# Supporting Information

# Indirect reciprocity with private, noisy and incomplete information

Christian Hilbe[1], Laura Schmid[1], Josef Tkadlec[1], Krishnendu Chatterjee[1], and Martin A. Nowak[2]

[1]IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

[2]Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology and Department of Mathematics, Harvard University, Cambridge MA 02138, USA

We are interested in the evolution of indirect reciprocity under incomplete and noisy information. To this end, we describe the dynamics on two separate timescales. First, we consider the *game dynamics*. Here, we take the players' strategies as given, and we compute how the players' reputations change over time, how often they cooperate, and which payoffs they obtain on average. Second, we describe the *evolutionary dynamics*. Here, we allow players to change their strategies over time, assuming that strategies that yield a comparably high payoff are more likely to be adopted. In the following, we describe these two dynamics in more detail.

## 1 Game dynamics

### 1.1 Description of the indirect reciprocity game

We consider a well-mixed population of size $N$ whose members engage in a series of cooperative interactions. In each time step, two members of the population are randomly drawn from the population, a donor and a recipient. The donor can then decide whether or not to pay a cost $c > 0$ to transfer a benefit $b > c$ to the recipient. We interpret paying the cost as cooperation ($C$) and refusing to do so as defection ($D$). The donor's decision is partially observable: the donor and the receiver always learn whether or not the donor decided to cooperate, whereas the other population members independently observe it with probability $q$. Observations may be subject to noise: with probability $\varepsilon$ the donor's action is misinterpreted such that a $C$ is taken for a $D$, or vice versa. All players are equally likely to misperceive an action, independent of whether or not they actively took part in the respective interaction. We assume that there are infinitely many rounds in which players are asked to cooperate, and that the players' payoffs for the indirect reciprocity game are defined as their average payoffs over time (explicit definitions will be provided in the next section).

To make their cooperation decisions, we assume that each player is equipped with a strategy and a private reputation repository. The reputation repository is used to keep track of the reputations of all population members (as illustrated in **Fig. 1**). In line with a large body of the previous literature (1–3), we assume that reputations are binary: players are either considered as 'good' or 'bad'. We represent the

1

state of player $i$'s reputation repository at time $t$ by an $N$-dimensional vector

$$\mathbf{m}_i(t) = \Big( m_{i1}(t), \ldots, m_{iN}(t) \Big).$$ [1]

The entries $m_{ij}$ of this vector are either 0 or 1, whereby $m_{ij}(t) = 1$ means that player $i$ assigns a good reputation to player $j$ at time $t$, whereas $m_{ij}(t) = 0$ means that $i$ views $j$ as bad. We assume that $\mathbf{m}_i(0) = (1, \ldots, 1)$ for all players $i$. That is, initially all players consider everyone as good.

A player's strategy then needs to tell the player (*i*) how to update the entries in her reputation repository after observing a donor's action, and (*ii*) whether to cooperate if the focal player finds herself in the role of the donor. The first aspect is determined by the player's assessment rule. Assessment rules take the form of an 8-dimensional vector

$$\alpha = (\alpha_{gCg},\ \alpha_{gCb},\ \alpha_{bCg},\ \alpha_{bCb},\ \alpha_{gDg},\ \alpha_{gDb},\ \alpha_{bDg},\ \alpha_{bDb}).$$ [2]

The entries of $\alpha$ can again take either the value 0 or 1. An entry $\alpha_{xAy} = 1$ means that it is regarded as good if a donor with reputation $x$ chooses action $A$ against a recipient with reputation $y$. Analogously, $\alpha_{xAy} = 0$ means that the focal player assigns a bad reputation to the corresponding donor. We note that after any interaction, players only update the donor's reputation; the reputation of the receiver remains unaffected. Assessment rules of the form [2] are called third-order assessment rules, since they depend on the donor's action, on the reputation of the donor, and on the reputation of the recipient. If the assessment is independent of the donor's current reputation, the assessment rule is referred to as second-order; and if the assessment is independent of both players' reputations, the assessment rule is referred to as first-order. The two unconditional cases $\alpha = \mathbf{1} := (1, \ldots, 1)$ and $\alpha = \mathbf{0} := (0, \ldots, 0)$, according to which donors are always considered as good or bad, respectively, are zeroth-order assessment rules.

If the focal player finds herself in the role of the donor, her decision is determined by her action rule. Action rules are given by a 4-dimensional vector

$$\beta = (\beta_{gg},\ \beta_{gb},\ \beta_{bg},\ \beta_{bb}).$$ [3]

An entry $\beta_{xy} = 1$ indicates that a respective donor chooses to cooperate if her own reputation is $x$ and the recipient's reputation is $y$. An entry $\beta_{xy} = 0$ indicates that she defects. A player's strategy $\sigma = (\alpha; \beta)$ is a combination of an assessment rule and an action rule. Since there are $2^8$ different assessment rules and $2^4$ different action rules, the space of third-order strategies contains $2^{12} = 4,096$ elements.

Herein, we have only considered the dynamics for a small subset of third-order strategies (but the methods that we introduce in the following sections equally apply to all other strategies). The subset we have considered consists of the so-called leading-eight strategies (4). Under the assumption that all relevant information is shared publicly and that all players agree on everyone's reputation, past research has shown that if the whole population adopts any of the leading-eight strategies, the population will be fully cooperative and no other strategy can invade (5). However, the assumption of publicly shared

information seems rather stringent; it implies that the reputation repositories of all players agree at all times, $\mathbf{m}_i(t) = \mathbf{m}_j(t)$ for all $i, j$ and $t$. Instead, we explore the robustness of the leading-eight strategies when information is incomplete, noisy, and private. In that case, different players may hold different beliefs about their co-players, such that for any pair of players $i$ and $j$ there may be times for which $\mathbf{m}_i(t) \neq \mathbf{m}_j(t)$. To probe the robustness of the leading-eight strategies, we let each of these strategies compete with the two unconditional strategies ALLC and ALLD. Herein, we define these two strategies as ALLC$= (\mathbf{1}; \mathbf{1})$ and ALLD$= (\mathbf{0}; \mathbf{0})$. We note that alternative strategy representations are possible: For example, we could define ALLC as the strategy that deems everyone as bad but cooperates anyway. Using such an alternative definition would not alter our qualitative results.

## 1.2 A formal description of the reputation dynamics

If the strategies $(\alpha^i; \beta^i)$ are given for all players $1 \leq i \leq N$, we can describe the reputation dynamics of the population as a Markov chain. As the state of the Markov chain at time $t$, we collect all individual reputation repositories into an $N \times N$ matrix

$$M(t) = \Big( m_{ij}(t) \Big). \tag{4}$$

We call $M(t)$ the image matrix of the population at time $t$. Given the players' strategies and the current image matrix $M = M(t)$ we can in principle calculate the transition probability $h_{M,M'}$ that describes how likely we are to find the population in state $M'$ at time $t+1$. Because in every single time-step at most one column of the image matrix can change (the column with respect to the player who was randomly chosen to be the donor), we observe that $h_{M,M'} = 0$ if $M$ and $M'$ differ in more than one column. Otherwise, when the image matrices $M$ and $M'$ differ in at most one column, the exact value of $h_{M,M'}$ depends on the players' strategies, on the observation probability $q$, and on the error probability $\varepsilon$.

Over time, the average probability to observe a given image matrix $M$ approaches an invariant distribution $\mathbf{v} = (v_M)$ of the Markov chain $H = (h_{M,M'})$, such that $\mathbf{v}H = \mathbf{v}$ and $\sum_M v_M = 1$. The entries $v_M$ give the expected frequency with which we observe the image matrix $M$ over the course of the indirect reciprocity game. Given the invariant distribution $\mathbf{v}$, we can calculate the average probability that player $i$ considers $j$ as good as

$$\bar{m}_{ij} = \sum_M v_M \cdot m_{ij}. \tag{5}$$

Moreover, we can calculate the average probability $\bar{x}_{ij}$ with which player $i$ cooperates with player $j$ over the course of the game as

$$\bar{x}_{ij} = \sum_M v_M \beta^i_{m_{ii}, m_{ij}}. \tag{6}$$

Using this average cooperation probability, we define player $i$'s expected payoff $\bar{\pi}_i$ as

$$\bar{\pi}_i = \frac{1}{N-1} \sum_{j \neq i} b\bar{x}_{ji} - c\bar{x}_{ij}. \tag{7}$$

3

Unfortunately, we note that the above approach quickly becomes computationally infeasible as the population becomes large. For a population of size $N$, the image matrix has $N^2$ entries, implying that there are $2^{N^2}$ possible image matrices.

Instead of calculating numerically exact payoffs according to Eq. [7], we have thus simulated the reputation dynamics to obtain estimates $\hat{x}_{ij}$ for the average probability that player $i$ cooperates with $j$. These estimates are then plugged into Eq. [7] to obtain estimated payoffs

$$\hat{\pi}_i = \frac{1}{N-1} \sum_{j \neq i} b\hat{x}_{ji} - c\hat{x}_{ij}. \tag{8}$$

By the theory of Markov chains, the estimated payoffs according to [8] converge to the true values [7] if the simulation is iterated for sufficiently many time steps. Unless noted otherwise, we have simulated the Markov chain for $2 \cdot 10^6$ time steps, using a population of size $N = 50$. The respective simulations have been run with MATLAB; all scripts are provided in the **Appendix**.

Depending on the players' strategies, the Markov chain of the reputation dynamics does not need to be ergodic. In those cases, the players' average payoffs according to Eq. [8] may depend on their initial reputations. All our figures are based on the assumption that all players begin with a good reputation. To explore the robustness of our results, we have re-run the simulations in **Fig. 2** for two alternative scenarios. First we have run ten independent simulations in which the initial reputations are assigned randomly. For each of these ten simulations, we have obtained the same result as reported in the main text. Second, we have run a simulation assuming that all players start with a bad reputation. Again, we obtain the same result as in **Fig. 2** for seven out of the eight cases. Only for L7 the result is different. According to the assessment rule of L7 (**Tab. 1**), a bad donor can only gain a good reputation by cooperating against a good recipient. However, since there are no good recipients to start with, it is impossible for donors to gain a good standing in the eyes of an L7 observer. Hence, all L7 players keep their initial bad assessment of all co-players. We conclude that under such an initial assignment of reputations, L7 does not yield any cooperation at all.

## 1.3   Recovery analysis after single disagreements

While we use simulations to explore the general reputation dynamics under noisy private information, we can derive analytical results in the limiting case that the players' actions are perfectly observable ($q = 1$) and that perception errors are rare ($\varepsilon \to 0$). To this end, we consider a homogeneous population of players who all apply the same leading-eight strategy (as given in **Tab. 1**). We assume that initially, all players perceive everyone as good; only player 1 perceives player 2 as bad, possibly because of an error (**Fig. S3a**). That is, the initial image matrix is given by $M(0) = M^0$ with entries

$$m_{ij}^0 = \begin{cases} 0 & \text{if } i = 1, j = 2 \\ 1 & \text{otherwise.} \end{cases} \tag{9}$$

4

We assume that in subsequent rounds, no further errors occur and that all individuals observe all co-players' interactions. We say that the population recovers from a single disagreement, if it returns to the state where all players have a good reputation starting from the state $M^0$. In the following we are interested in the following two quantities, depending on the applied leading-eight strategy $L_i$:

1. The population's recovery probability $\rho_i$
2. The expected time until recovery $\tau_i$ (conditioned on that the population actually recovers)

The following Proposition simplifies the corresponding analysis.

**Proposition 1.** *Consider the indirect reciprocity game for a population in which everyone applies the same leading-eight strategy $L_i$. Moreover, assume that the initial image matrix is $M(0) = M^0$ as defined in [9], and let $M(t)$ denote the image matrix at some subsequent time $t > 0$ according to the process with perfect observation and no noise, $q = 1$ and $\varepsilon = 0$. Then $M(t) \in \mathcal{M}$, where $\mathcal{M}$ is the set of all image matrices that satisfy the following three conditions*

$$(i) \ m_{ii} = 1 \ \textit{for all } i, \quad (ii) \ m_{ij} = 1 \ \textit{for all } i, j \geq 2, \quad (iii) \ m_{i1} = m_{j1} \ \textit{for all } i, j \geq 2. \qquad [10]$$

All proofs are provided in the **Appendix**. The above Proposition guarantees that in the process with perfect observation and no noise, (*i*) all players think of themselves as good, (*ii*) the players $2 \leq i, j \leq N$ consider each other as good, and (*iii*) all players $2 \leq i, j \leq N$ have the same opinion about player 1.

Proposition 1 is useful because it allows us to consider a simplified state space. Instead of considering the space of all image matrices $M$, in the following we consider the space of all tuples $(r, k)$ with $r \in \{0, 1\}$ and $k \in \{0, \ldots, N-1\}$. The value of $r$ refers to the reputation of player 1 from the perspective of all other players (due to Proposition 1 (*iii*) all other players agree on player 1's reputation). We use $r = 1$ to indicate that player 1 is perceived as good, and $r = 0$ to indicate that she is perceived as bad. The value of $k$ refers to the number of co-players that player 1 considers as good (due to Proposition 1, we can treat all other players as equivalent). In this reduced state space, the initial state thus corresponds to the pair $(1, N-2)$, and the full recovery state is $(1, N-1)$, see **Fig. S3a** for an illustration.

