SCIENTIFIC REP⚙RTS

A double-edged sword: Benefits and pitfalls of heterogeneous punishment in evolutionary inspection games

Matjaž Perc[1,2,3] & Attila Szolnoki[4]

As a simple model for criminal behavior, the traditional two-strategy inspection game yields counterintuitive results that fail to describe empirical data. The latter shows that crime is often recurrent, and that crime rates do not respond linearly to mitigation attempts. A more apt model entails ordinary people who neither commit nor sanction crime as the third strategy besides the criminals and punishers. Since ordinary people free-ride on the sanctioning efforts of punishers, they may introduce cyclic dominance that enables the coexistence of all three competing strategies. In this setup ordinary individuals become the biggest impediment to crime abatement. We therefore also consider heterogeneous punisher strategies, which seek to reduce their investment into fighting crime in order to attain a more competitive payoff. We show that this diversity of punishment leads to an explosion of complexity in the system, where the benefits and pitfalls of criminal behavior are revealed in the most unexpected ways. Due to the raise and fall of different alliances no less than six consecutive phase transitions occur in dependence on solely the temptation to succumb to criminal behavior, leading the population from ordinary people-dominated across punisher-dominated to crime-dominated phases, yet always failing to abolish crime completely.

In 1982 Wilson and Kelling[1] introduced the "broken windows theory", explaining how seemingly unimportant and harmless signals of urban disorder may over time elicit antisocial behavior and serious crime. The central premise of the theory is simple yet powerful, and it is reminiscent of preferential attachment or the Matthew effect[2,3] with a negative connotation. Just like the more connected nodes attract more new links during network growth[4,5], so does an unattended broken window invite bypassers to behave mischievously or even disorderly. Similarly, a graffiti might point to an unkept environment, signaling that more egregious damage will likely be tolerated as well. One broken window is thus likely to become many broken windows, and the inception of urban decay and criminal behavior is in place.

The simplicity of this widely adopted criminological theory invites mathematicians and physicists to adopt a complex systems approach[6] to study criminal behavior[7], in particular since the collective behavior of the system in this case can hardly be inferred from the relatively simple individual actions. Emergent phenomena such as pattern formation including percolation[8,9] and phase transitions are commonly associated with complex social and biological systems[10–13], and in this realm the mitigation of crime is certainly no exception. Recent research highlights that crime is far from being uniformly

[1]Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, SI-2000 Maribor, Slovenia. [2]Department of Physics, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia. [3]CAMTP – Center for Applied Mathematics and Theoretical Physics, University of Maribor, Krekova 2, SI-2000 Maribor, Slovenia. [4]Institute of Technical Physics and Materials Science, Research Centre for Natural Sciences, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary. Correspondence and requests for materials should be addressed to M.P. (email: matjaz.perc@uni-mb.si)

distributed across space and time[14,15], and this is confirmed also by the dynamic nucleation and dissipation of crime hotspots[16–19] and the emergence of complex criminal networks[20–23].

The emergence of crime can also be treated as a social dilemma[24–26], in as far that social order is the common good that is threatened by criminal activity, with competition arising between criminals and those trying to prevent crime. An adversarial evolutionary game with four competing strategies has recently been proposed[27], where paladins are model citizens that do not commit crimes and collaborate with authorities, while villains, at the other extreme of the spectrum, commit crimes and do not report them. Intermediate figures are informants who report on other offenders while still committing crimes, and apathetics who neither commit crimes nor report to authorities. Apathetics are similar to second-order free-riders in the context of the public goods game with punishment[28–31], in that they cooperate at first order by not committing crimes, but defect at second order by not punishing offenders. Simulations have revealed that in the realm of the adversarial game informants are key to the emergence of a crime-free society, and this has subsequently been confirmed also with human experiments[32].

In general, the mitigation of crime can be framed as an evolutionary game with punishment, although recent research has raised doubts on the use of sanctions as a means to promote prosocial behavior[33–37]. Rewards for not doing and reporting crime are a viable alternative, and in this case the "stick versus carrot" dilemma becomes an important consideration[38–41]. In the context of rehabilitating criminals, the question is also how much punishment for the crime and how much reward for eschewing wrongdoing in the future is in order for optimal results, as well as whether these efforts should be the responsibility of individuals or institutions[42–44] under the assumption of a limited budget[45].

It is at this intersection of statistical physics of complex system and evolutionary games that we aim to contribute in the present paper by considering a three-strategy spatial inspection game with uniform punishment as well as a five-strategy spatial inspection game with heterogeneous punishment. The inspection game is a recognized model in the sociological literature for the dynamics of crime[46,47]. The game addresses the question of why anybody would be willing to invest into costly punishment of criminals, given that individuals are tempted to benefit from the punishing activities of others without actively contributing to them. As soon as ordinary people are introduced who neither commit crimes nor contribute to their mitigation, one is thus faced with the second-order free-rider problem[30,48]. As we will show in what follows, this may introduce cyclic dominance that enables the coexistence of all three competing strategies in the uniform punishment model. More importantly, the consideration of heterogeneous punisher strategies drastically elevates the complexity of possible solutions, revealing on the one hand a more effective solution to the second-order free-rider problem, yet still failing to abolish crime completely. As a consequence, the diversity of punishment allows the formation of different alliances between competing strategies, which gives rise to a sophisticated range of solutions in dependence on the payoffs.

In the next Section we first present the details of the considered 3-strategy and 5-strategy spatial inspection game, and then demonstrate how systematic Monte Carlo simulations reveal the benefits and pitfalls of punishing criminal behavior. Simulation details are described in the Methods Section. We conclude by discussing the presented results and their wider implications.

Results

3-strategy and 5-strategy spatial inspection game. We first introduce a three-strategy version of the spatial inspection game, where in addition to criminals C and punishers P, also ordinary people O compete for space on a $L \times L$ square lattice with periodic boundary conditions. We use the latter as the simplest network to account for the fact that the interaction range among individuals in human societies

	O	C	P
O	0	$-\beta$	0
C	β	0	$\beta - 1$
P	$-\alpha$	$\gamma - \alpha$	$-\alpha$

is limited. The payoff matrix contains α as the punishment cost, β as the temptation to succumb to criminal behavior as well as the loss when being a victim of crime, and γ as the reward for punishing criminals. Moreover, when a criminal faces a punisher, it will receive $\beta - 1$, where -1 corresponds to the normalized punishment fine. These payoffs apply for each pairwise interaction between the players.

