



In and out of equilibrium II: Evolution in repeated games with discounting and complexity costs



Matthijs van Veelen ^{a,b,*}, Julián García ^c

^a Department of Economics, Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, the Netherlands

^b Tinbergen Institute, the Netherlands

^c Faculty of Information Technology, Monash University, Melbourne, 3800 VIC, Australia

ARTICLE INFO

Article history:

Received 13 June 2018

Available online 4 March 2019

JEL classification:

C73

Keywords:

Repeated games

Complexity costs

Neutrally stable strategy

Robust against indirect invasions

Uniform invasion barrier

Wright-Fisher process

ABSTRACT

We explore evolutionary dynamics for repeated games with small, but positive complexity costs. We begin by extending a folk theorem result by Cooper (1996) to continuation probabilities, or discount rates, smaller than 1. Then we show that All D has a uniform invasion barrier. Since none of the more cooperative equilibria are robust against indirect invasions, we might expect not to observe any cooperative equilibria when complexity costs are positive. The average level of cooperation in the dynamics, however, can hover anywhere between no cooperation at all, and the average level of cooperation in the absence of complexity costs, depending on how small complexity costs are and how large the population is.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

In the evolutionary game theory literature on repeated games, most papers do not account for complexity costs. There are however a few notable exceptions. Binmore and Samuelson (1992) and Volij (2002) both formulate static equilibrium refinements that do account for complexity costs. In both refinements complexity costs enter lexicographically, but in different ways, and these different refinements therefore lead to different results. Binmore and Samuelson (1992) show that in order to satisfy their equilibrium requirement, an equilibrium must be efficient. Volij (2002) on the other hand shows that All D is the only strategy that satisfies his. A third result takes a middle position. In Cooper (1996) complexity costs do not enter lexicographically, but are subtracted from the payoffs. His result fills up the entire spectrum between full cooperation and no cooperation at all, by proving that one can get arbitrarily close to any level of cooperation with a neutrally stable strategy (NSS), provided that complexity costs are not too large.

In this paper we will also consider complexity costs, but here we will explicitly aim at understanding evolutionary dynamics in repeated games with complexity costs. Because the lexicographic treatment of complexity costs precludes a straightforward link between equilibrium refinement and stability in the replicator dynamics, we will follow Cooper (1996), and also subtract (possibly small) complexity costs from the payoffs. This allows us to consider stability concepts that do have dynamic stability properties, such as ESS, NSS, and robustness against indirect invasions (RAII, van Veelen, 2012).

* Corresponding author.

E-mail addresses: c.m.vanveelen@uva.nl (M. van Veelen), julian.garcia@monash.edu (J. García).

First we extend the folk theorem result by Cooper (1996) to continuation probabilities, or discount factors, smaller than 1. With this extension, the result no longer holds true for the entire interval between the symmetric Pareto efficient payoff and the one-shot Nash equilibrium payoff, but only for parts of it. The perimeters of the payoffs that can be supported by an NSS vary with the continuation probability and can be visualized in a tree-shaped figure.

Another result that will help understand the dynamics is that with positive complexity costs, All D is evolutionarily stable, and even has a uniform invasion barrier. Because none of the other equilibria from the extension of Cooper's (1996) result are RAI, and these other equilibria therefore can be left through indirect invasions, this implies that All D is more stable than all other equilibria.

Depending on the size of the complexity costs and the size of the population, properties of the dynamics are expected to hover somewhere between what one would expect from the equilibrium analysis without complexity costs (García and van Veelen, 2016), and the analysis with complexity costs. Without complexity costs, no equilibrium is RAI, which implies that we expect indirect invasions to make the population transition between different equilibria with different levels of cooperation, ranging from full cooperation to no cooperation at all. With positive complexity costs, All D is more stable than cooperative equilibria.

Simulations show that for a fixed complexity cost larger than 0, one can indeed get the average payoff ever closer to the payoff of mutual defection by increasing the population size. On the other hand, one can, for a fixed population size, get the average payoff ever closer to the average payoff in the absence of complexity costs, by letting complexity costs approach 0. It is important to realize that the average payoff without complexity costs is approached *from below* if complexity costs go to 0. That implies that (small) complexity costs do not have a beneficial effect on the average amount of cooperation in the dynamics, contrary to what one might be tempted to think is implied by results in Binmore and Samuelson (1992). Complexity costs decrease cooperation, and very low complexity costs decrease cooperation by a very small amount.

Section 2 contains preliminaries. Section 3 extends Cooper (1996) to continuation probabilities smaller than 1. Section 4 shows that All D has a uniform invasion barrier, both with and without discounting. In Section 5 we look at dynamics with *small* complexity costs. Here we also consider the role of “almost neutral” transitions, that make the dynamics for vanishing complexity costs resemble the dynamics in the absence of complexity costs. In Section 6 we link up with the results in Binmore and Samuelson (1992) and Volij (2002). Section 7 concludes the paper.

2. Preliminaries

2.1. Strategies

We start by repeating a few formal definitions from García and van Veelen (2016). Consider a symmetric one-shot 2-player game g characterized by a set of players $I = \{1, 2\}$, an action space A , equal for both players, and a payoff function $\pi : A \times A \rightarrow \mathbb{R}$. Using a discount factor $\delta \in (0, 1)$, interpreted as a continuation probability, this one-shot game is turned into a repeated one, which will be called $\Gamma(\delta)$. A history at time t is a list of the actions played up to and including time $t - 1$, where an empty pair of brackets is used to denote the history ‘no history’. If $a_{t,i}$ is the action played by player i at time t , then these histories are:

$$h_1 = ()$$

$$h_t = ((a_{1,1}, a_{1,2}), \dots, (a_{t-1,1}, a_{t-1,2})), \quad t = 2, 3, \dots$$

Sometimes we will also write $(h_t, (a_{t,1}, a_{t,2}))$ for a history h_{t+1} . The set of possible histories at time t is:

$$H_1 = \{h_1\}$$

$$H_t = \prod_{i=1}^{t-1} (A \times A) \quad t = 2, 3, \dots$$

and the set of all possible histories is:

$$H = \bigcup_{t=1}^{\infty} H_t.$$

A pure strategy is a function that maps histories to the action space; $S : H \rightarrow A$. For two strategies, say S and T , the course of actions is determined by recursion; all actions at all stages are determined by the initiation

$$h_1^{S,T} = (), h_1^{T,S} = ()$$

and the recursion step for $t = 1, 2, \dots$

$$a_t^{S,T} = \left(S \left(h_t^{S,T} \right), T \left(h_t^{T,S} \right) \right), a_t^{T,S} = \left(T \left(h_t^{T,S} \right), S \left(h_t^{S,T} \right) \right)$$

$$h_{t+1}^{S,T} = \left(h_t^{S,T}, a_t^{S,T} \right), h_{t+1}^{T,S} = \left(h_t^{T,S}, a_t^{T,S} \right)$$

For $0 \leq \delta < 1$ the discounted, normalized payoffs to (a player that uses) strategy S against strategy T is given by:

$$\Pi(S, T) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \pi \left(a_t^{S,T} \right)$$

For $\delta = 1$ we will consider the limit-of-means for the infinitely repeated game:

$$\Pi(S, T) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} \pi \left(a_t^{S,T} \right)$$

With strategies that can be represented by finite state automata, the limit-of-means always exists.

2.2. Finite state automata

Because we need a measure of complexity, we will restrict attention to strategies that can be represented by finite state automata. This is in line with the literature on repeated games with complexity costs; see Rubinstein (1986), Abreu and Rubinstein (1988), Kalai and Stanford (1988), Binmore and Samuelson (1992), Schlag (1993), Cooper (1996), and Volij (2002).

A finite state automaton, or a Moore machine, representing strategy S , is a tuple $\{\{1, \dots, N_S\}, \lambda_S, \mu_S\}$, where N_S is the number of states, $\lambda_S : \{1, \dots, N_S\} \rightarrow \{C, D\}$ gives the output in every state, and $\mu_S : \{1, \dots, N_S\} \times \{C, D\} \rightarrow \{1, \dots, N_S\}$ gives the transitions. Sometimes the formal definition of a machine also specifies in which state the machine starts, but because the states can always be renumbered so that the starting state is the first, we assume, without loss of generality, that the machine starts in state 1. Also without loss of generality, we assume that if it plays against a copy of itself, it will transition from state i to state $i + 1$, until for the first time it goes back to a state it has already been in, or remains in the state it is currently in. This also only reflects a choice how to order and name the states an automaton visits before it revisits a state, when matched with a copy of itself.

Not all strategies can be represented by a finite state automaton. Different finite state automata, on the other hand, can represent one and the same strategy. We always choose a finite state automaton with the smallest possible number of states to represent a strategy. This we can do without consequences for the results, because a finite state automaton for which there is a smaller version that represents the same strategy would never be a best response to any strategy. The smallest finite state automaton that represents a strategy S is moreover unique, up to the renumbering of states (Hopcroft et al., 2006). Since the numbering of the states does not affect payoffs in any interaction, nor the size of the automaton, we can, again, do this without loss of generality (for the same reason, we could actually have chosen any representation with the minimum number of states, even if the smallest representation had not been unique). In Appendix A we moreover define a natural distance between strategies, with which one can show that the set of finite state automata is dense in the set of all strategies. This implies that restricting attention to strategies that can be represented by a finite state automaton still leaves us with a large and representative set of strategies. Repeated games with the strategy space restricted to finite state automata are sometimes also called machine games.

As in Cooper (1996) we define the fitness – or profits – of the players as the payoffs from the game minus the costs:

$$U(S, T) = \Pi(S, T) - k|S|,$$

where $|S|$ is the number of states and $k > 0$ is the per state cost. The only difference with Cooper (1996) is that we also allow $\Pi(S, T)$ to be normalized, discounted payoffs instead of the limit-of-means.

2.3. Stability concepts

We will reproduce the stability concepts we use in order of increasing strictness. Let \mathcal{S} be the space of all strategies that can be represented by finite state automata.

Definition 1. $S \in \mathcal{S}$ is a neutrally stable strategy (NSS) if for all $T \in \mathcal{S}$

- 1) $U(S, S) \geq U(T, S)$ and
- 2) if $U(S, S) = U(T, S)$ then $U(S, T) \geq U(T, T)$

This is the pure strategy version, as used by Cooper (1996), of neutral stability. The first condition makes an NSS a symmetric Nash equilibrium. The original definition (Maynard Smith, 1982), see also Weibull (1995) allows both for S and

the mutants T to be mixed strategies. We will follow Cooper (1996), and the literature on repeated games with complexity costs in general, by focusing on pure equilibria. The results moreover are unaffected if we allowed for mutants T that are mixed (see also footnote 6 in Cooper, 1996). The same applies to the other stability concepts.

