Structured populations facilitate cooperation in policed Public Goods Games

G. A. Kaiping\textsuperscript{a,}\textsuperscript{*}, T. J. Sluckin\textsuperscript{b}, S. J. Cox\textsuperscript{a}

\textsuperscript{a}Computational Engineering and Design Group, Building 176, University of Southampton, Boldrewood Innovation Campus, Burgess Road, Southampton SO16 7QF, UK
\textsuperscript{b}Mathematical Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, UK

Abstract

Societies consisting of cooperative individuals seem to require for their continuing success that defectors be policed. The precise connection between punishers and benefits, population structure, and division of labour, however, remains ill-understood. Many models assume costly “peer punishment” to enforce cooperation, but results in the economics literature suggest that this assumption may not be generally valid. In many human and animal societies, there is a division of labour between a purely supportive majority and a dedicated minority of police-like enforcers. Here we present several extensions to the Public Goods Game with punishment which allow for this possibility, and evaluate their influence on the level of cooperative behaviour. We find that a structure of separate subpopulations, which only interact through migration of individuals, can have a strong effect on the evolutionary dynamics of a system and significantly facilitate cooperation. Forcing defectors to contribute and enabling fitness transfers to punishers both have a weak positive effect on cooperation levels. In the presence of group competition, however, evolutionary effects can paradoxically hinder cooperation.

Highlights.

\begin{itemize}
  \item Group selection can happen through the migration of individuals, if migration rates are fitness dependent.
  \item Migration-based group selection can allow stable punishment to evolve.
  \item Monitors that force defectors to contribute to the public good are beneficial in well-mixed population, but detrimental under group selection.
  \item Mandatory fitness transfers to punishers support cooperation, while optional fitness transfers to punishers are significantly more deleterious.
\end{itemize}

Keywords: group selection, delegated punishment, island migration
PACS: 87.23.-n, 87.10.Mn

1. Introduction

The evolution of large-scale cooperation is a well-known puzzle, which has been much studied in the context of evolutionary game theory \cite{5}. It has been shown that in many circumstances, punishment can stabilise a cooperative system that would otherwise be dominated by defection. In two-player games, the iterated Prisoner’s Dilemma allows a very simple “tit for tat” strategy, which immediately punishes opponents after they defected. In multi-player contexts, the evolution of cooperation is usually studied in terms of the Public Goods Game \cite{40}. Here, usually one of two variants of punishment is adopted.

On the one hand, players can be allowed to punish defectors in addition to other strategy choices. Peer punishers were introduced to the literature by Boyd and Richerson \cite{11}, but the concept only became widespread after Fehr and Gächter \cite{14,15} published experimental results showing that humans seem to punish defectors even when it is not in their material (short-term!) self-interest. Peer punishment is usually considered to be costly for the punisher and even more expensive for the punished, and the costs are assumed to be proportional to the numbers of defectors and punishers respectively. Following Fehr and Gächter’s experiments, the concept of costly punishment has gained significant influence in both the modelling and the experimental literatures. Costly punishment options may be restricted to cooperators, who are called “peer punishers” if

\textsuperscript{*}Corresponding author

Preprint submitted to Elsevier

March 16, 2015
they choose that option \[12, 43\] or include other strategies like “immoralists” (defectors punishing other defectors) \[46\].

On the other hand, players can be allowed to pay a less costly sum before the Public Goods Game proper. If any player paid the sum, all defectors are punished by a fixed amount independent of the number of so-called “pool punishers”. If no pool punishers are present in a round, no punishment occurs. The sum paid to the punishment pool is not refunded however, even if no defectors were present. Economical experiments have shown \[23\] that humans prefer such delegated punishment to peer punishment.

Szolnoki et al. \[45\] studied the phase space of the Public Goods Game with peer and pool punishers, describing phases in which each of these two punishment strategies could prevail. Instead of being a separate strategy, pool punishment has also been realised in terms of institutions that deduct the cost for punishment from the public good \[36\].

However, neither ansatz seems to capture all key features of punishment in real-world systems. In pool punishment systems the punishment is an external mechanism not subject to the evolutionary pressures of the system, and can therefore not explain how such strategies arise in a given system. Peer punishment on the other hand is costly and beneficial to the whole system. As such, this method of punishment constitutes a second-order public good. Peer punishment would thus be subject to the same Tragedy of the Commons as the original Public Goods Game, with non-punishing cooperators dominating peer-punishers.

In this article, we address three specific points where the usual implementation of Public Goods Games with peer punishment \[40, 11\] deviates from the perceived reality \[32, 19\]. We present extensions to the game-theoretical model that can be considered to address these shortcomings, and study their impact on the evolutionary dynamics of the Public Goods Game. The points of criticism and corresponding extensions are as follows.

Firstly, it has been suggested that local structure may play an important role in the evolution of cooperation. Models with regular local structure have been known to show non-trivial interactions between cooperators and defectors for long time \[12, 22, 27, 44\]. Recently the so-called Simpson’s paradox \[42, 36, 35\] has been proposed as a driving mechanism for global prevalence of cooperators. This effect becomes relevant when the large growth rate of a cooperative sub-population becomes relevant, migration between subpopulations will not be random, but instead will be mediated through Link Dynamics updating. We will study the impact of this extension in section 3.1.

Secondly, recent research in both the economics \[31, 20\] and the ecology \[38, 30, 16\] literature have suggested that part of the fine, i.e. the fitness penalty inflicted on defectors by punishers, can be used to compensate the punisher. This allows lower or even negative punishment costs, allowing punishers to prevail in defector-rich environments. We will study the impact of low punishment cost in the context of our basic model in section 2. We will observe that direct benefits can enable monitors to invade non-cooperative populations.

Alternatively, some of the fitness penalty incurred by punished agents can reflect that they are forced to invest into the public good. This should indirectly benefit the punishing individuals. In order to implement this point in the Public Goods Game, we replace the peer punisher strategy by a strategy with a similar payoff structure, which we will refer to as Monitor. Just like peer punishers, monitors will be able to punish defectors by paying a cost proportional to the number of defectors. The punished defectors will, however, not just pay a fine, but also be forced to invest into the public good. We will see that such indirect benefits are marginally helpful for increasing cooperation in an isolated population, but can be detrimental in a structured population.

Lastly, just as humans prefer third parties to undertake the punishment of wrong-doers on their behalf, eusocial insects exhibit a division of labour between policing and contributing workers \[8, 51\]. Such a division of labour cannot arise in current peer punisher models of cooperation. In order to change this, we include the possibility that cooperators may explicitly transfer fitness (“tax”) to punishers. This tax payment is voluntary and initiated by the cooperators. We find that in this case, the impact of the extension on the evolutionary dynamics of the model is marginally supporting cooperation.

The organisation of this paper is as follows. In section 2 we will describe our base-line model for a Public Goods Game with peer punishment. We will quote some well-known results for this model and derive additional results that have previously been shown for similar models for the specific variant of the Public Goods Game with punishment that we take as the base line for this article.

We will then present three extensions to the base-line model in section 3. In the first part of this section, we will extend the baseline model by adding an island-based local structure. We will demonstrate that such a structure can have a large influence on the results of the model, and can demonstrate Simpson’s paradox. In an ecological context this signifies that in each individual sub-population defection dominates punishment and cooperation. Globally, however, in apparent contradiction to this, high levels of cooperation are maintained. We will then substitute punishers by monitors and study the effect of direct and indirect benefits of punishment to the monitors. As our last modification to the baseline model, we will add tax payment options to the model to see under what circumstances it increases levels of cooperation.

Finally, in section 4 we conclude with a general discussion of the properties of the models studied. We identify the parameter areas where the described extensions have the largest impact, and relate those areas to real-world systems that show similar modifications of the Public Goods Game.
1. overview of the parameters used in this article is given in Table defector, in order to inflict a fine interest of the experimental subjects. Players following strategy even if the punishment is designed to be not in the material self-

2. Public Goods Games with Peer Punishment

The Public Goods Game is a well-understood game-theoretical model, which is frequently used to study the social dilemma known as the Tragedy of the Commons [21]. The basic Linear Public Goods Game (LPGG), which we summarise in this section, is a symmetric $N$ player game. Agents playing the LPGG choose either to invest a fixed cost $c$ into a public venture, or to abstain from investing. Investing players are called cooperators (C), abstaining players are referred to as defectors (D). The investments into the public venture are then scaled up by a linear factor $1 < r < N$ and split between all players.

