Should law keep pace with society?

Daria Roithmayr *, Alexander Isaiov † ‡, and David Rand § ¶

*University of Southern California, Gould School of Law, Los Angeles, CA 90089, †Harvard University, Department of Physics, Cambridge MA 02138, ‡joint first author, §Yale University, Department of Psychology, Department of Economics, School of Management, New Haven, CT 06520, and ¶corresponding author

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Most modern societies have adopted centralized rules of legal punishment to promote collaborative behavior. Among other advantages, a centralized institutional punisher can unilaterally decide the rate at which legal rules evolve relative to the social behavior being regulated. Legal and political theorists disagree over whether or not law should evolve more slowly than social behavior. Some scholars argue that slower evolution promotes the stability of tradition. Others argue that more frequent adaptation permits law to remain relevant to contemporary decision-making. We investigate this question by modeling the co-evolution of law and social norms in a public goods game with centralized punishment. We vary the relative update rate of legal rules—the rate at which the State updates the legal punishment strategy relative to citizens’ updating of their contribution strategy—to observe the effect of such variance on citizen cooperation. We find that when States have unlimited resources, legal rules that evolve more slowly will maximize citizen cooperation: slower relative updating forces citizens to adapt to the State’s legal punishment rules. When States depend on citizens to finance their punishment activities, however, we find a Goldilocks effect. Citizen cooperation is maximized when legal rules evolve at a critical evolutionary rate that is slow enough to allow citizens to adapt, but fast enough to enable States to quickly respond to outbreaks of citizen lawlessness.

SOCIAL SCIENCES, BIOLOGICAL SCIENCES: Evolutionary analysis

Significance Statement

Although the majority of societies promote cooperation by using centralized institutional punishment, most research on punishment and cooperation focuses on punishment administered by individuals. We investigate a unique advantage of institutional punishment: the ability to control how quickly legal rules evolve relative to social behavior. We develop a co-evolutionary model of a public goods game with centralized punishment, which varies the pace at which legal rules evolve relative to the updating of citizens contribution strategy. We identify an ideal “Goldilocks” update rate for law: slow enough to force citizens to respond but fast enough to respond to outbreaks of citizen defection. We thus develop an argument that institutional punishment can promote cooperation in a way that individually administered punishment cannot.

Introduction

Human cooperation is a puzzle. Although a group of people collectively benefits from cooperation, individuals are better off if they can free-ride on others’ efforts. In modern society, institutional punishment plays an important role in sustaining cooperation, particularly for very large groups of people who are not culturally or genetically connected to one another. Most modern societies have adopted institutional rules of legal punishment administered by a centralized punisher, the state, to promote collaborative behavior by citizens [1–3].

Centralized punishment offers several unique advantages [4], including the ability to overcome coordination failures and free-riding problems [5], and the avoidance of anti-social punishment and retaliation [6–9]. Beyond these advantages, cen-

Reserved for Publication Footnotes
democracies, the ability to punish depends on citizens’ political support; rules of punishment that are illegitimate are disciplined at the ballot box. We investigate the effect of this state dependence on its citizenry by examining the effect of a budget constraint that funds punishment proportional to citizen contributions.

Model
Our co-evolutionary model considers a series of public goods games, with punishment within each game determined by centralized States. We consider $N$ separate States, $S_1, \ldots, S_N$, each containing a population of $n$ citizens.

In each iteration of the game, each citizen $j$ adopts a contribution strategy $c_{i,j} \in [0,1]$ which defines the amount of income the citizen pays to its State in the given period. All citizens have equal income, normalized to 1 unit per iteration. After citizens have chosen how much to contribute, every State $S_i$ chooses whether to punish each citizen based on the citizen’s contribution $c_{i,j}$. Each State $S_i$ adopts a punishment threshold $T_i$ such that all citizens contributing below this mandated threshold are punished. Punishment costs the State $\alpha$ per punished citizen, and inflicts a loss of $\gamma$ on each punished citizen.

