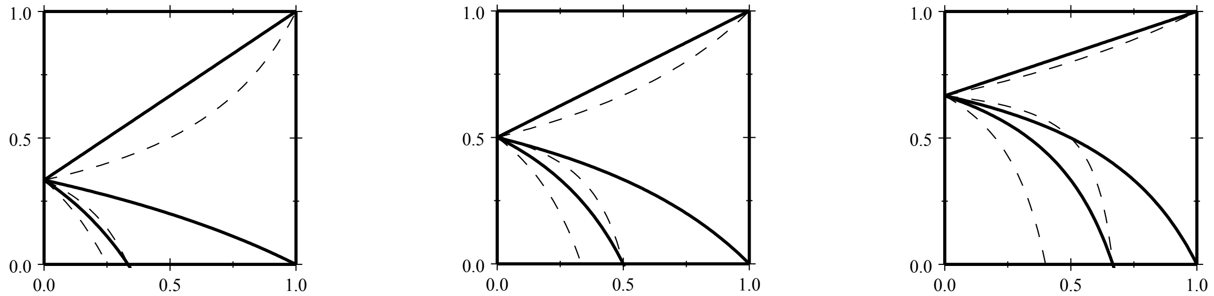


Fig. S7. The five regions for benefit-to-cost ratios of $3, 2$ and 1.5

These predictions can be seen as encompassing two classic results, one from biology, one from economics. The unrepeated, one-shot game has equal gains from switching, and the game is played between two players. Together this implies that for $\delta = 0$ there is a straightforward link to Hamilton's rule (26, 27, 4, 18). This link is visible on the vertical axis; all thresholds intersect with the vertical axis at $\alpha = c/b$. This intersection represents a steep transition; the prediction changes from full defection below the threshold to full cooperation above the threshold. If assortment parameter α is interpreted as relatedness, which is reasonable in this setting (1–5), then that gives us Hamilton's rule.

On the horizontal axis, the intersection with the first solid line represents the threshold for the continuation probability in the Folk Theorem for repeated games (28, 29). The Folk Theorem states that all levels of cooperation can be sustained as (subgame perfect) Nash equilibria for sufficiently large δ. The analysis here concerns neutrally stable strategies and strategies that are robust against indirect invasions. The sets of strategies that are neutrally stable, or robust against indirect invasions, are subsets of the set of Nash equilibria, but still cover all levels of cooperation.

5 Literature review of related papers on repetition and/or assortment

The classic paper by Axelrod & Hamilton (7) pointed out the importance of assortment as well as repetition. Their stability analysis however focused exclusively on repeated games in well-mixed populations, while the role of assortment as a way to get cooperation started was pointed out verbally. Also the conditions they used to determine whether strategies are evolutionarily stable define what we now would call neutral stability, or weak evolutionary stability (15, 19). The fact that neutral stability does not exclude neutral mutants, and the fact that the presence of neutral mutants would matter, was recognized in subsequent work (amongst others: 21–23, 30, 31). One paper shows that according to the standard definition of Maynard Smith & Price (16) no strategy is evolutionarily stable (21). A second uses another notion of evolutionary stability, and shows that also with that altered definition no ESS exists (22, see also 23). The dynamic implications of this alternative definition of evolutionary stability are explored in (25), where it is concluded that while the standard definitions of evolutionary stability (16, see also 15) and neutral stability (19, see also 15) maintain a solid link with dynamics, this is not the case with the alternative definition from (22). They also show that there is a multitude of strategies that are neutrally stable. In the current paper, we use the notion of a strategy that is *robust against indirect invasions* (RAII). This concept both acknowledges that there might be an important role to be played by neutral mutants, and it establishes a link with dynamics; the static notion of a strategy that is RAII links with stability in the replicator dynamics for sets of strategies in the same way as the static notion of an ESS links with stability in the replicator dynamics for single strategies (20). Other papers take a different approach and add stability of the solution to perturbations as a requirement for (evolutionary) stability (30, 31). By introducing a minimum probability of mistakes, and then letting that probability go to zero, (30) generates an evolutionary equivalent of Selten’s concept of a *trembling hand perfect equilibrium*. This is applied to repeated games in (31). None of these papers, however, combine repetition and population structure; the analysis in (7, 21–23, 30, 31) always assumes a well-mixed population.

Other, more recent papers that do combine games with (moderate levels of) assortment are (5) and (8–10). The first considers a game where all individuals play a sequence of one-shot games, being re-matched with new partners every round (5). In their setting there is a ‘global’ continuation probability, which is the probability with which everyone goes on to have another interaction (with a different partner). The paper looks at the evolution of generalized reciprocity, and considers three strategies: pure cooperators, pure defectors, and generalized reciprocators. Generalized reciprocators either help if they have been helped in the previous round, or they do not help if they did not receive help in the previous round. This strategy is similar to *TFT*, but also different, because in their setting partners change between rounds. They find that in order for generalized reciprocity to be an ESS, at least some assortment is needed.

