Collapse of cooperation in evolving games: Supplementary Information

Alexander J. Stewart[1], Joshua B. Plotkin[1,2]

[1] Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
[2] E-mail: jplotkin@sas.upenn.edu

1 Overview of Supplementary Information

In this supplement we prove that only three types of memory-1 strategies – self-alternators, self-cooperators, and self-defectors – can be evolutionary robust and prevalent in an arbitrary iterated two-player game with a 2×2 payoff matrix. We derive analytic expressions for the subsets of these strategy types that are evolutionary robust. We show that the volume of robust self-alternators, self-cooperators, and self-defectors provide good approximations for the time spent at each of these strategy types in an evolving population, for a fixed payoff matrix. As shown in Fig. 2 and Fig. 4 of the main text, this analysis enables us to explain strategy evolution in iterated two-player games even when payoff matrices are also allowed to evolve. In particular, this analysis predicts the collapse of cooperation in the Iterated Prisoner's Dilemma (Fig. 2), as well as the transition from the Iterated Prisoner's Dilemma to the Snowdrift or Staghunt game (Fig. 4).

We first define evolutionary robustness for an arbitrary iterated two-player game with a 2×2 payoff matrix in a well-mixed population of finite size N; and we state necessary and sufficient conditions for evolutionary robustness under the limits of either strong or weak selection on strategies. We then show that, under strong selection on strategies, only three subsets of memory-1 strategies can be evolutionary robust and prevalent: self-cooperate, self-defect and self-alternate. Within each of these three strategy types we derive the precise subset that is robust, and from this we calculate the volume of robust strategies of each type. Under weak selection, by contrast, we show that only the self-cooperate and self-alternate strategies can be evolutionary robust.

Finally we perform simulations under weak mutation for a variety of payoff mutation schemes, in addition to those used in the main text. These simulations demonstrate that the volume of robust strategies within each of these types continues to determine the outcome of payoff-strategy co-evolution in finite populations.

2 Iterated two-player games

We consider an iterated game with an infinite number of successive rounds between two players, X and Y. We study games with a 2×2 payoff matrix, so that in each round each player has two choices, denoted cooperate (c) or defect (d). The payoffs for the respective players are given in Table S1, in their most general form.

Table S1: Payoff matrix for an arbitrary 2×2 game

		Player Y	
		c	d
Player X	c	$R_{xy}(cc), R_{yx}(cc)$	$R_{xy}(cd), R_{yx}(dc)$
	d	$R_{xy}(dc), R_{yx}(cd)$	$R_{xy}(dd), R_{yx}(dd)$

2.1 Memory

In general, a player may have an arbitrarily long memory, such that her play in each round depends on the plays in all previous rounds of the current game. However, as per Press and Dyson [1], a player with memory-1 may treat all opponents as though they are also memory-1, regardless of the opponent's actual memory of the current game. And so the sets of scores s_{xx}, s_{xy} and s_{yx} for a player X with memory-1 facing an opponent Y with arbitrary memory can be understood by considering the scores received by X against an arbitrary memory-1 opponent instead. By contrast, the score a long-memory player Y received against himself, s_{yy}, may depend on his memory capacity. Nonetheless, since our results for strong selection do not depend on s_{yy}, we will show that a robust strategy for a memory-1 player X is robust against any opponent, regardless of his memory capacity. The same is true under weak selection: under weak selection, the robustness of a strategy may depend on s_{yy}, but nevertheless, as shown in [2] for the Prisoner's Dilemma, we can still derive conditions for robustness under weak selection that do not depend on the memory of Y. It is in this sense that our results on the evolutionary robustness of memory-1 strategies are without loss of generality – because such memory-1 strategies are evolutionary robust against all opponents, regardless of the opponent's memory capacity in the current game.

However, our results do not exclude the possibility that there exist long-memory strategies that are also evolutionary robust nor do they exclude the possibility that long memory strategies may gain an advantage under finite time horizons [3,4]. Furthermore, if players have a super long memory of previous interactions with different opponents they can "learn" the composition of the population and construct something similar to a "tag" based strategy [5–7] – that is, they can gain an advantage by only cooperating with players like themselves, while punishing players unlike themselves [5–7].

2.2 Equilibrium payoffs in Iterated Games

The longterm scores received by two memory-1 players in an iterated two-player game are calculated from the equilibrium rates of the different plays, (cc), (cd), (dc) and (dd), given by the stationary vector $\mathbf{v} = (v_{cc}, v_{cd}, v_{dc}, v_{dd})$ of the Markov matrix describing the iterated game [8]. The equilibrium score of player X against player Y is calculated according to

$$s_{xy} = \frac{\mathbf{v} \cdot \mathbf{R_{xy}}}{\mathbf{v} \cdot \mathbf{I}} = \frac{D\left(\mathbf{p_x}, \mathbf{p_y}, \mathbf{R_{xy}}\right)}{D\left(\mathbf{p_x}, \mathbf{q_y}, \mathbf{I}\right)} \tag{1}$$

where $\mathbf{I} = (1,1,1,1)$, $\mathbf{p_x}$ and $\mathbf{q_y}$ are the strategies of players X and Y, and $\mathbf{R_{xy}} = (R_{xy}(cc), R_{xy}(cd), R_{xy}(dc), R_{xy}(dd))$ is the payoff vector of player X against player Y, where the plays (cc), (cd) etc list the play of X first and the play of Y second. The determinant $D(\mathbf{p_x}, \mathbf{q_y}, \mathbf{f})$ gives the dot product between the stationary vector $\mathbf{v}$ and an arbitrary vector $\mathbf{f} = (f_{cc}, f_{cd}, f_{dc}, f_{dd})$ [1], where

$$D(\mathbf{p_x}, \mathbf{q_y}, \mathbf{f}) = \det \begin{bmatrix} -1 + p_{cc}q_{cc} & -1 + p_{cc} & -1 + q_{cc} & f_{cc} \\ p_{cd}q_{dc} & -1 + p_{cd} & q_{dc} & f_{cd} \\ p_{dc}q_{cd} & p_{dc} & -1 + q_{cd} & f_{dc} \\ p_{dd}q_{dd} & p_{dd} & q_{dd} & f_{dd} \end{bmatrix}. \tag{2}$$

In general Eq. 1 is sufficient to calculate the scores received by a pair of memory-1 players. However, there are certain pathological cases in which the Markov chain describing the iterated game has multiple absorbing states. The scores in these cases can be calculated by assuming that players execute their strategy with some small "error rate" ϵ [9], so that the probability of cooperation is at most $1 - \epsilon$ and at least ϵ. Assuming this, and taking the limit $\epsilon \to 0$ then gives the player's scores in the cases where multiple absorbing states exist.

2.3 Akin coordinate system

As shown by Akin in [8], manipulations of Eq. 1 produce an alternate coordinate system for the four-dimensional space of memory-1 strategies, useful for analysing the outcomes of the iterated Prisoner's Dilemma. In particular, for an arbitrary 2×2, two-player game, we can convert from the basis $(p_{cc}, p_{cd}, p_{dc}, p_{dd})$ to the basis $(\phi, \chi, \kappa, \lambda)$ [1, 2, 8], where the two coordinate systems are related by

$$\tilde{\mathbf{p}}_\mathbf{x} = -\phi \left[\mathbf{R}^\dagger_\mathbf{yx} - \chi\mathbf{R_{xy}} - (1-\chi)\kappa\mathbf{I} + \lambda\mathbf{L}\right]. \tag{3}$$

Here $\tilde{\mathbf{p}}_\mathbf{x} = (-1+p_{cc}, -1+p_{cd}, p_{dc}, p_{dd})$, $\mathbf{I} = (1,1,1,1)$, $\mathbf{L} = (0,1,1,0)$ and $\mathbf{R}^\dagger_\mathbf{yx} = (R_{yx}(cc), R_{yx}(dc), R_{yx}(cd), R_{yx}(dd))$, which is the payoff vector of Y when X is the focal player [1]. To convert directly between the two coordinate systems we have the equations

$$\begin{aligned} p_{cc} &= 1 - \phi\left(R_{yx}(cc) - \chi R_{xy}(cc) - (1-\chi)\kappa\right) \\ p_{cd} &= 1 - \phi\left(R_{yx}(dc) - \chi R_{xy}(cd) - (1-\chi)\kappa + \lambda\right) \\ p_{dc} &= \phi\left(\chi R_{xy}(dc) - R_{yx}(cd) + (1-\chi)\kappa - \lambda\right) \\ p_{dd} &= \phi\left((1-\chi)\kappa - R_{yx}(dd) + \chi R_{xy}(dd)\right). \end{aligned}$$

When considering the evolution of payoffs in general we must use the coordinate transform above for the general case in which $\mathbf{R_{xy}} \neq \mathbf{R_{yx}}$ (see below). When considering the evolution of strategies, we will be concerned with monomorphic populations in which $\mathbf{R_{xy}} = \mathbf{R_{yx}} = \mathbf{R}$, such that

$$\begin{aligned}
p_{cc} &= 1 - \phi(1-\chi)\left(R(cc) - \kappa\right) \\
p_{cd} &= 1 - \phi\left(R(dc) - \chi R(cd) - (1-\chi)\kappa + \lambda\right) \\
p_{dc} &= \phi\left(\chi R(dc) - R(cd) + (1-\chi)\kappa - \lambda\right) \\
p_{dd} &= \phi(1-\chi)\left(\kappa - R(dd)\right).
\end{aligned} \tag{4}$$

In this coordinate scheme the players' scores are related by [2,8]

$$s_{yx} - \chi s_{xy} - (1-\chi)\kappa + \lambda(v_{cd} + v_{dc}) = 0. \tag{5}$$

which, unlike ZD strategies, depends on the equilibrium rate of playing (cd) and (dc). The Akin coordinate system can be used to analyse arbitrary 2×2 player games of this type, unless $R(cc) = R(dc)$ or $R(cd) = R(dd)$. These two special cases are discussed separately, following our treatment of the more general case in which $R(cc) \neq R(dd)$ and $R(cd) \neq R(dc)$.

Finally, we note from Eq. 3, that in the limit $\phi \to 0$ (i.e when X tries to set her own score [1]) we are left with

$$\begin{aligned}
p_{cc} &= 1 + \phi\chi R_{xy}(cc) + \phi(1-\chi)\kappa \\
p_{cd} &= 1 + \phi\chi R_{xy}(cd) + \phi(1-\chi)\kappa - \phi\lambda \\
p_{dc} &= \phi\chi R_{xy}(dc) + \phi(1-\chi)\kappa - \phi\lambda \\
p_{dd} &= \phi(1-\chi)\kappa + \phi\chi R_{xy}(dd)
\end{aligned}$$

To produce a viable strategy in this limit we require $\phi\chi \to 0$ and $\phi(1-\chi)\kappa \to 0$ to ensure $p_{cc} \in [0,1]$ and $p_{dd} \in [0,1]$. This in turn implies $\phi\lambda \to 0$ to ensure $p_{cd} \in [0,1]$ and $p_{dc} \in [0,1]$, and we are left with the singular strategy $(1,1,0,0)$ [1] known as Repeat [8], regardless of the choice of payoffs.

2.4 Useful inequalities

In addition to the relationship Eq. 5 we make note of four inequalities which we will use to determine the memory-1 strategies that are evolutionary robust. We begin by noting that the equilibrium payoff for X playing against an opponent Y is given by [8]

$$s_{xy} = R(cc)v_{cc} + R(cd)v_{cd} + R(dc)v_{dc} + R(dd)v_{dd} \tag{6}$$

(i) From Eq. 6, the difference between the two players' scores can be written as

$$s_{xy} - s_{yx} = (v_{dc} - v_{cd})(R(dc) - R(cd))$$

which gives

$$s_{xy} - s_{yx} \leq (v_{cd} + v_{dc})|R(dc) - R(cd)| \tag{7}$$

where equality is achieved by an opponent Y for whom $v_{cd} = 0$ (e.g. an opponent who always cooperates).

(ii) Similarly, we must have

$$s_{xy} - s_{yx} \geq -(v_{dc} + v_{cd})|R(dc) - R(cd)| \tag{8}$$

where equality is achieved by an opponent Y for whom $v_{dc} = 0$ (e.g. an opponent who never cooperates).

(iii) From Eq. 6, the sum of the two players' scores is

$$s_{xy} + s_{yx} = 2(v_{cc} + (v_{dc} + v_{cd}))(R(cc) - R(dd)) - (v_{dc} + v_{cd})(2R(cc) - (R(cd) + R(dc))) + 2R(dd)$$

and, since $v_{cc} + (v_{dc} + v_{cd}) \leq 1$, we have

$$s_{xy} + s_{yx} \leq 2R(cc) - (v_{dc} + v_{cd})(2R(cc) - (R(cd) + R(dc))) \tag{9}$$

where equality is achieved when $v_{dd} = 0$ (e.g. by an opponent who never defects once they have been defected against).

(iv) Finally, we also have

$$s_{xy} + s_{yx} \geq 2R(dd) - (v_{dc} + v_{cd})(2R(dd) - (R(cd) + R(dc))) \tag{10}$$

where equality is achieved when $v_{cc} = 0$ (e.g. by an opponent who always defects once they have been cooperated with).

2.5 Evolution in a population of players

We study evolution in a well-mixed, finite population of N haploid, memory-1 players. Evolution is described by the "imitation" process of [10]. Under this model, which is similar to the Moran process, pairs of individuals, X and Y, are drawn randomly from a population of size N at each time step. Player X adopts the strategy of player Y with a probability $(1 + \exp[\sigma(s_x - s_y)])^{-1}$ that depends on their respective total payoffs, s_x and s_y, summed across pairwise matchups with all players in the population. Here σ denotes the strength of selection.

We study evolution in the limit of weak mutation. This means that, at any point in time, the population is monomorphic for some payoff matrix $\mathbf{R} = (R(cc), R(cd), R(dc), R(dd))$ and some strategy $\mathbf{p} = (p_{cc}, p_{cd}, p_{dc}, p_{dd})$. Given a population monomorphic for the resident type X, a mutation producing type Y will fix with probability [10]

$$\rho(X,Y) = \left(1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} e^{-\sigma[(j-1)s_{yy} + (N-j)s_{yx} - js_{xy} - (N-j-1)s_{xx}]}\right)^{-1},$$

or otherwise will be lost.