To reflect the public goods that States provide, the sum of citizen contributions in a State $i$, $\sum_{j=1}^{n} c_{i,j}$ is multiplied by a factor $r$ with $1 < r < n$. This amount is then divided equally among all citizens in the State. Thus, citizens face a social dilemma in deciding whether to contribute to the common pool.

The payoff of citizen $j$ in State $i$ is thus given by:

$$\pi_{i,j} = 1 - c_{i,j} + \sum_{k=1}^{n} \frac{r \cdot c_{i,k}}{n} - \gamma I_{c_{i,j} < T_i}$$  \hspace{1cm} [1]$$

where $I_{c_{i,j} < T_i}$ is an indicator function that returns 1 when $c_{i,j} < T_i$ and 0 otherwise.

A State’s payoff is the aggregate payoffs of its citizens minus the cost of administering punishment to any non-compliant citizens. Thus, the payoff of State $i$ is given by:

$$\Pi_i = \sum_{j=1}^{n} \pi_{i,j} - \alpha \sum_{j=1}^{n} I_{c_{i,j} < T_i}$$  \hspace{1cm} [2]$$

After each iteration of game play, strategy updating occurs. In any given iteration, either citizens or States update their strategies. Both citizens and States update their strategies by comparing payoffs to their peers to determine whether to switch their strategies: when a given citizen updates, she compares her payoff to that of a randomly selected peer citizen from the State, and adopts the peer’s contribution strategy if the peer has a higher payoff. If both have equal payoff, the updating citizen adopts the peer’s strategy with a probability of 0.50. If the peer has a lower payoff, the citizen does not change strategy. When a given State updates, it compares its payoff to a randomly selected peer State, and adopts the peer State’s punishment threshold using the same deterministic payoff comparison process.

The relative evolutionary rate (or update rate) for States, or the number of citizen updates per State update, is given by $\theta$. Thus, every $\theta$ iterations, a State update occurs: all States compare themselves to randomly picked peer States and update simultaneously. During all other iterations, citizen updates occur: one citizen per State compares her payoffs with a peer citizen and chooses whether to switch strategies. Thus, higher values of $\theta$ represent States that update relatively more slowly than their citizens.

Both citizen and State updates are susceptible to mutation. Each time a State considers changing strategy, it picks a random strategy from a uniform distribution on $[0,1]$ with probability $\mu_s$. Similarly, each time a citizen considers changing strategy, the citizen picks a random strategy from a uniform distribution with probability $\mu_c$.

To incorporate the fact that modern states depend on citizen financial support, our later simulations modify the model to introduce a “budget constraint” parameter, $\delta$, which makes the amount of punishment depend proportionally on citizen contributions: punishment cost and effect are scaled linearly with $\delta$.

Thus in the presence of a constraint, the actual cost imposed on a punished citizen contributing under the threshold changes from $\gamma$ to:

$$\gamma \left( \sum_{j=1}^{n} \frac{c_{i,j} \delta}{n} \right) + \gamma (1 - \delta)$$  \hspace{1cm} [3]$$

and the actual cost to the State of punishing such a citizen changes from $\alpha$ to:

$$\alpha \left( \sum_{j=1}^{n} \frac{c_{i,j} \delta}{n} \right) + \alpha (1 - \delta)$$  \hspace{1cm} [4]$$

The budget constraint models the State’s need to raise capital or rely on political support to maintain a fully-functioning enforcement mechanism. Varying levels of $\delta$ from 0 to 1 correspond to different sensitivities to citizen contributions. Setting $\delta = 0$ gives the baseline model in which punishment is unconstrained by citizen contribution.

