ARTICLE IN PRESS

Evolution and Human Behavior xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens

Evolving institutions for collective action by selective imitation and self-interested design

Sergey Gavrilets[a,*], Mahendra Duwal Shrestha[b]

[a] Department of Ecology and Evolutionary Biology, Department of Mathematics, National Institute for Mathematical and Biological Synthesis, Center for the Dynamics of Social Complexity, University of Tennessee, Knoxville, TN 37996, USA
[b] Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA

ARTICLE INFO

Keywords:
Punishment
Cooperation
Conflict
Cultural evolution
Design
Cultural group selection

ABSTRACT

Human behavior and collective actions are strongly affected by social institutions. A question of great theoretical and practical importance is how successful social institutions get established and spread across groups and societies. Here, using institutionalized punishment in small-scale societies as an example, we contrast two prominent mechanisms - selective imitation and self-interested design - with respect to their ability to converge to cooperative social institutions. While selective imitation has received a great deal of attention in studies of social and cultural evolution, the theoretical toolbox for studying self-interested design is limited. Recently Perry, Shrestha, Vose, and Gavrilets (2018) expanded this toolbox by introducing a novel approach, which they called foresight, generalizing standard myopic best response for the case of individuals with a bounded ability to anticipate actions of their group-mates and care about future payoffs. Here we apply this approach to two general types of collective action – "us vs. nature" and "us vs. them" games. We consider groups composed by a number of regular members producing collective good and a leader monitoring and punishing free-riders. Our results show that foresight increases leaders' willingness to punish free-riders. This, in turn, leads to increased production and the emergence of an effective institution for collective action. We also observed that largely similar outcomes can be achieved by selective imitation, as argued earlier. Selective imitation by leaders (i.e. cultural group selection) outperforms self-interested design if leaders strongly discount the future. Foresight and selective imitation can interact synergistically leading to a faster convergence to an equilibrium. Our approach is applicable to many other types of social institutions and collective action.

1. Introduction

Cooperating human groups can acquire material benefits that would be completely out of reach (or too costly) for single individuals. For this to happen, however, group members have to be able to effectively coordinate their actions, resolve potential conflicts, and eliminate or minimize free-riding. The collective action problem (i.e., free-riding of group members) is generic for both human and non-human animal groups and can easily undermine within-group cooperation (Hardin, 1982; Olson, 1965; Pecorino, 2015; Sandler, 1992). Collective action problems can be (partially) resolved by several mechanisms including kin cooperating with each other (and gaining compensatory benefits through indirect fitness), direct and indirect reciprocity, punishment, group selection, selective incentives, within-group heterogeneity as well as social norms and social institutions regulating individual and group behavior (Gavrilets, 2015b; Hardin, 1968, 1982; McElreath & Boyd, 2007; North, 1990; Nowak, 2006; Olson, 1965; Ostrom, 2000; Pecorino, 2015; Sandler, 1992).

A question of crucial theoretical and practical importance is how social institutions for collective action become effective and stable. Institutions that regulate social life, including those that reconcile disputes, manage the commons, and ensure norm compliance are ubiquitous and a key feature enabling the success of our species (Alesina & Giuliano, 2015; Miller, 2019; North, 1990; Powers, van Schaik, & Lehmann, 2016; Richerson & Boyd, 2005; Singh, Wrangham, & Glowacki, 2017). Yet, we know little about how these institutions develop. To what extent do they reflect the interests and intentionality of their members? Do the design features reflect just transmission processes such as selective imitation or are they better explained by individual wielding agency over their shape through expectations about the future and the behavior of others?

One powerful method of optimizing individual strategies is random

* Corresponding author.
E-mail address: gavrila@tiem.utk.edu (S. Gavrilets).

https://doi.org/10.1016/j.evolhumbehav.2020.05.007
Received 26 August 2019; Received in revised form 31 March 2020; Accepted 10 May 2020
1090-5138/

Please cite this article as: Sergey Gavrilets and Mahendra Duwal Shrestha, Evolution and Human Behavior, https://doi.org/10.1016/j.evolhumbehav.2020.05.007

innovation coupled with selective imitation by payoff-biased social learning. Under this method, individuals observe and evaluate actions and payoffs of others and adapt strategies resulting in a higher payoff. This mirrors how one typically thinks of biological evolution: a blind process of mutation introduces variation and then natural selection favors mutants with higher fitness. Selective imitation is mathematically analogous to natural selection acting in biological systems but it can operate on much faster time-scales (Richerson et al., 2016). Selective imitation can also drive cultural group selection, resulting in the spread of beneficial institutions across different groups (Richerson et al., 2016; Richerson & Boyd, 2005; Turchin, 2016). Some researchers view cultural group selection as the most important (or even the only) mechanism that can account for institutionalized cooperation in human societies (Chudek, Zhao, & Henrich, 2013; Richerson et al., 2016; Turchin, 2016). However, both humans and non-human animals exhibit bounded rationality (de Waal, 2016; Gigerenzer & Selten, 2001). Moreover the power and usefulness of selective imitation within the context of collective action can be questioned. At the individual level, because free-riders often end up with a higher payoff than cooperators, their strategies are more likely to be copied which would undermine cooperation (Burton-Chellew, El Mouden, & West, 2017; Molleman, van den Berg, & Weissing, 2014; van den Berg, Molleman, & Weissing, 2015). Moreover, because individuals differ in a variety of characteristics, a strategy that is advantageous for one will not necessarily be beneficial or even feasible for another. At the group level, selective imitation of institutions requires a flow of information between (potentially competing) groups and the intimate knowledge of relevant details. Even if these are readily available, institutions might not be transferable "off the shelf" because of differences among groups in their social or ecological environment (Aoki, 2001; Powers et al., 2016; Singh et al., 2017). Additionally, models of selective imitation typically fail to more broadly consider how within-group variation in interests and power constrains the form of emerging institutions (Cofnas, 2018; Singh et al., 2017; Smith, 2020) and pay only cursory attention to how new rules emerge treating innovation as pretty much a random process.

An alternative view emphasizes the power of within-group design processes driven by the motivation of the whole group or some of its segments to increase their material payoffs or some more general utility. For example, Ostrom (1990) has identified a number of "design principles" for stable and successful management of common resources by local communities. Early eighteenth-century pirates designed democratic institutions (with constitutions, separation of power, and checks and balances) making pirate predatory groups very efficient (Defoe, 1724; Leeson, 2009). Similar examples are known among contemporary prison gangs (Skarbek, 2012). Singh et al. (2017) argue for the importance of self-interested design in the creation of institutions and put forward a "self-interested enforcement" hypothesis, which proposes that many observed group-level traits and institutions reflect the differences in relative enforcement capabilities of different group segments. We note that the idea of self-interested design also captures key aspects of human sociality – that we can in fact take guesses about the future and the future behavior of our peers.

One way to contrast selective imitation and self-interested design as the mechanisms of social evolution is through mathematical modeling. Inspired by the work of Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985) (see also earlier work by Rashevsky, 1949, 1951, 1965a, 1965b, Landahl, 1950, Truco, 1954), there is now a great diversity of mathematical models and theoretical approaches dealing with social learning and imitation. In contrast, there is no established mathematical framework for modeling the evolution of group-level traits and institutions by self-interested design. A powerful and well studied method of optimizing individual behavior is myopic best response (Sandholm, 2010), which is an example of the bounded rationality approach (Gigerenzer & Selten, 2001). Under this method, individuals attempt to optimize their behavior under the assumption that everybody else keeps their strategies. If each group member is using myopic optimization, the group can end up at a Nash equilibrium (Hofbauer & Sandholm, 2002; Xu, 2016). However, myopic best response can fail in social dilemmas or when there is a collective action problem, because self-interested individuals will be motivated to free-ride on the effort of others.

Recently Perry et al. (2018), Perry and Gavrilets (2020), Gavrilets (2020) introduced a novel strategy updating method which generalizes myopic best response for individuals with a bounded ability to i) anticipate future actions of their group-mates and ii) consider their effects on future payoffs. The method, which we called (one-step) foresight, attempts to capture some aspects of human decision-making which are well established empirically. One such aspect is the "theory of mind", i.e. the ability to reason about the knowledge and thought processes of others in the social context (Premack & Wodruff, 1979). The "theory of mind" exists in humans (Tomasello, Carpenter, Call, Behne, & Moll, 2005) and also appears to exist in great apes (de Waal, 2016; Krupenye, Kano, Hirata, Call, & Tomasello, 2016). Foresight with respect to the effects of punishment is also well established in experimental studies of cooperation as a very powerful driver of individual behavior: the threat of punishment immediately makes subjects more cooperative (Fehr & Gächter, 2002; Spitzer, Fischbacher, Herrnberger, Grön, & Fehr, 2007) and subjects expect that individuals who were punished earlier will be more cooperative in the future (Krasnow, Cosmides, Pedersen, & Tooby, 2012; see also Axelrod, 1986's discussion of "deterrence" as a mechanism for maintaining cooperative norms). More generally, foresight is related to our ability to represent mentally what might happen in the future (captured in the notion of prospection, Szpunar, Spreng, & Schactera, 2014). Humans are routinely engaged in making intertemporal choices when they have to trade off costs and benefits at different points in time (Berns, Laibson, & Loewenstein, 2007; Frederick, Loewenstein, & O'Donoghue, 2002). Intertemporal choices imply a degree of self-control (Hayden, 2019), which is also found in other animals (MacNulty, Tallian, Stahler, & Smith, 2014; Miller, 2019). Consideration of future interactions and their impact on individual payoffs are important in many game-theoretic models (e.g., Axelrod, 1984; Jehiel, 1995, 2001; O'Donoghue & Rabin, 1999, 2001; Sandholm, 2010). Similarly there exist theoretical approaches aiming to capture humans' ability to predict behavior of others, such as "beauty contests" games (Duffy & Nagel, 1997) and level-k and cognitive hierarchy models (Nagel, 1995a; Stahl & Wilson, 1995b). Foresight is a simple way to bring these two aspects together in application to cooperation and punishment.

Foresight was initially introduced within the context of collective action in heterogeneous groups in the presence of peer punishment (Perry et al., 2018). In particular, we showed that foresight can allow groups to overcome the first- and second-order free-riding problems leading to successful cooperation. We also demonstrated the emergence of a division of labor in which some individuals (e.g., more powerful) specialized in punishment while others (e.g., less powerful) mostly contributed to the production of collective goods. The power of foresight in ensuring cooperation via peer punishment suggests that it may also promote the evolution of social institutions.

Here our focus will be on one particularly important example of early social institutions which is leadership (Dogan, Glowacki, & Rusch, 2018; Gächter & Renner, 2018; Garfield, Hubbard, & Hagen, 2019; Glowacki & von Rueden, 2015; Hooper, Kaplan, & Boone, 2010; Smith et al., 2016; Wiessner, 2019). Leaders can coordinate the actions of group members making their efforts more efficient, monitor and punish free-riders, reward contributors, and foster pro-social norms and values. Leaders and followers emerge naturally as a result of heterogeneity in preferences, motivation, personality, physical characteristics, information available, and other features affecting individual performance in different activities (Gavrilets, Auerbach, & van Vugt, 2016; Perry et al., 2018; Smith et al., 2016). In some small-scale societies, leaders get an equal share of the collective goods produced by the group while in others they get extra benefits (Garfield et al., 2019; Glowacki & von

Rueden, 2015; Smith et al., 2016). Leadership can be informal or institutionalized, e.g., when there are rules or shared expectations establishing who is a leader, what they do, and what their privileges are.

Withing the context of collective action, the evolution of leadership in small-scale societies was modeled by Hooper et al. (2010) and Powers and Lehmann (2013, 2014). Isakov and Rand (2012) and Roithmayr, Isakov, and Rand (2015) studied theoretically institutionalized punishment in large-scale societies. Perry and Gavrilets (2020) modeled interactions between a subordinate producing a collective good and a leader tasked with monitoring and punishing the subordinate.

Here we build on these earlier approaches to study collective actions under institutionalized punishment in small-scale societies. Specifically, we will assume that the division of labor between leaders who identify and punish cheaters and the rest of the group who produce collective goods is already established and collectively endorsed (Garfield et al., 2019, Wiessner, 2019) and will study its evolution. Following our earlier work (Gavrilets, 2015a, 2015b; Gavrilets & Richerson, 2017; Perry et al., 2018; Perry & Gavrilets, 2020), we will consider two general types of collective action - "us vs. nature" and "us vs. them" (see below). Our main goal is to compare selective imitation and self-interested design (implemented via foresight) with respect to their ability to identify and converge to cooperative social institutions.

2. Models and results

Consider a population comprised of a number of groups. Each group has n regular members and an additional entity which we will call a "leader". (The "leader" does not have to be a single individual but can be a group, such as elders in a small-scale society, or a formal institution.) Regular group members have an opportunity to participate in collective actions producing shared benefits. Leaders monitor their efforts, punish free-riders, and collect tax.

We will consider separately and contrast two types of mathematical models aiming to describe two most general kinds of collective action that early human groups were most definitely engaged in: "us vs. nature" games and "us vs. them" games (Gavrilets, 2015a, 2015b; Gavrilets & Fortunato, 2014), Whitehouse et al., 2017, Gavrilets & Richerson, 2017). The former describe collective actions such as defense from predators, cooperative hunting, cooperative breeding, habitat improvements, building dams or fences to drive animals, etc. The success of a particular group in solving these problems does not depend much on the actions of neighboring groups. As groups become more cooperative, their payoffs usually rise (Kropotkin, 1902). In contrast, "us vs. them" games describe direct conflicts and/or other costly competition with other groups over territory, mating opportunities, access to trade routes, etc. (Darwin, 1871). The success of one groups in an "us vs. them" game against less cooperative groups means failure or reduced success for other competing groups. For each group, becoming more cooperative does not necessarily means higher payoff in the long run as other groups respond in kind (Konrad, 2009; Tullock, 1980). Such an "arms race" will stop once the costs become too high. "Us vs. them" games are more conducive for the evolution of cooperation than "us vs. nature" games but can result is the waste of resources (Gavrilets, 2015a, 2015b).

2.1. Basic model: groups without leaders

To get a better intuition about our general model, we first consider groups without leaders. We assume an individual's effort in a collective action (specified by a binary variable $x = 0$ or 1) is costly while any benefit produced and retained by a group is shared equally among all of them; this creates an incentive to free-ride (Olson, 1965). Without leaders, the payoff of an individual making effort x in a collective action is

$$\pi_s(x) = bP - cx, \tag{1}$$

where b and c are the benefit and cost parameters. The function P gives the normalized value of the resource produced or secured by the group. Let $X = \Sigma x$ be the total group effort. In "us vs. nature" games, we define $P = X/(X + X_0)$, where X_0 is a half-success parameter (Gavrilets, 2015a, 2015b). If $X = X_0$, the probability of group success P is equal to one half. The larger X_0, the more group effort is required to secure the reward. [A linear public goods game is a special case of model (1) with $P = X/X_0$, where $X_0 = n$.] In "us vs. them" games, we define $P = X/\overline{X}$, where $\overline{X}$ is the average group effort over all G competing groups in the system. Note that "us vs. nature" games are a special case of the generalized Volunteer's Dilemma (Archetti, 2009; Diekmann, 1985) while "us vs. them" games are common in the theory of between-group contests (Konrad, 2009; Rusch & Gavrilets, 2017).

In the Supplementary Information (SI), we provide details on group behavior in these two models (see also Gavrilets, 2015a, 2015b; Gavrilets & Fortunato, 2014). In "us vs. nature" games, groups always evolve to an equilibrium at which the group effort can be approximates as

$$X^* = X_0(\sqrt{R} - 1), \tag{2}$$

if $R \equiv b/(cX_0) > 1$, and is 0 otherwise. Note that R is the ratio of the individual benefit b to the group cost cX at half-success effort $X = X_0$. In this model, groups cooperate only if R is sufficiently large. In "us vs. them" games, the equilibrium value of the group effort can be approximated

$$X^* = \frac{G-1}{G}\ \frac{b}{c}, \tag{3}$$

so that the group effort is always positive. [Numerical simulations of "us vs. them" games also show a possibility of non-equilibrium dynamics at which the average effort is close to the one predicted by Eq. (2b); see the SI]. In both types of games, in general each group is a mixture of contributing and free-riding individuals; groups size n has no effect on group effort X^*, which however increases with the benefit-to-cost ratio b/c, as expected. [Of course, X^* cannot exceed n.]

2.2. Full model: institutionalized punishment

Next we consider the full model with leaders added to groups. The collectively endorsed role of leaders is to identify and punish free-riders. That is, punishment in our model works via the institution of leadership (Glowacki & von Rueden, 2015; Hooper et al., 2010; Isakov & Rand, 2012; Roithmayr et al., 2015) rather than been administered by peers (Boyd & Richerson, 1992; Hauert, Traulsen, Brandt, Nowak, & Sigmund, 2007).

We assume that a leader makes a costly monitoring effort y ($0 \le y \le 1$). As a result of this effort each free-rider in the group is identified with probability y. The leader then punishes each identified free-rider by reducing their payoff by κ at a cost δ to the leader. We define the leader's cost of monitoring as c_yny, i.e. it grows linearly with the group size; c_y is a cost of monitoring parameter. [We note that realistically in small-scale societies the costs of monitoring and collectively endorsed punishment can be low.] The leader's benefit comes from a constant tax ρ which they collect from the group's production. Our model of institutionalized punishment can be viewed as a multiplayer extension of the inspection game (Fudenberg & Tirole, 1992; Perry & Gavrilets, 2020).

In this model, the expected payoff of a regular member is

$$\pi_s(x,y,X) = (1-\rho)bP(X) - cx - \kappa y(1-x), \tag{3a}$$

where the last term is the expected cost of being punished. The expected payoff of the leader is

$$\pi_l(y,X) = \rho nbP(X) - c_y ny - \delta(n-X)y, \tag{3b}$$

where the first term is the tax collected from n regular members and the last term is the expected cost of punishing $(n - X)y$ identified free-riders.

Given a fixed level of monitoring y, equilibrium values of the group effort X can be approximated using the results on groups without leaders as the presence of a leader merely decreases the benefit and cost terms from b and c in Eq. (1) to $\tilde{b} = b(1 - \rho)$ and $\tilde{c} = c - \kappa y$ (compare Eqs. (10) and (3a)). That is, the presence of leaders effectively decreases the benefit and cost terms for regular members. Therefore the latter can be motivated to produce given a sufficiently high level of monitoring y. In contrast, for leaders, π_l always decreases with y, so they will chose a zero effort. As a result, the only equilibrium in this model is the state with no production and no monitoring (c.f. Perry & Gavrilets, 2020). This effect is analogous to a well-known second-order free-rider problem in models of peer-punishment (Boyd, Gintis, Bowles, & Richerson, 2003; Boyd & Richerson, 1992).

To study the dynamics of the full model with changing efforts x and y we use stochastic agent-based simulations. We assume time to be discrete and focus on a sequence of collective actions occurring synchronously in all groups.

2.2.1. Strategy revision

After each collective action, each individual is independently given an opportunity to revise their efforts x and y with probability q. We allow for random innovation (analogous to random mutation in biology), selective imitation, and self-interested optimization at rates E_1, E_2, and E_3 per time step per individual, respectively ($E_1 + E_2 + E_3 = q$). We also allow for errors in decision-making (see below). We assume that under selective imitation individuals can imitate their peers: regular members can imitate other regular members in their own or other groups; leaders can imitate other leaders. We consider two self-interested optimization strategies. The first is the standard myopic best response (Hofbauer & Sandholm, 2002; Sandholm, 2010). The second self-interested optimization strategy is one-step foresight (Gavrilets, 2020; Perry et al., 2018; Perry & Gavrilets, 2020) which we describe next.

2.2.2. Foresight

Here for simplicity we will apply foresight only to leaders. Our justification is that leaders have more information and more power in using it than regular members. [Perry & Gavrilets, 2020 showed in a simpler model that allowing for one-step foresight in regular members does not have any effect.] The foresight mechanism includes two components: i) consideration of future benefits and ii) the forecast of actions of others.

With respect to the former, the idea is that a major goal of punishment is often to modify the transgressor's future behavior (Axelrod, 1986; Cushman, 2015; Ellsworth & Ross, 1983; Krasnow et al., 2012). This suggests that expected *future* payoffs are usually a part of the punisher's utility function. In our implementation of (one-step) foresight for leaders, the leaders attempt to maximize their utility function u_l, which we define as a sum of the expected payoff $\pi_l(y,X)$ after the current round of strategy updates and the forecasted payoffs after the next round of strategy updating $\pi_l(y',X')$:

$$u_l = (1 - \omega)\ \pi_l(y, X) + \omega\ \pi_l(y', X'), \tag{4a}$$

where $0 \leq \omega \leq 1$ is an exogenous constant parameter weighting the importance of future payoffs. Research in behavioral economics shows that people usually discount future payoffs (Frederick et al., 2002; O'Donoghue & Rabin, 1999, 2001) which implies that $\omega \leq 0.5$.