This game is often extended by a peer punisher (P) strategy. The concept was introduced to the experimental economics literature in 2000 by Ernst Fehr and Simon Gächter [14]. They showed experimentally that humans tend to punish defectors even if the punishment is designed to be not in the material self-interest of the experimental subjects. Players following strategy P do not only provide for the public good, but additionally punish defectors. After the first stage payoff has been declared as described above, each punisher pays a punishment cost $\gamma$ per defector, in order to inflict a fine $\beta$ each on every defector. An overview of the parameters used in this article is given in Table [11].

The payoff structure of the resulting Public Goods Game with peer punishment (PGG/P) with $n_i$ players following strategy $s$ respectively is therefore given by

\[
\pi_D = \Pi - \frac{\beta n_p}{N},
\]

\[
\pi_C = \Pi - c,
\]

\[
\pi_P = \Pi - c - \gamma \frac{n_D}{N},
\]

where

\[
\Pi = rc \frac{n_C + n_P}{N}
\]

is the individual’s share in the public good. The payoffs are structured such that in a well-mixed population in the absence of punishers (and even in the presence of punishers, as long as $\frac{c}{N} < \frac{\gamma}{\beta}$), defectors dominate cooperators. When all three strategies are present, cooperators dominate punishers. Following Boyd and coworkers [10, 34], the more cooperative strategies are assumed to be error-prone, i.e. every cooperator has a low probability $\epsilon$ to behave like a defector in each time step, and every punisher punish with a probability of $1 - 2\epsilon$, but has a probability $\epsilon$ each to behave like a cooperator or a defector. If we denote the realised frequencies of strategies with $n_i$ and the general frequencies with $\bar{\pi}_s$, i.e. $n_C = (1 - \epsilon)\bar{n}_C + \epsilon\bar{n}_D$ etc., then for the expected payoffs $\bar{\pi}_s$ we have

\[
\bar{\pi}_D = \frac{\Pi - (1 - 2\epsilon)\beta \bar{n}_P}{N},
\]

\[
\bar{\pi}_C = \Pi - (1 - \epsilon)c - \epsilon(1 - 2\epsilon)\beta \bar{n}_P,
\]

\[
\bar{\pi}_P = \Pi - (1 - \epsilon)c - (1 - 2\epsilon)\gamma \frac{\bar{n}_D}{N} = \epsilon(1 - 2\epsilon)\beta \frac{\bar{n}_P}{N}.
\]

We will hereafter use these transformed payoffs, instead of explicitly modelling errors in the execution of strategies.

With this payoff structure, cooperators dominate punishers even in the absence of defection. It is therefore evident that in this model defection is the only evolutionarily stable strategy [10]. However, mutation permits other stable equilibria to exist [49].

Just as in previous implementations of the Public Goods Game with punishment, we investigate the evolutionary context of populations of fixed size with overlapping generations. In every time step, one individual of the population dies and one individual procreates. The probabilities for birth and death are obtained from the payoffs using an updating rule. One such updating rule that has already been used in the study of the PGG/P [10] is the Link Dynamics updating rule [45, 28].

The present authors have shown elsewhere [28] that different updating rules lead to non-trivially different equilibria in evolutionary systems with more than two strategies. In this model, both the dynamics for low payoffs – due to punishment

Figure 1: A cartoon of the evolutionary dynamics of the basic Public Goods Game with Punishment, with defectors (red), cooperators (black) and punishers (green).
costs and fines – and the dynamics for high payoffs – where defectors fare better than cooperators in high-cooperation low-punishment populations – are relevant for this model. We therefore need to employ an updating rule that implements both variable birth and variable death probabilities, dependent on the payoffs. Link Dynamics fulfills these requirements, and will therefore be the updating rule we use for this model throughout this article. According to this updating rule, in every time step two individuals \(i\) and \(j\) are chosen at random from the population. Fitnesses \(f_i\) and \(f_j\) are calculated. With probability 

\[
\frac{f_j}{f_j + f_i},
\]

\(i\) is replaced by a copy of \(j\) and in the other case, with probability 

\[
\frac{f_i}{f_j + f_i},
\]

\(j\) is replaced by a copy of \(i\).

However, the payoffs from the Public Goods Game can be negative. We transform payoffs \(\pi_s\) from the Public Goods Game into fitnesses by an exponential function. The resulting replacement probabilities are

\[
p(i \rightarrow j) = p(i \text{ replaced by a copy of } j) = \frac{1}{1 + \exp(w(\pi_i - \pi_j))}.
\]

The constant \(w > 0\) measures the selection strength. This is also the replacement probability used in the Fermi process updating rule [1]. However, in the Fermi process, replacement is not guaranteed: In the other case, with a probability of 

\[
1 - p(i \rightarrow j),
\]

no replacement happens in the time step.

As behavioural experiments have shown that there is a strong tendency to explore strategy space in a given social context [41, 25], we assume mutation rates \(\mu \gg 0\).

We first consider an analytic approximation of his model. The replicator-mutator equation [47, 7, 33] is a well-studied ordinary differential equation. Here we follow the derivation of Traulsen and coworkers [47, 48]. For frequencies \(x_s = n_s/N\) \((s = 1, \ldots, d)\), the probability density \(\rho(x)\) of strategies can be approximated by a Fokker-Planck equation [24, 47, 48, 46]

\[
\dot{\rho}(x) = -\sum_s \frac{\partial}{\partial x_s} \rho(x) a_s(x) + \frac{1}{2} \sum_{s,t} \frac{\partial^2}{\partial x_s \partial x_t} \rho(x) b_{st}(x),
\]

where the drift-vector terms

\[
a_s = \sum_t x_t x_s (1 - \mu) \left[ p(t \rightarrow s) - p(s \rightarrow t) \right] + \frac{\mu}{d - 1} (x_t - x_s)
\]

express the evolutionary dynamics of the system and the diffusion matrix \(b\) represents the noise due to mutation. It is known that for large \(N\), the dynamics of the system converge towards the replicator-mutator equation \(x_i(t) = a_i(x(t))\) [47].

Now, using the transition probabilities given by the Link Dynamics updating rule, eq. (9), we obtain the replicator-mutator equation as

\[
x_i = a_i(x)
\]

\[
= \sum_t x_t x_i (1 - \mu) \left[ 1 - \frac{2}{1 + \exp(\beta(\pi_s - \pi_i))} \right] + \frac{\mu}{d - 1} (x_t - x_i),
\]

where the \(1 - \frac{2}{1 + \exp(\beta(\pi_s - \pi_i))}\) term represents the updating rule.
For low mutation rates \( \mu \), the model has one stable equilibrium \( X \) in the defection corner, very near to the pure defection state. Trajectories may pass close to the pure punishment or the pure cooperation state, but all trajectories will ultimately converge towards \( X \). The underlying vector field, and examples of trajectories in the state space of the replicator-mutator approximation for low \( \mu \), can be seen in Fig. 2a.

With increasing \( \mu \) the model gains a second attractor \( Y \) near the cooperator corner, very close to the punisher-cooperator line (c.f. Fig. 2b). \( R_Y \), the basin of attraction of \( Y \), consists of the punisher corner, while the basin of attraction of \( X \), \( R_X \), is reduced to a smaller including the whole defector-cooperator line. The separatrix between \( R_X \) and \( R_Y \) starts at the unstable equilibrium between punishers and defectors, splitting the state space roughly parallel to the defector-cooperator line until it ends on the punisher-cooperator line very close to \( Y \).

For very large mutation rates \( \mu \) (Fig. 2c), the system loses the original high-defection attractor \( X \) and only the highly cooperative equilibrium \( Y \), now moved significantly away from the punisher-cooperator line due to noise from mutation, remains. The separatrix has essentially shifted outside the state space, all trajectories now converge towards \( Y \).

The observations reported here for the behaviour of the replicator-mutator equation are consistent with the observations given by Traulsen et al. [48] for strong selection and error-free strategies (\( w \to \infty, \epsilon = 0 \)).