To explore the dynamics of this system, numerical simulations were performed using Matlab. A population of $N = 10$ States and $n = 10$ citizens was initialized with all citizens contributing $c = 0$ and all States having punishment thresholds of $T = 0$. The following parameters were used: $r = 3$, $\alpha = 0.5$, $\gamma = 1.25$ and $\mu = 0.05$. We then varied the update rate $\theta$ from 1 to 1000, and the constraint $\delta$ from 0 to 0.8. Every simulation was run for $10^6$ generation and results were averaged over the second half of the simulation. Results were further averaged over 10 simulation runs. Additional simulations presented in the Supporting Information demonstrate robustness of our results to variation in the State size $N$ and the mutation rates.
Results

Effect of States’ Update Rate on Evolutionary Dynamics of Cooperation.

Unconstrained $\delta=0$

Fig. 1. Effect of States’ Relative Update Rate on Cooperation. Counts represent the fraction of time that the system spends at the relevant contribution level, averaged over the last half in each of ten runs of $1 \times 10^6$ generations.) $N, n = 10, r = 3, \alpha = 0.5, \gamma = 1.25, \mu_s, \mu_c = 0.05$

Fig. 1 depicts how the distribution of citizen cooperation rates changes with State update rate $\theta$ in a setting where States have unlimited resources with which to finance punishment. (For corresponding distributions of State punishment threshold $T$ and citizen and State payoffs relative to $\theta$, see SI Fig. S1-S4). Fig. 1 shows that more slowly-updating legal rules promote greater citizen cooperation in the absence of a budget constraint. As $\theta$ increases and the States’ relative update rates become slower, citizen cooperation shifts dramatically and in a discontinuous way from no cooperation by the citizens to a high level of cooperation.

This shift in cooperation occurs because slowly the evolutionary rate of legal punishment in essence prevents the States from abandoning a costly legal punishment strategy before citizens have learned to adapt. Depriving the States of the opportunity to update effectively functions as a pre-commitment device: a relatively slower updating rate forces the States to continue punishing at significant cost where those States might otherwise have shifted to a strategy of non-punishment, to avoid the cost of punishment. A driver playing the game of chicken can win by throwing her steering wheel out the window; playing against an opponent who is disabled from changing strategy, players who swerve will be favored by natural selection. Likewise, by adopting a slower rate of updating, the State forces its citizens to adapt to State punishment rules before the States update their punishment strategy. Relatedly, some scholars in the biology literature have suggested that slower relative evolution rates produce the same results as giving the slower-evolving species the first move in a sequential game [15].

Effect of State Update Rate on Citizen Cooperation with Budget Constraint. We now consider the dynamics of the model in a setting where the State operates under a budget constraint. This budget constraint can represent a variety of ways in which the States depend on their citizens for effective punishing. For example, States that deploy centralized punishment often depend on citizen revenues to finance punishment activities. Democratic states often depend on citizen political support for particular legal punishment rules like the death penalty; in the absence of such support, government cannot continue such punishment strategies.

As described above, the States’ resources to finance legal punishment are set to be proportional to citizen contributions. For example, a constraint of 1 indicates that all States’ revenues for punishment activities are directly proportional to citizen contributions. In contrast, a constraint of 0.2 indicates that 20 percent of States’ revenues are proportional to citizen contributions and the rest are exogenously supplied with no limitation.

Fig. 2 illustrates how citizen cooperation varies with State update rate $\theta$ under constraints that vary from 0.2 to 0.8. (For corresponding distributions of State punishment threshold $T$ and citizen and State payoffs, see SI Fig. S1-S4). When $\delta = 0.2$ (Fig. 2a), the dynamics mirror those of unconstrained States at $\delta = 0$. Because the baseline punishment strength is set to $\gamma = 1.25$, States are still able to fully disincentivize defection even when all citizens contribute nothing; with no citizen support, the cost of being punished is 1, which is enough to wipe out any gains from non-cooperation. Thus, when $\delta = 0.2$, citizen cooperation increases as the States’ relative update rate $\theta$ increases.

When States are more heavily constrained, $\delta = 0.4$ (Fig. 2b), we see important differences. As States update more slowly relative to citizens, we now observe a bimodal distribution, in which citizens spend time at very high and very low levels of cooperation, with little to no time at intermediate levels of contribution. When the States are very severely constrained, $\delta = 0.6$ and $\delta = 0.8$, (Figs. 2c and 2d), citizens spend almost all their time not contributing. Thus, unlike the unconstrained State, increasing $\theta$ in the presence of constraints is not uniformly good; past the $\theta$ at which cooperation first arises, slowing the State also increases the probability of citizens contributing nothing.