The leader expects that their action y this round will affect their subordinates' effort X' in the next round. At the same time, their y has no effect on the benefit $\rho nbP(X)$ to be produced by the subordinates this round, or the cost of the inspection in the next round, $[c_y n + \delta(n - X']$ y' (see Eq. (3b)). Therefore the leader's utility function (4a) reduces to a sum of the costs of inspection and punishment this round and the benefit next round

$$u_l = (1 - \omega)(-c_y ny - \delta(n - X)y) + \omega\rho nbP(X'). \tag{4b}$$

2.2.3. Forecasting the group's effort

To evaluate utility function (4b), we need to specify the leaders' forecast for the group effort X and X'. Our assumption is that leaders know (from previous experience or previous leaders) how regular group members typically behave in response to a given level of monitoring y. To capture this assumption mathematically we have used two approaches. In the first approach, leaders predict X on the basis of the ESS Eq. (2) appropriately adjusted for the corresponding level of monitoring. [That is, to predict X and X', we use Eq. (2) with b and c substituted for $(1 - \rho)b$, $c - \kappa\tilde{y}$ and $(1 - \rho)b$, $c - \kappa y$, respectively. Here $\tilde{y}$ is the previous monitoring effort of the leader.] In the second approach, instead of using Eq. (2), we pre-compute the average total group effort X as observed in numerical simulations for different values of parameters b, X_0, n, λ and $c = 1$ in the model without leaders. We interpret these functions as capturing the leader's knowledge of subordinates' group behavior. We then adjust X values appropriately for the corresponding level of monitoring (as described above). The numerical results observed were similar for both approaches; the results shown below correspond to the first approach. Note that both the leaders' and regular members' attempts to optimize their actions are subject to stochasticity as described below.

Given a certain level of monitoring y, the regular group members will attempt to optimize their behavior which will lead to a certain group effort $X^*(y)$. (As stated above, there can be multiple equilibria of X.) Then, a leader capable of predicting their group behavior is expected to make a minimum effort still assuring that the current group effort X^* is stable. We can then expect multiple Nash equilibria differing in the amount of production, monitoring, and payoffs (see the SI).

2.2.4. Errors

To deal with errors in decision-making, which are unavoidable in almost all real situations, we use a Quantal Response Equilibrium-like approach (Goeree, Holt, & Palfrey, 2016) with logit errors and a non-negative precision parameter λ (see the SI). If $\lambda = 0$, the agents choose a strategies with uniform probabilities. If $\lambda \to \infty$, the agent always chooses the best response and the dynamics converges to a Nash equilibrium (Hofbauer & Sigmund, 1998; Sandholm, 2010). QRE approach generalizes classical Nash equilibria. Other ways to describe errors are possible and have received considerable attention (e.g., Young, 1998). The advantage of QRE is that in this approach error probabilities depend on error costs.

2.2.5. Simulations

We considered different combinations of various strategy revision methods. Each agent updated its strategy randomly and independently with probability $q = 0.25$ per time step. The innovation probability was fixed at $E_1 = 0.01$. Innovation for regular group members meant flipping x (between 0 and 1). Innovation for leaders was implemented by changing their strategy y to a number drawn randomly and independently from a PERT distribution defined on the unit interval with the mode at the previous value y (see the SI). Selective imitation took place with probability E_2. Under selective imitation the agent compares their payoff with that of a randomly chosen "model" and adopts the model's strategy with a probability dependent on the difference in the payoffs. Self-interested optimization happens at rate E_3. Myopic optimization for regular group members was based on evaluating the expected payoffs of the two strategies ($x = 0$ and $x = 1$) and choosing the one with the higher payoff subject to errors. To implement optimization for leaders (myopic best response or foresight), we first generated a single "candidate strategy" for each leader using the same approach as for innovation. Then the leader evaluates the expected payoff (or utility) of the old and new strategies and chooses the one with the higher payoff (or utility) subject to errors.

Simulations were run for 5000 time steps. Convergence to a stochastic equilibrium was confirmed by visual inspection of trajectories. The equilibrium values were estimated as the averages over the last 50 samples; samples were made every 10 time steps; 20 independent runs for each parameter combination. Characteristic time-scale τ for convergence to an equilibrium was evaluated as the time for the average y to reach half of its equilibrium value for the first time.

In the graphs below, the tax will be specified by parameter $\theta = \rho n/(1 - \rho)$ which is the leader-to-regular member share ratio (e.g., with $\theta = 1$, the leader's and a regular member's shares of the reward are the same; with $\theta = 2$, the leader gets twice as much as a regular member). To better see the effects of leadership and institutionalized punishment, we focused on relatively small values of the benefit parameter b under which groups without leaders would not make any effort in "us vs. nature" games and a relatively low effort in "us vs. them" games.

2.3. Results for the full model

We observed that if both types of agents used only innovation and selective imitation, cooperation and punishment were practically absent. As discussed in the Introduction, with selective imitation, regular group members tend to imitate higher-fitness defectors which suppresses group production and removes the incentives for leaders to contribute. If leaders used innovation and myopic best response, not much punishment happened and the levels of cooperation were relatively low independently of how regular group members updated their strategies. This happens because leaders reduce monitoring and punishment to avoid associated costs.

High levels of monitoring and cooperation were observed when regular group members used myopic best response and leaders employed either selective imitation or foresight. Next we focus on and contrast these two scenarios with respect to equilibrium levels of production by regular group members x, monitoring (and punishment) by leaders y, their average payoffs π_s and π_l as well as the time τ to reach an equilibrium.

Our simulations show that if benefits of cooperation are sufficiently large relative to its cost, both methods of strategy update in leaders result in similarly high monitoring and production. If the benefit of cooperation is sufficiently small relative to its cost, both methods result in the absence of monitoring and production. The differences between the methods mostly reveals itself at intermediate benefits (see the SI). We now consider this situation in more detail.

Fig. 1 illustrates the effects of the foresight parameter ω and relative frequencies E_2 and E_3 of selective imitation and foresight in the leader's decision making for different levels of taxation (and, consequently, inequality). [We also allowed for random innovation at a constant small rate $E_1 = 0.01$.] In each graph, the leftmost set of bars (labeled $\theta = 0$) corresponds to the case of groups with no leaders. In this case, group efforts are practically absent in "us vs. nature" games and are relatively low in "us vs. them" games due to the collective action problem (Gavrilets, 2015a, 2015b; Gavrilets & Fortunato, 2014). Adding leaders capable and allowing for institutionalized punishment leads to the establishment of a certain level of monitoring and punishment accompanied by a significant increase in group effort X in both games.

Then the foresight parameters $\omega = 0.5$ so that there is no discount of future payoffs, relative frequencies of selective imitation E_2 and foresight E_3 have weak effects in both games. As ω decreases, selective imitation leads to higher monitoring by leaders and consequently harsher punishment and more production by subordinates (especially in "us vs. them" games when $E_2 : E3$ ratio is small). In "us vs. nature" games increased monitoring and production lead to higher payoffs for both types of players. In contrast, in "us vs. them" games, we observe "overproduction" and reduced payoffs for subordinates. In "us vs. nature" games, foresight leads to higher monitoring and cooperation than selective imitation if the cost of punishment κ is low. With larger κ, selective imitation results in higher monitoring and cooperation. In "us vs. them" games, if there are differences between selective imitation and foresight, they are manifested in higher monitoring and cooperation under more frequent selective imitation. It thus appears that when leaders are more powerful, selective imitation is a better approach to finding appropriate strategies. In "us vs. nature" games, the speed of convergence to a stochastic equilibrium is usually the fastest when both mechanisms operate at comparable frequencies. In "us vs. them" games, foresight results in faster convergence if $\omega = 0.5$ but is slower convergence if ω is smaller. These are typical results (see the SI.)

One consistent difference is that under selective imitation the actual spread of innovations across the whole population in a particular run can happen more rapidly but there is more variation in the onset of the transition to higher monitoring across different runs (Fig. 3).

Note that an increase in cooperation is observed even when $\theta = 1$, so that the leader and regular group members get an equal share of the reward. Increasing θ does not necessarily increase production but does affect the payoff in an obvious way (i.e. decreases it for commoners and increases it for leaders).

Fig. 2 illustrates the effects of some parameters in "us vs. nature" games. (Similar results are observed for "us vs. them" games, see the SI.) Increasing benefit b and punishment κ and decreasing group size n, the half-effort parameter X_0, and precision parameter λ increase cooperation and punishment. These results are intuitive and similar to those in the models of peer punishment (Boyd et al., 2003; Boyd & Richerson, 1992; Traulsen, Röhl, & Milinski, 2012). The payoffs of leaders increase with the group size n (because large group means larger overall tax). The same happens for regular group members in "us vs. nature" games because with a fixed X_0, a larger group size means a smaller individual effort will be sufficient to result in a specific group effort. In "us vs. them" games, the payoff of regular group members only weakly depends on the group size. Increasing punishment κ increases payoffs in "us vs. nature" games but decreases it in "us vs. them" games where it causes "overproduction" (Figs.S6-S16; cf. Konrad (2009), Gavrilets (2015b)). Increasing taxation θ decreases commoners' production and payoff; decreasing precision λ causes higher efforts (by errors) which is expected.

Intuitively, our results can be understood in the following way. Consider first the case of selective imitation and innovation with no foresight (i.e., $E_3 = 0$). Earlier Isakov and Rand (2012) Roithmayr et al. (2015) studied similar models of institutionalized punishment. They showed that selective imitation can lead to the evolution of punishment if leaders update their strategy at a much slower rate than subordinates. A low update rate prevents leaders from abandoning a costly punishment strategy before subordinates have learned to contribute to avoid punishment. In contrast, in our model of selective imitation, subordinates do not have to learn from others via incremental improvements to adapt but rather they use the best response to the current strategy of the leader. But the overall effect is similar - monitoring and punishment evolve by selective imitation in leaders. Let $X_{BR} = X_{BR}(y)$ be the best response value of the total group effort to a given y. Then selective imitation in leaders will optimize the leader's payoff given by Eq. (3b) with X substituted for X_{BR}:

$$\pi_l(y, X_{BR}) = \rho nbP(X_{BR}) - c_y ny - \delta(n - X_{BR})y, \tag{5}$$

Next consider the case of foresight with no selective imitation (i.e., $E_2 = 0$). For a leader using strategy y who is able to predict the subordinates' total production $X_{FR}(y)$, the leaders utility function u_l is given by Eq. (4b) with X and X' substitute for $X_{FR}(y)$:

$$u_l = (1 - \omega)(-c_y ny - \delta(n - X_{FR})y) + \omega\rho nbP(X_{FR}). \tag{6}$$

Both $\pi_l(y, X_{BR})$ and u_l as functions of y are represented by a difference between a benefit term which asymptotically approaches a fixed limit (ρnb and $\omega\rho nb$, respectively) and a cost terms which increases linearly with y. Both functions can have multiple local maxima. In the case of selective imitation, a local maximum can be found by trial-and-

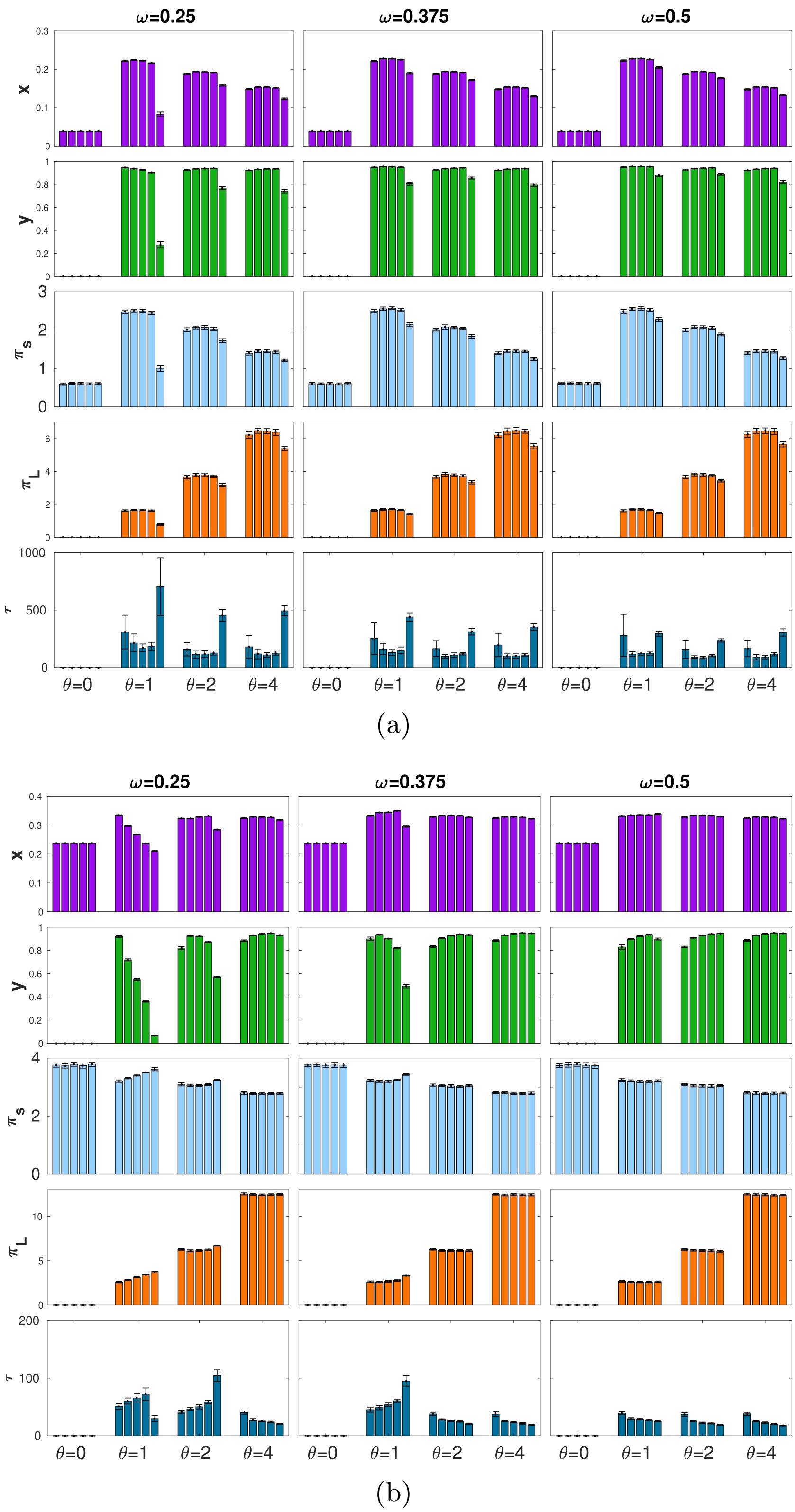


Fig. 1. Average values of efforts x, y, payoffs π_s, π_l and the time to equilibrium τ for 5 different combinations of frequencies of selective imitation E_2 and foresight E_3 in leaders. The frequencies are $E_2 : E_3 = 0.23 : 0.01$ (leaders mostly use selective imitation, left-most bar in each set of five bars), $0.16 : 0.08$,$0.12 : 0.12$,$0.08 : 0.16$ and $0.01 : 0.23$ (leaders mostly use foresight, right-most bar in each set of five bars). The frequency of random mutation $E_1 = 0.01$. (a) "Us vs. nature" games for $K = 1, \lambda = \infty$, $\kappa = 0.5$, $n = 24$, $b = 17$, $X_0 = 24$. (b) "Us vs. them" games for $K = 1$, $\lambda = \infty$, $\kappa = 0.5$, $n = 16$, $b = 4$. Other parameters are at default values (see the SI). Commoners always use innovation at rate $E_1 = 0.01$ and myopic best response at rate $E_3 = 0.24$. Initial values of x and y are assigned randomly and independently; for regular group members, $x = 0$ or 1 with equal probabilities, for leaders, y is chosen from a uniform distribution on [0,0.05].

error and then it can spread across system by imitation. In the case of foresight, a maximum can be discovered by leaders via a process of mental scenario building by considering several candidate strategies and comparing their expected utilities. Assume that $\omega = 0.5$ so that leaders do not discount future. Then if the leader's prediction X_{FR} is the same as the group's best response X_{BR}, the functions in the right-hand side of Eqs. (5) and (6) differ only by a factor 0.5. This explain why our numerical results with $\omega = 0.5$ do not show much difference between selective imitation and foresight in leaders. With smaller ω, leaders using foresight discount future benefits and reduce their monitoring efforts which leads to lower production.

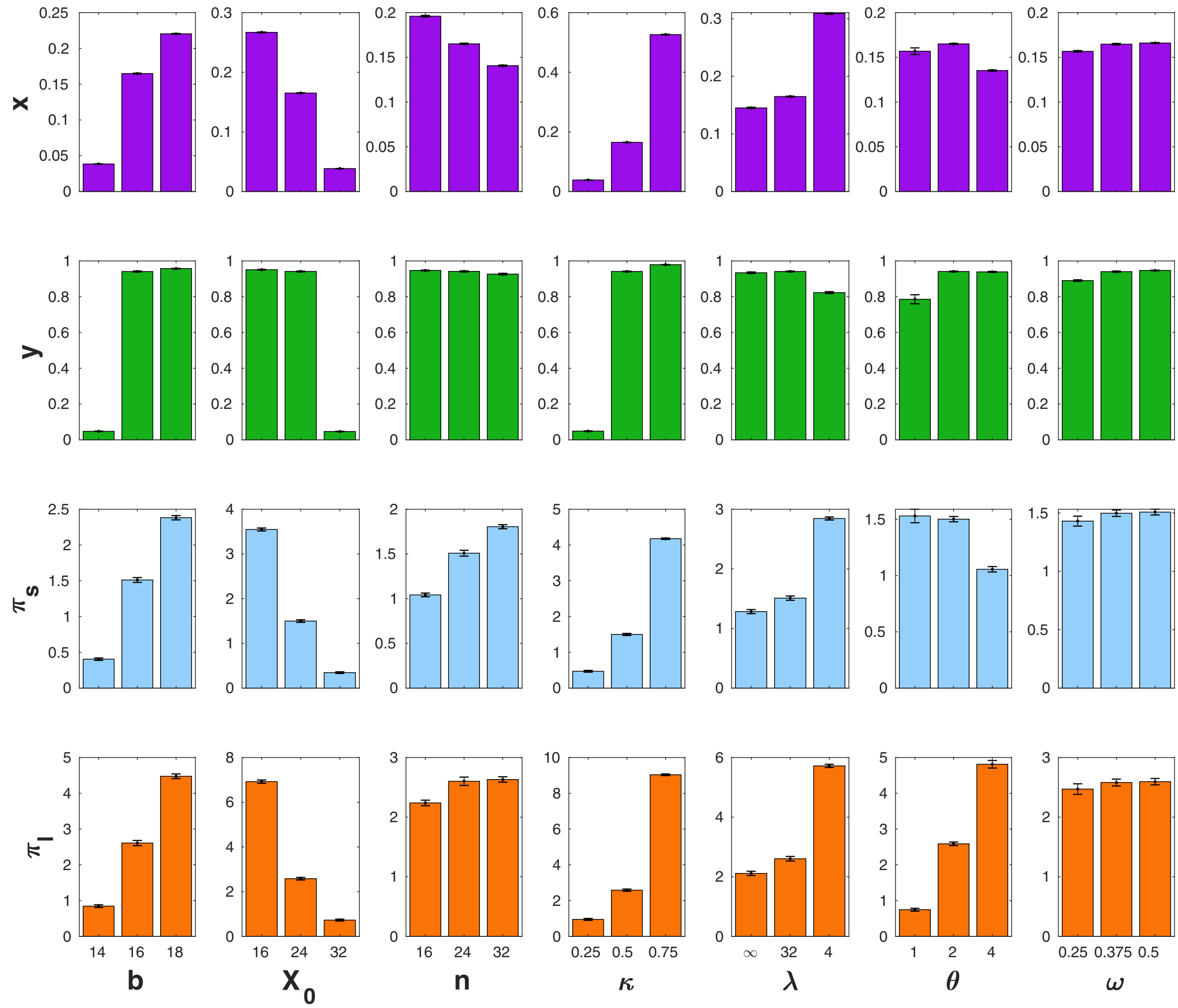


Fig. 2. Effects of parameters benefit b, half-effort X_0, group size n, punishment strength κ, precision λ, tax θ, and the foresight parameter ω on the average efforts of regular group members x and leaders y and their payoffs π_s and π_l in "us vs. nature" games. Parameters are changed one at a time relative to a "default" set with $b = 16$, $n = 24$, $X_0 = 24$, $\theta = 2$, $\kappa = 0.5$, $\lambda = \infty$ and $\omega = 0.375$. Frequencies of updating events: $E_1 = 0.01$, $E_2 = E_3 = 0.12$ in leaders and $E_1 = 0.01$, $E_2 = 0$, $E_3 = 0.24$ in regular group members.

