Interim Report   **IR-13-038**

Evolution of extortion in iterated prisoner's dilemma games

Christian Hilbe
Martin A. Nowak
Karl Sigmund (ksigmund@iiasa.ac.at)

Approved by

Ulf Dieckmann
Director, Evolution and Ecology Program

June 2015

The Evolution of Extortion
in Iterated Prisoner's Dilemma Games

Christian Hilbe[1], Martin A. Nowak [2], & Karl Sigmund [3,4]

[1] Max Planck Institut for Evolutionary Biology, Ploen
[2] Program for Evolutionary Dynamics, Harvard University,
One Brattle Square, Cambridge, MA 02138, USA
[3] Faculty of Mathematics, University of Vienna, A-1090 Vienna, Austria
[4] International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

February 18, 2013

Corresponding author:

Karl Sigmund
Faculty of Mathematics, University of Vienna, A-1090 Vienna, Austria
e-mail: karl.sigmund@univie.ac.at, phone: +43 (0)1 4277 506 12, fax: +43 (0)1 4277 9 506

Iterated games are a fundamental component of economic and evolutionary game theory. They describe situations where two players interact repeatedly and have the possibility 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 the 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 can also do well in co-evolutionary arms races between two distinct populations: significantly, they benefit the population which evolves at the slower rate - an instance of the so-called Red King effect. This may affect the evolution of interactions between host species and their endosymbionts.

Introduction

The Iterated Prisoner's Dilemma (IPD) has a long history as a model for the cultural and biological evolution of cooperation (1–13). A new class of so called *zero-determinant (ZD)* strategies has recently attracted considerable attention (14–16). Such strategies allow players to unilaterally enforce a linear relation between the own and the co-player's payoff. A subset consists of the so-called *equalizer* strategies: these assign to the co-player's score a predetermined value, independent of the co-player's strategy, see also (17). Another subset consists of the *extortion* strategies: they guarantee that the own surplus exceeds the co-player's surplus by a fixed percentage. Press and Dyson (14) have explored the power of ZD-strategies to manipulate any 'evolutionary' opponent, i.e., any co-player able to learn, and to adapt.

In their commentary to Press and Dyson, Stewart and Plotkin (15) 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 suc-

cessful, 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 and to cooperate. Extortion is therefore evolutionarily unstable (16). However, we shall see that if the two players engaged in an IPD belong to distinct populations, the evolutionary prospects of extortion improve significantly. Moreover, extortioners can do well even in a single population, if its size is very small.

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 catalyzers for cooperation, not as a long-term outcome. However, if the IPD 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 (18).

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 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 own cost c, with $0 < c < b$. Then $T = b$, $R = b - c$, $P = 0$ and $S = -c$.

In the Iterated Prisoner's Dilemma game (IPD), the two players are required to play an infinite number of rounds, and their payoffs P_I resp. P_{II} 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_R, p_S, p_T and p_P 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 (19)). An important class of memory-one strategies consists of *reactive* strategies, which only depend on the co-player's move in the previous round (not the own). Then $p_R = p_T =: p$ and $p_P = p_S =: q$, so that a reactive strategy corresponds to a point (p, q) in the unit square (20).

We will first define and characterize zero-determinant 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

3

can catalyze the emergence of cooperation. The same result will then be obtained if we consider *all* memory-one strategies: in particular, extortion strategies can only get a foothold if the population is very small. If the IPD is played between members of two distinct populations, ZD-strategies can emerge in the population which evolves more slowly. In particular, extortion strategies can allow host species to enslave their endosymbionts.

