

From extortion to generosity, evolution in the Iterated Prisoner's Dilemma

Alexander J. Stewart and Joshua B. Plotkin¹

Department of Biology, University of Pennsylvania, Philadelphia, PA 19104

Edited* by William H. Press, University of Texas at Austin, Austin, TX, and approved July 25, 2013 (received for review April 3, 2013)

Recent work has revealed a new class of “zero-determinant” (ZD) strategies for iterated, two-player games. ZD strategies allow a player to unilaterally enforce a linear relationship between her score and her opponent's score, and thus to achieve an unusual degree of control over both players' long-term payoffs. Although originally conceived in the context of classical two-player game theory, ZD strategies also have consequences in evolving populations of players. Here, we explore the evolutionary prospects for ZD strategies in the Iterated Prisoner's Dilemma (IPD). Several recent studies have focused on the evolution of “extortion strategies,” a subset of ZD strategies, and have found them to be unsuccessful in populations. Nevertheless, we identify a different subset of ZD strategies, called “generous ZD strategies,” that forgive defecting opponents but nonetheless dominate in evolving populations. For all but the smallest population sizes, generous ZD strategies are not only robust to being replaced by other strategies but can selectively replace any noncooperative ZD strategy. Generous strategies can be generalized beyond the space of ZD strategies, and they remain robust to invasion. When evolution occurs on the full set of all IPD strategies, selection disproportionately favors these generous strategies. In some regimes, generous strategies outperform even the most successful of the well-known IPD strategies, including win-stay-lose-shift.

evolution of cooperation | altruism | evolutionary stability | Nash

Press and Dyson (1) recently revealed a remarkable class of strategies, called “zero-determinant” (ZD) strategies, for iterated two-player games. ZD strategies are of particular interest in the Iterated Prisoner's Dilemma (IPD), the canonical game used to study the emergence of cooperation among rational individuals (2–9). By allowing a player to unilaterally enforce a linear relationship between her payoff and her opponent's payoff, Press and Dyson (1) argue, ZD strategies provide a sentient player unprecedented control over the long-term outcome of IPD games. In particular, Press and Dyson (1) highlighted a subset of ZD strategies, called “extortion strategies,” that grants the extorting player a disproportionately high payoff when employed against a naive opponent who blindly adjusts his strategy to maximize his own payoff.

A natural response to Press and Dyson (1) is to ask: What are the implications of ZD strategies for an evolving population of players (10)? Although several recent studies have begun to explore this question (11, 12), they have focused almost exclusively on extortion strategies. Extortion strategies are not successful in evolving populations unless the population size is very small. Like all strategies that prefer to defect rather than to cooperate, extortion strategies are vulnerable to strategies that reward cooperation but punish defection. However, there is more to ZD strategies than just extortion, and recent work has uncovered some ZD strategies that promote cooperation in two-player games (10, 13). Here, we consider the full range of ZD strategies in a population setting and show that when it comes to evolutionary success, it is generosity, not extortion, that rules.

We begin our analysis by considering populations restricted to the space of ZD strategies. We show that evolution within ZD always leads to a special subset of strategies, which we call “generous” ZD. Generous ZD strategies reward cooperation but punish defection

only mildly, and they tend to score lower payoffs than those of defecting opponents. Next, we build on recent work by Akin (13), who identified generous strategies beyond those contained within ZD. We demonstrate that a large proportion of these generous strategies are robust to replacement in an evolving population. At worst, the robust generous strategies can be replaced neutrally. Conversely, we demonstrate that most generous strategies can readily replace resident nongenerous strategies in a population. As a result, generous strategies are just as, or sometimes even more, successful than the most successful of well-known IPD strategies in evolving populations. Finally, we show that populations evolving on the full set of IPD strategies spend a disproportionate amount of time near generous strategies, indicating that they are favored by evolution.

Methods and Results

In the Prisoner's Dilemma, two players, X and Y , must simultaneously choose whether to cooperate (c) or defect (d). If both players cooperate (cc), they each receive payoff R . If X cooperates and Y defects (cd), X loses out and receives the smallest possible payoff, S , whereas Y receives the largest possible payoff, T . If both players defect (dd), both players receive payoff P . Payoffs are specified so that the reward for mutual defection is less than the reward for mutual cooperation (i.e., $T > R > P > S$). It is typically assumed that $2R > T + S$, so that it is not possible for total payoff received by both players to exceed $2R$. In what follows, we will consider the payoffs $T = B$, $R = B - C$, $P = 0$, and $S = -C$, which comprise the so-called “donation game” (12).

The IPD consists of infinitely many successive rounds of the Prisoner's Dilemma. Press and Dyson (1) showed that it is sufficient to consider only the space of memory-1 strategies (i.e., strategies that specify the probability of a player cooperating in each round

Significance

Cooperative behavior seems at odds with the Darwinian principle of survival of the fittest, yet cooperation is abundant in nature. Scientists have used the Prisoner Dilemma game, in which players must choose to cooperate or defect, to study the emergence and stability of cooperation. Recent work has uncovered a remarkable class of extortion strategies that provide one player a disproportionate payoff when facing an unwitting opponent. Extortion strategies perform very well in head-to-head competitions, but they fare poorly in large, evolving populations. Rather we identify a closely related set of generous strategies, which cooperate with others and forgive defection, that replace extortionists and dominate in large populations. Our results help to explain the evolution of cooperation.