To enable a more sophisticated response to the second-order free-rider problem, we also consider an extended model with heterogeneous punishment. Similarly to other diversity-motivated social problems[49–51], we expect that such a model will provide further insights and a more adequate answer to the free-rider problem. In the proposed five-strategy version of the spatial inspection game punishers are divided into three categories, namely L, M and H, depending on the cost they are willing to bear for punishing criminals. The extended payoff matrix contains the same three main parameters as the
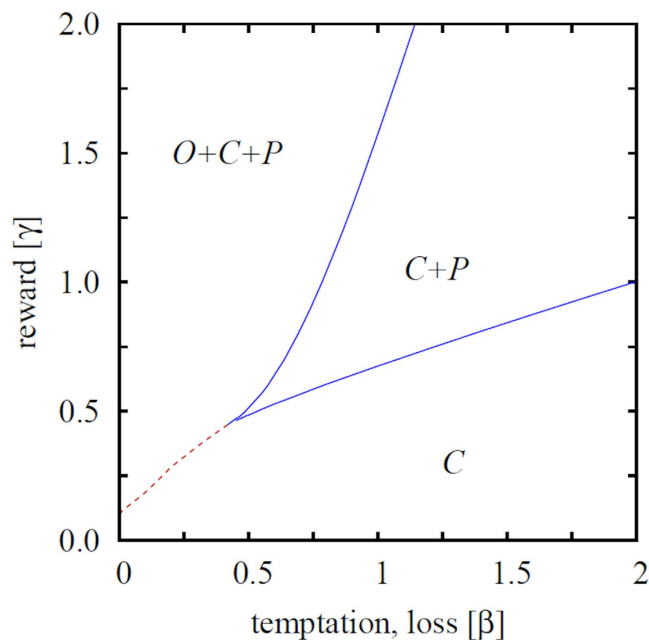
Figure 1. Phase diagram of the three-strategy spatial inspection game with uniform punishment.
Depicted are strategies remaining on the square lattice after sufficiently long relaxation times as a function of the temptation/loss β and the reward for punishing criminals γ, as obtained for the the punishment cost $\alpha = 0.5$. Here C marks the parameter region where the population terminates in a homogeneous "all-criminal" phase, $C+P$ marks the region where criminals and punishers coexist, while in the $O+C+P$ region all three strategies are present in the stationary state due to cyclic dominance. Solid blue lines denote continuous phase transitions, while the dashed red line denotes the border of cyclic dominance between competing strategies.

	O	C	L	M	H
O	0	$-\beta$	0	0	0
C	β	0	$\beta - \frac{1}{3}$	$\beta - \frac{2}{3}$	$\beta - 1$
L	$-\frac{1}{3}\alpha$	$\frac{1}{3}(\gamma - \alpha)$	$-\frac{1}{3}\alpha$	$-\frac{1}{3}\alpha$	$-\frac{1}{3}\alpha$
M	$-\frac{2}{3}\alpha$	$\frac{2}{3}(\gamma - \alpha)$	$-\frac{2}{3}\alpha$	$-\frac{2}{3}\alpha$	$-\frac{2}{3}\alpha$
H	$-\alpha$	$\gamma - \alpha$	$-\alpha$	$-\alpha$	$-\alpha$

three-strategy payoff matrix, with the key difference being that punishers L and M are willing to bear only 1/3 and 2/3 of the full punishment cost α, respectively. Naturally, they also receive a proportionally smaller reward γ. Punishers H correspond to punishers P in the three-strategy model in terms of their commitment to sanctioning criminals, but we introduce a different notation for convenience.

Both the uniform three-strategy and the heterogeneous five-strategy spatial inspection game are studied by means of Monte Carlo simulations, as described in the Methods section.

Evolutionary dynamics. We begin by presenting the complete $\beta - \gamma$ phase diagram at a representative value of the punishment cost α in Fig. 1. It can be observed that criminals dominate if the reward for their punishment γ is small. If the reward exceeds a certain value at a fixed temptation/loss β, then the punishers become viable. At moderate β values, however, their presence is also accompanied by the emergence of ordinary players. The stability of the $O+C+P$ phase is due to cyclic dominance between the three competing strategies[13]. In particular, within the $O+C+P$ region ordinary people outperform the punishers, the punishers defeat the criminals, while the criminals beat ordinary people, thus closing the $O \rightarrow P \rightarrow C \rightarrow O$ loop of dominance. Conversely, for larger values of β, in particular if $\beta > \alpha$, the pure C phase becomes the two-strategy $C+P$ phase via a second-order continuous phase transition as γ increases. Moreover, at sufficiently large values of the reward γ, the three-strategy $O+C+P$ phase and the two-strategy $C+P$ phase are separated by a second-order continuous phase transition.
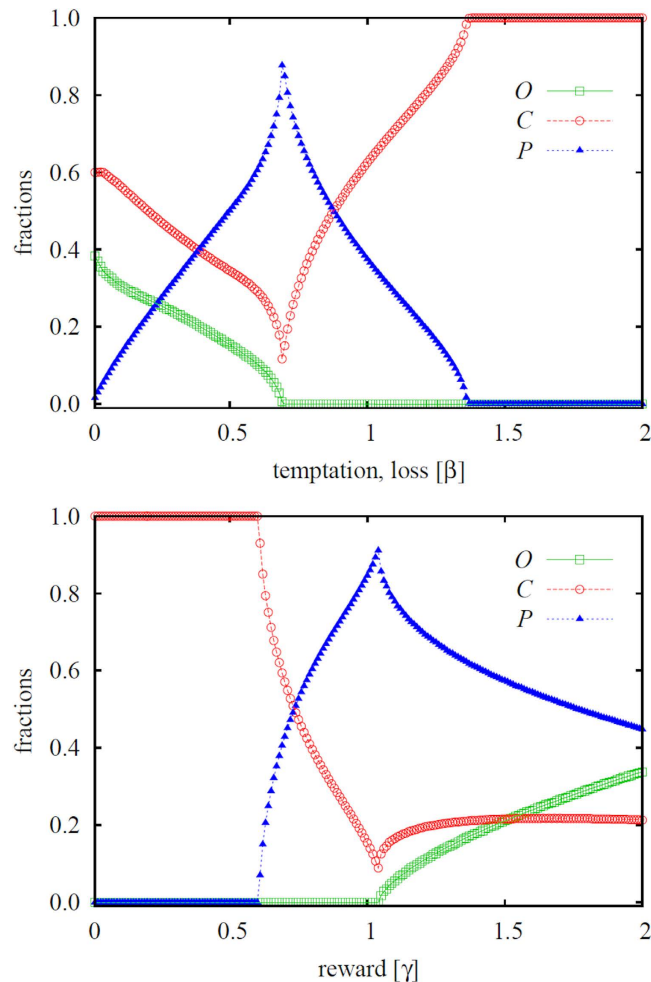
Figure 2. Two characteristic cross-sections of the phase diagram depicted in Fig. 1. Left panel shows the fraction of the three strategies in dependence on the temptation/loss β at $\gamma = 0.8$. Starting at the three-strategy $O + C + P$ phase, the fraction of ordinary people and the criminals decreases steadily with increasing the value of β until eventually O die out and the two-strategy $C + P$ phase is reached. Immediately thereafter the fraction of criminals starts rising as the value of β increases further, with the second continuous phase transition marking the emergence of the pure C phase. Right panel shows the fraction of the three strategies in dependence on the reward for punishing criminals γ at $\beta = 0.8$. In this case we start at the pure C phase, which turns to the two-strategy $C + P$ phase as soon as γ is large enough to sustain the punishers. As γ increases further ordinary people become viable too through a second continuous phase transition, ultimately yielding the three-strategy $O + C + P$ phase that is maintained by cyclic dominance. In both panels the punishment cost is $\alpha = 0.5$.