Methods and Results

Definitions. Press and Dyson (14) define the class of 'zero-determinant' strategies ZD as those memory-one strategies (p_R, p_T, p_S, p_P) satisfying, for some reals α, β, γ, the equations

$$
\begin{aligned}
p_R - 1 &= \alpha R + \beta R + \gamma & [1a] \\
p_S - 1 &= \alpha S + \beta T + \gamma & [1b] \\
p_T &= \alpha T + \beta S + \gamma & [1c] \\
p_P &= \alpha P + \beta P + \gamma. & [1d]
\end{aligned}
$$

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

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

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

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

Thus player I can assign to the co-player any payoff between P and R. (Indeed, since the p_i have to be between 0 and 1, it follows that $\beta < 0$ and $P \leq P_{II} \leq R$). The so-called χ-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 the own 'surplus' (over the maximin value P) is the χ-fold of the co-player's surplus. Fig. 1 shows examples of these different ZD-strategies.

Press and Dyson speak of zero-determinant strategies because they use for their proof of [2] an ingenious method based on determinants. In Appendix A, we present a more elementary proof, following (17). Within the four-dimensional unit cube of all memory-one strategies (p_R, p_S, p_T, p_P), the ZD strategies form a three-dimensional subset $\mathcal{ZD}$ containing the two-dimensional subsets $\mathcal{EQ}$ and $\mathcal{EX}$ of equalizers resp. extortioners (see Appendix B). In Fig. 2 we sketch these sets for the reactive strategies.

Extortion within one population. In order to investigate the role of extortion in the context of evolutionary games, we concentrate on the donation game (in the SI 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 χ-extortion strategy E_χ fares against some of the most important memory-one strategies, namely $TFT = (1, 0, 1, 0)$, $AllD = (0, 0, 0, 0)$, $AllC = (1, 1, 1, 1)$ and the Win-Stay-Lose-Shift strategy $WSLS$ which is encoded by $(1, 0, 0, 1)$, and hence cooperates if and only if the co-player's move, in the previous round, was the same as the own move, see (11). 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 with strategy j is given by the (i, j)-th element of the following matrix:

	TFT	$WSLS$	E_χ	$All\ C$	$All\ D$	
TFT	$(b-c)/2$	$(b-c)/2$	0	$b - c$	0	
$WSLS$	$(b-c)/2$	$b - c$	$\frac{b^2-c^2}{b(1+2\chi)+c(2+\chi)}$	$(2b-c)/2$	$-c/2$	[4]
E_χ	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	
$All\ C$	$b - c$	$(b-2c)/2$	$\frac{b^2-c^2}{b\chi+c}$	$b - c$	$-c$	
$All\ D$	0	$b/2$	0	b	0	

Let us start with the pairwise comparisons. E_χ is neutral with respect to $AllD$. 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.) $AllC$ 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 that 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 $AllC$. No Nash equilibrium involves E_χ. If $b < 2c$, there are two Nash equilibria: a mixture of TFT, $AllC$ and $AllD$ and a mixture of $WSLS$, $AllC$ and $AllD$. If $b > 2c$, there exist four Nash equilibria. In particular, $WSLS$ is then a strict Nash equilibrium.

The replicator dynamics (19, 21, 22) displays for the payoff matrix continuous families of fixed points and periodic orbits, and hence is far from being robust. The same applies to most other deterministic game dynamics. It seems more reliable to consider a stochastic process which 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 (22, 23, 26). Selection is here viewed 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 the SI, the details of the imitation process matter little.) Additionally, mutations occur with a small probability $\mu > 0$ (corresponding to the adoption of another strategy, each alternative being equally likely).

