# Best response dynamics improve sustainability and equity outcomes in common-pool resources problems, compared to imitation dynamics

Isaiah Farahbakhsh *  Chris T. Bauch †  Madhur Anand ‡

**Abstract**

Shared resource extraction among profit-seeking individuals involves a tension between individual benefit and the collective well-being represented by the persistence of the resource. Many game theoretic models explore this scenario, but these models tend to assume either best response dynamics (where individuals instantly switch to better paying strategies) or imitation dynamics (where individuals copy successful strategies from neighbours), and do not systematically compare predictions under the two assumptions. Here we propose an iterated game on a social network with payoff functions that depend on the state of the resource. Agents harvest the resource, and the strategy composition of the population evolves until an equilibrium is reached. The system is then repeatedly perturbed and allowed to re-equilibrate. We compare model predictions under best response and imitation dynamics. Compared to imitation dynamics, best response dynamics increase sustainability of the system, the persistence of cooperation while decreasing inequality and debt corresponding to the Gini index in the agents' cumulative payoffs. Additionally, for best response dynamics, the number of strategy switches before equilibrium fits a power-law distribution under a subset of the parameter space, suggesting the system is in a state of self-organized criticality. We find little variation in most mean results over different network topologies; however, there is significant variation in the distributions of the raw data, equality of payoff, clustering of like strategies and power-law fit. We suggest the primary mechanisms driving the difference in sustainability between the two strategy update rules to be the clustering of like strategies as well as the time delay imposed by an imitation processes. Given the strikingly different outcomes for best response versus imitation dynamics for common-pool resource systems, our results suggest that modellers should choose strategy update rules that best represent decision-making in their study systems.

*Department of Applied Mathematics, University of Waterloo, 200 University Ave W, Waterloo, Ontario N2L 3G1, Canada  email: infarahb@uwaterloo.ca

†Department of Applied Mathematics, University of Waterloo, 200 University Ave W, Waterloo, Ontario N2L 3G1, Canada  email: cbauch@uwaterloo.ca

‡School of Environmental Sciences, University of Guelph, 50 Stone Rd E, Guelph, Ontario N1G 2W1, Canada  email: manand@uoguelph.ca

## 1 Introduction

Common-pool resources are resources such as forests and fisheries which are both available for public extraction and finite, therefore being very susceptible to overuse by profit-seeking individuals [17, 44, 46, 3]. This fragility in the face of individual self-interest has led to the study of common-pool resources in many diverse fields such as economics, sociology, applied mathematics, and ecology [55, 8, 7, 59, 17, 50, 52, 57, 44, 54, 25]. A pervasive idea regarding the outcome of common-pool resources is known as *the tragedy of the commons*. The underlying argument is that given a resource shared among rational individuals, each individual can increase their personal profit by increasing their level of resource extraction. There is an associated cost to the health of the resource; however, this cost is shared by all individuals accessing the commons and is consequently less than the expected profit of increased extraction. The conclusion of the tragedy of the commons is that any common-pool resource is doomed to depletion in the absence of control by a central government or private ownership [24]. Following the inception of this paradigm, there have been many studies investigating its validity. Scholars have now found that many human communities are in fact able to sustainably harvest common-pool resources without a centralized governing body [27, 44, 43, 17, 31]. One important reason for these successes are the value systems that these communities follow regarding proper resource use. These value systems, known as social norms, can encourage individuals to harvest sustainably and also punish those who do not. These social norms are dynamic and can evolve among individuals and communities due to external and internal pressures [17, 43, 44, 35, 36, 27].

Evolutionary game theory offers a framework to model the propagation of social norms in which individuals are represented by players of a game, and social norms are the game strategies. There have been many types of models exploring the persistence of common-pool resources as well as the tragedy of the commons in general. These models range from systems of ordinary differential equations assuming a well-mixed population to discrete spatial models where the agents interact on a network or a lattice [18, 20, 23, 39, 38, 45, 49, 57, 59, 30, 47, 28, 40, 13, 56, 34, 15, 51]. In spatial models, cooperative strategies, which are not favoured in well-mixed models, are shown to be much more resilient, challenging the base assumptions of the tragedy of the commons. This can be explained through the formation of cooperative spatial clusters, where defection becomes disadvantageous at the boundaries and cannot invade [39, 38, 51].

Although there have been many studies on abstract tragedy of the commons games, there is little literature presenting models which structure a human population on a network from which the population harvests an ecological resource. Models that explore this interaction can give significant insight into common-pool resources since most human interactions are determined by social networks rather than physical location within a community [16, 11]. One study that has explored this analyzed an empirical network common-pool resource model where cooperators maximize their payoffs over a longer time horizon than defectors. Using a single strategy update rule, more links between cooperators, as well as smaller networks, were found to promote cooperation and efficient resource management [28]. A second study investigated agents on a lattice informed by both social norms and organizational rules when harvesting a forest; however, the long term dynamics were governed by sampling strategies from a normal distribution [1]. A third study modelling agents on a network found that in most cases increasing the probability of rewiring cooperator-defector links increased cooperation in the system [32]. In the previous studies, as well as the majority of game theoretic models, one mechanism for the evolution of strategies among agents is assumed. Rarely in the literature are the dependence of model predictions on this choice of mechanism compared. In evolutionary game theory, there are two evolutionary dynamics widely used. One involves a given agent comparing their payoff with their neighbours and changing their strategy to that of their highest-earning neighbour ("imitation dynamics") [26]. The other involves a given agent comparing the expected payoff of their current strategy to that of a different strategy for the next time step and choosing the strategy with the highest expected payoff ("best-response dynamics") [18]. Both of these mechanisms are supported by the literature; however, their direct comparison in a single model could shed light on the qualitative differences resulting from these contrasting psychological inclinations [22, 12]. The closest we have seen to this is a study that systematically compared imitation dynamics and strategy evolution using genetic algorithms with an N-player Prisoner's Dilemma game on a lattice. The authors found evolution to significantly promote cooperation and increase strategy convergence rates in their system. The genetic algorithm incorporated aspects of imitation and the model did not include payoffs coupled to a common-pool resource [13]. In our model, we are using best-response dynamics instead of genetic modification as well as an explicit common-pool resource, allowing us to further separate social learning from independent prediction in a human-environment system.

Another concept that is important in the literature regarding evolutionary game-theoretic models is the systems' sensitivity to external perturbations [18, 26]. This is often investigated by changing the strategy of an agent regardless of its perceived benefit when the system is at equilibrium. Then, the resulting time steps that the system takes in order to reach equilibrium are counted, and this number of time steps is referred to as a cascade. What has been found in previous studies is that the cascade size can vary quite drastically covering many orders of magnitude and in many cases, it behaves in line with the criteria for self-organized criticality proposed by Bak et al. [4, 26, 18]. Self-organized criticality posits that many complex systems with local interactions tune themselves into a critical state that displays scale-invariant spatial or temporal characteristics. In this state, small perturbations can cascade throughout the system with the cascade size-distribution fitting a power law [4]. Power law distributions are scale-invariant, having the form,

$$p(x) \propto x^{-\alpha} \tag{1}$$

which forms a straight line when graphed on a log-log plot. These distributions are found in many physical systems and are most prominent when the system is in a critical state and extremely sensitive to perturbations [29].