Let $f^i(r, k; r', k')$ denote the transition probabilities for the reduced state space; the value of $f^i(r, k; r', k')$ corresponds to the probability that a population of $L_i$ players moves from state $(r, k)$ to $(r', k')$ in one round. We can deduce these probabilities as follows:

*Transition $(1, k) \rightarrow (1, k+1)$.* This case can only occur if a player $i > 1$ is chosen to be the donor who is perceived as bad by player 1. Given that the current state is $(1, k)$, it follows from Proposition 1 that the donor considers everyone as good, and hence she cooperates. If player 1 considers the receiver to be good, this leads her to assign a good reputation to the donor, independent of the applied leading-eight strategy $L_i$. Otherwise, if player 1 considers the receiver to be bad, the donor only obtains a good reputation for $L_1$, $L_2$, $L_3$, and $L_5$. Therefore, the corresponding transition

probability is

$$f^i(1,k;\ 1,k+1) = \begin{cases} \frac{N-k-1}{N} & \text{if } i \in \{1,2,3,5\} \\ \frac{N-k-1}{N}\frac{k+1}{N-1} & \text{if } i \in \{4,6,7,8\}. \end{cases}$$

[11]

*Transition $(1,k) \to (1,k-1)$.* This case can only occur if a player $i > 1$ is randomly chosen to act as the donor who is perceived as good by player 1. Similar to before, player $i$ will always cooperate, which is only considered as bad by player 1 if the receiver is considered as bad by player 1 and if the applied strategy is either $L_2, L_5, L_6$, or $L_8$. Therefore, the transition probability is

$$f^i(1,k;\ 1,k-1) = \begin{cases} 0 & \text{if } i \in \{1,3,4,7\} \\ \frac{k}{N}\frac{N-k-1}{N-1} & \text{if } i \in \{2,5,6,8\}. \end{cases}$$

[12]

*Transition $(1,k) \to (0,k)$.* This case can only occur if player 1 is chosen to be the donor, and if player 1 defects against the receiver (which in turn requires player 1 to consider the receiver as bad). The corresponding transition probability is

$$f^i(1,k;\ 0,k) = \frac{1}{N}\frac{N-k-1}{N-1}.$$

[13]

*Transition $(0,k) \to (0,k+1)$.* This case requires that a player $i>1$ is chosen to be the donor who is considered as bad by player 1. This donor cooperates, unless the randomly chosen receiver happens to be player 1 (who is bad from the perspective of all other players). Thus, player 1 considers the donor as good after this round unless the receiver is player 1, or the receiver is a group member that is considered as bad by player 1 and the applied leading-eight strategy is $L_4, L_6, L_7$, or $L_8$. Hence, we obtain

$$f^i(0,k;\ 0,k+1) = \begin{cases} \frac{N-k-1}{N}\frac{N-2}{N-1} & \text{if } i \in \{1,2,3,5\} \\ \frac{N-k-1}{N}\frac{k}{N-1} & \text{if } i \in \{4,6,7,8\}. \end{cases}$$

[14]

*Transition $(0,k) \to (0,k-1)$.* This case requires that a player $i > 1$ is chosen to be the donor who player 1 considers as good. To become bad in player 1's eyes, this donor then either needs to defect against player 1, or he needs to cooperate against a receiver who is considered as bad by player 1 (provided that the applied leading-eight strategy is $L_2, L_5, L_6$, or $L_8$). The transition probability becomes

$$f^i(0,k;\ 0,k-1) = \begin{cases} \frac{k}{N}\frac{1}{N-1} & \text{if } i \in \{1,3,4,7\} \\ \frac{k}{N}\frac{N-k}{N-1} & \text{if } i \in \{2,5,6,8\}. \end{cases}$$

[15]

*Transition $(0,k) \to (1,k)$.* This case requires player 1 to be the donor, and that player 1 cooperates with

6

her co-player. The corresponding probability is

$$f^i(0, k;\ 1, k) = \frac{1}{N}\frac{k}{N-1}. \tag{16}$$

All other transitions from $(r, k)$ to $(r', k')$ have transition probability $f^i(r, k;\ r', k') = 0$. We note that for this reduced Markov chain, the recovery state $(1, N-1)$ is absorbing. Once this state is reached, it cannot be left anymore, because $f^i(1, N-1;\ 1, N-2) = 0$ and $f^i(1, N-1;\ 0, N-1) = 0$ for all $i$. However, for populations that apply one of the four leading-eight strategies $L_4$, $L_6$, $L_7$, $L_8$, there exists another absorbing state, which is $(0, 0)$. This state corresponds to a full segregation state: player 1 considers everyone else as bad, whereas all other players consider player 1 to be bad (**Fig. S3a**). Whether the population is able to recover from a single disagreement thus depends on how likely the full segregation state is reached. We obtain the following three cases.

**Proposition 2** (Recovery probabilities)**.**
*Suppose the population applies the leading-eight strategy $L_i$.*

    *1. For $i \in \{1, 2, 3, 5\}$, the recovery probability is $\rho_i = 1$.*

    *2. For $i \in \{4, 7\}$, the recovery probability satisfies $1 - 2/(N-1)! \le \rho_i < 1$.*

    *3. For $i \in \{6, 8\}$, the recovery probability is $\rho_i = 1 - 1/N$.*

With respect to the time it takes the population to recover from a single disagreement, we obtain a similar case distinction.

**Proposition 3** (Expected time until recovery)**.**
*Suppose the population applies the leading-eight strategy $L_i$.*

    *1. For $i \in \{1, 3, 4, 7\}$, the expected time until recovery $\tau_i$ is of order $\Theta(N)$.*

    *2. For $i \in \{2, 5\}$, the expected time until recovery $\tau_i$ is of order $\Theta(N \log N)$.*

    *3. For $i \in \{6, 8\}$, the expected time until recovery is $\tau_i = N \cdot H_N - N$, where $H_N = \sum_{n=1}^N \frac{1}{n}$ is the $N$-th harmonic number.*

While some of the results in Proposition 3 only address the asymptotic behavior as the population becomes large, we can numerically compute exact solutions for the expected recovery time for all population sizes, based on the transition probabilities $f^i(r, k;\ r', k')$ specified above. The respective solutions are shown in **Fig. S3b**.

    The above results suggest that with respect to recovery from single disagreements, there are four groups of leading-eight strategies. The strategies $L_1$ and $L_3$ are most robust: after a single disagreement, they are guaranteed to recover, and the expected recovery time is linear in the population size. These two strategies have in common that a cooperating donor is always perceived as good (even if the donor or the receiver had a bad reputation to start with). Due to this property, a single individual who is perceived as

bad can immediately regain a good reputation once it is chosen to act as a donor (which takes an expected time of $N$ rounds).

A similar observation holds for the two strategies $L_4$ and $L_7$: a single individual with bad reputation can always regain a good reputation by cooperating, independent of the identity of the receiver. Unlike the first two strategies, however, $L_4$ and $L_7$ can end up in the full segregation state. However, this requires a rather specific sequence of events: before player 2 can regain his good reputation in the eyes of player 1, it needs to be the case that player 1 is chosen to act as the donor with respect to player 2 as the recipient. In such a scenario, player 1 would defect, yielding her a bad reputation from the viewpoint of all other players. To end up in the full segregation state, all other players would then have to be chosen to act as a donor against player 1. As the population grows bigger, this sequence of events becomes increasingly unlikely, and the recovery probability quickly approaches 1.

Recovery is guaranteed for the two strategies $L_2$ and $L_5$, but in contrast to the first four strategies, recovery may take substantially longer. The longer recovery time is due to the property of these strategies that cooperation against an individual perceived as bad is itself perceived as bad. Thus, if player 2 is randomly chosen to be the recipient of the next indirect reciprocity interaction, the donor is guaranteed to obtain a bad reputation in the eyes of someone (either because the donor is an individual $i \geq 2$, who is then perceived as bad by player 1; or because the donor is player 1, who is then perceived as bad by everyone else).

The above recovery analysis confirms that $L_6$ and $L_8$ perform worst in the presence of single disagreements. Recovery is not guaranteed, and even if it occurs, it may take a substantial time until recovery. Both aspects are due to the relative ease with which these two strategies assign a bad reputation to other population members: a donor who cooperates against a receiver who is perceived as bad is guaranteed to end up with a bad reputation (independent of whether the donor was perceived as good or bad before). Under these conditions, bad reputations can proliferate comparably quickly.

## 2 Evolutionary dynamics

### 2.1 Description of the evolutionary process

While the previous section has assumed that the players' strategies are fixed, we now describe a simple process that allows us to explore which dynamics arises once players may change their strategies over time. To allow for a transparent treatment, we assume that strategy adaptation occurs on a time scale that is slow compared to the reputation dynamics.

Specifically, we assume that the strategies in the population change according to a simple imitation process (6). For this process, we again consider a population of $N$ individuals where each individual $i$ may have its own strategy $(\alpha^i;\ \beta^i)$. In each time step of the evolutionary process, one individual $i$ is chosen at random to update her strategy. There are two ways to do so: with probability $\mu$ (corresponding to the mutation rate), she adopts a random strategy (with all remaining strategies having the same probability to be picked). With the remaining probability $1-\mu$, she seeks for a role model instead, by randomly choosing another individual $j$ from the population. Suppose player $i$'s payoff is given by $\hat{\pi}_i$

and $j$'s payoff is $\hat{\pi}_j$, where $\hat{\pi}_i$ and $\hat{\pi}_j$ are given by the estimated payoffs according to Eq. [8]. We assume that the probability that $i$ adopts $j$'s strategy is given by the Fermi function (7),

$$P(\hat{\pi}_j, \hat{\pi}_i) = \frac{1}{1 + \exp\left(-s(\hat{\pi}_j - \hat{\pi}_i)\right)}.$$ [17]

The parameter $s \geq 0$ is called the strength of selection. It determines how much players value payoffs when adopting new strategies. For $s = 0$ we obtain $P(\hat{\pi}_j, \hat{\pi}_i) = 1/2$, and imitation occurs essentially at random. For $s > 0$, the imitation probability $P(\hat{\pi}_j, \hat{\pi}_i)$ is monotonically increasing in the payoff difference $\hat{\pi}_j - \hat{\pi}_i$. That is, the more there is to gain for player $i$, the more likely she is to abandon her old strategy and to imitate player $j$'s strategy instead.

This elementary updating process, involving mutation and imitation, is then iterated over many time steps. As a result, we obtain an ergodic process on the space of all possible population compositions. This process has again a unique invariant distribution, to which we refer as the selection-mutation equilibrium. The abundance of each strategy in this equilibrium, and the corresponding average cooperation rate, can always be obtained by simulating the above process over a sufficiently long time span. However, in certain special cases, the selection-mutation equilibrium can be computed more efficiently. These more efficient algorithms either require a small or intermediate-sized population in which only a few different strategies are available, or that mutations are rare. In the following two sections we describe these algorithms in more detail.

Here, we have phrased the evolutionary process as an imitation dynamics. However, similar results can be obtained if one assumes that strategies are genetically encoded and inherited in a birth-death process (8), where the fitness of an individual is given by an exponential function of its payoffs, $\exp(s\hat{\pi}_i)$.

## 2.2 Selection-mutation equilibrium in populations with few strategies

Throughout the main text, we have focused on a population of intermediate size $N$ that has access to at most $k = 3$ different strategies. The possible states of the population are elements of the set

$$\Delta_N^k := \left\{ \mathbf{n} = (n_1, \ldots, n_k) \in \mathbb{N}^k \mid \sum_{i=1}^k n_i = N \right\}.$$ [18]

The entries $n_i$ of each vector represent how many players currently apply strategy $i$. The number of such population compositions is

$$|\Delta_N^k| = \binom{N+k}{k}.$$ [19]

The evolutionary process described in Section 2.1 defines a Markov chain with state space $\Delta_N^k$. For two population states $\mathbf{n}, \mathbf{n}' \in \Delta_N^k$, the transition probability to move from $\mathbf{n}$ to $\mathbf{n}'$ in one step of the process

is given by

$$
w_{\mathbf{n},\mathbf{n}'} = \begin{cases} \frac{n_i}{N}\left(\frac{\mu}{k-1} + (1-\mu)\frac{n_j}{N}P(\hat{\pi}_j, \hat{\pi}_i)\right) & \text{if } n_i' = n_i - 1,\ n_j' = n_j + 1,\ n_l' = n_l \text{ for } l \notin \{i,j\} \\[2ex] 1 - \sum_{j \neq i} \frac{n_i}{N}\left(\frac{\mu}{k-1} + (1-\mu)\frac{n_j}{N}P(\hat{\pi}_j, \hat{\pi}_i)\right) & \text{if } \mathbf{n} = \mathbf{n}' \\[2ex] 0 & \text{otherwise.} \end{cases}
$$

[20]

Provided that neither the population size $N$ nor the number of strategies $k$ is prohibitively large, the transition matrix $W = (w_{\mathbf{n},\mathbf{n}'})$ can be computed explicitly. By computing the normalized left eigenvector of $W$ (with respect to eigenvalue 1), we obtain the selection-mutation equilibrium over an evolutionary timescale. We have implemented this algorithm using MATLAB; the corresponding code is provided in the **Appendix**.