Any such stochastic process yields a steady state distribution of strategies. We find that while extortioners are never the most abundant strategy, they can play the role of a catalyzer. Indeed, if only $AllD$ and $WSLS$ are available, a population may be trapped in a non-cooperative state for a considerable time, leading to a mutation-selection equilibrium that clearly favors defectors (see Fig. 3A). In such a case, extortioners (Fig. 3B) and TFT (Fig. 3C. and (10)) offer an escape: both strategies can subvert an $AllD$ population through neutral drift. Once defectors are rare, $WSLS$ outperforms TFT, and it also prevails against extortioners if the population is sufficiently large (in a direct competition, $WSLS$ always gets a higher payoff than E_χ if $M > 1 + \chi$). Thus, in large populations, extortioners and TFT-players tip the mutation-selection balance towards $WSLS$, and therefore increase the level of cooperation. Further expansion of the strategy space through adding $AllC$ has only a small effect on the steady state (Figs. 3D and 3E), 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 (25), assuming that mutants can pick up any memory-one strategy, with a uniform probability distribution on the four dimensional 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 δ-neighborhood of A (with respect to Euclidean distance) by A_δ, and let $\mu(A_\delta)$ denote the fraction of time that the evolving population visits A_δ. We say that A_δ is favored by evolution if the evolutionary process visits A_δ more often than expected under neutral evolution, i.e., if $\mu(A_\delta)$ is larger than the volume of the intersection of A_δ 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 zero-determinant strategies, are favored by evolution if the population is reasonably large (see Fig. 4A). By contrast, very small population sizes promote the evolution 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 with players using the own strategy (23); this effect may even result in the evolution of spite (28, 29). We address this point in more detail in the SI (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 in (12). The same holds for most ZD-strategies. By contrast, $WSLS$ players do well against their like, and therefore stably dominate the evolutionary dynamics for larger population sizes, at least when $b > 2c$ or, for more general PD games, when $2R > T + P$ (see Fig. 4B and (19)). As a (possibly surprising) consequence, larger populations also yield higher average payoffs (Fig. 4C). In the SI, 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 (for instance, hosts and their symbionts), or two classes of a single species, old and young, for example, buyers and sellers, or rulers and subjects, engaged in an IPD game which, of course, is now unlikely to be symmetric. In such situations, extortioners may evolve even in large populations. Indeed, extortioners provide incentives to cooperate: as shown by Press and Dyson (14), $AllC$ is always a best response to an extortion strategy. In a single population of homogeneous players, this is not turned to advantage, as 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 co-players of their own kind.

7

Nevertheless, their success may be short-lived, since they will be tempted to adopt the even more profitable $AllD$-strategy as a reaction to the $AllC$ co-players which they have produced, which in turns leads to the disappearance of the $AllC$-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 co-evolutionary model of Damore and Gore (32). 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 a relative evolutionary rate of one, 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 relative evolutionary rate, by allowing symbionts to adapt more quickly. Fast adaptation results in a short-term increase of the symbionts' payoffs, since 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 in the long run, for the host, to be slow to evolve; for the parameters in Fig. 5B, the resulting equilibrium allocates them on average a surplus more than ten times larger than the surplus achieved by the symbionts.

Discussion

Our main results show that within one population, extortion strategies can act as catalyzers 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 the SI, we emphasize this robustness. We could also assume that the players alternate their moves in the donation game (30, 31); or that the underlying PD game is asymmetric (the definitions have to be modified in an straightforward way). As noted in (14), some results hold also for non-PD games; this deserves further investigation.

In orthodox game theory, strategy A dominates B if no matter what the co-

player does, A yields at least the payoff of B. When Press and Dyson (14) argue that extortioners dominate their co-players, they mean that no matter what the co-player does, the extortioner gets more. This is not quite the same, and we display in section 2 of the SI an example which highlights the difference. Adami and Hintze (16) stress a similar point in their title: 'Winning isn't everything'. Moreover, when Press and Dyson 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 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, see (3, 9). In the context of evolutionary game theory, whole populations are engaged in the game. For very small population size extortion strategies still offer good prospects. This is not surprising, as the limiting case, a population size $M = 2$, reduces to the scenario analyzed by Press and Dyson (14). 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 game 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 dis-proportionally often. This does not preclude, of course, that certain elements of their class are favored by selection. Thus Generous TFT $(1, 1 - c/b, 1, 1 - c/b)$ does well. So do other, less known strategies. In particular, Stewart and Plotkin highlighted a class of strategies defined, instead of Eq. [3], by $P_I - R = \chi(P_{II} - R)$ (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 Adami and Hintze (16), the evolutionary stability of several ZD strategies was tested by replicator dynamics and agent-based simulations, which independently confirms the result that these strategies do not prevail in large populations. They used a population size $M = 1024$, and payoff values $R = 3, S = 0, T = 5$ and $P = 1$, i.e., a Prisoner's Dilemma game which cannot be reduced to a donation game. Adami and Hintze 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, since 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, which has to evolve and risks being faked. When we consider interactions between different populations, a cheater-proof tag is provided for free.