The model presented in this paper explores a human population arranged on various network topologies that harvest a well-mixed ecological resource represented with a logistic growth difference equation. Both strategy evolving mechanisms are compared over identical model parameters. Topologies are also compared since spatial structure has often been found to influence outcomes in tragedy of the commons scenarios [20, 23, 53, 50]. Through running this model across a large parameter space, insight will be gained regarding the mechanisms which lead to cooperation and the persistence of resources in common-pool resource systems.

# 2 Methods

## 2.1 Harvesting model

The model simulated a network of $N$ individuals or nodes harvesting from a generalized common-pool resource. At initialization, each node is randomly given one of two harvesting strategies; cooperation or defection, with equal probability. At each time step, every node simultaneously harvests the resource with cooperators harvesting less than their 'maximal equal share' and defectors harvesting more than their 'maximal equal share'. The 'maximal equal share' is defined by $\frac{R_t}{N}$ where $R_t$ is the resource at time $t$. When harvesting, cooperators inflict a punishment proportional to the depletion of the resource to any defective nodes to which they are directly linked. This proportionality is justified because enforcement by social norms is often more severe when the resource is close to depletion [42, 10, 37]. This punishment also incurs a small cost to the cooperators' own harvest. Each nodes respective net profit at each time step is given by the following payoff functions,

$$\pi_c = \frac{1}{N}(cR_t - a \cdot p \cdot n_d(1 - R_t)), \quad c, a < 1 \tag{2}$$

$$\pi_d = \frac{1}{N}(dR_t - p \cdot n_c(1 - R_t)), \quad d \geq 1 \tag{3}$$

where $\pi_c$ and $\pi_d$ are the cooperators and defectors payoffs respectively, $c$ and $d$ are the proportions of the 'maximal equal share' that the cooperative and defective nodes harvest respectively, $p$ is the magnitude of the punishment inflicted by cooperative nodes on defective nodes, $a$ is the relative cost of that punishment for the cooperative nodes, $n_c$ and $n_d$ are the amount of cooperative and defective nodes directly connected to any given node in the network.

The resource is updated at each time step using a logistic difference equation from which the net harvest of all the nodes in the network is subtracted. It is modeled by,

$$R_{t+1} = R_t(1 + F(1 - R_t)) - \frac{R_t}{N}(dN_d + cN_c) \tag{4}$$

where $N_d$ and $N_c$ are the total number of defectors and cooperators in the network respectively.

After all nodes harvest and update their payoffs, one randomly selected node changes its strategy if it is perceived to increase its individual payoff. This process is repeated until an equilibrium is reached where no node can change its strategy to increase its perceived payoff. Once equilibrium is reached, the system is then perturbed by randomly choosing a node and changing its harvesting strategy regardless of any perceived profit increase. The system is then left to re-equilibrate. This process continues until the resource is depleted or a predetermined number of perturbations is reached. If at any point during the simulation, the resource drops below a critical level, $\epsilon$, the resource is extinct and the simulation ends. This represents the resource reaching a depleted level from which it cannot recover. The value for $\epsilon$ scales with the network size and is given by $\frac{1}{N \cdot 10000}$. A flowchart of this model is shown in Figure S1.

## 2.2 Strategy propagation

Two different mechanisms for the propagation of harvesting strategies among nodes were compared. One is best response dynamics and the other we call imitation. In the case of imitation, the node that is selected for strategy reassignment compares its cumulative payoff with that of its neighbouring nodes. If any neighbouring node has a higher cumulative payoff, the selected node will change its strategy to that of its highest earning neighbour. If there are multiple highest earning neighbours, the selected node will randomly choose between the highest earning neighbours with equal probability. In the case of best response dynamics, the selected node will change its strategy to that which is most profitable for the next time step. It does so by simulating harvesting from the updated resource with the strategy it wasn't previously using. If the expected payoff from that simulation is greater than the selected node's mean payoff over all previous time steps, the node will change its' harvesting strategy to the new one.

## 2.3 Network simulations

Four network topologies were tested in this model; lattice, random, scale-free and small-world networks. These were chosen as they are very common network topologies used in the literature, and both scale-free and small-world networks share similarities with social networks [19, 58]. In the square lattice, the von Neumann neighbourhood was used with periodic boundary conditions. For a random network of $N$ nodes, the Erdös-Rényi model was used where, for average degree of connectivity, $k$, an edge was created between each node with a probability of $\frac{k}{N-1}$ [21]. For scale-free networks, the Barabási–Albert model [5] was used and for small-world networks, the Watts-Strogatz model was used with $\beta$, the probability of rewiring an edge in the ring equal to 0.08 [58].

## 2.4 Parameters

Each network was generated with $N = 15 \times 15 = 225$ nodes and average degree of connectivity, $k = 4.0$. These values were chosen to allow for direct comparison with the square lattice (as $N$ is a perfect square) using a von Neumann neighbourhood while at the same time being within the range of those used in Ebel & Bornholdts Prisoner's Dilemma network study [18]. For each network topology, all combinations of parameters shown in Table S1 were run. Cooperators harvest was limited to $c \leq 0.6$ since cooperators are conservationists, thus committing to harvest significantly less than their fair share of the resource. For values of $c > 0.6$, the resource was depleted very quickly while being initialized with half of the nodes harvesting at $d \geq 1.0$. Defectors harvest was limited to $d \leq 1.9$ since values larger than this result in immediate resource depletion. Fecundity was limited to $0.5 \leq F \leq 0.9$ to agree with values that allowed for the persistence of the resource within realistic constraints. For punishment, values larger than 0.3 were not included since they made punishment and its cost too severe to be realistic. For the cost of punishment, $a = 0.1$ such that this cost would not be high enough to disincentivize cooperation. This parameter space resulted in 450 parameter sets that were simulated for each network topology. Baseline parameter values were chosen based on mid-range parameter values which allowed for the persistence of the resource over multi-parameter variations (see Table S1). This was done as resource persistence is a pre-condition for gathering much of the data that will be analysed. The resource was initialized at $R = 1.0$ to simulate a community harvesting a resource that has not been harvested previously. For each parameter set and topology, 50 networks were generated and each network was perturbed 100 times.

## 2.5 Outcome metrics

From the simulations, the cascade size, defined as the amount of strategy switches taken for the system re-equilibrate after a perturbation, was recorded. The perturbed node's degree and clustering coefficient were collected as well. For each cascade, the average number of strategy switches per time step was also logged. The rest of the data was collected every time the system reached equilibrium after a perturbation. This data consisted of the level of the resource and the number of cooperators at equilibrium. Average local clustering coefficient and network transitivity over the whole network and over the subnetworks of same strategy-types were recorded as well as network modularity. The Gini index for total node payoff was also recorded over the whole network to investigate the role of the distribution of wealth in the system [45]. The Gini index was calculated using,

$$\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}|P_i - P_j|}{2N\sum_{i=1}^{N}P_i} \tag{5}$$

where $P_m$, $m = \{1, 2, ..., N\}$ is the cumulative payoff of a given node, $m$. [48].

Finally, the sustainability level of each parameter combination was calculated. Here, sustainability is defined by whether the resource can persist and is quantified by the frequency of simulations in which the resource is not depleted to a value less than $\epsilon$. Therefore, this sustainability level is a value between 0 and 1, with 1 being maximally sustainable and 0 being completely unsustainable.