Author contributions: A.J.S. and J.B.P. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: jplotkin@sas.upenn.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1306246110/-DCSupplemental.

in terms of the payoff she received in the previous round. Memory-1 strategies consist of four probabilities, $\mathbf{p} = \{p_{cc}, p_{cd}, p_{dc}, p_{dd}\}$. In particular, Press and Dyson (1) showed that the long-term payoff to a memory-1 player pitted against an arbitrary opponent is the same as her payoff would be against some other memory-1 opponent. Thus, we limit our analysis to memory-1 players without loss of generality (*Materials and Methods*).

Evolutionary Game Theory. In the context of evolutionary game theory, we consider a population of N individuals who are each characterized by a strategy \mathbf{p} . We say strategy \mathbf{p} receives long-term payoff $\pi(\mathbf{p}, \mathbf{q})$ against an opponent with strategy \mathbf{q} . The success of a strategy depends on its payoff when pitted against all individuals in the population (14–17). Traditionally, the evolutionary outcome in such a population has been understood in terms of evolutionary stable strategies (ESSs). A strategy \mathbf{p} is an ESS if its long-term payoffs satisfy $\pi(\mathbf{p}, \mathbf{p}) > \pi(\mathbf{p}, \mathbf{q})$, or $\pi(\mathbf{p}, \mathbf{p}) = \pi(\mathbf{p}, \mathbf{q})$ and $\pi(\mathbf{p}, \mathbf{q}) > \pi(\mathbf{q}, \mathbf{q})$, for all opponents $\mathbf{q} \neq \mathbf{p}$.

The ESS condition provides a useful notion of stability in the context of an infinite population. However, in a finite population, the concept must be generalized to consider whether selection favors both invasion and replacement of a resident strategy by a mutant strategy (18, 19). In a finite, homogeneous population of size N , a newly introduced neutral mutation (i.e., a mutation that does not change the payoff to either player) will eventually replace the entire population with probability $\rho = 1/N$. A deleterious mutation, which is opposed by selection, will fix with probability $\rho < 1/N$, whereas an advantageous mutation, which is favored by selection, will fix with probability $\rho > 1/N$. We say that a resident strategy \mathbf{p} in a finite population of size N is “evolutionary robust” against a mutant strategy \mathbf{q} if the probability of replacement satisfies $\rho \leq 1/N$; in other words, the robust strategy cannot be selectively replaced by the mutant strategy. In the limit of infinite population size, $N \rightarrow \infty$, the condition $\rho < 1/N$ reduces to the ESS condition.

When selection is weak ($\sigma \ll 1$; *Materials and Methods*), we can write down an explicit criterion for robustness: A resident Y is evolutionary robust against a mutant X if and only if

$$s_{xx}(N-2) + s_{yy}(2N-1) \leq s_{yx}(N+1) + 2s_{yy}(N-2), \quad [1]$$

where we denote the long-term payoff of player X against player Y by s_{xy} . We restrict our analysis to memory-1 players. In the two-player setting, this restriction does not sacrifice generality because, as per Press and Dyson (1), the payoff received by a memory-1 strategy Y can be determined independent of an opponent’s memory. However, in an evolutionary setting, Y ’s success depends also on the payoff her opponent receives against himself. Nonetheless, we will show that our results for generous strategies hold against all opponents, no matter how long their memories, provided the standard IPD assumption $2R > T + S$ holds.

Zero-Determinant Strategies, Extortion, and Generosity. Among the space of all memory-1 IPD strategies, Press and Dyson (1) identified a subspace of ZD strategies that ensure a fixed, linear relationship between two players’ long-term payoffs. If player Y facing player X employs a ZD strategy of the form

$$\begin{aligned} p_{cc} &= 1 - \phi(1 - \chi)(B - C - \kappa) \\ p_{dc} &= 1 - \phi[\chi C + B - (1 - \chi)\kappa] \\ p_{cd} &= \phi[\chi B + C + (1 - \chi)\kappa] \\ p_{dd} &= \phi(1 - \chi)\kappa, \end{aligned}$$

their payoffs will satisfy the linear relationship

$$\phi[s_{yy} - \chi s_{yx} - (1 - \chi)\kappa] = 0. \quad [2]$$

The parameters χ and κ must lie in the range $0 \leq \kappa \leq B - C$ and $\max(\frac{\kappa - B}{\kappa + C}, \frac{\kappa + C}{\kappa - B}) \leq \chi \leq 1$ to produce a feasible strategy. Eq. 2 defines the full space of ZD strategies introduced by Press and Dyson (1). Within this space, two particular subsets are of special interest: the extortion strategies, described by Press and Dyson (1), for which $\kappa = P = 0$ and $\chi > 0$, and the generous strategies, described in our commentary (10), for which $\kappa = R = B - C$ and $\chi > 0$.

Extortion strategies ensure that either the extortioner Y receives a higher payoff than her opponent X , $s_{yx} > s_{xy}$, or that both players otherwise receive the payoff for mutual defection, $s_{yx} = s_{xy} = 0$. In contrast, generous strategies ensure that both players receive the payoff for mutual cooperation, $s_{yx} = s_{xy} = B - C$, or that the generous player Y otherwise receives a lower payoff than her opponent, $s_{yx} < s_{xy}$.

Recent work has focused on the evolutionary prospects of extortioners (11, 12) and has found that such strategies are unsuccessful, except in very small populations. In fact, as we will show below, selection favors replacement of extortioners by generous strategies, and generous strategies are robust to replacement by extortioners. Moreover, the success of generous strategies persists when evolution proceeds within the full space of IPD strategies.