The "strong-selection" limit of this process is defined by taking $N \to \infty$ while keeping σ fixed, so that even small differences in the longterm payoff of the iterated game convey a significant advantage to the player with the greater payoff. Alternatively, the "weak-selection" limit arises when $N\sigma \ll 1$, by taking a Taylor expansion of the fixation probability $\rho(X,Y)$ and neglecting terms $O\left((N\sigma)^2\right)$ and larger, in which case even deleterious strategies may reach high frequency through genetic drift. We consider both of these regimes of selection in our analyses below.

2.5.1 Mutations affecting payoffs

In addition to mutations to strategies, we also consider mutations that affect the payoffs recieved by a mutant and by her opponents. When such mutations occur, the payoff vectors for a given game may depend on the genotypes of both players, so that $\mathbf{R_{xy}} \neq \mathbf{R_{yx}}$. Under the public goods game that we study in the main text, a resident genotype X pays a cost C_x when she cooperates. We assume that a cost C_x generates a benefit B_x when both players cooperate or a payoff $2\alpha_x B_x$ when the focal player cooperates alone. In the simplest form of the public goods game, when $\alpha_x = 1/2$, the amount of benefit produced does not depend on how the opponent behaves. We assume that B_x satisfies $B_x - C_x = f(C_x)$ for some function $f(C)$, which is typically chosen to be monotonically increasing with C. Given the choice of function $f(C)$, the genotype of player X is comprised of her strategy vector $\mathbf{p}$, her contributed cost C_X, and her synergy factor α_X. In general, when two players who contribute different costs to the public good face each other we assume that they equally share the payoffs for mutual cooperation, so that the payoff vectors for X and Y have the form

$$
\begin{aligned}
\mathbf{R_{xx}} &= \Big(f(C_x), \alpha_x f(C_x) - (1-\alpha_x)C_x, \alpha_x f(C_x) + \alpha_x C_x, 0\Big) \\
\mathbf{R_{xy}} &= \left(\frac{f(C_x)+f(C_y)}{2}, \alpha_x f(C_x) - (1-\alpha_x)C_x, \alpha_y f(C_y) + \alpha_y C_y, 0\right) \\
\mathbf{R_{yx}} &= \left(\frac{f(C_x)+f(C_y)}{2}, \alpha_y f(C_y) - (1-\alpha_y)C_y, \alpha_x f(C_x) + \alpha_x C_x, 0\right) \\
\mathbf{R_{yy}} &= \Big(f(C_y), \alpha_y f(C_y) - (1-\alpha_y)C_y, \alpha_y f(C_y) + \alpha_y C_y, 0\Big)
\end{aligned}
$$

In any given matchup between players X and Y the resulting longterm scores in the iterated game, s_{xx}, s_{xy}, s_{yx} and s_{yy} can be calculated from Eq. 1 and the fixation probability of Y can be calculated as described above. In particular, mutations that increase the benefit for mutual cooperation are favored by selection in a population of self-cooperators.

2.6 Evolutionary robustness of strategies

We will use the above relations to determine which strategies are evolutionary robust in a population of N players.

The concept of evolutionary robustness [2] is similar to the notion of evolutionary stability [11,12]. An evolutionary stable strategy $\mathbf{p_x}$ is one that satisfies either $s_{xx} > s_{yx}$, or else $s_{xx} = s_{yx}$ and $s_{xy} > s_{yy}$, for all opponents $\mathbf{p_y} \neq \mathbf{p_x}$ [11,12]. This means that a strategy is evolutionary stable provided (i) it cannot be selectively invaded by any other strategy ($s_{xx} > s_{yx}$), or (ii) it can selectively invade ($s_{xy} > s_{yy}$) any strategy that can neutrally invade it ($s_{xx} = s_{yx}$). However, as shown in [2,8], evolutionary stable strategies rarely exist within the full space of memory-1 strategies, because many strategies can neutrally invade each other. Therefore, we analyze the outcomes of evolution in a population using the notion of evolutionary robustness [2].

In general, a strategy is defined to be evolutionary robust if, when resident in a population, there is no mutant strategy that is favored to spread by natural selection when rare [2]. In particular, under strong selection a strategy X is evolutionary robust iff, when resident in a population of size N, it cannot be selectively invaded by any mutant strategy Y – that is, iff $s_{xx} \geq s_{xy}$ for all Y. The condition for evolutionary robustness under strong selection is thus identical to that of a Nash equilibrium [8]. Under weak selection, by contrast, a resident strategy X is evolutionary robust iff the fixation probability of any new mutant strategy Y satisfies $\rho_{yx} \leq 1/N$ (see [2]). As shown previously [2, 8], evolutionary robustness, as opposed to evolutionary stability, is useful for characterizing the strategies that dominate in evolving populations. In the remainder of the supplement we first derive results for evolutionary robustness under strong selection, which are used in the main text. We then derive conditions for evolutionary robustness under weak selection.

3 Necessary conditions for memory-1 strategies to be robust under strong selection

We start by proving that strategies in the interior of the four-dimensional memory-1 strategy space cannot be evolutionary robust – that is, they can always be selectively invaded by some other strategy. In fact, we will show that nothing on the interior can be robust with the exception of the "equalizers", which have $\chi = \lambda = 0$, and always set their opponent's score to $s_{yx} = \kappa$ (the special case of the equalizers is treated separately below). To prove that no strategy in the interior of memory-1 strategy space (except the equalizers) can be robust under strong selection (i.e a Nash equilibrium), we show that any such strategy can by invaded by a one of its "neighbors" in the alternate coordinate system $(\phi, \chi, \kappa, \lambda)$.

First, consider a resident strategy X characterised by $(\phi_x, \chi_x, \kappa_x, \lambda_x)$ and a mutant strategy Y characterised by $(\phi_y, \chi_y, \kappa_y, \lambda_y)$. From Eq. 5, with $y = x$, we find that the payoff of the resident against itself is

$$s_{xx} = \kappa_x - \frac{\lambda_x}{1-\chi_x}(v_{cd} + v_{dc}).$$

Similarly, from Eq. 5 we find that the payoff of Y against X is

$$s_{yx} = \frac{(1-\chi_x)\kappa_x - \lambda_x(w_{cd} + w_{dc}) + \chi_x((1-\chi_y)\kappa_y - \lambda_y(w_{cd} + w_{dc}))}{(1-\chi_x\chi_y)}$$

where $\mathbf{v}$ is the stationary vector for X playing against itself and $\mathbf{w}$ is the stationary vector for X playing against Y.

3.1 A strategy with $0 \le p_{cc} < 1$ and $0 < p_{dd} \le 1$ can always be invaded unless $p_{cd} = 0$ and $p_{dc} = 1$

First suppose that X satisfies $0 < p_{cc} < 1$ and $0 < p_{dd} < 1$. Recall that, in order for Y to selectively invade X under strong selection requires $s_{yx} > s_{xx}$.

Suppose Y is chosen such that $\chi_x = \chi_y$ and $\phi_x = \phi_y$. Finally, choose λ_y such that $(1-\chi_x)\kappa_x - \lambda_x = (1-\chi_y)\kappa_y - \lambda_y$ i.e. so that p_{cd} and p_{dc} are unaltered by the mutation. This gives

$$s_{yx} = \frac{(\kappa_x + \chi_x\kappa_y)}{1+\chi_x} - \frac{\lambda_x}{1-\chi_x}(w_{cd} + w_{dc}) - \frac{\chi_x(\kappa_y - \kappa_x)}{(1+\chi_x)}(w_{cd} + w_{dc})$$

We can then write

$$s_{yx} - s_{xx} = \frac{\chi_x(\kappa_y - \kappa_x)}{1+\chi_x}\left[1 - (w_{cd} + w_{dc})\right] + \frac{\lambda_x}{1-\chi_x}\left[(v_{cd} + v_{dc}) - (w_{cd} + w_{dc})\right]$$

and so Y selectively invades X iff

$$\frac{\chi_x(\kappa_y - \kappa_x)}{1+\chi_x}\left[1 - (w_{cd} + w_{dc})\right] > \frac{\lambda_x}{1-\chi_x}\left[(w_{cd} + w_{dc}) - (v_{cd} + v_{dc})\right]$$

This inequality can always be satisfied unless $v_{cd} + v_{dc} = 1$, in which case both sides vanish and the mutation is neutral. To see this, we use Eq. 2 to calculate

$$v_{cd}+v_{dc} = \frac{2(1-p_{cc})(1+p_{cc}-p_{dd})p_{dd}}{(1-p_{cc})((1-p_{cd}-p_{dc})(1+p_{cc})+2p_{cd}p_{dc})+2(1-p_{cc}^2+p_{cd}p_{dc})p_{dd}-(1-2p_{cc}+p_{cd}+p_{dc})p_{dd}^2}$$

where we note that $v_{cd}+v_{dc}=1$ implies $p_{dc}=1-p_{cd}$ and $p_{cd}p_{dc}=0$, so that either $p_{cd}=1$ and $p_{dc}=0$ or vice versa. We now assume that the mutant is such that $\kappa_y=\kappa_x+\eta$ where η is small. We can then write

$$(w_{cd}+w_{dc})-(v_{cd}+v_{dc}) = (v_{cd}+v_{dc})\frac{\eta(1-(p_{cc}+p_{dd}))(1+p_{cc}-p_{dd})}{2(1-p_{cc})(1+p_{cc}-p_{dd})p_{dd}} - \frac{\eta(1-(p_{cc}+p_{dd})+(p_{cc}-p_{dd})(p_{cd}+p_{dc}-(p_{cc}+p_{dd})))}{2(1-p_{cc})(1+p_{cc}-p_{dd})p_{dd}} + O\left(\eta^2\right)$$

or, more conveniently

$$(w_{cd}+w_{dc})-(v_{cd}+v_{dc}) = A\eta$$

where A depends on the resident strategy and is zero when $v_{cd}+v_{dc}=1$. The condition for invasion of X by Y then becomes

$$\frac{\chi_x}{1+\chi_x}\left[1-(v_{cd}+v_{dc})\right]\eta > \frac{\lambda_x}{1-\chi_x}A\eta$$

This inequality permits the following cases:

- If $\chi_x \neq 1$, $v_{cd}+v_{dc}\neq 1$ and either $|\chi_x|>0$ or $|\lambda_x|>0$ the inequality can always be satisfied (since the sign of η can always be reversed by taking $\kappa_y>\kappa_x$ or $\kappa_y<\kappa_x$).
- If $\chi_x=1$, this implies $p_{cc}=1$ and $p_{dd}=0$, which violates our assumption that $0\leq p_{cc}<1$ and $0<p_{dd}\leq 1$ (see below)
- If $\chi_x=\lambda_x=0$ the strategy X is an equalizer and all mutants are neutral (we discuss the equalizers further below)
- If $v_{cd}+v_{dc}=1$ then $p_{cd}=0$ and $p_{dc}=1$ or else $p_{cd}=1$ and $p_{dc}=0$, which implies $A=0$, and both sides of the inequality vanish.

The final two cases consist of $p_{cd}=0$ and $p_{dc}=1$ or $p_{cd}=1$ and $p_{dc}=0$. When these strategies play against themselves, the play is always *cd*. We call strategies with $p_{cd}=0$ and $p_{dc}=1$ "self-alternating" strategies since, when playing against themselves, the play alternates between *cd* and *dc*. Self-alternating strategies can be robust (see below)

Strategies of the form $p_{cd}=1$ and $p_{dc}=0$ cannot be robust: if we assume, without loss of generality, that $R(dc)>R(cd)$, then an opponent with $p_{dc}=0$ and $p_{cd}<1$ scores $R(dc)$ at equilibrium, whereas such an alternator scores $(1/2)(R(cd)+R(dc))$ against itself. Therefore the mutant can selectively invade and this strategy type cannot be robust. The only exception occurs in the special case $R(cd)=R(dc)$, which we discuss further below).

Note that the above expressions hold for $0<p_{cc}<1$ and $0<p_{dd}<1$. When $p_{cc}=0$ or $p_{dd}=1$, changing κ_x so that $\kappa_y=\kappa_x+\eta$ necessitates that we also decrease ϕ_x so that $\phi_y=(1-\eta^*)\phi_x$, where η^* must be chosen so that a viable strategy is produced when $\eta<0$. In this case it is simple to show that $(w_{cd}+w_{dc})-(v_{cd}+v_{dc})=A^*\eta$, and the same argument holds as for mutations that change κ_x alone. Thus strategies with $p_{cc}=0$ or $p_{dd}=1$ can be invaded under the same conditions as strategies with $0<p_{cc}<1$ and $0<p_{dd}<1$.

3.2 A strategy with $0 < p_{cd} < 1$ and $0 < p_{dc} < 1$ can always be invaded unless $p_{cc} = 1$ or $p_{dd} = 0$

Now suppose $0 < p_{cd} < 1$ and $0 < p_{dc} < 1$. We consider a mutant Y such that $\kappa_x = \kappa_y$, $\chi_x = \chi_y$ and $\phi_x = \phi_y$. This has the effect that p_{cc} and p_{dd} remain constant under mutation. We then have

$$s_{yx} = \kappa_x - \frac{(\lambda_x + \chi_x\lambda_y)(w_{cd} + w_{dc})}{(1 - \chi_x^2)}$$

and Y can selectively invade iff

$$\frac{\lambda_x}{1-\chi_x}\left[(v_{cd} + v_{dc}) - (w_{cd} + w_{dc})\right] > \frac{\chi_x}{1-\chi_x}\left[\lambda_y(w_{cd} + w_{dc}) - \lambda_x(v_{cd} + v_{dc})\right]$$

This inequality can always be satisfied unless $v_{cd} + v_{dc} = 0$, in which case both sides vanish and the mutation is neutral. To see this, we once again write

$$v_{cd}+v_{dc} = \frac{2(1-p_{cc})(1+p_{cc}-p_{dd})p_{dd}}{(1-p_{cc})((1-p_{cd}-p_{dc})(1+p_{cc})+2p_{cd}p_{dc})+2(1-p_{cc}^2+p_{cd}p_{dc})p_{dd}-(1-2p_{cc}+p_{cd}+p_{dc})p_{dd}^2}$$

and note that $v_{cd} + v_{dc} = 0$ implies either $p_{cc} = 1$, $p_{dd} = 0$ or else $p_{cc} = 0$ and $p_{dd} = 1$. As in the previous case, we assume that the mutant is such that $\lambda_y = \lambda_x + \eta$ where η is small. We can then write

$$(w_{cd} + w_{dc}) - (v_{cd} + v_{dc}) =$$
$$(v_{cd}+v_{dc})\frac{\eta(1-p_{cd}-p_{dc}-(p_{cc}-p_{dd})(p_{cc}-p_{cd}-p_{dc}+p_{dd}))}{(1-p_{cc})((1-p_{cd}-p_{dc})(1+p_{cc})+2p_{cd}p_{dc})+2(1-p_{cc}^2+p_{cd}p_{dc})p_{dd}-(1-2p_{cc}+p_{cd}+p_{dc})p_{dd}^2} + O\left(\eta^2\right)$$

or, more conveniently

$$(w_{cd} + w_{dc}) - (v_{cd} + v_{dc}) = A\eta$$

where A depends on the resident strategy and is zero when $v_{cd} + v_{dc} = 0$. The condition for invasion of X by Y then becomes

$$-\lambda_x\frac{1+\chi_x}{1-\chi_x}A\eta > \frac{\chi_x}{1-\chi_x}\eta(v_{cd} + v_{dc})$$

This inequality permits the following cases:

- If $\chi_x \neq 1$, $v_{cd} + v_{dc} \neq 0$ and either $|\chi_x| > 0$ or $|\lambda_x| > 0$ the inequality can always be satisfied (since the sign of η can always be reversed by taking $\lambda_y > \lambda_x$ or $\lambda_y < \lambda_x$).
- If $\chi_x = 1$, this implies $p_{cc} = 1$ and $p_{dd} = 0$, in which case the strategy has (at least) two absorbing states when playing against itself. We treat this case separately below.
- If $\chi_x = \lambda_x = 0$ the strategy X is an equalizer and all mutants are neutral (we discuss the equalizers further below)
- If $v_{cd} + v_{dc} = 0$ then $p_{cc} = 1$, $p_{dd} = 0$ or else $p_{cc} = 0$ and $p_{dd} = 1$, which implies $A = 0$, and both sides of the inequality vanish.

