Graduated Punishments in Public Good Games

van der Made, Allard

Published in:
Southern Economic Journal

DOI:
10.1002/soej.12308

Southern Economic Journal 2019, 85(3), 939–959
DOI: 10.1002/soej.12308

Graduated Punishments in Public Good Games

Allard van der Made

I explain the ubiquitous use of graduated punishments by studying a repeated public good game in which a social planner imperfectly monitors agents to detect shirkers. Agents' cost of contributing is private information and administering punishments is costly. Using graduated punishments can be optimal for two reasons. It increases the price of future wrongdoing (temporal spillover effect) and it can lead to bad types revealing themselves (screening effect). The temporal spillover effect is always present if graduated punishments prevail, but screening need not occur if agents face a finite horizon. Whether or not a screening effect is exploited has a substantial impact on both outcomes and actual punishments. If the temporal spillover effect is sufficiently strong, then first-time shirkers are merely warned.

JEL Classification: D82, H41, K49

1. Introduction

A host of social situations involve collective action problems: from the point of view of the collective it is best if everybody acts in the interest of the group, yet it is individually optimal to act differently. Examples include tax avoidance, the tragedy of the commons, using polluting production technologies, and vote abstention. In many instances groups or societies have managed to induce individuals to behave in the interest of the collective. One important factor ensuring that individuals do so is the presence of a monitoring institution that is able to punish (alleged) wrongdoers. Successful punishment schemes often exhibit graduated sanctions: repeat offenders are punished more severely than first-time offenders (e.g., Ostrom 1990; Ellickson 1991; Wade 1994; Ostrom 2000; Agrawal 2003). Graduated sanctions also appear in many judiciary systems.[1] In its most extreme form graduated punishments are such that first-time offenders receive a mere warning.

I present a theory that explains the prevalence of graduated punishments and show that using graduated punishments is often optimal if monitoring is imperfect, administering punishments is costly, and agents differ with respect to how tempted they are to choose the selfish action. Graduated punishments can prevail both in a setting where agents are only supposed to contribute to a public good twice (today and tomorrow, say) and in an infinite-horizon setting in which agents are supposed to contribute in each period of their life.

By considering more general punishment strategies than the extant literature does I am able to bring together the two main rationales for using graduated punishments suggested by scholars: it

IEEF, University of Groningen, Nettelbosje 2, 9747AE, Groningen, The Netherlands; E-mail: a.van.der.made@rug.nl.
Received October 2015; accepted July 2018.

[1] For example, various state governments in the United States have enacted *Three Strikes Laws.* Such laws require state courts to hand down a mandatory and extended period of incarceration to persons who have been convicted of a serious criminal offense on three or more separate occasions.

increases the price of future wrongdoing (the *temporal spillover effect*) and it can lead to bad types revealing themselves (the *screening effect*). This synthesis leads to the following novel insights. I show that if graduated punishments are used, then a temporal spillover effect is necessarily present but screening need not occur. There are substantial differences between both the outcomes and the punishments associated with graduated punishments with and without a screening effect. This framework also allows one to quantify the temporal spillover effect. This enables me to endogenize the use of warnings. Furthermore, I uncover key differences between settings with a finite horizon and those with an infinite horizon. These stem from the fact that screening always occurs if graduated punishments are used in the latter setting.

In this model, a social planner faces a repeated public good problem. It is socially efficient if all agents contribute to the public good in each period, but an agent incurs a cost each time he contributes. The planner monitors the behavior of individual agents, but this monitoring is imperfect: some noncontributors (shirkers) dodge being detected and some contributors are found guilty of shirking. The planner can administer punishments to alleged shirkers and keeps track of who has been punished at least once. Administering punishments is costly for society.[2] The individually borne cost of contributing differs among agents and is either high or low. An agent's type is private information. The planner maximizes welfare, that is, the social benefits of the contributions to the public good minus all costs.