Evolution of Generosity Within ZD Strategies. We start by identifying the subset of ZD strategies that is evolutionary robust against all IPD strategies in a population of size N . Substituting Eq. 2 into Eq. 1 shows that a resident ZD strategy Y with $\kappa_y = B - C$ is robust against any mutant IPD strategy X if and only if $\chi_y \geq (N + 1)/(2N - 1)$ (*Materials and Methods*). Conversely, provided that $N > 2$, any resident ZD strategy Y with $\kappa_y < B - C$ can be selectively replaced by another strategy, namely, by a ZD strategy with $\kappa_x = B - C$ and $\chi_x > (N + 1)/(2N - 1)$ (*Materials and Methods*). Hence, those ZD strategies with $\kappa = B - C$ and $\chi \geq (N + 1)/(2N - 1)$ are precisely the ZD strategies that are evolutionary robust against all IPD strategies. We denote this set of robust ZD strategies as ZD_R :

$$ZD_R = \left\{ (\kappa, \chi, \phi) \mid \kappa = B - C, 1 > \chi \geq \frac{N + 1}{2N - 1} \right\}.$$

Here, ϕ is left unconstrained, but it must lie in the range required to produce a feasible strategy, $0 < \phi \leq \chi B / (\chi C + B)$.

The robust ZD strategies are what we call “cooperative,” meaning they satisfy $\kappa = B - C$. Any cooperative player will agree to mutual cooperation when facing another cooperative player, and so they each receive payoff $B - C$. If a cooperative strategy further satisfies the condition $\chi > 0$, we say that the strategy is generous, meaning that any deviation from mutual cooperation causes the generous player’s payoff to decline more than that of her opponent. The robust ZD strategies are all generous.

We now consider evolution in a population of $N > 2$ players restricted to the space of ZD strategies. Because selection favors replacement of any noncooperative ZD strategy by some member of ZD_R , we expect evolution within the space of ZD strategies to tend towards generous strategies, and thereafter to remain at generous strategies, because ZD_R is robust. This expectation is confirmed by Monte Carlo simulations of well-mixed populations of IPD players (Fig. 1). Following Hilbe et al. (12) and Traulsen et al. (20), we modeled evolution as a process in which individuals copy successful strategies with a probability that depends on their relative payoffs (*Materials and Methods*). As Fig. 1 shows, evolution within the set of ZD strategies proceeds from extortion ($\kappa = 0$ and $\chi > 0$) to generosity ($\kappa = B - C$ and $\chi > 0$). In fact, even populations initiated with $\chi < 0$ evolve to generosity (Fig. S1).

Good Strategies. The generous ZD strategies identified above are best understood by comparison with the space of “good” strategies recently introduced by Akin (13). A good strategy stabilizes cooperative behavior in the two-player IPD: By definition, if both

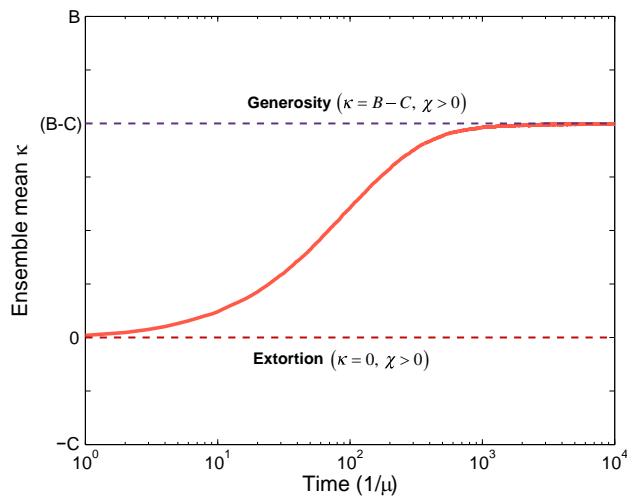


Fig. 1. Evolution from extortion to generosity within the space of ZD strategies. Populations were simulated in the regime of weak mutation. The figure shows the ensemble mean value of κ in the population, plotted over time. The expression $\kappa=0$ corresponds to the extortion strategies of Press and Dyson (1), whereas $\kappa=B-C$ corresponds to the generous ZD strategies. Each population was initialized at an extortion strategy E_χ , with χ drawn uniformly from the range $\chi \in (0,1)$. Given a resident strategy in the population, mutations to κ were proposed as normal deviates of the resident strategy, truncated to constrain $\kappa \in [0, B-C]$, whereas mutations to $\{\chi, \phi\}$ were drawn uniformly from $\max\left(\frac{\kappa-B}{\kappa+C}, \frac{\kappa+C}{\kappa-B}\right) \leq \chi \leq 1$ with ϕ drawn uniformly within the feasible range, given κ and χ . A proposed mutant strategy replaces the resident strategy with a fixation probability dependent on their respective payoffs, as in the work of Hilbe et al. (12) and Traulsen et al. (20). The mean κ among 10^3 replicate populations is plotted as a function of time. Parameters are $B=3$, $C=1$, $N=100$, and selection strength $\sigma=1$.

players adopt good strategies, each receives payoff $B-C$ and neither player can gain by unilaterally changing strategy. All good strategies are cooperative (i.e., they have $\kappa=B-C$). Moreover, the generous ZD strategies described above are precisely the intersection of good strategies (13) and ZD strategies (1) (Fig. 2).