Additional simulations demonstrate qualitatively equivalent results using larger States with 50 citizens each (Fig. S5).
as well as lower citizen and State mutation rates ($\mu_S = \mu_C = 0.005$; Fig. S6).

What accounts for these dynamics? More particularly, why does slowing the rate at which the State updates its legal rules have negative effects on cooperation in the presence of a non-negative budget constraint? The answer can be traced to the fact that budget constraints, which essentially link the States’ ability to punish to citizen cooperation, require the States to be able to respond quickly when a significant number of citizens switch to non-cooperation. Even in unconstrained States, from time to time, citizens shift from high levels of contribution to very low levels of contribution, defecting en masse. This dramatic shift is particularly likely when a State mutates to a new legal rule requiring substantially higher contributions. On those occasions, the State begins to punish all of its citizens, who are currently contributing a middling amount which is below the new threshold but greater than zero. Thus, a citizen mutant who contributes 0 will earn higher payoffs than the middling cooperators. Because both strategies are punished, a citizen earns higher payoffs if she avoids the cost of contribution. Accordingly, such a defector mutant can invade.

In unconstrained States, these rebellions are self-correcting. Eventually, a mutant citizen with a compliant high contribution strategy that avoids State punishment will arise and earn higher payoffs than the punished defectors. Other citizens will then copy the innovator mutant’s compliant co-operative strategy, to complete the recovery.

For constrained States, however, this recovery is significantly slowed. Owing to the budget constraints that tie punishment ability to citizen contributions, an individual State’s ability to punish non-contributing citizens is severely hindered when its citizens stop contributing. Once citizens stop contributing, the State is no longer able to fully enforce punishment to achieve the high level of cooperation that corresponds to its threshold. As a result, a compliant mutant that might arise would not be able to out-compete defectors because defectors are not adequately punished to render their payoffs less than the mutant’s.

Instead, to restore recovery, a constrained State must initially lower its threshold and then “bootstrap” its way back up to full cooperation. Eventually, a mutant State will arise that adopts a sufficiently low threshold that the State will be able to incentivize citizens. Even when citizens have defected en masse, a constrained State ($\delta < 1$) still has the power to impose a small punishment to enforce a lower threshold, such that a mutant compliant citizen contributing at a slightly greater level will earn higher payoffs than defectors. Once such a mutant fixes, contribution levels increase, and these contributions can then finance a larger punishment, favoring citizens who contribute an even larger amount, and so on.

Fig. 3 provides evidence for such bootstrapping. The key to the bootstrapping dynamic is that under a budget constraint, States with no citizen cooperation (and thus no State payoff) must lower its threshold, and make initially small and then progressively larger increases in punishment threshold $T$. The smaller increases at the beginning of the recovery generate enough citizen cooperation (and State revenue) to give States the financial ability to punish at a higher threshold, which in turn will generate even more cooperation and revenue. In this way, the State is able to bootstrap up to higher levels of punishment. Unconstrained States, on the other hand, can move directly to punish at a high $T$ without the need for intermediate “stepping stone” thresholds.

Consistent with this, Fig. 3A demonstrates that States with higher budget constraints are more likely to change their threshold during State update rounds, especially among States that are in the process of increasing their threshold. Fig. 3B then shows that among the low threshold States ($T < 0.1$) that increase their threshold, the average size of the threshold increase is larger for unconstrained and smaller for constrained States. Constrained States that are increasing their threshold change the threshold more often and in smaller increments, working their way up to the larger threshold. States without a budget constraint move in far larger increments, jumping immediately to large $T$, and thus do not change as often. Thus, as shown in Fig. 3C, States under higher constraint spend a greater fraction of their time at intermediate levels of $T$.