For a more quantitative view, we present in Fig. 2 characteristic cross-sections of the phase diagram shown in Fig. 1. These cross-sections confirm that criminals can dominate in the high temptation/loss region or in the low reward region. Moreover, it can be observed that larger rewards are beneficial for the punishers, but only up to a certain point. If γ increases beyond a critical point ordinary people emerge, and as second-order free-riders they flourish on the expense of those that punish criminal behavior. We emphasize that, interestingly, the payoffs of ordinary people are independent of γ, yet still their fraction increases as γ increases. This counterintuitive result is due to cyclic dominance, where feeding the prey, in this case the punishers who do get larger payoffs for larger γ values, directly benefits the predator, which in this case are the ordinary people[52,53]. We can thus conclude that the real obstacle in the fight against criminal behavior is the possibility of ordinary people to free-ride on the efforts of punishers. A similar conclusion has been reached before for the evolution of cooperation in the public goods game with punishers, where the free-riding problem of defectors is simply deferred to the second-order free-riding problem of cooperators[28].

As a natural response of punishers to the harmful exploitation of ordinary people, we next consider the five-strategy spatial inspection game with heterogeneous punishment. In particular, strategies L and M try to eschew the exploitation by reducing the amount they contribute for sanctioning to 1/3 and 2/3
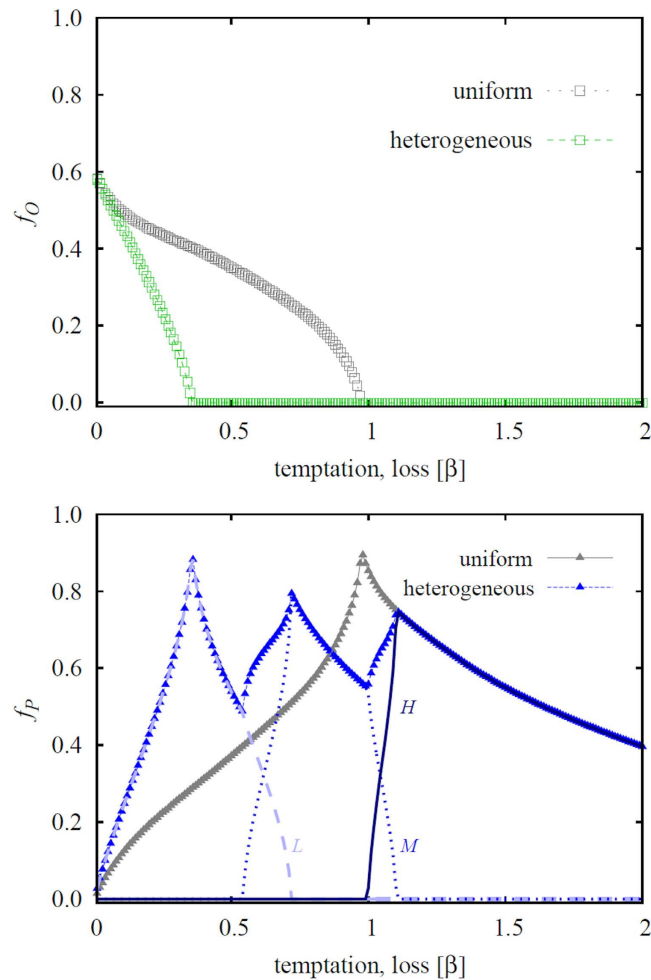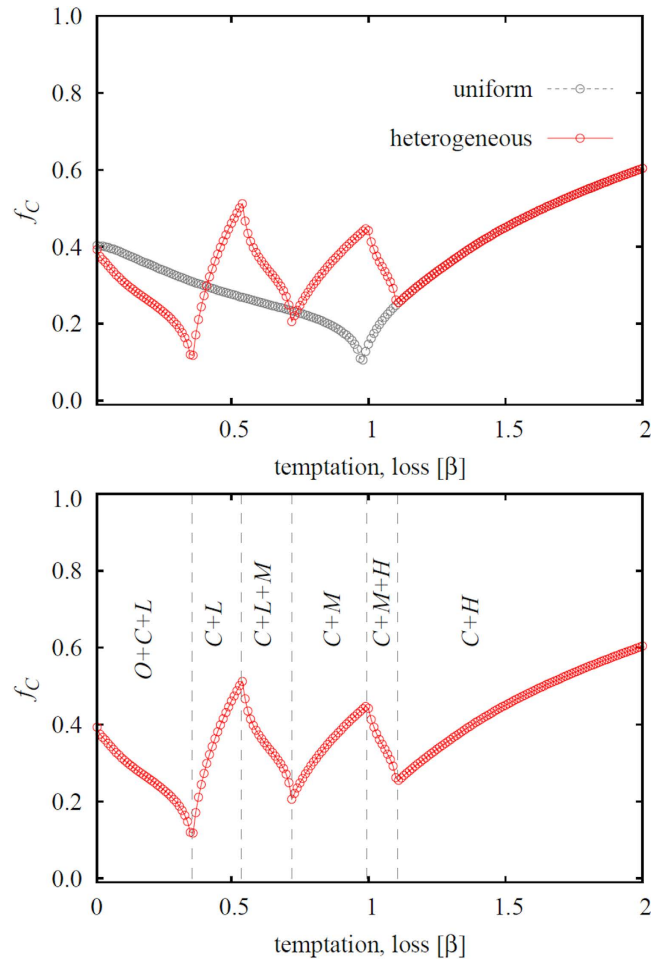
Figure 3. Left panel shows the fraction of ordinary people in dependence on the temptation/loss β, as obtained for the three-strategy spatial inspection game with uniform punishment and the five-strategy spatial inspection game with heterogeneous punishment (see legend). It can be observed that heterogeneous punishment is indeed more effective in eliminating second-order free-riding by ordinary people than uniform punishment. Right panel shows the fraction of punishers in dependence on the temptation/loss β for the uniform punishment model and the aggregate fraction of all punishers in the heterogeneous punishment model, as well as the fraction of punishers *L*, *M* and *H* individually (see legend). The success of heterogeneous punishment to eliminate second-order free-riding is somewhat relativized, as higher punishment levels will not necessarily lead to lower criminal levels (see Fig. 4 for an explanation). The origin of the zig-zag outlay of the aggregate fraction of all punishers is analyzed in Fig. 5. In both panels the punishment cost is $\alpha = 0.5$ and the reward for punishing criminals is $\gamma = 1.5$.