We can identify the space of memory-1 good strategies as those of the form

$$\begin{aligned} p_{cc} &= 1 - \phi(1-\chi)(B-C-\kappa) \\ p_{dc} &= 1 - \phi[\chi C + B - (1-\chi)\kappa + \lambda] \\ p_{cd} &= \phi[\chi B + C + (1-\chi)\kappa - \lambda] \\ p_{dd} &= \phi(1-\chi)\kappa, \end{aligned}$$

where $-1 \leq \chi \leq 1$ and $-(\chi B + C) \leq \lambda \leq (B + \chi C)$ are required to produce a feasible strategy. Sufficient conditions for set G of good strategies are (SI Text):

$$\begin{aligned} G &= \{(\kappa, \chi, \phi, \lambda) | \\ &\kappa = B - C, \lambda > -(B - C)\chi, \lambda > -(B + C)\chi\}, \end{aligned}$$

where the parameter ϕ is left unconstrained except that it must produce a feasible strategy. Numerics indicate these sufficient conditions are also necessary (SI Text). Note that the good strategies with $\lambda=0$ correspond precisely to the generous ZD strategies.

It is interesting to note that in addition to tit-for-tat and generous tit-for-tat, which are ZD, the set of good strategies contains win-lose-stay-shift, which is widely known as one of the most evolutionary successful IPD strategies (7). Nonetheless, even though win-lose-stay-shift is good, it is not generous (it has $\chi = -C/B < 0$; Fig. 3). Because it lacks generosity, win-lose-stay-shift can, in fact, be outcompeted in evolving populations, as we shall see below.

Evolutionary of Generosity Within Good Strategies. In this section, we ask which good strategies are evolutionary robust, and we find that the robust good strategies are always generous (i.e., have $\chi > 0$, regardless of λ). In the case of ZD, the conditions for evolutionary robustness do not depend on the parameter ϕ . Similarly, we will derive conditions for the robustness of good strategies that hold regardless of ϕ .

Application of Eq. 1 allows us to derive the conditions for a good strategy to be evolutionary robust against all IPD strategies in a population of size N (SI Text). The resulting set, G_R , of evolutionary robust good strategies satisfies

$$\begin{aligned} G_R &= \{(\kappa, \chi, \phi, \lambda) | \kappa = B - C, \chi < 1, \\ &\lambda > \frac{B-C}{3N} [N + 1 - (2N - 1)\chi], \\ &\lambda > \frac{B+C}{N-2} [N + 1 - (2N - 1)\chi]\}. \end{aligned} \quad [3]$$

Here, ϕ is left unconstrained, except that it must produce a feasible strategy. These analytical conditions for robustness are confirmed by Monte Carlo simulations (Fig. S2). Setting $\lambda=0$ in the equation above recovers the conditions we previously derived for the robustness of ZD strategies. As in the case of ZD, the robust good strategies are exclusively limited to generous strategies (i.e., strategies with $\kappa=B-C$ and $\chi > 0$; Fig. 3).

Interestingly, the strategy win-stay-lose-shift does not lie within the region of robust good strategies (Fig. 3). As a concrete demonstration of this result, we have identified a specific strategy that selectively replaces win-lose-stay-shift in a finite population (SI Text and Fig. S3). Furthermore, even under strong selection, and under increased mutation rates, win-stay-lose-shift can be dominated by some strategies (Fig. S3).

Evolutionary Success of Generosity. We have shown that generous strategies are evolutionarily robust, and eventually dominate in a population, when players are confined to the space of ZD strategies. We have also shown that among the good strategies, which stabilize cooperative behavior, the evolutionary robust strategies, G_R , are also generous. To complement these results, we now systematically query the evolutionary success of generous

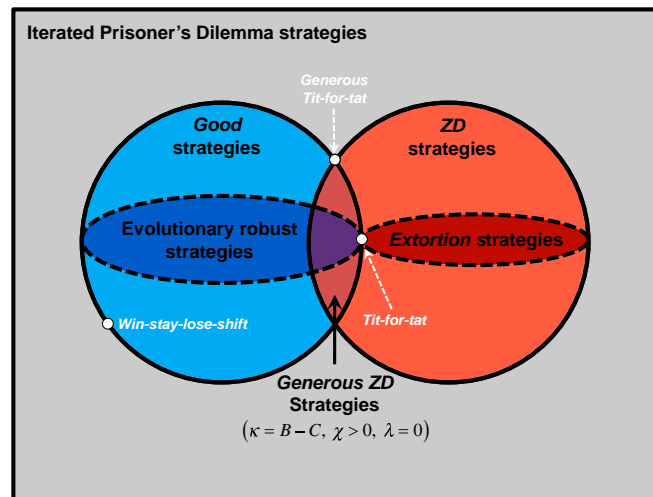


Fig. 2. Relationship between ZD and good strategies in the IPD. The intersection between ZD and good is precisely the set of generous ZD strategies. Not all good strategies are generous. As a result, only a strict subset of good strategies is evolutionary robust, just as a strict subset of ZD strategies is evolutionary robust. Extortion strategies are neither generous nor evolutionary robust. Also shown are the locations of the classic IPD strategies (19) win-stay-lose-shift, tit-for-tat, and generous tit-for-tat.