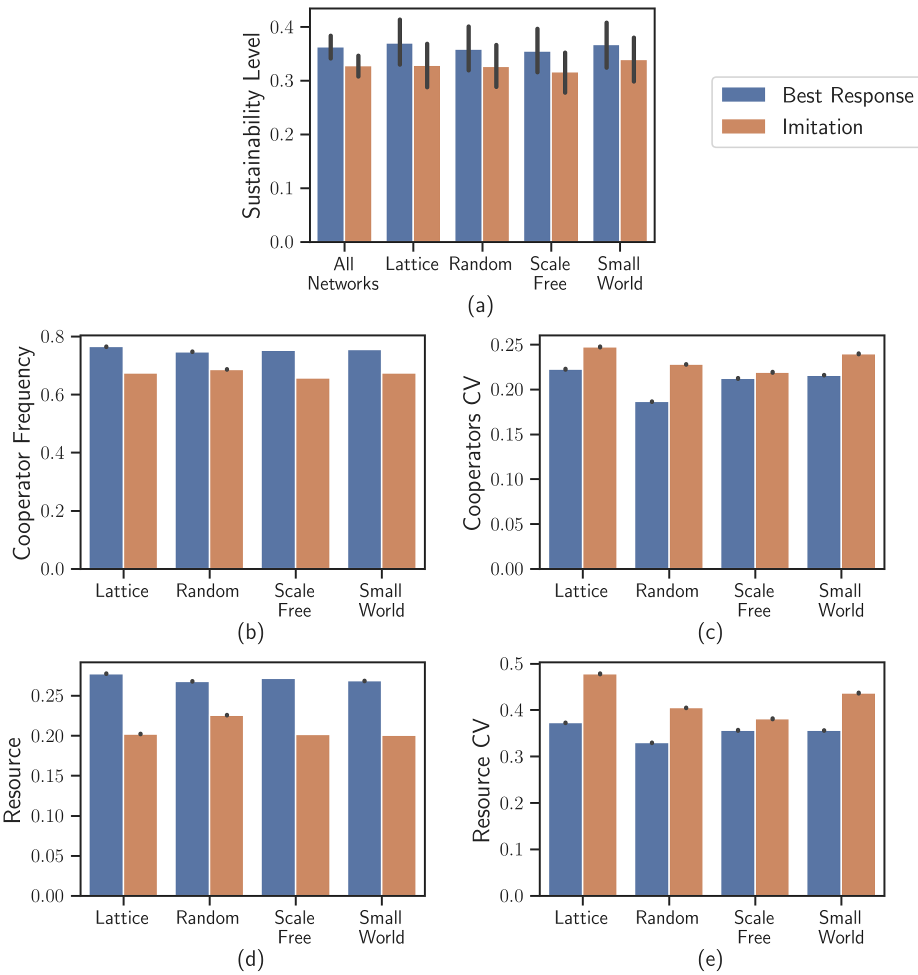


**Figure 1:** Mean metrics relating to the systems' sustainability level defined as the frequency of simulations in which the resource is not depleted below $\epsilon$. Best response dynamics promote higher sustainability (a), cooperator frequency (b) and resource level (d) at equilibrium while also having lower coefficients of variation (CV) (c,e). Error bars represent the 95% confidence interval.

# 3 Results

## 3.1 Sustainability of the resource

Over the parameter space and all topologies, the systems with best response dynamics were more sustainable than those employing imitation. However, these differences were not statistically significant, due to many simulations having extreme values for sustainability. Two important variables related to sustainability are the number of cooperators and the level of the resource. These variables were recorded when the system was at equilibrium and the analysis of their means and coefficients of variation (CV) over all topologies were conducted as seen in Figure 1. Similar to the sustainability of the system, the mean cooperator frequency and resource level were both significantly higher for best response dynamics; however, there was a greater differential between the mean resource levels for each network topology. The CV of the cooperator frequency was low (relative to that of the resource) for both best response dynamics and imitation; however, the difference between them was statistically significant with imitation having a higher CV. Additionally, over all topologies imitation had a higher CV for the resource level. This difference can be explained by the underlying distributions of the data. Here, best response dynamics have single values with very high frequency surrounded by much lower frequency values whereas imitation dynamics result in distributions closer to a skewed Gaussian with many values having mid range frequencies over a larger range (Figure S2). The mechanism driving these distributions is that best response dynamics converge to a global optimum rapidly as each node evolving its strategy at a given time step is always able to access the strategy it is not using. In contrast, imitation dynamics allow only for sampling strategies that are in use by connected nodes. Therefore, the system can drastically change between perturbations as seen in Figure 2.

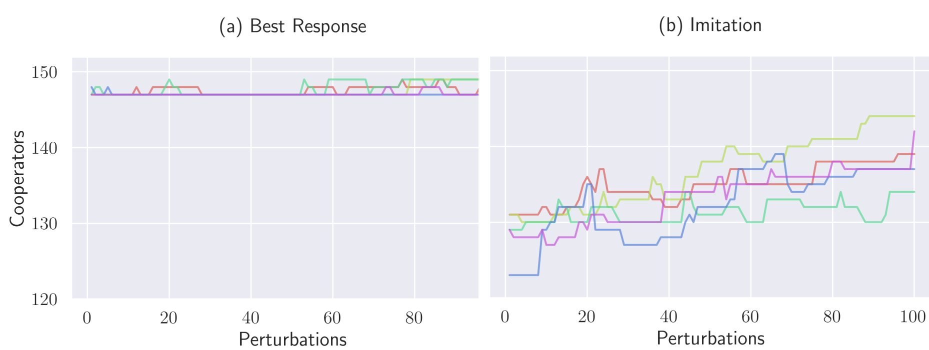


**Figure 2:** Cooperators at equilibrium after perturbations for 5 runs on a scale-free network with identical parameters. Each colour represents a different stochastic realization of the model. In the best response system (a), the number of cooperators does not vary significantly after the first perturbation; however, in the imitation system (b) the there is much more variance.

### 3.2 Multi-parameter analysis

To gain insight into the different qualitative regimes within each network topology, the data was visualized in parameter planes visualized as heatmaps. The parameters varied in the heatmaps were $c$, the cooperators harvest and $d$, the defectors harvest. This is because these two parameters cover a much larger range than the others as they can vary greatly among human systems and have much more relaxed realism constraints [33, 41]. Examining the heatmaps for the mean sustainability over each topology, the plots show that as both harvesting rates are increased, the system becomes less sustainable to the point where not a single simulation can persist (Figure 3a,b). Intuitively, an increase in sustainability would be correlated with an increase in cooperation. However, the heatmaps for the number of cooperators at equilibrium on all network topologies in the imitation system show that the number of cooperators increase with both $c$ and $d$ such that they are maximal at the point in which the system cannot persist due to over-extraction of resources (Figure 3d). In this case, stress in the system encourages cooperation. For increasing values of $c$, an increase in cooperation is expected because the payoff, and therefore the incentive to cooperate, has increased. With high values of $d$, the increase in number cooperators is less obvious. When $d$ increases, the incentive to defect is strong, which could cause a large-scale propagation of defection in the network. However, with a large defectors' harvest, the resource decreases very quickly and once it reaches a very low level, the cooperators' punishment becomes very large, to the point that even with few cooperators in the system, defectors adjacent to them cannot harvest profitably and are likely to switch to the cooperator strategy.