Because punishing agents is costly, using a punishment that is sufficiently severe to deter all agents from shirking need not be optimal. Indeed, in a one-shot setting this is only optimal if the number of high-cost types is sufficiently large. If this number is small, then the social costs of erroneously administering severe punishments to a large group of low-cost types outweigh the benefits of deterring a small group of high-cost types from shirking. The planner consequently uses a low punishment that only induces low-cost types to contribute. If agents are not only supposed to contribute today, but also in the future, then the planner can often improve upon the outcome of the one-shot setting by using graduated punishments.

Using graduated punishments instead of a uniform punishment improves welfare for two reasons. First, the mere threat of becoming "branded" a shirker makes agents reluctant to shirk today if those who have already been punished in the past receive a harsher punishment than those who have never been punished before. As monitoring is imperfect and hence contributors are occasionally punished, being caught shirking today always increases expected future punishments, even if the agent plans to contribute in future periods. This temporal spillover effect (Mungan 2014) enables the planner to reduce the punishment for first-time shirkers below the low punishment of the one-shot setting.

The temporal spillover effect (TSE) is particularly strong if expected future punishments are relatively important vis-à-vis the costs that an agent (potentially) incurs today. It can even lead to first-time offenders receiving no punishment at all, that is, they are merely warned. This occurs if it is impossible to reduce the punishment for first-time shirkers by the desired amount as this would lead to a negative punishment. Warnings consequently arise endogenously in this framework. As expected future punishments are relatively large if the probability that the planner falsely judges someone guilty of shirking is large, the model predicts that warnings are often used if it is difficult to gauge agents' behavior.

The second effect causing graduated punishments to emerge is the screening effect. By imposing a mild sanction today the planner is able to (imperfectly) sort agents by type: only low types

[2] These costs include the administrative and legal costs associated with penalizing someone.

contribute if today's punishment is relatively low.[3] This implies that the bulk of agents who are punished are high types whereas agents who are not punished are predominantly low types. As a consequence, the planner can in the future (again imperfectly) tailor punishments to types by imposing a harsh punishment on alleged repeat offenders—that is, agents who have been punished in the past—and a moderate one on alleged first-time offenders. This enables the planner to induce a given level of contributions in a more cost-efficient way.

Whether graduated punishments are optimal and how they are implemented depends on the horizon agents' face. In the two-period setting one cannot create a TSE to incentivize agents to contribute in the last period, but the planner can treat first-time offenders differently across periods. This possibility opens up a plethora of strategies. For example, if the welfare gains stemming from contributions are very low, then it can be optimal to create a very strong TSE in period 1 by not punishing first-time offenders in period 2. If the welfare gains are not very low, then there are only two ways to design a graduated punishments scheme: one that only induces the high types who are punished in period 1 to contribute in period 2 and one that also induces high types to contribute in period 1. The first option exploits both a TSE and a screening effect whereas the second option only exploits a TSE. As the first option leads to less contributions by high types, it is preferred if the fraction of high types is relatively low. In the infinite-horizon setting endgame effects are absent and the planner simply chooses between a uniform punishment and graduated punishments that are such that only the high types who have been punished in the past contribute. In the infinite-horizon setting both the TSE and the screening effect are always exploited whenever graduated punishments are used.

Using graduated punishments is not always optimal.[4] If society consists mainly of high types, then using graduated punishments would yield a very low level of public good provision and it is better to use a high uniform punishment that deters all agents from shirking. By contrast, if the vast majority of the agents are low types, then ignoring the high types and using a low uniform punishment is only optimal if the planner is unable to fully exploit the TSE and therefore has to resort to warnings. The reason is that compared to a uniform punishment an ideal graduated punishment scheme merely shifts punishments administered to low types between periods without affecting the aggregate punishments administered to them. However, such an ideal graduated punishment scheme is only feasible if the TSE can be fully exploited.

The remainder of this article is organized as follows. In section 2, I relate my work to the literature. Section 3 introduces the main ingredients of the model and contains the derivation of the optimal punishment scheme of the one-shot setting. In section 4, I analyze the two-period setting. Section 5 deals with the infinite-horizon setting. Section 6 offers concluding remarks. Proofs and further technicalities are relegated to Appendix A and Supporting Information Online Appendices.