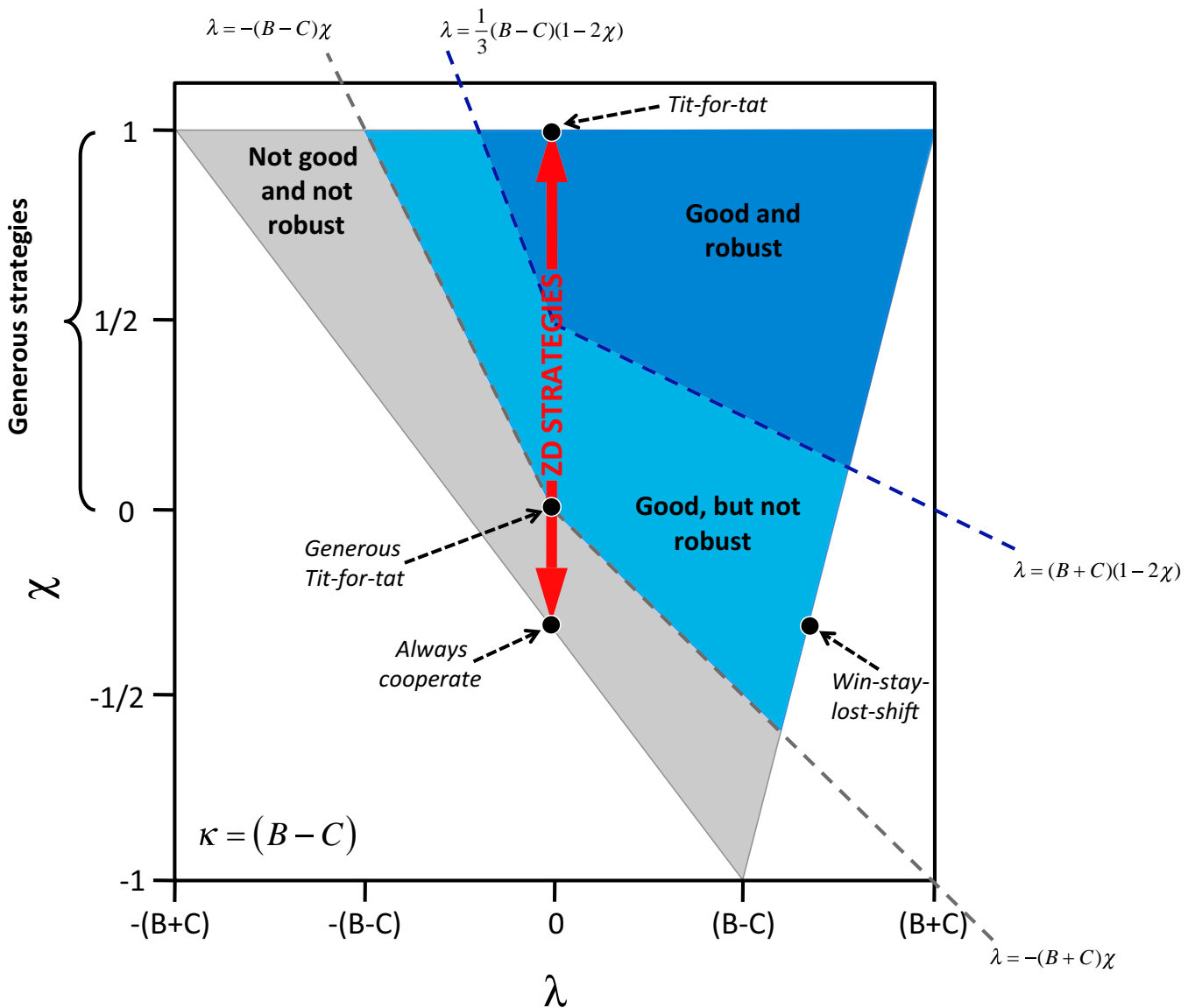


Fig. 3. Space of all cooperative IPD strategies, projected onto the parameters χ and λ . The boundary of the simplex delineates the set of feasible strategies with $\kappa = B - C$. Strategies colored light blue or dark blue are good, whereas strategies colored dark blue are both good and evolutionary robust, under weak selection. Setting $\lambda = 0$ recovers the space of cooperative ZD strategies (red line). Note that all robust strategies are generous (i.e., $\chi > 0$, $\kappa = B - C$). Each point in the figure, $\{\chi, \lambda\}$, has an associated range of ϕ values, and thus corresponds to multiple IPD strategies. However, the evolutionary robust good strategies resist replacement by any other strategy, regardless of the choice of ϕ . The figure illustrates the robust region for a large population size, whereas Eq. 3 gives the exact N -dependent conditions for robustness. Also shown are the locations of several classic IPD strategies. Tit-for-tat ($\chi = 1$) and generous tit-for-tat ($\chi = 0$) are limiting cases of generous ZD strategies, but they are not robust. Likewise, win-stay-lost-shift is good but not robust.

strategies in general by allowing a population to explore the full set of memory-1 strategies $\mathbf{p} = \{p_{cc}, p_{cd}, p_{dc}, p_{dd}\}$ and quantifying how much time the population spends near generosity.

Following Hilbe et al. (12) and Imhof and Nowak (21), we performed simulations in the regime of weak mutation, so that the population is monomorphic for a single strategy at all times. Mutant strategies, drawn uniformly from the space $\{p_{cc}, p_{cd}, p_{dc}, p_{dd}\}$, are proposed at rate μ . A proposed mutant either immediately fixes or is immediately lost from the population, according to its fixation probability calculated relative to the current strategy in the population (12, 20). Over the course of this simulation, we quantified how much time the population spends in a δ -neighborhood of ZD, ZD_R , G , and G_R strategies, as well as extortion strategies (Fig. 4). The δ -neighborhood of a strategy set is defined as those strategies within Euclidean distance δ of it, among the space of all memory-1 strategies. If the proportion of time spent in the δ -neighborhood is

greater than would be expected by random chance (which is proportional to the volume of the δ -neighborhood), evolution is said to favor that set of strategies.

