

# Evolution of extortion in Iterated Prisoner's Dilemma games

Christian Hilbe<sup>a</sup>, Martin A. Nowak<sup>b</sup>, and Karl Sigmund<sup>c,d,1</sup>

<sup>a</sup>Evolutionary Theory Group, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany; <sup>b</sup>Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138; <sup>c</sup>Faculty of Mathematics, University of Vienna, A-1090 Vienna, Austria; and <sup>d</sup>International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

Edited by Kenneth Wachter, University of California, Berkeley, CA, and approved March 15, 2013 (received for review August 29, 2012)

Iterated games are a fundamental component of economic and evolutionary game theory. They describe situations where two players interact repeatedly and have the ability to use conditional strategies that depend on the outcome of previous interactions, thus allowing for reciprocation. Recently, a new class of strategies has been proposed, so-called “zero-determinant” strategies. These strategies enforce a fixed linear relationship between one’s own payoff and that of the other player. A subset of those strategies allows “extortioners” to ensure that any increase in one player’s own payoff exceeds that of the other player by a fixed percentage. Here, we analyze the evolutionary performance of this new class of strategies. We show that in reasonably large populations, they can act as catalysts for the evolution of cooperation, similar to tit-for-tat, but that they are not the stable outcome of natural selection. In very small populations, however, extortioners hold their ground. Extortion strategies do particularly well in coevolutionary arms races between two distinct populations. Significantly, they benefit the population that evolves at the slower rate, an example of the so-called “Red King” effect. This may affect the evolution of interactions between host species and their endosymbionts.

replicator dynamics | adaptive dynamics

The Iterated Prisoner’s Dilemma (IPD) has a long history as a model for the cultural and biological evolution of cooperation (1–9). A new class of so-called “zero-determinant” (ZD) strategies has recently attracted considerable attention (10–12). Such strategies allow players to enforce a linear relation unilaterally between one player’s own payoff and the coplayer’s payoff. A subset consists of the so-called “equalizer” strategies, which assign to the coplayer’s score a predetermined value, independent of the coplayer’s strategy (13). Another subset consists of the extortion strategies, which guarantee that one player’s own surplus exceeds the coplayer’s surplus by a fixed percentage. Press and Dyson (10) have explored the power of ZD strategies to manipulate any “evolutionary” opponent (i.e., any coplayer able to learn and to adapt).

In Stewart and Plotkin’s (11) commentary to the article by Press and Dyson (10), they ask: “What does the existence of ZD strategies mean for evolutionary game theory: Can such strategies naturally arise by mutation, invade, and remain dominant in evolving populations?” In evolutionary game theory, it is the population that adapts: More and more players switch to the more successful strategies. From the outset, it may seem that the opportunities for extortion strategies are limited. If a strategy is successful, it will spread, and therefore be more likely to be matched against its like, but any two extortioners hold each other down to surplus zero. In a homogeneous population of extortioners, it is thus better to deviate by cooperating. Extortion is therefore evolutionarily unstable (12). However, we shall see that if the two players engaged in an IPD game belong to distinct populations, the evolutionary prospects of extortion improve significantly.

In the following, we investigate the impact of ZD strategies on evolutionary game theory. We show that in large, well-mixed populations, extortion strategies can play an important role, but only as catalysts for cooperation and not as a long-term outcome. However, if the IPD game is played between members of two

separate populations evolving on different time scales, extortion strategies can get the upper hand in whichever population evolves more slowly and enable it to enslave the other population, an interesting example of the so-called “Red King” effect (14).

The Prisoner’s Dilemma (PD) game is a game between two players *I* and *II* having two strategies each, which we denote by *C* (“to cooperate”) and *D* (“to defect”). It is assumed that the payoff for two cooperating players, *R*, is larger than the payoff for two defecting players, *P*. If one player cooperates and the other defects, the defector’s payoff *T* is larger than *R* and the cooperator’s payoff *S* is smaller than *P*. Thus, the game is defined by  $T > R > P > S$ . An important special case is the so-called “donation game,” where each player can “cooperate” (play *C*) by providing a benefit *b* to the other player at his or her cost *c*, with  $0 < c < b$ . Then,  $T = b$ ,  $R = b - c$ ,  $P = 0$ , and  $S = -c$ .

In the IPD game, the two players are required to play an infinite number of rounds, and their payoffs *P<sub>I</sub>* respectively (resp.) *P<sub>II</sub>* are given by the limit in the mean of the payoffs per round. An important class of strategies consists of so-called “memory-one” strategies. They are given by the conditional probabilities *p<sub>R</sub>*, *p<sub>S</sub>*, *p<sub>T</sub>*, and *p<sub>P</sub>* to play *C* after experiencing outcome *R*, *S*, *T* resp. *P* in the previous round. [In addition, such a strategy has to specify the move in the first round, but this has only a transient effect and plays no role in the long run (15)]. An important class of memory-one strategies consists of reactive strategies, which only depend on the coplayer’s move in the previous round (not one’s own move). Then,  $p_R = p_T =: p$  and  $p_P = p_S =: q$ , such that a reactive strategy corresponds to a point  $(p, q)$  in the unit square (16).

We will first define and characterize ZD strategies, equalizers, and extortioners. We then investigate, in the context of evolutionary game theory, the contest between extortioners and four of the most important memory-one strategies. We will show that extortion cannot be an outcome of evolution but can catalyze the emergence of cooperation. The same result will then be obtained if we consider all memory-one strategies. Hence, extortion strategies can only get a foothold if the population is very small. If the IPD game is played between members of two distinct populations, ZD strategies can emerge in the population that evolves more slowly. In particular, extortion strategies can allow host species to enslave their endosymbionts.