![Fig. 3](image-url)

**Owing to the need for bootstrapping to address outbreaks of citizen defection, slowing the State’s relative evolutionary rate becomes a double-edged sword. As States update slowly enough to force citizens to respond to punishment, citizens respond by contributing at progressively greater levels. But less**
frequent updating also gives constrained States progressively fewer opportunities to bootstrap between punishment levels in order to recover cooperation after an outbreak of lawless defection.

Put differently, States whose ability to punish depends on citizen contribution are vulnerable to citizen rebellion. Slowly evolving States that update their rules less frequently lose the flexibility and agility to respond when citizens have defected en masse and citizen norms have deviated significantly from the legal rule.

Importantly, this dynamic creates an optimal update rate for legal rules that maximizes cooperation under budget constraints. This optimal update rate is slow enough to force citizens to respond to legal punishment by adopting and copying a compliant strategy and at the same time is fast enough to maximize the States’ ability to respond quickly to citizen rebellions via bootstrapping.

This optimal evolutionary rate can be seen in Fig. 4, where average cooperation and variance in cooperation are shown as a function of $\theta$. At the optimal $\theta$, the curves for constrained States flatten dramatically, and further slowing generates no additional increase in mean citizen cooperation (and often a slight decrease in cooperation). This is because a higher $\theta$ makes things simultaneously better and worse: slower State relative updating promotes greater contribution by citizens but also promotes greater defection because the States are not able to respond quickly to mass citizen defections.

In addition, Fig. 4 shows that further slowing also generates a dramatic increase in the variance of cooperation rates, as citizens cycle between periods of high cooperation and longer periods of mass defection.

Thus, if the States’ relative update rate is too fast, States are unable to force citizens to respond. If the States’ update rate is too slow, States are unable to address mass outbreaks of citizen defection, and citizens experience volatile cycles between social order and long periods of mass defection. The ideal or critical State update rate maximizes cooperation while simultaneously minimizing citizen defection and volatility.

**Discussion**

In the last several years, a large body of cross-disciplinary research has investigated the role that peer punishment plays in generating cooperation. Most existing evolutionary game theory literature exploring punishment focuses on informal punishment administered by peer individuals—vigilante punishment that depends on social norms to enforce cooperation [16–21].

Our co-evolutionary model contributes to a nascent literature that studies the emergence of centralized institutions, which offer unique advantages in the regulation of human behavior. For example, recent work explores the relative cost advantage of institutional punishers who spread the cost of punishment among themselves [22, 23].

Our model extends the study of institutions beyond the simple cost-sharing function of group-based punishment, to investigate the unique advantages that legal punishment offers relative to decentralized punishment. As an institution, legal punishment differs from vigilante punishment precisely because a centralized institution can generate meta-rules that affect the creation, extinction and alteration of primary legal punishment rules. These meta-rules include rules about how, and how often, legal punishment rules change [24]. In particular, lawmakers can deliberately adopt procedural rules to slow or accelerate the rate at which legal rules are amended in order to promote the possibility of successful long-term citizen collaboration.

We show that a state’s ability to unilaterally vary the evolutionary rate of legal punishment rules can itself generate cooperation. Our model demonstrates that, in the absence of budgetary constraints, a state can force its citizens to adapt to rules of punishment by restricting the opportunity for legal punishment rules to evolve. Beyond longer time horizons and a check on impulsive decision-making, relatively slow updating provides the state with an additional benefit: a first-mover advantage that comes from the ability to tie its own hands.

A similar phenomenon has been observed in biological contexts, where species that evolve relatively more slowly in symbiotic or parasitic relationships can force their partners or hosts to play a more cooperative strategy. In particular, when a species evolves more slowly than its partner, the partner will adaptively respond to the species’ strategy rather than the other way around [15, 21, 24]. Likewise, when legal punishment rules evolve more slowly than citizen contribution strategies, citizens will adaptively respond to punishment rules.