of the full cost, respectively. However, their reward is proportionally smaller as well (see the extended payoff matrix in Section 2 for details). Due to the large number of competing strategies and the resulting multitude of possible subsystem solutions we focus on the most important parameter region where ordinary players survive in the uniform, three-strategy, model. Accordingly, we explore a representative cross section when the reward is high enough for punishing strategies to survive, and we explore how the system responds to the diversity of punishment.

Results presented in the left panel of Fig. 3 confirm the effectiveness of resorting to heterogeneous punishment in that second-order free-riders are able to survive only in a significantly narrower interval of the temptation/loss β if compared to the uniform punishment model. Furthermore, results presented in the right panel of Fig. 3 also give credence to the expectation that the reduced viability of ordinary people will promote the evolution of punishers. More precisely, we find that the uniform punishment strategy is significantly less effective than heterogeneous punishment for almost the entire range of the temptation/loos β, except for a narrow interval in the $\beta > \alpha$ region. As we will show in Fig. 4 below, this fact has important consequences for the mitigation of criminal behavior in the population.

Another peculiarity that can be observed in the right panel of Fig. 3 is the zig-zag outlay of the aggregate fraction of all punishers in the five-strategy model. Yet this can be understood thoroughly simply by

Figure 4. Top panel shows the fraction of criminals in dependence on the temptation/loss β, as obtained for the three-strategy spatial inspection game with uniform punishment and the five-strategy spatial inspection game with heterogeneous punishment (see legend). It can be observed that heterogeneous punishment is more effective than uniform punishment in eliminating crime only in the low β limit, which also agrees with the region in which second-order free-riding is deterred more efficiently (see Fig. 3). In general, however, uniform punishment works just as well or better than heterogeneous punishment in abating crime. Bottom panel again shows the fraction of criminals, along with the different phases that contain the C strategy. Despite the multitude of consecutive phase transitions in dependence on solely a single parameter, criminal behavior is never completely eliminated. In both panels the punishment cost is $\alpha = 0.5$ and the reward for punishing criminals is $\gamma = 1.5$.

looking at the fraction of punishers L, M and H individually. The mentioned panel reveals clearly that low values of β are able to sustain only those punishers who are willing to invest the lowest cost towards sanctioning criminals. The rank of the most viable punishers subsequently increases from L over M to H as we increase β, and the solution of the five-strategy model thus eventually becomes identical to the the solution of the three-strategy model. Remarkably, we can observe six consecutive phase transitions $[(O + C + L) \rightarrow (C + L) \rightarrow (C + L + M) \rightarrow (C + M) \rightarrow (C + M + H) \rightarrow (C + H) \rightarrow C]$ as we increase a single parameter, β. It is worth pointing out that the reported increment of the punisher rank with increasing the temptation/loss β resonates with the outcome of a recent human experiment[54], where, in the realm of a social dilemma, it was shown that if cooperation is likely one should punish mildly.