## Methods and Results

**Definitions.** Press and Dyson (10) define the class of ZD strategies as those memory-one strategies  $(p_R, p_T, p_S, p_P)$  satisfying, for some real values  $\alpha, \beta, \gamma$ , the equations

$$p_R - 1 = \alpha R + \beta T + \gamma P \quad [1A]$$

Author contributions: C.H., M.A.N., and K.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: karl.sigmund@univie.ac.at.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1214834110/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1214834110/-DCSupplemental).

$$p_S - 1 = \alpha S + \beta T + \gamma$$

$$p_T = \alpha T + \beta S + \gamma$$

$$p_P = \alpha P + \beta P + \gamma.$$

We note that  $1-p_R$  and  $1-p_S$  are the probabilities to switch from  $C$  to  $D$ , whereas  $p_T$  and  $p_P$  are the probabilities to switch from  $D$  to  $C$ . Press and Dyson (10) showed that if player  $I$  uses such a ZD strategy, then

$$\alpha P_I + \beta P_{II} + \gamma = 0, \quad [2]$$

no matter which strategy player  $II$  is using. Equalizer strategies are those ZD strategies for which  $\alpha=0\neq\beta$ , then

$$P_{II} = -\gamma/\beta. \quad [3]$$

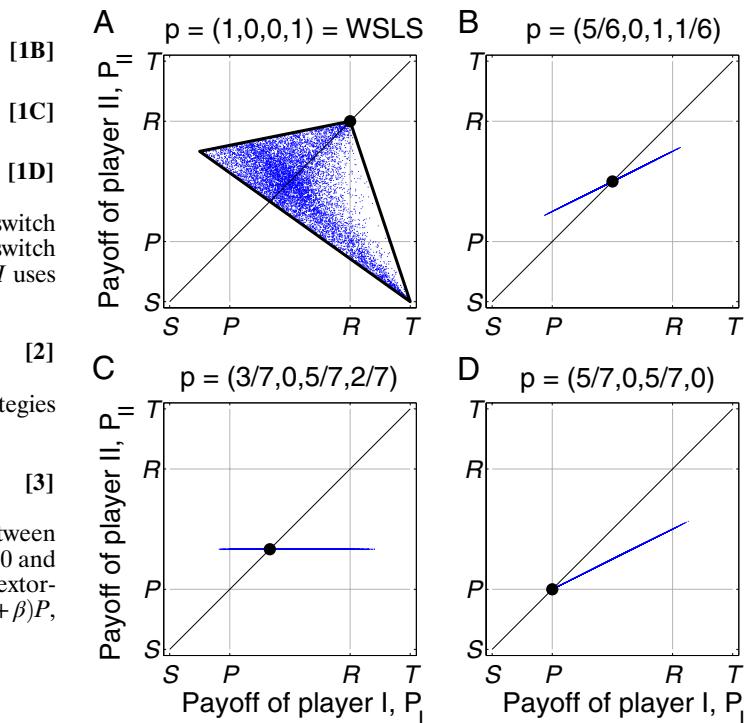
Thus, player  $I$  can assign to the coplayer any payoff between  $P$  and  $R$ . (Indeed, because the  $p_i$  values have to be between 0 and 1, it follows that  $\beta < 0$  and  $P \leq P_{II} \leq R$ ). The so-called “ $\chi$ -extortion” strategies are those ZD strategies for which  $\gamma = -(\alpha + \beta)P$ , with  $\chi := -\beta/\alpha > 1$ . Then,

$$P_I - P = \chi(P_{II} - P).$$

In this case, player  $I$  can guarantee that his or her own “surplus” (over the maximin value  $P$ ) is the  $\chi$ -fold of the coplayer’s surplus. Fig. 1 shows examples of these different ZD strategies.

Press and Dyson (10) speak of ZD strategies because they use for their proof of Eq. 2 an ingenious method based on determinants. In Appendix A, we present a more elementary proof, following the method of Boerlijst et al. (13). Within the 4D unit cube of all memory-one strategies  $(p_R, p_S, p_T, p_P)$ , the ZD strategies form a 3D subset  $\mathcal{ZD}$  containing the 2D subsets  $\mathcal{EQ}$  and  $\mathcal{EX}$  of equalizers resp. extortioners (Appendix B). In Fig. 2, we sketch these sets for the reactive strategies.

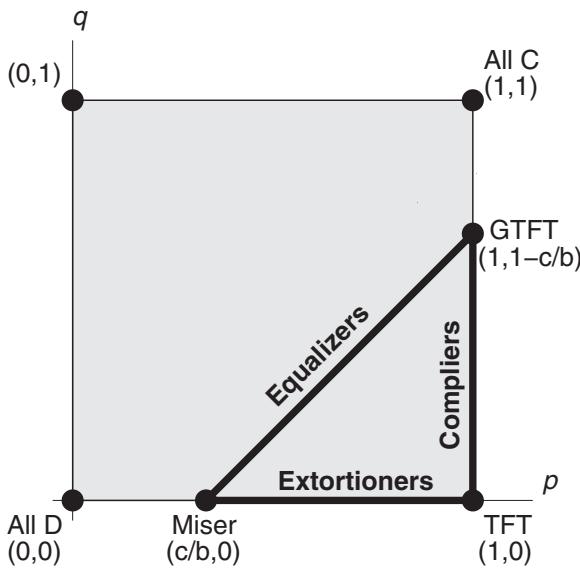
**Extortion Within One Population.** To investigate the role of extortion in the context of evolutionary games, we concentrate on the donation game (in SI Text, we provide further results for the general IPD, which show that the main conclusions are independent of special characteristics of the donation game). We first consider how a  $\chi$ -extortion strategy  $E_\chi$  fares against some of the most important memory-one strategies, namely, tit for tat [ $TFT = (1, 0, 1, 0)$ ], always defect [ $All D = (0, 0, 0, 0)$ ], always cooperate [ $All C = (1, 1, 1, 1)$ ] and the win-stay-lose-shift strategy