The replicator-mutator equation gives insight into the stable states of the evolutionary process. To understand the full dynamics of the system, it is however necessary to understand the long-term fluctuations of the process. This requires a study of the stochastic model. It can be described as a time-independent discrete Markov chain. When mutation rates are non-zero, this chain is both irreducible and aperiodic, and therefore has a unique equilibrium distribution.

We consider this equilibrium distribution for the same parameter values as the replicator-mutator equation. For example, the Markov transition matrix is small, containing \( \frac{(N+1)\times(N+2)}{2} \) states with each row having at most 7 entries, the equilibrium distribution can be easily found using standard methods of sparse numerical linear algebra.

The behaviour shown for low mutation rates is significantly different from what might be expected, given the results of the replicator-mutator approximation. For low and medium mutation rates \( \mu \), pure defection is a distinct mode of the equilibrium distribution, even though the \( R_X \) can be significantly smaller than \( R_Y \), when both exist. As Figs. 3a, 3b show, all other states in the state space are less probable than this mode by at least an order of magnitude. The other pure strategy states, pure cooperators and pure punishers, are local modes. If the Markov chain temporarily fixes in a pure strategy state, it will not leave that state until a mutation occurs, which is generally less likely than other state transitions. For the same reason, states in the interior of the state space have a much lower probability of presence than states on the boundary. However, the probability of the pure punishers state is larger than the probability of pure cooperators, even though the equilibrium \( Y \) of the approximation contains more cooperators than punishers.
For high mutation rates, the picture changes. If the probability of mutations is of the same order of magnitude as other state transitions, the long-term evolutionary dynamics of the system are no longer dominated by fixation and extinction of strategies, but by the attractors of the evolutionary system. Consequently, high mutation rates lead to higher probabilities for the stable equilibria and lower probabilities for the evolutionarily unstable states in the system. This can be seen in Fig. 3. The Markov states near the only stable equilibrium $Y$ of the replicator-mutator approximation for high $\mu$ have a higher probability, with the state corresponding to $Y$ being a mode. States on the punisher-defector line, where an unstable equilibrium is located, have relatively low probability.

![Figure 4](image1.png)

Figure 4: Frequencies of strategies for varying $\mu$, for $x = 0$. The lines show the means of the Markov stable state distributions, areas $\pm 1\sigma$ around the mean are shaded. For low mutation rates, the system is nearly exclusively defective. For high mutation rates, cooperators and punishers form a majority, before cooperators fall again as $\mu \to 1$, where relative frequencies approach $1/3$ each.

The mean frequency of strategies in the equilibrium distribution behaves as might be expected from the previous observations of the replicator-mutator approximation and the stochastic model. As shown in Fig.

![Figure 5](image2.png)

Figure 5: Frequencies of strategies for varying punishment cost $\gamma$, for $x = 0$. Lines show the means of the Markov stable state distributions, areas $\pm 1\sigma$ around the mean are shaded. While costly punishment permits only defection to prevail, for low values of $\gamma$ the evolutionary dynamics are dominated by punishers, and for negative values of $\gamma$ defection rates are very low.

Punishment was therefore made even less effective in an environment with many defectors. While the human strategy of punishing even under detrimental circumstances \[14\] \[5\] \[49\] seems currently to be stable, the question of how this behaviour has evolved remains open. It is possible that in order to evolve stable costly punishment, intermediate states would be necessary. In these intermediate states, policing would have a direct or indirect fitness benefit.

In fact, it has been argued \[31\] \[36\] that stable, self-organised institutions with cooperation enforced through policing by peers do generally require the incorporation of some type of explicit reward for the police. This countervails, and may even replace, the high punishment costs typically encountered in simple versions of the PGG. These explicit rewards – *bounties*, in common parlance – are not restricted to humans. Punishment options in other species may also involve benefits to the individual punisher, or even to the whole population. An example of such a direct benefit can be found in eusocial hymenoptera. Male eggs laid not by the queen, but by other workers \[16\] \[30\], are eaten by other specialist workers. By doing so, they not only punish the miscreants, but also thereby gain valuable protein.

2.1. Bounties and cheap punishment

In the original Fehr-Gächter experiment \[14\], punishment was deliberately designed never to be in the material self-interest of the experiment subjects. Furthermore, the fine $\beta$ was not constant, but proportional to the Public Goods Game payoff.
Bounties and other direct benefits for policing agents are easily incorporated in the Public Goods Game as described in eq. (1) by allowing the punishment cost parameter $\gamma$ to be a small positive or even negative number. It is obvious that such a change will increase the fitness of precisely the peer punishers, thus *ceteris paribus* increasing the amount of punishment in the system. This effect can indeed be seen in the model, as Fig. 5 shows. Costly punishment permits only defection to prevail.

For low values of $\gamma$, the mean frequency of punishers increases. Because cooperators dominate punishers but are in turn dominated by defectors, this also leads to an increase of cooperators. This increase is however not as steep as the increase in punishers.

For punishment cost $\gamma < 0$, punishers have an evolutionary advantage over cooperators and over defectors. Consequently the dynamics of the system are dominated by punishers, and defection rates are very low. Even though the number of individuals generally following a defection strategy $n_D$ is negligible, due to erroneous strategies the effective amount of defectors $n_D = \epsilon N$ is still positive in that case. This means that even in that case, there are direct fitness benefits to punishers, and those benefits are larger the lower $\gamma$ becomes. Therefore the proportion of cooperators in the population decreases as $\gamma$ decreases.

This explains the effect of direct benefits of punishment to the punisher. Deterring of defectors may be an indirect benefit of punishment, but more immediate indirect benefits of the punishment process require a change in the payoff structure of the model. We will explore such an extension in §3.2.

### 3. Modifications to the basic Peer Punishment Model

The Public Goods Game is one of the simplest paradigms exhibiting $N$-player cooperation dilemmas, and as such is similar to the Prisoners’ Dilemma and the Hawk-Dove game for two players. Cooperation is however much more frequent in the real world than these simple models would predict [59]. The prevalence of cooperation can be seen in the real-world in all scales, from bacterial microfilms [35], via state-forming insects [8, 16, 30, 38, 50], all the way to human society. To describe this, both experimentalists and modellers have tried to extend the Public Goods Game and make it more applicable as a model for real-world interactions [17, 9, 36, 4]. In this section we will consider three new extensions to the mathematical model inspired by real-life observations, focusing on the implementation of punishment in the PGG.

Fowler [17] and other workers [48, 13, 23] have observed that cooperation within a Public Goods Game context is increased if at each step player participation is no longer compulsory. If in the Public Goods Game players can opt out (following a *loner* (L) strategy), gaining a small fixed payoff $0 < \sigma < (r-1)c$ instead, the evolutionary system does allow the evolution of cooperative states of the population from non-cooperative states. However, the cyclic dominance of cooperators over loners, loners over defectors, and defectors over cooperators means that cooperation is not an evolutionarily stable state (ESS) of such a system. Furthermore, the reliance on fundamentally unsocial strategies to explain social behaviour is conceptually unsatisfactory.

In some instances, such a fixed payoff $\sigma$ can be used to represent participation in a less risky, separate game. Instead of such an indirect representation, the population can be explicitly divided into separate, weakly connected subpopulations.
This will lead to indirect competition between subgroups in the population, and can therefore be expected to increase levels of cooperation in the model. This will be the first extension we study in the following section, followed by two other extensions of this basic Public Goods Game that could be reasonably assumed to change the equilibrium distribution of the Markov chain towards more cooperation. We will see that population structure has a large effect, both on its own and in conjunction with the other extensions.

3.1. Island-Model Local Structure with Constant Population Size

![Figure 8: Frequencies of strategies dependent on the expected number of cross-steps](image)

Figure 8: Frequencies of strategies dependent on the expected number of cross-subpopulation updating steps \(x\). Lines show the average relative frequencies of strategies from numerical simulations with \(10^6\) time steps, the areas corresponding to \(\pm 1\sigma\) around the means are shaded. The values for \(x = 0\) correspond to the basic model described in [12]. While low migration rates increase defection, increasing \(x\) further leads to an increase in both cooperators and, with a delay, but going on for longer, punishers. For high values of \(x\), the frequencies of cooperative strategies decrease again, converging towards approximately a 2 : 1 split of the total population between defectors and punishers.