It is already known that except for very small populations, a population spends far less time near extortion strategies than expected by random chance and that the same is true for the set of all ZD strategies (11, 12). Thus, in general, extortion and ZD strategies are disfavored by evolution in populations. This has led to the view that ZD strategies are of importance only in the setting of classical two-player game theory, and not in evolving populations (11, 22). In Fig. 4, we repeat this analysis but additionally report the δ -neighborhoods of ZD_R , G , and G_R strategies. We find that, except in very small populations, selection strongly favors G , G_R , and especially ZD_R strategies. In particular, the population spends more than 100-fold longer in the neighborhood of ZD_R strategies than expected by random chance. Thus, ZD contains a

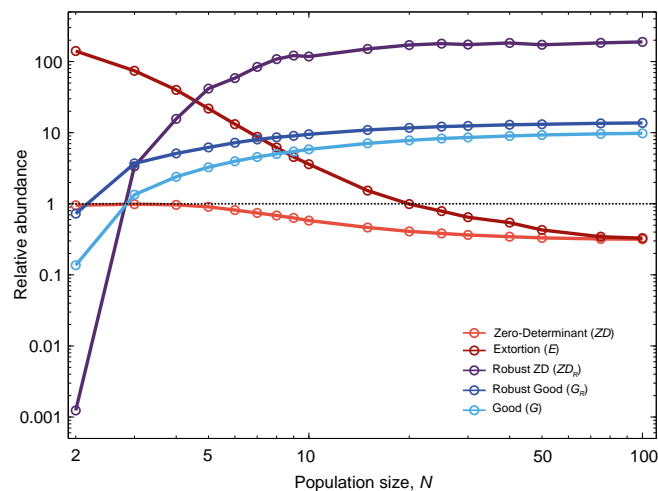


Fig. 4. Generous strategies are favored by selection in evolving populations. We simulated a population under weak mutation, proposing mutant strategies drawn uniformly from the full set of memory-1 IPD strategies. We calculated the time spent in the δ -neighborhood (12) of ZD and extortion strategies, as well as robust ZD strategies, good strategies, and robust good strategies, relative to their random (neutral) expectation. For small population sizes, extortioners are abundant and generous strategies are nearly absent. As population size increases, the frequency of generous strategies and good strategies is strongly amplified by selection, whereas extortion strategies, and ZD strategies in general, are disfavored, as previously reported (12). Simulations were run until the population fixed 10^7 mutations. Parameters are $B=3$, $C=1$, $\delta=0.05$, and selection strength $N\sigma=100$.

subset of strategies that is remarkably successful in evolving populations, in contrast to the claims of Adami and Hintze (11).

We also analyzed the time spent near each individual good strategy, under both weak and strong selection. We found that the strategies most strongly favored by selection are virtually all generous (Fig. S4). The remaining good strategies are typically moderately favored by selection, with the exception of those near win-stay-lose-shift, which are also strongly favored.

Success of Generous Strategies Against Classic IPD Strategies. To complement the weak-mutation studies described above, we also compared the performance of generous ZD strategies against several classic IPD strategies, in a finite population of players (12, 18–20), assuming either strong or weak mutation (i.e., high or low mutation rates). We performed Monte Carlo simulations of populations constrained to different subsets of strategies, similar to those of Hilbe et al. (12). In these simulations, a pair of individuals is chosen from the population at each time step, and the first individual copies the strategy of the second with a probability that depends on their respective payoffs (Table S1), as above. Mutations also occur, with probability μ , so that the mutated individual randomly adopts another strategy from the set of strategies being considered. We ran simulations at a variety of population sizes, ranging from $N=2$ to $N=1,000$.

At very small population sizes, defector strategies tend to dominate (Fig. S5), reflecting the fact that extortion pays in the classic two-player setting (1). However, as the population size increases, good strategies, such as win-stay-lose-shift, and generous ZD quickly begin to dominate (Fig. S5). Which strategy does best depends on the population size, the mutation rate, and the set of available strategies (Figs. S5 and S6). In some regimes, generous ZD strategies even outperform win-stay-lose-shift (Fig. S5).

Discussion

We have shown that generous strategies tend to dominate in evolving populations of IPD players. This is a surprising result

because, when faced with a defector strategy, generous strategies must, by definition, suffer a greater reduction in payoff than their opponent suffers. One might expect such strategies to be vulnerable to replacement by defector strategies, whereas, in fact, we have shown that the reverse is true. Likewise, one might expect generous strategies to be unsuccessful at displacing resident strategies in a population. However, simulations reveal (Figs. S7 and S8) that most generous strategies can selectively replace almost all other IPD strategies.

How can we account for the remarkable evolutionary success of generosity? First, it is important to note that the most successful generous strategies are not too generous. For example, in a large population, evolutionary robust ZD strategies must have $\chi > 1/2$; that is, they must reduce their payoff when faced with a defector opponent but not by too much. Second, although generous strategies score less than defector strategies in head-to-head matches, they are able to limit the difference between their own payoff and their opponent's payoff (Materials and Methods). As a result, they tend to have a consistent probability $\rho > 1/N$ of replacing a diversity of resident IPD strategies (Fig. S8), allowing them to succeed in an evolutionary setting.