We continue with the results presented in Fig. 4, where we compare the effectiveness of uniform and heterogeneous punishment to deter criminal behavior. To a degree unexpected, it can be observed that the possibility to resort to different levels of punishment does not necessarily work better than uniform punishment in reducing crime. On the contrary, the fraction of C players is generally higher over a large interval of β values when the heterogeneous punishment model is used. More precisely, the fraction of criminals is lower only in the low temptation/loss region where L punishers can adjust to this favorable condition. This observation is related to the failure of heterogeneous punishment to eliminate second-order free-riding more effectively than uniform punishment, and it indicates that sophisticatedly adjusted punishers may win a battle against ordinary people, but loose the main war against the actual
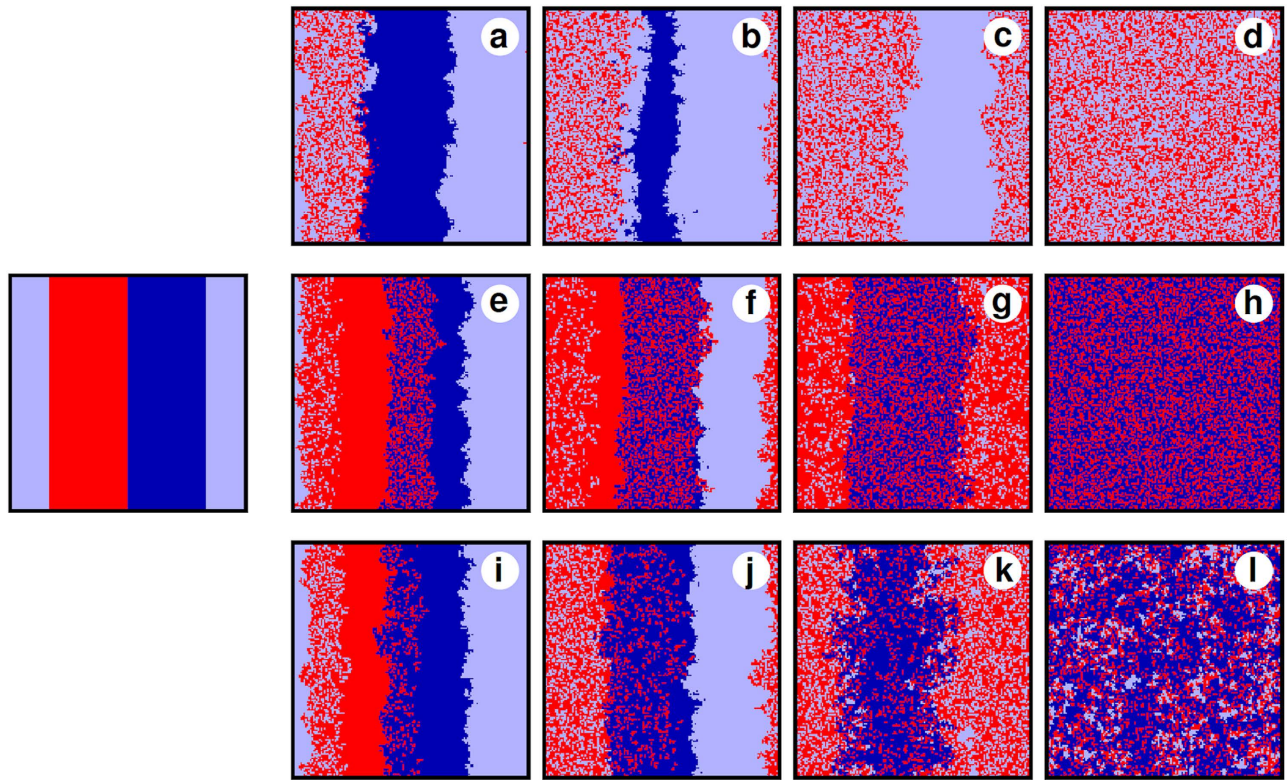
Figure 5. Time evolution of strategy distributions in the population, as obtained with the heterogeneous punishment game starting from the same prepared initial state (leftmost panel) for $\gamma = 1.5$ and three different values of the temptation/loss: (**a**)–(**d**) $\beta = 0.5$, (**e**)–(**h**) $\beta = 0.9$, and (**i**)–(**l**) $\beta = 0.7$. The resulting three different stationary states are reached within 400 MCS, which are depicted in the rightmost panels. Colors red, light blue and dark blue depict the location of C, L and M players, respectively. For visual clarity, we have used a small 150×150 system size. See main text for a detailed description of the different evolutionary outcomes.

enemy, the criminals. While punishers can lower the amount they invest towards sanctioning criminals, such a reduced effort also yields smaller rewards. Interestingly, the positive side of lower costs can be utilized only if the heterogeneity of punishers is maintained. The said effect becomes visible if we mark the borders of different phases on the curve of criminals, as shown in the right panel of Fig. 4. As it is illustrated, the fraction of criminals can be a decaying function even if we increase the temptation/loss β, but only as long as different types of punishers exist and compete against the criminals. As soon as evolution favors a single punisher type, an effective response to an increase of the value of β becomes absent. Lastly, we note that the conclusions attained with the results presented in Figs 3 and 4 remain generally valid also for all high temptation values.

To obtain a better understanding of the origin of the zig-zag outlay of criminals depicted in Fig. 4, we monitor the time evolution of the distribution of strategies in the population for three different combinations of payoff parameters, as shown in Fig. 5. We emphasize that the main mechanism responsible for the formation of different stationary states is due to the different motion of interfaces that separate the possible solutions of the system. Accordingly, we follow the evolution of interfaces starting from a prepared initial state, but for clarity only two types of punishers are present because this minimal model is sufficient to capture the essence of the emerging effect. The extrapolation to the full five-strategy model, however, is straightforward. For comparison, we use an identical prepared initial state, as shown in the leftmost panel, for three representative values of β. As in previous figures, red color depicts C players while light and dark blue depict the L and M punishers, respectively. Before discussing each specific case, we note that, individually, L always beats M due to the lower cost of inspection. When the temptation/loss is low, as shown in panels (a)–(d), M can beat C very efficiently, while L is unable to do the same but simply coexists with the criminals. The superiority of L over M, however, will result in a shrinking area of the M domain, as shown in panel (b). Ultimately, this fact leads to the extinction of strategy M, despite the fact that it is more successful in deterring criminals than strategy L. As soon as M die out, as shown in panel (c), criminals can exploit the milder punishment from strategy L and spread towards the stationary state, as shown in panel (d). A seemingly surprising and counterintuitive result is that

criminals, who can coexist with L players but are defeated by M players, are able to survive while their "predators" (M) go extinct. But in fact, the evolution depicted in the panels (a)–(d) simply illustrates the actual consequence of second-order free-riding. Namely, L players exploit the more altruistic M players by contributing less to sanctioning criminals. In the absence of L players, however, the common enemy (C) can spread relatively free and reach a significantly high level ($f_C \approx 0.46$).