The final three cases consist of $p_{cc} = 1$, $p_{dd} = 0$ or $p_{cc} = 0$ and $p_{dd} = 1$. We call $p_{cc} = 1$ and $p_{dd} \neq 0$ a "self-cooperating" strategy since it always cooperates when playing itself. These strategies can be robust (see below). We call $p_{cc} \neq 1$ and $p_{dd} = 0$ a "self-defecting" strategy since it always defects when playing itself. These strategies can be robust (see below).

Strategies with $p_{cc} = 0$ and $p_{dd} = 1$ cannot be robust: as shown in the previous section, such strategies are vulnerable to mutations that change κ_x and ϕ_x simultaneously, and can always be invaded.

3.3 Strategies belonging to multiple types can only be robust in the absence of noise

It now remains to consider strategies that simultaneously satisfy more than one of the necessary conditions for robustness outlined above. When each strategy of a single type plays against itself, the associated Markov chain describing the iterated game has either a single absorbing state (in the case of self-cooperate or self-defect), or else displays stable periodic behavior (in the case of self-alternating strategies of either type). When a strategy belonging to multiple types plays against itself, there exist multiple absorbing states (or stable periodic behaviors). For example, tit-for-tat is a memory-1 strategy of the form $(1, 0, 1, 0)$ and is therefore a self-cooperator, a self-defector and a self-alternator. Such strategies admit two possible behaviors.

- In the presence of a small amount of "noise" [9], such that players execute their strategy with some small error rate, the score received by a strategy against itself is the average of the scores received by each of the possible strategy-types it belongs to.
- In the absence of noise, the equilibrium play arrived is determined by the play in the first round (e.g if tit-for-tat players always cooperate in the first round, they self-cooperate at equilibrium).

In the first case (with noise) a strategy with multiple types cannot be robust in general, unless it plays like a single type against itself. To see this, suppose that the equilibrium payoffs received by strategy type i against itself is R_i^*. If a strategy belongs to k types, the equilibrium score a strategy receives against itself is

$$s_{xx} = \sum_{i=1}^{k} \tau_i R_i^*$$

where $\tau_i > 0$ is the proportion of time spent in state i at equilibrium. Now consider a mutant Y which shares only a single type j with resident strategy X, such that $R_j^* > R_i^*$, $\forall i \neq j$. If there is no other absorbing state for the Markov chain describing the iterated game between X and Y, then he resulting score is

$$s_{yx} = R_j^*$$

which means $s_{yx} > s_{xx}$ and X can always be invaded. However, an exception occurs in the special payoffs for all of the types to which X belongs are exactly equal. Note also that if $\tau_i = 0$ for all but one type, the strategy behaves as though it belongs to a single type, and it may therefore be robust. It therefore remains for us to calculate τ_i for the four cases of strategies that belong to multiple types in 2×2 two-player games. There four cases are as follows:

- $\mathbf{p} = (1, p_{cd}, p_{dc}, 0)$ with $p_{cd} > 0$ and $p_{dc} < 1$ (self-cooperate AND self-defect)
- $\mathbf{p} = (1, 0, 1, p_{dd})$ or $\mathbf{p} = (1, 1, 0, p_{dd})$ with $p_{dd} > 0$ (self-cooperate AND self-alternate)
- $\mathbf{p} = (p_{cc}, 0, 1, 0)$ or $\mathbf{p} = (p_{cc}, 1, 0, 0)$ with $p_{cc} < 1$(self-defect AND self-alternate)

- $\mathbf{p} = (1, 0, 1, 0)$ of $\mathbf{p} = (1, 1, 0, 0)$ (self-cooperate AND self-defect AND self-alternate)

In the case of self-alternating strategies above we consider for completeness both those strategies with $p_{cd} = 0, p_{dc} = 1$ and those strategies with $p_{cd} = 1, p_{dc} = 0$. The times spent at each of the three possible equilibria are $\tau_c = v_{cc}$, $\tau_d = v_{dd}$ and $\tau_a = v_{cd} + v_{dc}$. Assuming a small amount of noise ϵ in the execution aof a particular play, these times can be calculated for each of the four cases from Eq. 2, to give the following:

3.3.1 Equilibrium play against self for $\mathbf{p} = (1, p_{cd}, p_{dc}, 0)$

$$\begin{aligned}
\tau_c &= \frac{p_{cd}p_{dc}}{1 - p_{cd} - p_{dc} + 2p_{cd}p_{dc}} + O(\epsilon) \\
\tau_d &= \frac{1 - p_{cd} - p_{dc} + p_{cd}p_{dc}}{1 - p_{cd} - p_{dc} + 2p_{cd}p_{dc}} + O(\epsilon) \\
\tau_a &= O(\epsilon) .
\end{aligned}$$

If $p_{cd} > 0$ and $p_{dc} > 0$ we have $\tau_c > 0$ and $\tau_d > 0$ and the arguement given in section 3.3 above holds. To see this, choose a mutant of the form $\mathbf{q} = (1, p_{cd}, p_{dc}, q_{dd})$. Calculating the proportion of time spent cooperating between the resident and such a mutant we find $\tau_c = 1 - O(\epsilon)$, so that the mutant recieves the payoff for mutual cooperation and it can selectively invade.

However, if $p_{cd} = 0$ or $p_{dc} = 0$, and $\epsilon \to 0$ we find $\tau_d \to 1$. These strategies behave like self-defectors, and they must be analysed as such. The robustness of such strategies can be determined from the robsutness conditions given in the main text (and below Eq. 12). These conditions reveal that such strategies can be robust only if $p_{dc} = 0$ and $\frac{R(dd)-R(cd)}{R(dc)-Rdd} > 0$.

3.3.2 Equilibrium play against self for $\mathbf{p} = (1, 0, 1, p_{dd})$ or $\mathbf{p} = (1, 1, 0, p_{dd})$

Both strategies of these forms have the same equilibrium rates of play, given by

$$\begin{aligned}
\tau_c &= \frac{1}{3 - p_{dd}} + O(\epsilon) \\
\tau_d &= O(\epsilon) \\
\tau_a &= \frac{2 - p_{dd}}{3 - p_{dd}} + O(\epsilon) .
\end{aligned}$$

In this case, regardless of p_{dd}, we have $\tau_c > 0$ and $\tau_a > 0$, and it is simple to verify that such a strategy can always be selectively invaded in the way described above in section 3.3, by choosing a mutant of the form $\mathbf{q} = (1, q_{cd}, q_{dc}, p_{dd})$ if $2R(cc) > R(cd) + R(dc)$, or by choosing $\mathbf{q} = (q_{cc}, 0, 1, p_{dd})$ otherwise.

3.3.3 Equilibrium play against self for $\mathbf{p} = (p_{cc}, 0, 1, 0)$ or $\mathbf{p} = (p_{cc}, 1, 0, 0)$

Both strategies of these form have the same equilibrium rates of play, given by

$$\begin{aligned}\tau_c &= O(\epsilon)\\ \tau_d &= \frac{1}{2+p_{cc}}+O(\epsilon)\\ \tau_a &= \frac{1+p_{cc}}{2+p_{cc}}+O(\epsilon)\,.\end{aligned}$$

In this case, regardless of p_{cc}, we have $\tau_d > 0$ and $\tau_a > 0$, and it is simple to verify that such a strategy can always be selectively invaded in the way described above in section 3.3, by choosing a mutant $\mathbf{q} = (p_{cc}, 0, 1, q_{dd})$.

3.3.4 Equilibrium play against self for $\mathbf{p} = (1, 0, 1, 0)$ or $\mathbf{p} = (1, 1, 0, 0)$

Both of these form of strategies have the same equilibrium rates of play, given by

$$\begin{aligned}\tau_c &= \frac{1}{4}\\ \tau_d &= \frac{1}{2}\\ \tau_a &= \frac{1}{4}.\end{aligned}$$

This strategy is tit-for-tat. It has $\tau_c > 0$, $\tau_d > 0$ and $\tau_a > 0$, and it is simple to verify that such a strategy can always be selectively invaded in the way described above in section 3.3, by choosing a mutant $\mathbf{q} = (1, q_{cd}, q_{dc}, q_{dd})$ if $2R(cc) > R(cd) + R(dc)$, or by choosing $\mathbf{q} = (q_{cc}, 0, 1, q_{dd})$ otherwise

Finally, in the case that there is no noise in the execution of each play, a strategy with multiple types belongs to the class determined by its initial play. Strategies belonging to multiple types can ensure they always receive the payoff of a particular type, j, by choosing their initial play appropriately. To ensure mutual cooperation for example, players belonging to multiple types (including the self-cooperators) must always cooperate on the first move. In the case of alternators, ensuring opposite play in the opening moves of the game may require more complex communication between players, but is nonetheless possible in principle. In the case of no noise, a strategy belonging to multiple types can be robust provided it receives the maximum payoff among the types it belongs to, $R_j^* > R_i^*$, $\forall i \neq j$, when playing against itself. If it does not receive the maximum payoff, the strategy can be invaded just as described when noise is present. If the payoff is maximized, the strategy will be robust provided it also satisfies the robustness conditions described below.

4 Necessary and sufficient conditions for evolutionary robustness under strong selection

As discussed above, a memory-1 strategy that is evolutionary robust in an arbitrary two-player, 2 × 2 game must belong to one of the following four types:

- the self-cooperators $\mathcal{C} = \{(p_{cc}, p_{cd}, p_{dc}, p_{dd})\,|p_{cc} = 1\}$,
- the self-defectors $\mathcal{D} = \{(p_{cc}, p_{cd}, p_{dc}, p_{dd})\,|p_{dd} = 0\}$,

- the self-alternators $\mathcal{A} = \{(p_{cc}, p_{cd}, p_{dc}, p_{dd})\, | p_{cd} = 0,\ p_{dc} = 1\}$,
- the equalizers $\mathcal{E} = \{(\phi, \chi, \kappa, \lambda) | \lambda = \chi = 0\}$.

We now derive sufficient conditions for strategies of each of these types to be robust.

4.1 Self-cooperators with $\lambda \geq -\chi(R(cc) - (R(cd) + R(dc)))$ and $\lambda \geq -\chi|R(dc) - R(cd)|$ are robust

The self-cooperators $\mathcal{C}$ satisfy $p_{cc} = 1$ and score $s_{xx} = R(cc)$ against themselves, which corresponds to $\kappa = R(cc)$. In the context of the Iterated Prisoner's Dilemma, these are precisely the "Good" strategies of [8] and discussed it [2]. In order to invade a resident strategy X, a mutant Y must have

$$s_{yx} > R(cc).$$

From Eq. 5, this condition becomes

$$\chi s_{xy} > \chi R(cc) + \lambda(v_{cd} + v_{dc})$$

We can also combine Eq. 5 with Eq 9 (where the extremal case of Eq. 9 corresponds here to an opponent who plays $(0, 1, 1, 1)$, so that $v_{dd} = 0$) to give

$$s_{xy}(1 + \chi) \leq (1 + \chi)R(cc) - (v_{cd} + v_{dc})(2R(cc) - (R(cd) + R(dc)) + \lambda)$$

and we therefore find that Y can invade X if

$$-\chi(R(cc) - (R(cd) + R(dc)) > \lambda$$

Similarly, we can combine Eq 5 with Eq 8 (where the extremal case of Eq. 8 corresponds here to an opponent who plays $(0, 0, 0, 0)$, so that $v_{dc} = 0$) to give

$$-\chi|R(dc) - R(cd)| > \lambda$$

as a condition for Y to invade X. Converting back to our original coordinate system, this implies a self-cooperator X is robust iff:

$$\begin{aligned} p_{dc}(R(dc) - R(cc)) &\leq (R(cc) - R(cd))(1 - p_{cd}) \\ p_{dd}(R(dc) - R(cc)) &\leq (R(cc) - R(dd))(1 - p_{cd}). \end{aligned} \tag{11}$$

These expressions are given in [8] for the self-cooperators under the Prisoner's Dilemma.

The evolutionary robust self-cooperating strategies are thus described by the set

$$\mathcal{C}_r = \left\{ \mathbf{p} \mid p_{cc} = 1, p_{dc} \leq \frac{R(cc) - R(cd)}{R(dc) - R(cc)}(1 - p_{cd}), p_{dd} \leq \frac{R(cc) - R(dd)}{R(dc) - R(cc)}(1 - p_{cd}) \right\}.$$

These analytic expressions for the robust self-cooperating strategies are confirmed by Monte-Carlo simulations (Fig. S1).