In this subsection we investigate the effect of a metapopulation consisting of \(M\) well-mixed subpopulations that are weakly linked to each other. The model works as follows. There are \(M\) subpopulations of \(N\) individuals each, yielding a total population of \(MN\) individuals. The individuals in each sub-population participate in a Public Goods Game with punishment as described above. For each time step, in every sub-population one individual is replaced by a copy of a different individual according to the Link Dynamics updating rule given in eq. (9). In addition, however, after all subpopulations have been updated, additional Link Dynamics cross updating steps between any two random individuals from the total population can occur. Fig. 7 shows an example with one of these cross-steps per time step. This means that a random individual \(i\) from the total population compares its payoff from the local game with an individual \(j\) from anywhere in the total population. One of them replaces the other with probability \(p(i \rightarrow j)\) as given in eq. (9).

In order to approximate random migration, the number of such cross steps per time step will be Poisson distributed, with mean \(x \geq 0\). After each cross step, the payoffs in the population that has gained a new member are recalculated.

Boyd et al. [10] consider a similar model using group selection in which, by contrast, whole subpopulations compete with and replace each other. They find that for small groups (roughly less than 32 individuals), high levels of cooperation can be maintained. Here we find that even when only single individuals are copied between subpopulations, instead of entire subpopulations replacing each other, high cooperation levels can be maintained in a significant part of the phase space.

We find the following results for the PGG/P in this weakly coupled system of subpopulations. For \(x = 0\), no cross steps occur, so the basic Public Goods Game in \(M\) separate parallel instances is recovered. The model therefore shows pure defection with an admixture of noise due to mutation as the only ESS, as can be seen in Fig. 12a.

We now explore now the case of \(x > 0\). Mutations will lead to random low fluctuations in individual populations. Consider a sub-population that contains a mixture of \(n_C\) cooperators and \(n_D\) defectors only, due to such fluctuations. According to eq. (1), the payoffs of these strategies are \(\pi_C = rc\frac{N}{2} - c\) and \(\pi_D = rc\frac{N}{2}\) respectively. When a cooperator and a defector within this population are selected in a Link Dynamics updating step, the probability that the cooperator dies and is replaced by a defector is therefore, according to eq. (9),

\[
p(C \rightarrow D) \approx \frac{1}{1 + \exp(cw)} < \frac{1}{2} \quad .
\]

This is different if such a cooperator instead competes with a defector \(D'\) from a sub-population containing only defectors. If \(\pi_C > \pi_D = 0\), i.e. in case \(n_C > N/r\), the cooperator has an advantage over the defector. On the other hand, even for \(n_C \leq N/r\) such a cross-step can be beneficial to cooperation. Because \(p(C \rightarrow D') < p(C \rightarrow D)\), the chance of losing a cooperator in such a cross-step is less than it is for a within-population updating step. Furthermore, if

\[
p(D' \rightarrow C) \approx \frac{1}{1 + \exp(wc[1 - \frac{\pi_D}{\pi_C}]1)} > \mu \quad .
\]

such a cross-updating step is able to introduce cooperators to pure-defector subpopulations with higher probability than mutation. The condition given in eq. (14) holds generally unless mutation rates are particularly high, in which case cooperation is prevalent anyway (as observed in the previous chapter, cf. Fig. 4), and unless selection is especially strong.

It is therefore reasonable to assume that, for some regions of the parameter space, the increase in fluctuations due to the cross-steps should counterbalance the dominance of defectors within subpopulations. This is indeed the case. Figure 8 shows the mean frequency of the three different strategies during simulation runs, for varying values of \(x\).

When cross-steps are rare, the local evolutionary dynamics are dominated by the local behaviour of the system. The addition of a very low number of cross-steps effects that defectors are re-introduced to subpopulations that may have lost
Parameter | Symbol | Default value
--- | --- | ---
Contribution | $c$ | 1
Selection strength | $w$ | 1
Linear scaling of public good | $r$ | 3
Sub-population size | $N$ | 10
Sub-population count | $M$ | 20
Fine | $\beta$ | 16
Punishment cost | $\gamma$ | 5
Mutation rate | $\mu$ | 0.06
Error rate | $\epsilon$ | 0.01
Mean of cross steps | $x$ | 10
Tax payment | $t$ | $1 - 2\epsilon$ for $\epsilon \gamma \approx 0.025$

Object | Symbol
--- | ---
Effective absolute frequency of strategy $s$ | $n_s$
Absolute frequency of players generally following strategy $s$ | $\bar{n}_s$
Payoff for players realising strategy $s$ | $\pi_s$
Mean payoff for players generally following strategy $s$ | $\bar{\pi}_s$
Proportion of public good paid to an individual | $\Pi_s$, $\bar{\Pi}_s$
Defector strategy | D, red
Cooperator strategy | C, black
Punisher strategy | P, green
Monitor strategy | M, bright green
Taxpayer strategy | T, blue

Table 1: Symbols used throughout this article

Figure 7: Caricature of the evolutionary dynamics in a structured population, showing link dynamics steps ($\leftrightarrow$) between defectors (red), cooperators (black) and punishers (green) over time. In structured populations, payoff-dependent replacement of individuals (curved arrows), including rare cross-steps, lead to an evolutionary dynamic locally favouring defection. When fitnesses are compared across subpopulations, however, even for strong selection punishers can replace cooperators and cooperators can replace unpunished defectors, which is not possible in the base game. This allows islands of cooperative players to prevail in the long run.

On the other hand, such cooperative subpopulations provide room for punishers to evolve, with only minor disadvantage to the cooperators due to random $\epsilon$-errors in the realisation of strategies. Should punishers fixate in a sub-population, that sub-population cannot be easily invaded by defectors. The mean frequency of punishers therefore rises even further than the mean frequency of cooperators.

In these simulations, both a typical run (Fig. 9b) and the general trend (Fig. 12b) show punishers to be much more common than cooperators. The explanation for this is that a sub-population of cooperators can easily be invaded by defectors. By contrast a sub-population containing a sufficiently large proportion of punishers is able to repulse a defector invasion. While a high local frequency of cooperators gains high fitness and can therefore invade a defector-rich sub-population in the cross-steps, the converse is false.

When $x$ becomes sufficiently large that the local group composition changes significantly (or in the extreme case, completely) from one updating step to the next, the situation changes. Now we have two apparently contradictory intuitions. On the one hand, one might say that the effect discussed in the small
Figure 9: Strategy frequencies over for the first 10,000 time steps in the total population for typical Monte Carlo simulation runs, with parameters according to Table 1. An increase in punishers is generally followed by an increase in cooperators, which in turn leads to an increase in defectors.

The case would be magnified, leading to an even larger degree of fluctuations and therefore cooperation to prevail, just as in the case of high mutation for the basic model (Figs. 2, 3). On the other hand, the total population may now be regarded as having lost its local structure, and may be treated as a well-mixed population of size $MN$. In this case, defectors are expected to prevail by the usual PGG/P argument.

Neither of these two pictures is entirely correct, as can be seen in Fig. 8. For high values of $x$, the frequency of cooperators decreases again, because high mixing leads to both cooperators and defectors present in subpopulations, such that on average defectors have a significant advantage. Punishers are also generally rarer than defectors. The average number of punishers is nonetheless far bigger than it would be in a well-mixed population, because even for large $x$ defectors cannot easily invade a punisher-dominated sub-population.

Figure 10: Frequencies of strategies for varying punishment cost $\gamma$, for $x = 10$. Lines show the average frequencies of strategies per sub-population from numerical simulations with $10^6$ time steps, the areas corresponding to $\pm 1\sigma$ around the mean are shaded. While costly punishment permits only defection to prevail, for low values of $\gamma$ the evolutionary dynamics are dominated by punishers, and for negative values of $\gamma$ defection rates are very low. In comparison with Fig. 5 cooperators have a higher peak frequency in this model with structured populations, and out-compete punishers even for low punishment costs.
For the remainder, we will generally consider the case of $x = 10$ when studying the interaction of a structured population with other modifications. First we will consider how such a structured population interacts with different punishment costs.

For low punishment costs, fluctuations play a significant role again. For the basic model we had found (Fig. 5) that low and negative punishment costs $\gamma$ lead to a steep increase in punishers. This is true for the structured model, as well. This can be seen in Fig. 10. In addition to this the structured model also shows an increase in the cooperator frequency for low $\gamma$. This can be easily explained by the fact that the basic model will remain for an expected time of $\mu^{-1}$ in the pure punisher state, while the cross-steps in the structured population model allow cooperators to spread into punishment-dominated subpopulations much faster.