### 2.3 Selection-mutation equilibrium in the limit of rare mutations

Alternatively, in arbitrarily large populations with an arbitrary number of available strategies, we can still calculate exact strategy abundances in the selection-mutation equilibrium if mutations are sufficiently rare (9–11). In that case, the population will find itself in a homogeneous state most of the time, in which all individuals adopt the same strategy $i$. Only occasionally, a mutant strategy $j$ arises. This mutant then either reaches fixation in the population, or it goes extinct. The mutant's fixation probability $p_{ij}$ can be calculated explicitly (12, 13),

$$
p_{ij} = \frac{1}{1 + \sum_{l'=1}^{N-1} \prod_{l=1}^{l'} \exp\left(-s(\hat{\pi}_j(l) - \hat{\pi}_i(l))\right)}.
$$

[21]

Here, $\hat{\pi}_j(l)$ and $\hat{\pi}_i(l)$ refer to the payoffs of a mutant and a resident, provided that the number of mutants in the population is $l$. Using Eq. [21], the selection-mutation equilibrium of the evolutionary process can be computed by considering a reduced Markov chain with $k$ states (corresponding to the homogeneous populations in which every player applies the same strategy). The probability to move from state $i$ to state $j$ is given by $p_{ij}/(k-1)$. The invariant distribution of this reduced Markov chain approximates the invariant distribution of the evolutionary process as $\mu$ becomes small (9). Again, we have implemented this algorithm in MATLAB; the code is given in the **Appendix**.

### Appendix

### Proofs of the recovery analysis

*Proof of Proposition 1.* For the proof, we consider the Markov chain $H = (h_{M,M'})$, as defined in Section 1.2, for the limiting case $\varepsilon = 0$ and $q = 1$. We show that the set $\mathcal{M}$ of image matrices that satisfy the three properties in [10] is invariant. That is, let $M \in \mathcal{M}$ be arbitrary and suppose $h_{M,M'} > 0$ for some image matrix $M'$. Then also $M' = (m_{ij}')$ satisfies the three characteristic properties [10].
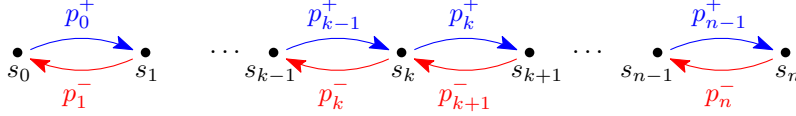
**Figure A1: One-dimensional random walk.** A discrete-time Markov chain with $n+1$ states labelled $s_0 \ldots, s_n$ arranged in a line and with transition probabilities $p_k^+ : s_k \to s_{k+1}$ and $p_k^- : s_k \to s_{k-1}$.

(*i*) $m'_{ii} = 1$ for all $i$. Since $M \in \mathcal{M}$, initially all players consider themselves as good. All leading-eight strategies have the property that the strategy's action rule prescribes an action that lets a good donor maintain her good reputation in her own eyes, independent of which reputation she assigns to the recipient. Thus all players keep considering themselves as good after one interaction; either they do not need to make a decision (because they were not chosen to act as the donor), or they choose an action they themselves evaluate as good.

(*ii*) and (*iii*) $m'_{ij} = 1$ and $m'_{i1} = m'_{j1}$ for all $i, j \geq 2$. Since $M \in \mathcal{M}$, all players $i, j \geq 2$ initially agree on the reputations of all population members. Because they all apply the same assessment rule and observation errors are excluded, they also agree on how the donor's action in the subsequent interaction needs to be assessed. This shows $m'_{il} = m'_{jl}$ for all $i, j \geq 2$ and all $l$. Moreover, since all players $i, j \geq 2$ consider each other as good initially, and since their common action rule only lets them choose actions that let them keep their good reputation, we conclude $m'_{ij} = 1$ for $i, j \geq 2$.

$\square$

*Proof of Propositions 2 and 3.* In the following, we provide the proofs of Propositions 2 and 3 by considering each of the four different cases {L1, L3}, {L2, L5}, {L4, L7}, {L6, L8} individually. For a given leading-eight strategy $L_i$, we consider the respective Markov chain $M_i$ with $2N$ states $s_{r,k}$. The value of $r \in \{0, 1\}$ corresponds to the reputation of player 1 from the perspective of all other players. The value of $k \in \{0, 1, \ldots, N-1\}$ corresponds to the number of co-players that player 1 considers as good. Then the recovery probability $\rho_i$ is the probability that a random realization of the process (a *random trace*) starting in a state $s_{1,N-2}$ reaches the absorbing state $s_{1,N-1}$. The expected recovery time $\tau_i$ is the expected number of steps until a random trace starting in a state $s_{1,N-2}$ reaches the state $s_{1,N-1}$, conditioned on it doing so. The main idea of the proof is to make use of a coupling argument. That is, to each two-dimensional random walk defined by the Markov chain $M_i$, we associate a simpler, one-dimensional random walk. For the one-dimensional random walk, the absorption probabilities and conditional absorption times can be calculated explicitly, and they can serve as upper (or lower) bounds for the respective quantities in the two-dimensional random walk. To this end, let us recall the following result.

**Proposition 4** (One-dimensional random walk, see Chapter 7.7, proof of Theorem 7.1 in (14))**.**
*Let $M$ be a discrete-time Markov chain with $n+1$ states $s_0, \ldots, s_n$ and transition probabilities $p_i^+ : s_i \to s_{i+1}$ ($i = 0, \ldots, n-1$) and $p_i^- : s_i \to s_{i-1}$ ($i = 1, \ldots, n$), as illustrated in Fig. A1.*

11
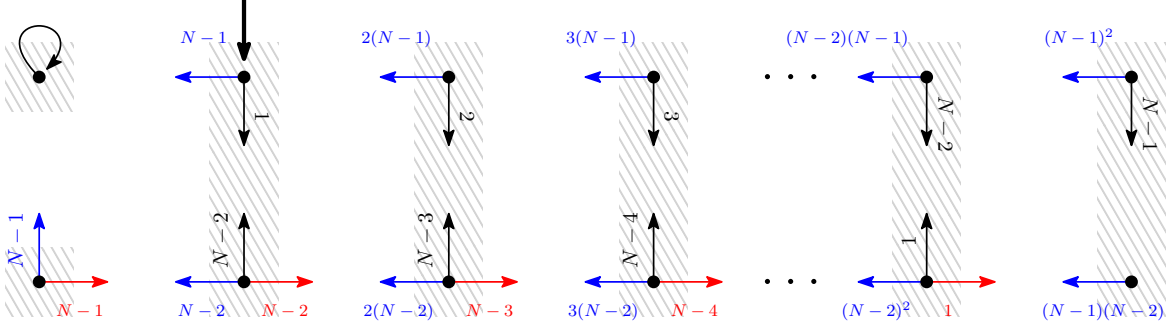
**Figure A2:** Markov chain $M_1 = M_3$. The transition probabilities are normalized by $N(N-1)$. Self-loops in non-absorbing states are not shown.

1. *If both states $s_0$ and $s_n$ are absorbing (i.e. $p_0^+ = p_n^- = 0$) then the probability $\rho(k)$ of reaching $s_n$ before reaching $s_0$ when starting at $s_k$ is given by*

$$\rho(k) = \frac{1 + r_1 + r_1 r_2 + \cdots + r_1 \ldots r_{k-1}}{1 + r_1 + r_1 r_2 + \cdots + r_1 \ldots r_{n-1}}, \qquad [22]$$

*where $r_i = p_i^- / p_i^+$ for each $i = 1, \ldots, n-1$.*

2. *If there is no absorbing state except, possibly, $s_0$ then the expected number of time steps $t_{k,k-1}$ to reach state $s_{k-1}$ from state $s_k$ is given by*

$$t_{k,k-1} = \frac{1}{p_k^-} + \frac{p_k^+}{p_{k+1}^- p_k^-} + \cdots + \frac{p_k^+ \cdots p_{n-1}^+}{p_n^- \cdots p_k^-} = \frac{1}{p_k^-}\left(1 + \sum_{i=1}^{n-k} \prod_{j=1}^{i} \frac{p_{k+j-1}^+}{p_{k+j}^-}\right). \qquad [23]$$

*Moreover, the expected number of time steps $t_{k,l}$ to reach state $s_l$ from state $s_k$, with $l < k$, is*

$$t_{k,l} = t_{k,k-1} + t_{k-1,k-2} + \cdots + t_{l+1,l} = \sum_{i=l+1}^{k} t_{i,i-1}. \qquad [24]$$

We use standard notation $o(\cdot)$ and $\Theta(\cdot)$ for strict asymptotic upper bound and for asymptotically tight bound, ignoring the constant factors. Hence, for example, we have $1/n = o(1)$ and $2n+1 = \Theta(n)$, because for large $n$ we have $1/n \ll 1$ whereas $\frac{2n+1}{n}$ tends to a constant. See Section 1.3 of (15) for a detailed treatment. For asymptotic results, we assume $N \geq 4$.

<u>Recovery analysis for $L_1 = L_3$.</u> The Markov chain $M_1 = M_3$ is depicted in Fig. A2.

- Recovery probability: $M_1$ has a unique absorbing state $s_{1,N-1}$. Moreover, starting from any other state, there is a positive probability to reach the state $s_{1,N-1}$. Hence, it follows that $\rho_1 = 1$.

- Recovery time: To get a lower bound on the recovery time, we note that reaching $s_{1,N-1}$ requires at least one non-self-looping transition. Starting from the state $s_{1,N-2}$, the expected time until the first
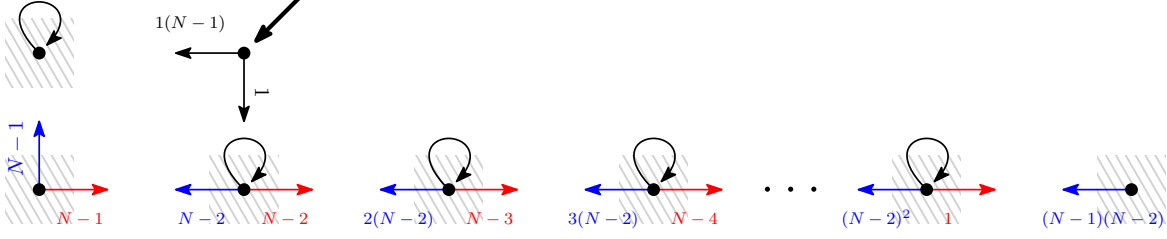
**Figure A3:** Markov chain $M_1'$. The transition probabilities are normalized by $N(N-1)$. By coupling, $\tau_1' \geq \tau_1$.

non-self-looping transition occurs is $\frac{N(N-1)}{N-1+1} = N-1$. Hence $\tau_1 \geq N-1$.

In the following, we show $\tau_1 \leq N+7$, which establishes the desired $\tau_1 = \Theta(N)$. To this end, consider the Markov chain $M_1'$ obtained from $M_1$ by erasing states $s_{1,k}$ for $k \leq N-2$, and by replacing transitions $s_{0,k} \to s_{1,k}$ with self-loops (see Fig. A3).

First, we argue that $\tau_1 \leq \tau_1'$, where $\tau_1'$ is the expected number of steps to reach $s_{1,N-1}$ from $s_{1,N-2}$ in $M_1'$. To this end, consider an arbitrary trace $T$ in $M_1$. If $T$ never takes any transition $s_{0,k} \to s_{1,k}$, for $k < N-1$, then we can associate the identical trace $T'$ in $M_1'$ to $T$. Otherwise, if there is a moment when $T$ moves from $s_{0,k}$ to $s_{1,k}$ for some $k < N-1$, the associated trace $T'$ has a self-loop at $s_{0,k}$. Since for each $k < N-1$ the transition $s_{1,k} \to s_{1,k+1}$ has higher probability than the transition $s_{0,k} \to s_{0,k+1}$, we can then couple traces $T$ and $T'$ such that, from that point on, $T$ is not to the right of or below $T'$ (formally, at any time after that moment, the $k$-coordinate of $T$ is larger or equal to the $k$-coordinate of $T'$, and the $r$-coordinate of $T$ is larger or equal to the $r$-coordinate of $T'$). In particular, if $T'$ has reached $s_{1,N-1}$ then so did $T$, and the inequality $\tau_1 \leq \tau_1'$ follows.

It remains to prove that $\tau_1' \leq N+7$. We use Proposition 4. Let $x$ be the expected number of steps to reach $s_{1,N-1}$ from $s_{0,N-2}$ in $M_1'$. Then $\tau_1' = (N-1) + \frac{N-1}{N} \cdot 0 + \frac{1}{N} \cdot x$. By Eq. [24], $x$ takes the form $x = t_{2,1} + t_{1,0}$. The quantities $t_{2,1}$ and $t_{1,0}$ are calculated using Eq. [23]. For the first quantity, we get

$$
\begin{aligned}
t_{2,1} &= N(N-1)\left(\frac{1}{N-2} + \frac{N-2}{2!(N-2)^2} + \frac{(N-2)(N-3)}{3!(N-2)^3} + \cdots + \frac{(N-2)!}{(N-1)!(N-2)^{N-1}}\right) \\
&\leq \frac{N(N-1)}{N-2}\left(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{(N-1)!}\right) \leq \frac{N(N-1)}{N-2} \cdot (e-1),
\end{aligned}
$$

where we have made use of the definition of Euler's number, $e := \sum_{i=0}^{\infty} 1/i!$. For the second quantity, we similarly obtain

$$
\begin{aligned}
t_{1,0} &= N(N-1)\left(\frac{1}{N-1} + \frac{N-1}{(N-1)(N-2)} + \frac{(N-1)(N-2)}{(N-1)2!(N-2)^2} + \cdots + \frac{(N-2)!}{(N-1)(N-1)!(N-2)^{N-1}}\right) \\
&\leq \frac{N(N-1)}{N-2}\left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{(N-1)!}\right) \leq \frac{N(N-1)}{N-2} \cdot e
\end{aligned}
$$

13

Hence $x = t_{2,1} + t_{1,0} \leq \frac{N(N-1)}{N-2} \cdot (2e-1)$ and, for $N \geq 4$,

$$\tau_1' = N - 1 + \frac{1}{N} \cdot x \leq N - 1 + \frac{N-1}{N-2} \cdot (2e-1) = N + 2(e-1) + \frac{2e-1}{N-2} < N + 7,$$

as desired.