**Fig. 1.** Payoffs  $P_I$  and  $P_{II}$  of players  $I$  and  $II$  if both players use memory-one strategies in an IPD game (with  $T = 3$ ,  $R = 2$ ,  $P = 0$ , and  $S = -1$ ). In each graph, the strategy of player  $I$  is fixed to some  $p$ , whereas the strategy  $q$  of the coplayer  $II$  can vary, sampling the 4D cube of memory-one strategies (the blue dots correspond to  $10^4$  different realizations of  $q$ ). (A) In general, the payoff-pairs of the two players cover a 2D area, as here, when player  $I$  applies the strategy of WSLS (i.e.,  $p_R = p_P = 1$  and  $p_S = p_T = 0$ ). (B) However, if player  $I$  adopts a ZD strategy, the possible payoff-pairs are restricted to a line. Two special classes of ZD strategies were highlighted by Press and Dyson (10): equalizers [strategies that set the coplayer’s score to a fixed value (the line of payoffs has slope zero)] (C), and extortioners [strategies that guarantee the surplus of player  $I$  is the  $\chi$ -fold of the surplus of player  $II$  (i.e.,  $P_I - P = \chi(P_{II} - P)$ , with  $\chi > 1$  (the line of payoffs has a positive slope and intersects the diagonal at  $P$ )] (D).

WSLS, which is encoded by  $(1, 0, 0, 1)$ , and hence cooperates if and only if the coplayer’s move in the previous round was the same as one’s own move (7). We note that TFT is a ZD strategy and can be viewed as a limiting case of an extortion strategy, with  $\chi=1$ . For the donation game, the payoff for a player using strategy  $i$  against a player using strategy  $j$  is given by the  $(i,j)$ th element of the following matrix:

	<i>TFT</i>	<i>WSLS</i>	$E_\chi$	<i>All C</i>	<i>All D</i>	
<i>TFT</i>	$(b-c)/2$	$(b-c)/2$	0	$b-c$	0	
<i>WSLS</i>	$(b-c)/2$	$b-c$	$\frac{b^2-c^2}{b(1+2\chi)+c(2+\chi)}$	$(2b-c)/2$	$-c/2$	
$E_\chi$	0	$\frac{(b^2-c^2)\chi}{b(1+2\chi)+c(2+\chi)}$	0	$\frac{(b^2-c^2)\chi}{b\chi+c}$	0	
<i>All C</i>	$b-c$	$(b-2c)/2$	$\frac{b^2-c^2}{b\chi+c}$	$b-c$	$-c$	
<i>All D</i>	0	$b/2$	0	$b$	0	



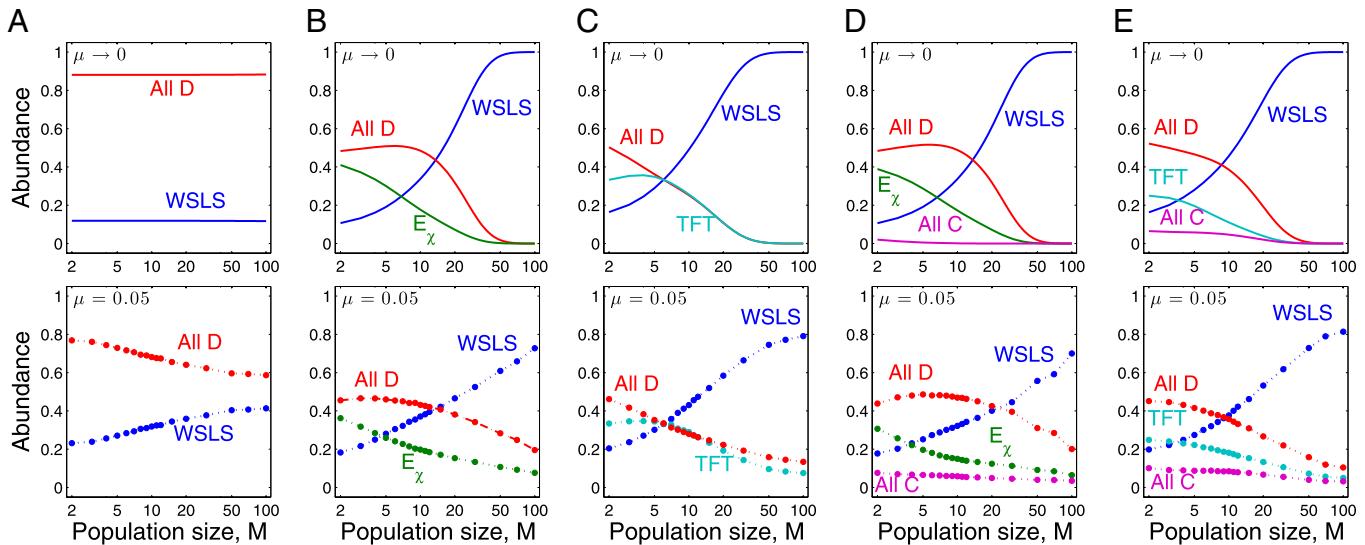
**Fig. 2.** Reactive strategies ( $p_R = p_T = p$ ,  $p_S = p_P = q$ ) for the donation game. All reactive strategies (the square  $0 \leq p, q \leq 1$ ) are ZD strategies. The equalizer strategies are those on the segment between “generous TFT” (GTFT) ( $p=1$ ,  $q=1-c/b$ ) (16) and “Miser” ( $p=c/b$ ,  $q=0$ ) (30), the extortion strategies are those between Miser and TFT ( $p=1$ ,  $q=0$ ), and the “compliant” strategies (Discussion and ref. 11) are those between GTFT and TFT.