When punishment costs become negligible, due to $\gamma$ and therefore also $n_D$ becoming low, both the cooperators and punishers increase more slowly. For negative punishment costs, punishers again have an advantage over cooperators, just as in the basic model.

Naïvely, high values of $r$ might be considered beneficial to cooperation, because for $r > N$, an individual’s payoff is higher if it contributes to the public good than if it does not. However, this is not the case. In the basic model, the public good was an additive constant for all individuals in the population. It did therefore not influence the replacement probabilities at all. However, when subgroups are in competition with each other, the fitness different individuals gain from the public good depends on the composition of their sub-population. This means that while the parameter $r$ was irrelevant for the basic model (cf. the inset in Fig. 11), behaviour in this extended model depends on the value of $r$. The effect is however counter-intuitive, because the higher the linear scaling $r$ is, the more is defection favoured in the model. Only for very small values of the scaling factor does the model favour cooperation.

While this model does not incorporate variable population size, some of the effects seen here are reminiscent of the statistical Simpson’s paradox [42]. We will discuss this further in §4

![Figure 11: Frequencies of strategies under variation of the public good factors $r$. Lines show the average frequencies of strategies per sub-population from numerical simulations with $10^6$ time steps, the areas corresponding to $\pm 1\sigma$ around the mean are shaded. Defection dominates for high values of $r$, while cooperative strategies are favoured only when investments into the public good are only marginally beneficial. The inset shows the equivalent graph for the well-mixed model: frequencies are independent of $r$.](image)

![Figure 12: Relative frequencies of strategies, without and with cross-steps, as generated by numerical simulations with $10^6$ time steps. The mean frequencies are marked with a blue cross. Mixed states are much more frequent for the model with migration, leading to less defection in the mean. The simulated results are indistinguishable from the numerically computed eigenvector of the Markov chain, Fig. 3a.](image)
### 3.2. Replacing Peer Punishers with Monitors

In the previous subsection, we have seen that high levels of cooperation can only develop when the interactions between subpopulations happen somewhat rarely. If punishers were to increase the fitness of a population compared to cooperators, a structured population might lead to punishment being stable for a wider range of parameters. This corresponds to a change in the payoff structure of the model. We will in the following study such a modification, which turns punishers into what we will refer to as monitors (M).

As we have seen in §2.1, low punishment costs can allow the cooperative strategies to prevail. As we found in the previous section (and can be seen in Fig. 5), this holds true even more for the case of structured populations. In addition to bounty-style direct benefits to the policers, punishment may also confer indirect benefits through forcing the culprit to con-
tribute to the public good. Prisoners, for example, are a notorious source of cheap labour. This effect is however not restricted to Western human societies. Guala [20] cites the example of a hunter who has to give up his catch, which he gained by breaking the social norms of the hunt, to the whole group. It has also been argued that policing is more likely to increase the efficiency of a colony of eusocial insects than to decrease it [38].

We here introduce a model in which players can monitor other players’ behaviours, and force them to contribute to the public good if it appears that they would otherwise defect. This is the modification we will focus on below.

For this new variant of the Public Goods Game, which we will refer to using the shorthand PGG/M, the payoffs for the different pure strategies are as follows.

\[
\Pi_M = rc \frac{n_C + n_M + n_D n_M}{N}, \quad (16)
\]
\[
\pi_C = \Pi_M - c, \quad (17)
\]
\[
\pi_D = \Pi_M - \beta n_M \frac{n_M}{N} - c n_M, \quad (18)
\]
\[
\pi_M = \Pi_M - c - \gamma n_D \frac{n_D}{N} \quad . \quad (19)
\]

Just as with the PGG/P, individuals do realise noisy versions of the pure strategies, and just as punishers in the basic model, monitors will realise the M pure strategy with probability \((1 - 2\epsilon)\), and defect and cooperate with probability \(\epsilon\) each.

Eq. 16 strongly resembles the base payoff structure for the PGG/P (eq. (1)), but there are two principal differences. The first difference concerns the fine for defection, which is effectively increased. Defectors in the PGG/P have to pay a fine of \(\beta n_M/N\) per punisher, while in the Public Goods Game with monitors, the effective fine is slightly higher at \(\beta' = \beta + c\). Algebraically, this difference could thus be easily absorbed by a reparametrisation, but from a game theoretical point of view, this reparametrisation would be misleading. The comparison between the two models would become unclear, because the interpretation of the terms is different. In both cases, \(\beta\) models a plain fitness penalty and the corresponding good is “destroyed”. Only in PGG/M, however, is there an investment (of magnitude \(c n_M/N\)) in the public good. The second difference concerns precisely this reinvestment. Whereas the public good in the basic PGG/P game amounts to \(\Pi = rc(n_C + n_P)/N\), it is increased for the PGG/M from \(\Pi\) to \(\Pi_M = \Pi + \frac{c n_M}{N}\).

Figure 15: Equilibria, trajectories and the underlying vector field of the replicator-mutator approximation for two variants of the basic Public Goods Game. If policing involves forcing defectors to contribute to the public good (“monitors”), the vector field and the shape of the basins of attraction change slightly. In particular, \(Y\) moves away from the basin of attraction of \(X\).
(a) Frequencies of strategies for varying punishment cost $\gamma$, for the monitor model in a well-mixed population ($x = 0$). Lines show the means of the Markov stable state distributions, areas $\pm \sigma$ around the mean are shaded. Very similarly to the basic model (Fig. 5), costly punishment permits only defection to prevail, whereas for low values of $\gamma$ the evolutionary dynamics are dominated by punishers, and for negative values of $\gamma$ defection rates are very low.

(b) Gain in the expected equilibrium frequency for the PGG/M, compared to the PGG/P. For negative and high punishment cost, the two models are very close. For medium punishment costs, around $2 < \gamma < 7$, however, monitoring is deleterious to defectors.

Figure 16: Frequencies of strategies for varying $\gamma$, for the PGG/M.

We find that in a single well-mixed population, the effect of this modification is very minor, both in the replicator-mutator equation and in the equilibrium distribution generated by the Markov chain. The two equilibrium distributions generated by the the Markov chains for the PGG/P and the PGG/M are shown in Fig. 18. The distributions are very similar, with only a minor difference in the mean.

When varying the punishment cost $\gamma$, the qualitative behaviour of the system is the same as for the basic model. Quantitatively the results change by a minor amount. For positive punishment costs, replacing punishers with monitors leads to a very small decrease in defection. The change due to the introduction of monitors has a single maximum in the parameter range where the system changes from being dominated by defectors to being dominated by cooperative strategies.

Figure 17: Comparison of frequency of Defectors depending on the expected number $x$ of cross-steps in games where policing is done by punishers (red) vs. monitors (orange). While for a low number of cross-steps monitoring is worse for cooperators than punishment, for high $x$ defectors fare slightly better when monitored than when just punished.

Figure 18: Stable state distribution of the Markov chain for the evolutionary Public Goods Game with monitors (PGG/M) in a well-mixed population ($x = 0$) for default parameters. The mean of this distribution is marked with a black $\times$. This distribution is very similar to the corresponding distribution for the PGG/P (Fig. 5), but that distribution contains slightly more defectors in the mean (blue $\pm$).
A fundamental problem of the concept of peer punishment is the fact that the punishment of defectors is in itself a public good \[52\]. This means that non-punishing cooperators are themselves defectors of this second-level game. Various other suggestions have been made for avoiding this second-level free rider dilemma \[49, 13\]. Of particular interest is the observation by Andreoni and Gee \[3\] that, given the opportunity, human experimental subjects prefer to delegate their punishment interests to external third parties instead of punishing themselves. This is not unexpected. The expected cost for an individual delegating the punishment to someone else should be lower than for a punisher, due to a lower risk of retribution from the punishment. On the other hand, we expect there to be a threshold in fines, above which more punishment will not lead to better behaviour. Therefore, a low number of punishers bearing the full punishment cost but gaining recompensation from other players appears \textit{a priori} a reasonable ansatz for higher stability of cooperation. Both experiments \[3, 2, 4, 49\] and models \[43\] have shown that pool punishment is a viable strategy, however, in none of these the punishers that obtain the taxes are modelled in none of these. This means that non-punishing cooperators are themselves defectors of this second-level game. Various other suggestions have been made for avoiding this second-level free rider dilemma \[49, 13\]. Of particular interest is the observation by Andreoni and Gee \[3\] that, given the opportunity, human experimental subjects prefer to delegate their punishment interests to external third parties instead of punishing themselves. This is not unexpected. The expected cost for an individual delegating the punishment to someone else should be lower than for a punisher, due to a lower risk of retribution from the punishment. On the other hand, we expect there to be a threshold in fines, above which more punishment will not lead to better behaviour. Therefore, a low number of punishers bearing the full punishment cost but gaining recompensation from other players appears \textit{a priori} a reasonable ansatz for higher stability of cooperation. Both experiments \[3, 2, 4, 49\] and models \[43\] have shown that pool punishment is a viable strategy, however, in none of these the punishers that obtain the taxes are modelled and able to co-evolve with the delegators.