Recovery analysis for $L_2 = L_5$. The Markov chain $M_2 = M_5$ is depicted in Fig. A4.

- Recovery probability: As before, since there is a unique absorbing state that can be reached from any initial state, $\rho_2 = 1$.

- Recovery time: To show that the expected number of steps $\tau_2$ is of the order $\Theta(N \cdot \log N)$, we consider two simpler Markov chains $M_2^+$ (resp. $M_2^-$) with expected number of steps $\tau_2^+$ (resp. $\tau_2^-$). We then prove that $\tau_2^+ \geq \tau_2 \geq \tau_2^-$ and that both $\tau_2^+$ and $\tau_2^-$ are of the order $\Theta(N \cdot \log N)$.

  Intuitively, $M_2^+$ is obtained by identifying states $s_{1,k}$ and $s_{0,k+1}$ for each $k < N-1$ (that is, we identify all states that have the same number of individuals that are universally considered as good). For the transitions of $M_2^+$, we take the transition probabilities as in Fig. A5(a). In doing so, we decrease the total probability to move either to the left or upwards in each state, while leaving the total probability to move to the right or downwards unchanged. As a consequence, any trace $T^+$ in $M_2^+$ can be coupled with a trace $T$ in $M_2$ such that if $T^+$ is in a state $s_{r^+,k^+}$ at time $t$, and $T$ is in a state $s_{r,k}$, then $r^+ + k^+ \leq r + k$. Hence, if $T^+$ has reached $s_{1,N-1}$ then so did $T$ and the inequality $\tau_2^+ \geq \tau_2$ follows. For $M_2^-$ we proceed analogously: we use the same identification of states, but this time the probability to move left in Fig. A5(b) is larger or equal to the corresponding probability to move left or up in Fig. A4. As a consequence, $\tau_2 \geq \tau_2^-$.

  Finally, we compute $\tau_2^+, \tau_2^-$. By Eq. [23], we obtain

$$\tau_2^+ = N(N-1) \left( \frac{1}{N-3} + \frac{N-1}{2!(N-3)^2} + \frac{1 \cdot 2 \cdot (N-1)(N-2)}{3!(N-3)^3} + \cdots + \frac{1 \cdot 2 \cdots (N-1) \cdot (N-1)(N-2) \cdots 1}{N!(N-3)^N} \right)$$

$$= \frac{N(N-1)}{N-3} \left( 1 + \frac{1}{2} \cdot \frac{N-1}{N-3} + \frac{1}{3} \cdot \frac{(N-1)(N-2)}{(N-3)^2} + \cdots + \frac{1}{N} \cdot \frac{(N-1)!}{(N-3)^{N-1}} \right)$$

$$\leq \frac{N(N-1)^2(N-2)}{(N-3)^3} \left( 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} \right) = \Theta(N \log N),$$
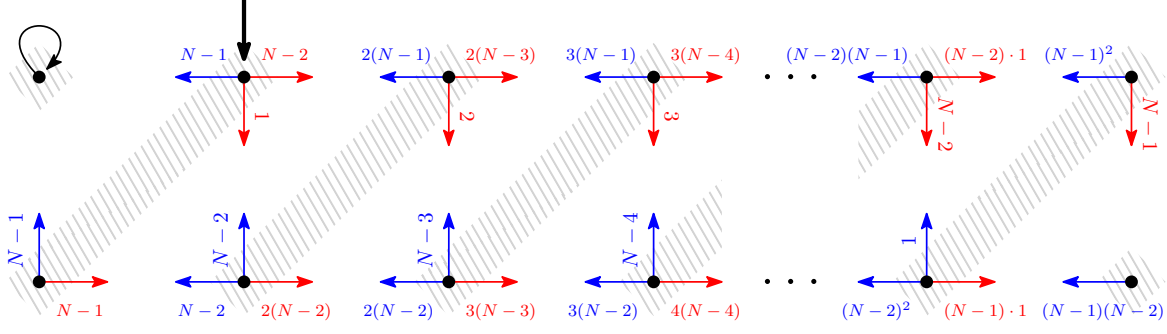
**Figure A4:** Markov chain $M_2 = M_5$. The transition probabilities are normalized by $N(N-1)$. Self-loops in non-absorbing states are not shown.
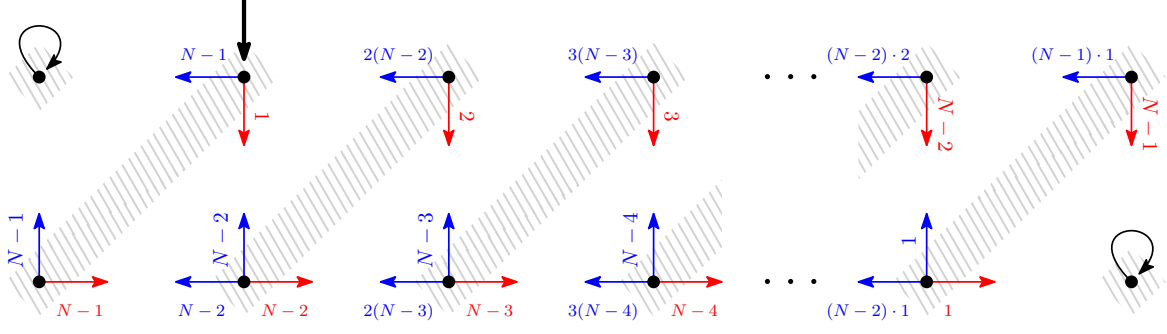


(a) $M_2^+$



(b) $M_2^-$
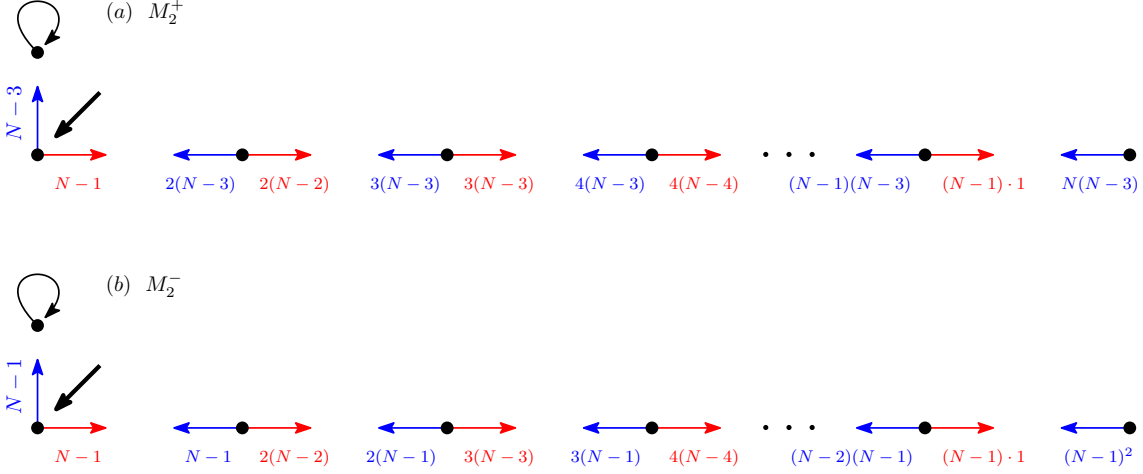
**Figure A5:** Markov chains $M_2^+$ and $M_2^-$. The transition probabilities are normalized by $N(N-1)$. By coupling, $\tau_2^+ \geq \tau_2 \geq \tau_2^-$.

$$\tau_2^- = N(N-1)\left(\frac{1}{N-1} + \frac{N-1}{2!(N-1)^2} + \frac{1 \cdot 2 \cdot (N-1)(N-2)}{3!(N-1)^3} + \cdots + \frac{1 \cdot 2 \cdots (N-1) \cdot (N-1)(N-2) \cdots 1}{N!(N-1)^N}\right)$$

$$\geq N\left(1 + \frac{1}{2} \cdot \frac{N-1}{N-1} + \frac{1}{3} \cdot \frac{(N-1)(N-2)}{(N-1)^2} + \cdots + \frac{1}{\sqrt{N}} \cdot \frac{(N-1)\ldots(N-\sqrt{N})}{(N-1)^{\sqrt{N}}}\right)$$

$$\geq N\left(1 + \frac{1}{2}\frac{N-1}{N-1} + \frac{1}{3}\frac{N-3}{N-1} + \cdots + \frac{1}{\sqrt{N}}\frac{N-\frac{1}{2}N}{N-1}\right) \geq N \cdot \frac{1}{2} \cdot H_{\sqrt{N}} \geq \frac{1}{4}N\log N = \Theta(N\log N),$$

where we only took the sum of the first $\sqrt{N}$ terms and used that $1 + 2 + \cdots + \sqrt{N} \approx \frac{1}{2}N$ and that $\log\sqrt{N} = \frac{1}{2}\log(N)$.

Recovery analysis for $L_4 = L_7$. The Markov chain $M_4 = M_7$ is depicted in Fig. A6.

• Recovery probability: $M_4$ has two absorbing states, $s_{0,0}$ and $s_{1,N-1}$, and both states can be reached with positive probability from any other state. Nevertheless, we show in the following that $\rho_4$ quickly

**Figure A6:** Markov chain $M_4 = M_7$. The transition probabilities are normalized by $N(N-1)$. Self-loops in non-absorbing states are not shown.
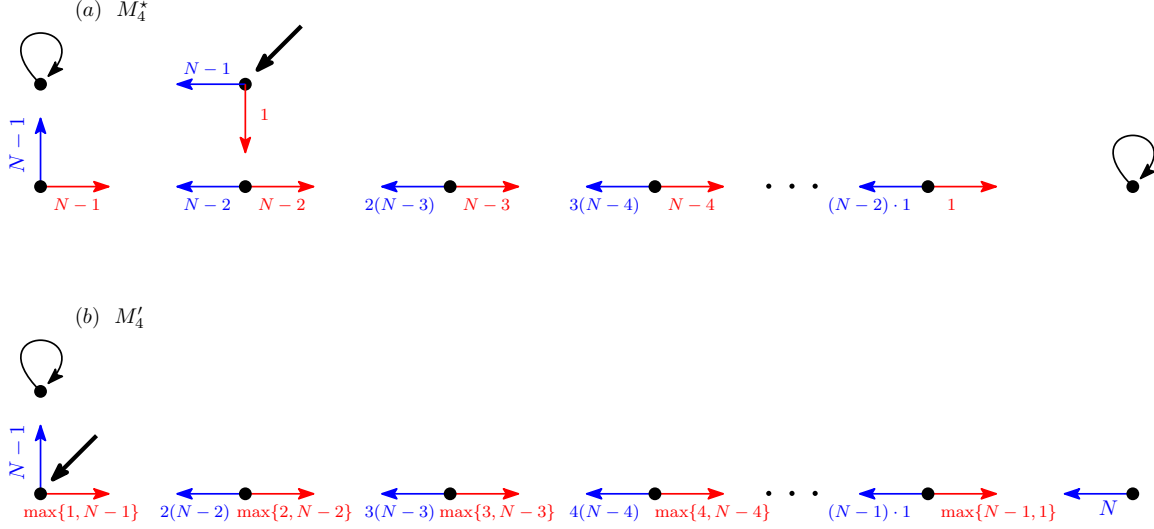


**Figure A7:** Markov chains $M_4^\star$ and $M_4'$. The transition probabilities are normalized by $N(N-1)$. By coupling, $\tau_4 \leq \tau_4'$.

approaches 1 as the population size increases, because $\rho_4 \geq 1 - 2/(N-1)!$.

To obtain this lower bound for $\rho_4$, we consider the Markov chain $M_4^\star$ that is obtained from $M_4$ by erasing states $s_{1,k}$ for $k < N-1$, and replacing the corresponding transitions of the form $s_{0,k} \to s_{1,k}$ by self-loops. Formally, we define $M_4^\star$ to be as in Fig. A7(a). For the coupling, note that any time a trace $T$ in $M_4$ takes a transition of the form $s_{0,k} \to s_{1,k}$, the corresponding trace $T^\star$ in $M_4^\star$ waits until (if ever) $T$ reaches state $s_{r,k}$ with $r = 0$ again. At that point, $T$ is to the left of $T^\star$ or in the same state. Hence if $T^\star$ ever reaches $s_{1,N-1}$ then so did $T$ and we have $\rho_4 \geq \rho_4'$.

Denoting by $x$ the probability of reaching $s_{0,0}$ from $s_{0,N-2}$ (before reaching $s_{1,N-1}$), we have $1 - \rho_4' = \frac{N-1}{N} \cdot 0 + \frac{1}{N} \cdot x = x/N$ and by Proposition 4 we compute

$$x = \frac{1+1}{1+1+1!+2!+3!+\cdots+(N-2)!} \leq \frac{2}{(N-2)!}$$

implying that $\rho_4 \geq \rho_4' = 1 - x/N \geq 1 - 2/(N-1)!$ as desired.