3. Discussion

Here, we modeled social dynamics in groups composed by a number of regular group members investing in the production or appropriation of collective goods and a leader whose collectively endorsed role was to identify and punish free-riders. We postulated that the institution of monitoring and punishment by leaders is already established and studied how it could become efficient. Overall our results provide theoretical support to empirical research in small-scale societies showing that leadership in the form of institutionalized punishment can solve the collective action problem (Garfield et al., 2019; Glowacki & von Rueden, 2015).

We used our models to put the ongoing debate on the relevance and generality of cultural group selection and self-interested design in the evolution of social institutions (Cofnas, 2018; Richerson et al., 2016; Singh et al., 2017; Smith, 2020) on more solid, quantitative grounds. We implemented cultural group selection via standard selective payoff-biased imitation (Richerson, Bettinger, & Boyd, 2005). Our results agree with conclusions from earlier models that selective imitation can lead to effective leadership (Hooper et al., 2010; Isakov & Rand, 2012; Powers & Lehmann, 2013, 2014; Roithmayr et al., 2015). We implemented self-interested design via a novel strategy updating method - foresight.

Foresight generalizes standard myopic best response for the case of individuals caring about their future payoffs and capable of anticipating to a certain extent future actions of their group-mates (Perry et al., 2018; Perry & Gavrilets, 2020). Relative to the myopic best response, foresight is more realistic and, simultaneously, not too taxing on cognitive abilities of individuals and required information. We have shown that foresight makes monitoring and punishment a utility-increasing option (cf. Perry et al., 2018; Perry & Gavrilets, 2020). This, in turn, leads to increased production, cooperation, and the emergence of an effective institution for collective action by self-interested design (Singh et al., 2017). Richerson et al. (2016) questioned the existence of "the alternatives to [cultural group selection that] can easily account for the institutionalized cooperation that characterizes human societies" (p.16). Our results here offer one such alternative.

Our main and unexpected conclusion is that both mechanisms generally lead to very similar outcomes with respect to the levels of cooperation, punishment, and payoffs. If the benefits of cooperation are high (low) enough, the groups cooperate (or not) under both mechanisms. The differences between the mechanisms appear under intermediate benefits and only when the leaders strongly discount future payoffs (i.e. foresight parameter ω is small). The intuition behind is relatively simple, at least in hindside. Assuming best response in regular group members, both strategies update protocols in leaders optimize mathematically similar payoff and utility functions. As a result,

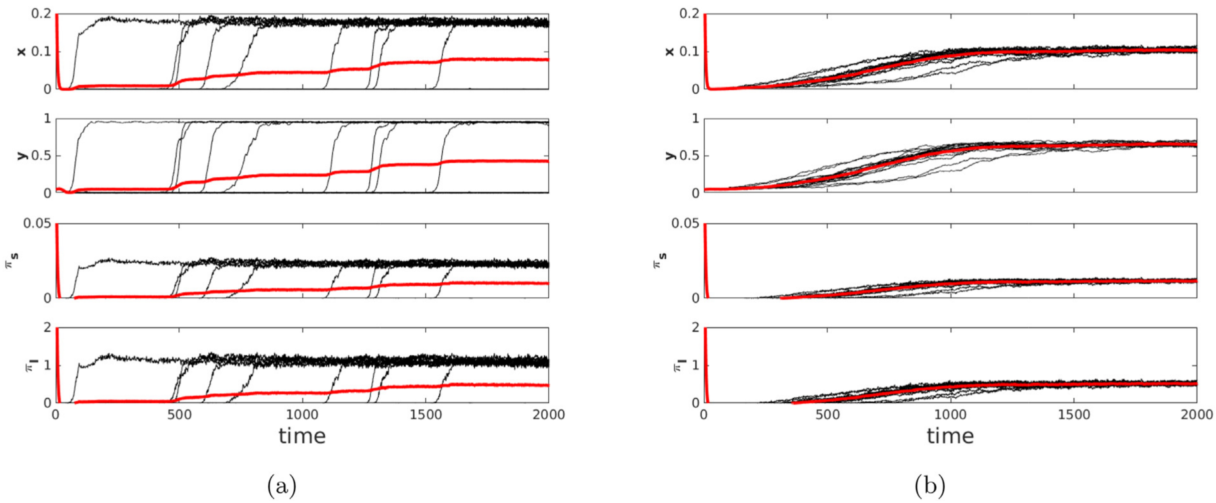


Fig. 3. The dynamics of average values of x, y, π_s and π_l in 20 independent runs. (a) "Us vs. nature" game with $\lambda = \infty$, $\kappa = 0.75$, $b = 7$, $n = 24$, $X_0 = 16$, $\theta = 2$ and mostly selective imitation in leaders ($E_2 = 0.23$, $E_3 = 0.01$). Note that transition to high monitoring by leaders was observed in only 9 runs but if does happen, the transition is very rapid. (b) Same as (a) but with mostly foresight in leaders ($E_2 = 0.01$, $E_3 = 0.23$). Transition to high monitoring by leaders was observed in all 20 runs. Initial values of x (0 or 1) are drawn randomly and independently with equal probabilities. Initial values of y of all leaders are drawn randomly and independently from a uniformly distribution on [0.00,0.05] in 20 runs. Red lines show the averages over 20 runs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

"optimum" solutions are similar. (These optimum solutions are different from the Nash equilibria in the corresponding one-shot games as both selective imitation and foresight change the structure of the corresponding game.) We did observe that under some relatively narrow ranges of parameters there were additional differences between the mechanisms. Specifically, in "us vs. nature" games, when leaders are not too powerful, foresight can lead to higher monitoring and cooperation whereas with more powerful leaders, selective imitation can outperform foresight. In "us vs. them" games, selective imitation can lead to higher monitoring and cooperation. We remind that our conclusions are based on the assumption that regular group members use myopic best response. As discussed above, if they rely exclusively on selective imitation, not much cooperation happens (unless they update their strategies much more often than leaders (Isakov & Rand, 2012, Roithmayr et al., 2015)). That is, in the models considered here successful group cooperation requires self-interested optimization in regular group members.

In the foresight model, leaders are only able to probabilistically forecast their group's response to punishment, but cannot immediately identify long-term optimum strategies. As a result, convergence to such strategies happens only asymptotically. Same behavior is observed under selective imitation. Our results also show that foresight and selective imitation can interact synergistically leading to faster convergence to an equilibrium if leaders use both of them. What seems to happen is that self-interested design leads to a faster establishment of a social innovation in a single group while selective imitation speeds up its spread across other groups.

Although both selective imitation and foresight can result in similar outcomes, their prerequisites differ. Selective imitation is a cognitively simple optimization method based on learning from others with whom the focal agent (i.e. an individual or a group) shares important characteristics (so that the strategy used by the "model" remains feasible and successful for the "mimic"). The agent using selective imitation aims to be as successful as its model. Foresight and, more generally, self-interested design also use social information and learning about behavior of others. However they are not restricted to interactions with similar agents, and agents using them can become more successful than their social partners. Cognitive skills needed for foresight, as modeled here, are not too demanding. Predicting others' behavior requires some "theory of mind" which can be formed on the basis of previous observations or just by asking a question "what would I do if I were in their place". With respect to group traits (such as social institutions), foresight could "work" within a single group. In contrast, selective imitation requires multiple groups, the transfer of relevant information between them, and (cultural) group selection. Take a closer look at the relative efficiency of foresight, selective imitation, fictitious play, and reinforcement in some simple models allowing for analytical, rather than just numerical, investigation.

Here we allowed for random innovation, selective imitation, and two types of self-interested design - myopic best response and foresight. Selective imitation can happen readily for regular group members who can constantly interact with each other. However selective payoff-biased imitation among them will not lead to increased production as higher-payoff individuals are defectors. Selective imitation between leaders does lead to increased production but leaders can only truly learn from other leaders when contact is made. Myopic best response in subordinates can lead to production as subordinates attempt to avoid punishment. However myopic best response in leaders will not lead to monitoring because the leaders will be motivated to avoid its immediate costs. One-step foresight in subordinates does not lead to production (Perry & Gavrilets, 2020). Intuitively, while the myopic best response for subordinates depends on the leaders' strategy, the myopic best response for leaders is always to do nothing. Therefore if subordinates care about the future payoff and assume that leaders use the myopic best response, the subordinates will predict that there will be no monitoring and punishment and thus will choose to ignore the leaders. In contrast, one-step foresight in leaders does lead to increased monitoring and subsequently in production.

In our simulations, we assumed comparable rates of strategy revision for both mechanisms. However imitation of institutions from other groups is likely to be a rarer event than attempts to improve poorly functioning institutions by local "means". This implies that the relative rate of social evolution by cultural group selection will likely be slower than that by self-interested design. If however selective imitation is unconstrained, the timing of adoption of a new effective institution by different groups will be more similar than that under self-interested design because it will spread in an infection-like fashion. (See Fig. 3 showing that transitions from low to high values of inspection frequency y in individual runs are much more rapid under selective imitation.)

Our comparison of "us vs. them" and "us vs. nature" games parallel earlier conclusions: the former games are more conducive for the evolution of cooperation than the latter but they can also easily lead to over-production of individual efforts and wasted payoffs. The effects of parameters on the dynamics of punishment and cooperation are also pretty much in line with intuition and earlier results (Gavrilets, 2015a, 2015b; Gavrilets & Fortunato, 2014; Gavrilets & Richerson, 2017).

An interesting (but not surprising) feature of our models is the existence of multiple equilibria which appear under both selective imitation and self-interested design if the errors in decision-making are small. Multiple equilibria imply significant influence of initial conditions and stochastic exogenous or endogenous (e.g., due to errors in payoff or utility evaluation) events on both transient and long-term dynamics of social institutions.

Earlier work has shown that leaders emerge naturally under many circumstances including leaders who specialize in punishment (Garfield et al., 2019; Perry et al., 2018; Perry & Gavrilets, 2020; Smith et al., 2016). Here we studied the *evolution* of the institute of leadership rather than its *emergence*. In our paper, the evolving part of the institution was the level of monitoring and punishment levels as in Isakov and Rand (2012), Roithmayr et al. (2015). In Hooper et al. (2010), Powers and Lehmann (2014) the evolving trait was the tax imposed by the leaders while in Powers and Lehmann (2013)] it was the proportion of public goods invested into the group's growth rate. Hooper et al. (2010), Powers and Lehmann (2013, 2014) assumed that players culturally inherited their strategies directly from parents (subject to rare random mutation). In Isakov and Rand (2012), Roithmayr et al. (2015) players used payoff-biased imitation. In contrast, we have considered and compared a number of different strategy update methods. We also note that recent experiments strongly points at institutionalized punishment as a more efficient and preferred form of punishment than peer punishment and pool punishment (Fehr & Williams, 2017).

In small-scale societies leadership may or may not be associated with receiving a higher proportion of the group-produced collective goods Glowacki & von Rueden, 2015, Smith et al., 2016, Garfield et al., 2019). In our models, the relevant parameter was θ - the leader's share of the reward relative to that of a regular group member. The effects of θ are nonlinear: under some conditions larger θ leads to more monitoring, punishment, and productions while under other conditions large θ reduce incentives for subordinates to produce collective goods. We have not considered other potentially beneficial effects of leadership such as coordination, norm promotion, or engineering specific religious doctrines incentivizing collective actions – leaving this for future work.

Our strategy updating method foresight is related to several existing theoretical approaches. In particular, consideration of future benefits is common in behavioral economics. Our utility function under foresight (Eq. (4a)) can be viewed as a special simple case of the beta-delta model of hyperbolic discounting (Frederick et al., 2002; O'Donoghue & Rabin, 1999, 2001; Phelps & Pollak, 1968) corresponding to just two time steps. Predicting behavior of peers plays a central role in "beauty contests" games (Duffy & Nagel, 1997) and level-k and cognitive hierarchy models (Nagel, 1995b; Stahl & Wilson, 1995a). In the former, the players attempt to predict the average choice of other players (or some other statistic of interest). In the latter, the players follow a particular system of reasoning which forms a hierarchy. In a simple case, level-0 individuals change their behavior completely randomly. Level-1 individuals assume that all others are of level-0 type and optimize their behavior accordingly. Level-2 individuals assume that all others are of level-1, etc. Such models typically focus on the coexistence of different types of individuals, often within the context of dyadic interactions. Foresight is different from it in two aspects. First, in our model, level-0 individuals do not change their strategies between the rounds and level-1 types use standard myopic optimization. [We note that randomly uniform choice of strategies is not an appropriate level-0 model in studies of cooperation as it would predict 50% cooperation rate.] Individuals employing foresight then corresponds to level-2 reasoning types as they assume that others use myopic best response. Second and most important, we explicitly include future benefits (which are completely ignored in level-k models) in the utility function of level-2 individuals. In classical game theory, players have complete information and are fully rational. This implies that they should be able to predict what exactly will happen (say, plays and payoffs) in any future moment. Relaxing this assumption, Jehiel (1995, 2001) assumed that each player in a dyadic game can predict what exactly will happen over a limited number of steps in the future. A player's prediction of what is to come beyond his horizon of foresight is given by an exogenous noise. In contrast, in our approach players use the theory of mind in an attempt to predict what will happen rather than know this for sure. We also focus on group level behavior and collective actions rather than on dyadic games. Using agent-based simulations, De Weerd, Verbrugge, and Verheij (2013) and Weerd, Verbrugge, and Verheij (2014) have shown that possessing theory of mind is beneficial when participating in dyadic games. In contrast, our work is a step towards understanding how it affects group behavior.

There is a number of important directions for extending our work such as explicitly considering the dynamics of population densities (as in Powers and Lehmann (2013, 2014)), allowing for the simultaneous presence of competition of groups with and without leaders (as in Hooper et al. (2010), Powers and Lehmann (2013, 2014)), and allowing for changeable rather than fixed taxes as well as for a market for leaders (as in Hooper et al. (2010)). Foresight can be applied to any other game with repeated interactions (dyadic or group-level). We assumed only one level of iterated reasoning generalizing myopic optimization. However, many people may use higher-order theories of mind. Note that some theoretical results suggest that using more levels of iterated reasoning does not necessarily result in higher payoffs (De Weerd et al., 2013; Jehiel, 1995, 2001; Weerd et al., 2014). We assumed that individuals attempt to maximize their future material payoff and neglected any effects of past history. An alternative is that players condition their actions on the memory of past events (Press & Dyson, 2012; Stewart & Plotkin, 2012). People are often motivated by social norms and values (Bicchieri, 2006; Gavrilets & Richerson, 2017) and their current actions may be influenced by what happened to them or their groups in the past (Whitehouse et al., 2017). While we have only considered selective imitation by payoff-biased social learning, cooperation can also be maintained by conformity- and prestige-biased social learning (Henrich & Boyd, 1998; Henrich & Gil-White, 2001). These mechanisms could act synergistically with those studied here. All these are important factors that need to be considered in future work.

Human capacity for cultural learning and selective imitation has no doubt greatly contributed both to our uniqueness as a species (Boyd, Richerson, & Henrich, 2011; Henrich, 2016) and to cooperative social institutions we have built (Richerson et al., 2016). However our abilities to innovate and to design and enforce certain rules and social institutions benefiting our societies or some of their segments have also left a significant footprint in our history and will certainly continue to be important in the future (Singh et al., 2017).

Acknowledgements

We thank L. Glowacki, K. Panchanathan, L. Perry, P. J. Richerson, M. Singh and reviewers for comments and suggestions. Supported by the U. S. Army Research Office grants W911NF-14-1-0637 and W911NF-17-1-0150, the Office of Naval Research grant W911NF-18-1-0138, the National Institute for Mathematical and Biological Synthesis through NSF Award #EF-0830858, and by the University of Tennessee, Knoxville.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://

doi.org/10.1016/j.evolhumbehav.2020.05.007.

References

Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature, 53*, 898–944.

Aoki, K. (2001). Theoretical and empirical aspects of gene-culture coevolution. *Theoretical Population Biology, 59*, 253–261.

Archetti, M. (2009). Cooperation as a volunteer's dilemma and the strategy of conflict in public goods games. *Journal of Evolutionary Biology, 22*, 2192–2200.

Axelrod, R. (1984). *The evolution of cooperation.* New York, NY: Basic Books.

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review, 80*(4), 1095–1111.

van den Berg, P., Molleman, L., & Weissing, F. J. (2015). Focus on the success of others leads to selfish behavior. *Proceedings of the National Academy of Sciences USA, 112*(9), 2912–2917.

Berns, G. S., Laibson, D., & Loewenstein, G. (2007). Intertemporal choice - toward an integrative framework. *Trends in Cognitive Sciences, 11*(11), 482–488.

Bicchieri, C. (2006). The grammar of society. *The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences USA, 100*, 3531–3535.

Boyd, R., & Richerson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology, 13*, 171–195.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process.* Chicago: University of Chicago Press.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioral Ecology and Sociobiology, 65*(3, SI), 431–444.

Burton-Chellew, M. N., El Mouden, C., & West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society of London B, 284*(1853).

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach.* Princeton, NJ: Princeton University Press.

Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation and the shaping of human social psychology. In R. Joyce, K. Sterelny, & B. Calcott (Eds.). *Signaling, commitment, and emotion* (pp. 425–458). Cambridge: MIT Press.

Cofnas, N. (2018). Power in cultural evolution and the spread of prosocial norm. *The Quarterly Review of Biology, 93*, 297–318.

Cushman, F. (2015). Punishment in humans: From intuitions to institutions. *Philosophy Compass, 10*(2), 117–133.

Darwin, C. (1871). *The descent of man, and selection in relation to sex.* London: John Murray.

De Weerd, H., Verbrugge, R., & Verheij, B. (2013). How much does it help to know what she knows you know? An agent-based simulation study. *Artificial Intelligence, 199–200*, 67–92.

Defoe, D. (1724). *A general history of the pyrates.* Mineola: Dover.

Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution, 29*, 605–610.

Dogan, G., Glowacki, L., & Rusch, H. (2018). Spoils division rules shape aggression between natural groups. *Nature Human Behavior, 2*, 322–326.

Duffy, J., & Nagel, R. (1997). On the robustness of behaviour in experimental "beauty contest" games. *The Economic Journal, 1075*, 1684–1700.

Ellsworth, P. C., & Ross, L. (1983). Public opinion and capital punishment: A close examination of the views of abolitionists and retentionists. *Crime & Delinquency, 29*(1), 116–169.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in human. *Nature, 415*, 137–140.

Fehr, E., & Williams, T. (2017). *Creating an efficient culture of cooperation.* Department of Economics, University of Zurich page Working Paper No. 267.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40*, 351–401.

Fudenberg, D., & Tirole, J. (1992). *Game theory.* Cambride, MS: The MIT Press.

Gächter, S., & Renner, E. (2018). Leaders as role models and "belief managers" in social dilemmas. *Journal of Economic Behavior and Organization, 154*, 321–334.

Garfield, Z. H., Hubbard, R. L., & Hagen, E. H. (2019). Evolutionary models of leadership: Tests and synthesis. *Human Nature, 30*, 23–58.

Gavrilets, S. (2015a). Collective action and the collaborative brain. *Royal Society Interface, 12*, Article 20141067.

Gavrilets, S. (2015b). Collective action problem in heterogeneous groups. *Philosophical Transactions of the Royal Society of London B, 370*, Article 20150016.

Gavrilets, S. (2020). Foresight, punishment, and cooperation. In M. J. Gelfand, C. Y. Chiu, & Y. Y. Hong (Eds.). *Advances in culture and psychology. Vol.* New York, NY: Oxford University Press.

Gavrilets, S., Auerbach, J., & van Vugt, M. (2016). Convergence to consensus in heterogeneous groups and the emergence of informal leadership. *Scientific Reports, 6* (article number: 29704).

Gavrilets, S., & Fortunato, L. (2014). A solution to the collective action problem in between-group conflict with within-group inequality. *Nature Communications, 5*. https://doi.org/10.1038/ncomms4526) article 3526.

Gavrilets, S., & Richerson, P. J. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences of the United States of America, 114*, 6068–6073.