2. Relation to the Literature

Various theoretical explanations for the ubiquitous use of graduated punishments have been proposed. Funk (2004) and Miceli and Bucci (2005) argue that the dire labor market prospects of convicted criminals makes committing crimes relatively more attractive for those who already have

[3] From now on I abbreviate low-cost type and high-cost type to low type and high type.

[4] Of course, if the planner would never falsely judge someone guilty of shirking, then it is optimal to use a high uniform punishment: this leads to contributions by all agents at zero punishment costs.

a criminal record. This effect can be negated by punishing repeat offenders harsher than first-time offenders. If offenders learn how to evade apprehension, as in Mungan (2010), then the *expected* punishment a repeat offender faces is lower than the expected punishment a first-time offender faces should the *actual* punishment remain the same. It is then optimal to set the actual punishment for repeat offenders higher than that for first-time offenders. However, if law enforcers learn more from past offenses than the criminals themselves, then the optimal punishment for repeat offenders is lower than the one for first-time offenders.[5] In Miles and Pyne (2015), a criminal gradually discovers how good he is at it. As a criminal's own perception of this ability depends positively on the number of past convictions, one needs to use an escalating penalty schedule to ensure that expected punishments are sufficiently deterring.

Stigler (1970) argued informally that heavy penalties are unnecessary for first-time offenders if they are likely to have committed the offense by mistake and the probability of repetition is negligible. In Rubinstein (1979) offenses may also have been committed by accident. Convicting such "innocent offenders" is detrimental to welfare. It is then optimal to be lenient toward individuals with a "reasonable" criminal record and merely warn them. Erroneous convictions also play a central role in Chu, Hu, and Huang (2000). Because it is very unlikely that an innocent person commits an offense twice, the vast majority of the repeat offenders are real criminals, and hence it is optimal to punish them more severely than first-time offenders.[6] As punishing those who did commit crimes is costless, their planner does not face a trade-off between crime prevention and cost minimization comparable to my trade-off between public good provision and cost minimization.

Emons (2007) considers a two-period model with a homogeneous population in which wealth-constrained agents either commit a crime in both periods (i.e., pursue a criminal career) or always abide the law. Each period a law-abiding agent might commit a crime by accident. To deter agents from pursuing a criminal career the planner, who aims to minimize apprehension costs, either uses a maximal fine for first-time offenders and no fine for repeat offenders or no fine for first-time offenders and a maximal fine for repeat offenders. Using a maximal fine for repeat offenders is optimal if the benefits of pursuing a criminal career are large compared to an agent's initial wealth. Mungan (2014) also considers a two-period setting with a homogeneous population. He shows that if individuals occasionally behave irrationally by committing a crime without taking into account its consequences, then it can be optimal to use graduated punishments. This result hinges on the presence of a TSE. Yet, in Mungan (2014), this effect stems from the fact that agents anticipate that they might behave irrationally in the future, not by monitoring mistakes.

Polinsky and Rubinfeld (1991) study a setting with perfect monitoring. An individual's gain from committing a crime is partly socially acceptable and partly illicit. The planner maximizes aggregate acceptable gains minus harms stemming from criminal activities by choosing fines for first and second offenses. Individuals who commit crimes in the first period are likely to enjoy high illicit gains. This fact allows the planner to sort agents by "illicit type." Using higher fines for second offenses reduces both underdeterrence vis-à-vis low uniform fines and overdeterrence vis-à-vis high uniform fines, making graduated fines often socially optimal.

Sorting also plays a prominent role in Miceli (2013). In his model, there is a maximal sanction which is too low to deter all agents from committing a crime. If the number of undeterrable offenders is large, then it is optimal to impose high sanctions on repeat offenders and no sanction

[5] Dana (2001) provides ample arguments in favor of a higher probability of detection for repeat offenders.