- Recovery time: since reaching $s_{1,N-1}$ requires taking at least one non-self-looping transition, we immediately get $\tau_4 \geq \frac{N(N-1)}{N-1+1} = N - 1$. In the following, we focus on the upper bound. To that end, we define a simpler Markov chain $M_4'$. Intuitively, $M_4'$ is obtained by identifying states $s_{1,k}$ and $s_{0,k+1}$ for each $k < N-2$. Since, for each $k = 1, \ldots, N-2$, the Markov chain $M_4$ has the property that

$$\mathbb{P}[s_{1,k-1} \to s_{1,k}] = \mathbb{P}[s_{0,k} \to s_{1,k}] + \mathbb{P}[s_{0,k} \to s_{0,k+1}],$$

the transition probabilities to the left are equal for each pair of identified states. To get an upper bound for $\tau_4$, for the transition probabilities to the right we take the larger of $\mathbb{P}[s_{1,k} \to s_{0,k}]$ and $\mathbb{P}[s_{0,k} \to s_{0,k-1}]$. Formally, we define $M_4'$ as in Fig. A7(b). Note that we have defined $\mathbb{P}[s_{0,0} \to s_{0,1}] = N$, such that $M_4'$ has only one absorbing state.

Next let us show that $\tau_4 \leq \tau_4'$. By construction, this is clear for traces that don't reach state $s_{0,0}$. For those that do, we need to show that the expected number $s$ of steps to reach $s_{1,N-1}$ from $s_{0,0}$ in $M_4'$ satisfies $s \geq \tau_4$. This follows from $\mathbb{P}[s_{0,0} \to s_{0,1}]$ in $M_4'$ is equal to $\mathbb{P}[s_{1,N-2} \to s_{1,N-1}] + \mathbb{P}[s_{1,N-2} \to s_{0,N-2}]$ in $M_4$: once the trace in $M_4'$ transitioned from $s_{0,0}$, the coupled trace in $M_4$ that started at $s_{1,N-2}$ either converged directly or transitioned to $s_{0,N-2}$ which is to the left of $s_{0,0}$.

Finally, we compute $\tau_4'$ using Eq. [23] again:

$$\tau_4' = N(N-1)\left(\frac{1}{N-1} + \frac{1}{2!(N-2)} + \frac{1}{3!(N-3)} + \cdots + \frac{1}{(N/2)!(N-N/2)} + \underbrace{X_1 + X_2 + \cdots + X_{N/2}}_{X_i \leq \frac{1}{(N/2)!}}\right)$$

$$\leq \frac{N(N-1)}{N/2}\left(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{(N/2)!}\right) + \underbrace{N(N-1) \cdot \frac{N}{2} \cdot \frac{1}{(N/2)!}}_{o(1)}$$

$$\leq 2(N-1) \cdot (e-1) + o(1) = \Theta(N).$$

where we used that each of the last $N/2$ terms in the first line is less than $1/(N/2)!$ and $\sum_{i=0}^{\infty} 1/i! = e$.

Recovery analysis for $L_6 = L_8$. The Markov chain $M_6 = M_8$ is depicted in Fig. A8.

- Recovery probability: Identifying the states $s_{1,k}$ and $s_{0,k+1}$ for $k < N-1$, we obtain an equivalent Markov chain $M_6'$ depicted in Fig. A9. Proposition 4 applied to $M_6'$ immediately implies

$$\rho_6 = 1 - \frac{1}{\underbrace{1 + 1 + \cdots + 1}_{N\times}} = 1 - \frac{1}{N}.$$

- Recovery time: For $M_6'$ we observe that $p_k^+ = p_k^- = \frac{k(N-k)}{N(N-1)}$. Thus, up to rescaling, the Markov chain $M_6'$ is equivalent to the Moran process under neutral drift (see e.g. 16, Section 4). Denoting by $t^a$ the expected time until the process in $M_6'$ reaches an absorbing state, and by $t^b$ the conditional time given

17
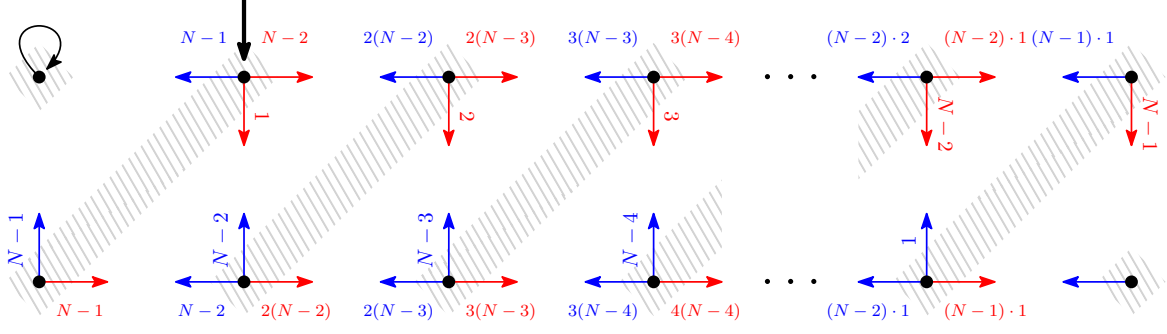
**Figure A8:** Markov chain $M_6 = M_8$. The transition probabilities are normalized by $N(N-1)$. Self-loops in non-absorbing states are not shown.
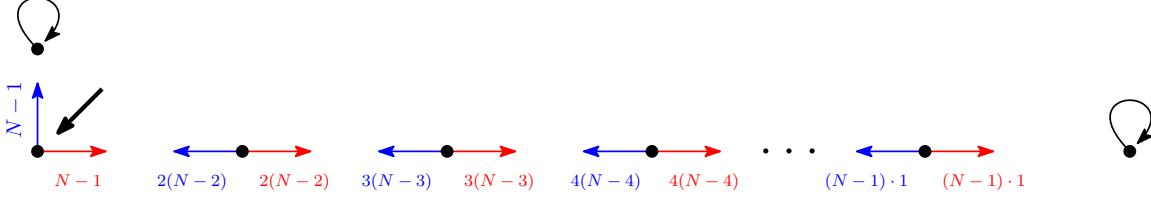


**Figure A9:** Markov chain $M_6'$ that is equivalent to $M_6 = M_8$. The transition probabilities are normalized by $N(N-1)$. We have $\rho_6 = \rho_6'$ and $\tau_6 = \tau_6'$.

that it reaches the non-recovering state $s_{0,0}$, equations [18] and [19] from (16) rescale to

$$t^a = (N-1) \cdot H_{N-1} \qquad \text{and} \qquad t^b = (N-1)^2.$$

The probability of absorbing at $s_{1,N-1}$ is $\rho_6 = 1 - \frac{1}{N}$. Therefore, $t^a = (1 - \frac{1}{N}) \cdot \tau_6 + \frac{1}{N} \cdot t^b$, or

$$\tau_6 = \frac{N \cdot t^a - t^b}{N-1} = N \cdot H_{N-1} - (N-1) = N \cdot H_N - N,$$

as desired.

$\square$

**MATLAB scripts**

In the following we provide the MATLAB scripts that have been used to generate the data shown in Figs. 1–4 of the main text, and Figs. S1–S9 in this Supporting Information.

Simulating the reputation dynamics for players with given strategies.

```
function [X,M,MEnd]=SimReputationDynamics(AssRule,ActRule,PopComp,ep,q,nIt);
% Simulates the reputation dynamics as described in Section 1.2.
```

18

```matlab
% X, Matrix containing the estimated pairwise cooperation frequencies
% M, Estimated average image matrix
% MEnd, Image Matrix in the end of the simulation run.
% AssRule, Matrix that contains all assessment rules present in population
% ActRule, Matrix that contains all action rules present in the population
% PopComp, Vector that specifies how many players use each strategy
% ep, Constant error probability to commit a perception error
% q, Constant probability to observe a third-party interaction
% nIt, number of iterations of the indirect reciprocity game.

%% Setting up the objects
N=sum(PopComp); % N, Size of population
nS=length(PopComp); % nS, Number of different strategies
X=zeros(nS,nS); M=zeros(N,N); MEnd=zeros(N,N); % Initializing the output
MC=ones(N,N); % Current image matrix; initially everyone is good
CP=zeros(N,N); % Matrix that counts all individual cooperation decisions
IN=zeros(N,N); % Matrix that counts all pairwise interactions
xP=ones(1,PopComp(1)); for j=2:nS, xP=[xP, j*ones(1,PopComp(j))]; end
    % N-dim vector.  The i-th entry is the strategy index of player i.
nCt=nIt/2; % initial transient time that is ignored for computing averages

%% Simulating the interactions
for t=1:nIt
 %% Choosing a donor and a recipient and letting them interact
 Do=randi(N); % Selecting a donor
 Re=Do; while Re==Do, Re=randi(N); end % Selecting a different receiver
 aD=ActRule(xP(Do),:); % Getting the donor's action rule
 stD=MC(Do,Do); stR=MC(Do,Re); % Defining the players' standings
 iA=(stD==0)*2+(stR==0)*1+1; cp=(rand(1)<aD(iA)); % cp=1 if donor cooperates

 %% Updating the donor's reputation
 for Obs=1:N % Going through all all individuals as potential observers
   if Obs==Do | Obs==Re | rand(1)<q % If individual observes interaction
     sO=AssRule(xP(Obs),:); % Getting the observer's assessment rule
     stD=MC(Obs,Do); stR=MC(Obs,Re); % Retrieving the players' standings
     if rand(1)<=1-ep % Image update without perception error
       iAs=4*(cp==0)+2*(stD==0)+1*(stR==0)+1; MC(Obs,Do)=sO(iAs);
     else % Image update with perception error
       iAs=4*(cp==1)+2*(stD==0)+1*(stR==0)+1; MC(Obs,Do)=sO(iAs);
     end
   end
 end
```

```matlab
  %% Updating the output matrices
  if t>nCt% After the initial transient time has passed
    CP(Do,Re)=CP(Do,Re)+cp; IN(Do,Re)=IN(Do,Re)+1; % Updating CP and IN
    if t==nCt+1 % Updating M
      M=MC;
    else
      M=(t-nCt-1)/(t-nCt)*M+1/(t-nCt)*MC;
    end
  end
end

%% Calculating output variables by averaging over all players with same strategy
MEnd=MC; % Image matrix in the end is the last current image matrix
Ms=M; M=-ones(nS,nS); % Storing M in Ms, re-initializing M
for i=1:nS % Going through all possible strategies
  for j=1:nS
    nD=find(xP==i); nR=find(xP==j); % Identifying the respective players
    if i~=j % Updating when the two strategies differ
      M(i,j)=mean(mean(Ms(nD,nR)));
      X(i,j)=sum(sum(CP(nD,nR)))/sum(sum(IN(nD,nR)));
    elseif PopComp(i)>1 % Updating within one strategy, excluding self-interactions
      M(i,j)=(sum(sum(Ms(nD,nR))) - sum(diag(Ms(nD,nR))))/(PopComp(i)*(PopComp(i)-1));
      X(i,j)=sum(sum(CP(nD,nR)))/sum(sum(IN(nD,nR)));
    end
  end
end
end
```