Let us start with the pairwise comparisons. The extortioner strategy  $E_\chi$  is neutral with respect to *All D*. It is weakly dominated by *TFT*, in the sense that a *TFT* player does not fare better than an extortioner against extortioners but that interactions with other *TFT* players are giving an advantage to *TFT*. *All C* players can invade extortioners, and vice versa: These two strategies can stably coexist in proportions  $c(\chi-1):(b+c)$ . Finally, *WSLS* dominates extortioners (in the sense that *WSLS* provides a better response

than extortion against itself and against extortioners). We note that the mixed equilibrium of extortioners and unconditional cooperators can be invaded by each of the other three strategies. The same holds for the mixed equilibria of extortioners and unconditional defectors if the frequency of extortioners is sufficiently high. In particular, *TFT* can always invade such a mixed equilibrium but can, in turn, be invaded by *WSLS* or *All C*. No Nash equilibrium involves  $E_\chi$ . If  $b < 2c$ , there are two Nash equilibria: a mixture of *TFT*, *All C*, and *All D*, and a mixture of *TFT*, *WSLS*, and *All D*. If  $b > 2c$ , there exist four Nash equilibria. In particular, *WSLS* is then a strict Nash equilibrium.

The replicator dynamics (17) displays for the payoff matrix continuous families of fixed points and periodic orbits, and hence is far from being structurally stable: Small changes in the dynamics can lead to vastly different outcomes. The same applies to most other deterministic game dynamics (18). It seems more reliable to consider a stochastic process that describes a finite, well-mixed population consisting of  $M$  players and evolving via copying of successful strategies and exploration (i.e., by a selection-mutation process) (19–21). Selection is viewed here as an imitation process; in each time step, two randomly chosen players  $A$  and  $B$  compare their average payoffs  $P_A$  and  $P_B$ , and  $A$  switches to  $B$ 's strategy with a probability given by  $(1+\exp[s(P_A-P_B)])^{-1}$ , where  $s \geq 0$  corresponds to “selection strength.” (As shown in SI Text, the details of the imitation process matter little.) Additionally, mutations occur with a small probability  $\mu > 0$  (corresponding to the adoption of another strategy, with each alternative being equally likely). Any such stochastic process yields a steady-state distribution of strategies.

We find that although extortioners are never the most abundant strategy, they can play the role of a catalyst. Indeed, if only *All D* and *WSLS* are available, a population may be trapped in a noncooperative state for a considerable time, leading to a mutation-selection equilibrium that clearly favors defectors (Fig. 3A). In such a case, extortioners (Fig. 3B) and *TFT* (Fig. 3C) offer an escape: These strategies can subvert an *All D* population through neutral drift and selection, respectively. Once defectors are rare, *WSLS* outperforms *TFT*, and it also prevails against extortioners if the population is sufficiently large (in a direct competition,



**Fig. 3.** Evolutionary competition between some important strategies in the IPD. For various population sizes  $M$ , the graphs show the frequency of each strategy in the mutation-selection equilibrium. We consider two mutation regimes: the limit of rare mutations  $\mu \rightarrow 0$  (Upper), for which the equilibrium can be calculated analytically using the method of Fudenberg and Imhof (23), and a regime with mutation rate  $\mu = 0.05$  (Lower), which is explored by individual-based simulations. For the copying process, we assume that individuals  $A$  and  $B$  are chosen randomly.  $A$  switches to  $B$ 's strategy with a probability given by  $(1+\exp[s(P_A-P_B)])^{-1}$ , where  $P_A$  and  $P_B$  are the corresponding payoff values and  $s \geq 0$  corresponds to “selection strength” (cf. ref. 21). (A) If *All D* competes with *WSLS*, the population is mostly in the defector's state, independent of population size and the mutation rate. (B and C) However, once  $E_\chi$  or *TFT* is added, *WSLS* succeeds if populations are sufficiently large. *TFT* works slightly better than  $E_\chi$ . (D and E) Adding *All C* only leads to minor changes in the stationary distribution, which now slightly favors  $E_\chi$ . The parameters are  $b=3$ ,  $c=1$ ,  $s=1$ , and  $\chi=2$ .

*WSLS* always gets a higher payoff than  $E_\chi$  if  $M > 1 + \chi$ ). Thus, in large populations, extortioners and *TFT* players tip the mutation-selection balance toward *WSLS*, and therefore increase the level of cooperation. Further expansion of the strategy space through adding *All C* has only a small effect on the steady state (Fig. 3D and E), slightly favoring extortioners.