For best response dynamics, there is a difference to this narrative as $c$ is increased. At high $c$ levels ($c > 0.4$), there is a drastic increase in cooperators and the average number of cooperators does in fact decrease with $d$. This could be caused by having a large enough incentive to cooperate at initialization such that there are never enough defectors to bring the resource to a level in which the cooperators' punishment causes defective harvesting to be unprofitable. This agrees with the resource having its maximal value at $c = 0.5$ as well as the system showing a near saturation of cooperators at $c = 0.6$ (Figure 3c,e). At this saturation of cooperators, the resource level decreases for low $d$, demonstrating that above $c = 0.5$, cooperators have taken over the system and any higher payoff for them results in greater resource extraction as defectors are not influencing the system. With imitation dynamics, the system does not experience the same regime change to a saturation of cooperators, most likely due to clusters of defective nodes which are immune to invasion by cooperators.

Regarding modularity in the best response system, for $c \leq 0.4$, modularity is maximized when both $d$ and $c$ are minimized (Figure 3g). This happens in order for both cooperators and defectors to maximize payoff as now punishment and cost of punishment play a much more significant role in reducing payoffs for nodes connected to others with opposing strategies. For $c > 0.4$, modularity approaches 0 since the network is approaching a globally homogeneous population. In the imitation system, these phenomena are not observed (Figure 3h).

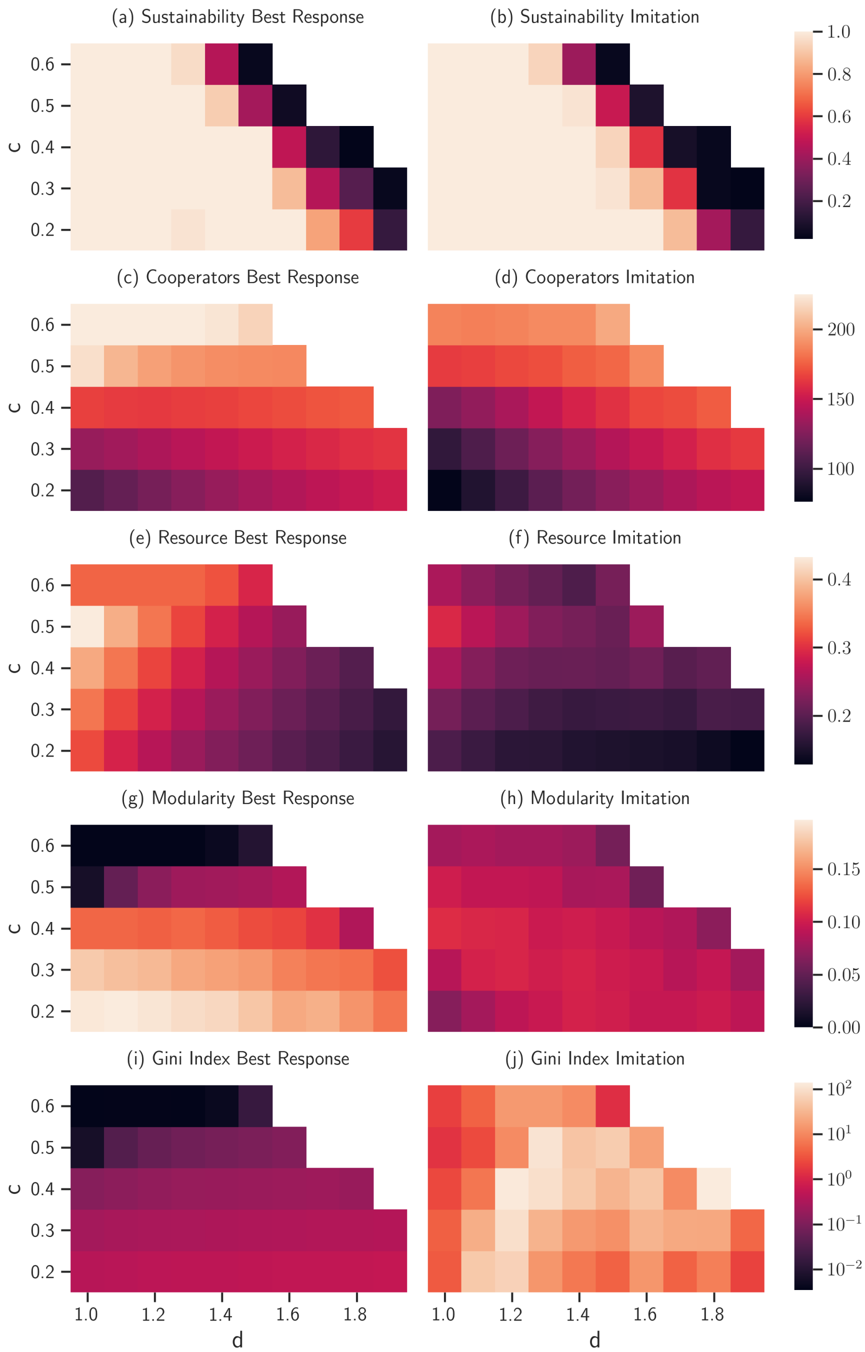


**Figure 3:** Comparing mean values for best response and imitation dynamics on the scale-free network for baseline parameters. The axes are $d$, corresponding to the defectors' harvest and $c$, corresponding to the cooperators' harvest. The values shown are sustainability (a, b), cooperators at equilibrium (c, d), resource at equilibrium (e, f), modularity at equilibrium (g, h) and the absolute value of the Gini index at equilibrium (i, j).

Instead, modularity is maximized for intermediate values of $d$ and $c$. The reason it is not maximized for low values of $d$ and $c$ could be because the number of cooperators is low enough that the system is approaching global homogeneity towards defection.

Finally, the distribution of payoffs across the network was quantified using the Gini index. This metric can be interpreted as zero being complete equality with the distance from zero measuring the extent of inequality among agents. Usually the index is in the range $[0, 1]$; however, including negative individual payoffs in its calculation can take it out of this range. Additionally, a negative Gini index means that the net total payoff of the system is negative. When comparing systems with heatmaps, the absolute value of the Gini index is used in order to have a suitable resolution to compare both systems. Under the baseline parameters, the best response system has much lower Gini index values (Figure 3i,j). The inequality in the system increases with $d$ and decreases with $c$. This is seen across all parameters for best response dynamics; however, the negative correlation with $c$ is much stronger for the complete parameter space. This correlation with $d$ and $c$ follows from the fact that as the two parameters approach each other so do $\pi_d$ and $\pi_c$, increasing equality in the system. There is also a sudden decrease in Gini index in the same high $c$ region where the number of cooperators and network modularity also experience a drastic transition. This regime experiences the highest equality because the system is saturated with cooperators and therefore all payoffs are equal. For imitation dynamics, the Gini index is much higher and does not show a strong correlation with any parameters or variables. The reason for the drastic inequalities in payoff could be due to defective nodes which only share edges with like strategy nodes. These nodes will not receive any punishment and will harvest every round at the highest possible level. Unless the neighbouring nodes are perturbed, they will continue to imitate the high earning defector which could result in accumulation of debt if they share edges with cooperators, especially when the resource is at a low level. As the modularity is still relatively low for these high magnitude Gini index values, these groups of like strategies must not reach a significantly large size. To differentiate these from clusters around multiple nodes large enough to significantly influence the modularity, we will call this phenomenon 'micro-clustering'.