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality : The adaptive toolbox.* Cambridge, Mass: MIT Press.

Glowacki, L., & von Rueden, C. (2015). Leadership solves collective action problems in small-scale societies. *Philosophical Transactions of the Royal Society of London B, 370*, Article 20150010.

Goeree, J., Holt, C., & Palfrey, T. (2016). *Quantal response equilibrium: A stochastic theory of games.* Princeton, NJ: Princeton University Press.

Hardin, G. (1968). Tragedy of commons. *Science, 162*(3859), 1243–1248.

Hardin, R. (1982). *Collective action.* Baltimore: John Hopkins University Press.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., & Sigmund, K. (2007). Via freedom to coercion: The emergence of costly punishment. *Science, 316*(5833), 1905–1907.

Hayden, B. Y. (2019). Why has evolution not selected for perfect self-control? *Philosophical Transactions of the Royal Society of London B, 374*(1766).

Henrich, J. (2016). *The secrete of our success.* Princeton, NJ: Princeton University Press.

Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior, 19*, 215–241.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*, 165–196.

Hofbauer, J., & Sandholm, W. (2002). On the global convergence of stochastic fictitious play. *Econometrica, 70*(6), 2265–2294.

Hofbauer, J., & Sigmund, K. (1998). *Evolutionary games and population dynamics.* Cambridge: Cambridge University Press.

Hooper, P. L., Kaplan, H. S., & Boone, J. L. (2010). A theory of leadership in human cooperative groups. *Journal of Theoretical Biology, 265*(4), 633–646.

Isakov, A., & Rand, D. (2012). The evolution of coercive institutional punishment. *Dynamic Games and Applications, 2*, 97–109.

Jehiel, P. (1995). Limited horizon forecast in repeated alternate games. *Journal of Economic Theory, 67*(2), 497–519.

Jehiel, P. (2001). Limited foresight may force cooperation. *Review of Economic Studies, 68*(2), 369–391.

Konrad, K. (2009). *Strategy and dynamics in contests.* Oxford: Oxford University Press.

Krasnow, M. M., Cosmides, L., Pedersen, E. J., & Tooby, J. (2012). What are punishment and reputation for? *PLoS One, 7*(9).

Kropotkin, P. (1902). *Mutual aid: A factor of evolution.* London: Heinemann.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science, 354*, 110–114.

Landahl, H. D. (1950). Mathematical theory of imitative behavior in a social group with finite imitation threshold. *Bulletin of Mathematical Biophysics, 12*, 207–213.

Leeson, P. (2009). *The invisible hook: The hidden economics pf pirates.* Princeton: Princeton University Press.

MacNulty, D. R., Tallian, A., Stahler, D. R., & Smith, D. W. (2014). Influence of group size on the success of wolves hunting bison. *PLoS One, 9*(11).

McElreath, R., & Boyd, R. (2007). *Mathematical models of social evolution. A guide for the perplexed.* Chicago: Chicago University Press.

Miller, S. (2019). Social institutions. In E. N. Zalta (Ed.). *The stanford encyclopedia of philosophy*(2019 ed.). Metaphysics Research Lab, Stanford University summer.

Molleman, L., van den Berg, P., & Weissing, F. J. (2014). Consistent individual differences in human social learning strategies. *Nature Communications, 5*.

Nagel, R. (1995a). Unraveling in guessing games: An experimental study. *The American Economic Review, 85*(5), 1313–1326.

Nagel, R. (1995b). Unraveling in guessing games: An experimental study. *The American Economic Review, 85*, 1313–1326.

North, D. C. (1990). *Institution, institutional change and economic performance.* NY: Cambridge University Press.

Nowak, M. (2006). *Evolutionary dynamics.* Harvard: Harvard University Press.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*, 103–124.

O'Donoghue, T., & Rabin, M. (2001). Choice and procrastination. *Quarterly Journal of Economics, 116*, 121–160.

Olson, M. (1965). *Logic of collective action: Public goods and the theory of groups.* Cambridge, MA: Harvard University Press.

Ostrom, E. (1990). *Governing the commons. The evolution of institutions for collective action.* Cambridge: Cambridge University Press.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives, 14*, 137–158.

Pecorino, P. (2015). Olson's logic of collective action at fifty. *Public Choice, 162*, 243–262.

Perry, L., & Gavrilets, S. (2020). Foresight in the game of leadership. *Scientific Reports, 10*, Article 2251.

Perry, L., Shrestha, M. D., Vose, M. D., & Gavrilets, S. (2018). Collective action problem in heterogeneous groups with punishment and foresight. *Journal of Statistical Physics, 172*, 293–312 https://link.springer.com/article/10.1007/s10955–018–2012–2.

Phelps, E. S., & Pollak, R. (1968). On second-best national saving and game- equilibrium growth. *Review of Economic Studies, 35*, 186–199.

Powers, S., & Lehmann, L. (2013). The co-evolution of social institutions, demography, and large-scale human cooperation. *Ecology Letters, 16*, 1356–1364.

Powers, S., & Lehmann, L. (2014). An evolutionary model explaining the neolithic transition from egalitarianism to leadership and despotism. *Philosophical Transactions of the Royal Society of London B, 281*, 20141349.

Powers, S. T., van Schaik, C. P., & Lehmann, L. (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. *Philosophical Transactions of the Royal Society of London B, 371*, Article 20150098.

Premack, D., & Wodruff, G. (1979). Does the chimpanzee have a theory of mind. *Behavioral and Brain Sciences, 1*, 515–526.

Press, W., & Dyson, F. (2012). Iterated Prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences USA, 109*, 10409–10413.

Rashevsky, N. (1949). Mathematical biology of social behavior: II. *Bulletin of Mathematical Biophysics, 11*, 255–271.

Rashevsky, N. (1951). *Mathematical biology of social behavior.* Chicago: University of Chicago Press.

Rashevsky, N. (1965a). A note on imitative behavior. *Bulletin of Mathematical Biophysics, 27*, 311–313.

Rashevsky, N. (1965b). On imitative behavior. *Bulletin of Mathematical Biophysics, 27*, 175–185.

Richerson, P., J., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., ... Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences, 39* article number UNSP e30.

Richerson, P. J., Bettinger, R. L., & Boyd, R. (2005). Evolution on a restless planet: Were environmental variability and environmental change major drivers of human evolution. In F. M. Woketits, & F. J. Ayala (Eds.). *Handbook of evolution. Vol. 2. The evolution of living systems* (pp. 223–242). Wiley.

Richerson, P. J., & Boyd, R. (2005). Not by genes alone. *How culture transformed human evolution*. Chicago: University of Chicago Press.

Roithmayr, D., Isakov, A., & Rand, D. (2015). Should law keep pace with society? Relative update rates determine the co-evolution of institutional punishment and citizen contributions to public goods. *Games, 6*(2), 124–149.

Rusch, H., & Gavrilets, S. (2017). The logic of animal intergroup conflict: A review. *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2017.05.004.

Sandholm, W. H. (2010). *Population games and evolutionary dynamics.* Cambridge, Massachusetts: MIT Press.

Sandler, T. (1992). *Collective action: Theory and applications.* Ann Arbor: University of Michigan Press.

Singh, M., Wrangham, R., & Glowacki, L. (2017). Self-interest and the design of rules. *Human Nature, 28*, 45–480.

Skarbek, D. (2012). Prison gangs, norms, and organizations. *Journal of Economic Behavior & Organization, 82*, 702–716.

Smith, D. (2020). Cultural group selection and human cooperation: A conceptual and empirical review. *Evolutionary Human Sciences, 2*, Article e2.

Smith, J. E., Gavrilets, S., Mulder, M. B., Hooper, P. L., ElMouden, C., Nettle, D., ... Smith, E. A. (2016). Leadership in mammalian societies: Emergence, distribution, power, and payoff. *Trends in Ecology & Evolution, 31*, 54–66.

Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007). The neural signature of social norm compliance. *Neuron, 56*, 185–196.

Stahl, D., & Wilson, P. (1995a). On players models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10*, 218–254.

Stahl, D. O., & Wilson, P. W. (1995b). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10*(1), 218–254.

Stewart, A. J., & Plotkin, J. B. (2012). Extortion and cooperation in the Prisoner's dilemma. *Proceedings of the National Academy of Sciences USA, 109*, 10134–10135.

Szpunar, K. K., Spreng, R. N., & Schactera, D. L. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences USA, 111*, 18414–18421.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(05), 675–691.

Traulsen, A., Röhl, T., & Milinski, M. (2012). An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proceedings of the Royal Society of London B, 279*, 3716–3721.

Truco, E. (1954). Studies in imitative behavior: A generalization of the Rashevsky model; its mathematical properties. *Bulletin of Mathematical Biophysics, 16*, 279–316.

Tullock, G. (1980). Efficient rent seeking. In J. M. Buchanan, R. D. Tollison, & G. Tullock (Eds.). *Toward a theory of the rent-seeking society* (pp. 97–112). College Station: Texas A & M University.

Turchin, P. (2016). *Ultrasociety: How 10,000 years of war made humans the greatest Cooperators on earth.* Beresta Books.

de Waal, F. (2016). *Are we smart enough to kow how smart animals are?* New York: W W Norton and Company.

Weerd, H. D., Verbrugge, R., & Verheij, B. (2014). *Theory of mind in the mod game: An agent-based model of strategic reasoning.* ECSI128–136.

Whitehouse, H., Jong, J., Buhrmester, M. D., Gómez, Á., Bastian, B., Kavanagh, C. M., ... Gavrilets, S. (2017). The evolution of extreme cooperation via shared dysphoric experiences. *Scientific Reports, 7*, 44292 article.

Wiessner, P. (2019). Collective action for war and for peace. A case study among the Enga of Papua New Guinea. *Current Anthropology, 60*.

Xu, Z. (2016). Convergence of best-response dynamics in extensive-form games. *Journal of Economic Theory, 162*, 21–54.

Young, E. P. (1998). *Individual strategy and social structure: An evolutionary theory of institutions.* Princeton, N.J: Princeton University Press.

Supplementary Information for
"Evolving institutions for collective action by selective imitation and self-interested design"

immediate

June 12, 2020

Contents

1 Model's variables, functions and parameters

Table S1 summarizes the model's variables, functions and parameters.

2 Nash, best response, ESS, and QRE equilibria in "us vs. nature" games without leaders

We consider a model in which the payoff to a regular group member making an effort x (= 0 or 1), who belongs to a group making the total effort X, is

$$\pi = b\frac{X}{X + X_0} - cx, \tag{S1}$$

where X_0 is the half-effort parameter.

Table S1: Model variables, functions and parameters.

	Symbols	Their meaning
Variables	x	regular group member's production effort ($x = 0$ or 1)
	y	leader's monitoring effort ($0 \leq y \leq 1$)
Functions	X	total effort of the group, $X = \sum x$
	$P(X)$	normalized value of the resource produced or secured by the group:
		$P = X/(X + X_0)$ in "us vs. nature" games; $P = X/\overline{X}$ in "us vs. them" games
	π_c	regular group member's expected payoff, $\pi_c = (1-\rho)bP(X) - cx - \kappa y(1-x)$
	π_l, π_l', π_l''	leader's expected payoffs, $\pi_l = \rho n b P(X) - c_y n y - \delta(n - X)y$,
		$\pi_l' = -c_y n y' - \delta(n - X')y', \pi_l'' = \rho n b P(X'')$
	u_l	leader's utility with foresight, $u_l = (1-\omega)\pi_l' + \omega\pi_l''$
Parameters	n	number of regular group members per group
	b, c	benefit and cost parameters for regular group members
	X_0	half-effort parameter in "us vs. nature" games
	ρ, θ	tax rate and the leader-to-regular group member share ratio; $\theta = \rho n/(1-\rho)$
	c_y	leader's cost of monitoring parameter
	κ	regular group member's cost of being punished
	δ	leader's cost of punishing a regular group member
	ω	weight of future payoffs in leaders' utility function
	λ	precision parameter in the QRE approach

Nash equilibria. Consider a state where the total group effort $X = 0$. The payoff to each individual is $\pi_0 = 0$. If a single individual switches to $x = 1$, his payoff will be $\pi_{0\to1} = b\frac{1}{1+X_0} - c$. Therefore the state $X = 0$ is a strict Nash equilibrium if $\pi_0 > \pi_{0\to1}$ or

$$b/c < 1 + X_0. \tag{S2}$$

Consider a state where the total group effort $X = n$. The payoff to each individual is $\pi_1 = b\frac{n}{n+X_0} - c$. If a single individual switches to $x = 0$, his payoff will be $\pi_{1\to0} = b\frac{n-1}{n-1+X_0}$. Therefore the state $X = n$ is a strict Nash equilibrium if $\pi_1 > \pi_{1\to0}$ which simplifies to

$$2n - 1 + X_0 + \frac{n(n-1)}{X_0} < b/c. \tag{S3}$$

Consider a state where the total group effort $0 < X < n$. The payoff to an individual contributing 0 is $\pi_0 = b\frac{X}{X+X_0}$. If this individual switches to $x = 1$, his payoff will be $\pi_{0\to1} = b\frac{1+X}{1+X+X_0} - c$. The individual will not be interested in switching if $\pi_0 > \pi_{0\to1}$. The payoff to an individual currently contributing 1 is $\pi_1 = b\frac{X}{X+X_0} - c$. If this individual switches to $x = 0$, his payoff will be $\pi_{1\to0} = b\frac{X-1}{X-1+X_0}$. The individual will not be interested in switching if $\pi_1 > \pi_{1\to0}$. Solving the two inequalities above, a state with total group effort X is a strict Nash equilibrium if

$$\frac{(X+X_0)(X+X_0-1)}{X_0} < b/c < \frac{(X+X_0)(X+X_0+1)}{X_0} \tag{S4a}$$

An alternative way to express this result is to say that if conditions (S2) and (S3) are not satisfied, then there exists an unique strict Nash equilibrium at which the group effort X is an integer within an unit-length interval $(I_c - 1/2, I_c + 1/2)$ centered on the value

$$I_c = \sqrt{\frac{1}{4} + rX_0} - X_0 \approx X_0 \left(\sqrt{\frac{r}{X_0}} - 1 \right), \tag{S4b}$$

where the benefit-to cost ratio $r = b/c$ and the approximation assumes that $r \gg 1/(4X_0)$. [This follows from the fact that, using variable $u = X + X_0$, conditions (S4a) can be rewritten as $u(u-1) < rX_0 < u(u+1)$, or equivalently, $\sqrt{\frac{1}{4} + rX_0} - \frac{1}{2} < u < \sqrt{\frac{1}{4} + rX_0} + \frac{1}{2}$.] Note that the approximate expression above is exactly the same as the ESS solution 2 in the main text.

Best response dynamics in stochastic agent-based simulations. In numerical implementation, we assumed that each agent updates its strategy independently with probability q. Each updating agent evaluates the expected payoffs π_0 and π_1 if choosing $x = 1$ and $x = 0$ under the assumption that all group-mates keep their strategies. Then the agent chooses to cooperate ($x = 1$) rather than defect ($x = 0$) with probability

$$p = \frac{1}{1 + \exp[\lambda(\pi_0 - \pi_1)]}, \tag{S5}$$

where λ is a non-negative precision parameter. This formulation follows the QRE approach with logit errors (Goeree *et al.*, 2016). If $\lambda \to \infty$, the agent always chooses the best response, if $\lambda = 0$, the agent choses $x = 0$ or $x = 1$ with equal probabilities. If $\lambda \to \infty$, the dynamics converges to a Nash equilibrium.

Figure S1 shows the equilibrium values while Figure S2 illustrates the transient dynamics.

Mixed Nash equilibria. Assume that each regular group member in a group contributes independently an effort 1 to a collective action with probability $0 \le p \le 1$. Following Archetti (2009), the probability that there are j contributors among $n-1$ group mates of the focal individual is

$$f_j = \binom{n-1}{j} p^j (1-p)^{n-1-j}, \tag{S6}$$

where $\binom{n-1}{j}$ is the corresponding binomial coefficient. Then the expected payoff to a focal individual if he cooperates is

$$\pi_1 = \sum_{j=0}^{n-1} f_j \left(b \frac{j+1}{j+1+X_0} - c \right). \tag{S7}$$

If the focal individual defects, his expected payoff is

$$\pi_0 = \sum_{j=0}^{n-1} f_j b \frac{j}{j+X_0}. \tag{S8}$$

The mixed Nash equilibrium for p can be found from equality $\pi_1 = \pi_0$. Using a symbolic

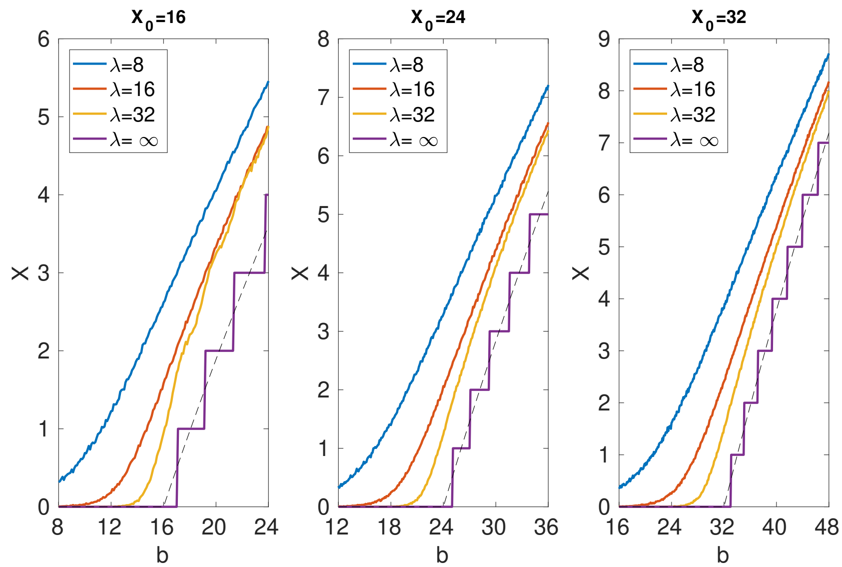


Figure S1: Average effort X of an acephalous group in stochastic simulations in "us vs. nature" games for different values of parameters X_0, b and λ. The dashed lines show the ESS prediction (2a). Other parameters: $n = 24, c = 1$, probability of strategy updating $q = 0.25$. The simulations were run for 2,000 times steps; 20 runs for each parameter combinations; the averages were evaluated over the last 1,000 time steps.

manipulation program and the Pfaff transformation, the latter can be simplified to

$$\frac{F(2, 1-n; X_0+2; p)}{1+X_0} = \frac{c}{b}, \tag{S9}$$

where $F(...)$ is the hyper-geometric function. Equation (S9) can be solved for p numerically (or graphically). One can show graphically that a positive solution exists only if $b/c > 1 + X_0$, i.e. if the benefit to cost ratio is sufficiently large.

QRE equilibria. Within the realm of the QRE approach with logit errors (Goeree *et al.*, 2016), an individual chooses to cooperate ($x = 1$) rather than defect ($x = 0$) with probability p given by equation (S5) above. The QRE solution for p satisfies the equality

$$\frac{\ln(p(1-p))}{\lambda} = \pi_0 - \pi_1 \tag{S10}$$

(Goeree *et al.*, 2016). Note that as $\lambda \to \infty$, the QRE solution converges to the mixed Nash equilibrium considered above. The QRE values can be found by numerically solving the above equation.

Potential games and stochastic equilibria. A game is an exact potential game if there is a real-valued function $\psi(z)$ defined on the space of strategies $z = (z_1, \ldots, z_n)$ such that whenever player i unilaterally changes its strategies from z_i to z_i', the corresponding change in his payoffs $\pi_i(z) - \pi_i(z')$ is equal exactly to the change in potential function, $\psi(z) - \psi(z')$ (Monderer and Shapley, 1996). The players of a potential game act as if they are jointly attempting to maximize the potential function. In any finite potential game, best response dynamics always converge to a Nash equilibrium (Hofbauer and Sandholm, 2002, Xu, 2016).