We found that generous ZD strategies are particularly successful when mutations arise at an appreciable rate. Under such circumstances, ZD_R strategies can dominate even win-stay-lose-shift, a perennial favorite in evolving populations (7, 11, 23, 24). Overall, selection strongly favors generous ZD strategies when evolution proceeds in the full space of memory-1 strategies. These results strongly contravene the view that ZD strategies are of little evolutionary importance (11, 22). In fact, we have shown that a subset of ZD strategies, the generous ones, is strongly favored in the evolutionary setting.

The discovery and elegant definition of ZD strategies remains a remarkable achievement, especially in light of decades worth of prior research on the Prisoner's Dilemma in both the two-player and evolutionary settings. ZD strategies comprise a variety of new ways to play the IPD, and Akin's generalization of cooperative ZD to good strategies (13) provides novel insight into how cooperation between two rational players can be stabilized. However, in an evolutionary setting, among both ZD and good strategies, it is the generous ones that are most successful.

Materials and Methods

Notation. For ease of analysis, the parameter χ we use throughout is the inverse of that used by Press and Dyson (1). In addition, to avoid confusion with δ -neighborhoods, we use λ in place of Akin's δ (13).

Evolutionary Simulations. We simulated a well-mixed population in which selection follows an "imitation" process (12, 20). At each discrete time step, a pair of individuals (X, Y) is chosen at random. X switches its strategy to imitate Y with probability $f_{X \rightarrow Y}$:

$$f_{X \rightarrow Y} = \frac{1}{1 + \exp[\sigma(s_X - s_Y)]},$$

where s_X and s_Y denote the average IPD payoffs of players X and Y against the entire population and σ denotes the strength of selection. When a mutant strategy X is introduced to a population otherwise consisting of a resident strategy Y , its probability of fixation, ρ , is given by

$$\rho = \frac{1}{\sum_{i=0}^{N-1} \prod_{j=1}^i e^{\sigma[(j-1)s_{YY} + (N-j)s_{YX} - \beta s_{YY} - (N-j-1)s_{XX}]}}, \quad [4]$$

Taylor expansion to first order about $\sigma=0$ gives Eq. 1, the condition for selective replacement of Y by X under weak selection.

Evolutionary Robustness of Cooperative ZD Strategies. Suppose that a resident strategy Y is cooperative and ZD. We will show that Y is evolutionary robust if and only if $\chi \geq (N+1)/(2N-1)$. From Eq. 1, we deduce that Y is robust against any mutant IPD strategy X if their payoffs satisfy

$$s_{xy}(2N-1) \leq s_{yx}(N+1) + (B-C)(N-2).$$

$$\chi_x > \frac{N+1}{2N-1} \quad [6]$$

Using Eq. 2 to substitute for s_{yx} yields the equivalent condition

$$[(B-C) - s_{xy}](\chi(2N-1) - (N+1)) \geq 0.$$

Furthermore, we know that $B-C \geq s_{xy}$ for any mutant X (because, otherwise, Eq. 2 would imply that both s_{yx} and s_{xy} exceed $B-C$, which contradicts the assumption $2R > T+S$). Therefore, the cooperative ZD strategy Y is robust if and only if $\chi \geq (N+1)/(2N-1)$.

Noncooperative ZD Strategies Can Be Selectively Replaced. Here, we show that a resident ZD strategy Y with $\kappa_y < B-C$ is selectively replaced by a ZD strategy X with $\kappa_x = B-C$ and $\chi_x > (N+1)/(2N-1)$. Because both players are ZD, their payoffs satisfy the equations

$$\begin{aligned} \chi_x s_{xy} - \chi_x \kappa_x &= s_{yx} - \kappa_x \\ \chi_y s_{yx} - \chi_y \kappa_y &= s_{xy} - \kappa_y, \end{aligned}$$

which result in the payoff matrix

	X	Y
X	κ_x	$\frac{(1-\chi_y)\kappa_y + \chi_y(1-\chi_x)\kappa_x}{(1-\chi_y\chi_x)}$
Y	$\frac{(1-\chi_x)\kappa_x + \chi_x(1-\chi_y)\kappa_y}{(1-\chi_y\chi_x)}$	κ_y

Substituting these payoffs into Eq. 1 shows that X can selectively replace Y if

$$\left[(2(1-\chi_x)(1+\chi_y - N\chi_y) + (1-\chi_y)(1+\chi_x - N\chi_x)) \right] \times (\kappa_x - \kappa_y) < 0. \quad [5]$$

By our assumptions on X and Y , $\kappa_x - \kappa_y > 0$ and $\chi_x > 1/(N-1)$. If $\chi_y > 1/(N-1)$, inequality 5 is satisfied, and so X is selected to replace Y . If $\chi_y < 1/(N-1)$, to determine the conditions for which X is selected to replace Y , we make the coordinate transformations $\chi_y^* = \chi_y - \frac{1}{N-1}$ and $\chi_x^* = \chi_x - \frac{1}{N-1}$, so that $\chi_x^* > 0$ and $\chi_y^* < 0$. The inequality 5 is then satisfied provided

$$-2\left(1 - \chi_x^* - \frac{1}{N-1}\right)\chi_y^* - \left(1 - \chi_y^* - \frac{1}{N-1}\right)\chi_x^* < 0.$$

Rearranging this gives

$$\chi_x^* > \frac{-2\left(1 - \frac{1}{N-1}\right)\chi_y^*}{\left(1 - 3\chi_y^* - \frac{1}{N-1}\right)},$$

which is hardest to satisfy when χ_y^* is at its minimum (i.e., $\chi_y^* = -1 - \frac{1}{N-1}$). This results in the inequality