Calculating the selection-mutation equilibrium of the evolutionary dynamics for k=3 different strategies.

```matlab
function [SMEq,Coop,AvFr,Pop]=CalcSelMutEquilibrium(b,c,mu,s,nr);
% Calculates the selection-mutation equilibrium as specified in Section 2.2
% Input parameters:  b benefit of cooperation, c cost of cooperation,
% mu mutation rate, s strength of selection
% nr is a number in {1,..,8}, specifying which leading-eight strategy is considered.
% Output:  SMEq vector that gives frequency of each population state in equilibrium.
% Coop Resulting average cooperation rate over course of the game
% AvFr Abundance of each of the three strategies over course of the game
% Pop List of possible population states

% Considers the competition between ALLC, ALLD and an L8-strategy L_nr
% Requires as input how often each of theses strategies would...
% cooperate against each other, as estimated by SimReputationDynamics
```

```matlab
% This program assumes these estimates are stored in the data file 'Data L[nr].mat'
% This file contains the matrix Pop and the matrix CP.
% The matrix Pop has (N+2)(N+1)/2 rows and 3 columns; each row (n1,n2,n3)...
% gives a possible composition of the population.
% The matrix CP has (N+2)(N+1)/2 rows and 9 columns.  Each row again corresponds...
% to a possible population composition.  The first three columns represent how...
% often the L[nr] strategy cooperates against each other strategy; columns 4-6...
% represent the corresponding frequencies for ALLC and columns 7-9 give the...
% corresponding values for ALLD.

%% Preparations
DatName=['Data L',num2str(nr),'.mat']; load(DatName); % Loading simulation data.
nS=size(Pop,1); % Number of population states.
N=max(Pop(:,1)); % Population size.
[Pay,Coop]=CalcPay(Pop,CP,b,c,nS,N); % Calculates respective payoffs and...
     % cooperation rates in each possible population state.

%% Create the transition matrix W as specified in Section 2.2
W=zeros(nS,nS);
for iP=1:nS % Run over all possible previous states
 n=Pop(iP,:); % Previous population composition, n=[nL,nC,nD]

 % Going through all reachable population compositions
 wmp0= n(1)/N * (mu/2+(1-mu)*n(2)/(N-1)/(1+exp(-s*(Pay(iP,2)-Pay(iP,1)))));
   % Minus one L-player, plus one C-player
 wm0p= n(1)/N * (mu/2+(1-mu)*n(3)/(N-1)/(1+exp(-s*(Pay(iP,3)-Pay(iP,1)))));
   % Minus one L-player, plus one D-player
 wpm0= n(2)/N * (mu/2+(1-mu)*n(1)/(N-1)/(1+exp(-s*(Pay(iP,1)-Pay(iP,2)))));
   % Minus one C-player, plus one L-player
 w0mp= n(2)/N * (mu/2+(1-mu)*n(3)/(N-1)/(1+exp(-s*(Pay(iP,3)-Pay(iP,2)))));
   % Minus one C-player, plus one D-player
 wp0m= n(3)/N * (mu/2+(1-mu)*n(1)/(N-1)/(1+exp(-s*(Pay(iP,1)-Pay(iP,3)))));
   % Minus one D-player, plus one L-player
 w0pm= n(3)/N * (mu/2+(1-mu)*n(2)/(N-1)/(1+exp(-s*(Pay(iP,2)-Pay(iP,3)))));
   % Minus one D-player, plus one C-player

 % Inserting the values into the transition matrix
 if n(1)>0 % if the number of L[nr] individuals can further decrease
   np=[n(1)-1,n(2)+1,n(3)]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
   W(iP,iN)=wmp0;
   np=[n(1)-1,n(2),n(3)+1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
   W(iP,iN)=wm0p;
 end
 if n(2)>0 % if the number of ALLC individuals can further decrease
```

```
    np=[n(1)+1,n(2)-1,n(3)]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=wpm0;
    np=[n(1),  n(2)-1,n(3)+1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=w0mp;
  end
  if n(3)>0 % if the number of ALLD individuals can further decrease
    np=[n(1)+1,n(2),n(3)-1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=wp0m;
    np=[n(1),n(2)+1,n(3)-1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=w0pm;
  end
  W(iP,iP)=1-wmp0-wm0p-wpm0-w0mp-wp0m-w0pm;
end

v=null(W'-eye(nS,nS)); % Calculating left eigenvector of W with respect to EV 1
SMEq=v/sum(v); % Selection-mutation equilibrium is the normalized vector v
Coop=Coop*SMEq; % Average cooperation rates across all population states
AvFr=Pop'*SMEq/N; % Average abundance of each strategy across all states
end


function [Pay,Coop]=CalcPay(Pop,CP,b,c,nS,N);
% Subroutine that calculates the payoff for each of the three strategies...
% and the average cooperation rate for all possible population compositions.
% Required input:  Pop, Matrix of all possible population compositions
% CP, Matrix containing the estimated pairwise cooperation frequencies
% b,c, benefit and cost of cooperation, nS, number of population states
% N, population size.

Pay=zeros(nS,3); Coop=zeros(1,nS);
for st=1:nS % going through all possible population compositions
 MC=[CP(st,1:3); CP(st,4:6); CP(st,7:9)]; % Reconstructing the image matrix
  for i=1:3, for j=1:3, % Replacing irrelevant NaN's in image matrix
   if MC(i,j)>=0; else MC(i,j)=-1; end, end
  end
 n=Pop(st,:); cp=0; % n, current population state, cp, current cooperation rate
  for iS=1:3
   if n(iS)>0
     n1=n(1)-(iS==1); n2=n(2)-(iS==2); n3=n(3)-(iS==3);
      % Remaining population from perspective of player iS
     Pay(st,iS)=(MC(1,iS)*n1+MC(2,iS)*n2+MC(3,iS)*n3)/(N-1)*b - ...
      (MC(iS,1)*n1+MC(iS,2)*n2+MC(iS,3)*n3)/(N-1)*c;
      % Estimated payoffs according to Eq.[8]
     cp=cp+n(iS)/N*(MC(iS,1)*n1+MC(iS,2)*n2+MC(iS,3)*n3)/(N-1);
      % Updating the average cooperation rate    end
```

```
  end
 Coop(st)=cp;
end
end
```

Calculating the selection-mutation equilibrium of the evolutionary dynamics for rare mutations.

```
function [SMEq,Coop,FXP]=CalcEqRareMutations(b,c,s,N,ep,q,nr,nIt);
% Calculates the selection-mutation equilibrium as specified in Section 2.3
% Input parameters:  b benefit of cooperation, c cost of cooperation,
% s strength of selection, nr index of considered leading-eight strategy.
% N population size, ep error probability, q observation probability
% nr index of the considered leading-eight strategy % nIt number of rounds used
in SimReputationDynamics
% Output:  SMEq vector for frequency of each homogeneous population in equilibrium.
% Coop Resulting average cooperation rate
% AvFr Abundance of each of the three strategies
% FXP Matrix that contains all pairwise fixation probabilities

%% Obtain the pairwise cooperation probabilities using SimReputationDynamics
AssR=[1 1 1 1 0 1 0 0;
 1 0 1 1 0 1 0 0; % Consistent Standing
 1 1 1 1 0 1 0 1; % Simple Standing
 1 1 1 0 0 1 0 1;
 1 0 1 1 0 1 0 1;
 1 0 1 0 0 1 0 1; % Stern
 1 1 1 0 0 1 0 0; % Staying
 1 0 1 0 0 1 0 0]; % Judging
 % Assessment rules of leading-eight strategies
ActR=[1 0 1 1;
 1 0 1 1;
 1 0 1 0;
 1 0 1 0;
 1 0 1 0;
 1 0 1 0;
 1 0 1 0;
 1 0 1 0];
 % Action rules of leading-eight strategies
AssC=[1 1 1 1 1 1 1 1]; AssD=[0 0 0 0 0 0 0 0];
ActC=[1 1 1 1]; ActD=[0 0 0 0];
 % Assessment and action rules for ALLC and ALLD
[CPC,kvec]=GetPairwiseRepData([AssR(nr,:); AssC],[ActR(nr,:); ActC],N,ep,q,nIt);
```

```matlab
  % Getting estimated cooperation probabilities when L[nr] competes with ALLC
  % for each combination [N-k k] of L[nr] and ALLC players, with k given by kvec
[CPD,kvec]=GetPairwiseRepData([AssR(nr,:); AssD],[ActR(nr,:); ActD],N,ep,q,nIt);
  % Getting estimated cooperation probabilities when L[nr] competes with ALLD
FXP=zeros(3,3); k=kvec';

%% Calculating pairwise fixation probabilities:  L[nr] vs ALLC
PayL=(CPC(:,1).*(N-k-1)+CPC(:,3).*k)/(N-1)*b-(CPC(:,1).*(N-k-1)+CPC(:,2).*k)/(N-1)*c;
PayC=(CPC(:,2).*(N-k)+CPC(:,4).*(k-1))/(N-1)*b-(CPC(:,3).*(N-k)+CPC(:,4).*(k-1))/(N-1)*c;
  % Calculating the expected payoffs of the two strategies
RhoC=CalcFixProb(PayC,PayL,s);
RhoL=CalcFixProb(PayL(end:-1:1),PayC(end:-1:1),s);
FXP(1,2)=RhoC; FXP(2,1)=RhoL;
  % Calculating and storing fixation probabilities using a subroutine

%% Calculating pairwise fixation probabilities:  L[nr] vs ALLD
PayL=(CPD(:,1).*(N-k-1)+CPD(:,3).*k)/(N-1)*b-(CPD(:,1).*(N-k-1)+CPD(:,2).*k)/(N-1)*c;
PayD=(CPD(:,2).*(N-k)+CPD(:,4).*(k-1))/(N-1)*b-(CPD(:,3).*(N-k)+CPD(:,4).*(k-1))/(N-1)*c;
  % Calculating the expected payoffs of the two strategies
RhoD=CalcFixProb(PayD,PayL,s);
RhoL=CalcFixProb(PayL(end:-1:1),PayD(end:-1:1),s);
FXP(1,3)=RhoD; FXP(3,1)=RhoL;
  % Calculating and storing fixation probabilities using a subroutine

%% Calculating pairwise fixation probabilities:  ALLC vs ALLD
PayC=b*(k-1)/(N-1)-c; PayD=b*k/(N-1); % Does not require estimated payoffs
RhoC=CalcFixProb(PayC,PayD,s);
RhoD=CalcFixProb(PayD(end:-1:1),PayC(end:-1:1),s);
FXP(2,3)=RhoD; FXP(3,2)=RhoC;

%% Calculating the output
T=FXP/2; for i=1:3, T(i,i)=1-sum(T(i,:)); end % Constructing the transition matrix
v=null(T'-eye(3)); SMEq=v'/sum(v); % Calculating the selection-mutation equilibrium
CPL=SimReputationDynamics(AssR(nr,:),ActR(nr,:),N,ep,q,nIt); % Cooperation of
Leading-Eight against itself
Coop=SMEq*[CPL;1;0];
end


function pij=CalcFixProb(PayM,PayR,s);
% Calculates fixation probability according to Eq.[21]
z=exp(-s*(PayM-PayR));
pij=1/(sum(cumprod(z))+1);
end
```

```matlab
function [CP,kvec]=GetPairwiseRepData(AssRule,ActRule,N,ep,q,nIt);
% Considers population with 2 strategies only, AssRule and ActRule have two rows
% kvec=1:N-1 counts how many players with strategy 2 are in population
% CP is an (N-1)x4 dim matrix.  Each row corresponds to a population composition
% as defined through the vector k.  Each entry cp_ij of a row specifies
% how often an i player would cooperate with a j player, in the following
% format:  cp11, cp12, cp21, cp22.

kvec=1:N-1; CP=zeros(N-1,4);
for i=1:N-1
  k=kvec(i);
  PopComp=[N-k k];
  X=SimReputationDynamics(AssRule,ActRule,PopComp,ep,q,nIt);
  CP(i,:)=[X(1,1), X(1,2), X(2,1), X(2,2)];
end
end
```

# References

[1] Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).

[2] Sigmund, K. *The Calculus of Selfishness* (Princeton Univ. Press, 2010).

[3] Sigmund, K. Moral assessment in indirect reciprocity. *Journal of Theoretical Biology* **299**, 25–30 (2012).

[4] Ohtsuki, H. & Iwasa, Y. The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology* **239**, 435–444 (2006).

[5] Ohtsuki, H. & Iwasa, Y. How should we define goodness? – Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology* **231**, 107–20 (2004).

[6] Traulsen, A., Nowak, M. A. & Pacheco, J. M. Stochastic dynamics of invasion and fixation. *Physical Review E* **74**, 011909 (2006).

[7] Szabó, G. & Tőke, C. Evolutionary Prisoner's Dilemma game on a square lattice. *Physical Review E* **58**, 69–73 (1998).

[8] Wu, B., Bauer, B., Galla, T. & Traulsen, A. Fitness-based models and pairwise comparison models of evolutionary games are typically different—even in unstructured populations. *New Journal of Physics* **17**, 023043 (2015).

[9] Fudenberg, D. & Imhof, L. A. Imitation processes with small mutations. *Journal of Economic Theory* **131**, 251–262 (2006).

[10] Imhof, L. A., Fudenberg, D. & Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences USA* **102**, 10797–10800 (2005).

[11] Wu, B., Gokhale, C. S., Wang, L. & Traulsen, A. How small are small mutation rates? *Journal of Mathematical Biology* **64**, 803–827 (2012).

[12] Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).

[13] Traulsen, A. & Hauert, C. Stochastic evolutionary game dynamics. In Schuster, H. G. (ed.) *Reviews of Nonlinear Dynamics and Complexity*, 25–61 (Wiley-VCH, Weinheim, 2009).

[14] Karlin, S. & Taylor, H. M. A. *A First Course in Stochastic Processes* (Academic, London, 1975), 2nd edition edn.

[15] Cormen, T. H. *Introduction to algorithms* (MIT Press, Cambridge, MA, 2009).