For a strategy to be robust against indirect invasions (van Veelen, 2012) it must not only be an NSS, but there must also not be a sequence of neutral mutants that open the door for each other, one after another, until some mutant strategy has an actual selective advantage. For the formal definition, we need to define, for a strategy S , the set of (evolutionary) worse, equal, and better performers against S .

$$S_W(S) = \{T \mid U(T, S) < U(S, S) \text{ or } (U(T, S) = U(S, S) \text{ and } U(T, T) < U(S, T))\}$$

$$S_E(S) = \{T \mid U(T, S) = U(S, S) \text{ and } U(T, T) = U(S, T)\}$$

$$S_B(S) = \{T \mid U(T, S) > U(S, S) \text{ or } (U(T, S) = U(S, S) \text{ and } U(T, T) > U(S, T))\}$$

Definition 2. $S \in \mathcal{S}$ is robust against indirect invasions (RAII) if

- 1) $S_B(S) = \emptyset$ and
- 2) $\nexists T^1, \dots, T^n, n \geq 2$, such that
$$\begin{cases} T^1 \in S_E(S) \\ T^i \in S_E(T^{i-1}), \quad 2 \leq i \leq n-1 \\ T^n \in S_B(T^{n-1}) \end{cases}$$

The first condition makes a strategy that is RAII an NSS. If a strategy is RAII, then it moreover gives us a set of strategies that is asymptotically stable in the replicator dynamics. van Veelen (2012) uses Theorem 3 in Balkenborg and Schlag (2001) to show that if S is RAII, then this strategy, together with its neutral mutants (but now including mixed strategies), and the neutral mutants of the neutral mutants, and so on, form a (minimal) ES-set. An ES-set is a setwise generalization of an ESS. An ESS is defined below, a definition of an ES-set can be found in Thomas (1985) or Weibull (1995). From Thomas (1985) we furthermore know that ES-sets are asymptotically stable in the replicator dynamics. The converse is also true; if X is an ES-set, and $S \in X$, then S is RAII.

The definition of an ESS is from Maynard Smith and Price (1973), and can also be found in Weibull (1995).

Definition 3. $S \in \mathcal{S}$ is an evolutionarily stable strategy (ESS) if for all $T \in \mathcal{S}$

- 1) $U(S, S) \geq U(T, S)$ and
- 2) if $U(S, S) = U(T, S)$ then $U(S, T) > U(T, T)$ for $T \neq S$

If S is an ESS, then $S_B(S) = \emptyset$ and $S_E(S) = \{S\}$, which implies that an ESS is automatically also RAII. We will call $\bar{\epsilon}_T(S) \in (0, 1]$ an invasion barrier for S against T if

$$(1 - \epsilon)U(S, S) + \epsilon U(S, T) > (1 - \epsilon)U(T, S) + \epsilon U(T, T)$$

holds for every $\epsilon \in (0, \bar{\epsilon}_T(S))$. If S is an ESS, then it has an invasion barrier against every strategy $T \neq S$; if $U(S, S) \geq U(T, S)$ and $U(S, T) > U(T, T)$, then $\bar{\epsilon}_T(S) = 1$ is an invasion barrier against T , and if $U(S, S) > U(T, S)$ and $U(S, T) < U(T, T)$, then $\bar{\epsilon}_T(S) = \frac{U(S, S) - U(T, S)}{U(S, S) - U(T, S) + U(T, T) - U(S, T)}$ is an invasion barrier against T .

Definition 4. $S \in \mathcal{S}$ has a uniform invasion barrier (UIB) if there is some $\bar{\epsilon}(S) \in (0, 1)$, such that for all $T \in \mathcal{S}, T \neq S$ and every $\epsilon \in (0, \bar{\epsilon}(S))$

$$(1 - \epsilon)U(S, S) + \epsilon U(S, T) > (1 - \epsilon)U(T, S) + \epsilon U(T, T)$$

With finitely many (pure) strategies, being ESS and having a UIB is the same, but with infinitely many strategies, an ESS might not have a UIB (see Bomze and Pötscher, 1989, Weibull, 1995, and van Veelen and Spreij, 2009).

2.4. The stage game

We will consider prisoners dilemmas with “equal gains from switching”, which implies that the payoffs can be defined with only two parameters.

$$\begin{bmatrix} & C & D \\ C & b & 0 \\ D & b+c & c \end{bmatrix}$$

The b can be interpreted as the benefits conferred to the other player by cooperating (the left column equals the right column plus b), while c is the cost of cooperating (the top row equals the bottom row minus c)

When describing a prisoners dilemma with equal gains from switching, the following matrix is also regularly used:

$$\begin{bmatrix} & C & D \\ C & b-c & -c \\ D & b & 0 \end{bmatrix}$$

The equilibrium analysis remains the same if a constant is subtracted from all entries, so these two matrices are equivalent. In the simulations probabilities are computed, which requires payoffs to be nonnegative. For consistency with the simulations we chose to use the matrix with nonnegative payoffs.

The results that follow can also be formulated and proven for prisoners dilemmas with unequal gains from switching, but this captures the essence and keeps the notation simple.

3. A “tree” full of NSS’es

Here we extend the result by Cooper (1996) to continuation probabilities smaller than 1. Theorem 7 gives perimeters of the combinations of payoffs and continuation probabilities that allow for NSS’es, provided that the complexity costs are not too large. For the proof, it will be useful to have a value function for the states of strategy S from the point of view of its opponent. Given a continuation probability, or discount factor $\delta \in (0, 1)$, the value to the opponent of strategy S of S being in state i is denoted by $V_S^*(i)$, $i = 1, \dots, N_S$, and is the solution to the following system of equations.

$$V_S(i) = \max_{a \in \{C, D\}} \{ \pi(a, \lambda_S(i)) + \delta V_S(\mu_S(i, a)) \} \quad i = 1, \dots, N_S$$

In the absence of complexity costs, strategy S is a best response to itself if $\lambda_S(i) \in \arg \max_{a \in \{C, D\}} \{ \pi(a, \lambda_S(i)) + \delta V_S^*(\mu_S(i, a)) \}$ for all states that are visited if S plays against a copy of itself. If S is indeed a best response to itself, then the value in the first state is the not normalized, discounted payoff of S playing against itself; $(1 - \delta) V_S^*(1) = \Pi(S, S)$.

If we include complexity costs, and if a strategy S is to be an NSS, or even an equilibrium, then that immediately puts some restrictions on the shape of S . As noted in the literature, from the definition of fitness function U it directly follows that an equilibrium cannot have unused states, when it plays against a copy of itself. If T is the same as S , but without the unused states, then $\Pi(T, S) = \Pi(S, S)$, while $|T| < |S|$, and therefore $U(T, S) > U(S, S)$. The absence of unused states in equilibrium, together with the way our definition of a finite state automaton orders the states, implies that S , when playing against itself, passes through states 1 to N_S , and then moves to one of the N_S states it has already been in before, or is currently in, and starts cycling.

If S is a best response against itself without complexity costs, it can cease to be a best response when complexity costs are positive. If T has fewer states than S , it is possible that $\Pi(T, S) < \Pi(S, S)$, but $U(T, S) > U(S, S)$. Although this is possible for given complexity costs $k > 0$, one can always choose k sufficiently low for the comparison of fitnesses – or profits – to align with a strict inequality in payoffs. Therefore, if S is the smallest best response against itself without complexity costs, there will be a k' such that S is also a best response against itself for all $0 \leq k < k'$.

If S is not a best response against itself without complexity costs, it is also not a best response against itself with complexity costs. The reason is that one can always construct a best response against S with a finite state automaton that has the same number of states as S , or fewer, since an opponent only needs to have one action for every state $i = 1, \dots, N_S$ that S can be in (see also Section 5.2 in García and van Veelen, 2018). This implies that if S is a best response against itself with complexity costs, it must also be a best response against itself without complexity costs.

The next lemma shows that if S is an equilibrium with complexity costs, the value to the opponent of strategy S of S being in state 1 must be lower than or equal to the value for any other state. In the proof, we first use the last implication above, and later we use the fact that, with positive complexity costs, there are no unused states in equilibrium, and when playing against itself, an equilibrium strategy therefore begins with visiting all of its states in succession.

Lemma 5. *If S is an equilibrium for $k > 0$, then $V_S^*(1) \leq V_S^*(i)$, $i = 1, \dots, N_S$.*

Proof. We begin with two simple observations. If $\lambda_S(i) = C$, and S is an equilibrium, then $V_S^*(i) = b + \delta V_S^*(\mu_S(i, C))$. This implies that $\frac{b}{1-\delta} - V_S^*(i) = \delta \left(\frac{b}{1-\delta} - V_S^*(\mu_S(i, C)) \right)$, and therefore that $V_S^*(\mu_S(i, C)) \leq V_S^*(i)$. If S always plays C against itself after visiting state i , then $V_S^*(\mu_S(i, C)) = \frac{b}{1-\delta} = V_S^*(i)$. If S does not always play C against itself after visiting state i , both $V_S^*(i) < \frac{b}{1-\delta}$ and $V_S^*(\mu_S(i, C)) < \frac{b}{1-\delta}$, and the inequality is strict.

If $\lambda_S(i) = D$, and S is an equilibrium, then $V_S^*(i) = c + \delta V_S^*(\mu_S(i, D))$. This implies that $V_S^*(i) - \frac{c}{1-\delta} = \delta (V_S^*(\mu_S(i, D)) - \frac{c}{1-\delta})$, and therefore that $V_S^*(\mu_S(i, D)) \geq V_S^*(i)$. If S always plays D against itself after visiting state i , then $V_S^*(\mu_S(i, D)) = \frac{c}{1-\delta} = V_S^*(i)$. If S does not always play D against itself after visiting state i , both $V_S^*(i) > \frac{c}{1-\delta}$ and $V_S^*(\mu_S(i, D)) > \frac{c}{1-\delta}$, and

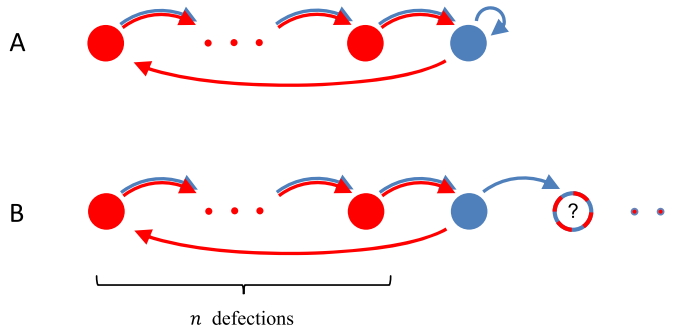


Fig. 1. Strategies that start with n defections (red) before they play cooperate (blue) for the first time, when they play against a copy of themselves, either have $n + 1$ states, or more than $n + 1$ states. Strategy A would attain the upper bound from Theorem 7 after n defections. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

the inequality is strict. Summarizing, if S is an equilibrium, then after a state in which C is played, the value does not go up, and after a state in which D is played, the value cannot go down.