[6] This observation echoes the argument put forward by Stigler (1970).

on first-time offenders. Like in my setting, the planner thereby saves on punishment costs and tailors period 2-punishments to types. However, the optimality of graduated punishments in Miceli (2013) stems from the presence of undeterrable offenders: if their number is small, then it is optimal to punish first-time offenders severely and not punish repeat offenders at all.

In Polinsky and Shavell (1998) monitoring is perfect, apprehending offenders are costly and punishments cannot exceed an upper bound. As using differentiated punishments creates a TSE, it can be optimal to set the punishment for first-time offenders in period 2 below the upper bound and the one for repeat offenders equal to the upper bound. This increases crime deterrence in period 1, but reduces deterrence in period 2. Polinsky and Shavell's TSE has no impact on the punishment that prevails in period 1.

Harrington (1988), studying the enforcement of compliance with environmental regulations, shows that a higher compliance rate (compared to a system with a uniform punishment) can be achieved if firms with relatively good compliance records are merely warned. This stems from the fact that using differentiated punishments creates a TSE. As Harrington (1988) assumes perfect monitoring, this result hinges on the presence of an upper bound on punishments. The regulator in Nyborg and Telle (2004) has insufficient budget to cover all sanctioning expenditures should too many firms violate regulations. There are consequently two equilibria: a "good equilibrium" in which all firms intend to comply and a "bad equilibrium" in which no firm complies. As firms occasionally violate regulations by accident, the economy can switch from the good equilibrium to the bad one. Warnings can be used to reduce the probability that switching occurs. Rousseau (2009) argues that the use of warnings reduces the number of erroneous convictions and at the same time mitigates overcompliance. Mungan (2013) shows that if some individuals do not know that a certain act is illegal and punishing uninformed individuals are costly, then it can be optimal to issue warnings to first-time offenders.[7]

Landsberger and Meilijson (1982) study how tax evasion is best combatted in a dynamic setting with an exogenously given penalty system and a homogeneous population. They show that if the tax authority does not have the means to audit everybody, then tax revenues are often maximal if those who have been caught evading taxes in the previous period are audited with a higher probability than those who have not been caught evading taxes in the previous period.

Endres and Rundshagen (2012) study punishments schemes in an infinite-horizon setting. They show that in the presence of an upper bound on punishments a given level of crime control can be achieved at the lowest costs by using graduated punishments.

3. The Environment

A social planner faces a public good problem. If a fraction π of a population with total mass 1 contributes to the public good, then the total social benefits of the public good amount to π. Contributing to the public good is costly. A fraction $1 - \rho$ of the population consists of agents who incur the high cost γ_H when contributing, where $\gamma_H < 1$. The remaining fraction ρ consists of agents who incur the low cost γ_L when contributing, where $\gamma_L \in (0, \gamma_H)$. Because $\gamma_H < 1$, it is socially optimal if all agents contribute. Yet, for both types of agents it is individually optimal to refrain from

[7] Punishing uninformed individuals in Mungan (2013) is akin to making a type II error and subsequently punishing an innocent person in my model.

contributing, that is, to shirk. An agent's type (low or high) is private information. All agents are risk-neutral. I use the subscript L (H) to refer to low (high) types.

The planner monitors agents' behavior, enabling her to punish alleged shirkers. The planner's monitoring technology is flawed: with probability $\epsilon_I \in \left(0,\frac{1}{2}\right)$, she fails to detect a shirker (a *type I error*) and with probability $\epsilon_{II} \in \left(0,\frac{1}{2}\right)$, she erroneously judges someone guilty of shirking (a *type II error*). So, only a fraction $1 - \epsilon_I$ of the shirkers are caught, whereas a fraction ϵ_{II} of the contributors are found guilty of shirking. The difference $\phi = 1 - \epsilon_I - \epsilon_{II}$ between these two fractions measures the quality of the monitoring technology: the larger ϕ, the less often monitoring mistakes are made. I assume that monitoring agents is free, but that administering punishments is costly. Specifically, if the planner administers a punishment that reduces an agent's utility by f, then society bears a cost of cf, where $c > 0$.[8]