Consistently across all network topologies for best response dynamics, the extrema for Gini index values are positive and less than 1, meaning that there is no excessive debt among individual nodes and all systems have a net positive cumulative payoff (Figure 4c). For imitation dynamics, the extrema of the Gini index surpass those of best response dynamics by several orders of magnitude and in fact vary significantly across network topologies. Regarding the frequency of debt in the system, global and individual debt is most prevalent in scale-free networks and rarest in small-world networks (Figure 4d). This could be explained by separately examining the clustering of individual strategies in each network quantified by the relative transitivity for each strategy type (Figure 4a,b). Transitivity offers a metric for clustering through the presence of triadic closure in the network. Due to the impossibility of triads in lattice networks, this topology was excluded from the transitivity analysis. When looking at same-strategy clustering, transitivity was calculated over the whole network as well as a subset of the network made up of nodes with a single strategy and the edges they share. Since the transitivity over the entire network varied significantly across topologies, to compare topologies, the transitivity of a given strategy type was scaled by the transitivity of the entire network. This gives us the relative transitivity, which measures whether or not the presence of triadic closure in a given strategy type is greater than that found in the entire network.

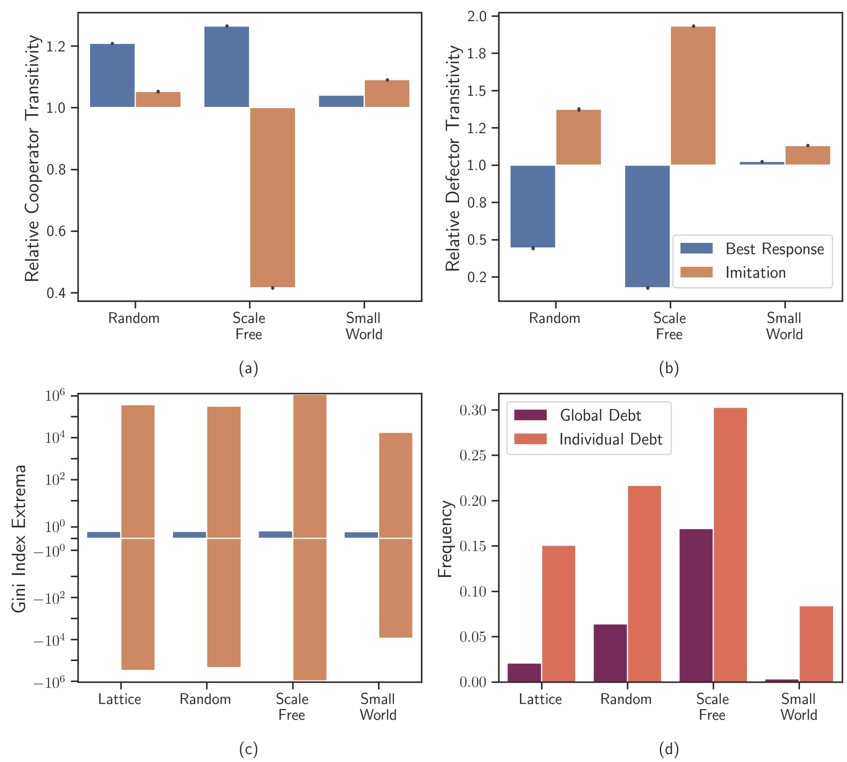


**Figure 4:** The clustering of like strategies as quantified by relative transitivity varies significantly across network topologies (top). Wealth distribution across network topologies (bottom). In imitation systems there is much more inequality than best response systems as seen in the Gini index extrema (c). There are also systems with debt in imitation dynamics (d) which is not seen in best response systems.

For scale-free networks with imitation dynamics, the transitivity of cooperators is significantly lower than that of the entire network (giving a relative transitivity less than 1) and the transitivity of defectors is significantly higher (giving a relative transitivity greater than 1). This is due to the scale-free degree distribution which allows a small number of very high-degree nodes which are not found in other the other network topologies. If these high-degree nodes are initialized as defectors, they can be central in much larger 'micro-clusters' than are possible on other network topologies. The central defecting node of this 'micro-cluster' is then able to influence a greater number of nodes to remain defectors when it is unprofitable at low resource levels, resulting in a larger number of nodes accumulating debt. Interestingly, in best response systems, these 'micro-clusters' found in scale-free networks would more likely be surrounding a cooperative node. In small-world networks, the transitivity of each strategy is very close to the transitivity of the whole network. This is because having nodes with degrees as high as those found in scale-free networks has a low probability in this system since the re-wiring probability is low ($\beta = 0.08$). Additionally, these networks have a low degree of separation between any given pair of nodes, making it difficult for 'micro-clusters' to form, which are strong against invasion from the opposite strategy as there is a greater probability of high-earning cooperators to be connected to defectors by a very short path.

### 3.3 Power-law analysis

In this section, all power-law fitting was conducted using the *powerlaw* Python package [2]. As smaller values in empirical data usually do not follow power laws and power-law fitting is often performed on the tails of probability distributions, the data was fit on cascades larger than the size of the system, $N = 225$ [14].

To quantify which cascade size distributions fit a power-law, the Kolmogorov–Smirnov goodness-of-fit test was performed on the data. As the Kolmogorov–Smirnov test is extremely sensitive to noise, instead of testing whether the distribution is a power law, the fit of a power law is compared to that of an exponential distribution. If the distribution is closer to that of an exponential, we can rule out that it is a power law fit, and if it is closer to a power law we can at least be certain that it is a heavy-tailed distribution [14, 2].

From the results of the Kolmogorov– Smirnov test, all of the cascade size distributions for imitation dynamics fit an exponential distribution more closely than a power law distribution, immediately ruling out the potential for self-organized criticality in these systems. However, for best response dynamics, about 21% of the parameter space had distributions that were closer to a power law for lattice (20.89%), random (22.44%) and small-world (21.77%) networks and therefore confirm the existence of heavy-tailed distributions suggesting the occurrence of self-organized criticality in these systems. For scale-free networks, only 6.67% of the parameter space had distributions closer to a power law. This could be due in part to the existence of larger 'micro-clusters' surrounding high degree nodes as these have a very low probability of invasion by the opposite strategy and could decrease the probability of very large cascade sizes leading to much fewer heavy-tailed cascade size distributions.