4.2 Self-defectors with $\lambda \geq \chi(R(cd) + R(dc) - R(dd))$ and $\lambda \geq \chi|R(dc) - R(cd)|$ are robust

The self-defectors $\mathcal{D}$ satisfy $p_{dd} = 0$ and score $s_{xx} = R(dd)$ against themselves, which corresponds to $\kappa = R(dd)$. In order to invade, a mutant Y must therefore have

$$s_{yx} > R(dd).$$

From Eq. 5, this condition becomes

$$\chi s_{xy} > \chi R(dd) + \lambda(v_{cd} + v_{dc})$$

We can also combine Eq. 5 with Eq 10 (where the extremal case of Eq. 10 corresponds here to an opponent who plays $(0, 0, 0, 1)$, so that $v_{cc} = 0$) to give

$$\chi(R(cd) + R(dd) - 2R(dd)) > \lambda$$

as a condition for Y to invade X. Similarly, we can combine Eq 5 with Eq 7 (where the extremal case of Eq. 7 corresponds here to an opponent who plays $(1, 1, 1, 1)$, so that $v_{cd} = 0$) to give

$$\chi(R(dc) - R(cd)) > \lambda$$

as a condition for Y to invade X. Converting back to our original coordinate system, this implies that a self-defector X is robust iff:

$$\begin{aligned} p_{dc}(R(cc) - R(dd)) &\leq (R(dd) - R(cd))(1 - p_{cc}) \\ p_{dc}(R(dc) - R(dd)) &\leq (R(dd) - R(cd))(1 - p_{cd}). \end{aligned} \tag{12}$$

The evolutionary robust self-defecting strategies are thus described by the set

$$\mathcal{D}_r = \left\{ \mathbf{p} \mid p_{dd} = 0, p_{dc} \leq \frac{R(dd) - R(cd)}{R(cc) - R(dd)}(1 - p_{cc}), p_{dc} \leq \frac{R(dd) - R(cd)}{R(dc) - R(dd)}(1 - p_{cd}) \right\}.$$

These analytic expressions for the robust self-defecting strategies are confirmed by Monte-Carlo simulations (Fig. S1).

4.3 Self-alternators with $p_{cd} = 0$, $p_{dc} = 1$, $(1-\chi)\kappa \leq (1+\chi)\frac{R(cd)+R(dc)}{2} - 2\chi R(cc)$ and $(1-\chi)\kappa \leq (1+\chi)\frac{R(cd)+R(dc)}{2} - 2\chi R(dd)$ are robust

The self-alternators $\mathcal{A}$ satisfy $p_{cd} = 0$ and $p_{dc} = 1$. Using Eq. 4, and converting to the alternate coordinate system, we have

$$\lambda = (1-\chi)\left(\kappa - \frac{R(cd)+R(dc)}{2}\right)$$

for strategies of this type. From Eq. 5, a resident strategy X of this type has

$$s_{xx} = \frac{R(cd)+R(dc)}{2}.$$

In order to selectively invade the resident, then, a mutant Y must satisfy

$$s_{yx} > \frac{R(cd)+R(dc)}{2}$$

Combining this with Eq 5 gives

$$\chi s_{xy} > \frac{R(cd)+R(dc)}{2}(1-(v_{cd}+v_{dc})) + \chi\frac{R(cd)+R(dc)}{2}(v_{cd}+v_{dc}) - (1-\chi)\kappa(1-(v_{cd}+v_{dc}))$$

We can also combine Eq. 5 with Eq. 9 (where the extremal case of Eq. 9 corresponds here to an opponent who plays $(1,1,1,1)$, so that $v_{dd} = 0$) and we find that Y can selectively invade if

$$(1+\chi)\left(\frac{R(cd)+R(dc)}{2} - \kappa\right) < 2\chi(R(cc)-\kappa)$$

We can also combine Eq. 5 with Eq. 10 (where the extremal case of Eq. 10 corresponds here to an opponent who plays $(0,0,0,0)$, so that $v_{cc} = 0$) and we find that Y can selectively invade if

$$(1+\chi)\left(\frac{R(cd)+R(dc)}{2} - \kappa\right) < 2\chi(R(dd)-\kappa)$$

and converting back to our original coordinate system, this a self-alternator X is robust iff:

$$\begin{aligned} p_{cc} &\leq 2\frac{R(dc)-R(cc)}{R(dc)-R(cd)} \\ p_{dd} &\leq \frac{R(dc)+R(cd)-2R(dd)}{R(dc)-R(cd)}. \end{aligned} \tag{13}$$

These evolutionary robust self-alternating strategies are thus described by the set

$$\mathcal{A}_r = \left\{\mathbf{p} \mid p_{cd} = 0, p_{dc} = 1, p_{cc} \leq 2\frac{R(dc)-R(cc)}{R(dc)-R(cd)}, p_{dd} \leq \frac{R(dc)+R(cd)-2R(dd)}{R(dc)-R(cd)}\right\}.$$

These analytic expressions for the robust self-alternating strategies are confirmed by Monte-Carlo simulations (Fig. S1).

4.4 All equalizers are robust, and all mutants invade equalizers neutrally

Finally, we must deal with the case of the equalizers [13], which have $\chi = \lambda = 0$. From Eq. 5, we see that such strategies satisfy $s_{yx} = \kappa$ against any invader Y. Thus, a population of equalizers is neutral against all possible invaders. The equalizer strategies are thus evolutionary robust. However, unlike the other sets of robust strategies ($\mathcal{C}_r,\ \mathcal{D}_r,\ \mathcal{A}_r$), which resist replacement by any other strategy type, equalizers never resist invasion, and so they tend to be quickly lost from a population through neutral drift. Therefore we exclude them from our further discussion of robust strategies and, indeed, we find that populations spend very little time ($< 0.01\%$) at the equalizers.

4.5 Games with $R(cc) = R(dd)$, or $R(cd) = R(dc)$

As mentioned previously, the Akin coordinate system cannot describe games with $R(cc) = R(dd)$ or $R(cd) = R(dc)$. However, we can nonetheless look at the robust strategies in the limit $R(cc) \to R(dd)$ and $R(dc) \to R(cd)$.

When $R(cc) \to R(dd)$ we find, from Eq 11 for the self-cooperators, $p(dd) = 0$ and

$$p_{dc}(R(dc) - R(cc)) \leq (R(cc) - R(cd))(1 - p_{cd})$$

as the conditions for robustness. For the self-defectors we find from Eq 12 $p_{cc} = 1$ and

$$p_{dc}(R(dc) - R(cc)) \leq (R(cc) - R(cd))(1 - p_{cd})$$

and the sets of robust self-defectors and self-cooperators are identical. We also find, for Eq 13 for the self-alternators

$$\begin{aligned} p_{cc} &\leq 2\frac{R(dc) - R(cc)}{R(dc) - R(cd)} \\ p_{dd} &\leq \frac{R(dc) + R(cd) - 2R(cc)}{R(dc) - R(cd)}. \end{aligned}$$

which can only be satisfied if $R(dc) + R(cd) > 2R(cc)$, i.e if self-alternators maximize their score when playing against themselves.

When $R(cd) \to R(dc)$ we find, from Eq 11 for the self-cooperators

$$\begin{aligned} p_{dc}(R(dc) - R(cc)) &\leq (R(cc) - R(dc))(1 - p_{cd}) \\ p_{dd}(R(dc) - R(cc)) &\leq (R(cc) - R(dd))(1 - p_{cd}). \end{aligned}$$

which can only be satisfied if $R(cc) > R(dc)$, in which case all self-cooperators are always robust. For the self-defectors we find from Eq 12

$$\begin{aligned} p_{dc}(R(cc) - R(dd)) &\leq (R(dd) - R(dc))(1 - p_{cc}) \\ p_{dc}(R(dc) - R(dd)) &\leq (R(dd) - R(dc))(1 - p_{cd}). \end{aligned}$$

which can only be satisfied $R(dd) > R(dc)$. Finally, for the self alternators we find from Eq. 13

$$\begin{aligned} 0 &\leq R(dc) - R(cc) \\ 0 &\leq R(dc) - R(dd). \end{aligned}$$

as conditions for robustness, i.e self-alternators can only be robust if $R(dc) > R(cc)$, in which case they are all always robust.

5 The volume of a robust strategy type

We can use Eqs. 11-13 to calculate the volumes associated with each robust strategy type. In the case of the self-alternators the volume of $\mathcal{A}_r$ is in fact a 2D surface of area

$$\left(2\frac{R(dc)-R(cc)}{R(dc)-R(cd)}\right)\times\left(\frac{R(dc)+R(cd)-2R(dd)}{R(dc)-R(cd)}\right)$$

where, in addition, we must constrain the area so that only strategies within the unit square are included. Similarly, $\mathcal{C}_r$ has cross-sections of area

$$\left(\frac{R(cc)-R(cd)}{R(dc)-R(cc)}(1-p_{cd})\right)\times\left(\frac{R(cc)-R(dd)}{R(dc)-R(cc)}(1-p_{cd})\right)$$

and its volume is calculated by integration, with the limits of integration chosen to include only strategies lying within the unit cube. Finally, $\mathcal{D}_r$ has cross-sections of area

$$\left(1-\frac{R(cc)-R(dd)}{R(dd)-R(cd)}p_{dc}\right)\times\left(1-\frac{R(dc)-R(dd)}{R(dd)-R(cd)}p_{dc}\right)$$

and its volume is calculated by integrating across those strategies lying within the unit cube.

5.1 Time spent at different strategy types

We now use our results on the volumes of robust strategies to approximate the time spent at the different strategy types – self-cooperators, self-defectors, and self-alternators – for fixed payoffs under strong selection. To make this analytical approximation we will assume that the population spends all of its time at these three strategy types, an approximation motivated by the fact that these types contain all the evolutionary robust strategies (except for the equilizers, which are quickly replaced through neutral drift). Indeed, Monte Carlo simulations confirm that populations spend $> 97\%$ of their time at self-alternators, self-cooperators or self-defectors, for values of payoffs ranging across an order of magnitude.

To approximate the amount of time a population spends in $\mathcal{C}$, $\mathcal{D}$ or $\mathcal{A}$, we simply the evolution of strategies in population as a three-state Markov chain (Fig. S2). We assume that the probability g of entering a strategy type is given by the probability that a robust strategy of that type replaces a randomly drawn memory-1 strategy. We assume that non-robust strategies can be neglected, because although they may be able to invade, they can quickly be reinvaded.

In order to calculate the probability of entering a strategy type under the "imitation" model of [10], we use the probability that a new mutant Y fixes in a population otherwise comprised of a resident X:

$$\rho(\mathbf{p_x},\mathbf{q_y})=\left(\sum_{i=0}^{N-1}\prod_{j=1}^{i}e^{\sigma[(j-1)s_{yy}+(N-j)s_{yx}-js_{xy}-(N-j-1)s_{xx}]}\right)^{-1}$$

The probability of the population adopting a self-alternator strategy under in this three-state chain is then

$$g_a=Z\delta^2V_a\int_{\mathbf{p}\in[0,1]^4}\int_{\mathbf{q}\in\mathcal{A}_r}\rho(\mathbf{p},\mathbf{q})\mathbf{dpdq}$$

where $\mathbf{q}$ is integrated over the set of robust self-alternating strategies, $\mathbf{p}$ is integrated over the full set of memory-1 strategy, $\rho(\mathbf{p},\mathbf{q})$ is the probability that a resident strategy $\mathbf{p}$ is replaced by a robust alternator $\mathbf{q}$, and V_A is the two-dimensional area comprised by robust alternators. The term $\delta^2 V_a$ denotes the volumes of all memory-1 strategies within Euclidean distance δ of the robust alternators, called the δ-neighborhood of the robust alternators [2,14]. The constant term Z normalizes the probability of adopting a strategy, so that $g_a + g_c + g_d = 1$.

Similarly, the probability of the system adopting a robust self-cooperator is

$$g_c = Z\delta V_c \int_{\mathbf{p}\in[0,1]^4} \int_{\mathbf{q}\in\mathcal{C}_r} \rho(\mathbf{p},\mathbf{q})\mathbf{dpdq},$$

and the probability of the system adopting a robust self-defector strategy

$$g_d = Z\delta V_d \int_{\mathbf{p}\in[0,1]^4} \int_{\mathbf{q}\in\mathcal{D}_r} \rho(\mathbf{p},\mathbf{q})\mathbf{dpdq}. \tag{14}$$

Once at a robust strategy, we know that, under strong selection, the system evolves through neutral invasion among strategies of the same type ($\mathcal{C}$, $\mathcal{D}$, or $\mathcal{A}$). The probability h of leaving a strategy type is therefore the probability that a randomly drawn memory-1 strategy replaces a randomly drawn resident of that type. For the self-alternators we have

$$h_a = \int_{\mathbf{q}\in\mathcal{A}} \int_{\mathbf{p}\in[0,1]^4} \rho(\mathbf{q},\mathbf{p})\mathbf{dpdq}$$

where q is integrated over all self-alternator strategies $\mathcal{A}$. Similarly we have

$$h_c = \int_{\mathbf{q}\in\mathcal{C}} \int_{\mathbf{p}\in[0,1]^4} \rho(\mathbf{q},\mathbf{p})\mathbf{dpdq}$$

for self-cooperators, where q is integrated over all self-cooperator strategies $\mathcal{C}$. And

$$h_d = \int_{\mathbf{q}\in\mathcal{D}} \int_{\mathbf{p}\in[0,1]^4} \rho(\mathbf{q},\mathbf{p})\mathbf{dpdq} \tag{15}$$

for self-defectors, where q is integrated over all self-defector strategies $\mathcal{D}$. The stationary distribution of this three-state Markov chain with these transition probabilities can be readily found to give

$$\Pi_a = \frac{g_a/h_a}{g_a/h_a + g_c/h_c + g_d/h_d}$$

for the probability of the system to be at an self-alternator strategy,

$$\Pi_c = \frac{g_c/h_c}{g_a/h_a + g_c/h_c + g_d/h_d}$$

for the probability of the system to be at a self-cooperator strategy, and

$$\Pi_d = \frac{g_d/h_d}{g_a/h_a + g_c/h_c + g_d/h_d}$$

for the probability of the system to be at a self-defector strategy.

As shown in Fig. 2 and Fig. 4 of the main text, the analytic expressions above for the amount of time spent at each strategy type, given the current payoff matrix, provide very good approximations for the actual occupancy times observed in Monte-Carlo simulations over all strategies, even as the payoff matrix evolves.