The planner maximizes welfare by choosing the punishments administered to alleged shirkers. These punishments are made public before the contribution stage. I assume that the planner can commit to the announced punishments and that the punishment costs never exhaust the planner's budget. Welfare W consists of the social benefits of the public good, the individually borne costs of contributing, and the cost of administering punishments. Hence:

$$W = \rho(1-\gamma_L)\delta_L + (1-\rho)(1-\gamma_H)\delta_H - F,$$

where $\delta_L = 1$ ($\delta_L = 0$) if low types (do not) contribute, $\delta_H = 1$ ($\delta_H = 0$) if high types (do not) contribute, and F denotes the social costs of administering punishments. These costs are

$$F = \rho(\delta_L \epsilon_{II} c + (1-\delta_L)(1-\epsilon_I)c)f_0 + (1-\rho)(\delta_H \epsilon_{II} c + (1-\delta_H)(1-\epsilon_I)c)f_0,$$

where f_0 is the punishment that alleged shirkers face.

I assume that the laissez-faire outcome in which punishments are zero and no agent contributes is never optimal, even if all agents are high types ($\rho = 0$). I therefore maintain the following condition throughout the article:

Condition 1. Laissez-faire is never optimal, specifically: $1-\gamma_H > \epsilon_{II} c \frac{\gamma_H}{\phi}$.

Before I study the two-period setting and the infinite-horizon setting I derive the planner's optimal strategy in the one-shot setting. The one-shot outcome serves as a benchmark for the other settings: as the planner can always obtain the one-shot outcome in each period by using the optimal punishment of the one-shot setting, the analysis of the one-shot setting yields a lower bound on the per-period welfare that can be attained in the other settings.

In the one-shot setting an agent contributes if the associated expected costs do not exceed the expected costs the agent faces when shirking.[9] A low type consequently contributes if $\gamma_L + \epsilon_{II} f_0 \leq (1 - \epsilon_I) f_0$, that is, if $f_0 \geq \frac{\gamma_L}{\phi}$. Similarly, a high type contributes if $f_0 \geq \frac{\gamma_H}{\phi}$. So, the planner chooses between the *low (uniform) punishment* ($f_0 = \frac{\gamma_L}{\phi}$) which only induces low types to contribute and the *high (uniform) punishment* ($f_0 = \frac{\gamma_H}{\phi}$) which ensures that high types also contribute. Comparing welfare associated with the two possibilities yields:

[8] Because punishing an innocent person is in general seen as a grave injustice, it seems natural to assume that society bears an extra cost mf, where $m > 0$, when an innocent person is punished. Including such "injustice costs" does not change our results qualitatively. Details are available from the author upon request.

[9] I assume that an agent contributes if he is indifferent between contributing and shirking.

PROPOSITION 1. In the one-shot setting the social planner opts for

$$\phi f_0^* = \begin{cases} \gamma_H & \text{if } \rho \leq \overline{\rho} \\ \gamma_L & \text{if } \rho > \overline{\rho}, \end{cases} \tag{1}$$

where $\overline{\rho} := 1 - \frac{\epsilon_{II} c \frac{\gamma_H - \gamma_L}{\phi}}{1 - \gamma_H + c\gamma_L} \in (0,1)$.

As administering punishments is costly, it need not be optimal to induce all agents to contribute by using the high punishment. If the number of high types is small, then the increase in contributions caused by moving from the low punishment to the high one is small. This move would also entail administering a higher, more costly punishment to a fraction ϵ_{II} of the low types. If $\rho > \overline{\rho}$, this negative effect dominates the positive effect of more contributions and hence the planner opts for the low punishment.

4. The Two-Period Setting

Agents are now supposed to contribute to the public good twice: in period 1 and in period 2. An agent's type remains constant across periods and is private information. The planner recalls in period 2 whether or not she has punished a given agent in period 1. Just like in the one-shot setting the planner makes a type i error with probability ϵ_i when investigating an agent's behavior, $i \in \{I, II\}$. Drawing the wrong conclusion regarding an agent's behavior in period 1 does not affect the probability with which she misjudges that agent's behavior in period 2.