Consider the general public goods game with payoffs in the form $\pi_i(z) = G(z) - c_i(z_i)$. [Note

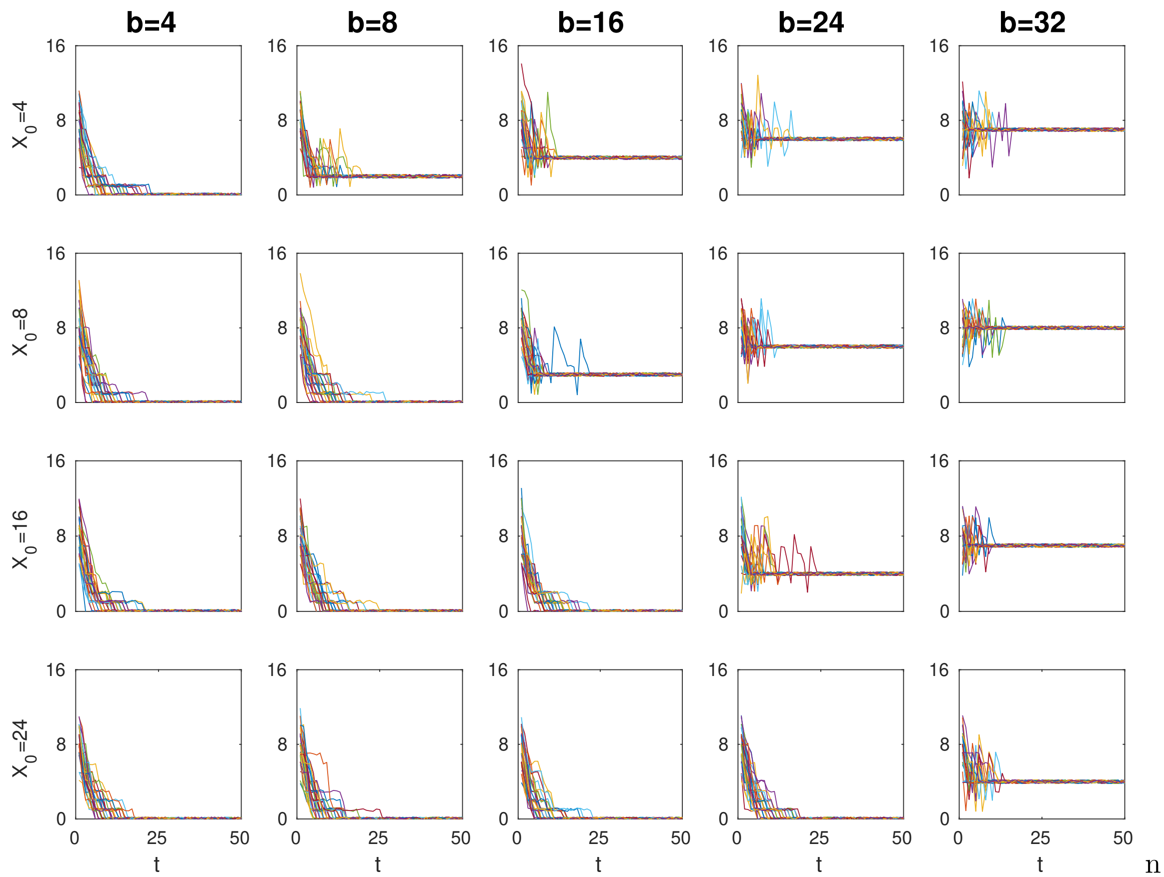


Figure S2: Dynamics of group efforts X under stochastic best response in "us vs. nature" games for different X_0 and b values. $G = 32, n = 24$. Single run for each parameter combination. Each graph shows G different lines representing efforts of G groups. Probability of updating is 0.25; $\lambda = \infty$.

that our "us vs. nature" game is a special case of this game.] Then the potential function

$$\psi(z) = G(z) - \sum c_i(z_i) \tag{S11}$$

(Myatt and Wallace, 2009).

With multinomial-logit quantal response updating, the ergodic probability that the system is found in a state z is

$$\Pr(z)\big|_{t\to\infty} = \frac{\exp(\lambda\psi(z))}{\sum_{z'}\exp(\lambda\psi(z'))} \tag{S12}$$

(Myatt and Wallace, 2009). The above equation allows one to find the equilibrium stochastic distribution of different strategies in "us vs. nature" games numerically.

3 Nash equilibria in "us vs. nature" games with leaders

Assume first that the monitoring effort of a leader is fixed at y. Comparing the expected payoffs of regular group members given by eq. (1) and (3a) of the main text, we see that a constant level of monitoring (and punishment) effectively means that the regular group members are engaged in a collective action with the benefit and cost parameters adjusted to $b(1-\rho)$ and $c-\kappa y$, respectively, so that the relevant benefit-to-cost ratio is

$$r = b(1-\rho)/(c-\kappa y).$$

As inequalities (S4a) show, each value of the group effort X is stable for a range of values of the benefit-to-cost ratio r. Specifically, as one increases r, the value of r at which the state with $X = i-1$ becomes unstable and the state with $X = i$ becomes a Nash equilibrium is

$$r_i = \frac{(X_0+i)(X_0+i-1)}{X_0}.$$

Because the leaders' effort is costly, the (Nash) equilibrium value of y will be a minimum value still compatible with stability of state $X = i$. This minimum value is

$$y_i = \frac{1}{\kappa}\left(c - \frac{b(1-\rho)}{r_i}\right).$$

This strict Nash equilibrium state with $X = i, y = y_i$ is meaningful if $0 \leq y_i \leq 1$. The leader's payoff at such a state is

$$\pi_{l,i} = bn\rho\frac{i}{X_0+i} - c_y n y_i - \delta(n-i)y_i.$$

For a given set of parameters, there can be multiple Nash equilibria (i, y_i) with different expected payoffs to the leader $\pi_{l,i}$ and different domains of attraction. The corresponding stochastic dynamics are expected to wander among these equilibria perhaps settling predominantly on one of them. We have explored these dynamics numerically in the main text.

4 Equilibria in "us vs. them" games without leaders

Consider first dyadic between-group conflicts. With two groups making efforts X and Y, respectively, in a conflict over a resource of value $2b$, the expected payoff to a member of the first group can be written as

$$\pi = \begin{cases} 2b\frac{X}{X+Y} - cx, & \text{if } X+Y>0, \\ b, & \text{if } X+Y=0. \end{cases}$$

To identify Nash equilibria in this model using the results from the previous section (specifically, inequalities S4a) we need to consider all combinations of X and Y where each variable takes values from $0,\ldots,n$.

The state $X = Y = 0$ is a Nash equilibrium if $b < c$ (because in this case, b is larger than $2b-c$).

From inequalities (S4a) with $X_0 = X$ and b substituted for $2b$, a symmetric state with $Y = X$ is a Nash equilibrium if

$$X - 1/2 \leq b/(2c) \leq X + 1/2. \tag{S13}$$

That is, $X = Y = 1$ is stable for $1 < b/c < 3$, $X = Y = 2$ is stable for $3 < b/c < 5$, and so on. Alternative, we can say that at the symmetric Nash equilibrium X is an integer closest to $b/(2c)$:

$$\frac{b}{2c} - \frac{1}{2} < X < \frac{b}{2c} + \frac{1}{2}.$$

One can also show that for $b/c = 2k+1$ where k is an integer, there are also additional Nash equilibria in the form $X = k, Y = k+1$ and $X = k+1, Y = k$ for $k = 0, 1, 2, \ldots$ One can show there are no other asymmetric Nash equilibria (i.e., with $|Y - X| > 1$).

In the case of G competing groups, let us define

$$\pi = \begin{cases} Gb\frac{X}{X+\sum Y} - cx, & \text{if } X + \sum Y > 0, \\ b, & \text{if } X + \sum Y = 0, \end{cases}$$

where $\sum Y$ is the sum of efforts of other $G-1$ groups.

State $X = 0$ is a Nash equilibrium if $Gb - c < b$, i.e. if

$$b/c < \frac{1}{G-1}.$$

From eq. (S4a) and b substituted for Gb, the symmetric state $X > 0$ with $X_0 = (G-1)X$ is a Nash equilibrium if

$$X + \frac{X-1}{G-1} < b/c < X + \frac{X+1}{G-1}.$$

Alternatively, we can write the above inequalities as

$$\left(1 - \frac{1}{G}\right)\frac{b}{c} - \frac{1}{G} < X < \left(1 - \frac{1}{G}\right)\frac{b}{c} + \frac{1}{G}.$$

As G increases, X becomes close to b/c. That is, with large G such an equilibrium exists only if b/c is an integer. With large G, the ranges of stability of these symmetric equilibria become very narrow.

There are also many asymmetric equilibria. For example, consider a focal group with $X = 0$ in a system where all other groups are making nonzero efforts. Nobody in the focal group will be willing to make an effort unless $b > c(\sum X + 1)$, where $\sum X$ is the sum of efforts over all groups.

Numerical results suggest there is a very large number of asymmetric equilibria with the average group effort $\overline{X}$ close to b/c. As the number of groups in the system grows, the system exhibits non-equilibrium dynamics (at least at the time scale of our simulations). See Figure S3 and Figure S4 here.

Mixed Nash equilibria. Assume that each individual in a group contributes independently an effort 1 to a collective action with probability $0 \leq p \leq 1$. The (mixed) Nash equilibrium value of p satisfies the equation for expectations

$$E(\pi_0) = E(\pi_1), \tag{S14a}$$

where $E(\pi_x)$ is the expected payoff of an individual making effort x. The expected payoff of a defector can be written as

$$E(\pi_0) = Gb\ E\left(\frac{i}{i+j}\right), \tag{S14b}$$

where $i \sim \text{Bin}(n-1, p)$ and $j \sim \text{Bin}(n-1+(G-1)n, p)$ are independent random variables drawn from the corresponding binomial distributions. The expected payoff of a cooperator can be written as

$$E(\pi_1) = Gb\ E\left(\frac{1+i}{1+i+j}\right) - c. \tag{S14c}$$

While the exact calculation of the expectations of the two ratios above does not seem to be possible,

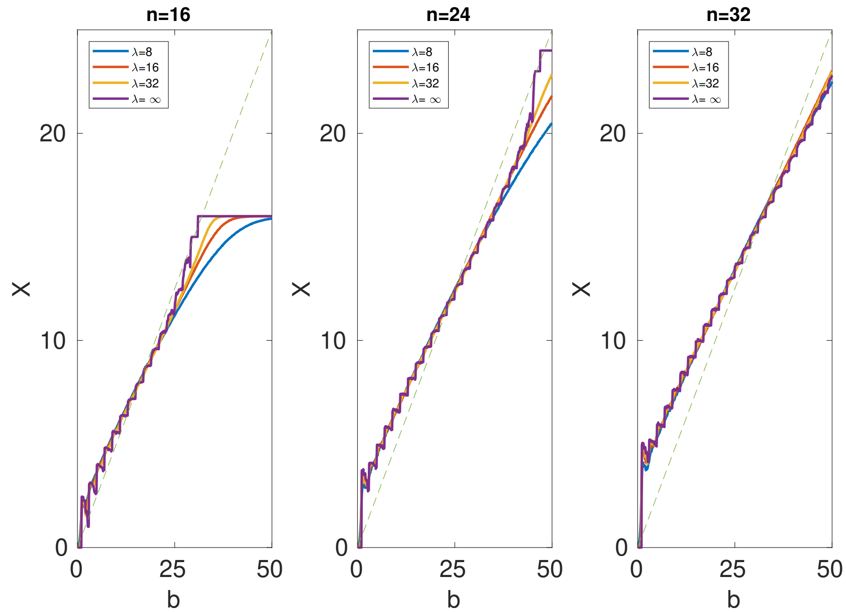


Figure S3: Average effort X of an acephalous group in stochastic simulations in "us versus them" games for different values of parameters n, b and λ. The dashed lines show the ESS prediction (2b). Other parameters: $c = 1$, probability of strategy updating $q = 0.25$, number of groups $G = 2$. Note that X levels off at the maximum possible size $X = n$.

we can approximate these expectations using a formula based on the second order Taylor expansion for the expectation of a ratio of two random variable (here x and y):

$$E(x/y) \approx E(x)/E(y) - \text{cov}(x, y)/E(y)^2 + \text{var}(y)E(x)/E(y)^3, \tag{S15}$$

where $\text{cov}(x, y)$ is the covariance of x and y, and $\text{var}(y)$ is the variance of y (Stuart and Ord, 2010). This approach is justified if the variance $var(y)$ and covariance $cov(x, y)$ are both much smaller than $E(y)^2$. In our case, this assumption is satisfied if the group size n is not too small. The relevant expectations are:

$$\begin{aligned} E(i) &= (n-1)p, \\ E(i+j) &= [n-1+(G-1)n]p, \\ \text{var}(i+j) &= [n-1+(G-1)n]p(1-p), \\ \text{cov}(i, i+j) &= (n-1)p(1-p). \end{aligned}$$

Making appropriate substitutions, one ends up with a cubic equation for p which can be solved numerically. Figure S5 shows the corresponding solutions for the total group effort $X^* = np^*$ as a function of the benefit b. These solutions are well approximated by the ESS values (**??**b) given in the main text.

Effects of ω on regular group members' payoffs. The leaders' effort y always increases with ω. To understand the effect of ω on regular group members in "us vs. nature" games, assume that regular group members use best response and that the regular group members effort is described by the ESS solution (2a) of the main text

$$X^* = X_0(\sqrt{R} - 1), \tag{S16}$$

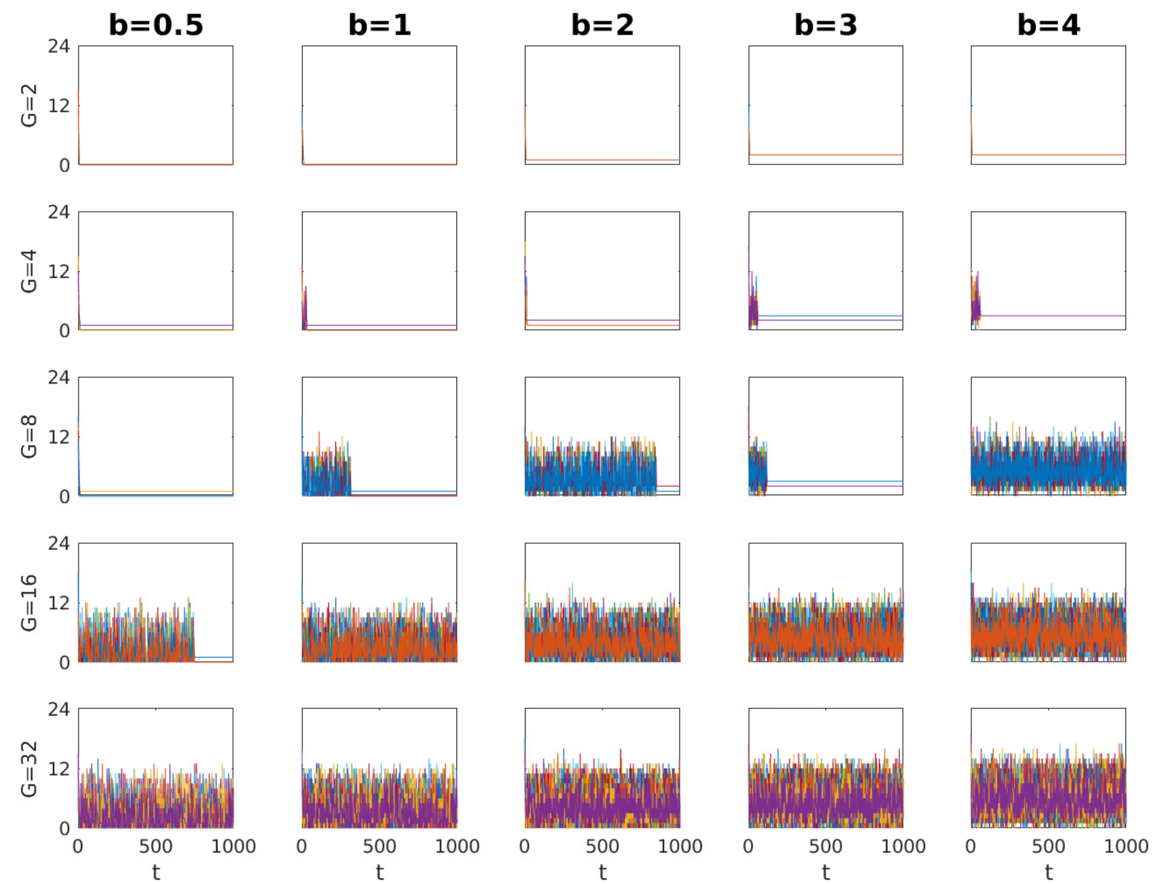


Figure S4: Dynamics of group efforts X under stochastic best response in "us versus them" games for different G and b values. $n = 24$. Single run for each parameter combination. Each graph shows G different lines representing efforts of G groups. Probability of updating is 0.25; $\lambda = \infty$.

if $R > 1$ and $X^* = 0$ otherwise. With punishment by leaders

$$R \equiv \frac{(1-\rho)b}{(c-\kappa y)X_0}.$$

Using equation (3a) of the main text, the total payoff of n regular group members is

$$\pi_c(y, X) = (1-\rho)nbP(X) - cX - \kappa y(n - X), \tag{S17}$$

Assume that $R > 1$. Taking the derivative, we find that

$$\frac{d\pi_c(y,X)}{dy} = \left(\sqrt{R}\frac{(n+1)X_0}{2} - n - X_0\right)\kappa$$

The expression in the right-hand side of the last equation is always positive if $X_0 > 2n/(n-1)$. Therefore if regular group members best response is positive, increasing y typically further increases their effort X and their payoffs.

If $R < 1$ and the best response for regular group members $X^* = 0$, then increasing y always

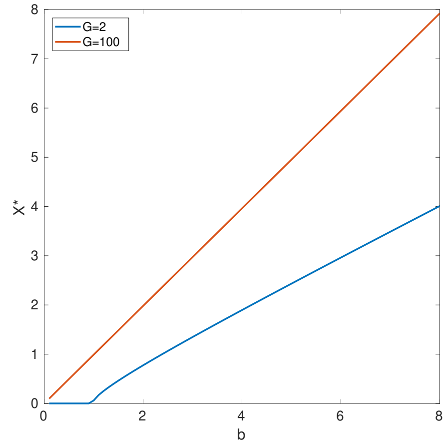


Figure S5: Mixed Nash equilibrium values in "us vs. them" games with $G = 2$ and 100 for $n = 24, c = 1$. Equilibrium values are found numerically by solving equations (S14) with help of approximation (S15).

decreases their payoffs.

5 Modeling innovation

In our numerical simulations, innovation for leaders was implemented by perturbing their strategy y by a number drawn randomly and independently from a PERT distribution defined on the unit interval with the mode at the current value. The PERT distribution belongs to a family of beta-distributions; it was offered as a simple alternative to using a truncated normal or a triangular distributions Clark (1962) and is now widely used in risk analysis.

The advantages of using the PERT distribution are that it is simple to use (no additional parameters relative, e.g. to stepwise mutation), smooth (relative to the triangular distribution), is biased towards previous values (relative to the uniform distribution), and has no discontinuous spikes (relative to truncated normal distribution). Our additional simulations using a truncated normal and a uniform distributions show that the equilibrium values are not affected much by the underlying distribution of innovations which mostly affect the time to convergence.

Consider a PERT distribution defined on the interval $[a, c]$ with the mode at b ($a \leq b \leq c$). Its cumulative distribution function is the incomplete beta function $I_z(\alpha, \beta)$ where

$$z = \frac{x-a}{c-a}, \ \alpha = 1 + 4\frac{b-a}{c-a}, \ \beta = 1 + 4\frac{c-b}{c-a}$$

In our case, $a = 0, b = y, c = 1$. To generate a random number from this distribution we used Matlab function $betaincinv(u, \alpha, \beta)$, where u is a random number from a uniform distribution.

6 Additional agent-based simulations

Figure S6 illustrates the effects of parameters in "us vs. them " games. This Figure is analogous for Figure 2 in the main text.