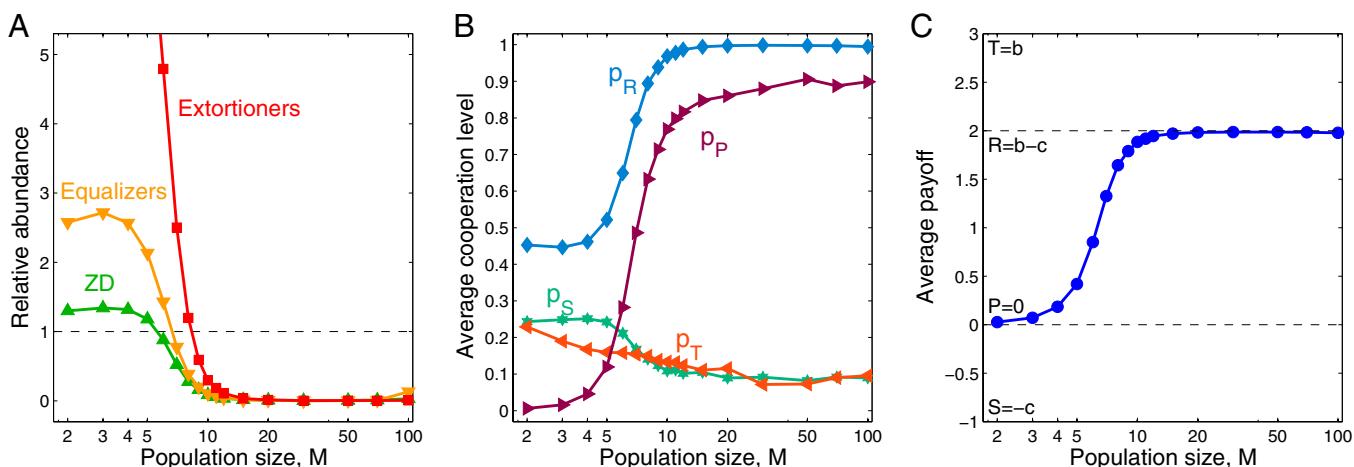
What happens when players are not restricted to the five specific strategies considered so far but can choose among all possible memory-one strategies? We study this by using the stochastic evolutionary dynamics of Imhof and Nowak (22), assuming that mutants can pick up any memory-one strategy, with a uniform probability distribution on the 4D unit cube. We further assume that the mutant reaches fixation, or is eliminated, before the next mutation occurs. Overall, this stochastic process leads to a sequence of monomorphic populations. The evolutionary importance of a given strategy can then be assessed by computing how often the state of the population is in its neighborhood. For a subset  $A$  of the set of memory-one strategies, we denote the  $\delta$ -neighborhood of  $A$  (with respect to Euclidean distance) by  $A_\delta$ , and let  $\mu(A_\delta)$  denote the fraction of time that the evolving population visits  $A_\delta$ . We say that  $A_\delta$  is favored by selection if the evolutionary process visits  $A_\delta$  more often than expected under neutral evolution, [i.e., if  $\mu(A_\delta)$  is larger than the volume of the intersection of  $A_\delta$  with the unit cube of all memory-one strategies]. We apply this concept to  $A = \mathcal{ZD}, \mathcal{EQ}, \mathcal{EX}$ .

Extensive simulations indicate that neither extortioners nor equalizers or ZD strategies are favored by selection if the population is reasonably large (Fig. 4A). By contrast, very small population sizes promote the selection of these behaviors. For extortioners, this result is intuitive: In small populations, the fact that self-interactions are excluded yields greater weight to interactions with players using the rival strategy rather than interactions with players using one's own strategy (19); this effect may even result in the evolution of spite (24, 25). We address this point in more detail in *SI Text (section 2)*. Essentially, both extortioners and equalizers suffer from not achieving maximal payoff  $b - c$  against themselves, which causes their inherent instability, as also stressed by Adami and Hintze (12). The same holds for most ZD strategies. By contrast, *WSLS* players do well against their like, and therefore prevail in the evolutionary dynamics for long periods if the population size is large, at least when  $b > 2c$  or, for more general PD games, when  $2R > T + P$  (15) (Fig. 4B). As a (possibly surprising) consequence, larger populations also yield higher

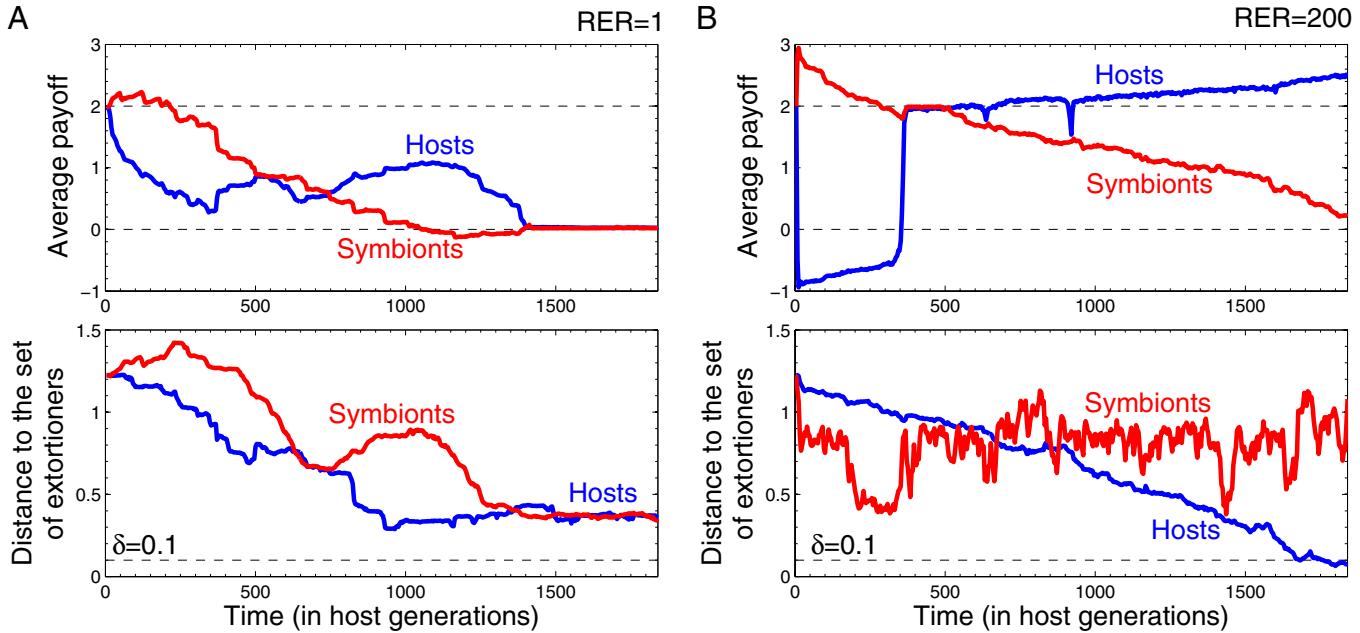
average payoffs (Fig. 4C). In *SI Text*, we show that these qualitative results are robust with respect to changes in parameter values, such as benefits and costs or the strength of selection. Hence, extortion is disfavored by evolution as soon as the population size exceeds a critical level.