Now let $\underline{V}_S = \min_{j \in \{1, \dots, N_S\}} V_S^*(j)$, and let state m be the first state for which this minimum is attained. If $m > 1$, then it follows from the second observation that the action in the preceding state cannot be D , and therefore that m is preceded by a state in which C is played. But if S is to be an equilibrium, then $b + \delta V_S^*(m) > c + b + \delta V_S^*(\mu_S(m - 1, D))$, which is impossible because $V_S^*(m) = \underline{V}_S \leq V_S^*(i)$ for all $i = 1, \dots, N_S$. \square

Theorem 3.2 in Schlag (1993) shows that if S is an equilibrium, it must defect in the first state. Our Lemma allows for an alternative proof.

Corollary 6. *If S is an equilibrium for $k > 0$, then $\lambda_S(1) = D$.*

Proof. If the output in every state is C , then the value in all states is $\frac{b}{1-\delta}$. Because this strategy always plays C , however, it can not be an equilibrium. If the output in every state is D , then the value in all states is $\frac{c}{1-\delta}$, and, obviously, $\lambda_S(1) = D$. If the output is not the same in every state, and $\lambda_S(1) = C$, then the first observation in the proof of Lemma 5 implies that $V_S^*(1) > V_S^*(2)$. But this is inconsistent with Lemma 5, according to which $V_S^*(1) \leq V_S^*(i)$, $i = 1, \dots, N_S$. Therefore $\lambda_S(1) = D$. \square

With all equilibria starting with at least one defection, we will go on to look at strategies that, when playing against themselves, start with $n > 0$ defections before they play cooperate for the first time. Theorem 7 gives the perimeters of the payoffs of equilibrium strategies for every $n > 0$. Together with the 1-state automaton All D, that covers all possible NSS'es (it is clear that if a strategy always plays D against itself, it cannot be an NSS if it has more than 1 state). Furthermore, no strategy other than All D is an equilibrium for all $k > 0$; even if a strategy is an equilibrium for some $k > 0$, there is always some threshold above which complexity costs are too high for it to be an equilibrium.

Theorem 7. *If strategy S is an equilibrium for $k > 0$, and if, for $n \geq 1$*

1. $\lambda_S(i) = D$ for $1 \leq i \leq n$ and $\lambda_S(n + 1) = C$
2. $\mu_S(i, D) = i + 1$ if $1 \leq i \leq n$

then

$$c + \frac{(1 - \delta) \delta^n}{1 - \delta^{n+1}} b \leq \Pi(S, S) \leq c + \delta^n (b - c)$$

Proof. If strategy S is an equilibrium for $k > 0$, then the following two must hold.

- (i) $V_S^*(1) = \sum_{i=1}^n \delta^{i-1} c + \delta^n b + \delta^{n+1} V_S^*(\mu_S(n + 1, C))$
- (ii) $b + \delta V_S^*(\mu_S(n + 1, C)) \geq b + c + \delta V_S^*(\mu_S(n + 1, D))$

Lemma 5 implies that $V_S^*(\mu_S(n + 1, D)) \geq V_S^*(1)$. We can therefore replace the former with the latter in (ii)

$$\delta V_S^*(\mu_S(n + 1, C)) \geq c + \delta V_S^*(1)$$

Using equation (i), we get

$$\begin{aligned}
 V_S^*(1) &\geq \frac{1-\delta^n}{1-\delta}c + \delta^n b + \delta^n [c + \delta V_S^*(1)] \\
 (1-\delta^{n+1}) V_S^*(1) &\geq \frac{1-\delta^{n+1}}{1-\delta}c + \delta^n b \\
 (1-\delta) V_S^*(1) &\geq c + \frac{(1-\delta)\delta^n}{1-\delta^{n+1}}b
 \end{aligned}$$

This gives the lower bound. The upper bound follows immediately from the maximum discounted, normalized payoff that is still attainable after n mutual defections (as for instance by strategy A in Fig. 1). □

Theorem 7 naturally implies that if $b - c < \frac{1-\delta}{1-\delta^{n+1}}b$, then the lower bound on the payoffs that can be sustained in equilibrium exceeds the upper bound, and therefore there are no equilibria that, when playing against themselves, start with exactly $n > 0$ defections before they play cooperate for the first time. As a consequence, for $\delta < \frac{c}{b}$ there are no equilibria at all other than All D; if $\delta < \frac{c}{b}$ then $\frac{1-\delta}{1-\delta^{n+1}}b > (1-\delta)b > (1-\frac{c}{b})b = b - c$ for all $n > 0$.

On the other hand, if $b - c > \frac{1-\delta}{1-\delta^{n+1}}b$, then for sufficiently low costs of complexity there is an NSS that attains the upper bound. This strategy is pictured as strategy A in Fig. 1, and the proof below follows the reasons given in Cooper (1996) for the case where $\delta = 1$.

Proposition 8. Let strategy S be defined by $N_S = n + 1$; $\lambda_S(i) = D$ if $1 \leq i \leq n$ and $\lambda_S(n + 1) = C$; $\mu_S(i, C) = \mu_S(i, D) = i + 1$ if $1 \leq i \leq n$, $\mu_S(n + 1, D) = 1$ and $\mu_S(n + 1, C) = n + 1$. Then S is an NSS if $b - c > \frac{1-\delta}{1-\delta^{n+1}}b$ and $k < \frac{\delta^n}{n}(\frac{\delta-\delta^{n+1}}{1-\delta^{n+1}}b - c)$.

Proof. All strategies T that do not play against S the way S does – that is, they do not start by playing exactly n defections before they play cooperate for the first time and ever after – earn a lower payoff against S than S does against itself; playing C when S is in state 1 to n cannot be optimal, and $b - c > \frac{1-\delta}{1-\delta^{n+1}}b$ implies that playing C when S is in state $n + 1$ gives the highest continuation payoff. Note that playing D also when S is in state $n + 1$ would make this AllD, which we consider in its 1-state implementation. While $b - c > \frac{1-\delta}{1-\delta^{n+1}}b$ implies that $\Pi(AllD, S) = c + \delta^n b \frac{1-\delta}{1-\delta^{n+1}} < c + \delta^n (b - c) = \Pi(S, S)$, adding the condition that $k < \frac{\delta^n}{n}(\frac{\delta-\delta^{n+1}}{1-\delta^{n+1}}b - c)$ implies that also $U(AllD, S) = \Pi(AllD, S) - k < \Pi(S, S) - (n + 1)k = U(S, S)$.

Having more than $n + 1$ states is superfluous, so the only strategies that remain are strategies T that also have $n + 1$ states and that, when playing against themselves, do start with exactly $n > 0$ defections before they play cooperate for the first time and ever after. For those strategies $U(T, S) = U(S, S) = U(S, T) = U(T, T)$. □

At the threshold, where $b - c = \frac{1-\delta}{1-\delta^{n+1}}b$, this strategy is not an equilibrium, because then $\Pi(AllD, S) = \Pi(S, S)$ while $|S| > |AllD|$.

The equilibria and the perimeters are summarized for an example in Fig. 2. For $b = 2$ and $c = 1$ we find that then there are no equilibria with $n = 1$ (the lower bound from Theorem 7 exceeds the upper bound for all $\delta < 1$). However, for every $n > 1$ we do get a new branch on the “tree” in the picture, where a branch is the space between the upper and the lower bound.

A few more details are worth mentioning. The first is that within such a branch, not the entire area between these bounds is filled with equilibria. Theorem 7 gives upper and lower bounds on equilibrium payoffs, given an initial sequence of n D’s and one C, when playing against a copy of itself. The upper bounds moreover are attained by strategy A in Fig. 1. If we specify the possible actions beyond these initial sequences further, we can find upper and lower bounds within the same branch. If we were to go beyond the moves up to and including the first C, when playing against a copy of itself, we can for instance distinguish two possibilities; after the first C, the strategy either 1) cooperates at least once, or it 2) defects at least once. By looking at the restrictions on the value functions in the same way as we did in the proof of Theorem 7, we find that in case 1) the normalized, discounted payoff is at least $\frac{c(1-\delta^n + \delta^{n+1} - \delta^{n+2}) + b(\delta^n - \delta^{n+2})}{1-\delta^{n+2}}$, while the upper bound is the same as it was in Theorem 7. The lower bound is found by first assuming there are $n + 2$ or more states, and combining $V^*(1) = \frac{1-\delta^n}{1-\delta}c + \delta^n b + \delta^{n+1}b + \delta^{n+2}V^*(\mu_S(n + 2, C))$ and $b + \delta V^*(\mu_S(n + 2, C)) \geq b + c + \delta V^*(\mu(n + 2, D)) \geq b + c + \delta V^*(1)$. If there are $n + 1$ states, then this must be strategy A from Fig. 1, and the upper bound from Theorem 7 is attained.

For case 2) we find that the normalized, discounted payoff is less than or equal to $c(1 - \delta^n + \delta^{n+1} - \delta^{n+2}) + b(\delta^n - \delta^{n+1} + \delta^{n+2})$, while the lower bound is still the same as it was in Theorem 7. This upper bound can be attained by a strategy that, when playing against a copy of itself, goes through the initial sequence of $n + 2$ states as specified, and then has a state $n + 3$ in which this strategy plays C, remains in that state if its opponent plays C, and returns to state 1 if its opponent plays D.

What these additional bounds amount to is pictured in Fig. 3. By specifying the possible initial sequence further, one can find ever more such bounds. The upper bounds give ever more payoffs that are attained, and the lower bounds exclude ever more payoffs, thereby carving out ever more areas within a branch without equilibria (see also footnote 2).