Therefore, extortion may evolve in endosymbiotic relationships due to the so-called Red-King effect (18, 33, 34): the species that evolves at a slower rate gains a disproportionate share of the benefits. 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 host's and the symbionts' payoffs are perfectly correlated. This ensures that only those symbiont mutants can succeed that are beneficial for the host. 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 $s_S(n + 1)$ by

$$s_R(n)(1-q_R(n))p_R + s_S(n)(1-q_S(n))p_S + s_T(n)(1-q_T(n))p_T + s_P(n)(1-q_P(n))p_P.$$

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 $\mathbf{s}(n) \cdot \mathbf{p} = \mathbf{s}(n) \cdot [\alpha \mathbf{g}_I + \beta \mathbf{g}_{II} + \gamma \mathbf{1} + \mathbf{g}_0$, where $\mathbf{g}_I := (R, S, T, P)$, $\mathbf{g}_{II} := (R, T, S, P)$, $\mathbf{1} = (1, 1, 1, 1)$ and $\mathbf{g}_0 = (1, 1, 0, 0)$. Thus $w(n) := p_C(n + 1) - p_C(n)$ is given by

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

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

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

and hence Eq. [2] holds, independently of the strategy of player II. The same proof works for any 2×2 game (even if it is asymmetric: one just has to replace

10

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

Appendix B: the sets $\mathcal{ZD}$, $\mathcal{EQ}$ and $\mathcal{EX}$: Elementary algebra shows that within the four-dimensional unit cube of all memory-one strategies (p_R, p_S, p_T, p_P), the ZD-strategies are characterized by

$$(1 - p_R)(S + T - 2P) + (1 - p_S)(P - R) + p_T(R - P) + p_P(S + T - 2R) = 0,$$

(a three dimensional 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 two-dimensional set) and χ-extortion strategies by $p_P = 0$ and

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

(for each χ a one-dimensional 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 $(1, 1 - c/b, 1, 1 - c/b)$, the set $\mathcal{EX}$ of extortion strategies 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 χ-extortioners those with $q = 0$ and $p = (b + \chi c)/(c + \chi b)$ (see Fig. 2).

References

[1] Rapoport A, Chammah A (1965) *The Prisoner's Dilemma*. (Ann Arbor, Univ Michigan Press).

[2] Trivers R (1971) The evolution of reciprocal altruism. *Quarterly Rev Biol* 46: 35-57.

[3] Aumann R (1981) Survey of repeated games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, (Mannheim, Wissenschaftsverlag).

[4] Axelrod R (1984) *The Evolution of Cooperation*. (Basic Books, New York).

[5] Molander P (1985) The optimal level of generosity in a selfish, uncertain environment.*Journ. Conflict Resolution* 29: 611-618.

[6] Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 50: 533-554.

[7] Milinski M (1987) Tit for tat in sticklebacks and the evolution of cooperation. *Nature* 325: 433-435

[8] May R M (1987) More evolution of cooperation. *Nature* 327: 15-17

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

[10] Nowak MA, Sigmund K (1992) Tit for tat in heterogeneous populations. *Nature* 355: 250-253

[11] 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: 56 - 58.

[12] Kendall G, Yao X, Chong, SY (eds) (2007) *The Iterated Prisoner's Dilemma: 20 Years On*. (Singapore: World Scientific).