Along with displaying a power law distribution, self-organized critical systems must display system-wide effects that are triggered by small perturbations [4]. In terms of this model, the extent of system-wide cascades can be examined through the maximum cascade size as well as the proportion of cascades larger

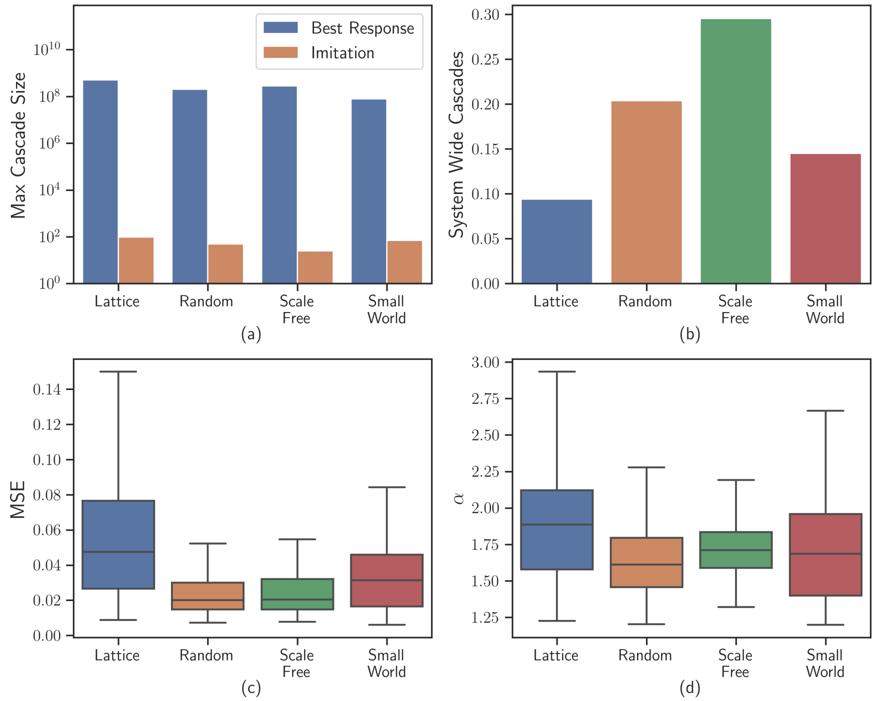


**Figure 5:** Metrics of self-organized criticality. Best response dynamics show evidence of self-organized criticality due to large maximum cascade sizes and a significant frequency of system wide cascades (top). All topologies in the best response systems show similarities in the mean standard error (MSE) and slope ($\alpha$) of the power-law fit (bottom).

than the number of nodes in the system (see Figure 5b). For the simulation presented in this paper, it would be the proportion of cascades larger than $N = 225$. For system wide cascades, systems with best response dynamics had a significantly larger maximum cascade size (by 6-7 orders of magnitude) as well as a significantly greater proportion of cascades of size greater than $N$ (Figure 5a). Interestingly, scale-free networks have a much higher proportion of cascade sizes larger than $N$ suggesting that the high-degree nodes unique to this system can in fact promote strategy switching up to a point (Figure 5b). This, along with the results from the Kolmogorov– Smirnov test strongly suggests that systems with best response dynamics exhibit self-organized criticality much more than those with imitation; however, the existence of high-degree nodes in scale-free networks acts contrary to this phenomenon.

Finally, there are similarities across all network topologies regarding fitting the cascade size distribution to a power-law. The mean standard error, which determines the closeness of the power-law fit over the entire parameter space is similar across all topologies despite scale-free networks scoring much lower as a power-law distribution in the Kolmogorov– Smirnov test (Figure 5c). Additionally, the slope of the power-law fit given by $\alpha$ does not significantly differ across topologies (Figure 5d). This implies an element of universality in the shape of the cascade distributions over all network types, even as the scale-free network distributions are less heavy-tailed. Regarding properties of the node being perturbed before each cascade, neither the local clustering coefficient nor the degree of the perturbed node were correlated with the cascade sizes.

# 4 Discussion

In this model, a tragedy of the commons was averted in many cases, with best response dynamics increasing the amount of sustainable runs compared to imitation dynamics. Just as importantly, outcomes were more equitable under best response dynamics, as measured by individual debt and the population's Gini index. Additionally, the cascade size distribution in systems with best response dynamics had a closer fit to a power-law and larger cascade sizes in general, demonstrating evidence of self-organized criticality.

Spatial structure, however, did not have a drastic effect on the persistence of the resource or number of cooperators at equilibrium. This agrees with the 2012 empirical study by Gracia-Lázarao et al. in which humans played a spatial Prisoner's Dilemma on a lattice and a scale-free network. In this study, there was no significant difference in cooperation among the two network topologies and in fact, the level of cooperation was comparable to that which arose in unstructured populations [23]. This spatial independence is surprising given the importance of local interaction in this model. However for other metrics, the network topology did have a significant effect, such as for the Gini index, the clustering of strategies and the goodness-of-fit of the cascade size distribution to a power-law. Most of these differences can be attributed to the existence of high-degree nodes in scale-free networks which can be central in 'micro-clusters' of the same strategy type.

The strategy propagation mechanism proved to be extremely important in determining system sustainability. In all cases, best response dynamics were much more successful than imitation at averting the tragedy of the commons. This, along with the rapid convergence of best response dynamics, complements similar findings when comparing strategy propagation in the N-player iterated Prisoner's Dilemma game [13]. These results hint at the importance of context-dependent foresight as opposed to pure imitation of peers regardless of context for the sustainability of common-pool resource-dependent human systems. They also suggest that in systems where imitation is important, more sustainable outcomes might be achievable by increasing the rate of social learning. This finding is consistent with other recent research social-climate systems [9]. In both strategy update rules, each agent is acting in its own self interest; however, the agent who makes choices based on its current state and relation to others benefits itself and the overall community more than a self-interested agent who defaults to following another individual's path to success.

Imitation can be seen as a representation of social learning through copying others. History has shown social learning to be highly adaptive, primarily through the propagation of information such as technology and survival skills. However, in the past, this social learning has been practised with a combined knowledge of an individual's context and social norms, as modelled with best response dynamics [6]. In this model, pure social learning is less sustainable in comparison to independent prediction modelled by best response dynamics. This could be due to the fact that best response dynamics take into account the current state of the resource while imitation dynamics do not. Consequently, agents using a best response strategy update rule are aware of the consequences of over-harvesting a resource near depletion. Therefore, they explicitly take into account the social norms of the community when making a decision on how to harvest. Imitating agents however, are only trying to increase their profitability using a strategy that was advantageous for others in the past, thus only indirectly accounting for social norms. This implicit time delay in decision making can have significant detriment to the wealth of the resource as well as the individual.

The hypothesis of the benefit of best response dynamics versus imitation dynamics for common-pool resource problems should be further tested with existing models in the literature that employ either update rule. An identical model with the other update rule can then be analysed and the results of both models can be compared to investigate whether best response dynamics do in fact prove more beneficial to the community than imitation over a diverse array of model assumptions. Finally, this model could be tested with empirical networks and case studies to verify whether the assumptions and conclusions gleaned from this study remain true in practice.