5.2 The evolution of strategy-types in the two-player public goods game

In the two-player public goods game, the volume of robust self-alternators $\mathcal{A}_r$ is a 2D surface of area

$$\left(\frac{2C-B}{C}\right) \times \left(\frac{B-C}{C}\right)$$

where, in addition, we must constrain the area so that only strategies within the unit square are included. Similarly, $\mathcal{C}_r$ has cross-sections of area

$$\left(\frac{B}{2C-B}(1-p_{cd})\right) \times \left(2\frac{B-C}{2C-B}(1-p_{cd})\right)$$

and its volume is calculated by integration, with the limits of integration chosen to include only strategies lying within the unit cube. Finally, $\mathcal{D}_r$ has cross-sections of area

$$\left(1-2\frac{B-C}{2C-B}p_{dc}\right) \times \left(1-\frac{B}{2C-B}p_{dc}\right)$$

As should be clear from this, the volume of robust self-cooperating strategies increases as $B-C$ increases, while the volume of self-defecting strategies decreases as $B-C$ increases. Thus we see why the collapse of cooperation occurs when B/C decreases. In general, mutations that increase $B-C$ are favored by selection in two-player public goods games, since they increase the payoff to self-alternators and self-cooperators and leave the payoff to self-defectors unchanged. Thus what matters is how $B-C$ changes with C. As discussed in the main text, if we assume $B-C=f(C)$ where $f(C)$ increases monotonically in C, then all that matters for the collapse of cooperation is whether $f(C)$ is sub- or super-linear (or linear) in C, since

$$\frac{B}{C} = 1 + \frac{f(C)}{C}$$

When $f(C)$ is sub-linear, cooperation will collapse, whereas if it is super-linear, cooperation will become more prevalent. If $f(C)$ is linear, cooperation will collapse if the intercept $k>0$ and will increase in prevalence if $k<0$. All of these cases are shown in Fig S6.

6 Necessary conditions for memory-1 strategies to be evolutionary robustness under weak selection:

We have so far assumed that selection is strong. However, we can relax this assumption, and consider instead the robustness of memory-1 strategies in the regime of weak selection,which arises when $N\sigma \ll 1$ [15]. For a population evolving under strong selection, a strategy X is evolutionary robust iff $s_{xx} \geq s_{xy}$ for all Y, i.e if no mutant is selected to invade. For a population evolving under weak selection, i.e. for which $N\sigma \ll 1$, even deleterious mutants may reach high frequency due to genetic drift. Therefore, in order to find the strategies that are evolutionary robust under weak selection, we must look at the probability of fixation, $\rho(\mathbf{p_x}, \mathbf{q_y})$. In particular, a strategy X is robust under weak selection iff $\rho(\mathbf{p_x}, \mathbf{q_y}) \leq 1/N$ for all mutants Y, where $1/N$ is the probability of neutral fixation. The expression for $\rho(\mathbf{p_x}, \mathbf{q_y})$ under weak selection can be Taylor expanded to give the following robustness condition: a strategy X is evolutionary robust iff

$$(N-2)(s_{yy} - 2s_{xx} + 2s_{yx} - s_{xy}) < 3(s_{xy} - s_{yx}) \tag{16}$$

where N is the population size [15]. First we derive necessary conditions for robustness. Recall that, for a resident strategy X we can write

$$s_{xx} = \kappa_x - \frac{\lambda_x}{1-\chi_x}(v_{cd} + v_{dc})$$

for the payoff of X against itself and the payoff of a mutant Y against X is

$$s_{yx} = \frac{(1-\chi_x)\kappa_x - \lambda_x(w_{cd} + w_{dc}) + \chi_x((1-\chi_y)\kappa_y - \lambda_y(w_{cd} + w_{dc}))}{(1-\chi_x\chi_y)}$$

Similarly we have

$$s_{yy} = \kappa_y - \frac{\lambda_y}{1-\chi_y}(v*_{cd} + v*_{dc})$$

for the payoff of Y against itself and the payoff of a mutant X against Y is

$$s_{xy} = \frac{(1-\chi_y)\kappa_y - \lambda_y(w_{cd} + w_{dc}) + \chi_y((1-\chi_x)\kappa_x - \lambda_x(w_{cd} + w_{dc}))}{(1-\chi_x\chi_y)}$$

6.1 A strategy with $0 \leq p_{cc} < 1$ and $0 < p_{dd} \leq 1$ can always be invaded unless $p_{cd} = 0$ and $p_{dc} = 1$

Consider, as before, a resident strategy with $p_{cc} < 1$ and $p_{dd} > 0$, along with a mutation that results in a small change to $\kappa_y = \kappa_x + \eta$, and a small change to λ_y so that $(1-\chi_x)\kappa_x - \lambda_x = (1-\chi_y)\kappa_y - \lambda_y$. We then have

$$s_{yx} = \kappa_x + \frac{\chi_x\eta}{1+\chi_x}(1-(w_{cd} + w_{dc})) - \frac{\lambda_x}{1-\chi_x}(w_{cd} + w_{dc})$$

as well as

$$s_{xy} = \kappa_x + \frac{\eta}{1+\chi_x}(1-(w_{cd} + w_{dc})) - \frac{\lambda_x}{1-\chi_x}(w_{cd} + w_{dc})$$

and

$$s_{yy} = \kappa_x + \eta(1 - (v *_{cd} + v*_{dc})) - \frac{\lambda_x}{1-\chi_x}(v *_{cd} + v*_{dc})$$

Also note that $(v *_{cd} + v*_{dc}) - (v_{cd} + v_{dc}) = A^*\eta$ where A^* is finite and is zero if $p_{cd} = 0$ and $p_{dc} = 1$ or $p_{cd} = 1$ and $p_{dc} = 0$. We can then write

$$s_{xy} - s_{yx} = \eta\frac{1-\chi_x}{1+\chi_x}(1 - (v_{cd} + v_{dc}))$$

and

$$s_{xx} - s_{yx} = A\eta\frac{\lambda_x}{1-\chi_x} - \eta\frac{\chi_x}{1+\chi_x}(1 - (v_{cd} + v_{dc}))$$

and

$$s_{yy} - s_{xy} = \eta\frac{\chi_x}{1+\chi_x}(1 - (v_{cd} + v_{dc})) - \eta\frac{\lambda_x}{1-\chi_x}(A^* - A)$$

where terms $O\left(\eta^2\right)$ and greater have been neglected. Replacing these expressions into Eq. 16 gives

$$\eta(N-2)\left[3\frac{\chi_x}{1+\chi_x}(1 - (v_{cd} + v_{dc})) - \frac{\lambda_x}{1-\chi_x}(A^* + A)\right] > 3\eta\frac{1-\chi_x}{1+\chi_x}(1 - (v_{cd} + v_{dc}))$$

This can always be satisfied unless $v_{cd} + v_{dc} = 1$ and $A^* + A = 0$, which occurs iff $p_{cd} = 0$ and $p_{dc} = 1$ or $p_{cd} = 1$ and $p_{dc} = 0$, i.e. if the resident is an self-alternating strategy.

6.2 A strategy with $0 < p_{cd} < 1$ and $0 < p_{dc} < 1$ can always be invaded unless $p_{cc} = 1$ or $p_{dd} = 0$

Similarly, we can consider mutations that change λ_x by a small amount, for strategies with $0 < p_{cd} < 1$ and $0 < p_{dc} < 1$. The resulting payoffs following such a mutation are

$$s_{xx} = \kappa_x - \frac{\lambda_x}{1-\chi_x}(v_{cd} + v_{dc})$$

for the payoff of X against itself and the payoff of a mutant Y against X is

$$s_{yx} = \kappa_x - \frac{\lambda_x}{1-\chi_x}(w_{cd} + w_{dc}) - \eta\frac{\chi_x}{1+\chi_x^2}(w_{cd} + w_{dc})$$

Similarly we have

$$s_{yy} = \kappa_x - \frac{\lambda_x}{1-\chi_x}(v *_{cd} + v*_{dc}) - \eta\frac{1}{1-\chi_x}(v *_{cd} + v*_{dc})$$

for the payoff of Y against itself and the payoff of a mutant X against Y is

$$s_{xy} = \kappa_x - \frac{\lambda_x}{1-\chi_x}(w_{cd} + w_{dc}) - \eta\frac{1}{1+\chi_x^2}(w_{cd} + w_{dc})$$

We can then write

$$s_{xy} - s_{yx} = -\eta \frac{1}{1+\chi_x}(v_{cd} + v_{dc})$$

and

$$s_{yy} - s_{xy} = \eta \frac{\lambda_x}{1-\chi_x}(A - A^*) - \eta \frac{\chi_x}{1+\chi_x^2}(v_{cd} + v_{dc})$$

where in this case $A^* = 0$ if $p_{cc} = 1$ or if $p_{dd} = 0$. We also have

$$s_{xx} - s_{yx} = \eta \frac{\lambda_x}{1-\chi_x} A + \eta \frac{\chi_x}{1+\chi_x^2}(v_{cd} + v_{dc})$$

Replacing these expressions into Eq. 16 gives

$$\eta(N-2)\left[\frac{\lambda_x}{1-\chi_x}(A + A^*) + 3\frac{\chi_x}{1+\chi_x^2}(v_{cd} + v_{dc})\right] < 3\eta \frac{1}{1+\chi_x}(v_{cd} + v_{dc})$$

which can always be satisfied unless $v_{cd} + v_{dc} = 0$ and $A + A^* = 0$, which occurs iff $p_{cc} = 1$ or $p_{dd} = 0$, i.e. if the resident strategy is either a self-cooperator or a self-defector.

6.3 When $N > 2$, a strategy cannot be robust unless it maximizes its score against itself

We have shown that only self-alternators, self-cooperators and self-defectors can be robust under weak selection. However, in a population of $N > 2$ we can also construct a strategy that will selectively replace any resident that does not achieve the maximum possible score against itself. To see this, consider a resident strategy of type i which scores $s_{xx} = R_i^*$ against itself. We can construct a mutant Y which is of both type i and another type j such that it scores $s_{yy} = \tau_i R_i^* + (1 - \tau_i) R_j^*$ against itself (where τ_i is the proportion of time spent in state i at equilibrium.). If R_i^* is not the maximum score a strategy can receive against itself, we can choose j such that $R_j^* > R_i^*$ and thus $s_{yy} > s_{xx}$. Such a mutant scores $s_{yx} = s_{xy} = R_i^*$ against the resident strategy. From Eq. 16 the condition for invasion is then

$$(N-2)(1-\tau_i)(R_j^* - R_i^*) > 0$$

which is satisfied by construction if $N > 2$. Thus a strategy can only resist invasion under weak selection if it maximizes its score against itself when $N > 2$. In the case $N = 2$ the condition for invasion from Eq. 16 is simply $s_{yx} > s_{xy}$, which allows strategies such as always defect to be robust.

7 Relaxation of assumptions

We have made the following simplifying assumptions in the main text and SI:

- Weak selection
- Non-local mutations to strategies
- Rapid mutations to payoffs
- Linear relationships between costs and benefits

We now relax each of these assumptions in turn.

7.1 The collapse of cooperation under weak selection

Sufficient conditions for a strategy to be robust under weak selection can be found using Eqs. 5-10 along with Eq. 16. The case $2R(cc) > R(dc) + R(cd)$, for example, in which only a subset of self-cooperators are robust, has been studied by [2]. For the two-player public goods game with $R(cc) = B-C$, $R(cd) = B/2 - C$, $R(dc) = B/2$ and $R(dd) = 0$, these conditions reduce to

$$\lambda \geq \frac{B-C}{3N}\left[N+1-(2N-1)\chi\right]$$

and

$$\lambda \geq \frac{C}{N-2}\left[N+1-(2N-1)\chi\right]$$

Using Eq. 1 to convert back to the standard coordinate system we have

$$\left[3NC+(2N-1)(B-C)\right](1-p_{cd}) \geq \left[3NC-(2N-1)(B-C)\right]p_{dc}$$

and

$$2(N-2)(B-C)(1-p_{cd}) \geq \left[3NC-(N-2)(B-C)\right]p_{dd}$$

Just as in the case of strong selection, the volume of robust self-cooperative strategies shrinks as the ratio of benefits to costs shrinks. And so this analysis predicts a collapse of cooperation as payoffs evolve towards higher values. This behavior is indeed confirmed by Monte-Carlo simulations (Fig. S9), illustrating that the collapse of cooperation occurs under both strong and weak selection.

7.2 Alternate mutation schemes

7.2.1 Varying γ

We have focused in the main text on a mutation scheme in which $\gamma = 0.1$, so that costs and benefits occur in the relationship $B - C = 1.25C + k$. The collapse of cooperation persists, to a lesser or greater extent, when larger or smaller values of $\gamma > 0$ are considered, as shown in Fig. 43a-c. These payoff-mutation schemes all correspond to a tradeoff in which larger benefits of mutual cooperation, $B - C$, are associated with larger costs of being defected against, C.

We can alternatively consider values $\gamma < 0$. In this case, as B and C increase, the benefit for mutual cooperation, $B - C$, decreases. And so larger benefits of mutual cooperation are no longer associated with larger costs of being defected against. When strategies and payoffs co-evolve under this mutation scheme selection leads to *decreasing* values of C, until C reaches zero (so that there is no longer a Prisoner's Dilemma). As might be expected, there is a collapse of defection in this case, with self-defectors replaced by self-cooperators as payoffs evolve (Fig. S5).

Finally, in the limiting case in which B can evolve independently and C remains fixed, self-cooperators become more successful as B evolves and there is no collapse of cooperation (Fig. S4c).

7.2.2 Varying k

We can alternatively consider values $k < 0$ in the linear relationship described above. In this case, as B and C increase, the ratio, B/C, increases. When strategies and payoffs co-evolve under this mutation scheme, as might be expected, there is a collapse of defection, with self-defectors replaced by self-cooperators as payoffs evolve (Fig. S6c).

7.2.3 Varying α

In the main text we have assumed a mutation scheme with $C > 0$ and $0 \leq \alpha \leq 1$, such that α either increases or decreases with increasing C, corresponding to increasingly synergistic or antagonistic interactions. If we relax this assumption and allow C and α to evolve independently, we find that both α and C tend to increase towards the Snowdrift game, as shown in Fig. S7a-b.

If we further relax these assumptions and allow α and C to adopt any value, we can produce all qualitatively different 2×2 two-player games Fig. S3). If α and C evolve independently under this mutation scheme, α and C increase and the game $R(dc) > R(cd) > R(cc) > R(dd)$ is produced (Fig S7c-d).