Figure S7 illustrates the possibility of multiple equilibria for "us vs. nature" games. In the simulations shown, the initial values of y were chosen in a uniform way across the range of values

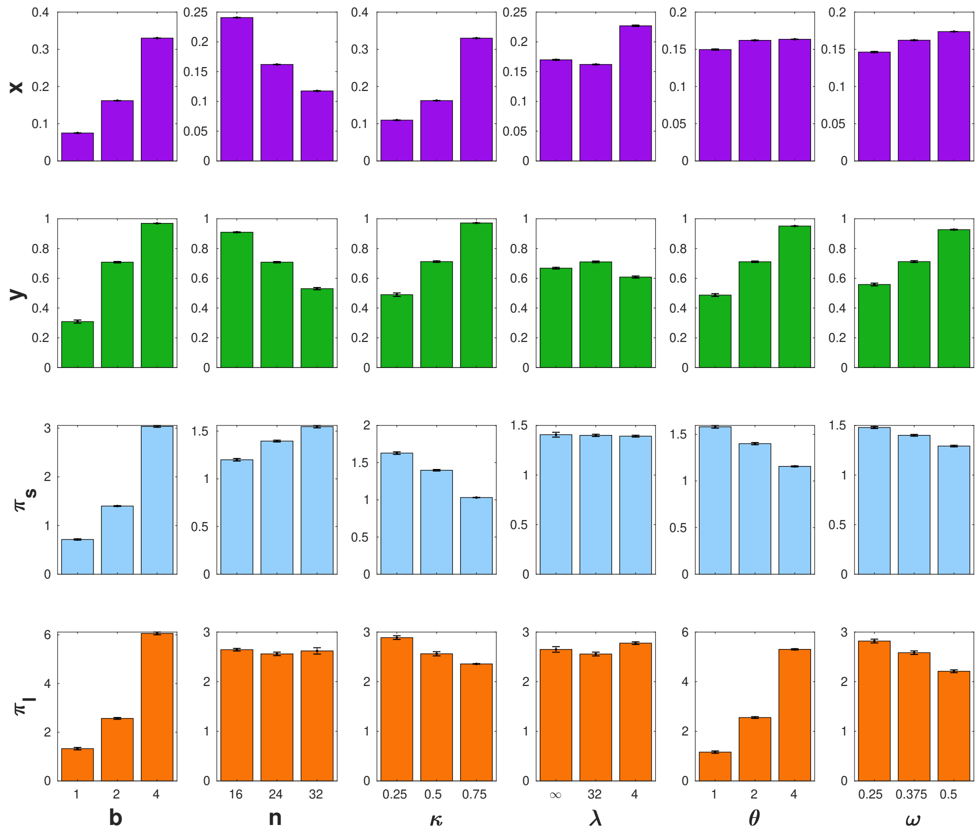


Figure S6: Effects of parameters benefit b, group size n, punishment strength k, and precision λ, tax θ, and the foresight parameter ω on the average efforts of regular group members x and leaders y and their payoffs π_c and π_l in "us vs. them" games. Parameters are changed one at a time relative to a "default" set with $b = 16, n = 24, \theta = 2, \kappa = 0.5, \lambda = \infty, \omega = 0.375. Frequencies of updating events$:$E_1 = 0.01, E_2 = E_3 = 0.12$ in leaders and $E_1 = 0.01, E_2 = 0, E_3 = 0.24$ in regular group members.

between 0 and 1 and the effects of stochasticity were reduced by making precision parameter λ infinite and greatly restricting the range of innovations.

The graphs S4-S15 below aim to show the transition from the state with no punishment and production (with smaller values of b) to punishment and production (with larger values of b). The values of b for which such a transition happens depend on the punishment parameter κ, with stronger punishment allowing for the transition at smaller benefits b. To make the patterns clearer, we only show the case of perfect precision $\lambda = \infty$. With smaller λ, there are naturally more noise in the data.

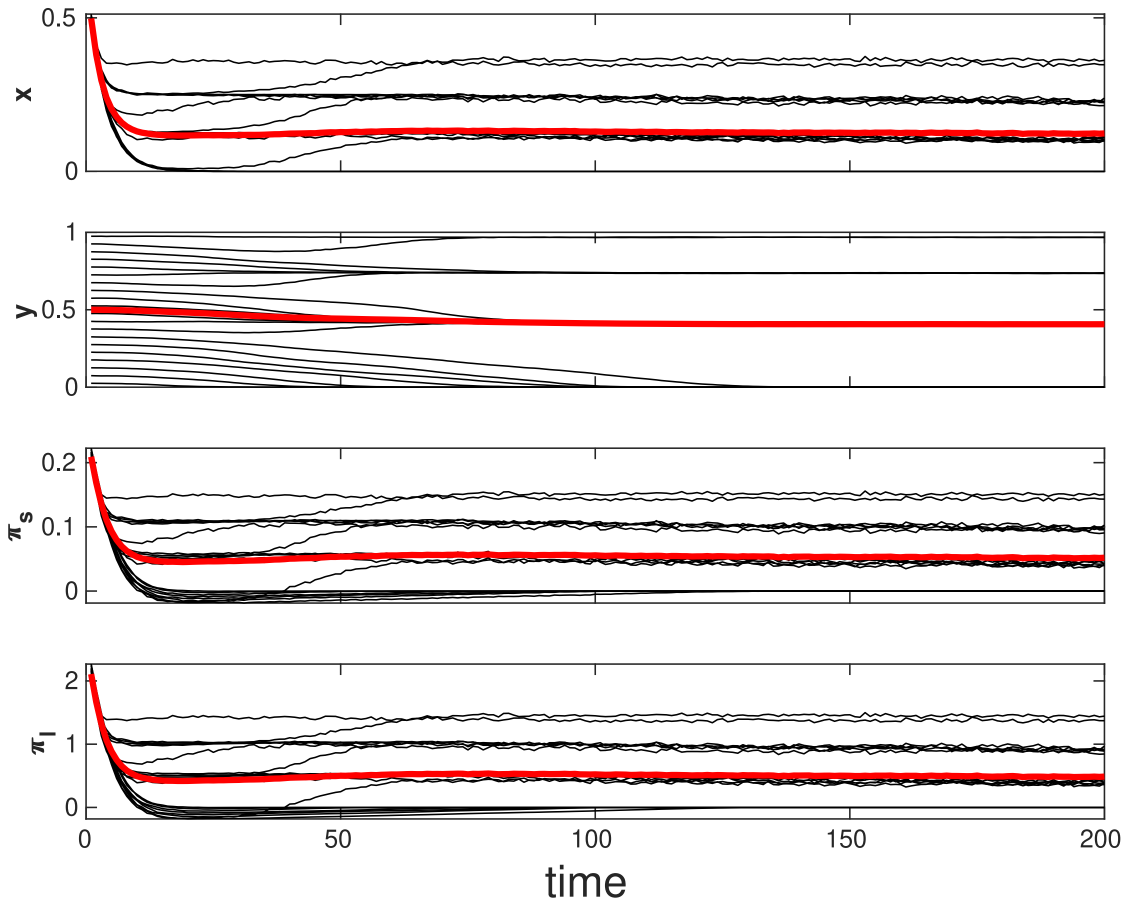
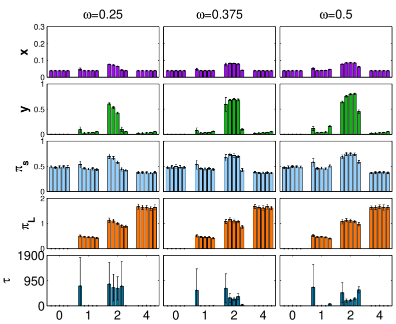
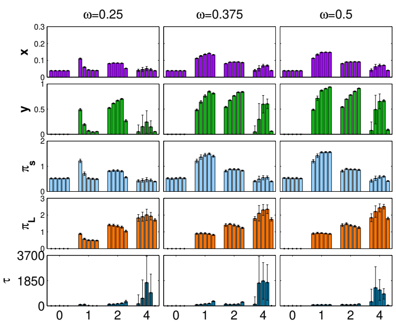


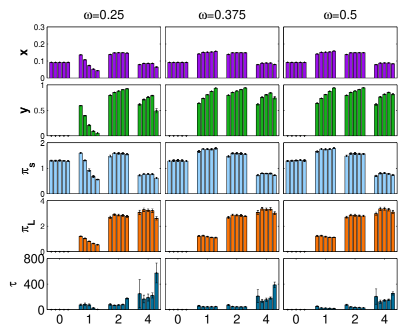
Figure S7: An example of multiple equilibria in "us vs. nature" games. Shown are the dynamics of average values of x, y, π_c and π_l in 20 independent runs. Parameters: $\lambda = \infty, \kappa = 0.75, b = 8, n = 8, X_0 = 8, \theta = 8$ and mostly selective imitation in leaders ($E_2 = 0.23, E_3 = 0.01$). Initial values of x (0 or 1) are drawn randomly and independently with equal probabilities. In each run, initial values of y of leaders are drawn randomly and independently from a uniform distribution within a narrow band; the bands for different runs did not overlap and covered the whole range of possible y values. Innovation was modeled by perturbing an original value of y by a random number drawn from a truncated normal distribution with zero mean and a small standard deviation $\sigma = 0.025$. Red lines show the averages over 20 runs. In the case shown, there are four different locally stable equilibria.



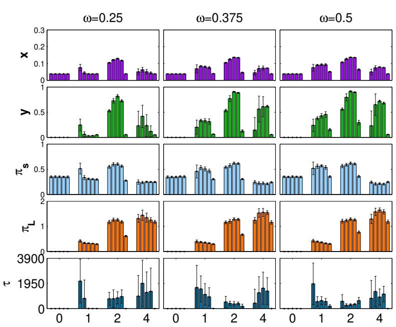
(a) $\kappa = 0.25, b = 15$



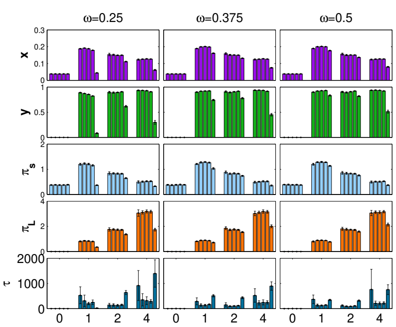
(b) $\kappa = 0.25, b = 16$



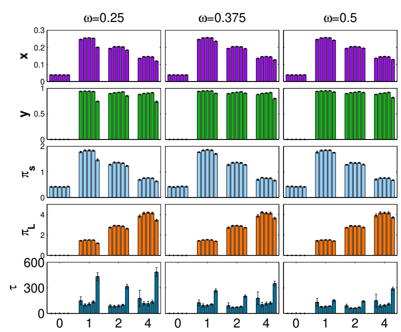
(c) $\kappa = 0.25, b = 17$



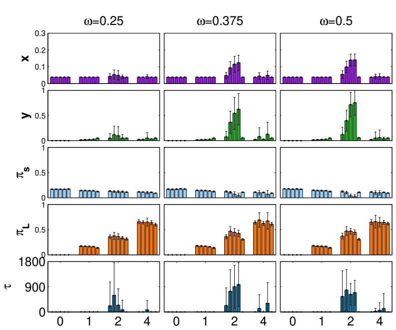
(d) $\kappa = 0.5, b = 11$



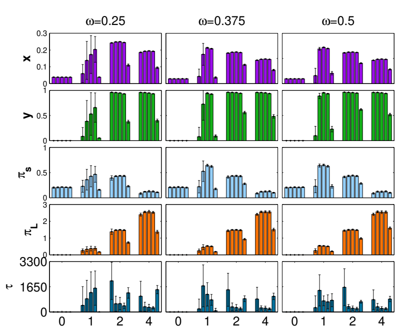
(e) $\kappa = 0.5, b = 12$



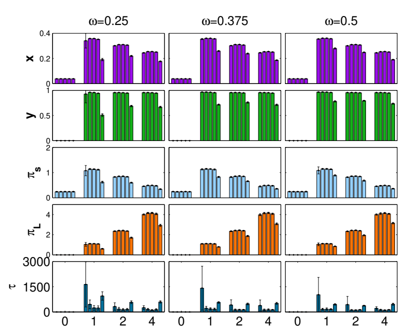
(f) $\kappa = 0.5, b = 13$



(g) $\kappa = 0.75, b = 6$

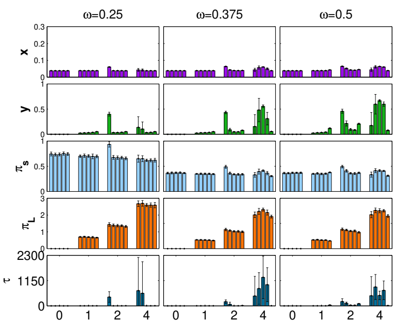


(h) $\kappa = 0.75, b = 7$

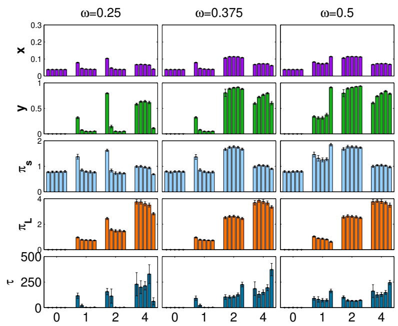


(i) $\kappa = 0.75, b = 8$

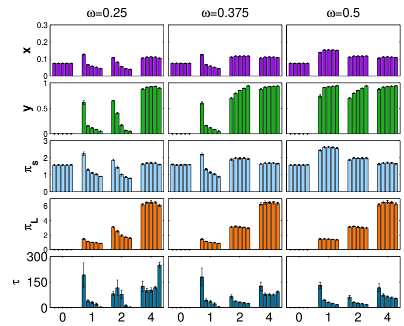
Figure S8: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 16, X_0 = 16$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



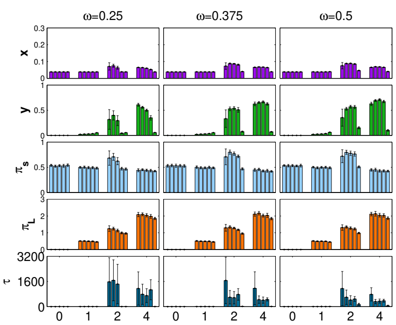
(a) $\kappa = 0.25, b = 15$



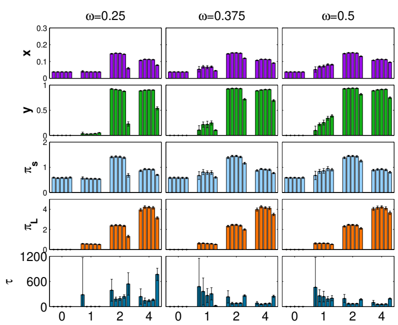
(b) $\kappa = 0.25, b = 16$



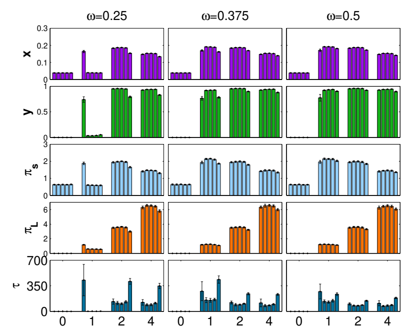
(c) $\kappa = 0.25, b = 17$



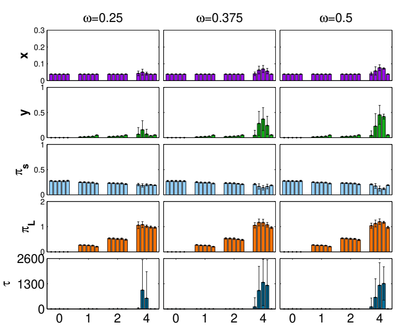
(d) $\kappa = 0.5, b = 11$



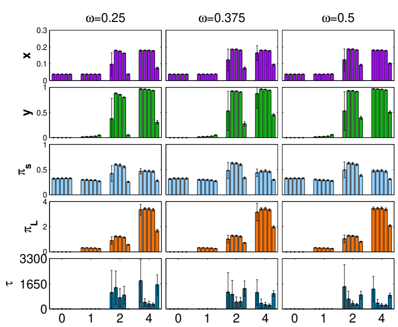
(e) $\kappa = 0.5, b = 12$



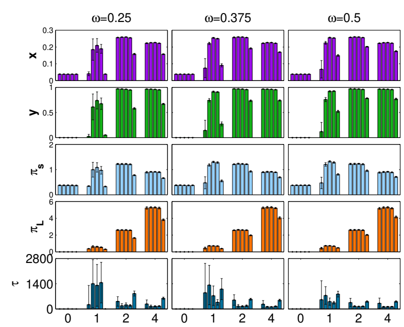
(f) $\kappa = 0.5, b = 13$



(g) $\kappa = 0.75, b = 6$

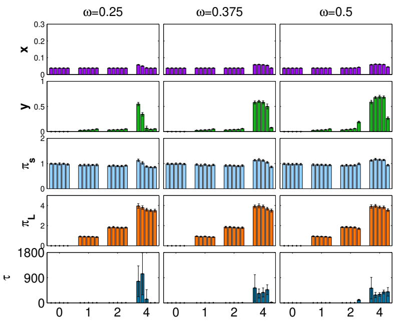


(h) $\kappa = 0.75, b = 7$

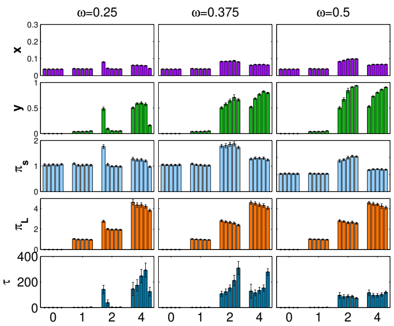


(i) $\kappa = 0.75, b = 8$

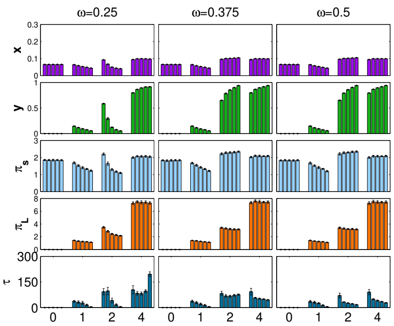
Figure S9: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 24, X_0 = 16$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