# 5 Supplementary Material

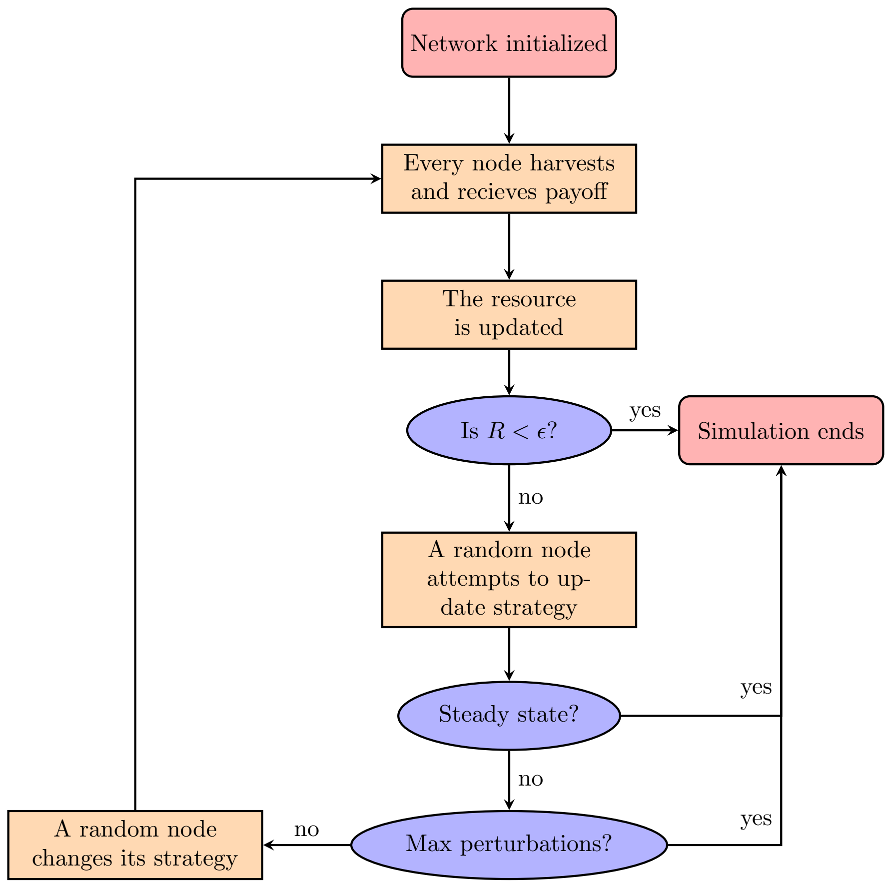


**Figure S1:** Flowchart of the model

| **Parameter** | **Minimum** | **Maximum** | **Step** | **Baseline** |
|---|---|---|---|---|
| Cooperators Harvest ($c$) | 0.2 | 0.6 | 0.1 | 0.3 |
| Defectors Harvest ($d$) | 1.0 | 1.9 | 0.1 | 1.4 |
| Fecundity ($F$) | 0.5 | 0.9 | 0.2 | 0.9 |
| Punishment ($p$) | 0.1 | 0.3 | 0.2 | 0.2 |
| Cost of Punishment ($a$) | – | – | – | 0.1 |

**Table S1:** Parameter ranges used in model simulations: minimum and maximum of parameter range, increments sampled, and baseline values.

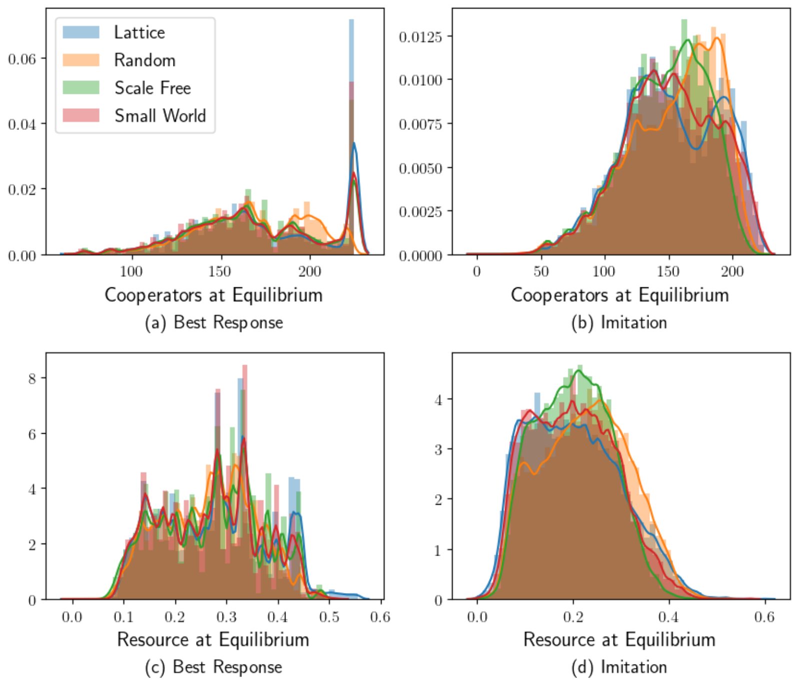


**Figure S2:** The frequency distribution of cooperators (a, b) and resource (c, d) at equilibrium, normalized with a Gaussian kernel density estimate

# Acknowledgments

This work was supported by the NSERC Alexander Graham Bell Scholarship and Ontario Graduate Scholarship to I.F., the NSERC Discovery Grant to C.T.B and M.A., and the James S. McDonnell Foundation Complex Systems Scholar Award to M.A.

# References

[1] Arun Agrawal, Daniel G Brown, Gautam Rao, Rick Riolo, Derek T Robinson, and Michael Bommarito II. Interactions between organizations and networks in common-pool resource governance. *Environmental Science & Policy*, 25:138–146, 2013.

[2] Jeff Alstott, Ed Bullmore, and Dietmar Plenz. powerlaw: a python package for analysis of heavy-tailed distributions. *PloS one*, 9(1):e85777, 2014.

[3] Mark Appiah, Dominic Blay, Lawrence Damnyag, Francis K Dwomoh, Ari Pappinen, and Olavi Luukkanen. Dependence on forest resources and tropical deforestation in ghana. *Environment, Development and Sustainability*, 11(3):471–487, 2009.

[4] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Physical review A*, 38(1):364, 1988.

[5] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[6] Robert Boyd, Peter J Richerson, and Joseph Henrich. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(Supplement 2):10918–10925, 2011.

[7] Hannelore Brandt, Christoph Hauert, and Karl Sigmund. Punishment and reputation in spatial public goods games. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1519):1099–1104, 2003.

[8] Chris Brozyna, Todd Guilfoos, and Stephen Atlas. Slow and deliberate cooperation in the commons. *Nature Sustainability*, 1(4):184, 2018.

[9] Thomas M Bury, Chris T Bauch, and Madhur Anand. Charting pathways to climate change mitigation in a coupled socio-climate model. *PLoS computational biology*, 15(6), 2019.

[10] Juan Camilo Cárdenas. Human behavior and the use of experiments to understand the agricultural, resource, and environmental challenges of the xxi century. *Agricultural Economics*, 47(S1):61–71, 2016.

[11] Lars Carlsson and Annica Sandström. Network governance of the commons. *International Journal of the Commons*, 2(1):33–54, 2008.

[12] Tanya L Chartrand and Jessica L Lakin. The antecedents and consequences of human behavioral mimicry. *Annual review of psychology*, 64:285–308, 2013.

[13] Raymond Chiong and Michael Kirley. Imitation vs evolution: Analysing the effects of strategy update mechanisms in n-player social dilemmas. In *IEEE Congress on Evolutionary Computation*, pages 1–8. IEEE, 2010.

[14] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[15] Peter J Deadman. Modelling individual behaviour and group performance in an intelligent agent-based simulation of the tragedy of the commons. *Journal of Environmental Management*, 56(3):159–172, 1999.

[16] Alain Degenne and Michel Forsé. *Introducing social networks*. Sage, 1999.

[17] Thomas Dietz, Elinor Ostrom, and Paul C Stern. The struggle to govern the commons. *Science*, 302(5652):1907–1912, 2003.