Interestingly, when M players are less successful in deterring C players, the outcome is completely the opposite, as shown in panels (e)–(h) of Fig. 5. Since the temptation/loss $\beta = 0.9$, C are able to coexist with M. The coexistence of C and L strategies is also still possible, and at the same time L continue to invade the pure M phase [the invasion ends in panel (f)]. However, L become ineffective against the $C + M$ alliance. Indeed, this two-strategy alliance is so powerful that it beats the other $C + L$ alliance completely. The competition between the two alliances starts in panel (g), and it terminates with the total victory of the $C + M$ alliance in panel (h). The conclusion is similar as in the preceding case. Namely, when the evolution selects only one type of punishers, then criminals have a reasonable chance to survive. Note that the fraction of criminals in the stationary state is again relatively high, $f_C \approx 0.40$, despite of substantial punishment.

The most favorable outcome can be obtained at an intermediate temptation/loss value, as shown in panels (i)–(l) of Fig. 5. The $\beta = 0.7$ value is still high enough to maintain the coexistence of the $C + M$ alliance, but it lessens its evolutionary advantage in that the $C + L$ alliance is able to survive. The stationary state thus contains three strategies, whereby a relatively small portion of the population, $f_C \approx 0.27$, is occupied by criminals. We thus conclude that, in the long-run, if different punisher strategies survive in the stationary state, heterogeneous punishment may be utilized successfully to mitigate crime better than uniform punishment. Note that f_C is a decreasing function of β in the three-strategy phase in Fig. 4, while it always increasing when homogeneous punishment is applied (in $C + L$, $C + M$, or in the $C + H$ phases). This is because heterogeneous punishment enables the validation of the most effective approach against crime: sometimes moderate efforts, yielding milder fines, serve the interest of whole population better than severe punishment. Even more importantly, the simultaneous presence of different types of punishers enables a synergy among them in that one strategy (in our case M) can lower the payoff of criminals significantly while the other strategy (L) can still enjoy a more competitive payoff due to a smaller cost. This multi-point effect is conceptually similar to when the duty of punishment is shared stochastically among cooperative players[45]. Of course, as we have already emphasized, these conclusions remain valid and can be extrapolated to a larger number of different punisher strategies.

Discussion

We have studied the effectiveness of punishment in abating criminal behavior in the spatial inspection game with three and five competing strategies, entailing criminals, ordinary people and punishers. In the five-strategy game, we have introduced three different types of punishers, depending on the amount they are willing to contribute towards sanctioning criminals. We have shown that cyclic dominance plays an important role in that it maintains the survivability of seemingly subordinate strategies through indirect support. For example, increasing the reward for punishing criminals might promote second-order free-riding of ordinary people, despite of the fact that it should in fact support the punishers. This is due to cyclic dominance, where directly promoting the prey, in this case the punishers, benefits the predator, which in this case are the ordinary people. Moreover, we have shown that the actual obstacle in the fight against criminal behavior is the possibility of ordinary people to free-ride on the efforts of punishers, which is also the main culprit behind the establishment of cyclic dominance. In general, sanctioning criminal behavior is thus a double-edged sword. The obvious benefit is that the evolution of crime is contained and is unable to dominate in the population. The pitfall is that, in conjunction with ordinary people, punishment creates conditions that support cyclic dominance, which prevents the complete abolishment of crime even if the sanctions are severe and effective.

In addition to these observations, we have shown that the possibility of heterogeneous punishment yields a highly ambiguous measure against criminal behavior. At specific parameter values it can happen that milder punishers play the role of second-order free riders, which ultimately prevents to eliminate crime completely [see panels (a)–(d) in Fig. 5]. Evidently, the reverse process is also possible in structured populations where the more altruistic punishers can separate from second-order free riders and win the indirect territorial battle[31,55]. But in the realm of the studied inspection game, we have also observed that the diversity of punishers can yield a more favorable social outcome even as the temptation to do crime is growing. In the latter case, the simultaneous presence of different punishers provides an advantageous coexistence: some punishers ensure a higher fine to criminal players while other punishers can benefit from a lower cost due to a less intensive engagement. Importantly, neither of these two options is effective on its own right, but together they improve the effectiveness of combating crime.

Notably, the emergence of cyclic dominance due to strategic complexity has been reported before, for example in public goods games with volunteering[56], peer punishment[31,57–59], pool punishment[43,44] and reward[39,60], but also in pairwise social dilemmas with coevolution[61,62]. Other counterintuitive phenomena that are due to cyclic dominance[63,64] include the survival of the weakest[52,65], the emergence of labyrinthine clustering[66], and the segregation along interfaces that have internal structure[67], to name but a few examples. Cyclical interactions are thus in many ways the culmination of evolutionary complexity[13], and we here show that they likely play a prominent role in deterring crime as well. However, while the

beneficial role of cyclic dominance for maintaining biodiversity is undeniable, one has to concur that it is a rather unsatisfactory outcome in terms of fighting criminal behavior. That is the sort of diversity in behavior that human societies could happily do without, yet it seems that this is precisely the trap the current system has fallen into. Indeed, data from the Federal Bureau of Investigation (see Fig. 2 in Ref. 7) indicate that crime, regardless of type and severity, is remarkably recurrent. Although positive and negative trends may be inferred, crime events between 1960 and 2010 fluctuate across time and space, and there is no evidence to support that crime rates are permanently decreasing. The search for more effective crime mitigation strategies is thus in order, in particularly for such where the permanent elimination of crime is not an a priori impossibility.

Methods

For both the 3-strategy and the 5-strategy spatial inspection game the Monte Carlo simulation procedure is the same. Initially all competing strategies are distributed uniformly at random on the square lattice. We note, however, that the reported final stationary states are largely independent of the initial fractions of strategies. Subsequently, in agreement with the random sequential update protocol, a randomly selected player x acquires its payoff Π_x by playing the game pairwise with all its four neighbors. Next, player x randomly chooses one neighbor y, who then also acquires its payoff Π_y in the same way as previously player x. Once both players acquire their payoffs, player x adopts the strategy s_y from player y with a probability determined by the Fermi function

$$W\left(s_y \rightarrow s_x\right) = \frac{1}{1 + \exp\left[\left(\Pi_x - \Pi_y\right)/K\right]},$$

(1)

where $K = 0.5$ quantifies the uncertainty related to the strategy adoption process[10,68]. In agreement with previous works, the selected value ensures that strategies of better-performing players are readily adopted by their neighbors, although adopting the strategy of a player that performs worse is also possible[69,70]. This accounts for imperfect information and errors in the evaluation of the opponent.