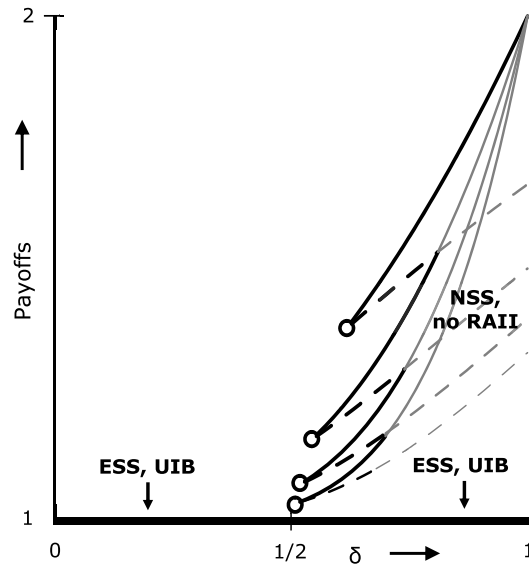


Fig. 2. Equilibria for $b = 2$, $c = 1$ and $k > 0$. The solid lines represent upper bounds, and the dotted lines represent lower bounds on the payoffs for equilibrium strategies starting with 2, 3, 4, and 5 defections (see Theorem 7). The upper bounds are attained by strategies that then go on to cooperate forever after, when matched with a copy of themselves. For $\frac{b}{c} = 2$, there are no equilibria that start with 1 defection before switching to cooperation. Many of the equilibrium strategies between the bounds are also NSS, although not every point there can be approximated arbitrarily closely with an equilibrium strategy. Also the complexity costs have to be sufficiently close to 0 to make a strategy an NSS. None of these NSS'es however are RAI. All D on the other hand is an ESS for all continuation probabilities and positive complexity costs, and also has a uniform invasion barrier (see Theorem 9 in Section 4).

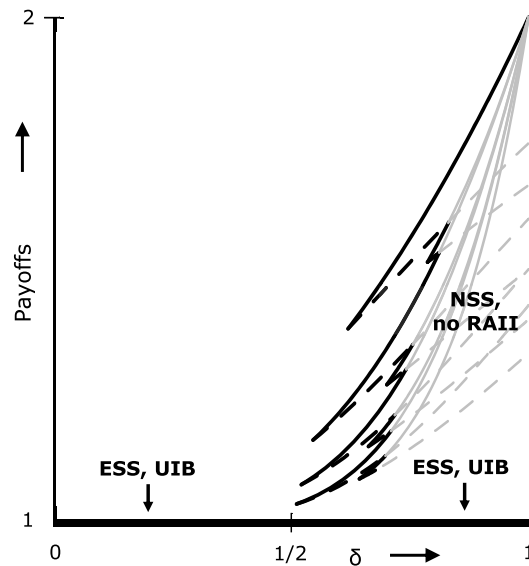


Fig. 3. Upper and lower bounds on equilibrium payoffs for strategies, the initial sequence of which is specified for one extra round beyond the initial sequences in Fig. 2 and Theorem 7. The tips of all branches and sub-branches again are not equilibria. This is not indicated in the picture by lack of space.

A second detail that is worth mentioning is that for continuation probabilities lower than $\frac{1}{2}$, not all normalized, discounted payoffs between c and b are feasible, even if we disregard equilibrium considerations. That can easily be seen by looking at an example; if $\delta = \frac{1}{3}$ then the only payoffs that remain are those that can be written as $c + x(b - c)$, where x must be an element of the Cantor set.

The NSS'es however are not RAI. That is easily seen if we realize that neutral mutants can be constructed without increasing the size of the automaton. If strategy S is an equilibrium with $\Pi(S, S) > c$, we can change the transitions in the following way: define T as equal to S , except that $\mu_T(i, D) = \mu_S(i, C)$ for all i . This is a neutral mutant – $U(T, S) = U(S, S) = U(S, T) = U(T, T)$ – and it is easily seen that $\Pi(AllD, T) > \Pi(T, T)$, and therefore a sequence of first T and then $AllD$ constitutes an indirect invasion. This observation follows the same logic as the proof of the second half of Claim 1 in

Volij (2002). Also here, the fact that no strategy S with $\Pi(S, S) > c$ is RAI implies that there is no ES-set other than {All D} (van Veelen, 2012).

4. All D has a uniform invasion barrier

With complexity costs $k > 0$, one can show that All D, in its simple, 1-state implementation, is an ESS for all $\delta \in [0, 1]$ and even has a uniform invasion barrier.

Theorem 9. *The strategy All D in its 1-state implementation has a uniform invasion barrier for all $\delta \in [0, 1]$ and costs $k > 0$.*

Proof. All other strategies S either also always play D against copies of themselves, or they do not. If they do, the matrix of fitnesses, or profits, is:

$$\begin{array}{cc} & \begin{array}{c} AllD \\ S \end{array} \\ \begin{array}{c} AllD \\ S \end{array} & \begin{array}{cc} c - k & c - k \\ c - N_S k & c - N_S k \end{array} \end{array}$$

If S is not also All D in its 1-state implementation, then $N_S > 1$, and

$$(1 - \epsilon)U(AllD, AllD) + \epsilon U(AllD, S) = c - k > c - kN_S = (1 - \epsilon)U(S, AllD) + \epsilon U(S, S)$$

for all $\epsilon \in (0, 1)$. All D's invasion barrier against S therefore is $\bar{\epsilon}_S(AllD) = 1$.

If S does not always play D against itself, then there is a state $n \geq 1$ in which it plays C for the first time, if it plays against itself or against All D. That implies that against itself, it attains at most a normalized, discounted payoff of $c + \delta^{n-1}(b - c)$. Against All D, S attains at most $c - (1 - \delta)\delta^{n-1}c$, while All D against S attains at least $c + (1 - \delta)\delta^{n-1}b$. The matrix of fitnesses, or profits, now contains lower and upper bounds.

$$\begin{array}{cc} & \begin{array}{c} AllD \\ S \end{array} \\ \begin{array}{c} AllD \\ S \end{array} & \begin{array}{cc} c - k & \geq c + (1 - \delta)\delta^{n-1}b - k \\ \leq c - (1 - \delta)\delta^{n-1}c - N_S k & \leq c + \delta^{n-1}(b - c) - N_S k \end{array} \end{array}$$

If $c + (1 - \delta)\delta^{n-1}b - k \geq c + \delta^{n-1}(b - c) - N_S k$, then both $U(AllD, AllD) > U(S, AllD)$ and $U(AllD, S) > U(S, S)$, and All D's invasion barrier against S is 1 again.

If $c + (1 - \delta)\delta^{n-1}b - k < c + \delta^{n-1}(b - c) - N_S k$, then $U(AllD, S)$ may be larger or smaller than $U(S, S)$, but for sure $\bar{\epsilon}_S(AllD) = \min\{1, \frac{(\frac{1}{\delta}-1)c + \frac{(N_S-1)k}{\delta^n}}{b-c}\}$ will be an invasion barrier. This can be seen by observing that

$$(1 - \epsilon)U(AllD, AllD) + \epsilon U(AllD, S) \geq c - k + \epsilon(1 - \delta)\delta^{n-1}b$$

and

$$(1 - \epsilon)U(S, AllD) + \epsilon U(S, S) \leq c - N_S k - (1 - \epsilon)(1 - \delta)\delta^{n-1}c + \epsilon\delta^{n-1}(b - c)$$

which makes

$$(1 - \epsilon)U(AllD, AllD) + \epsilon U(AllD, S) - (1 - \epsilon)U(S, AllD) + \epsilon U(S, S) \geq (N_S - 1)k + (1 - \delta)\delta^{n-1}c - \epsilon\delta^n(b - c)$$

The right hand side is larger than 0 if $\epsilon < \frac{(\frac{1}{\delta}-1)c + \frac{(N_S-1)k}{\delta^n}}{b-c}$, which makes $\bar{\epsilon}_S(AllD)$ an invasion barrier of All D against S indeed.

If S is Tit-for-tat, then the lower and upper bounds on the payoffs are met exactly, while also $n = 1$ and $N_S = 2$. This makes Tit-for-tat the strategy with the lowest maximum invasion barrier, and therefore $\bar{\epsilon}_{TF}(AllD) = \min\{1, \frac{(1-\delta)c+k}{\delta(b-c)}\}$ is also a uniform invasion barrier $\bar{\epsilon}(AllD)$ for All D. □

Please note that for $\delta < 1$ this uniform invasion barrier is bounded away from 0; the invasion barrier remains positive, also in the limit where k decreases to 0. Only for $\delta = 1$ does the bound on the uniform invasion barrier tend to 0 in the limit of $k \rightarrow 0$ (see also Theorems 3.4 and 3.5 in Schlag (1993) for similar observations).

If a strategy has a uniform invasion barrier, then it is an ESS (Weibull, 1995), and a strategy that is an ESS is also RAI (van Veelen, 2012, see also Section 2).

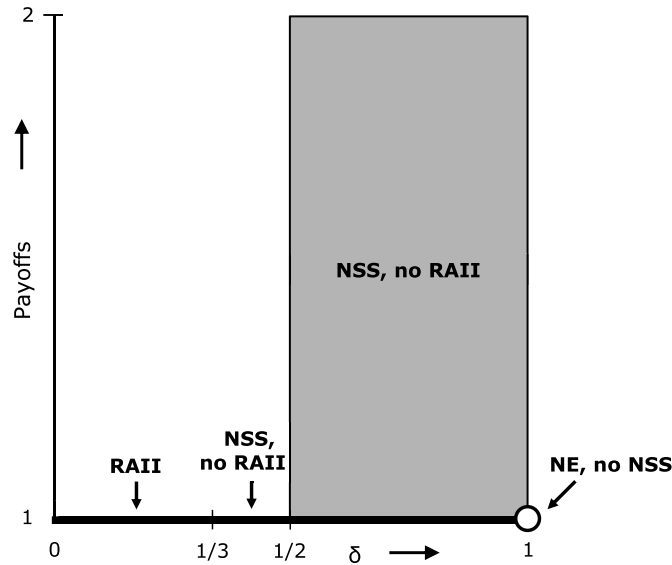


Fig. 4. Equilibrium properties for $b = 2$, $c = 1$, and $k = 0$. See also van Veelen et al. (2012) for the equilibrium analysis left of $\delta = 1/2$.

5. What if complexity costs are small?

Theorem 9 states that with complexity costs, All D in its simplest, one-state form, has a uniform invasion barrier. In Section 3 we have also seen that there is no strategy other than All D that is RAI, which means that there is also no ES-set – other than {All D}. This suggests that if complexity costs are positive, All D is more stable than any other equilibrium.

In the absence of complexity costs, on the other hand, we know that with sufficiently high continuation probability no strategy is RAI (García and van Veelen, 2016), while NSS'es exist with levels of cooperation ranging from full defection to full cooperation (Bendor and Swistak, 1995). This suggests that without complexity costs, and for δ sufficiently large, but smaller than 1, All D is neither more nor less stable than any other NSS.

Many results in the literature focus on *small* complexity costs. The results in Sections 3 and 4, using static equilibrium concepts, suggest that if we take the limit of complexity costs k decreasing to 0, there is a certain discontinuity; for all $k > 0$, All D is more stable than all other NSS'es, while at $k = 0$ it is not. More precisely, the lower bound for the invasion barrier from the proof of Theorem 9 implies that for $k \downarrow 0$ the uniform invasion barrier for All D is bounded away from 0, if $\delta < 1$, and yet at $k = 0$, there is no uniform invasion barrier at all, and All D is not even RAI. Moreover, there are windows of payoffs that cannot be supported by an NSS for any $k > 0$ – this is the space between the branches in Fig. 2 and Fig. 3 – while at $k = 0$ all symmetric payoffs can be approximated arbitrarily closely – see Fig. 4 – provided that δ is sufficiently high.