7.2.4 Changing the functional relationship between B and C

As shown above, the collapse of cooperation occurs in the public goods game occurs when the ratio B/C decreases as C increases. This in turn implies that the payoff for mutual cooperation, $B - C$ increases sub-linearly with C. In contrast, when $B - C$ increases super-linearly with C, the ratio B/C tends to increase and cooperation becomes increasingly prevalent. Examples of both cases are shown in Fig S6 with $B - C$ increasing as $\sqrt{C}$ (sub-linear case) and $B - C$ increasing as C^2 (super-linear case). Finally, if $B - C$ has an optimal value, the prevalence of cooperation depends on the ratio B/C at the optimum. The equilibrium frequencies of the robust strategy types are shown as a function of B/C in Fig S8.

7.2.5 Slow mutations to payoffs

In the main text we assumed that mutations to payoffs and mutations to strategies occur at equal rates. This assumption can be relaxed to allow for the scenario in which mutations to payoffs are relatively more rare. As shown in Fig. S10 the collapse of cooperation persists even when mutations to payoffs are relatively rare.

7.2.6 Local mutations to strategies

In the main text we assumed that mutations to strategies are global, such that a mutant was drawn uniformly from the space of all memory-1 strategies. This assumption can be altered to consider the scenario in which mutations to strategies increase or decrease each element of a memory-1 strategy, $\mathbf{p}$, by a small amount Δ, with the constraint that mutant probabilities lie in the range $[0, 1]$. As shown in Fig. S11 the collapse of cooperation persists when mutations to strategies are local.

Literature Cited

[1] Press WH, Dyson FJ (2012) Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc Natl Acad Sci U S A* 109:10409–13.

[2] Stewart AJ, Plotkin JB (2013) From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc Natl Acad Sci U S A* 110:15348–53.

[3] Lindgren K (1991) *Evolutionary phenomena in simple dynamics* eds Langton, Taylor, Farmer, Rasmussen (Addison-Wesley).

[4] Lindgren K, Nordahl M (1994) Evolutionary dynamics of spatial games. *Physica D* 75:292–309.

[5] Fu F, Hauert C, Nowak MA, Wang L (2008) Reputation-based partner choice promotes cooperation in social networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 78:026117.

[6] Riolo R, Cohn M, Axelrod R (2001) Evolution of cooperation without reciprocity. *Nature* 414:441–443.

[7] Lee C, Harper M, Fryer D (2014) The art of war: Beyond memory-one strategies in population games. *arXiv:1405.4327v1.*

[8] Akin E (2012) Stable cooperative solutions for the iterated prisoner's dilemma. *arXiv:1211.0969.*

[9] Fudenberg D, Maskin E (1990) Evolution and cooperation in noisy repeated games. *American Economic Review* 80:274–279.

[10] Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E Stat Nonlin Soft Matter Phys* 74:011909.

[11] Maynard Smith J (1973) The logic of animal conflict. *Nature* 246.

[12] Maynard Smith J (1982) *Evolution and the theory of games* (Cambridge University Press, Cambridge).

[13] Boerlijst MC, Nowak MA, Sigmund K (1997) Equal pay for all prisoners. *American Mathematical Monthly* 104:303–307.

[14] Hilbe C, Nowak MA, Sigmund K (2013) Evolution of extortion in iterated prisoner's dilemma games. *Proc Natl Acad Sci U S A.*

[15] Nowak MA (2006) *Evolutionary dynamics: exploring the equations of life* (Belknap Press of Harvard University Press, Cambridge, Mass.).

Supplementary figures

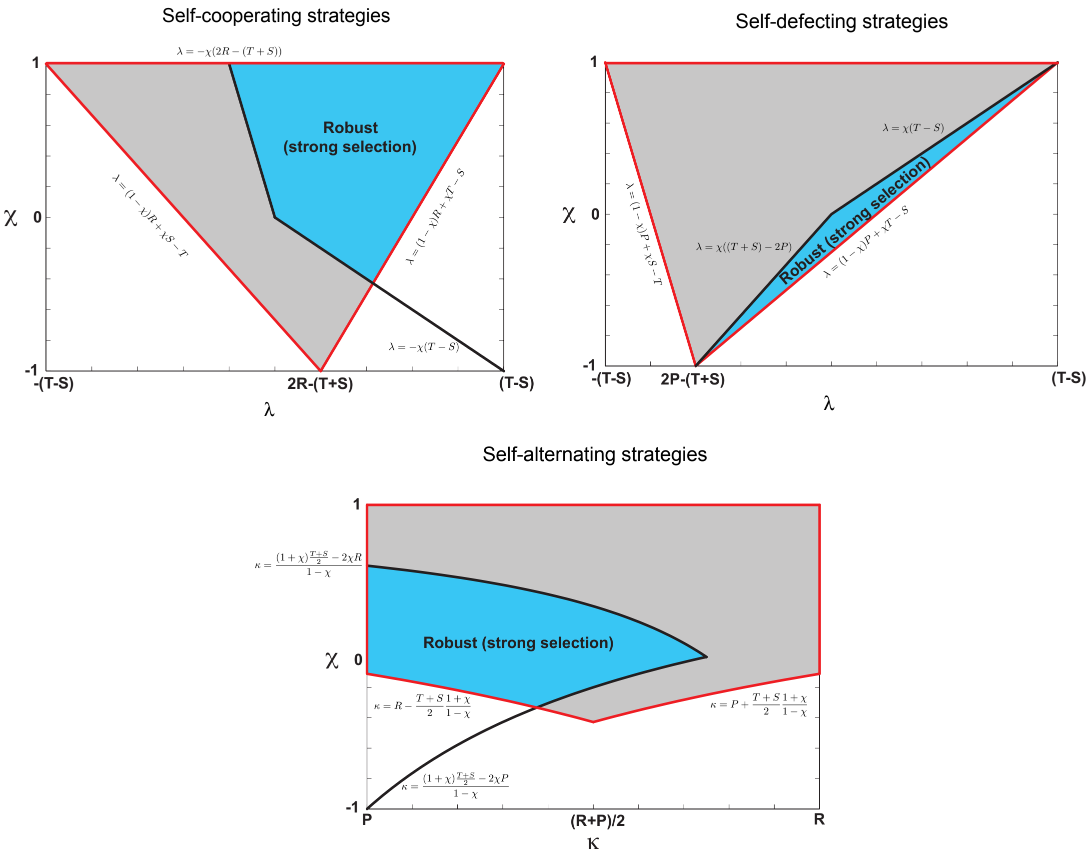

Figure S1 – Confirmation by Monte-Carlo simulation of analytical conditions for evolutionary robustness of strategies. For each of the three strategy types, self-cooperators ($p_{cc} = 1$), self-defectors ($p_{dd} = 0$), and self-alternators ($p_{cd} = 0$ and $p_{dc} = 1$), we compare analytic expression for evolutionary robustness (black lines) with numerical calculations of robustness (light blue regions). Coordinates (κ, χ) for the self-alternating strategies and (λ, χ) for self-cooperators and self-defectors were sampled in regular intervals of 0.01 within the space of all feasible strategies (outlined in red). For each sampled pair of co-ordinates (λ, χ) we also sampled 10^3 associated values of ϕ, ranging from $\phi \to 0$ to the maximum feasible ϕ. To determine numerically whether a focal strategy $X = (\lambda, \chi, \phi)$ is robust we computed the longterm payoffs s_{xx}, s_{yy}, s_{xy} and s_{yx} against 10^6 opponent strategies, Y, drawn uniformly from all memory-1 strategies. A focal strategy X was designated as robust if no strategy Y was found with a score $s_{yx} > s_{xx}$. Parameters are $N = 100$, $\sigma = 10$, $R(cc) = R = 3$, $R(cd) = S = 0$, $R(dc) = T = 5$ and $R(dd) = P = 1$.

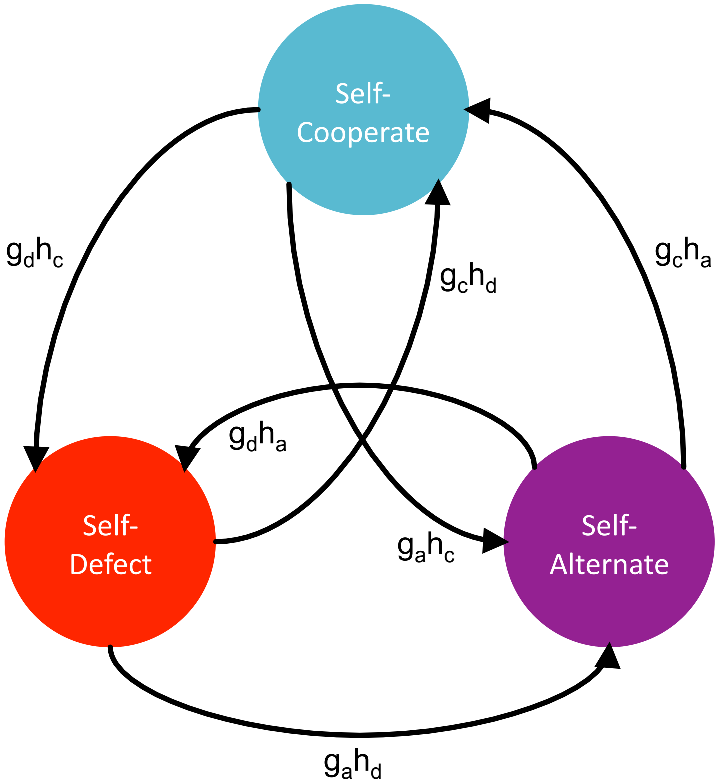

Figure S2 – A simplified, three-state Markov chain to describe evolution of strategies in two-player games. The transition rates are as given by Eqs. 14-15. In this simplified model we assume that the time spent away from these three strategy types can be neglected. This approximation is supported by simulations on the full space of strategies, which indicate that such populations occupy one of these three strategy types > 97% of the time.

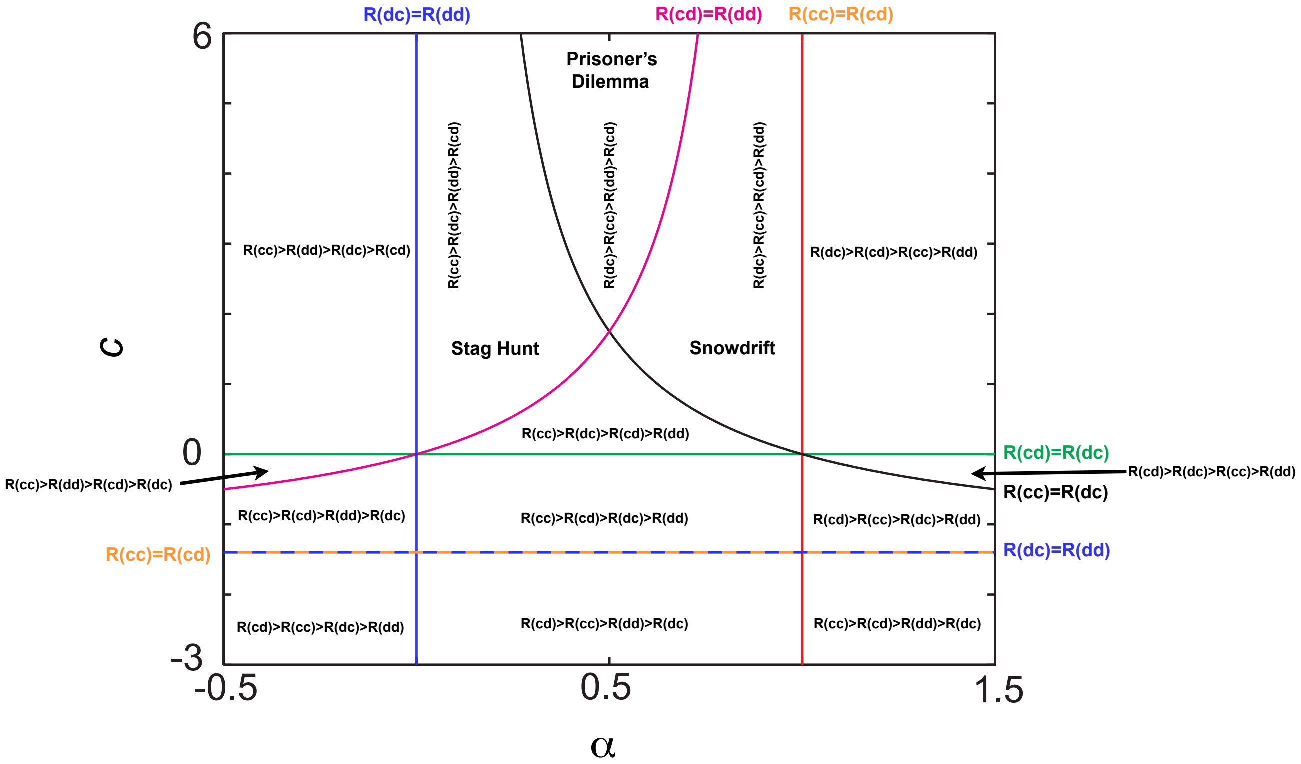

Figure S3 – All possible qualitative types of 2x2 two-player games can be parameterized by two parameters, *alpha* and C, under the payoff scheme $R(cc) = B - C$, $R(cd) = \alpha B - C$, $R(dc) = \alpha B$, and $R(dd) = 0$, where $B - C = \gamma C + k$ with $\gamma = 0.1$ and $k = 1.6$. As $\alpha \in \mathbb{R}$ and $C \in \mathbb{R}$ vary without constraint, all 12 qualitatively different games, as defined by the possible ordering of the payoffs $(R(cc), R(cd), R(dc), R(dd))$, can be produced. Although there are 24 possible orderings in total, the actual number of qualitatively different games is reduced by a factor of two, due to symmetry. Lines indicate the boundaries where two payoffs, as indicated, are equal; and each region is labelled according to the type of game it encodes. The three classic social dilemmas – the Prisoner's Dilemma, Snowdrift, and Stag Hunt games – are labelled and occur in the most biologically relevant region, $\alpha \in [0, 1]$ and $C \geq 0$.