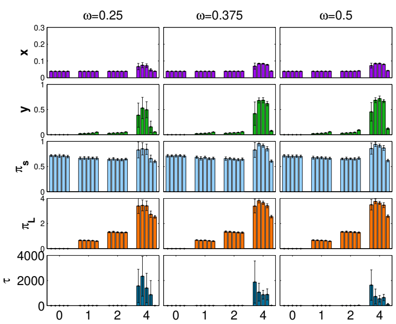
(a) $\kappa = 0.25, b = 15$



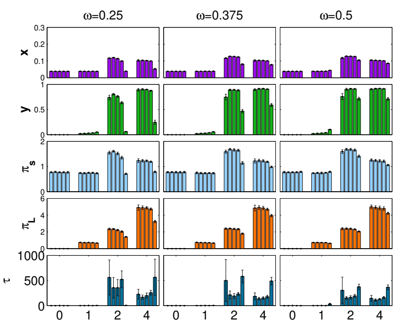
(b) $\kappa = 0.25, b = 16$



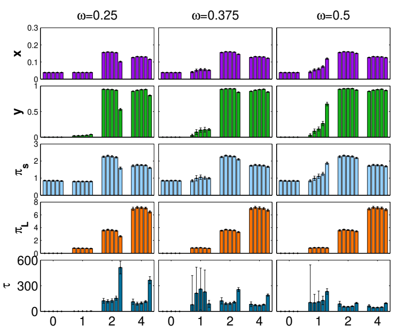
(c) $\kappa = 0.25, b = 17$



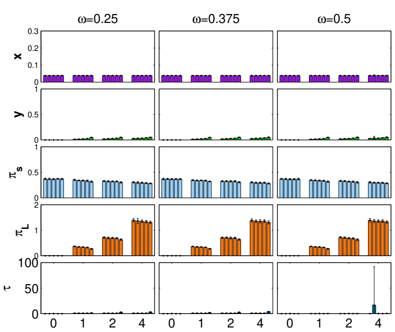
(d) $\kappa = 0.5, b = 11$



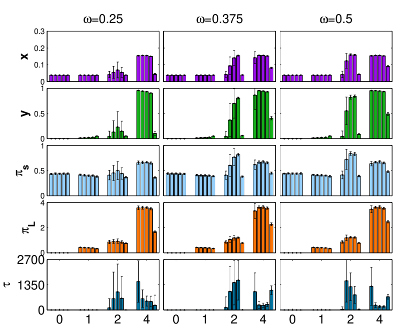
(e) $\kappa = 0.5, b = 12$



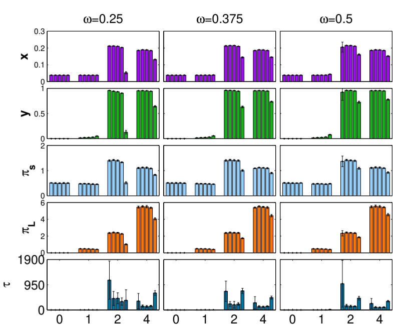
(f) $\kappa = 0.5, b = 13$



(g) $\kappa = 0.75, b = 6$

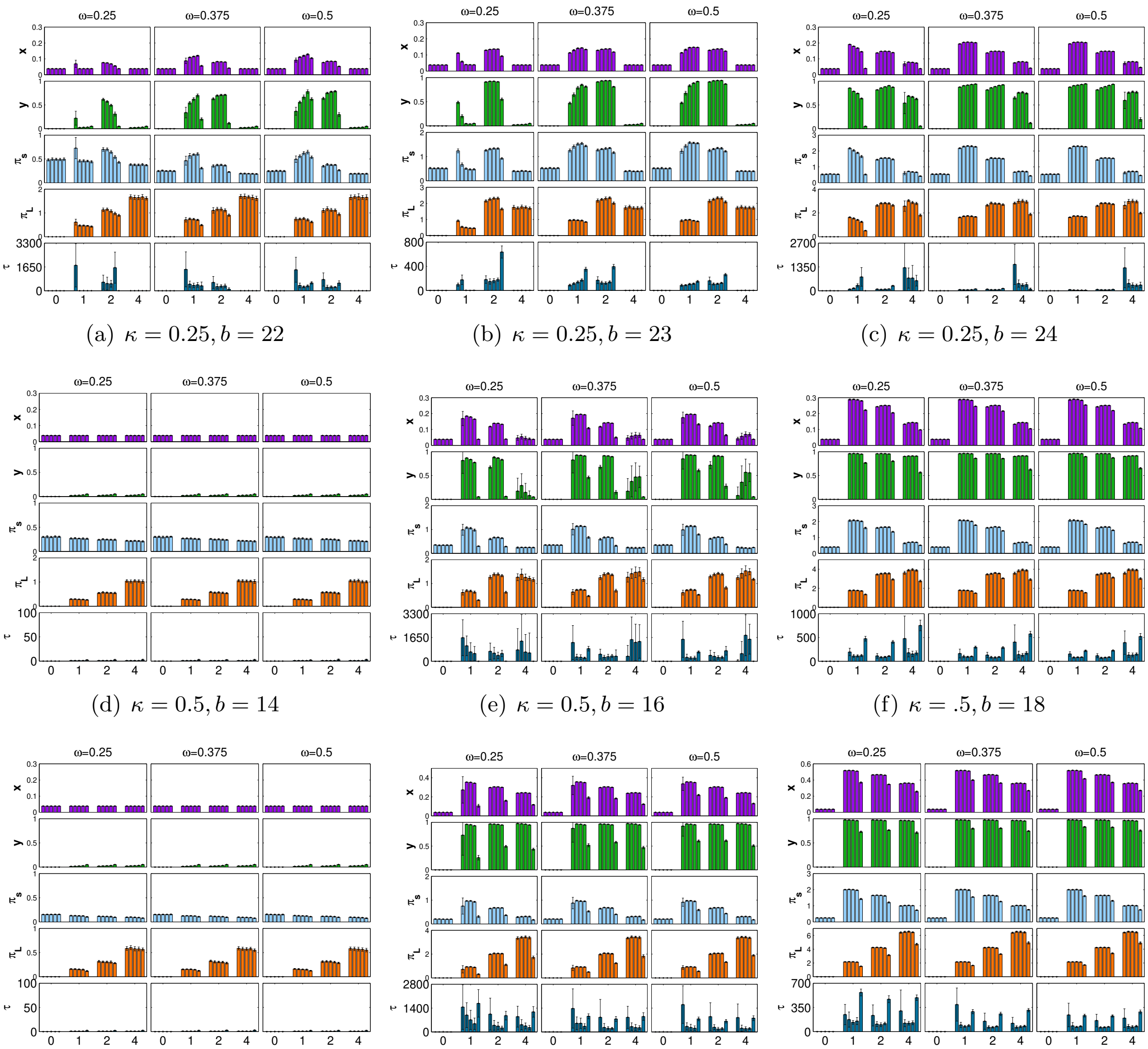


(h) $\kappa = 0.75, b = 7$



(i) $\kappa = 0.75, b = 8$

Figure S10: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 32, X_0 = 16$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.

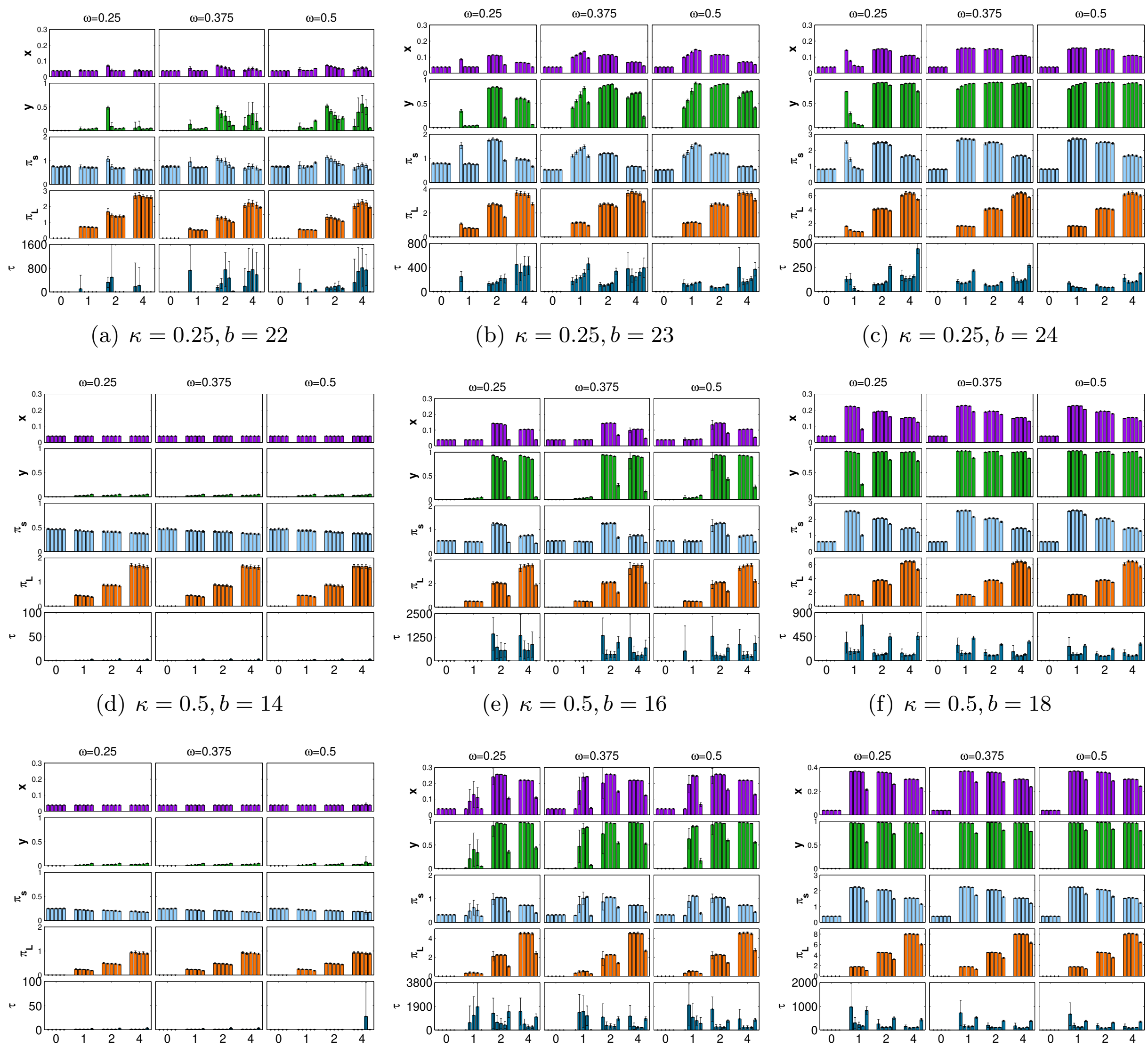


(a) $\kappa = 0.25, b = 22$ (b) $\kappa = 0.25, b = 23$ (c) $\kappa = 0.25, b = 24$

(d) $\kappa = 0.5, b = 14$ (e) $\kappa = 0.5, b = 16$ (f) $\kappa = .5, b = 18$

(g) $\kappa = 0.75, b = 8$ (h) $\kappa = 0.75, b = 10$ (i) $\kappa = 0.75, b = 12$

Figure S11: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 16, X_0 = 24$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.

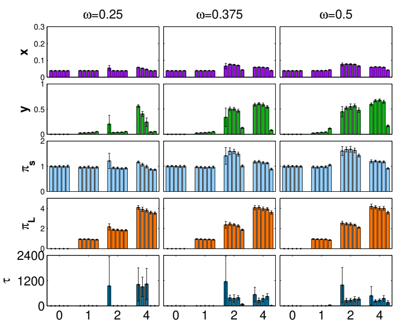


(a) $\kappa = 0.25, b = 22$ (b) $\kappa = 0.25, b = 23$ (c) $\kappa = 0.25, b = 24$

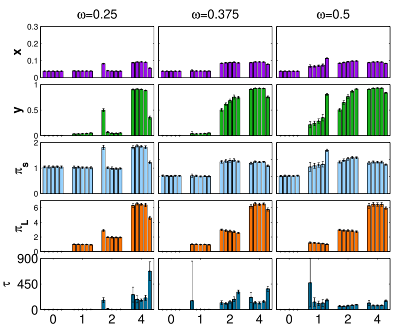
(d) $\kappa = 0.5, b = 14$ (e) $\kappa = 0.5, b = 16$ (f) $\kappa = 0.5, b = 18$

(g) $\kappa = 0.75, b = 8$ (h) $\kappa = 0.75, b = 10$ (i) $\kappa = 0.75, b = 12$

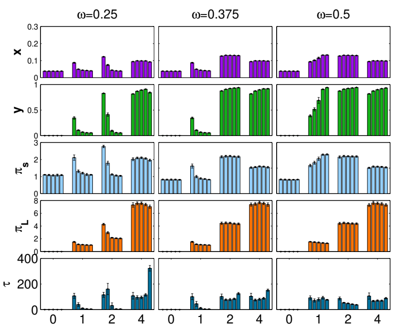
Figure S12: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 24, X_0 = 24$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



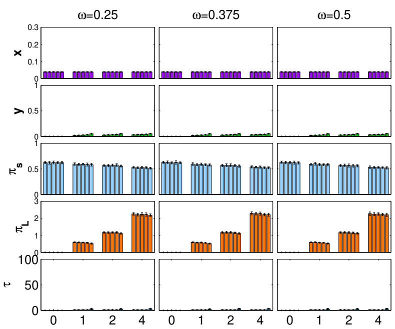
(a) $\kappa = 0.25, b = 22$



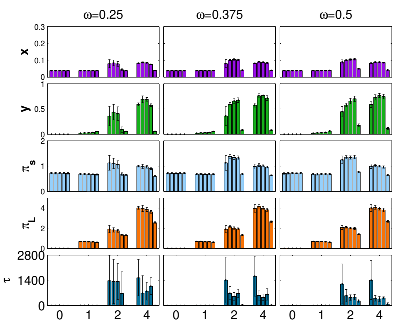
(b) $\kappa = 0.25, b = 23$



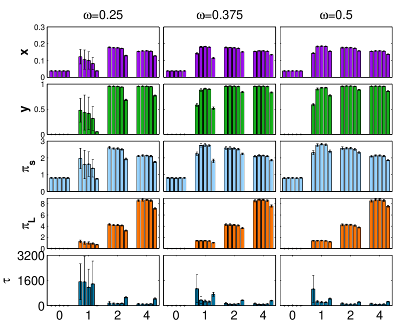
(c) $\kappa = 0.25, b = 24$



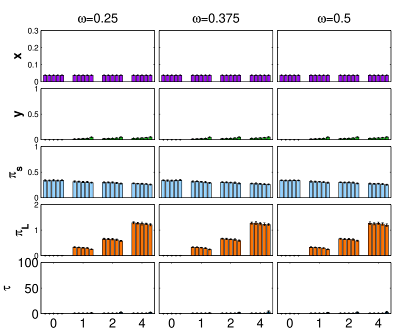
(d) $\kappa = 0.5, b = 16$



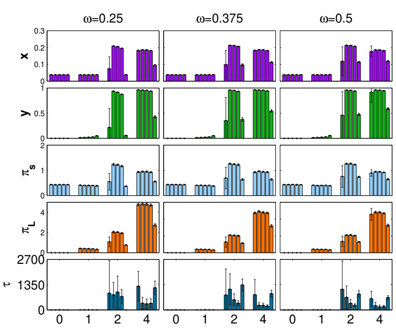
(e) $\kappa = 0.5, b = 17$



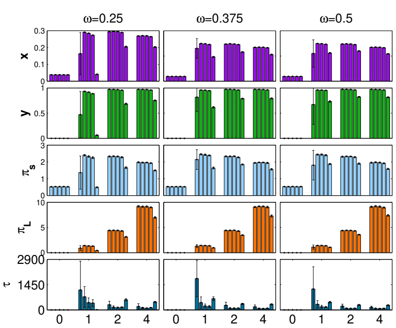
(f) $\kappa = 0.5, b = 18$



(g) $\kappa = 0.75, b = 8$

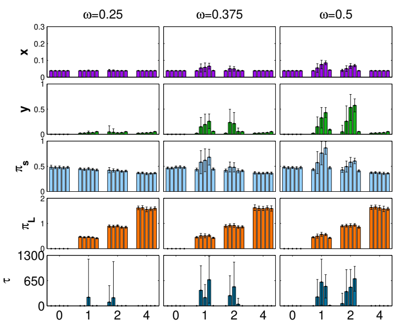


(h) $\kappa = 0.75, b = 10$

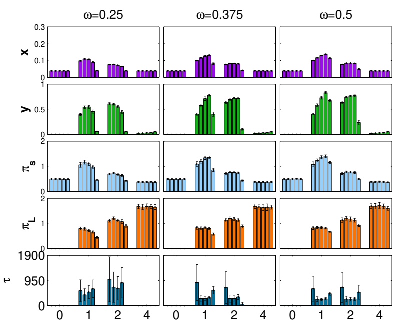


(i) $\kappa = 0.75, b = 12$

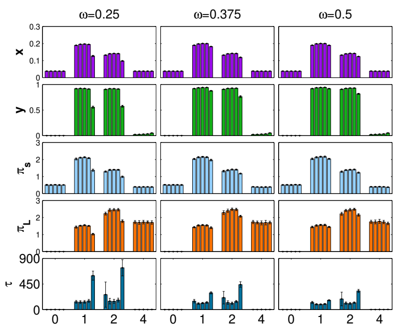
Figure S13: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 32, X_0 = 24$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



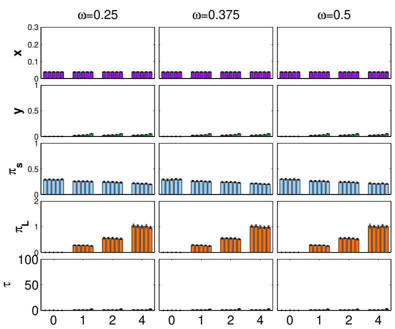
(a) $\kappa = 0.25, b = 28$



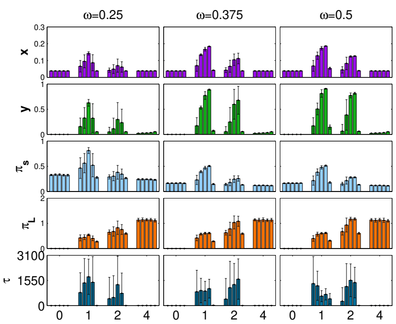
(b) $\kappa = 0.25, b = 29$



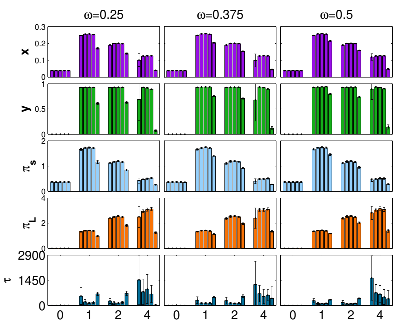
(c) $\kappa = 0.25, b = 30$



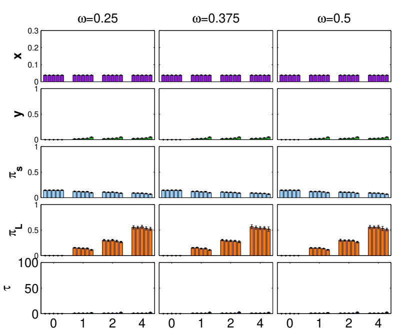
(d) $\kappa = 0.5, b = 18$



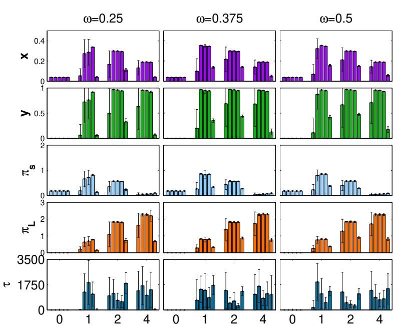
(e) $\kappa = 0.5, b = 20$



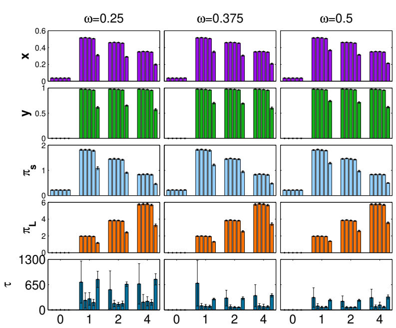
(f) $\kappa = 0.5, b = 22$



(g) $\kappa = 0.75, b = 10$

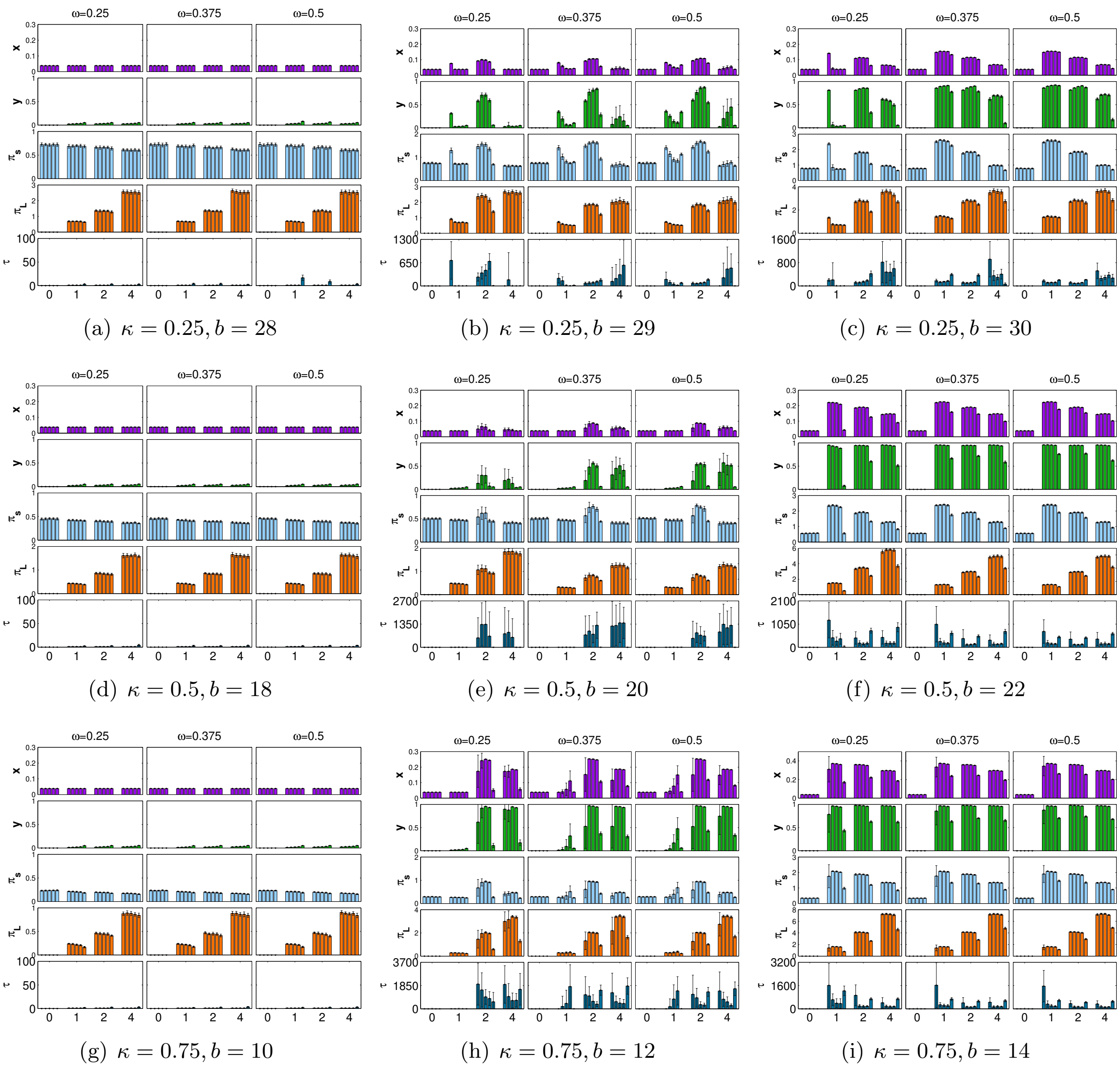


(h) $\kappa = 0.75, b = 12$



(i) $\kappa = 0.75, b = 14$

Figure S14: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 16, X_0 = 32$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.