Each full Monte Carlo step (MCS) consists of L^2 elementary steps as described above, which are repeated consecutively, thus giving a chance to every player to change its strategy once on average. We typically use lattices with 600×600 players, although close to the phase transition points up to 9000×9000 players had to be used in this case to avoid accidental extinctions, and thus to arrive at results that are valid in the large-size limit. The fractions of competing strategies f are determined in the stationary state after a sufficiently long relaxation time lasting up to 10^5 MCS. In general, the stationary state is reached when the average of the strategy fractions becomes time-independent. Moreover, to account for the differences in initial conditions and to further improve accuracy, the final results are averaged over up to 100 independent runs for each set of parameter values.

References

1. Wilson, J. Q. & Kelling, G. L. Broken windows. *The Atlantic Monthly* **249,** 29–38 (1982).
2. Rigney, D. *The Matthew effect: How advantage begets further advantage* (Columbia University Press, New York, 2013).
3. Perc, M. The Matthew effect in empirical data. *J. R. Soc. Interface* **11,** 20140378 (2014).
4. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509–512 (1999).
5. Albert, R. & Barabási, A. L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74,** 47–97 (2002).
6. Gell-Mann, M. Simplicity and complexity in the description of nature. *Eng. Sci.* **57,** 2–9 (1988).
7. D'Orsogna, M. R. & Perc, M. Statistical physics of crime: A review. *Phys. Life. Rev.* **12,** in press (2015).
8. Wang, Z., Szolnoki, A. & Perc, M. If players are sparse social dilemmas are too: Importance of percolation for evolution of cooperation. *Sci. Rep.* **2,** 369 (2012).
9. Yang, H.-X., Rong, Z. & Wang, W.-X. Cooperation percolation in spatial prisoner's dilemma game. *New J. Phys.* **16,** 013010 (2014).
10. Szabó, G. & Fáth, G. Evolutionary games on graphs. *Phys. Rep.* **446,** 97–216 (2007).
11. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81,** 591–646 (2009).
12. Pacheco, J. M., Vasconcelos, V. V. & Santos, F. C. Climate change governance, cooperation and self-organization. *Phys. Life. Rev.* **11,** 573–586 (2014).
13. Szolnoki, A. *et al.* Cyclic dominance in evolutionary games: a review. *J. R. Soc. Interface* **11,** 20140735 (2014).
14. Alves, L. G. A., Ribeiro, H. V., Lenzi, E. K. & Mendes, R. S. Distance to the scaling law: A useful approach for unveiling relationships between crime and urban metrics. *PLoS One* **8,** e69580 (2013).
15. Picoli, S., del Castillo-Mussot, M., Ribeiro, H. V., Lenzi, E. K. & Mendes, R. S. Universal bursty behaviour in human violent conflicts. *Sci. Rep.* **4,** 4773 (2014).
16. Short, M. B. *et al.* A statistical model of criminal behavior. *Math. Mod. Meth. Appl. Sci.* **18,** 1249–1267 (2008).
17. Short, M. B., Brantingham, P. J., Bertozzi, A. L. & Tita, G. E. Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proc. Natl. Acad. Sci. USA* **107,** 3961–3965 (2010).
18. Rodriguez, N. & Bertozzi, A. Local existence and uniqueness of solutions to a PDE model for criminal behavior. *Math. Mod. Meth. Appl. Sci.* **20,** 1425–1457 (2010).
19. Berestycki, H., Rodriguez, N. & Ryzhik, L. Reaction-diffusion model for criminal activity. *Multiscale Model. Sim.* **11,** 1097–1126 (2013).
20. Hegemann, R. A. *et al.* Geographical influences of an emerging network of gang rivalries. *Physica A* **390,** 3894–3914 (2011).
21. Catanese, S., Ferrara, E. & Fiumara, G. Forensic analysis of phone call networks. *Soc. Netw. Anal. Mining* **3,** 15–33 (2013).
22. Ferrara, E., Meo, P. D., Catanese, S. & Fiumara, G. Detecting criminal organizations in mobile phone networks. *Expert Syst. Appl.* **41,** 5733–5750 (2014).
23. Duijn, P. A. C., Kashirin, V. & Sloot, P. M. A. The relative ineffectiveness of criminal network disruption. *Sci Rep* **4,** 4238 (2014).