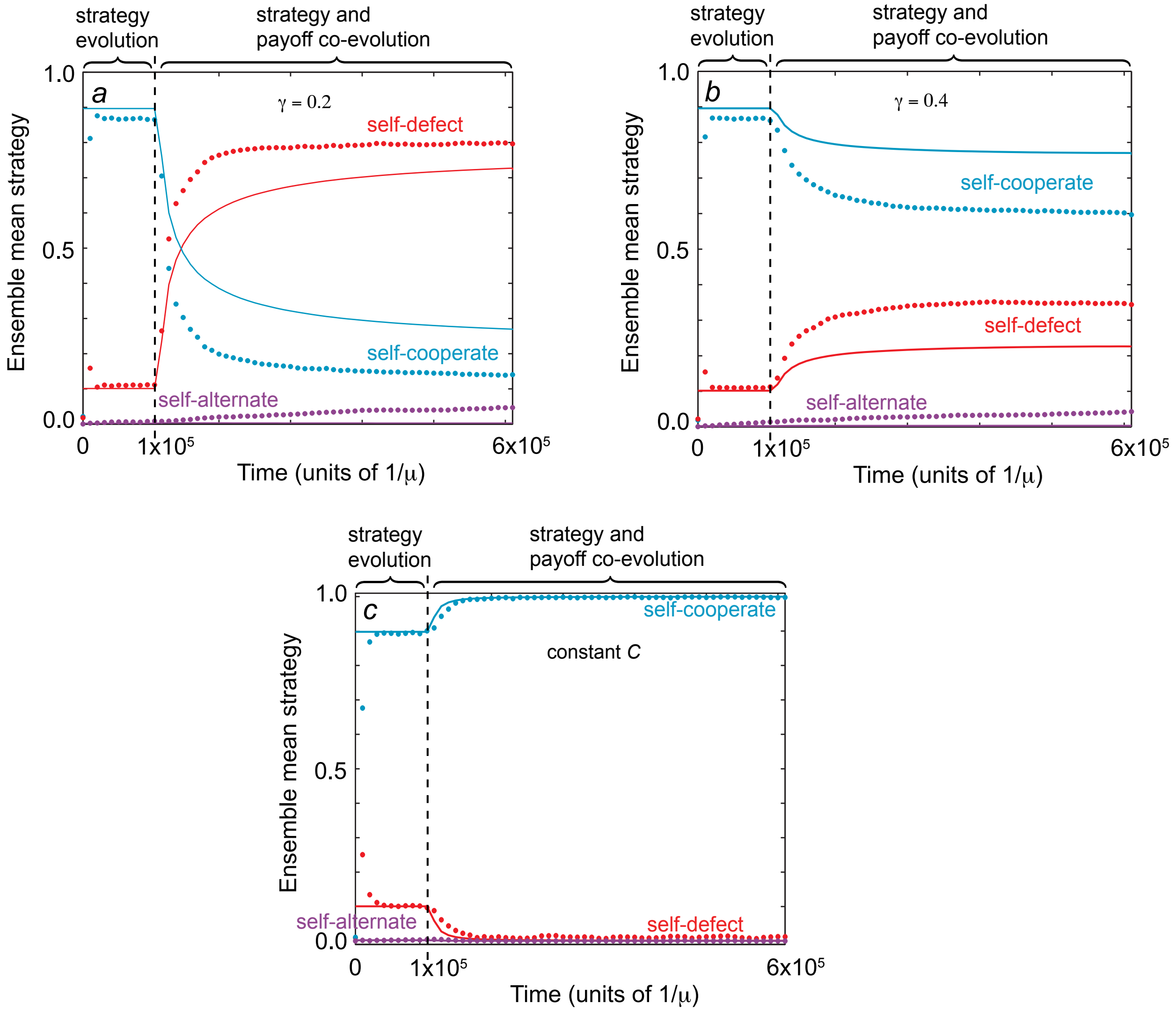

Figure S4 – The collapse of cooperation in the Prisoner's Dilemma under different mutation schemes. We simulated populations under weak mutation, proposing both mutant strategies and mutant payoffs at equal rates, $\mu/2$. Mutations to strategies were drawn uniformly from the full space of memory-1 strategies. Mutations to payoffs were drawn so that increasing benefits of cooperation incur increasing costs: mutations perturbing the benefit B by Δ were drawn uniformly from the range $\Delta \in [-0.1, 0.1]$, with the corresponding change to cost C chosen to enforce the relationship $B = \gamma C + k$ with (a) $\gamma = 0.2$, (b) $\gamma = 0.4$ or (c) allowing B to evolve with fixed $C = 1$. Evolution was modelled according to an imitation process under weak mutation [2, 10, 14]. Self-cooperative strategies are initially robust and dominate the population, but they are quickly replaced by self-defectors as payoffs evolve. Dots indicate the proportion of 10^5 replicate populations, at each time point, within distance $\delta = 0.01$ of the three strategy types self-cooperate, self-defect, and self-alternate. Lines indicate analytic predictions for the frequencies of these strategy types, which depend upon the corresponding volumes of robust strategies. Simulations were run until each population experienced 5×10^5 mutations. Populations of size $N = 100$ were initiated with $B = 3$ and $C = 1$, and evolved under selection strength $\sigma = 1$ (corresponding to strong selection on strategies).

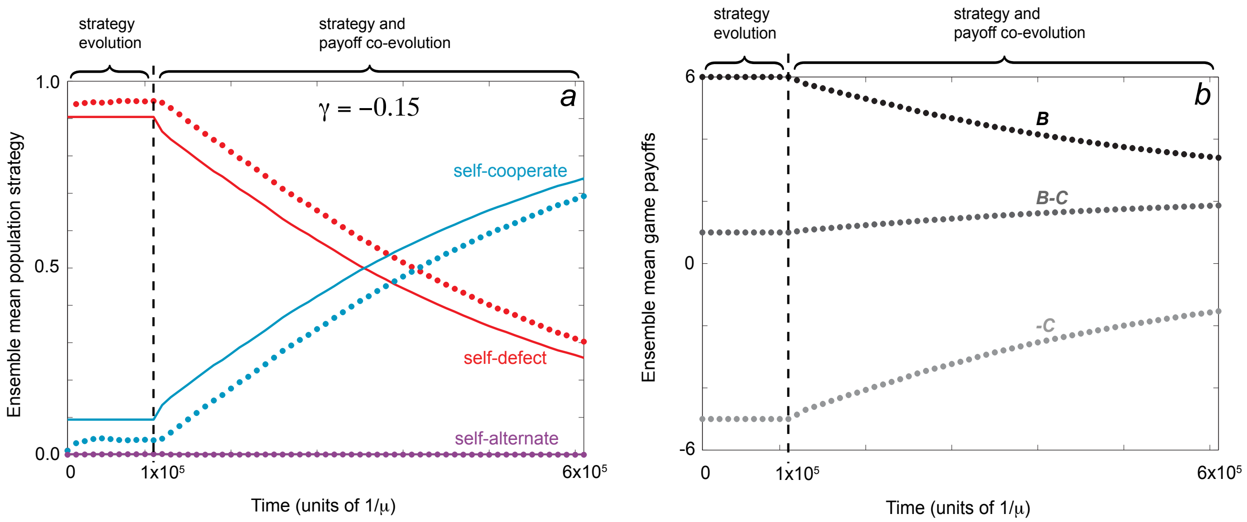

Figure S5 – Co-evolution of strategies and payoffs when B and C are allowed to evolve with $\gamma = -0.15$, so that the payoff for mutual cooperation, $B - C$, increases as B and C decrease. (a) Populations were initialized at $B = 6$ and $C = 5$, under which self-defect dominates. Once strategies and payoffs start to co-evolve, self-cooperate begins to increase and eventually comes to dominate. (b) Benefits B and costs C evolve towards lower values. We simulated populations under weak mutation as in Fig. 2a. Lines indicate analytic predictions for the frequencies of these strategy types, which depend upon the corresponding volumes of robust strategies. Simulations were run until each population had experienced 5×10^5 mutations. Populations of size $N = 100$ were initiated with $B = 3$ and $C = 1$, and evolved under selection strength $\sigma = 1$ (corresponding to strong selection on strategies).

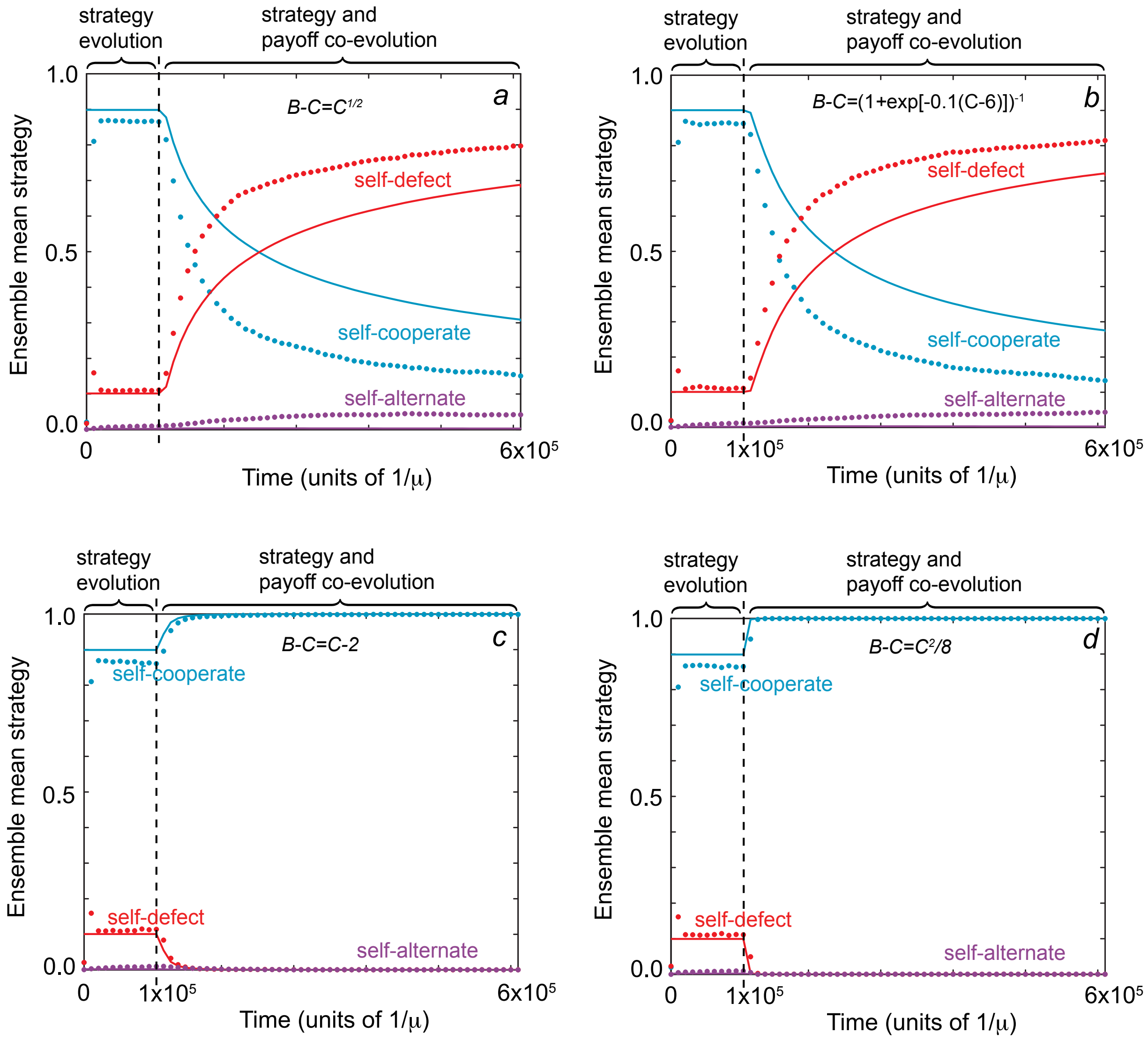

Figure S6 – The collapse of cooperation in the Prisoner's Dilemma under different mutation schemes. We simulated populations under weak mutation, proposing both mutant strategies and mutant payoffs at equal rates, $\mu/2$. Mutations to strategies were drawn uniformly from the full space of memory-1 strategies. Mutations to payoffs were drawn so that increasing benefits of cooperation incur increasing costs: mutations perturbing the benefit B by Δ were drawn uniformly from the range $\Delta \in [-0.1, 0.1]$, with the corresponding change to cost C chosen to enforce various payoff relationships. (a) $B - C = \sqrt{C}$, an example in which the benefits of mutual cooperation grow sub-linearly with the costs. Here self-cooperative strategies are initially robust and dominate the population, but they are quickly replaced by self-defectors as payoffs evolve. (b) $B - C = B_{max}/(1 + \exp[h(C_0 - C)])$, an example in which the benefits of mutual cooperation saturate with increasing costs (where B_{max} is the maximum value of B, chosen so that when $C = 4$, $B = 6$. (c) $B - C = \gamma C + k$, with $\gamma = 1$ and $k = -2$, an example in which the benefits of mutual cooperation grow linearly with the costs, with a negative intercept. (d) $B - C = C^2/8$, an example in which the benefits of mutual cooperation grow super-linearly with the costs. Here self-cooperative strategies are initially robust and become more prevalent as payoffs evolve. Here self-cooperative strategies are initially robust and dominate the population, but they are quickly replaced by self-defectors as payoffs evolve. Here self-cooperative strategies are initially robust and become more prevalent as payoffs evolve. Evolution was modelled according to an imitation process under weak mutation [2, 10, 14]. Dots indicate the proportion of 10^5 replicate populations, at each time point, within distance $\delta = 0.01$ of the three strategy types self-cooperate, self-defect, and self-alternate. Lines indicate analytic predictions for the frequencies of these strategy types, which depend upon the corresponding volumes of robust strategies. Simulations were run until each population experienced 5×10^5 mutations. Populations of size $N = 100$ were initiated with $B = 6$ and $C = 4$ and $h = 0.1$ in (c). Populations evolved under selection strength $\sigma = 1$ (corresponding to strong selection on strategies).