Recalling who has been punished in period 1 enables the planner to use differentiated punishments in period 2, one for agents who have not been punished in period 1 (f_2) and one for agents who have been punished in period 1 ($\hat{f}_2$). The planner can only use one punishment (f_1) in period 1. She announces all punishments at the start of the game. Each agent uses backward induction to arrive at his optimal strategy. The timing of the game is as follows:

0. The planner announces the punishments.
1a. Each agent either contributes or shirks.
1b. The planner carries out investigations and administers punishments.
2a. Each agent either contributes or shirks.
2b. The planner carries out investigations and administers punishments.

Payoffs are realized at the end of each period. An agent minimizes his total expected costs. The planner maximizes total welfare $\mathcal{W}$, the sum of welfare in period 1 (W_1) and welfare in period 2 (W_2), with respect to the punishment scheme $\boldsymbol{f} = \left(f_1, f_2, \hat{f}_2\right)$.

Using differentiated punishments can be optimal for two reasons. First, if $\hat{f}_2 > f_2$, then the possibility of receiving the severe punishment $\hat{f}_2$ instead of the mild punishment f_2 in period 2 can deter agents from shirking in period 1. So, the threat of increased future punishments incentivizes agents to contribute today. This rationale for punishing repeat offenders more severely is the TSE. Second, if low types and high types behave differently in period 1, then the planner can distill information regarding an agent's type from his behavior in period 1 and subsequently use this to tailor period 2-punishments to types. This is the screening effect.

The TSE can even increase total welfare if the population is homogeneous ($\rho = 0$ or $\rho = 1$). By setting $f_2 = 0$ and using a sufficiently large $\hat{f}_2$ the planner is able to markedly reduce f_1 without losing the period 1-contributions. The downside of this is that agents who are not punished in period 1 do not contribute in period 2. Yet, if the net benefits of contributions are sufficiently small compared to the costs of administering punishments, then this strategy is optimal and graduated punishments are consequently even used if the population is homogeneous.[10] To exclude this possibility, we need to impose the following condition:

Condition 2. The net benefits of contributions are not too small, specifically:

$$(1-\epsilon_{II})(1-\gamma_H) > (1+\phi)\epsilon_{II} c \frac{\gamma_H}{\phi}.$$

I maintain Condition 2 throughout the remainder of this section. It not only ensures that the TSE does not suffice to end up with differentiated punishments if the population is homogeneous, it also implies that it is socially optimal to incentivize low types to always contribute.[11]

It remains to determine which of the high types are incentivized to contribute. There are three groups of high types: high types in period 1 (group $\mathcal{H}_1$), high types in period 2 who were not punished in period 1 (group $\mathcal{H}_2$), and high types in period 2 who were punished in period 1 (group $\hat{\mathcal{H}}_2$). The two extreme cases—none of the groups contribute and all of the groups contribute—boil down to the outcomes of the one-shot setting. By using the uniform punishment $f_1 = f_2 = \hat{f}_2 = f_0^*$, the planner obtains the optimal one-shot outcome in both periods. It turns out that if the planner does not use this uniform punishment, then she uses a punishment scheme that either induces only group $\hat{\mathcal{H}}_2$ (*outcome* $\{\hat{\mathcal{H}}_2\}$) or both group $\mathcal{H}_1$ and group $\hat{\mathcal{H}}_2$ (*outcome* $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$) to contribute.

Let's explain why only these two possibilities can be optimal.[12] First, if it is optimal to induce all high types to contribute in period 2, then it must also be optimal to induce them to contribute in period 1. To see this note that the planner obtains outcome $\{\mathcal{H}_2, \hat{\mathcal{H}}_2\}$ by using the low uniform punishment in period 1 and the high uniform punishment in period 2. Of course, the resulting total welfare is lower than that of the *optimal* uniform punishment in both periods.