If we look at simulation result for stochastic finite population dynamics, however, we find no such discontinuity. In the remainder of this section we will therefore first describe the simulations. They suggest that for a fixed population size, a vanishing k will make the average payoff converge to the average payoff in the absence of complexity costs. For a fixed k , on the other hand, the simulations suggest that ever larger population sizes make the average payoff converge to the payoff of mutual defection. After describing the simulation results, we will discuss the effect of complexity costs on different transitions in the dynamics in order to provide an intuition for why the properties of these finite population dynamics hover between the prediction for $k = 0$ and the prediction for $k > 0$ the way they do.

5.1. The simulations

The simulations here are a variant of those reported in García and van Veelen (2016) for the case without complexity costs. In those simulations, a population state consists of N individuals that are all endowed with finite state automata. These individuals are randomly matched in pairs to play a repeated prisoner's dilemma. The payoffs from those encounters are used in the update step, in which all individuals in the new generation are drawn, one by one, and independently, from a distribution where the probability of being the offspring of individual j from the old generation is proportional to the payoff of j . After the new generation has been drawn, any individual mutates with a small probability. This completes the cycle for a generation, and this cycle is then repeated a large number of times. The update procedure makes this a Wright-Fisher process.

With complexity costs, we would like to also explore what happens with large populations. Running these simulations for large population sizes, however, is prohibitively time-consuming. Therefore we have chosen to assume that mutation rates are sufficiently low for mutants to almost always go extinct, or go to fixation, before the next mutant arrives. This allows us

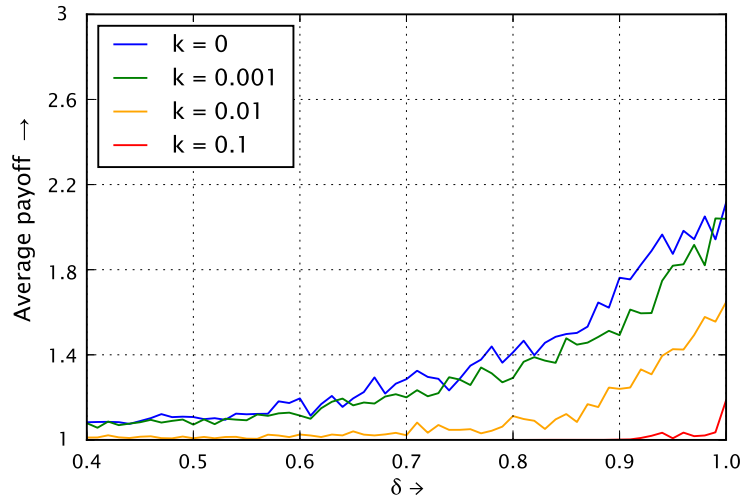


Fig. 5. Average payoffs with different complexity costs for $b = 3$, $c = 1$, and population size 250. Payoffs do not include complexity costs.

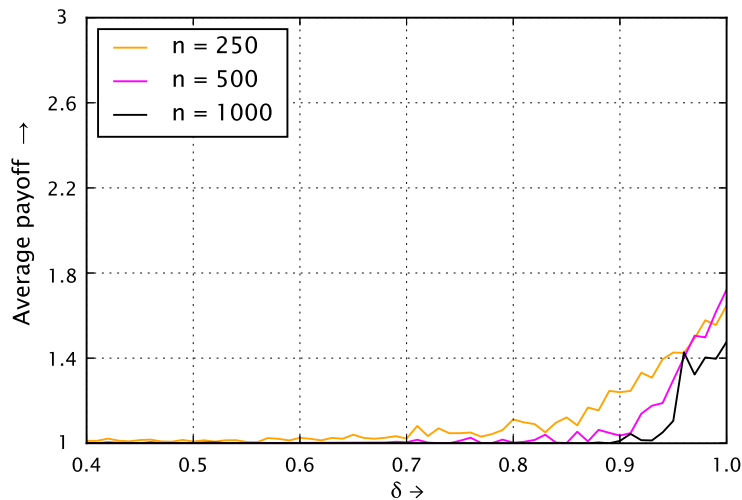


Fig. 6. Average payoffs with different population sizes for $b = 3$, $c = 1$, and complexity costs 0.01. Payoffs do not include complexity costs.

to determine the fate of mutants using analytical fixation probabilities, which requires much less computing time (and even then, runs with large populations take rather long). Further details of the simulation program are given in Appendix B. The simulation runs are sufficiently long for the population to pass through a variety of temporarily stable regimes with different payoffs. The average payoffs reflects the relative amounts of time the population spends in different regimes. Section 5.2 informally discusses how changes in the relative likelihood of different transitions changes these averages.

With a fixed population size, but varying complexity costs, we find that a decrease in complexity costs increases the average payoff (Fig. 5). Moreover, when k decreases to 0, the average payoff approaches the average payoff at $k = 0$ from below, suggesting that there is indeed no discontinuity at $k = 0$ in the expected level of cooperation.

For fixed costs k , on the other hand, an increase in population size shifts the balance back towards what is suggested by the equilibrium analysis with positive costs (Fig. 6). As the population size grows, the average payoff drops ever more towards the payoff of mutual defection.

5.2. Neutral and almost neutral transitions for $\delta < 1$

In order to understand how the static concepts and the finite population dynamics relate, we can first consider the effect of a decrease in k on the stability of All D. For $k > 0$, strategies S that also always play D against copies of themselves (and therefore also against All D), but that have more states, have an invasion barrier of 1. As noted in the proof of Theorem 9, this is because $(1 - \epsilon)U(AllD, AllD) + \epsilon U(AllD, S) = c - k$, while $(1 - \epsilon)U(S, AllD) + \epsilon U(S, S) = c - N_S k$, and hence the expected fitness, or profit, of All D in its 1-state implementation is always $(N_S - 1)k$ larger than the expected fitness of S , regardless of the share ϵ of strategy S . But while the invasion barrier against S is informative about how large ϵ can be

for All D to still outperform S, it has no implications for the magnitude of the difference in fitness. Here this difference vanishes as k goes to 0, and in the Wright–Fisher, or the Moran process, the fixation probability of such an “almost neutral” mutant therefore converges to the fixation probability of a neutral mutant (see Lemma 12 in Appendix C). Some of these mutants would open the door for other, more cooperative mutants with a strict advantage (García and van Veelen, 2016), and therefore we should expect indirect invasions, through “almost neutral” mutants, and into cooperation, to become important in the dynamics as $k \downarrow 0$, in spite of the fact that All D has an invasion barrier of 1 for all $k > 0$ against all strategies S that also always play D against copies of themselves, but that have more than 1 state.¹

The detrimental effect that a decrease in k has on the stability of All D is mirrored by the mitigating effect that a decrease in k has on the instability of cooperative strategies that have unused punishment states, like Tit-for-tat. Such strategies cannot be equilibria as long as $k > 0$, but they can for $k = 0$. What makes Tit-for-tat a disequilibrium strategy for $k > 0$ is the presence of All C, which has fewer states, but also always plays C against a copy of itself, and against Tit-for-tat. Their expected payoffs against various mixes of the two are $(1 - \epsilon)U(TFT, TFT) + \epsilon U(TFT, AllC) = b - 2k$ for Tit-for-tat, and $(1 - \epsilon)U(AllC, TFT) + \epsilon U(AllC, AllC) = b - k$ for All C, and All C therefore does strictly better than Tit-for-tat for all $\epsilon \in (0, 1)$. The difference in this case is k , and although All C does better for all ϵ , how much better it does vanishes as k goes to 0. What is a direct invasion of All C into Tit-for-tat for all $k > 0$ therefore gets closer and closer to being a neutral transition if $k \downarrow 0$ (see again Lemma 12 in Appendix C).

Now All C only harms cooperation because it opens the door for All D. In the limit of $k \downarrow 0$, the relevant way out of Tit-for-tat therefore becomes an indirect invasion through an “almost neutral” mutant, even though All C is not neutral yet for any $k > 0$. Therefore, even though Tit-for-tat is not even an equilibrium for any $k > 0$, if k decreases to 0, its stability in the finite population dynamics gets arbitrarily close to the stability in the dynamics of Tit-for-tat at $k = 0$, where Tit-for-tat is an NSS, but not RAll (just like all other NSS'es).

Tit-for-tat is an example of a strategy with payoffs that are not within the bounds given by Theorem 7 for $\delta < 1$. There are many other strategies that are also equilibria at $k = 0$, but have unused punishment states, and therefore are not equilibria for any $k > 0$. These strategies can have payoffs in the entire interval between the payoff of mutual cooperation and mutual defection, and also for those, the relevant path out is first a mutant that loses the unused punishment state, and then a mutant that takes advantage of the lack of punishment. With $k > 0$, the first mutant is the reason why such a strategy is not an equilibrium, but as k decreases, this becomes an “almost neutral” transition, which at $k = 0$ would be part of the indirect invasion. The possibility of such an indirect invasion is what stops the strategy from being RAll at $k = 0$, but it does not stop it from being an NSS at $k = 0$.

Besides the almost neutral transitions, there is a second effect of a decrease in k . While strategies with unused states cannot be equilibria, there are strategies without unused states that can – for instance strategy A in Fig. 1. In order for such strategies to actually be equilibria, complexity costs k still need to be sufficiently low. For any given δ , the thresholds that complexity costs k should be under will differ across strategies, and therefore, as k decreases, ever more strategies will become NSS'es. Even though none of them are RAll, these equilibria might still be visited by the dynamics. The growing set of equilibria as k decreases is visualized in a time lapse.²

While a decrease in complexity costs k for a given population size makes almost neutral transitions converge to neutral transitions, the simulations with fixed k and a variety of population sizes suggest that an increase in the population size can reinforce fitness differences, making fixation of mutants with even a slight disadvantage ever more unlikely, relative to the benchmark of neutral transitions (which is $\frac{1}{n}$, where n is the population size). The effect of a decrease in k on fixation probabilities of almost neutral mutants therefore can be counterbalanced by an increase in n .