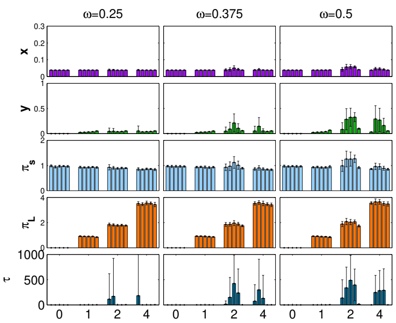


(a) $\kappa = 0.25, b = 28$ (b) $\kappa = 0.25, b = 29$ (c) $\kappa = 0.25, b = 30$

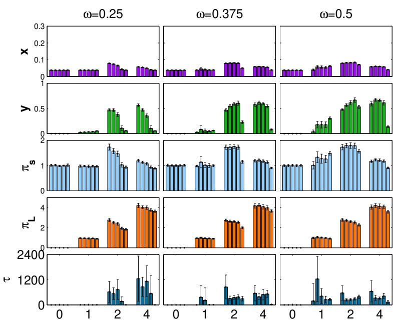
(d) $\kappa = 0.5, b = 18$ (e) $\kappa = 0.5, b = 20$ (f) $\kappa = 0.5, b = 22$

(g) $\kappa = 0.75, b = 10$ (h) $\kappa = 0.75, b = 12$ (i) $\kappa = 0.75, b = 14$

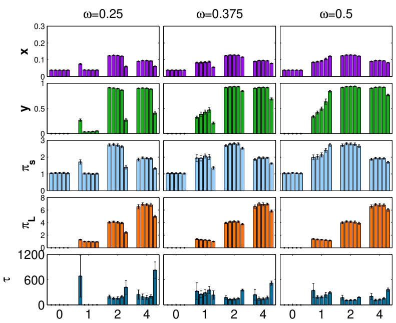
Figure S15: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 24, X_0 = 32$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



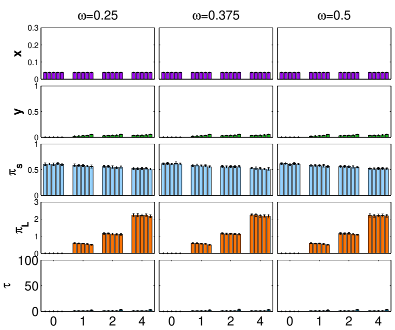
(a) $\kappa = 0.25, b = 28$



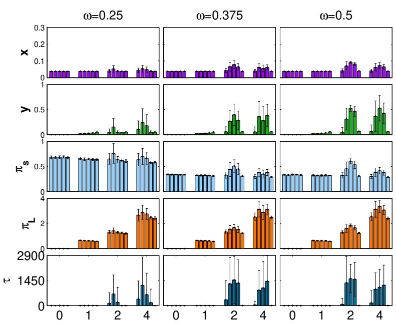
(b) $\kappa = 0.25, b = 29$



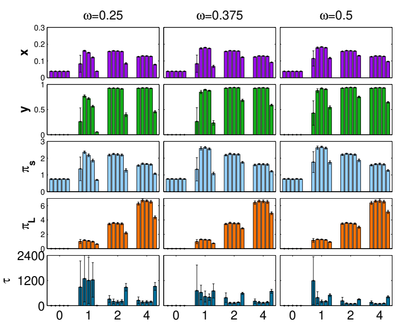
(c) $\kappa = 0.25, b = 30$



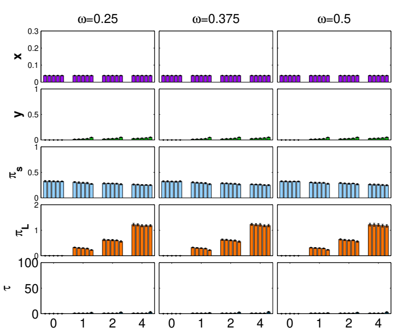
(d) $\kappa = 0.5, b = 18$



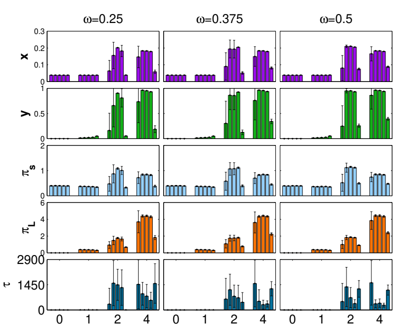
(e) $\kappa = 0.5, b = 20$



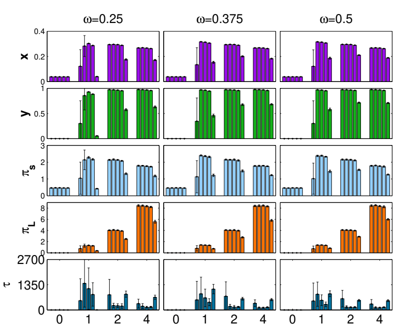
(f) $\kappa = 0.5, b = 22$



(g) $\kappa = 0.75, b = 10$

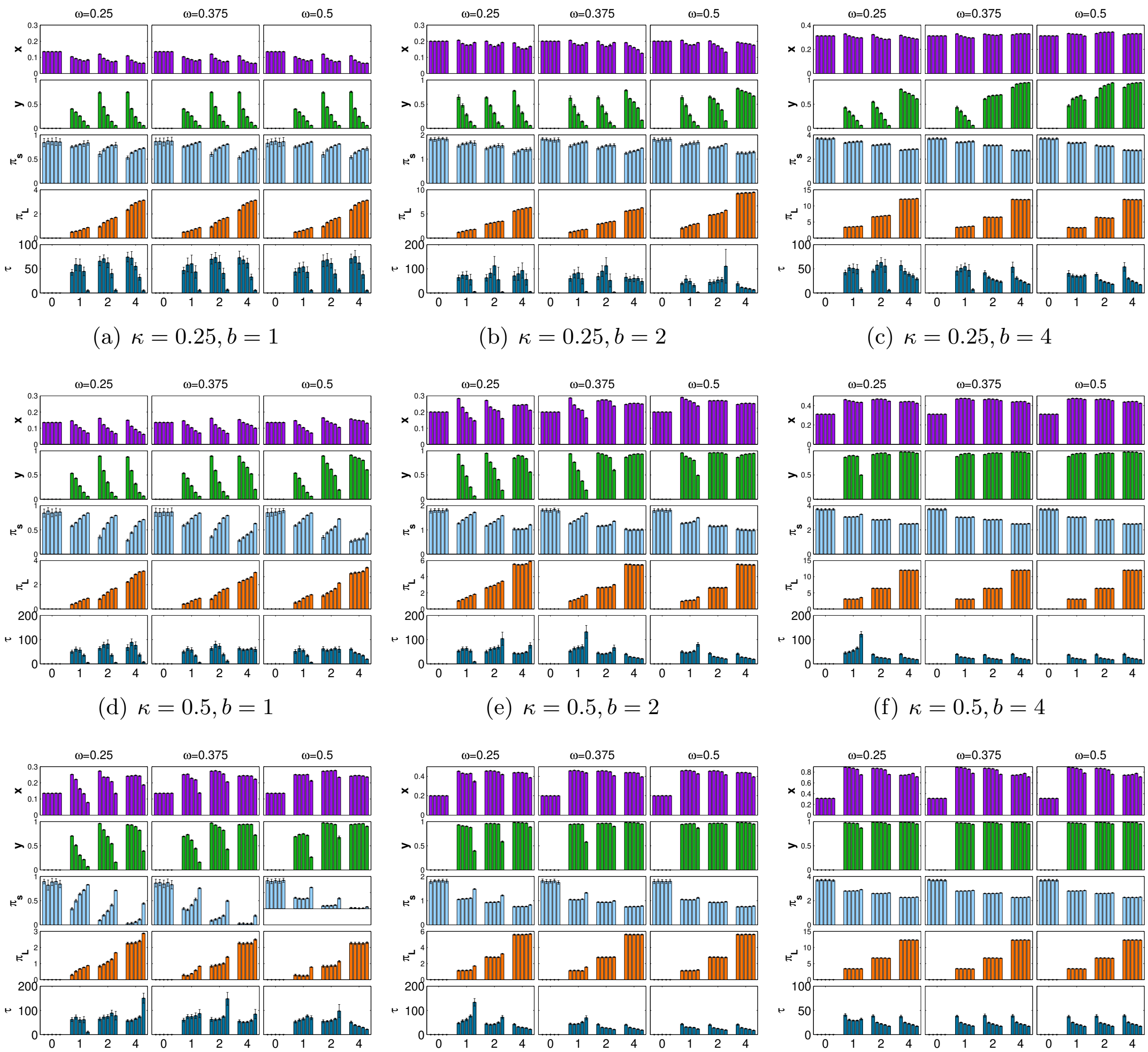


(h) $\kappa = 0.75, b = 12$



(i) $\kappa = 0.75, b = 14$

Figure S16: Equilibrium values of x, y, π_c, π_l, τ in "us vs. nature" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 32, X_0 = 32$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.

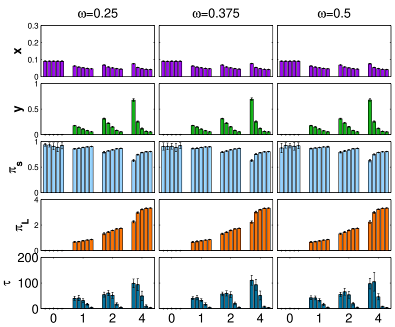


(a) $\kappa = 0.25, b = 1$ (b) $\kappa = 0.25, b = 2$ (c) $\kappa = 0.25, b = 4$

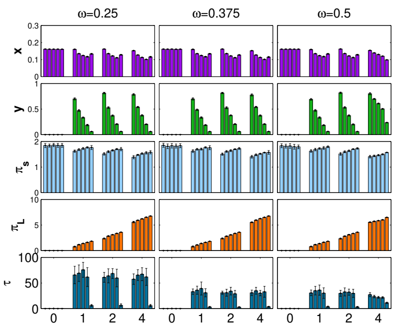
(d) $\kappa = 0.5, b = 1$ (e) $\kappa = 0.5, b = 2$ (f) $\kappa = 0.5, b = 4$

(g) $\kappa = 0.8, b = 1$ (h) $\kappa = 0.8, b = 2$ (i) $\kappa = 0.8, b = 4$

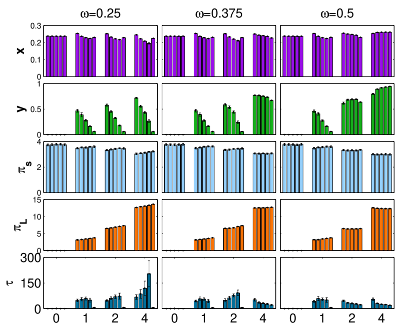
Figure S17: Equilibrium values of x, y, π_c, π_l, τ in "us vs. them" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 16$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



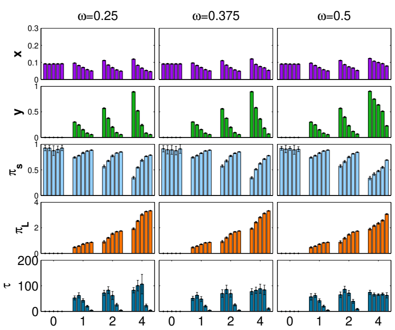
(a) $\kappa = 0.25, b = 1$



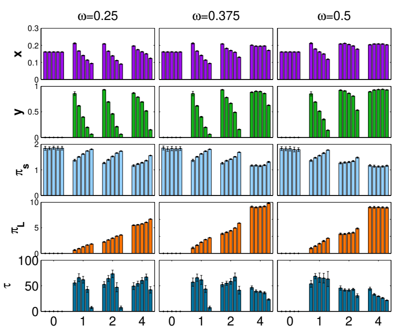
(b) $\kappa = 0.25, b = 2$



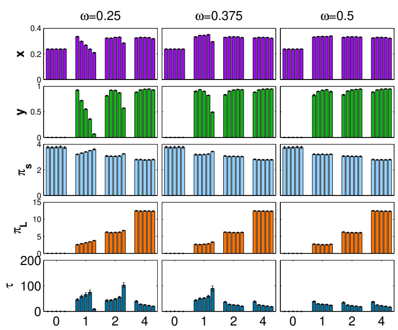
(c) $\kappa = 0.25, b = 4$



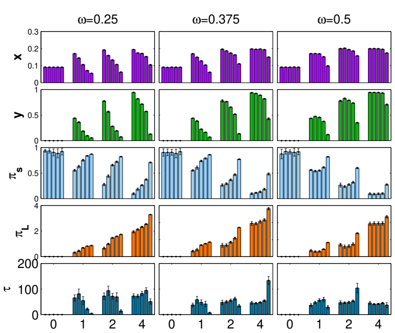
(d) $\kappa = 0.5, b = 1$



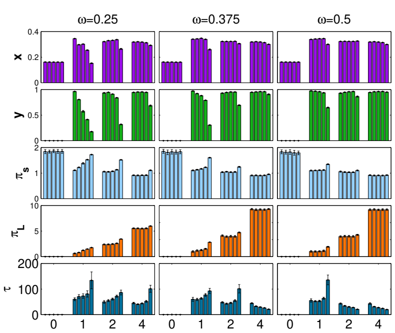
(e) $\kappa = 0.5, b = 2$



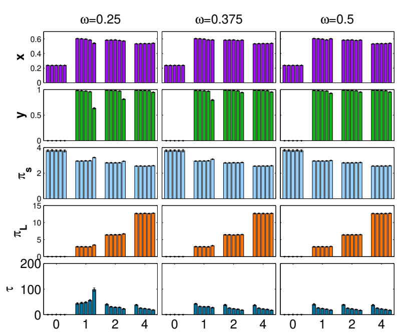
(f) $\kappa = 0.5, b = 4$



(g) $\kappa = 0.8, b = 1$

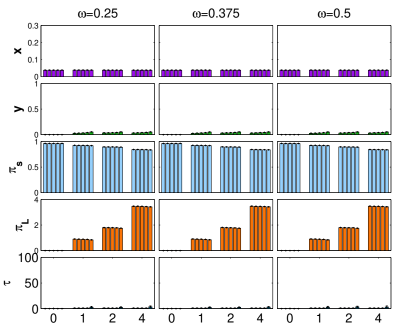


(h) $\kappa = 0.8, b = 2$

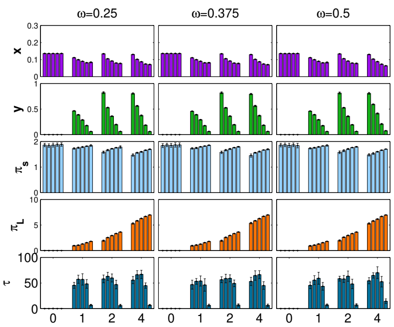


(i) $\kappa = 0.8, b = 4$

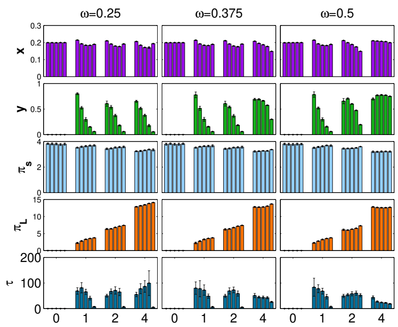
Figure S18: Equilibrium values of x, y, π_c, π_l, τ in "us vs. them" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 24$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.



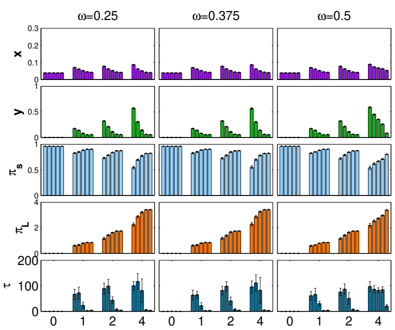
(a) $\kappa = 0.25, b = 1$



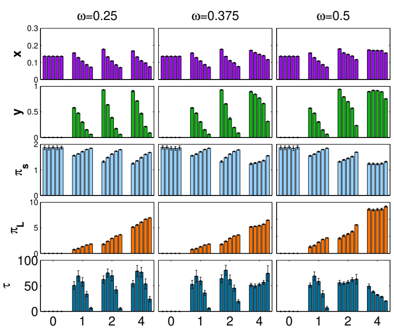
(b) $\kappa = 0.25, b = 2$



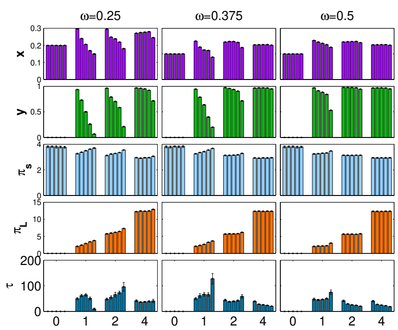
(c) $\kappa = 0.25, b = 4$



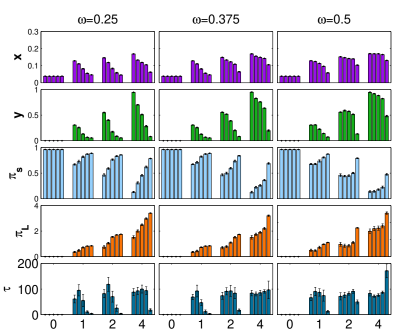
(d) $\kappa = 0.5, b = 1$



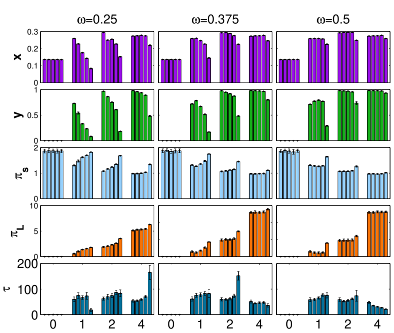
(e) $\kappa = 0.5, b = 2$



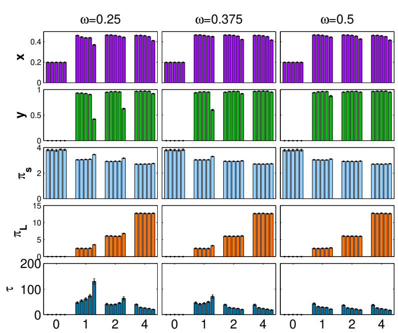
(f) $\kappa = 0.5, b = 4$



(g) $\kappa = 0.8, b = 1$



(h) $\kappa = 0.8, b = 2$



(i) $\kappa = 0.8, b = 4$

Figure S19: Equilibrium values of x, y, π_c, π_l, τ in "us vs. them" games for different tax θ, benefit b, and cost of punishment κ parameters. Within each set, different bars correspond to different combinations of the frequencies of selective imitation E_2 and foresight E_3 in leaders. Specifically, from the left-most bar to the right-most bar the ratio $E_2 : E_3$ is equal to $0.23 : 0.01$ (i.e, predominantly, selective imitation), $0.16 : 0.08, 0.12 : 0.12, 0.08 : 0.16, 0.01 : 0.23$ (predominantly foresight). The frequency of random mutation $E_1 = 0.01$. Other parameters:$\lambda = \infty, n = 32$ and initial values $y = 0$. With $\theta = 0$, the leader's effort is set to 0.

References

Archetti, M. (2009). Cooperation as a volunteer's dilemma and the strategy of conflict in public goods games. *Journal of Evolutionary Biology*, **22**, 2192–2200.

Clark, C. (1962). The PERT model for the distribution of an activity. *Operations Research*, **10**, 405–406.

Goeree, J., Holt, C., and Palfrey, T. (2016). *Quantal Response Equilibrium: A Stochastic Theory of Games*. Princeton University Press, Princeton, NJ.

Hofbauer, J. and Sandholm, W. (2002). On the global convergence of stochastic fictitious play. *Econometrica*, **70**(6), 2265–2294.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, **14**, 124–143.

Myatt, D. P. and Wallace, C. (2009). Evolution, teamwork, and collective action: production targets in the private provisioning of public goods. *The Economic Journal*, **119**, 61–90.

Stuart, A. and Ord, K. (2010). *Kendall's Advanced Theory of Statistics. Volume 1*. Wiley, London.

Xu, Z. (2016). Convergence of best-response dynamics in extensive-form games. *Journal of Economic Theory*, **162**, 21–54.