[13] Trivers R (2006) Reciprocal altruism: 30 years later. In *Cooperation in Primates and Humans: Mechanisms and Evolution* (Kappeller PM, van Schaik CP eds.) (Berlin, Springer).

[14] Press WH, Dyson FD (2012) Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proc Nat Acad Sci USA* 109:

[15] Stewart AJ, Plotkin JB (2012) Extortion and Cooperation in the Prisoner's Dilemma. *Proc Nat Acad Sci USA* 109: 10134-10135.

[16] Adami C, Hintze A (2012) Winning isn't everything: Evolutionary stability of Zero Determinant strategies. *preprint* http://arxiv.org/abs/1208.2666.

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

[18] Bergstrom CT, Lachman M (2003) The Red King effect: when the slowest runner wins the race *Proc Nat Acad Sci USA* 100: 593-598.

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

[20] Nowak MA, Sigmund K (1990) The evolution of stochastic strategies in the Prisoner's Dilemma. *Acta Applicandae Math* 20: 247-265.

[21] Helbing D (1992) A mathematical model for behavioural changes by pair interactions. Haag G, Mueller U, Troitzsch KG (eds), *Economic Evolution and Demographic Change. Formal Models in Social Sciences* (Berlin, Springer) 330-348.

[22] Nowak MA (2006) *Evolutionary Dynamics* (Cambridge MA, Harvard Univ Press).

[23] Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428: 646-650.

[24] Imhof L, Fudenberg D, Nowak MA (2005) Tit-for-tat or win-stay, lose-shift? *Journal of Theoretical Biology* 247: 574-580.

[25] Imhof LA, Nowak MA (2010) Stochastic evolutionary dynamics of direct reciprocity. *Proc R Soc B* 277: 463-468.

[26] Traulsen A, Nowak MA, Pacheco J (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E* 74 (1): 011909. http://dx.doi.org/10.1103/PhysRevE.74.011909

[27] Fudenberg D, Imhof LA (2006) Imitation processes with small mutations. Journal of Economic Theory 131: 251-262

[28] Rand DG, Armao JJ, Nakamaru M, Ohtsuki H (2010) Anti-social punishment can prevent the co-evolution of punishment and cooperation. *J Theor Biol* 265: 624-632.

[29] Hilbe C, Traulsen A (2012) Emergence of responsible sanctions without second-order free-riders, anti-social punishment or spite. *Scientific Reports* 2: 458.

[30] Nowak MA, Sigmund K (1994) The alternating Prisoner's Dilemma. *J Theor Biol* 168: 219-226.

[31] Frean MR (1994) The Prisoner's Dilemma without synchrony. *Proc Roy Soc London B* 257: 75-79.

[32] Damore JA, Gore J (2011) A slowly evolving host moves first in symbiotic interactions. *Evolution* 65:2391-2398.

[33] Doebeli M, Knowlton N (1998) The evolution of interspecific mutualisms. *Proc Nat Acad Sci USA* 95: 8676-8680.

[34] Frean MR, Abraham ER (2004) Adaption and enslavement in endosymbiont-host associations. *Physical Review E* 69: 051913.

Figure 1: The payoffs P_I and P_{II} of players I and II if both players use memory-one strategies in an iterated Prisoner's Dilemma (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 co-player II can vary, sampling the four-dimensional 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 two-dimensional area, as here, when player I applies the strategy of win-stay lose-shift, i.e., $p_R = p_P = 1$ and $p_S = p_T = 0$. **(B)** However, if player I adopts a zero-determinant strategy, then the possible payoff-pairs are restricted to a line. Two special classes of zero-determinant strategies were highlighted by Press and Dyson (14): **(C)** Equalizers, i.e., strategies that set the co-player's score to a fixed value (the line of payoffs has slope zero), and **(D)** Extortioners, strategies which guarantee that the surplus of player I is the χ-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).
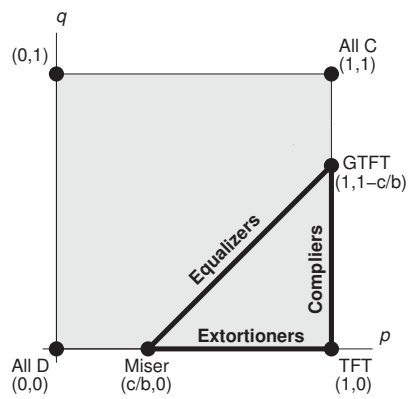
15