5.3. Additional transitions for $\delta = 1$

With $\delta < 1$, the static equilibrium concepts suggest that for $k > 0$, All D is more stable than other, more cooperative equilibria, while they are equally stable at $k = 0$. With $\delta = 1$ the same equilibrium concepts suggest that, while All D is still more stable than any other equilibrium for $k > 0$, it is less stable than more cooperative ones at $k = 0$. The reason is that, at $\delta = 1$, and without complexity costs, $U(AllD, AllD) = c = U(TFT, AllD)$ and $U(AllD, TFT) = c < b = U(TFT, TFT)$, and therefore All D is not an NSS. For $k > 0$, TFT does have a positive invasion barrier, also at $\delta = 1$, but $\delta = 1$ is the only place where this invasion barrier converges to 0 for $k \downarrow 0$ (see Theorem 9). Direct invasions into All D therefore should become

¹ As the proof of Theorem 9 indicates, strategies S that do not always play D when playing against a copy of themselves may have maximum invasion barriers below 1. If they do, then the proof also indicates why their maximum invasion barriers will decrease as $k \downarrow 0$, but not to 0. For those strategies, and for $\epsilon \in (0, \bar{\epsilon}(AllD))$, the difference between $(1 - \epsilon)U(AllD, AllD) + \epsilon U(AllD, S)$ and $(1 - \epsilon)U(S, AllD) + \epsilon U(S, S)$ does not converge to 0. So in spite of these strategies having an invasion barrier that is lower than the invasion barrier of the “almost neutral” mutants, they are not what makes All D unstable as $k \downarrow 0$.

² See <https://youtu.be/8hWuAltD7gQ>. This time lapse depicts the payoffs of a specific subset of automata, which are denoted using three parameters, n , m and j , with $2 \leq n \leq 15$, $1 \leq m \leq 15$, and $2 \leq j \leq n + 1$. When playing against copies of themselves, these strategies first go through n states in which they defect, then they go through m states in which they cooperate, and then they return to state j , from where they start cycling. Punishment always means a return to state 1. Formally: the output in the different states is $\lambda_{(n,m,j)}(i) = D$ for $i = 1, \dots, n$, and $\lambda_{(n,m,j)}(i) = C$ for $i = n + 1, \dots, n + m$. The transitions are $\mu_{(n,m,j)}(i, C) = \mu_{(n,m,j)}(i, D) = i + 1$ for $i = 1, \dots, n$; $\mu_{(n,m,j)}(i, D) = 1$ for $i = n + 1, \dots, n + m$; $\mu_{(n,m,j)}(i, C) = i + 1$ for $i = n + 1, \dots, n + m - 1$; and $\mu_{(n,m,j)}(i, C) = j$ for $i = n + m$. For any given δ , ever more of them become equilibria as complexity costs decrease. This is only a subset of all automata, though, because not all automata fit this pattern. The time lapse therefore is not a complete description of all equilibrium payoffs.

ever more relevant in the dynamics as complexity costs vanish. Please note that, while All D ceases to be an NSS, there is still a range of NSS'es with payoffs between c and b , even though none of them are RAll. The fact that All D is not an NSS therefore does not mean that fully efficient equilibria are more stable than those that are not; it only means that fully inefficient equilibria are less stable than a range of less inefficient ones. More details about the stability of different equilibria for $\delta = 1$ are in the next section.

6. MESS and LESS

In this section we would like to revisit two equilibrium concepts that also take complexity costs into account. First we reproduce the definitions of a *modified evolutionarily stable strategy* (MESS) from Binmore and Samuelson (1992), and a *lexicographic evolutionarily stable strategy* (LESS) from Volij (2002). Because their results consider games that are repeated indefinitely, we will also restrict attention to $\delta = 1$ here.

Definition 10 (MESS, Binmore and Samuelson, 1992). A finite automaton S is a MESS if for any mutant T

- 1) $\Pi(S, S) \geq \Pi(T, S)$
- 2) if $\Pi(S, S) = \Pi(T, S)$ then $\Pi(S, T) \geq \Pi(T, T)$
- 3) if $\Pi(S, S) = \Pi(T, S)$ and $\Pi(S, T) = \Pi(T, T)$ then $|S| \leq |T|$

An automaton S is a strict MESS if the last inequality is strict (Samuelson and Swinkels, 2003).³

Definition 11 (LESS, Volij, 2002). A finite automaton S is a LESS if for any mutant T

- 1) $\Pi(S, S) \geq \Pi(T, S)$
- 2) if $\Pi(S, S) = \Pi(T, S)$ then $|S| \leq |T|$
- 3) if $\Pi(S, S) = \Pi(T, S)$ and $|S| = |T|$ then $\Pi(S, T) > \Pi(T, T)$

One could define a neutral LESS by allowing the last inequality to be non-strict. The difference between these stability concepts – and especially between a *strict* MESS and a LESS – is explained clearly in Samuelson and Swinkels (2003).

These two definitions lead to very different predictions. For symmetrized, infinitely repeated games, where payoffs are evaluated according to the limit of means, Binmore and Samuelson (1992) find that strategies that are a MESS must be *utilitarian* (where a utilitarian strategy playing against itself maximizes the sum of payoffs in the underlying game). Applied to standard prisoners dilemmas, where the sum of payoffs is maximized if both cooperate, this implies that, after an initial phase, a MESS always cooperates, when playing against a copy of itself, thereby achieving fully efficient (limit-of-means) payoffs.

The main result in Volij (2002, Claim 1), on the other hand, is that only All D is a LESS, and that there is no ES-set other than {All D}, if the definition of an ES-set does not look at payoffs only, but instead uses lexicographic preferences where complexity matters after payoffs.⁴

The static equilibrium concepts that were used in Sections 3 and 4 do not use lexicographic preferences, and are linked directly to stability properties in the replicator dynamics. An ESS is asymptotically stable in the replicator dynamics (Taylor and Jonker, 1978; Weibull, 1995), and being RAll would imply that a strategy is an element of an ES-set (van Veelen, 2012), which is also asymptotically stable in the replicator dynamics (Thomas, 1985). The lexicographic treatment of complexity costs in the definitions of a MESS and a LESS on the other hand prevents that one can make a similar link for either of these two stability concepts. One could however imagine that being MESS or LESS might nonetheless be informative about dynamics for small complexity costs, given that complexity costs matter after payoffs in both of them – albeit in different

³ Samuelson and Swinkels (2003) also allow for mixed strategies in their definition of a MESS and LESS. They also indicate that they follow the standard practice (cf. Rubinstein, 1986, and Abreu and Rubinstein, 1988) of working with pure MESS or LESS to avoid difficulties in interpreting the complexity of mixed strategies. We do the same, also because the main results from Binmore and Samuelson (1992) and Volij (2002) concern pure strategies.

⁴ Volij (2002) actually shows that in the repeated prisoners dilemma, there is no evolutionary stable family of strategies (ESF) – where he allows for, and uses, lexicographic preferences – other than the singleton set {All D}. Applied to preferences that are not lexicographic, an ESF, as defined in Volij (2002) is a (Balkenborg & Schlag) ES-set, and therefore also a (Thomas) ES set (see Thomas, 1985; Balkenborg and Schlag, 2001; van Veelen, 2012), but the reverse is not true; in the following game, the set of all mixtures of the first two strategies is an ES-set, but not an ESF.

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The proof of Claim 1 in Volij (2002) however also works perfectly well for showing that there is no ES-set, other than the singleton set {All D}, if the definition of an ES-set does not look at payoffs only, but instead uses lexicographic preferences, where complexity matters after payoffs. The proof even provides the indirect invasions that imply that, with positive complexity costs instead of lexicographic preferences, no strategy other than All D is RAll, and therefore there are also no ES-sets in the standard definition, without lexicographic preferences, apart from the singleton set {All D}.

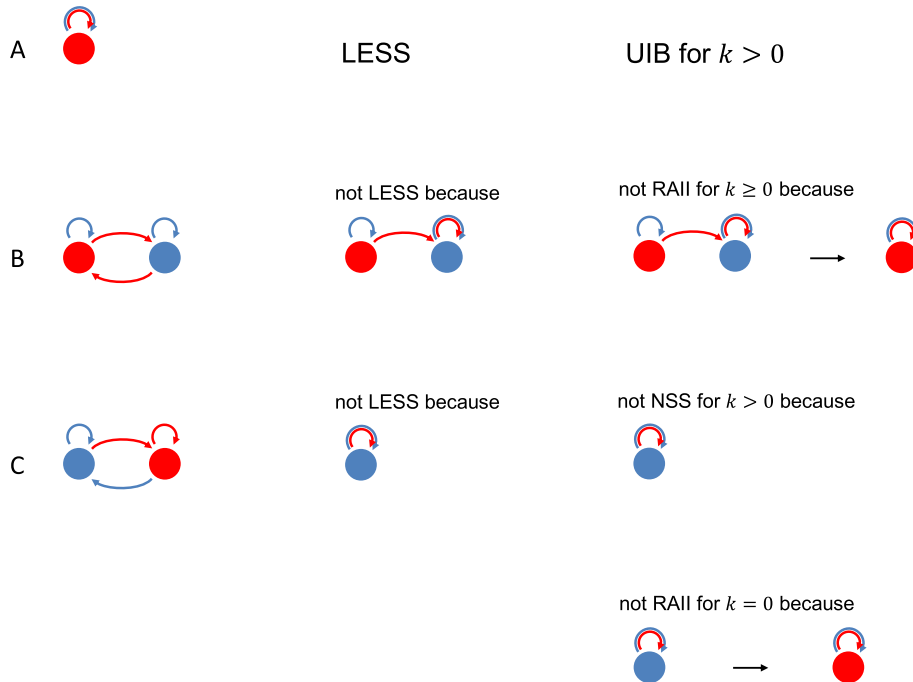


Fig. 7. Strategy A is All D in its 1-state implementation. It is the only LESS, and it also is the only strategy that has a uniform invasion barrier (UIB) for $k > 0$. Strategy B is called Tat-for-tit by Binmore and Samuelson (1992). The reason why it is not a LESS, is that there is another strategy that does equally well, both against Tat-for-tit and against itself, and has equally many states. The reason why it is not RAI, is that same strategy that stops it from being a LESS also opens the door for All D. Strategy C is Tit-for-tat, and it is not a LESS, because it has an unused state. The reason why it is not an NSS, and therefore also not RAI at $k > 0$, is that then the same strategy, with the unused state removed, does strictly better. Also at $k = 0$ Tit-for-tat is not RAI, because the strategy with the unused state removed now does equally well, and opens the door for All D.

ways. One possibility would be that depending on exactly how small these complexity costs are, the level of cooperation in the dynamics could range anywhere between these extremes, all the way from fully defecting (LESS) to fully cooperative (MESS).

In the simulations, we have seen that the range the average payoffs are in does indeed start at the payoffs of mutual full defection. If we look at the average payoffs in simulations for fixed N , $\delta = 1$ and varying k , we find that they do get ever closer to the payoff of mutual defection as k increases (Fig. 5).