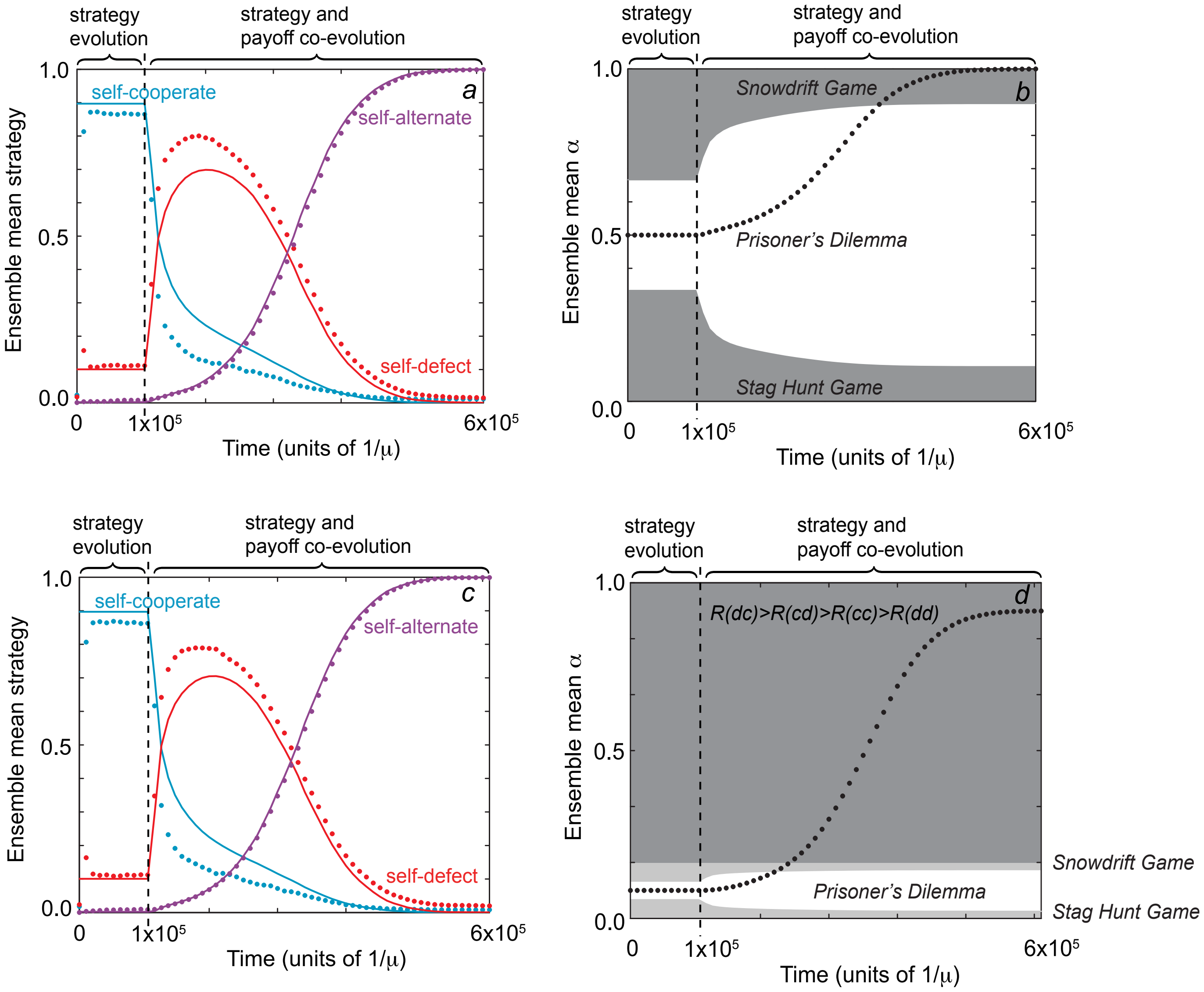

Figure S7 – Evolution away from the Prisoner's Dilemma. We simulated a population under weak mutation, proposing mutant strategies drawn uniformly from the full space of memory-1 strategies. Alongside mutations to the payoffs B and C, as in Fig. 2, we also allowed mutations to the additional payoff parameter α, which determines whether mutual cooperation is synergistic ($\alpha < 1/2$) or antagonistic ($\alpha > 1/2$) for the amount of public benefit produced. (a) Evolution with $\alpha \in [0, 1]$ (top panels) produces a rapid loss of cooperation and increase in defection, as in Fig. 2, followed by an increase in alternating strategies. Points indicate the proportion of simulated populations within a distance $\delta = 0.01$ of the three strategy types self-cooperation, self-defect, and self-alternate; lines indicate analytic predictions. (b) Following the collapse of cooperation, the Prisoner's Dilemma (unshaded region) is replaced by the Snowdrift game (shaded region), with $1 > \alpha > C/B$. Parameters values as in Fig. 2a-b. (c) Evolution with $\alpha \in \mathbb{R}$ (bottom panels) unconstrained produces a rapid loss of cooperation and increase in defection, as in Fig. 2, followed again by an increase in alternating strategies. Points indicate the proportion of simulated populations within a distance $\delta = 0.01$ of the three strategy types; lines indicate analytic predictions. (d) Following the collapse of cooperation, the Prisoner's Dilemma (unshaded region) is replaced first by the Snowdrift game (shaded region), with $\alpha > C/B$, and then by the game $R(dc) > R(cd) > R(cc) > R(dd)$, with $\alpha > 1$. Parameters values as in Fig. 2a-b. Populations of size $N = 100$ were initiated with $B = 6$ and $C = 4$ and $\alpha = 1/2$, and evolved under selection strength $\sigma = 1$ (corresponding to strong selection on strategies), with $\gamma = 0.1$.

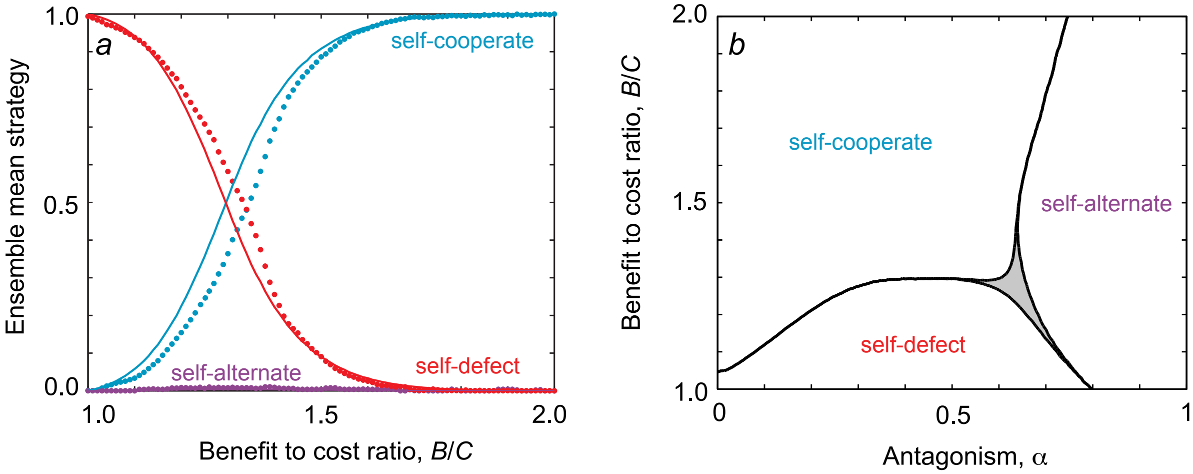

Figure S8 – Cooperation in the Prisoner's Dilemma as a function of the ratio B/C of costs to benefits. (a) We simulated populations under weak mutation, as in Fig. 2a, except that we exogenously varied the ratio B/C of costs to benefits (x-axis) (keeping $\alpha = 0.5$ fixed) and plotted the equilibrium frequency of self-cooperators, self-defectors, and self-alternators (points). These simulated equilibrium frequencies are compared to analytic predictions (lines) based on the volumes of robust strategies of each type. As B/C decreases, so too does the equilibrium frequency of self-cooperators. When there is an optimal value of $B - C$ then the ratio B/C at the optimum determines the long-term prevalence of self-cooperative strategies. (b) Analytical results for the equilibrium frequency of self-cooperators, self-defectors and self-alternators in terms of the ratio of costs to benefits, B/C (y-axis) and the degree of antagonism α (x-axis). Labelled regions show parameter values for which a given strategy type occurs $> 50\%$ of the time, while in the small gray region no strategy type is present $> 50\%$ of the time

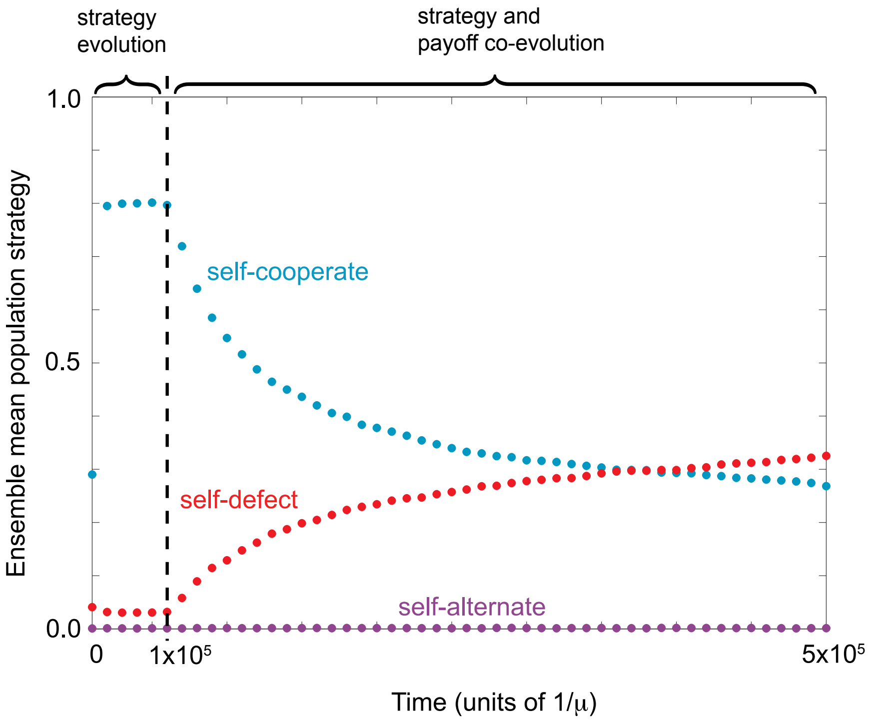

Figure S9 – The collapse of cooperation in the Prisoner's Dilemma under weak selection. We simulated populations under weak mutation as in Fig. 2a, except with $N = 100$ and $\sigma = 0.01$ (weak selection). Self-cooperative strategies are initially robust and dominate the population, but they are quickly replaced by self-defectors as payoffs evolve.

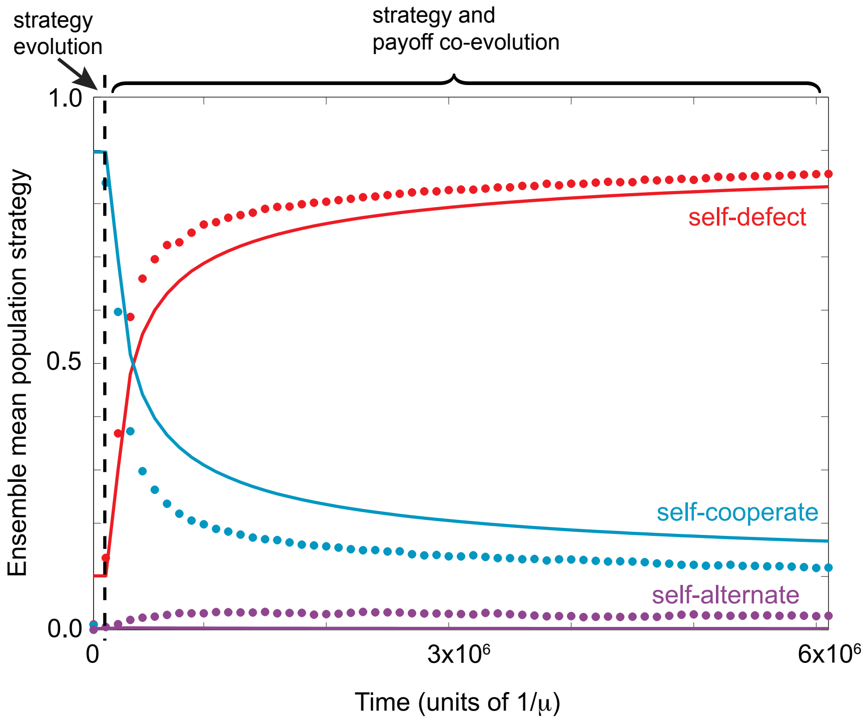

Figure S10 – Slow mutations to payoffs and the collapse of cooperation in the Prisoner's Dilemma. We simulated populations under weak mutation as in Fig. 2a, except that mutations altering strategies occur at 10^3-times the rate of mutations altering payoffs. Self-cooperative strategies are initially robust and dominate the population, but they are quickly replaced by self-defectors as payoffs evolve.

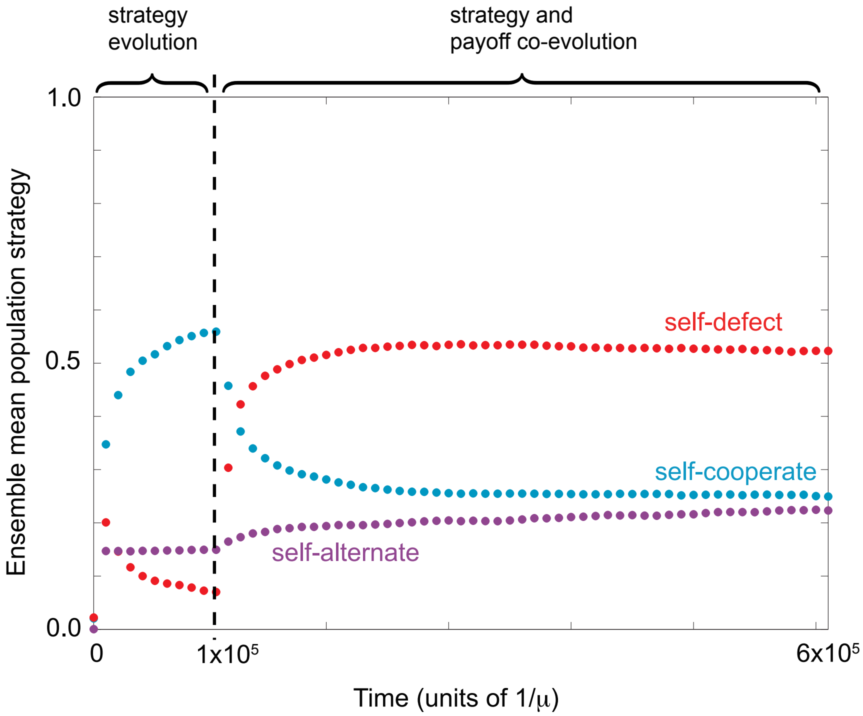

Figure S11 – Local mutations to payoffs and the collapse of cooperation in the Prisoner's Dilemma. We simulated populations under weak mutation as in Fig. 2a, except that mutations altering strategies are now "local" so that mutations perturbing each of the four probabilities ($p_{cc}, p_{cd}, p_{dc}, p_{dd}$ by an amount Δ were drawn uniformly from the range $\Delta \in [-0.01, 0.01]$. Self-cooperating strategies are initially robust and dominate the population, but they are quickly replaced by self-defectors as payoffs evolve.