Figure 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' ($p = 1$, $q = 1 - c/b$, (20)) and 'Miser' ($p = c/b$, $q = 0$, (31)), the extortion strategies those between 'Miser' and TFT ($p = 1$, $q = 0$), and the 'compliant' strategies (see (15) and Discussion) those between 'generous TFT' and TFT.
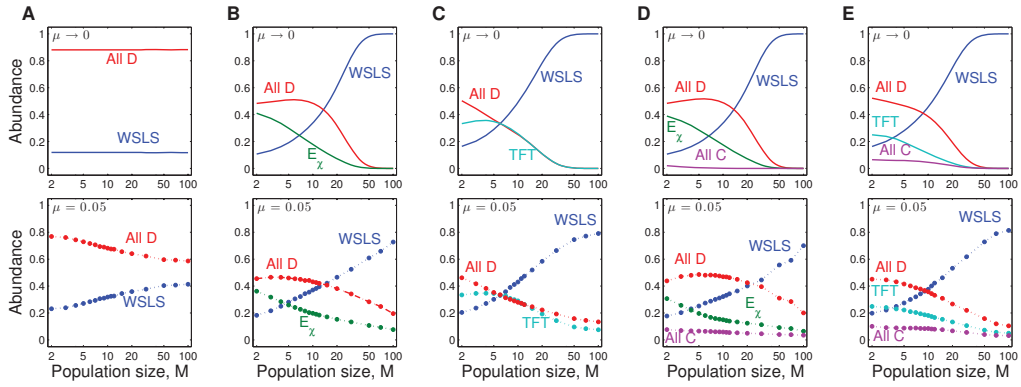
Figure 3: Evolutionary competition between some important strategies in the IPD. For various population sizes M, the graphs show the frequency of each strategy in the mutation-selection equilibrium. We consider two mutation regimes, the limit of rare mutations $\mu \to 0$ (top row), for which the equilibrium can be calculated analytically, using (26) and a regime with mutation rate $\mu = 0.05$ (bottom row) 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. (26). If $AllD$ competes with $WSLS$ the population is mostly in the defector's state, independently of population size and the mutation rate (A). However, once E_χ or TFT is added, $WSLS$ succeeds if populations are sufficiently large (B and C). TFT works slightly better than E_χ. Adding $AllC$ only leads to minor changes in the stationary distribution, which now slightly favour E_χ (D and E). Parameters: $b = 3$, $c = 1$, $s = 1$, and $\chi = 2$.
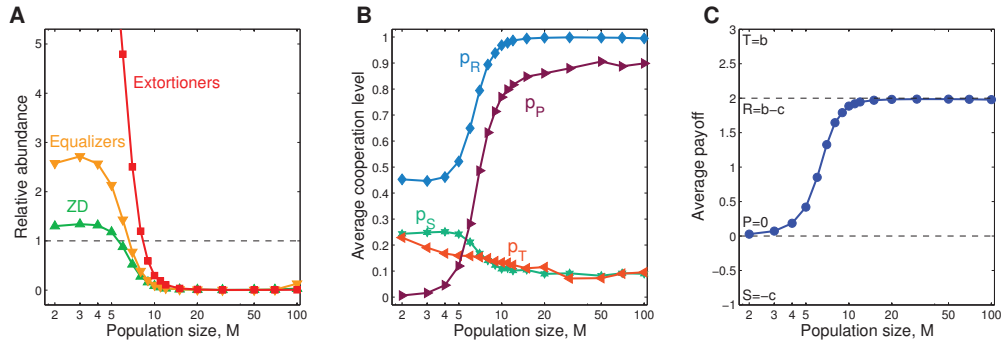
17