For $k \downarrow 0$, on the other hand, we do *not* find that the average payoffs approach the payoff of full mutual cooperation (see Fig. 5). The simulations for fixed N , $\delta = 1$ and varying k suggest that at $k = 0$ the average payoff is a value smaller than the payoff of mutual cooperation, and that this value is approached from below if k decreases to 0.

In the remainder of this section we will indicate which mutants matter for whether or not strategies have a UIB, are NSS, RAI, LESS and MESS, and which of those mutants are relevant to the dynamics.

6.1. LESS

The reason why no strategies other than All D are a LESS, is that for any strategy S with more than one state, one can either find a mutant T for which $\Pi(S, S) = \Pi(T, S)$ and $|S| > |T|$, or one can find a mutant T for which $\Pi(S, S) = \Pi(T, S)$, $|S| = |T|$ and $\Pi(S, T) = \Pi(T, T)$. If S has an unused state, then a strategy T that is the same as S , but with the unused state removed, would make a mutant of the first kind. If S does not have an unused state, then one can always make inconsequential changes to the transition function of the automaton in order to make a mutant of the second kind. This can be done by taking any state $1 \leq i \leq N_S$, and if the output in that state happens to be $\lambda_S(i) = C$, then make T the same as S , but change the transition out of state i when the opponent plays D from $\mu_S(i, D) = j$ to $\mu_T(i, D) = k$ for some $k \neq j$. Similarly, if $\lambda_S(i) = D$ and $\mu_S(i, C) = j$, then one can make T the same as S , but change to $\mu_T(i, C) = k$ for $k \neq j$.

Because it is always possible to make some mutant of one type or the other that suspends punishment in a state where $\lambda_S = C$, the same strategies that are not LESS are also not RAI (see also Fig. 7). And since All D is the only strategy that is both a LESS and has a UIB for $k > 0$, these two approaches select the same strategy.

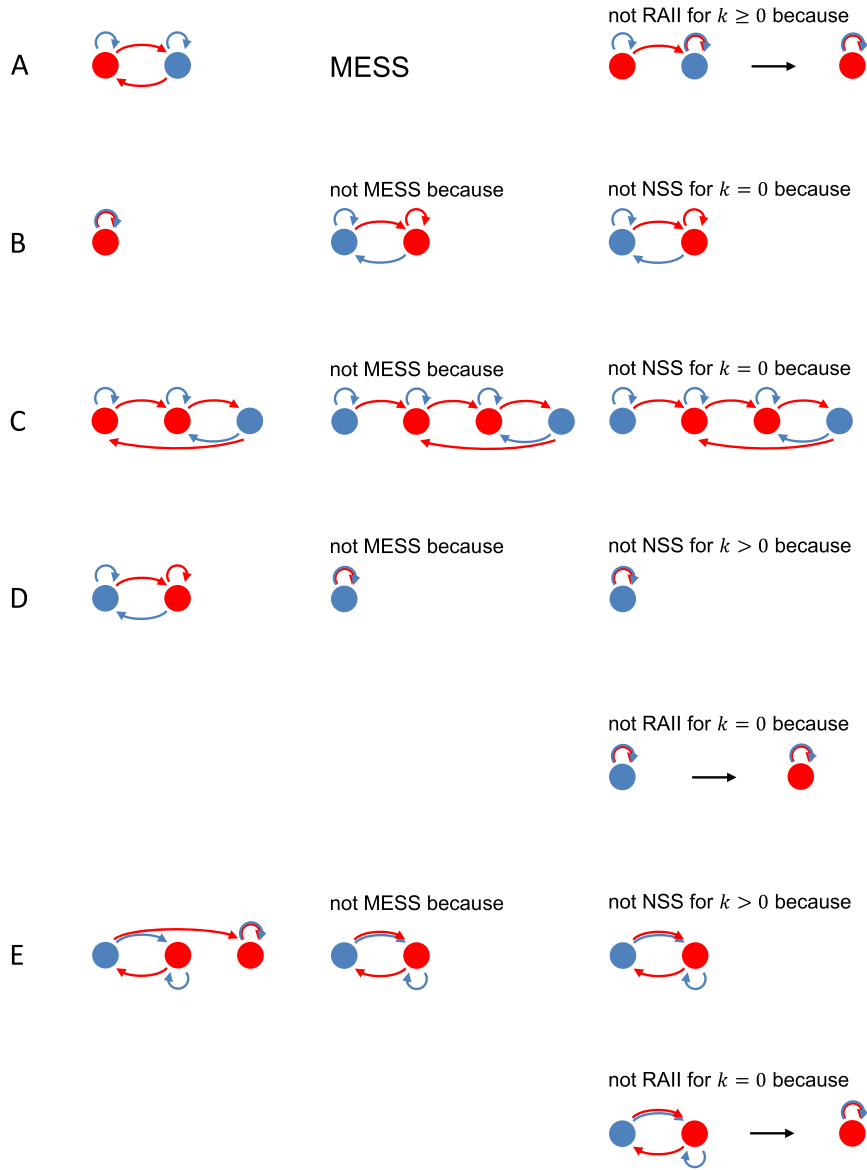


Fig. 8. Strategy A is called Tat-for-tit by Binmore and Samuelson (1992), and it is a MESS. Strategies B and C have no unused states, and become vulnerable to direct invasions in the limit of $k \downarrow 0$. Strategies D and E do have unused punishment states, and therefore they are also not MESS'es. In the limit of $k \downarrow 0$, however, strategies D and E become as stable, or unstable, as strategy A, because at $k = 0$ all three of them allow for indirect, but not direct invasions (see also Lemma 12 in Appendix C).

6.2. MESS

Applying the definition of a MESS separates fully efficient equilibria from those that are not fully efficient. There are strategies, such as strategy A in Fig. 8, that are a MESS, and achieve the payoff of full mutual cooperation. All strategies with less than fully efficient limit-of-means payoffs when playing against a copy of themselves are not a MESS.

Applying the definition of an NSS at $\delta = 1$ and $k = 0$, on the other hand, separates All D from a range of more cooperative equilibria. Any strategy that always defects can be directly invaded at $k = 0$, and is therefore not an NSS, while there is a range of more cooperative NSS'es (Bendor and Swistak, 1995). None of these are RAI (García and van Veelen, 2016), and at $k = 0$, as well as in the limit of $k \downarrow 0$, the dynamics make finite populations move between these different NSS'es. The key to this difference, and to why the dynamics do not line up with the prediction of full efficiency for small complexity costs, lies in the way the definition of a MESS treats strategies with unused states.

Strategies with unused states cannot be a MESS. If a strategy S does not have unused states, it is always possible for a mutant T to both figure out whether it is playing against S or against another T , and best respond to S if it turns that is what the opponent is. Moreover, one can always make T such that it cooperates fully when it figures out that it is playing

against a copy of itself (see Binmore & Samuelson, 1992). Such mutants prevent strategies without unused states that are not utilitarian from being a MESS. Since they also constitute direct invasions at $k = 0$, they also prevent them from being an NSS at $k = 0$. Strategies B and C in Fig. 8 are examples. These direct invasions become ever more likely as complexity costs decrease to 0. The concept of a MESS therefore captures that these mutants destabilize all strategies that are not fully cooperative, and that do not have unused punishment states.

If a strategy S on the other hand does have one or more unused punishment states, it might be impossible for a mutant T to both figure out if its opponent is S or T , and respond optimally if its opponent is S . As complexity costs go to 0, the disadvantage of having unused states also goes to 0. For strategies with unused punishment states, mutants with the unused state removed become ever closer to neutral as complexity costs decrease to 0 (strategies D and E in Fig. 8). Thereby they become as stable as they are for $k = 0$. Some of these do not cooperate fully, when playing against themselves, and cannot be figured out without triggering everlasting punishment (Strategy E in Fig. 8). Although the presence of unused states rules them out for being a MESS, at $k = 0$ they are as stable or instable as MESS'es, since it takes an indirect invasion to leave them in either case.

For $k \downarrow 0$ stability properties for strategies with and without unused states therefore move in opposite directions. Strategies that do not have unused punishment states become disequilibrium states, because at $k = 0$ it is free to build a machine that can figure them out. Strategies that do have unused punishment states on the other hand become equilibria at $k = 0$. Furthermore, at $k = 0$ we also know that no strategy is RAll, so the remaining NSS'es are all susceptible to indirect invasions, the fully cooperative ones as well as the not fully cooperative ones (while the fully defecting ones are not even NSS'es). This explains why we do not see average behavior converge to full cooperation if $k \downarrow 0$.

7. Conclusion and discussion

If we include complexity costs, we find that there are strategies with positive levels of cooperation that are NSS'es for sufficiently low complexity costs and sufficiently high continuation probability. This extends a folk theorem result in Cooper (1996) for infinitely repeated prisoners dilemmas without discounting to repeated prisoners dilemmas with continuation probabilities smaller than 1. The NSS'es with positive levels of cooperation are however not robust against indirect invasions – just like they were not in the absence of complexity costs. All D, on the other hand, in its 1-state implementation, is an ESS for continuation probabilities smaller than 1, and even has a uniform invasion barrier. This was not the case without complexity costs.

By choosing complexity costs and population size, one can push the average payoffs either towards the payoffs of mutual defection – which is what the static analysis with positive complexity costs would suggest – or towards the average payoffs that one would get in the absence of complexity costs. For fixed complexity cost larger than 0, one can get the average payoffs ever closer to the payoffs of mutual defection by increasing the population size. For a fixed population size, on the other hand, we can get the average payoffs ever closer to the average payoffs in the absence of complexity costs by letting the complexity costs approach 0. The dynamics with small complexity costs therefore are somewhere between what one would expect from the static analysis with complexity costs $k > 0$ and the static analysis without complexity costs ($k = 0$). The relevance of the latter can be understood if we realize that with small complexity costs, we can get “almost neutral” mutants opening doors for mutants with a selective advantage, which implies that we would expect transitions that, in the absence of complexity costs, would classify as indirect invasions.

Appendix A. A distance for strategies

We would like to have a distance that reflects how differently, or similarly, two strategies play. We therefore will use a function that, for any given history, returns 0 if the two strategies play the same action, and 1 if they play a different one. Let $f : H \times S \times S \rightarrow \{0, 1\}$ be defined by

$$f(h_t, S, T) = \begin{cases} 0 & \text{if } S(h_t) = T(h_t) \\ 1 & \text{if } S(h_t) \neq T(h_t) \end{cases}$$

We assume that the action space A is finite, and that it has k elements, a_1, \dots, a_k . The number of possible histories in H_t – the set of all histories at time t – is k^{2t-2} . For any given t , we will give all histories equal weights. Because a history at time t is reached at all with probability δ^{t-1} , the set H_t as a whole is weighted with that probability. Therefore we define the distance between S and T , both $S, T \in \mathcal{S}$, as follows:

$$d(S, T) = (1 - \delta) \sum_{t=1}^{\infty} \rho^{t-1} \sum_{h_t \in H_t} |f(h_t, S, T)| \text{ with } \rho = \frac{\delta}{k^2} \text{ and } \delta \in (0, 1).$$

With this definition, $d(S, S) = 0$, and, if we take the example of the repeated prisoner's dilemma, $d(AllC, AllD) = 1$, which implies that both are independent of δ . Distances between many pairs of strategies however will depend on δ .