The second considers a repeated N-player public goods game (8). This game is preceded by a communication stage, in which individuals signal their intent to punish defectors, and every repetition of the public goods game is followed by a punishment stage, in which individual defectors can be punished. They consider competition between two types – punishers and non-punishers – and find two equilibria: one with non-punishers only, and one with a mixture of the two types. They also find that assortment is not necessary for mixed equilibria to exist, but that assortment is needed to destabilize the all-defector equilibrium.

Our approach differs from those taken in both of these more recent studies. We consider direct reciprocity rather than generalized reciprocity, examining repeated interactions between the same pairs of players; we consider a repeated two-player Prisoner’s

Dilemma, rather than an N-player game followed by a set of pairwise punishment games; and we examine an open-ended, infinite strategy space, rather than considering the interaction of 2 or 3 specific strategies. For this rich strategy space we have analytical results, as well as simulations that match the predicted dynamics.

The game in (9) and (10) is similar to the game in the current paper, but also there attention is restricted to 2 and 3 strategies, respectively.

References

1. Eshel I, Cavalli-Sforza LL (1982) Assortment of encounters and evolution of cooperativeness *Proc Natl Acad Sci USA* 79:1331-1335.
2. Grafen A (1985) A geometric view of relatedness. *Oxford Surveys in evolutionary biology* 2:28-90.
3. Bergstrom T (2003) The algebra of assortative encounters and the evolution of cooperation. *Int Game Theory Rev* 5:211-228.
4. van Veelen M (2009) Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. *J Theor Biol* 259:589-600.
5. Rankin DJ, Taborsky M (2010) Assortment and the evolution of generalized reciprocity, *Evolution* 63:1913–1922.
6. Tarnita CE, Ohtsuki H, Antal T, Fu F, Nowak MA (2009) Strategy selection in structured populations, *J Theor Biol* 259:570-581.
7. Axelrod R, Hamilton WD (1981) The evolution of cooperation, *Science* 211:1390-1396.
8. Boyd R, Gintis H, Bowles S (2010) Coordinated punishment of defectors sustains cooperation and can proliferate when rare, *Science* 328:617-620.

9. Ohtsuki H, Nowak MA (2007). Direct reciprocity on graphs. *J Theor Biol* 247:462-470.

10. Tarnita CE, Wage N, Nowak MA (2011). Multiple strategies in structured populations. *Proc Natl Acad Sci USA* 108:2334-2337

11. Fletcher, JA, Doebeli, M (2009) A simple and general explanation for the evolution of altruism. Proc R Soc B 276:13-19

12. van Veelen M, García J (2010) In and out of equilibrium: evolution of strategies in repeated games with discounting, TI discussion paper 10-037/1.

13. Nash JF (1950) Equilibrium points in N-person games, *Proc Natl Acad Sci USA* 36:48–49.

14. Nash JF (1951) Non-cooperative Games, *Annals of Mathematics* 54:286–95.

15. Weibull JW (1995). *Evolutionary Game Theory*. (MIT Press, Cambridge MA)

16. Maynard Smith J, Price GR, (1973) The logic of animal conflict. *Nature* 246:15-18.

17. Taylor P, Jonker L (1978) Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40:145-156.

18. van Veelen M (2011) The replicator dynamics with n player games and population structure. *J Theor Biol* 276:78-85.

19. Maynard Smith J (1982) *Evolution and the theory of games*. (Cambridge University Press, Cambridge UK).

20. Van Veelen M (2012), Robustness against indirect invasions. *Games Econ Beh* 74:382-393.

21. Selten R, Hammerstein P (1984). Gaps in Harley's argument on evolutionarily stable learning rules and in the logic of "tit for tat", *Beh Brain Sci* 7:115-116

22. Boyd R, Lorberbaum JP (1987) No pure strategy is stable in the repeated prisoner's dilemma game, *Nature* 327: 58-59.

23. May R (1987) More evolution of cooperation, *Nature* 327:15-17.

24. Binmore KG, Samuelson L (1992) Evolutionary stability in repeated games played by finite automata, *J Econ Theory* 57:278-305.

25. Bendor J, Swistak P (1995) Types of evolutionary stability and the problem of cooperation. *Proc Natl Acad Sci USA* 92:3596-360.

26. Hamilton WD (1964) The genetical evolution of social behaviour, I & II. *J Theor Biol* 7:1–52.

27. van Veelen M (2007) Hamilton's missing link. *J Theor Biol* 246:551-554

28 Friedman J (1971) A noncooperative equilibrium for supergames. *Rev Econ Stud* 38:1-12.

29. Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information, *Econometrica* 54:533-554.

30. Selten R (1983) Evolutionary stability in extensive two-person games, *Math Soc Sci* 5: 269-363.

31. Kim Y-G (1994) Evolutionarily stable strategies in the repeated prisoner's dilemma, *Math Soc Sci* 28:167-197.