At the same time, our model also illustrates the potential real-world limit of states’ first-mover advantage. Most democratic states rely on citizens for budget revenues or political legitimacy. Our model demonstrates that this dependence on the citizenry might make legal punishment vulnerable to sudden shifts in which citizens en masse abandon compliance with the rules. As the model demonstrates, any procedural rule that makes it more difficult to revise law simultaneously forces citizens to adapt to legal rules more responsively and lessens law’s ability to adapt responsively to outbreaks of citizen lawlessness.

Thus, we identify a Goldilocks effect, in which states maximize cooperation at an evolutionary rate that is both slow enough to force citizens to adaptively respond and fast enough to respond to outbreaks of citizen defection without triggering volatile cycling between high cooperation and mass defection. Such stability likely permits economic and social actors to engage in long-term planning.

Quantitative empirical evidence suggests the existence of a real-world Goldilocks effect in the context of amendment rates for national constitutions. Elkins et al. [26] document a non-monotonic relationship between a constitution’s amendment
rate and the constitution’s life expectancy. Extremely high and low amendment rates are both associated with shorter life spans. Interestingly, India’s constitution, which has adopted different amendment procedures for different topics, possesses an amendment rate that is “just right”—neither too flexible or inflexible.

The existence of an optimal relative update rate for legal rules has significant implications for both evolutionary game theory and legal scholarship. The optimal “Goldilocks” rate of evolution reflects the twin sources of power available in modern states: the power of the states to coerce cooperation and the power of citizens to rebel.

More generally, we hope our work will spur more theoretical study of the operation of institutional forms of punishment, and the relationship between the evolutionary rate of legal rules and efficient regulation of human behavior. Such work is particularly important in areas like financial regulation and tax law, where citizen behavior changes at a very rapid pace, and lawmakers often argue that law should adapt quickly in response.

Our work can also guide an emerging experimental literature that investigates how to design centralized punishment institutions to best promote cooperation [28, 29]. Further theoretical and experimental inquiry can help us understand how best to design institutions that appropriately balance the restraining pull of stability and tradition with the adaptive push of transformation and change.


11. The Records of the Federal Convention of 1787 (Max Farrand ed. 1911) at 121 (Madison’s notes, June 5, 1787).
Supporting Information

for

Should law keep pace with society?

Daria Roithmayr, Alexander Isakov, David G. Rand
Fig. S1. Effect of States' Relative Update Rate on Cooperation. Counts represent the fraction of time that the system spends at the relevant level, averaged over the last half in each of ten runs of $1 \times 10^6$ generations in all SI figures. $N = n = 10$; $r = 3$; $\alpha = 0.5$; $\gamma = 1.25$; $\mu_s = \mu_c = 0.05$. (This figure reproduces Figs 1 and 2 from the main text, and is presented here for completeness.)
Fig. S2. Effect of States' Relative Update Rate on State Strategy (Punishment Threshold). N = n = 10; r = 3; alpha = 0.5; gamma = 1.25; mu_s = mu_c = 0.05
Fig. S3. Effect of States' Relative Update Rate on Citizen Payoff. N = n = 10; r = 3; alpha = 0.5; gamma = 1.25; mu_s = mu_c = 0.05
Fig. S4. Effect of States' Relative Update Rate on State Payoff. N = n = 10; r = 3; alpha = 0.5; gamma = 1.25; mu_s = mu_c = 0.05
Fig. S5. Robustness to state size (n=50). Effect of States' Relative Update Rate on Cooperation. N = 10, n = 50; r = 3; alpha = 0.5; gamma = 1.25; mu_a = mu_b = 0.05. Final panel shows the average value from each of the earlier panels.
Fig. S6. Robustness to mutation rate (\(\mu_s = \mu_c = 0.005\)). Effect of States' Relative Update Rate on Cooperation. \(N = 10\), \(n = 10\); \(r = 3\); \(\alpha = 0.5\); \(\gamma = 1.25\); \(\mu_s = \mu_c = 0.005\). Final panel shows the average value from each of the earlier panels.