[16] Altrock, P. M. & Traulsen, A. Fixation times in evolutionary games under weak selection. *New Journal of Physics* **11**, 013012 (2009).
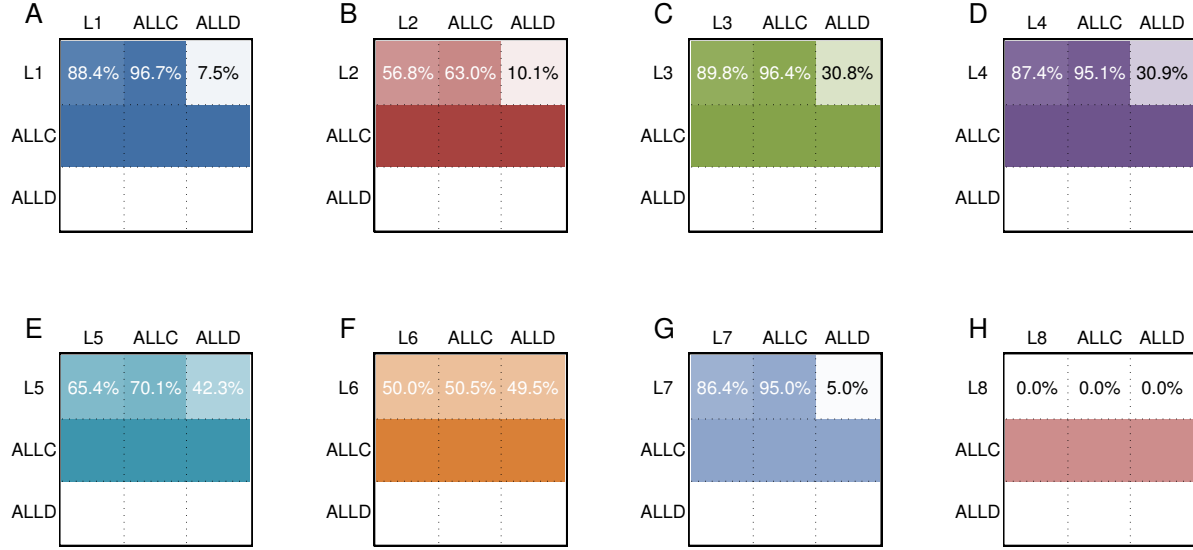
**Figure S1: The average images players have of each other converge over time.** While **Fig. 2** provides a snapshot of the reputation dynamics at a particular point in time, here we show how often players consider any other player as good on average. While ALLC players deem everyone as good and ALLD players deem everyone as bad, independent of the error rate, the leading-eight strategies differ in their assessments. Typically, a leading-eight player is most likely to consider ALLC players as good, followed by other leading-eight players and by ALLD. Whether the leading-eight strategy can resist invasion by ALLC or ALLD depends on the differences in the average image, and on the cost and benefit of cooperation. This figure uses the same parameters as in **Fig. 2**. The depicted numbers represent average values over the second half of $2 \cdot 10^6$ rounds of the game. Results for all strategies but L7 are robust with respect to starting from a different initial image matrix.
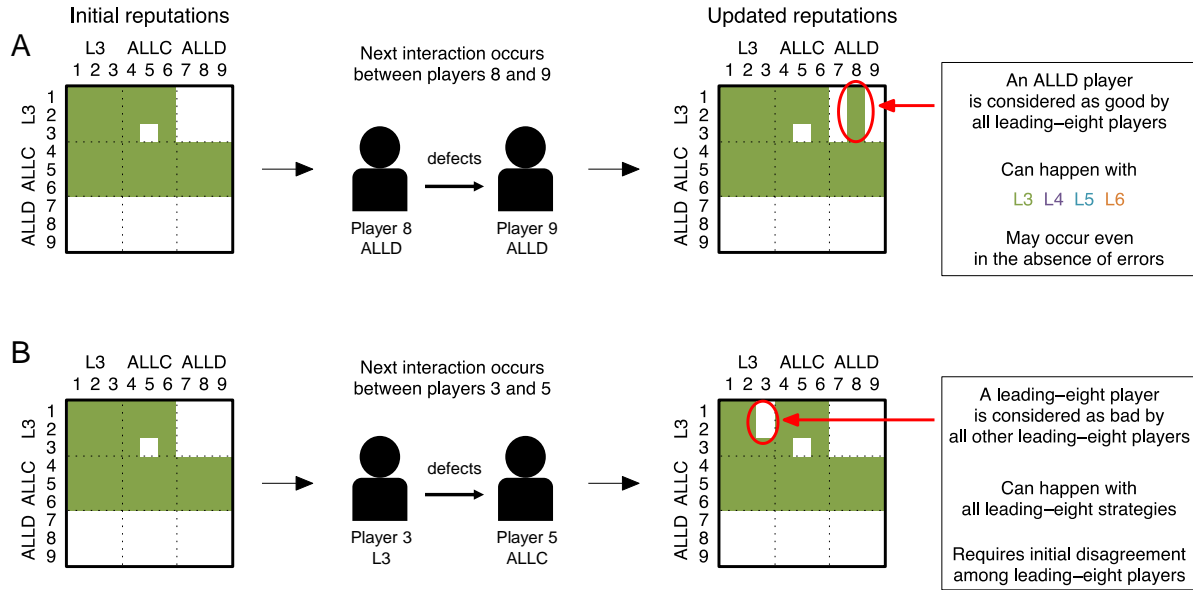
**Figure S2: The emergence of apparent misjudgments among leading-eight players.** As the game proceeds, it may occur that an ALLD player is considered as good by all leading-eight players (A), or that a leading-eight player is considered as bad by all other leading-eight players (B). Here we present stylized situations in which these cases occur. (A) For half of the leading-eight strategies (L3 − L6), an ALLD player can acquire a good reputation by defecting against another ALLD player. (B) When information is private and noisy, leading-eight strategies may disagree on the current reputation of a co-player. Here we show a case in which initially, player 3 is the only leading-eight player who deems player 5 as bad. If these two players are chosen for the next interaction, with player 3 as the donor, player 3 defects and acquires a bad reputation among her fellow leading-eight strategists.
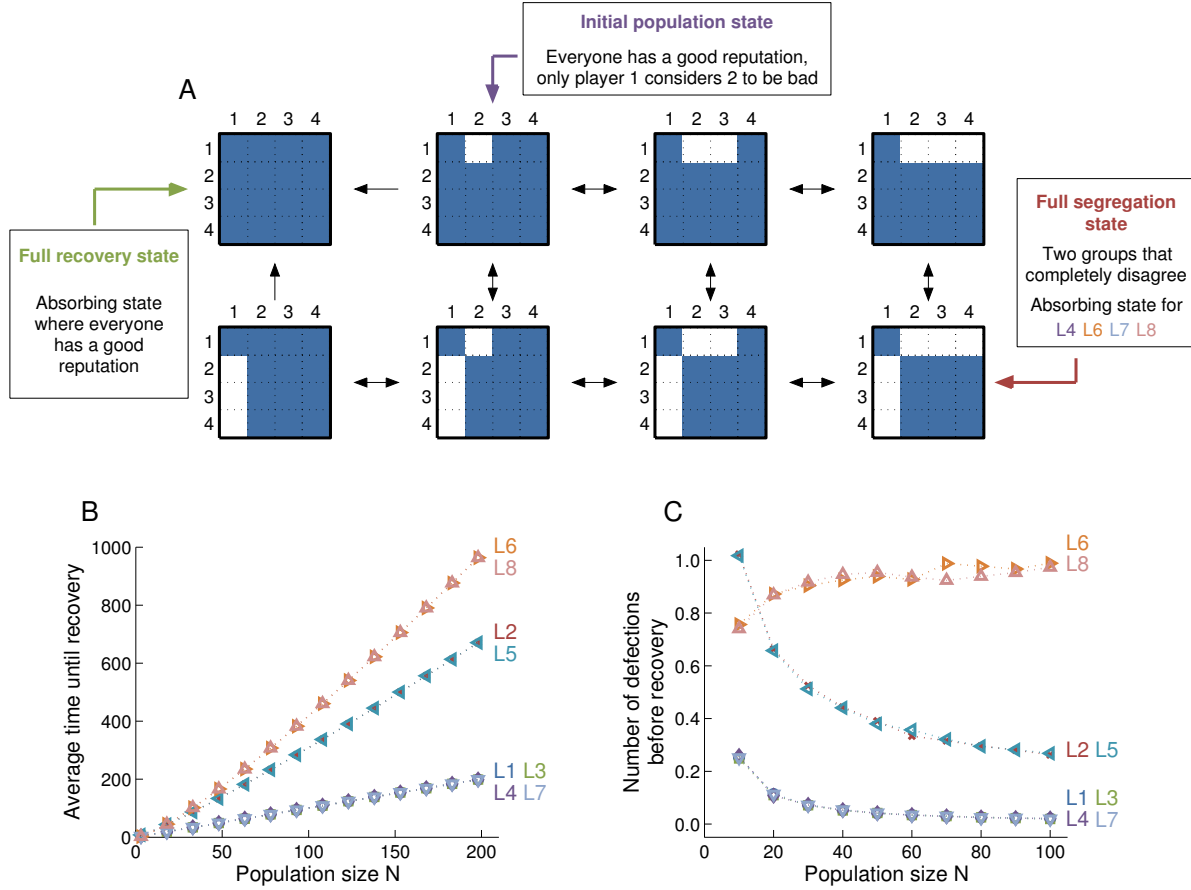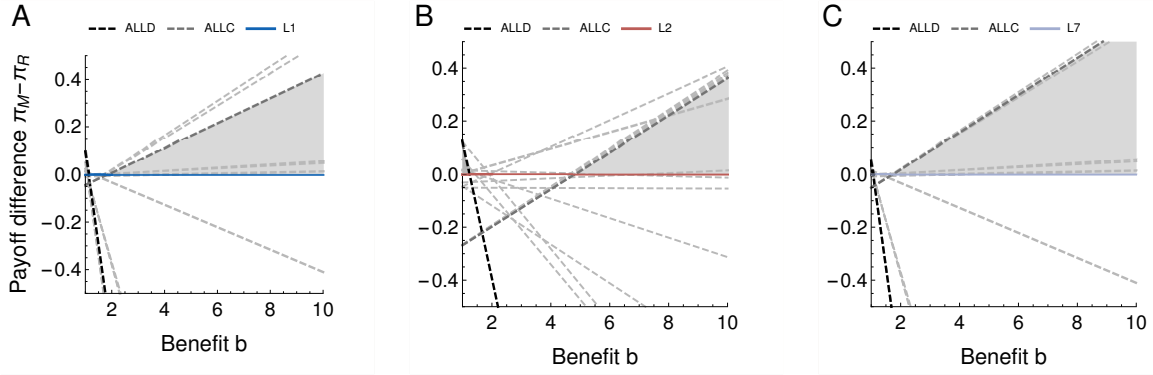
**Figure S3: The leading-eight strategies differ in their ability to recover from a single disagreements.** (A) We consider a homogeneous population in which everyone applies the same leading-eight strategy. Initially, all players are assumed to have a good reputation; only player 1 considers player 2 as bad. We derive analytical results for the reputation dynamics under the assumption that no more perception errors occur. We say the population recovers from a single disagreement, if it reaches the state where all players have a good reputation again. This is an absorbing state for all leading-eight strategies. We are interested in how likely the population recovers, and how long it takes until recovery. We find that only for four of the eight strategies, recovery is guaranteed (see SI Section 1.3). (B) With respect to the time until recovery, there are three cases. For four of the leading-eight strategies, recovery occurs quickly, and the recovery time is approximately linear. For all other strategies, including L2, we show that recovery may take substantially more time, being of order $N \log N$. (C) Also with respect to the expected number of defections until recovery, we observe three different qualitative behaviors. For four of the leading-eight strategies, a single perception error typically triggers no further defections, provided the population is sufficiently large. On the other extreme, for L6 and L8 we observe that any initial disagreement triggers on average one further defection.

29
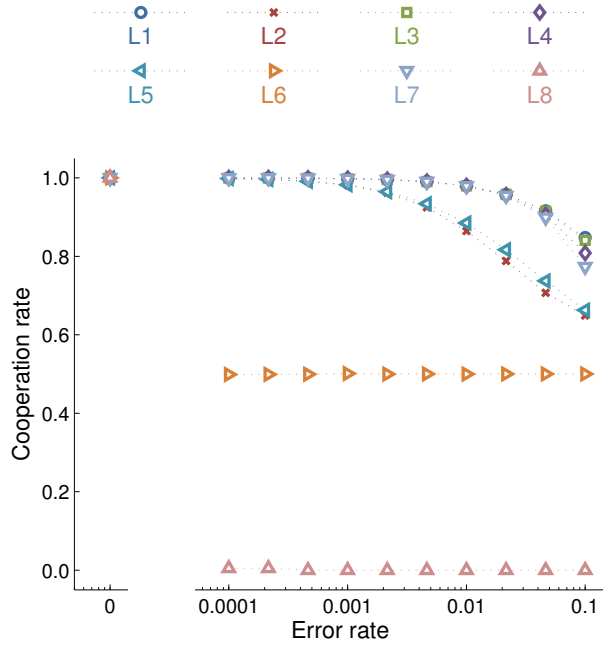
**Figure S4: Even rare perception errors render full cooperation impossible under "Stern" (L6) or "Judging" (L8).** We have simulated how often individuals cooperate if everyone in the population applies the same leading-eight strategy. The emerging cooperation rates depend on which leading-eight strategy is chosen and how often perception errors occur. For L6 and L8 we find that even rare perception errors undermine cooperation. Among the other six strategies, L2 and L5 are more susceptible to the noise introduced by perception errors than the other four strategies. Simulations were run with a population size of $N = 50$, assuming complete observation, $q = 1$.
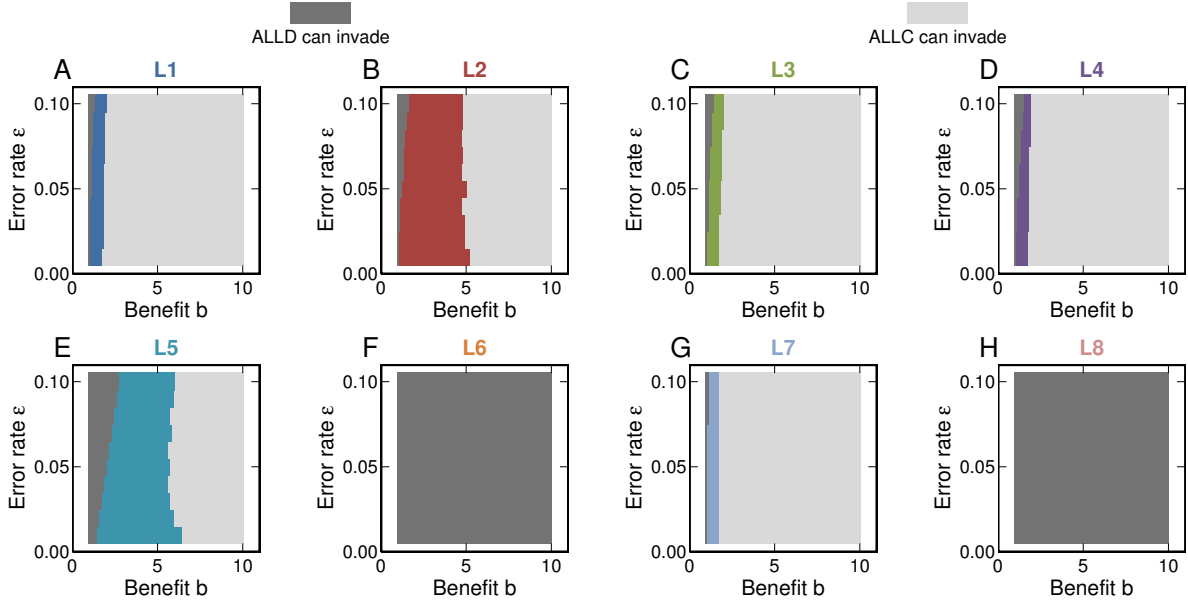
**Figure S5: Stability of each leading-eight strategy against invasion by ALLC and ALLD.** We consider a population of $N-1$ players adopting one of the leading-eight strategies, and we explore whether either a single ALLD mutant or an ALLC mutant gains at least the payoff of the residents. To this end, we vary the benefit of cooperation ($x$-axis) and the probability of perception errors ($y$-axis). Parameter regions in which ALLD can invade are depicted in dark grey, whereas parameter regions in which ALLC can invade are shown as light grey. Only in the colored region, the respective leading-eight strategy is stable against invasion by either ALLC or ALLD. Except for L2 and L5, we find that most leading-eight strategies are either unstable (F and H), or they only resist invasion in a small subset of the parameter space (A,C,D,G). Parameters: $N=50$, $c=1$, $q=0.9$.