**Extortion Between Two Populations.** Let us now consider two species (e.g., hosts and their symbionts) or two classes of a single species (e.g., old and young, buyers and sellers, rulers and subjects) engaged in an IPD game, which, of course, is now unlikely to be symmetrical. In such situations, extortioners may evolve even in large populations. Indeed, extortioners provide incentives to cooperate: As shown by Press and Dyson (10), *All C* is always a best response to an extortion strategy. In a single population of homogeneous players, this is not turned to advantage, because the extortioners' success leads to more interactions with their own kind. If extortioners evolve in one of two separate populations, they will not have to interact with coplayers of their own kind. Nevertheless, their success may be short-lived because they will be tempted to adopt the even more profitable *All D* strategy as a reaction to the *All C* coplayers who they have produced, which, in turns, leads to the disappearance of the *All C* players.

Extortioners can only achieve a lasting (rather than short-lived) success if the rate of adaptation for the host population is much slower than that for the symbionts. To elucidate this point, we extend our previous analysis by revisiting a coevolutionary model of Damore and Gore (26). These authors consider host–symbiont interactions where each host interacts with its own subpopulation of endosymbionts. Let us assume that these interactions are given by an IPD game. Members of both species reproduce with a probability proportional to their fitness (which is an increasing function of their payoffs) by replacing a randomly chosen organism of their species. However, the two populations of hosts and symbionts may evolve on different time scales, as measured by their relative evolutionary rate (RER). For an RER of 1, hosts and symbionts evolve at a similar pace in the evolutionary arms race and no population is able to extort the other (Fig. 5A). This changes drastically as soon as we increase the RER, by allowing symbionts to adapt more quickly. Fast adaptation results in a short-term increase of the symbionts' payoffs, because they can quickly adjust to their respective host. In the long term, however, this induces hosts to adopt extortion strategies (Fig. 5B), thereby forcing their symbionts to cooperate. Thus, it pays off in the long run for



**Fig. 4.** Statistics of the evolutionary dynamics for memory-one strategies for a range of different population sizes. We have calculated the relative abundance of extortioners, equalizers, and ZD strategies (i.e., the time spent in a  $\delta$ -neighborhood) divided by the volume of the intersection of that neighborhood with the set of memory-one strategies (A), the average strategy of the population (B), and the average payoff (C). Extortioners, equalizers, and ZD strategies are only favored for small population sizes. As the population size increases, individuals tend to apply *WSLS*-like strategies and to cooperate only after mutual cooperation or mutual defection. As a result, the average payoff increases with population size. For the simulations,  $10^7$  mutant strategies were randomly drawn from the space of memory-one strategies. As in the study by Imhof and Nowak (22), the switch from a monomorphic population using strategy  $X$  to a monomorphic population using strategy  $Y$  occurred with the probability of fixation of a single  $Y$  mutant in a population of  $X$  residents. The parameters are  $b = 3$ ,  $c = 1$ ,  $\delta = 0.1$ , and  $s = 100$ .



**Fig. 5.** Evolution of extortion in host–symbiont interactions. The graphs show two typical simulation runs for a population of 40 hosts, with each having a subpopulation of 20 symbionts. For each simulation run, one graph (*Upper*) shows the average payoff for each population, whereas the other graph (*Lower*) shows the Euclidean distance of each population to the set of extortioners (which can be 1.5275 at most). In the initial population, all individuals cooperate unconditionally. Further evolution depends on the RER. (A) If RER = 1, both species converge toward *All D* and no population is able to extort the other. (B) For RER = 200, symbionts evolve much more quickly. In the short term, they can thus increase their average payoff by switching to a noncooperative strategy. However, in the long term, hosts apply extortion strategies to force their symbionts to cooperate. Eventually, the hosts' payoff exceeds  $b - c$ , whereas the symbionts' payoff is close to zero. To model the evolutionary process, we followed the method of Damore and Gore (26). Whenever a symbiont reproduces, its offspring remains associated with the same host. Whenever the host reproduces, the new host offspring acquires its symbionts from other hosts (horizontal transmission). Mutations occur with probability  $\mu = 0.05$ , by adding Gaussian noise to an entry of the memory-one strategy of the parent ( $\sigma = 0.05$ ). The process is run for 2,000 host generations (corresponding to more than  $10^6$  reproduction events for RER = 1 and more than  $3 \times 10^8$  reproduction events for RER = 200). The other parameters are  $b = 3$ ,  $c = 1$ , and  $s = 10$ .

the host to be slow to evolve; for the parameters in Fig. 5B, the resulting equilibrium allocates them, on average, a surplus more than 10-fold larger than the surplus achieved by the symbionts.

## Discussion

Our main results show that within one population, extortion strategies can act as catalysts for cooperation but prevail only if the population size is very small, and that in interactions between two populations, extortion can emerge if the rates of evolution differ. This holds not only for the donation game (and therefore whenever  $R + P = T + S$ ) but in considerably more general contexts. In the last part of *SI Text*, we emphasize this robustness. We could also assume that the players alternate their moves in the donation game (27, 28) or that the underlying PD game is asymmetrical (the definitions have to be modified in a straightforward way). As noted by Press and Dyson (10), some results hold also for non-PD games; this deserves further investigation.