A necessary condition for stability in such a taxpayer model would be the requirement that a stable mixed equilibrium between taxpayers and punishers exists. This equilibrium should not only be stable in the absence of defectors, but should also be resistant against invasion by defectors. For this extension, we modify eq. \[1\] to allow these conditions to be fulfilled. Consider

\[ \Pi_T = rN \frac{n_T + n_p}{N} \]
\[ \pi_D = \Pi_T - p_N \]
\[ \pi_T = \Pi_T - c - t \]
\[ \pi_p = \Pi_T - c - \gamma n_D + \frac{n_T}{n_p} \]

Given that strategies are followed with an error rate \( \epsilon \), the equilibrium condition \( \pi_p = \pi_T \) translates to

\[ \epsilon \frac{N}{2} \left( \frac{n_T}{n_p} + \frac{n_p}{n_T} - 1 \right) \]
\[ \bar{\pi}_P = (1 - 2\epsilon)\pi_P + \epsilon\pi_T + \epsilon\pi_D = \bar{\pi}_T = (1 - \epsilon)\pi_T + \epsilon\pi_D . \]  
(24)

\[ \frac{\epsilon \bar{\pi}_D}{N} - t = \frac{\epsilon \bar{\pi}_T}{N} = t . \]  
(25)

If we choose this equilibrium to be at \( \bar{n}_P = \bar{n}_T \), we obtain the condition

\[ t = \frac{1 - 2\epsilon}{2 - 3\epsilon} \epsilon \gamma . \]  
(26)

In order to possibly be a stable equilibrium in the state space, \( \bar{n}_P = \bar{n}_T = \frac{N}{2} \) needs to fulfill two further conditions. Firstly, the equilibrium needs to be stable on the taxpayer-punisher line. This means that we have \( \bar{\pi}_T > \bar{\pi}_D \) for \( \bar{n}_T < \bar{n}_P \) and vice versa. This is the case in eq. (22) and eq. (23). Secondly, a single mutant defector must have lower fitness than the punishers and taxpayers near the equilibrium. We have

\[ 0 < \bar{\pi}_T - \bar{\pi}_D \]  
(27)

\[ 0 < \beta(1 - \epsilon) \frac{\bar{n}_P}{N} - c - \frac{1 - 2\epsilon}{2 - 3\epsilon} \epsilon \gamma \]  
(28)

which for large \( N \) and small \( \epsilon \) reduces to \( \beta > 2c \), which should be fulfilled in most cases. The replicator-mutator approximation of the Public Goods Game variant described in this section should therefore have an attractor near the middle of the taxpayer-punisher line, slightly offset from that exact location due to the inclusion of mutation.

The vector field that underlies the replicator-mutator approximation of the model for low mutation rates, shown in Fig. 21a, does indeed indicate that trajectories will go towards a roughly equal number of punishers and cooperators when started anywhere in that basin of attraction \( R_Y \). For higher mutation rates, the equilibrium \( Y \) does however move towards a higher proportion of taxpayers, as can be seen in Fig. 21b.

We have however seen that the replicator-mutator approximation is in some cases not very good, and will therefore compare the equilibrium states of the Markov chain. In the equilibrium distribution for a well-mixed population, the effects are minimal. Fig. 22a shows the steady state generated by the Markov chain for well-mixed populations. It is very similar to the corresponding distribution for the basic model, but the mean is marginally more cooperative. For large \( N \) (Fig. 22b), the same holds, but the cooperator-taxpayer axis contains a clear local mode near the middle.

\[ 10^5 \quad 10^4 \quad 10^3 \quad 10^2 \quad 10^1 \]  
(a) In a well-mixed population (\( x = 0 \))

\[ 10^5 \quad 10^4 \quad 10^3 \quad 10^2 \quad 10^1 \]  
(b) In a structured population (\( x = 10, M = 20 \)), as generated by a Monte Carlo simulation over \( 10^6 \) time steps.

Figure 22: Equilibrium distribution of the Markov chain for the PGG/P with mandatory fitness transfer from cooperators (“taxpayers”) to punishers. The mean of this distribution is marked with a black +. In both cases, the distribution shown here is very similar to the corresponding distribution for the PGG/P (Fig. 3 and 12b), but those distributions contain slightly more defectors in the mean (blue +).
Figure 23: Stable state distribution of the Markov chain of the evolutionary game in a well-mixed larger population ($x = 0, N = 25$). The mean of this each distribution is marked with a blue +. While the probability of presence for the states on and near the punisher-taxpayer line go up by an order of magnitude compared to the punisher-cooperator line, this has hardly any effect on the mean.

While the previous section has shown that in the presence of indirect group competition, monitoring is detrimental to cooperation levels, even though only marginally, this is not the case for tax payment. For the simulations of the taxpayer model in structured populations, a similar picture emerges as for well-mixed populations. The probability distributions are still very similar, and the mean of the distribution is changed towards more cooperation. For the population structure model, the gain is however not as high as for the well-mixed population for the given value of $t$.

Under varying tax values $t$, the qualitative behaviour of the well-mixed and the structured model are similar. For low $t$, defection constitutes the majority in the system. As higher $t$ lead to an increase in punishers and a decrease in defection, taxpayers have a maximum in the expected frequency. For high $t$, the system settles into a phase in which punishers are frequent, taxpayer are a small minority and defection rates are very low. The major difference between well-mixed and structured populations is that in the well-mixed model, the baseline frequency of taxpayers is lower than in the metapopulation model.

We have so far assumed that tax payment is obligatory. It would however be natural to assume that tax payment is optional unless enforced. Therefore we now consider what happens if both taxpayers and cooperators are strategies in the game that compete with each other. The payoffs change in the obvious way. The payoffs for the different strategies are unchanged as given by eqs. (21)–(23) and eq. (3). The public good, however, is now given by

$$\Pi = rc(n_C + n_T + n_P)/N,$$

and punishers never accidentally realise the taxpayer strategy. The cooperation thus introduced is detrimental to the success of cooperative strategies. As Fig. 24a shows, the 4-strategy game with defectors, cooperators, taxpayers and punishers is less cooperative than the basic game, both for well-mixed and for structured populations. This also holds for other values of $t$, as Fig. 25b shows. For well-mixed populations, the presence of tax-payers paying high taxes at their own serious detriment is beneficial to the policing individuals, and the frequency of defectors diminishes slightly. For a metapopulation the same can be said, but the positive effect for cooperation of some individuals paying high taxes is much more pronounced.
Only for very high values of $t$ is the optional tax payment model with both taxpayers and cooperators actually more beneficial for cooperation than the basic model.

4. Discussion

In this article, we have studied modifications of the linear Public Goods Game with punishment. We have tried to understand the apparent preference of humans for police, where a minority of dedicated individuals, supported by the non-acting majority, seek out defectors and force them to rethink their bad behaviour. We found that two modifications of the payoff structure, which try to capture such behaviour, have only marginal benefits for cooperation.

For well-mixed populations, both forcing defectors to contribute to the public good by monitoring their behaviour and allowing individuals to pay (non-enforced) taxes to the policers have slight positive effects on the levels of cooperation. These effects, however marginal, are generally strongest for those parameter ranges where the system is in a transition from one regime to another, i.e. at the phase transitions.