- Press WH, Dyson FJ (2012) Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proc Natl Acad Sci USA* 109(26):10409–10413.
- Rapoport A, Chammah AM (1965) *Prisoner's Dilemma: A Study in Conflict and Cooperation* (Univ of Michigan Press, Ann Arbor, MI).
- Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211(4489):1390–1396.
- Axelrod R (1984) *The Evolution of Cooperation* (Basic Books, New York).
- Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 50:533–554.
- Fudenberg D, Maskin E (1990) Evolution and cooperation in noisy repeated games. *Am Econ Rev* 80:274–279.
- Nowak M, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364(6432):56–58.
- Trivers R (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46(1):35–71.
- Boerlijst MC, Nowak MA, Sigmund K (1997) Equal pay for all prisoners. *Am Math Mon* 104:303–307.
- Stewart AJ, Plotkin JB (2012) Extortion and cooperation in the Prisoner's Dilemma. *Proc Natl Acad Sci USA* 109(26):10134–10135.
- Adami C, Hintze A (2012) Winning isn't everything: Evolutionary stability of zero determinant strategies. *Nature Communications* 4, 10.1038/ncomms3193.
- Hilbe C, Nowak MA, Sigmund K (2013) Evolution of extortion in iterated Prisoner's Dilemma games. *Proc Natl Acad Sci USA* 110(17):6913–6918.

as a sufficient condition for X to replace Y selectively. This sufficient condition is met by our assumption on X . Thus, noncooperative ZD strategies can always be selectively replaced, provided $N > 2$.

Generous Strategies Limit the Difference Between Their Payoff and Their Opponent's Payoff. Consider Eq. 2 for a generous ZD strategy Y facing an arbitrary opponent X :

$$\phi[s_{xy} - \chi s_{yx} - (1-\chi)(B-C)] = 0.$$

Rearranging this expression gives the difference in the players' payoffs:

$$s_{xy} - s_{yx} = (1-\chi)((B-C) - s_{yx}).$$

Increasing χ reduces the difference between two players' payoffs, regardless of the opponent's strategy. This is also true for generous good strategies, which satisfy

$$s_{xy} - s_{yx} = (1-\chi)((B-C) - s_{yx}) - \lambda(v_{cd} + v_{dc}),$$

where v_{cd} denotes the equilibrium rate of the play (cd) and v_{dc} the equilibrium rate of the play (dc) (13). The ability of generous strategies to limit the difference in payoffs with arbitrary opponents accounts for their remarkable consistency as invaders, as exemplified in Fig. S8. On the other hand, a nongenerous strategy, such as win-stay-lose-shift, is subject to larger differences between one player's payoff and her opponent's payoff, leading to less consistent success as an invader (Fig. S8).

Long-Memory Strategies. Our results for the evolutionary success of generous strategies in a finite population also hold against longer memory opponents. As per Press and Dyson (1), from the perspective of a memory-1 player, a long-memory opponent is equivalent to a memory-1 opponent. Thus, the payoff s_{xy} can be determined by considering only the set of memory-1 strategies. However, the payoff a long-memory opponent receives against itself, s_{xx} , may depend on its memory capacity. Nonetheless, under the standard IPD assumption $2R > T+S$, the highest total payoff for any pair of players in the IPD is $2R$; thus, $s_{xx} \leq R$. This condition on s_{xx} is the only condition required to derive our results on the robustness of ZD and good strategies (SI Text), and so our results continue to hold even against long-memory invaders.

ACKNOWLEDGMENTS. We thank William Press, Freeman Dyson, Karl Sigmund, and Christian Hilbe, and two anonymous referees for constructive feedback. We gratefully acknowledge support from the Burroughs Wellcome Fund, the David and Lucile Packard Foundation, the James S. McDonnell Foundation, the Alfred P. Sloan Foundation, the Foundational Questions in Evolutionary Biology Fund (Grant RFP-12-16), the US Army Research Office (Grant W911NF-12-1-0552), and Grant D12AP00025 from the US Department of the Interior.

- Akin E (2012) Stable cooperative solutions for the iterated prisoner's dilemma. *arXiv*:1211.0969.
- Maynard Smith J, Price GR (1973) The logic of animal conflict. *Nature* 246:15–18.
- Maynard Smith J (1982) *Evolution and the Theory of Games* (Cambridge Univ Press, Cambridge, U.K.).
- Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics* (Cambridge Univ Press, Cambridge, U.K.).
- Boyd R, Gintis H, Bowles S (2010) Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science* 328(5978):617–620.
- Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428(6983):646–650.
- Nowak MA (2006) *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap Press of Harvard Univ Press, Cambridge, MA).
- Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E Stat Nonlin Soft Matter Phys* 74(1 Pt 1):011909.
- Imhof LA, Nowak MA (2010) Stochastic evolutionary dynamics of direct reciprocity. *Proc Biol Sci* 277(1680):463–468.
- Ball P (2012) Physicists suggest selfishness can pay. Available at www.nature.com/news/physicists-suggest-selfishness-can-pay-1.11254. Accessed August 22, 2013.
- Imhof LA, Fudenberg D, Nowak MA (2007) Tit-for-tat or win-stay, lose-shift? *J Theor Biol* 247(3):574–580.
- Iliopoulos D, Hintze A, Adami C (2010) Critical dynamics in the evolution of stochastic strategies for the iterated prisoner's dilemma. *PLOS Comput Biol* 6(10):e1000948.