If we take for $\mathcal{S}_t \subset \mathcal{S}$ the set of strategies in \mathcal{S} that all play a_1 for all histories h_u with $u > t$, then it is a finite set; it has $k^{\left(\sum_{v=1}^t k^{2v-2}\right)} = k^{\left(\frac{k^{2t}-1}{k^2-1}\right)}$ elements. The set $\bigcup_{t=1}^{\infty} \mathcal{S}_t$ is therefore countable, but it is easy to see that it is dense in \mathcal{S} . If we

now restrict attention to the repeated prisoner's dilemma, and consider the set of all finite state automata, then this set is also countable, and dense in S .

Appendix B. The simulation procedure

The simulations presented in Section 5.1 capture evolutionary dynamics in finite populations for low mutation rates. For sufficiently low mutation rates the population almost always consists of at most two strategies, since new mutants mostly arrive after the fate of previous mutants has been determined. We use this fact to simulate efficiently. The simulation procedure is as follows.

1. First the program samples the time required for a new mutant to arrive. This follows a geometric distribution with low success probability.
2. A mutant strategy is generated by following the procedure in García and van Veelen (2016), applied to the resident.
3. For this mutant and the resident, expected payoffs against themselves and each other are computed (see Section 2).
4. The fixation probability p of the mutant is computed for the Wright-Fisher process. This requires building the transition matrix of the Markov chain without mutations – see Imhof and Nowak (2006).
5. A Bernoulli trial with probability p determines whether the mutant becomes the new resident, or the current resident stays.

This process is repeated a large number of times. The result of any given simulation run is a list of resident strategies, together with the time they spent as residents. For every resident we can compute the payoff against itself, and combine that with their durations in order to compute the average level of cooperation in the run.

This simulation algorithm more efficiently explores the same strategy space as the simulations in García and van Veelen (2016). In those simulations, randomness plays a role, not only in generating mutants, but also in the payoff computation. All individuals in any given generation are randomly matched in pairs to play the repeated prisoners dilemma once. For every pair separately, every round the game repeats with probability δ , so the realized number of repetitions often differs between pairs. For large population sizes, in order for it to be reasonably likely that the fate of a mutant has been decided before the next mutant arrives, mutation rates have to be small. As a result of both the large population size itself and the small mutation rate, these simulations become prohibitively time-consuming. A lot of computation time is spent idling, waiting for the next mutant to arrive, or during fluctuations in the interior of the simplex before fixation happens. The algorithm used here is computationally more efficient at exploring an unbounded strategy space. It does this while faithfully retaining the demographic noise that is key to the indirect invasions that drive the dynamics.

Appendix C. Almost neutral transitions

The Wright-Fisher process, with resident S and mutant T , is a Markov chain with transition probabilities that follow a binomial distribution, where the probability of a success is determined by the payoffs. In state $i = 0, \dots, n$, the number of individuals playing strategy T is i , and the number of individuals playing strategy S is $n - i$. Both state 0 and state n are absorbing. If S and T earn the same payoffs against themselves and against each other, that is, if $\Pi(S, S) = \Pi(S, T) = \Pi(T, S) = \Pi(T, T) = \Pi$, and if $N_T < N_S$, then the transition probability from state i to state j equals

$$T_{i \rightarrow j}(k) = \binom{n}{j} \left(\frac{i(\Pi - N_T k)}{n\Pi - iN_T k - (n-i)N_S k} \right)^j \left(\frac{(n-i)(\Pi - N_S k)}{n\Pi - iN_T k - (n-i)N_S k} \right)^{n-j}$$

These transition probabilities are continuous in k , and

$$T_{i \rightarrow j}(0) = \binom{n}{j} \left(\frac{i}{n} \right)^j \left(\frac{n-i}{n} \right)^{n-j}$$

From Theorem 1.1 in Durrett (2008), we know that with the transition probabilities for the neutral Wright-Fisher process at $k = 0$, the fixation probability (the probability of eventually absorbing in state n , when starting in state 1) equals $\frac{1}{n}$. We would like to show that the fixation probability $\rho_T(k)$ of T approaches $\rho_T(0) = \frac{1}{n}$ in the limit of $k \downarrow 0$.

Lemma 12. *If transition probabilities $T_{i \rightarrow j}(k)$ are continuous in k , and if $\lim_{k \downarrow 0} T_{i \rightarrow j}(k) = \binom{n}{j} \left(\frac{i}{n} \right)^j \left(\frac{n-i}{n} \right)^{n-j}$ then $\lim_{k \downarrow 0} \rho_T(k) = \rho_T(0) = \frac{1}{n}$.*

Proof. The number of paths that start in state 1, and reach state n for the first time after m steps is $(n-1)^{m-1}$; the first step starts in state 1, the last step goes to state n , and at all times in between, the system must be in one of the $n-1$ transient states. Define $p_{m,l}(k)$ as the probability of going to fixation in m steps, along path l , where $l = 1, 2, \dots, (n-1)^{m-1}$ enumerates

all possible paths given m , and $m = 1, 2, 3, \dots$. Any such $p_{m,l}(k)$ is continuous in k , as it is a product of continuous one-step transition probabilities. Since all states $0 < i < n$ are transient, we know that for the neutral process

$$\lim_{M \rightarrow \infty} \sum_{m=1}^M \sum_{l=1}^{(n-1)^{t-1}} p_{m,l}(0) = \frac{1}{n}.$$

Now take an ϵ , and choose M sufficiently large, so that

$$\sum_{m=1}^M \sum_{l=1}^{(n-1)^{m-1}} p_{m,l}(0) > \frac{1}{n} - \frac{\epsilon}{2}.$$

By continuity of $p_{m,l}(k)$, we can choose $k = \eta$ sufficiently close to 0, so that

$$\sum_{m=1}^M \sum_{l=1}^{(n-1)^{m-1}} p_{m,l}(\eta) > \sum_{m=1}^M \sum_{l=1}^{(n-1)^{m-1}} p_{m,l}(0) - \frac{\epsilon}{2}.$$

Thereby

$$\rho_T(\eta) = \sum_{m=1}^{\infty} \sum_{l=1}^{(n-1)^{m-1}} p_{m,l}(\eta) > \sum_{m=1}^M \sum_{l=1}^{(n-1)^{m-1}} p_{m,l}(\eta) > \frac{1}{n} - \epsilon$$

By considering paths that start in state 1 and absorb in state 0, instead of state n , one can show in a similar way that for any ϵ , one can also find a number of steps N and a $k = \zeta$ such that

$$1 - \rho_T(\zeta) > \frac{n-1}{n} - \epsilon$$

These observations together imply that $\lim_{k \downarrow 0} \rho_T(k) = \rho_T(0) = \frac{1}{n}$ \square

The Moran process allows for a relatively simple explicit expression for the fixation probability as a function of its transition probabilities (Nowak, 2006), and in this case it is a more straightforward observation that the fixation probabilities converge to $\frac{1}{n}$.

References

- Abreu, D., Rubinstein, A., 1988. The structure of Nash equilibrium in repeated games with finite automata. *Econometrica* 56, 1259–1281.
- Balkenborg, D., Schlag, K.H., 2001. Evolutionarily stable sets. *Int. J. Game Theory* 29, 571–595.
- Bendor, J., Swistak, P., 1995. Types of evolutionary stability and the problem of cooperation. *Proc. Natl. Acad. Sci. USA* 92, 3596–3600.
- Binmore, K.G., Samuelson, L., 1992. Evolutionary stability in repeated games played by finite automata. *J. Econ. Theory* 57, 278–305.
- Bomze, I.M., Pötscher, B.M., 1989. *Game Theoretical Foundations of Evolutionary Stability*. Springer Verlag.
- Cooper, D.J., 1996. Supergames played by finite automata with finite costs of complexity in an evolutionary setting. *J. Econ. Theory* 68, 266–275.
- Durrett, R., 2008. *Probability Models for DNA Sequence Evolution*. Probability and Its Applications, 2nd edition. Springer, New York.
- García, J., van Veelen, M., 2016. In and out of equilibrium I: evolution of strategies in repeated games with discounting. *J. Econ. Theory* 161, 161–189.
- García, J., van Veelen, M., 2018. No strategy can win in the repeated prisoner's dilemma: linking game theory and computer simulations. *Front. Robotics Artif. Intell.* 5, 102.
- Hopcroft, J.E., Motwani, R., Ullman, J.D., 2006. *Introduction to Automata theory, Languages and Computation*. Pearson.
- Imhof, L.A., Nowak, M.A., 2006. Evolutionary game dynamics in a Wright-Fisher process. *J. Math. Biol.* 52, 667–681.
- Kalai, E., Stanford, W., 1988. Finite rationality and interpersonal complexity in repeated games. *Econometrica* 56, 397–410.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press.
- Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. *Nature* 246, 15–18.
- Nowak, M.A., 2006. *Evolutionary Dynamics*. Harvard University Press, Cambridge, MA.
- Rubinstein, A., 1986. Finite automata play the repeated prisoner's dilemma. *J. Econ. Theory* 39, 83–96.
- Samuelson, L., Swinkels, J.M., 2003. Evolutionary stability and lexicographic preferences. *Games Econ. Behav.* 44, 332–342.
- Schlag, K.H., 1993. *Dynamic Stability in the Repeated Prisoners' Dilemma Played by Finite Automata*. Discussion Paper B-243.
- Taylor, P.D., Jonker, L., 1978. Evolutionary stable strategies and game dynamics. *Math. Biosci.* 40, 145–156.
- Thomas, B., 1985. On evolutionarily stable sets. *J. Math. Biol.* 22, 105–115.
- van Veelen, M., 2012. Robustness against indirect invasions. *Games Econ. Behav.* 74, 382–393.
- van Veelen, M., García, J., Rand, D.G., Nowak, M.A., 2012. Direct reciprocity in structured populations. *Proc. Natl. Acad. Sci. USA* 109, 9929–9934.
- van Veelen, M., Spreij, P., 2009. Evolution in games with a continuous action space. *Econ. Theory* 39, 355–376.
- Volij, O., 2002. In defense of defect. *Games Econ. Behav.* 39, 309–321.
- Weibull, J.W., 1995. *Evolutionary Game Theory*. MIT Press, Cambridge.