**Figure S6: Stability against ALLC and ALLD does not imply stability against all mutant invasions.** We consider a resident population that either applies the leading-eight strategy L1 (A), L2 (B), or L7 (C). For each resident population, we consider 14 possible mutant strategies, including ALLD (black), ALLC (dark grey) and the twelve strategies that differ from the resident in only one bit (light grey). For different benefit values, we plot the payoff advantage $\pi_M - \pi_R$ of a single mutant among $N-1$ residents. If for a given $b$-value there is a line in the upper half of the panel, the resident strategy can be invaded by the respective mutant. We find that even in the parameter region where a leading-eight strategy can resist invasion by ALLC and ALLD, other mutant strategies may be able to invade. For example, L2 can resist invasion by ALLD for $b \gtrsim 1.5$, and invasion by ALLC if $b \lesssim 5$. In between, for $1.5 \leq b \leq 5$ there is a different mutant strategy that can invade. This mutant coincides with L2, except that it assesses a good donor who defects against a bad recipient as bad. Similar cases of successful mutants different from ALLC and ALLD also exist for L1 and L7. Parameters: $N = 50$, $c = 1$, $q = 0.9$, $\varepsilon = 0.05$.
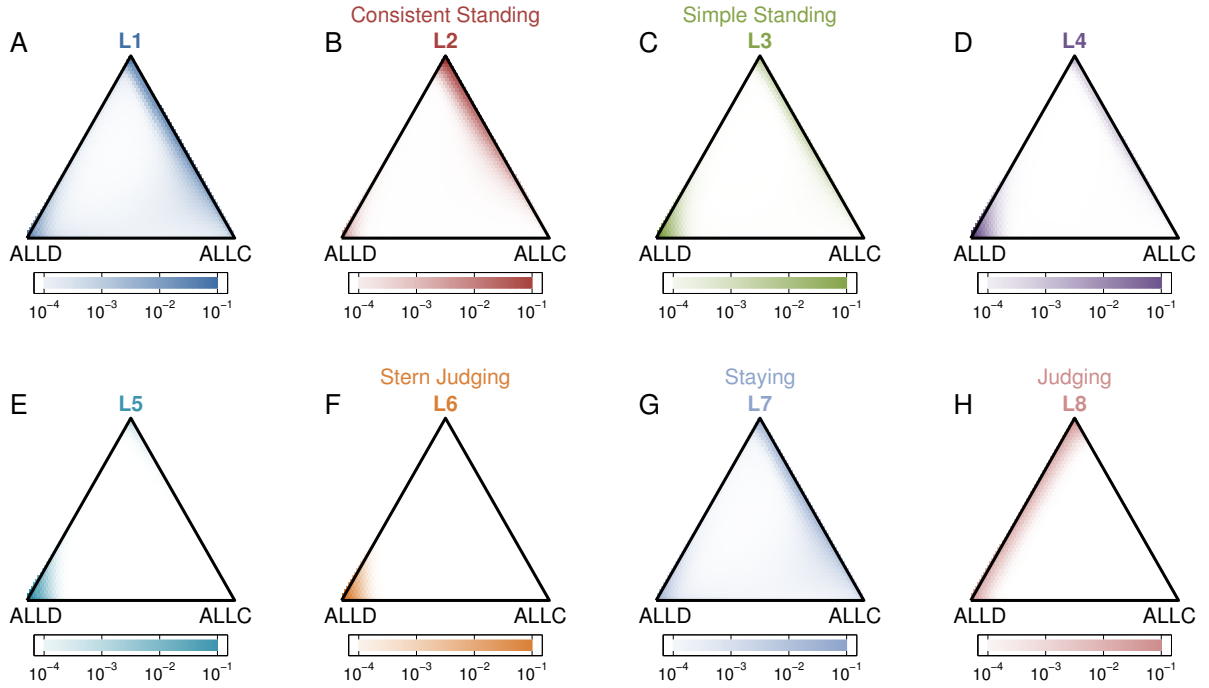
**Figure S7: When mutations are sufficiently frequent, three of the leading-eight strategies can maintain cooperation in coexistence with ALLC.** In the main text we have presented evolutionary results in the limit of rare mutations. In this limit, populations are homogeneous most of the time, rendering stable coexistences of multiple strategies impossible. Here, we present results for non-vanishing mutation rates when players can choose between a leading-eight strategy, ALLC, and ALLD. The possible population compositions are represented by a triangle. The corners of the triangle correspond to homogeneous populations, whereas interior points yield the corresponding mixed populations. The colors indicate how often the respective region of the state space is visited by the evolutionary process. (A,B,G) For the considered parameter values, we find that in three cases, a stable coexistence between a leading-eight strategy and ALLC can maintain cooperation for a considerable fraction of time. (C–F) In four cases, we find that populations typically find themselves in the vicinity of ALLD. (H) In the presence of noise, there is neutral drift between L8 and ALLD. Along that edge, no player cooperates. Parameters are the same as in **Fig. 3**, but with a strictly positive mutation rate $\mu = 0.01$.
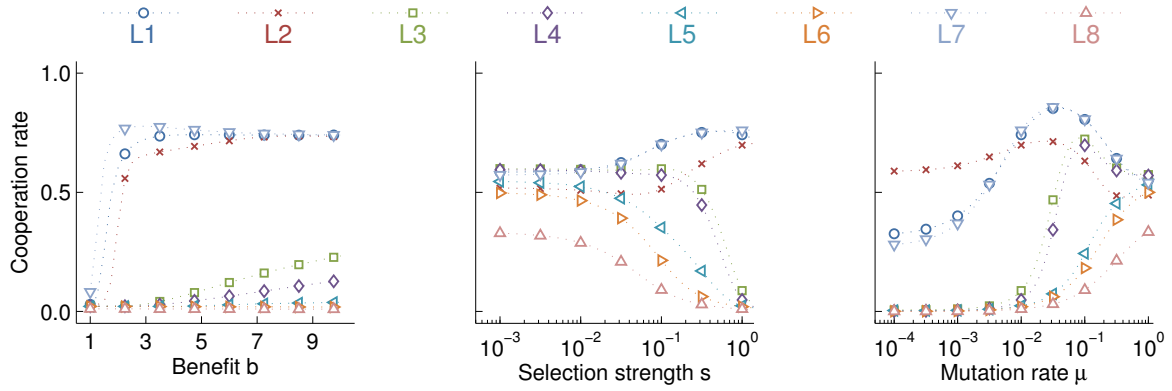
**Figure S8: Cooperation is most likely to evolve for high benefit-to-cost ratios and when mutations are sufficiently frequent.** To explore the robustness of our findings, we have systematically varied three key parameters of our model, the benefit $b$, the selection strength $s$, and the mutation rate $\mu$. For each combination of parameter values, we have simulated the evolutionary process between ALLC, ALLD, and $L_i$, for each of the leading-eight strategies $L_i$. The figure shows the resulting cooperation rates. We recover that for usual $b/c$-ratios, sufficiently strong selection, and rare mutations, there are only three leading-eight strategies that can maintain some cooperation. Surprisingly, the maximum amount of cooperation is achieved for intermediate mutation rates, $0.01 \le \mu \le 0.1$. Here we observe relatively stable coexistences between ALLC and either L1 or L7. Baseline parameters: $b = 5$, $s = 1$ and $\mu = 0.01$; all other parameters are the same as in **Fig. S7**.
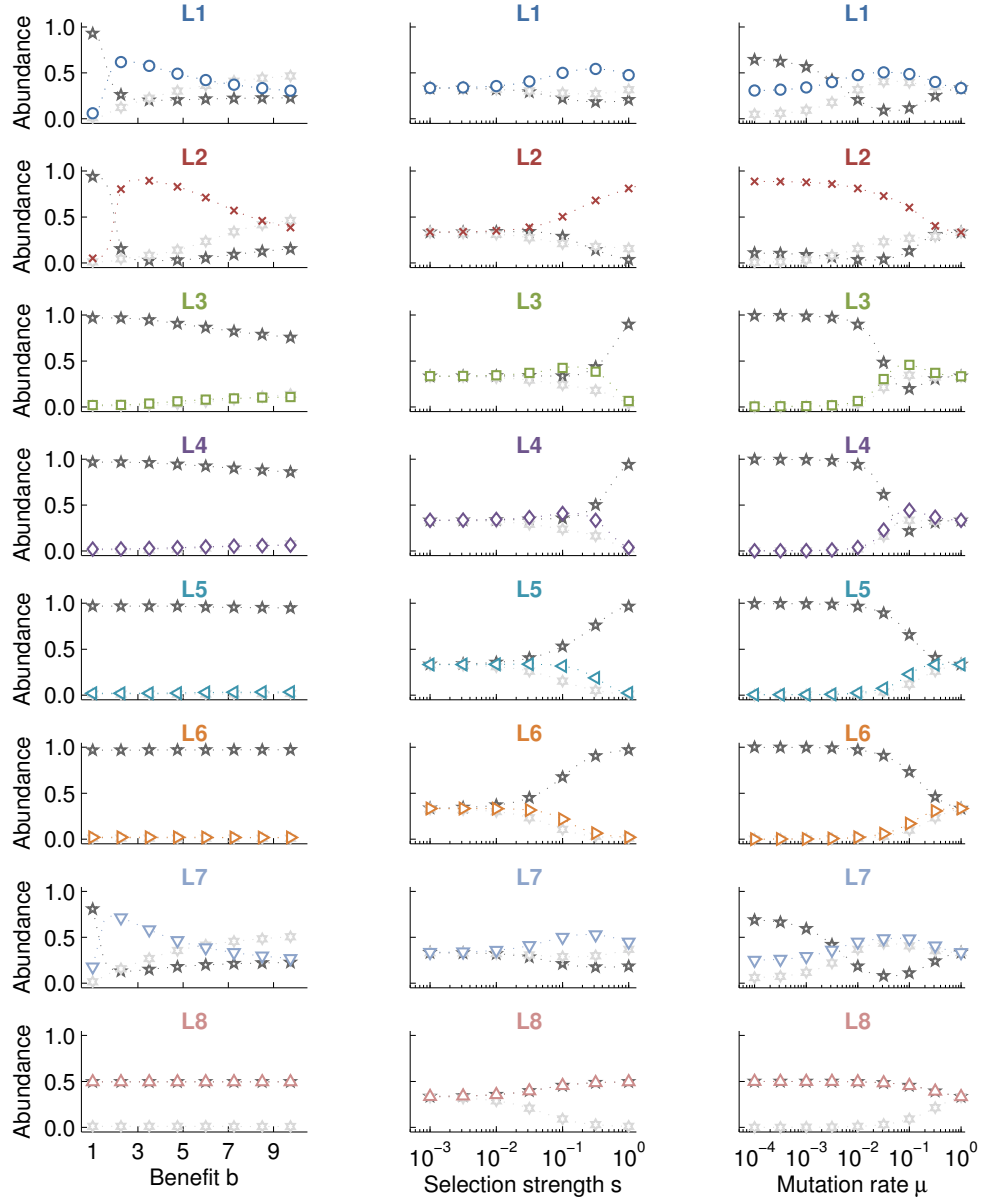
**Figure S9: Evolutionary abundance of the leading-eight strategies across different parameter regimes.** This figure considers the same scenario as **Fig. S8**, but it depicts the abundance of each strategy in the selection-mutation equilibrium. The abundance of ALLC is depicted in light grey, ALLD is shown in dark grey, and for the leading-eight strategy we use the respective color. In three cases (for L1, L2, L7) we observe that ALLC is played with positive frequency even as the selection strength $s$ increases. In these cases, ALLC typically coexists with a majority of players who apply the leading-eight strategy.