24. Santos, F. C., Pacheco, J. M. & Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. USA* **103,** 3490–3494 (2006).
25. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314,** 1560–1563 (2006).
26. Rand, D. A. & Nowak, M. A. Human cooperation. *Trends Cog. Sci.* **17,** 413–425 (2013).
27. Short, M. B., Brantingham, P. J. & D'Orsogna, M. R. Cooperation and punishment in an adversarial game: How defectors pave the way to a peaceful society. *Phys. Rev. E* **82,** 066114 (2010).
28. Fehr, E. Don't lose your reputation. *Nature* **432,** 449–450 (2004).
29. Sigmund, K. Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **22,** 593–600 (2007).
30. Panchanathan, K. & Boyd, R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* **432,** 499–502 (2004).
31. Helbing, D., Szolnoki, A., Perc, M. & Szabó, G. Evolutionary establishment of moral and double moral standards through spatial interactions. *PLoS Comput. Biol.* **6,** e1000758 (2010).
32. D'Orsogna, M. R., Kendall, R., McBride, M. & Short, M. B. Criminal defectors lead to the emergence of cooperation in an experimental, adversarial game. *PLoS One* **8,** e61458 (2013).
33. Dreber, A., Rand, D. G., Fudenberg, D. & Nowak, M. A. Winners don't punish. *Nature* **452,** 348–351 (2008).
34. Herrmann, B., Thoni, C. & Gachter, S. Antisocial punishment across societies. *Science* **319,** 1362–1367 (2008).
35. Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D. & Nowak, M. A. Positive interactions promote public cooperation. *Science* **325,** 1272–1275 (2009).
36. Rand, D. G. & Nowak, M. A. The evolution of antisocial punishment in optional public goods games. *Nat. Commun.* **2,** 434 (2011).
37. Vukov, J., Pinheiro, F., Santos, F. & Pacheco, J. Reward from punishment does not emerge at all costs. *PLoS Comput. Biol.* **9,** e1002868 (2013).
38. Hilbe, C. & Sigmund, K. Incentives and opportunism: from the carrot to the stick. *Proc. R. Soc. B* **277,** 2427–2433 (2010).
39. Szolnoki, A. & Perc, M. Evolutionary advantages of adaptive rewarding. *New J. Phys.* **14,** 093016 (2012).
40. Szolnoki, A. & Perc, M. Correlation of positive and negative reciprocity fails to confer an evolutionary advantage: Phase transitions to elementary strategies. *Phys. Rev. X* **3,** 041021 (2013).
41. Berenji, B., Chou, T. & D'Orsogna, M. R. Recidivism and rehabilitation of criminal offenders: A carrot and stick evolutionary game. *PLoS One* **9,** e85531 (2014).
42. Gurerk, O., Irlenbusch, B. & Rockenbach, B. The competitive advantage of sanctioning institutions. *Science* **312,** 108–111 (2006).
43. Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466,** 861–863 (2010).
44. Szolnoki, A., Szabó, G. & Perc, M. Phase diagrams for the spatial public goods game with pool punishment. *Phys. Rev. E* **83,** 036101 (2011).
45. Chen, X., Szolnoki, A. & Perc, M. Probabilistic sharing solves the problem of costly punishment. *New J. Phys.* **16,** 083016 (2014).
46. Becker, G. S. Crime and punishment: An economic approach. *J. Pol. Econ.* **76,** 169–217 (1968).
47. Tsebelis, G. Penalty has no impact on crime: A game theoretic analysis. *Ration. Soc.* **2,** 255–286 (1990).
48. Perc, M., Donnay, K. & Helbing, D. Understanding recurrent crime as system-immanent collective behavior. *PLoS One* **8,** e76063 (2013).
49. Perc, M. & Szolnoki, A. Social diversity and promotion of cooperation in the spatial prisoner's dilemma game. *Phys. Rev. E* **77,** 011904 (2008).
50. Santos, F. C., Santos, M. D. & Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454,** 213–216 (2008).
51. Santos, F. C., Pinheiro, F., Lenaerts, T. & Pacheco, J. M. Role of diversity in the evolution of cooperation. *J. Theor. Biol.* **299,** 88–96 (2012).
52. Frean, M. & Abraham, E. D. Rock-scissors-paper and the survival of the weakest. *Proc. R. Soc. Lond. B* **268,** 1323–1327 (2001).
53. Szabó, G., Szolnoki, A. & Izsák, R. Rock-scissors-paper game on regular small-world networks. *J. Phys. A: Math. Gen.* **37,** 2599–2609 (2004).
54. Jiang, L.-L., Perc, M. & Szolnoki, A. If cooperation is likely punish mildly: Insights from economic experiments based on the snowdrift game. *PLoS One* **8,** e64677 (2013).
55. Helbing, D., Szolnoki, A., Perc, M. & Szabó, G. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. *New J. Phys.* **12,** 083005 (2010).
56. Szabó, G. & Hauert, C. Phase transitions and volunteering in spatial public goods games. *Phys. Rev. Lett.* **89,** 118101 (2002).
57. Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Via freedom to coercion: The emergence of costly punishment. *Science* **316,** 1905–1907 (2007*).
58. Amor, D. R. & Fort, J. Effects of punishment in a mobile population playing the prisoner's dilemma game. *Phys. Rev. E* **84,** 066115 (2011).
59. Bednarik, P., Fehl, K. & Semmann, D. Costs for switching partners reduce network dynamics but not cooperative behaviour. *Proc. R. Soc. B* **281,** 20141661 (2014).
60. Szolnoki, A. & Perc, M. Reward and cooperation in the spatial public goods game. *EPL* **92,** 38003 (2010).
61. Szolnoki, A. & Perc, M. Resolving social dilemmas on evolving random networks. *EPL* **86,** 30007 (2009).
62. Szolnoki, A., Wang, Z., Wang, J. & Zhu, X. Dynamically generated cyclic dominance in spatial prisoner's dilemma games. *Phys. Rev. E* **82,** 036110 (2010).
63. Ni, X., Wang, W.-X., Lai, Y.-C. & Grebogi, C. Cyclic competition of mobile species on continuous space: Pattern formation and coexistence. *Phys. Rev. E* **82,** 066211 (2010).
64. Wang, W.-X., Ni, X., Lai, Y.-C. & Grebogi, C. Pattern formation, synchronization, and outbreak of biodiversity in cyclically competing games. *Phys. Rev. E* **83,** 011917 (2011).
65. Berr, M., Reichenbach, T., Schottenloher, M. & Frey, E. Zero-one survival behavior of cyclically competing species. *Phys. Rev. Lett.* **102,** 048102 (2009).
66. Juul, J., Sneppen, K. & Mathiesen, J. Labyrinthine clustering in a spatial rock-paper-scissors ecosystem. *Phys. Rev. E* **87,** 042702 (2013).
67. Avelino, P. P., Bazeia, D., Losano, L., Menezes, J. & de Oliveira, B. F. Interfaces with internal structures in generalized rock-paper-scissors models. *Phys. Rev. E* **89,** 042710 (2014).
68. Blume, L. E. The statistical mechanics of strategic interactions. *Games Econ. Behav.* **5,** 387–424 (1993).
69. Vukov, J., Szabó, G. & Szolnoki, A. Cooperation in the noisy case: Prisoner's dilemma game on two types of regular random graphs. *Phys. Rev. E* **73,** 067103 (2006).
70. Szolnoki, A., Perc, M. & Szabó, G. Topology-independent impact of noise on cooperation in spatial public goods games. *Phys. Rev. E* **80,** 056109 (2009).

Acknowledgments

Author Contributions

M.P. and A.S. designed and performed the research as well as wrote the paper.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Perc, M. & Szolnoki, A. A double-edged sword: Benefits and pitfalls of heterogeneous punishment in evolutionary inspection games. *Sci. Rep.* **5**, 11027; doi: 10.1038/srep11027 (2015).