Figure 4: Statistics of the evolutionary dynamics for memory-one strategies for a range of different population sizes. We have calculated (A) the relative abundance of extortioners, equalizers, and ZD strategies, i.e. the time spent in a δ-neighborhood, divided by the volume of the intersection of that neighborhood with the set of memory-one strategies; (B) the average strategy of the population; (C) the average payoff. 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 (25), the switch from a monomorphic population using strategy X to a monomorphic population using strategy Y occured with the probability of fixation of a single Y mutant in a population of X-residents. Parameters: $b = 3$, $c = 1$, $\delta = 0.1$ and $s = 100$.
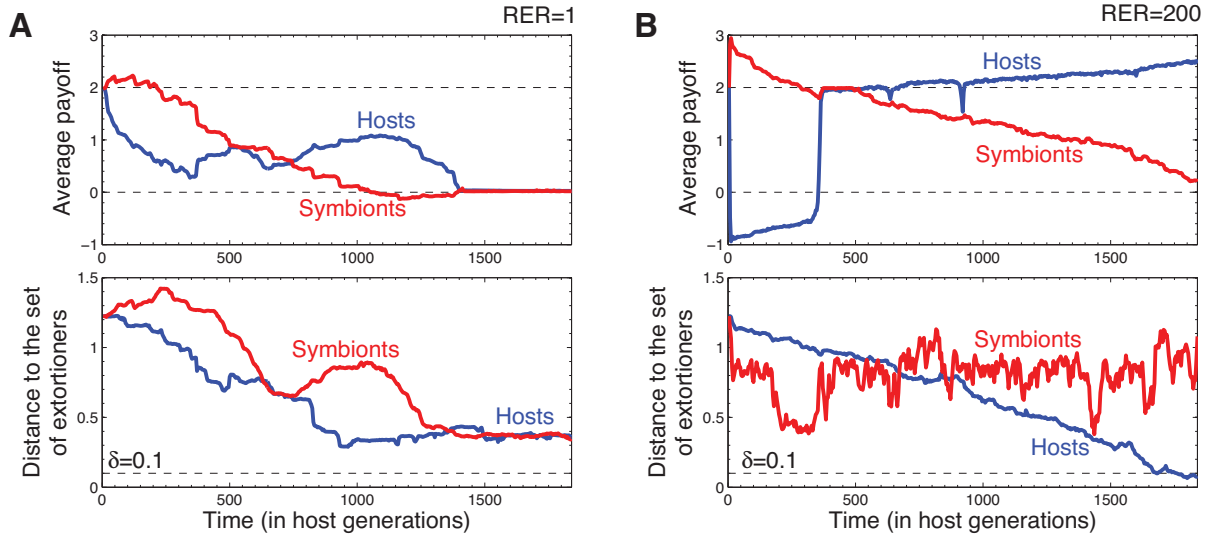
18

Figure 5: Evolution of extortion in host-symbiont interactions. The graphs show two typical simulation runs for a population of 40 hosts, each having a subpopulation of 20 symbionts. For each simulation run, the upper graph shows the average payoff for each population, whereas the lower graph shows the Euclidean distance of each population to the set of extortioners (which can be at most 1.5275). In the initial population all individuals cooperate unconditionally. The further evolution depends on the relative evolutionary rate (RER): (A) If $RER = 1$, both species converge towards $AllD$, 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 non-cooperative 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 (32): 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 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 3×10^8 reproduction events for $RER = 200$). The other parameters were $b = 3$, $c = 1$ and $s = 10$.