On the other hand, we have shown – continuing previous investigations into the structural constraints in evolutionary models [29] – that mutation and migration can play an important role for the presence of cooperation. The fact that random mutations lead to the emergence of cooperative stable states had been shown before [48], and has been found again here in a slightly different model. There is a wide scope for general model assumptions to influence the results of the evolutionary model. It is known [18] that the structure of mutations influences the dynamics of an evolutionary system. Similarly, the structured population presented here incorporates migration, but the specific structure of migration is important for the results.

Using a variant of migration of offspring, where the success of migration depends on the fitness according to the parent population of the individual migrating, we have implemented a model where cooperation can prevail through group competition, but selection acts only on individuals. We have observed that the migrating cross-steps lead to non-trivial interactions between the subgroups. Among other things we have seen that fitness-dependent migration steps can reverse a trend compared to well-mixed populations.

It is a well known paradox in statistics that a trend that occurs in different groups of data can turn to the opposite when the groups are combined. This effect is known as Simpson’s Paradox [29]. The terminology was introduced to the literature on the evolution of cooperation by Sober et al. [42]. The argument is as follows. Given a set of subpopulations, defection is a dominant strategy in each of them. The relative frequency of defectors, averaged over all subpopulations, will therefore be high. If, on the other hand, subpopulations with high amounts of cooperation have a higher population capacity and contain more individuals [15, 36, 42, 23], the relative frequency of defectors in the whole metapopulation will be large. The paradox is contained in the following statement: Defection dominates cooperation in every single sub-population, but in the metapopulation, cooperation prevails.
In the model presented here, the rise in cooperation does not stem from different sub-population sizes, as is the case in previous models, but from the fact that copying steps between these subpopulations are biased. While the results presented in this part are consistent with the ideas presented in the literature so far, the model presented here has two methodological advantages compared with models with variable sub-population size.

Firstly, the constant sub-population size reduces the parameter space of the model. If the population capacity of a subpopulation depends on the amount of public good that subpopulation produces, this dependency must be described by additional parameters of the model. If an effect can already be shown using the simpler model presented here, it is clear that such an effect is more universal than if it can only be shown for certain choices of that dependency.

Secondly, the state space of a model with constant population size is comparatively small. For a sub-population size of $N = 10$ and three different strategies, the Markov transition matrix is a sparse matrix between $\frac{(N+1)(N+2)}{2} = 66$ different states (up to symmetry), and can be easily calculated and manipulated by computational tools. In the model presented here, the interactions between all subpopulations are symmetric. For time-continuous Markov chains, there are methods to solve similar systems explicitly [6]. Using such methods, it may be possible to study the long-time limit of the behaviour of systems like the one presented here explicitly, without the need for time-extensive simulations. Models of symmetrical metapopulations with fixed sub-population size might therefore be able to shed more light on social dilemmas.

We have seen that fitness-based migration can be a proxy mechanism for group selection. In animal as well as human societies, we perceive a division of labour between a majority of cooperators that merely support a dedicated minority of police-like enforcers, which actually administer punishments. The goal of this work was to gain insight on this split of labour.

Monitoring the population for defection and preventing it, and even more so obligatory taxation, increase the stability of an equilibrium between cooperators and monitors. Assuming such an equilibrium to represent a division of labour between taxpayers and punishers, we would assume that a population in this stable state had an advantage over other populations, and that such compositions would therefore prevail in the long run. The models presented here have only begun to address this question. We do find certain contexts in which a division of labour appears to exist and favour cooperators, but the level of explanation is at best partial.

A challenge in the investigation of this type of structure is the understanding of division of labour games. Symmetric 2-player games can be easily classified, and one of those classes constitutes division of labour or anti-coordination games. Generalising this to $N$-player games is much harder.

We expect that in settings where punishment provides high indirect benefits and is coercive, but not evolutionarily unstable, the Simpson’s Dilemma effect described above could be able to sustain the second-level Public Goods Game. Different interactions between punishers and taxpayers could also be able to provide such a setting.

Further research is also necessary to understand how stable the evolutionary systems presented here are. Evidence from our Monte Carlo simulations suggests that subpopulations may show a strong cyclic behaviour.

In the model with competition between taxpayers and cooperators, we have seen that purely optional payoff transfers are often deleterious to cooperation, whereas mandatory tax payments are more promising. Future research should therefore investigate the effects of punishing tax-evaders. While punishing non-punishers, as free riders in the second-level Public Goods Game, constitutes a third-level public good, this is less an issue for tax payment, which in some sense occurs on a level between the ordinary Public Goods Game and the punishment game.

5. Conclusions

In this article, we have seen that group selection, which is known to improve cooperativity in an evolutionary system, can happen through the migration of individuals, if migration rates are fitness dependent. We have studied two natural extensions of the Public Goods Game payoff structure. Monitors, i.e. peer punishers that, in addition, force defectors to contribute to the public good, are marginally beneficial to cooperation in well-mixed populations, but deleterious in structured populations. Taxpayers, which in addition to cooperation pay a small fixed amount of payoff to punishers, have a slight benefit to cooperation both in the well-mixed case as well as with group selection. However, when taxpayers and pure cooperators are both allowed strategies, the system will show less cooperation than with only pure cooperators.

Finally we speculate as to the relevance of our study in the context of the current political climate. Human societies are of course far more complex than caricatured by any of these models. It is unclear how complex a model needs to be to capture the essence of specific aspects of human societies. In particular, our results suggest that for the provision of a public good, optional contribution can be worse than no contribution at all. In order to avoid this, future studies should focus on reward structures and the indirect effects of optional – or weakly enforced – contributions.

6. Acknowledgements

We would like to thank Guy S. Jacobs for valuable discussions and comments. PhD funding for this project was provided by EPSRC. We thank Microsoft for ongoing support.

6.1. Contributions

SJC, TJS and GAK defined the problem. GAK carried out the calculations and simulations. GAK, TJS and SJC analysed the results. GAK and TJS wrote up the results. The final paper was edited and reviewed by GAK, TJS, and SJC.
7. References

References

[1] Philipp M Altrock and Arne Traulsen. Fixation times in evolutionary games under weak selection. New J. Phys., 11(1):013012, January 2009.

[2] James Andreniou and Laura K. Gee. Gun for hire: Does delegated enforcement crowd out peer punishment in giving to public goods? Working Paper 17033, National Bureau of Economic Research, May 2011.

[3] James Andreoni and Laura K. Gee. The hired gun mechanism. Working Paper 17032, National Bureau of Economic Research, May 2011.

[4] James Andreniou and Laura K. Gee. For hire: Delegated enforcement and peer punishment in public goods provision. Journal of Public Economics, 96(11–12):1036–1046, December 2012.

[5] R. Axelrod and W. D. Hamilton. The evolution of cooperation. Science, 211(4489):1390–1396, March 1981.

[6] Yujuan Bao, Ilker N. Bozkurt, Tugrul Dayar, Xiaobai Sun, and Kishor S. Trivedi. Decompositional analysis of Kronkecker structured Markov chains. Electronic Transactions on Numerical Analysis, 31:271–294, 2008.

[7] Immanuel M. Bomze and Reinhard Bürger. Stability by mutation in evolutionary games. Games and Economic Behavior, 11(2):146–172, November 1995.

[8] Andrew F. G. Bourke. Social evolution: Community policing in insects. Current Biology, 17(13):R539–R552, July 2007.

[9] Samuel Bowles and Herbert Gintis. Homo reciprocans. The Handbook of Experimental Economics, 7(4):e35287, April 2012.

[10] Robert Boyd and Peter J. Richerson. The evolution of altruistic punishment. PNAS, 106(6):3531–3533, March 2009.

[11] Robert Boyd and Peter J. Richerson. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Eihology and Sociobiolgy, 13(3):171–195, May 1992.

[12] H. Brandt, C. Hauert, and K. Sigmund. Punishment and reputation in spatial public goods games. Proceedings of the Royal Society B: Biological Sciences, 270(1519):1099–1104, May 2003.

[13] Hannelore Brandt, Christoph Hauert, and Karl Sigmund. Punishing and abstaining for public goods. PNAS, 103(2):495–497, January 2006.

[14] Ernst Fehr and Simon Gächter. Cooperation and punishment in public goods experiments. American Economic Review, 90(4):980–994, September 2000.

[15] Ernst Fehr and Simon Gächter. Altruistic punishment in humans. Nature, 415(6868):125–128, January 2002.