In orthodox game theory, strategy *A* dominates strategy *B* if *A* yields at least the payoff of *B* no matter what the coplayer does. When Press and Dyson (10) argue that extortioners dominate their coplayers, they mean that no matter what the coplayer does, the extortioner gets more. This is not quite the same, and we display in *SI Text (section 2)* an example that highlights the difference. Adami and Hintze (12) stress a similar point in their title: “Winning isn’t everything.” Moreover, when Press and Dyson (10) speak of evolutionary players, they refer to players who adapt their strategy in the course of an IPD game, whereas in evolutionary game theory, it is the population that evolves. Thus, Press and Dyson (10) analyzed ZD strategies in the context of classical game theory, with two players locked in contest: Extortion strategies play an important role in this context, as do the more orthodox trigger strategies (3, 6). In the context of evolutionary game theory, whole populations are engaged in the game. For a very small population size, extortion strategies still offer good prospects. This

is not surprising, because the limiting case, a population size  $M = 2$ , reduces to the scenario analyzed by Press and Dyson (10). In larger populations (with our parameter values for  $M > 10$ ), the outcome is different. However, evolutionary game theory can reflect features of classical game theory if the two interacting players belong to two separate evolving populations.

Extortion strategies are only a small subset of ZD strategies. We have seen that within large populations, the class of ZD strategies is not favored by selection, in the sense that its neighborhood is not visited disproportionately often. This does not preclude, of course, that certain elements of the class are favored by selection. Thus, generous *TFT* ( $1, 1 - c/b, 1, 1 - c/b$ ) does well, as do other less known strategies. In particular, Stewart and Plotkin (11) highlighted a class of strategies defined, instead of Eq. 3, by  $P_I - R = \chi(P_H - R)$  (with  $\chi > 1$ ). A player using this strategy does not claim a larger portion of the surplus but a larger share of the loss (relative to the outcome  $R$  of full cooperation). Remarkably, these “compliant” strategies do as well as *WSLS*. They are the only ZD strategies that are best replies against themselves.

In the study by Adami and Hintze (12), the evolutionary stability of several ZD strategies was tested by replicator dynamics and agent-based simulations, which independently confirm the result that these strategies do not prevail in large populations. They used a population size of  $M = 1,024$ , and payoff values of  $R = 3$ ,  $S = 0$ ,  $T = 5$ , and  $P = 1$  (i.e., a PD game that cannot be reduced to a donation game). Adami and Hintze (12) also discuss the evolutionary success of “tag-based” strategies, which use extortion only against those opponents who do not share their tag. These strategies are not memory-one strategies because they depend not only on the previous move; rather, they use memory-one strategies in specific contexts, which depend on the tag. Such a tag is an additional trait that has to evolve and risks being faked.

In interactions between different populations, a cheater-proof tag is provided for free and extortion may accordingly evolve.

In endosymbiotic relationships, as we have seen, the species that evolves at the slower rate gains a disproportionate share of the benefit, an instance of the Red King effect (14, 29, 30). This requires two conditions to be met: Individuals need to come from different populations, and these populations have to evolve on different time scales. If these conditions are fulfilled, extortioner hosts can manipulate their symbionts' evolutionary landscape in such a way that the hosts' and the symbionts' payoffs are perfectly correlated. This ensures that only those symbiont mutants that are beneficial for the host can succeed. In this sense, such hosts apply an evolutionary kind of mechanism design; they create an environment that makes the symbionts' cooperation profitable for the symbionts but even more profitable for themselves.

### Appendix A: Proof of Eq. 2

Let us denote by  $P_I(n)$  and  $P_{II}(n)$  the players' payoffs in round  $n$ ; by  $s_i(n)$  the probability that  $I$  experiences outcome  $i \in \{R, S, T, P\}$  in that round; and by  $q_i(n)$  the conditional probability, given outcome  $i$ , that  $II$  plays  $C$  in round  $n+1$ . By conditioning on round  $n$ , we see that  $s_R(n+1)$  is given by

$$s_R(n)q_R(n)p_R + s_S(n)q_S(n)p_S + s_T(n)q_T(n)p_T + s_P(n)q_P(n)p_P,$$

and  $s_S(n+1)$  is given by

$$\begin{aligned} s_R(n)(1-q_R(n))p_R + s_S(n)(1-q_S(n))p_S \\ + s_T(n)(1-q_T(n))p_T + s_P(n)(1-q_P(n))p_P. \end{aligned}$$

Hence, the probability that  $I$  plays  $C$  in round  $n+1$  [i.e.,  $p_C(n+1) = s_R(n+1) + s_S(n+1)$ ], is given by  $s(n) \cdot \mathbf{p} = \mathbf{s}(n) \cdot [\alpha \mathbf{g}_I + \beta \mathbf{g}_{II} + \gamma \mathbf{g}_0]$ , where  $\mathbf{g}_I := (R, S, T, P)$ ,  $\mathbf{g}_{II} := (R, T, S, P)$ ,  $\mathbf{g}_0 := (1, 1, 1, 1)$ , and  $\mathbf{g}_0 = (1, 1, 0, 0)$ . Thus,  $w(n) := p_C(n+1) - p_C(n)$  is given by

$$\alpha s(n) \cdot \mathbf{g}_I + \beta s(n) \cdot \mathbf{g}_{II} + \gamma s(n) \cdot 1,$$

which is just  $\alpha P_I(n) + \beta P_{II}(n) + \gamma$ . Summing  $w(n)$  over  $n=0, 1, \dots, N-1$  and dividing by  $N$ , we obtain

$$\frac{p_C(N) - p_C(0)}{N} \rightarrow \alpha P_I + \beta P_{II} + \gamma;$$

hence, Eq. 2 holds, independent of the strategy of player  $II$ . The same proof works for any  $2 \times 2$  game (even if it is asymmetric).