[18] Holger Ebel and Stefan Bornholdt. Coevolutionary games on networks. *Physical Review E*, 66(5):056118, 2002.

[19] Holger Ebel, Lutz-Ingo Mielsch, and Stefan Bornholdt. Scale-free topology of e-mail networks. *Physical review E*, 66(3):035103, 2002.

[20] Víctor M Eguíluz, Martín G Zimmermann, Camilo J Cela-Conde, and Maxi San Miguel. Cooperation and the emergence of role differentiation in the dynamics of social networks. *American journal of sociology*, 110(4):977–1008, 2005.

[21] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[22] Carlos Gracia-Lázaro, José A Cuesta, Angel Sánchez, and Yamir Moreno. Human behavior in prisoner's dilemma experiments suppresses network reciprocity. *Scientific reports*, 2:325, 2012.

[23] Carlos Gracia-Lázaro, Alfredo Ferrer, Gonzalo Ruiz, Alfonso Tarancón, José A Cuesta, Angel Sánchez, and Yamir Moreno. Heterogeneous networks do not promote cooperation when humans play a prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 109(32):12922–12926, 2012.

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

[25] Benjamin Kerr, Claudia Neuhauser, Brendan JM Bohannan, and Antony M Dean. Local migration promotes competitive restraint in a host–pathogen'tragedy of the commons'. *Nature*, 442(7098):75–78, 2006.

[26] Timothy Killingback and Michael Doebeli. Self-organized criticality in spatial evolutionary game theory. *Journal of theoretical biology*, 191(3):335–340, 1998.

[27] Henry F Lyle and Eric A Smith. The reputational and social network benefits of prosociality in an andean community. *Proceedings of the National Academy of Sciences*, 111(13):4820–4825, 2014.

[28] Jorge Marco and Renan Goetz. Tragedy of the Commons and Evolutionary Games in Social Networks: The Economics of Social Punishment. ETA: Economic Theory and Applications 259486, Fondazione Eni Enrico Mattei (FEEM), July 2017.

[29] Dimitrije Marković and Claudius Gros. Power laws and self-organized criticality in theory and nature. *Physics Reports*, 536(2):41–74, 2014.

[30] Naoki Masuda and Kazuyuki Aihara. Spatial prisoner's dilemma optimally played in small-world networks. *Physics Letters A*, 313(1-2):55–61, 2003.

[31] J Terrence McCabe. Turkana pastoralism: A case against the tragedy of the commons. *Human ecology*, 18(1):81–103, 1990.

[32] Yong Min, Yuchen Du, and Cheng Jin. The effect of link rewiring on a coevolutionary common pool resource game. *Physica A: Statistical Mechanics and its Applications*, 512:935–944, 2018.

[33] Emilio F Moran and Elinor Ostrom. *Seeing the forest and the trees: human-environment interactions in forest ecosystems*. Mit Press, 2005.

[34] Joao A Moreira, Jorge M Pacheco, and Francisco C Santos. Evolution of collective action in adaptive social structures. *Scientific reports*, 3(1):1–6, 2013.

[35] Bryce Morsky and Erol Akcay. Evolution of social norms and correlated equilibria (vol 116, pg 8834, 2019). *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 117(14):8213–8213, 2020.

[36] Rachata Muneepeerakul and John M Anderies. The emergence and resilience of self-organized governance in coupled infrastructure systems. *Proceedings of the National Academy of Sciences*, 117(9):4617–4622, 2020.

[37] Munyaradzi J Mutenje, Gerald F Ortmann, and Stuart RD Ferrer. Management of non-timber forestry products extraction: Local institutions, ecological knowledge and market structure in south-eastern zimbabwe. *Ecological Economics*, 70(3):454–461, 2011.

[38] Martin A Nowak, Sebastian Bonhoeffer, and Robert M May. Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences*, 91(11):4877–4881, 1994.

[39] Martin A Nowak and Robert M May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826, 1992.

[40] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.

[41] Per Olsson, Lance Gunderson, Steve Carpenter, Paul Ryan, Louis Lebel, Carl Folke, and Crawford S Holling. Shooting the rapids: navigating transitions to adaptive governance of social-ecological systems. *Ecology and society*, 11(1), 2006.

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

[43] Elinor Ostrom. Collective action and the evolution of social norms. *Journal of economic perspectives*, 14(3):137–158, 2000.

[44] Elinor Ostrom, Joanna Burger, Christopher B Field, Richard B Norgaard, and David Policansky. Revisiting the commons: local lessons, global challenges. *science*, 284(5412):278–282, 1999.

[45] Julien Perolat, Joel Z Leibo, Vinicius Zambaldi, Charles Beattie, Karl Tuyls, and Thore Graepel. A multi-agent reinforcement learning model of common-pool resource appropriation. In *Advances in Neural Information Processing Systems*, pages 3643–3652, 2017.

[46] Andrew A Rosenberg. Managing to the margins: the overexploitation of fisheries. *Frontiers in Ecology and the Environment*, 1(2):102–106, 2003.

[47] Francisco C Santos and Jorge M Pacheco. Risk of collective failure provides an escape from the tragedy of the commons. *Proceedings of the National Academy of Sciences*, 108(26):10421–10425, 2011.

[48] Amartya Sen, Master Amartya Sen, Sen Amartya, James E Foster, James E Foster, et al. *On economic inequality.* Oxford University Press, 1997.

[49] Rajiv Sethi and Eswaran Somanathan. The evolution of social norms in common property resource use. *The American Economic Review*, pages 766–788, 1996.

[50] Hendrik Santoso Sugiarto, John Stephen Lansing, Ning Ning Chung, CH Lai, Siew Ann Cheong, and Lock Yue Chew. Social cooperation and disharmony in communities mediated through common pool resource exploitation. *Physical review letters*, 118(20):208301, 2017.

[51] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics reports*, 446(4-6):97–216, 2007.

[52] György Szabó and Christoph Hauert. Phase transitions and volunteering in spatial public goods games. *Physical review letters*, 89(11):118101, 2002.

[53] Attila Szolnoki and Matjaž Perc. Antisocial pool rewarding does not deter public cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 282(1816):20151975, 2015.

[54] Edward W Tekwa, Andrew Gonzalez, and Michel Loreau. Spatial evolutionary dynamics produce a negative cooperation–population size relationship. *Theoretical population biology*, 125:94–101, 2019.

[55] Raja Rajendra Timilsina, Koji Kotani, and Yoshio Kamijo. Sustainability of common pool resources. *PloS one*, 12(2):e0170981, 2017.

[56] Daniele Vilone, José J Ramasco, Angel Sánchez, and Maxi San Miguel. Social imitation versus strategic choice, or consensus versus cooperation, in the networked prisoner's dilemma. *Physical Review E*, 90(2):022810, 2014.

[57] Friedrich Burkhard von der Osten, Michael Kirley, and Tim Miller. Sustainability is possible despite greed-exploring the nexus between profitability and sustainability in common pool resource systems. *Scientific reports*, 7(1):2307, 2017.

[58] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440, 1998.

[59] Joshua S Weitz, Ceyhun Eksin, Keith Paarporn, Sam P Brown, and William C Ratcliff. An oscillating tragedy of the commons in replicator dynamics with game-environment feedback. *Proceedings of the National Academy of Sciences*, 113(47):E7518–E7525, 2016.