[16] Kevin R. Foster and Francis L. W. Ratnieks. Convergent evolution of worker policing by egg eating in the honeybee and common wasp. Proceedings of the Royal Society of London B: Biological Sciences, 268(1463):169–174, January 2001.

[17] James H. Fowler. Altruistic punishment and the origin of cooperation. PNAS, 102(19):7047–7049, May 2005.

[18] Julián García and Arne Traulsen. The structure of mutations and the evolutionary cooperation. PLoS ONE, 7(4):e35287, April 2012.

[19] Herbert Gintis. Beyond homo economicus: evidence from experimental economics. Ecological Economics, 35(3):311–322, December 2000.

[20] Francesco Guala. Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. Behavioral and Brain Sciences, 35(01):1–15, February 2012.

[21] Garrett Hardin. The tragedy of the commons. Science, 162(3859):1243–1248, December 1968.

[22] Christoph Hautert. Spatial effects in social dilemmas. Journal of Theoretical Biology, 240(4):627–636, June 2006.

[23] Christoph Hautert, Silvia De Monte, Josef Hoffbauer, and Karl Sigmund. Volunteering as red queen mechanism for cooperation in public goods games. Science, 296(5570):1129–1132, May 2002.

[24] Dirk Helbing. Boltzmann-like and Boltzmann-Fokker-Planck equations as a foundation of behavioral models. Physica A: Statistical Mechanics and its Applications, 196(4):546–573, July 1993.

[25] Dirk Helbing, Martin Schönhof, Hans-Ulrich Stark, and Janusz A. Holyst. How individuals learn to take turns: emergence of alternating cooperation in a congestion game and the prisoner’s dilemma. Adv. Complex Syst., 08(01):87–116, March 2005.

[26] Dirk Helbing, Attila Szolnoki, Matjaž Perc, and György Szabó. Evolutionary establishment of moral and double moral standards through spatial interactions. PLoS Comp Biol, 6(4):e1000758, 2010.

[27] Dirk Helbing, Attila Szolnoki, Matjaž Perc, and György Szabó. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. New J. Phys., 12(8):083005, August 2010.

[28] G. A. Kaiping, G. S. Jacobs, S. J. Cox, and T. J. Sluckin. Nonequivalence in a congestion game and the prisoner’s dilemma. Phys. Rev. E, 90(4):042726, October 2014.

[29] Gary Malinas and John Bigelow. Simpson’s paradox. In Edward N. Zalta, editor, The Stanford Encyclopedia of Philosophy. Winter 2012 edition, 2012.

[30] B. P. Oldroyd and F. L. W. Ratnieks. Evolution of worker sterility in honey-bees (apis mellifera): how anarchistic workers evade policing by laying eggs that have low removal rates. Behav Ecol Sociobiol, 47(4):268–273, March 2000.

[31] Elinor Ostrom. Governing the commons: The evolution of institutions for collective action. Cambridge university press, 1990.

[32] Elinor Ostrom, James Walker, and Roy Gardner. Covenants with and without a sword: Self-governance is possible. American Political Science Review, 86(2):404–417, June 1992.

[33] Karen M. Page and Martin A. Nowak. Unifying evolutionary dynamics. Journal of Theoretical Biology, 219(1):93–98, November 2002.

[34] Karlisch Panchanathan and Robert Boyd. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature, 432(7016):499–502, November 2004.

[35] Alexandra S. Penn, Tim C.R. Conibeer, Richard A. Watson, Alex R. Kraajjeveld, and Jeremy S. Webb. Can Simpson’s paradox explain cooperation in pseudomonsa aeruginosa biofilms? FEMS Immunol Med Microbiol, 65(2):226–235, July 2012.

[36] Simon T Powers and Laurent Lehmann. The co-evolution of social institutions, demography, and large-scale human cooperation. Ecol Lett, 16(11):1356–1364, 2013.

[37] William H. Press and Freeman J. Dyson. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. PNAS, 109(26):10469–10473, June 2012.

[38] Francis L. W. Ratnieks. Reproductive harmony via mutual policing by workers in eusocial hymenoptera. The American Naturalist, 132(2):217–236, August 1988.

[39] Matt Ridley. The Origins of Virtue. Penguin, new ed. edition, October 2001.

[40] Alvin E Roth and John Henry Kagel. The Handbook of Experimental Economics. Princeton University Press, 1995.

[41] Dirk Semmann, Hans-Jürgen Krambeck, and Manfred Milinski. Volunteering leads to rock-paper-scissors dynamics in a public goods game. Nature, 425(6956):390–393, September 2003.

[42] Elliott Sober. Unio Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, 1998.

[43] Arne Traulsen, György Szabó, and Lilla Czakó. Competition of individual and institutional punishments in spatial public goods games. Physical Review E, 84(4):046106, October 2011.

[44] Arne Traulsen, György Szabó, and Matjaž Perc. Phase diagrams for the spatial public goods game with pool punishment. Physical Review E, 83(3):036101, March 2011.

[45] Shaolin Tan, Jinhu Lu, and Gianluca Setti. An analysis of fixation probability of a mutant on a class of weighted networks under neutral selection. In IECON 2011 – 37th Annual Conference on IEEE Industrial Electronics Society, pages 4024–4028. IEEE, 2011.

[46] Arne Traulsen, Jens Clausen, and Christoph Hauert. Coevolutionary dynamics: From finite to infinite populations. Physical Review Letters, 95(23):238701, December 2005.

[47] Arne Traulsen, Jens Christian Clausen, and Christoph Hauert. Coevolutionary dynamics in large, but finite populations. Phys. Rev. E, 74(1):011901, July 2006.

[48] Arne Traulsen, Christoph Hauert, Hannelore De Silva, Martin A. Nowak, and Karl Sigmund. Exploration dynamics in evolutionary games. PNAS, 106(3):709–712, January 2009.

[49] Arne Traulsen, Torsten Röh, and Manfred Milinski. An economic experiment reveals that humans prefer pool punishment to maintain the commons. Proc. R. Soc. B, page rspb20120937, July 2012.

[50] T. Wenselheer, H. Helenaré, A. Hart, and F. L. W. Ratnieks. Worker
reproduction and policing in insect societies: an ESS analysis. Journal of Evolutionary Biology, 17(5):1035–1047, 2004.

[51] Tom Wenseleers, Adam G. Hart, and Francis L. W. Ratnieks. When resistance is useless: Policing and the evolution of reproductive acquiescence in insect societies. The American Naturalist, 164(6):E154–E167, December 2004.

[52] Toshio Yamagishi. The provision of a sanctioning system as a public good. Journal of Personality and Social Psychology, 51(1):110–116, 1986.

\[
\begin{align*}
D & \quad C \\
P & \quad D
\end{align*}
\]

\[
\begin{align*}
0.01 & \\
0.06 & \\
0.3 & \\
\end{align*}
\]

Figure 3: Stable state distribution of the Markov chain for the evolutionary PGG/P for \( x = 0 \) (i.e. a well-mixed population), different mutation rates \( \mu \) and otherwise default parameters (see Table[1]). The mean of the distribution is marked with a blue +. Note that the gradient is logarithmic. When \( \mu \) is low (a), pure defection is much more prevalent than other states, by orders of magnitude, and states on the border of the state space generally have a higher probability than those on the inside. Only for high mutation rates (c) is there a mode of the distribution near the equilibrium of the replicator-mutator equation.

21
Figure 21: Equilibria, trajectories and the underlying vector field of the replicator-mutator approximation Public Goods Game with non-optional tax payment. When cooperators automatically pay a fixed tax to punishers, the cooperative equilibrium $Y$ exists already for smaller mutation rates $\mu$, and its basin of attraction $R_Y$ includes the whole punisher-taxpayer line, reducing the basin of attraction $R_X$ compared to the basic model.

Figure 24: Equilibrium distribution of the Markov chain for the PGG/P with optional fitness transfer from cooperators ("taxpayers") to punishers. Colours denote the cumulative relative frequency of all states with the same sum of cooperators plus taxpayers. The means of the distributions shown are marked by a blue *. A black + marks the mean of the equilibrium distribution generated by the corresponding evolutionary game for mandatory tax payments, while the mean of the basic model with only $D, C$ and $P$ is marked by a blue x. Competition between cooperators and taxpayers leads to an increase in defection.