1. Rapoport A, Chammah A (1965) *The Prisoner's Dilemma* (Univ of Michigan Press, Ann Arbor, MI).
2. Trivers R (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46(1):35–57.
3. Aumann R (1981) Survey of repeated games. *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern* (Wissenschaftsverlag, Mannheim, Germany).
4. Axelrod R (1984) *The Evolution of Cooperation* (Basic Books, New York).
5. Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 50:533–554.
6. Fudenberg D, Maskin E (1990) Evolution and cooperation in noisy repeated games. *Am Econ Rev* 80:274–279.
7. Nowak MA, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364(6432):56–58.
8. Kendall G, Yao X, Chong SY, eds (2007) *The Iterated Prisoner's Dilemma: 20 Years On* (World Scientific, Singapore).
9. Trivers R (2006) Reciprocal altruism: 30 years later. *Cooperation in Primates and Humans: Mechanisms and Evolution*, eds Kappeller PM, van Schaik CP (Springer, Berlin).
10. 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.
11. Stewart AJ, Plotkin JB (2012) Extortion and cooperation in the Prisoner's Dilemma. *Proc Natl Acad Sci USA* 109(26):10134–10135.
12. Adam C, Hintze A (2012) Winning isn't everything: Evolutionary stability of Zero Determinant strategies. Available at <http://arxiv.org/abs/1208.2666>. Accessed October 12, 2012.
13. Boerlijst MC, Nowak MA, Sigmund K (1997) Equal pay for all prisoners. *Am Math Mon* 104:303–307.
14. Bergstrom CT, Lachmann M (2003) The Red King effect: When the slowest runner wins the coevolutionary race. *Proc Natl Acad Sci USA* 100(2):593–598.
15. Sigmund K (2010) *The Calculus of Selfishness* (Princeton Univ Press, Princeton).
16. Nowak MA, Sigmund K (1990) The evolution of stochastic strategies in the Prisoner's Dilemma. *Acta Appl Math* 20:247–265.
17. Taylor PD, Jonker L (1978) Evolutionarily stable strategies and game dynamics. *Math Biosci* 40:145–156.
18. Hofbauer J (2011) Deterministic evolutionary game dynamics. *Evolutionary Game Dynamics*, ed Sigmund K (American Mathematical Society, Providence, RI), pp 61–79.
19. Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428(6983):646–650.
20. Nowak MA (2006) *Evolutionary Dynamics* (Harvard Univ Press, Cambridge, MA).
21. 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, 10.1103/PhysRevE.74.011909.
22. Imhof LA, Nowak MA (2010) Stochastic evolutionary dynamics of direct reciprocity. *Proc Biol Sci* 277(1680):463–468.
23. Fudenberg D, Imhof LA (2006) Imitation processes with small mutations. *J Econ Theory* 131:251–262.
24. Rand DG, Armao JJ, 4th, Nakamaru M, Ohtsuki H (2010) Anti-social punishment can prevent the co-evolution of punishment and cooperation. *J Theor Biol* 265(4):624–632.
25. Hilbe C, Traulsen A (2012) Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Sci Rep* 2:458.
26. Damore JA, Gore J (2011) A slowly evolving host moves first in symbiotic interactions. *Evolution* 65(8):2391–2398.
27. Nowak MA, Sigmund K (1994) The alternating Prisoner's Dilemma. *J Theor Biol* 168:219–226.
28. Frean MR (1994) The prisoner's dilemma without synchrony. *Proc Biol Sci* 257(1348):75–79.
29. Doebeli M, Knowlton N (1998) The evolution of interspecific mutualisms. *Proc Natl Acad Sci USA* 95(15):8676–8680.
30. Frean MR, Abraham ER (2004) Adaptation and enslavement in endosymbiont-host associations. *Phys Rev E Stat Nonlin Soft Matter Phys* 69(5 Pt 1):051913.

one just has to replace  $\mathbf{g}_{II}$  with the corresponding payoff vector). In many cases, however, there will be no solutions to Eq. 1 that are feasible (i.e., probabilities between 0 and 1).

### Appendix B: Sets $\mathcal{ZD}$ , $\mathcal{EQ}$ , and $\mathcal{EX}$

Elementary algebra shows that within the 4D unit cube of all memory-one strategies  $(p_R, p_S, p_T, p_P)$ , the ZD strategies are characterized by

$$\begin{aligned} (1-p_R)(S+T-2P)+(1-p_S)(P-R)+p_T(R-P) \\ +p_P(S+T-2R)=0, \end{aligned}$$

(a 3D subset of the cube). Equalizers are characterized, in addition, by

$$(R-P)(p_S-p_T-1)=(T-S)(p_R-p_P-1),$$

(they form a 2D set), and  $\chi$ -extortion strategies are also characterized by  $p_P=0$  and

$$p_T[P-S+(T-P)\chi]=(1-p_S)[T-P+(P-S)\chi]$$

(for each  $\chi$ , a 1D set). In the special case of the donation game, these equations reduce to

$$p_R+p_P=p_S+p_T,$$

$$(b-c)(p_S-p_T-1)=(b+c)(p_R-p_P-1),$$

$$p_T(c+\chi b)=(1-p_S)(b+\chi c),$$

respectively. The set  $\mathcal{EQ}$  of equalizers is spanned by  $(1, 1, 0, 0)$ ,  $(c/b, 0, c/b, 0)$ ,  $(\frac{2c}{b+c}, 0, 1, \frac{b-c}{b+c})$ , and  $(1, 1-c/b, 1, 1-c/b)$ , and the set  $\mathcal{EX}$  of extortion strategies is spanned by  $(1, 1, 0, 0)$ ,  $(c/b, 0, c/b, 0)$ , and  $(1, 0, 1, 0)$ . All reactive strategies are ZD strategies, the reactive equalizers are those satisfying  $p-q=c/b$ , and the reactive  $\chi$ -extortioners are those with  $q=0$  and  $p=(b+\chi c)/(c+\chi b)$  (Fig. 2).

**ACKNOWLEDGMENTS.** We thank M. Abou Chakra, A. Traulsen, J.A. Damore and R. Trivers for useful discussions. K.S. acknowledges support from the Foundational Questions in Evolutionary Biology Fund (Grant RFP-12-21).