Outcome $\{\mathcal{H}_2\}$ is always worse than outcome $\{\hat{\mathcal{H}}_2\}$. By moving from outcome $\{\mathcal{H}_2\}$ to outcome $\{\hat{\mathcal{H}}_2\}$ one increases the fraction of high types contributing in period 2 from ϵ_I to $1 - \epsilon_I$. This move is accomplished by interchanging the values of f_2 and $\hat{f}_2$, thereby causing a TSE for low types which in turn creates room for a reduction in f_1 and hence lower punishment costs in period 1. As the increase in period 2-punishment costs is negligible compared to the positive effects of the move, outcome $\{\mathcal{H}_2\}$ is never optimal.

Outcome $\{\mathcal{H}_1, \mathcal{H}_2\}$ is always worse than the high uniform punishment. Moving from the former to the latter obviously increases aggregate contributions. Surprisingly, this move can be accomplished without significantly changing the punishments. This stems from the fact that the punishments supporting outcome $\{\mathcal{H}_1, \mathcal{H}_2\}$ cause a *negative* TSE: to ensure that group $\hat{\mathcal{H}}_2$ does not contribute whereas group $\mathcal{H}_2$ does $\hat{f}_2$ must be smaller than f_2. This creates an incentive to shirk in period 1. To counteract this effect the planner needs to set all punishments as close to the high uniform

---

[10] I would like to thank an anonymous referee for pointing out this possibility.
[11] See Appendix A for details.
[12] Detailed derivations can be found in the Supporting Information Appendix.

punishment as possible. One therefore gets the contributions from group $\hat{\mathcal{H}}_2$ almost for free by instead using the high uniform punishment.

Last, outcome $\{\mathcal{H}_1\}$ is always worse than outcome $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$. To induce group $\hat{\mathcal{H}}_2$ to contribute one has to increase $\hat{f}_2$. The resulting difference between $\hat{f}_2$ and f_2 causes a TSE and hence allows the planner to decrease f_1. This decrease is so large that the total punishment costs required to obtain outcome $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ are lower than those required to obtain outcome $\{\mathcal{H}_1\}$. So, one gets the contributions from $\hat{\mathcal{H}}_2$ while saving costs.

Let's turn attention to the outcomes than can be optimal, starting with the optimal punishment schemes associated with outcome $\{\hat{\mathcal{H}}_2\}$:

Lemma 1. Outcome $\{\hat{\mathcal{H}}_2\}$ is obtained at minimal punishment costs by using a punishment scheme $\boldsymbol{f}$ satisfying $\phi f_2 = \gamma_L$ and

i. $\phi f_1 + \epsilon_{II}\phi\hat{f}_2 = (1+\epsilon_{II})\gamma_L$ with $\phi f_1 \geq 0$ and $\phi\hat{f}_2 \geq \gamma_H$ if $\gamma_L \geq \epsilon_{II}(\gamma_H - \gamma_L)$;
ii. $\phi f_1 = 0$ and $\phi\hat{f}_2 = \gamma_H$ if $\gamma_L < \epsilon_{II}(\gamma_H - \gamma_L)$.

Any optimal punishment scheme supporting outcome $\{\hat{\mathcal{H}}_2\}$ is such that $f_1 < f_2 < \hat{f}_2$. The difference between f_2 and $\hat{f}_2$ causes a TSE, allowing the planner to reduce f_1 below the low uniform punishment without losing the period 1-contributions from the low types. If $\gamma_L \geq \epsilon_{II}(\gamma_H - \gamma_L)$, then $f_1 = \frac{\gamma_L}{\phi} - \epsilon_{II}\left(\hat{f}_2 - f_2\right)$. The reduction $\epsilon_{II}\left(\hat{f}_2 - f_2\right)$ is the increase in a low type's expected period 2-punishment stemming from being punished in period 1. If $\gamma_L < \epsilon_{II}(\gamma_H - \gamma_L)$, then the planner cannot fully exploit the TSE: $\frac{\gamma_L}{\phi} - \epsilon_{II}\left(\hat{f}_2 - \frac{\gamma_L}{\phi}\right)$ is negative for any $\hat{f}_2 \geq \frac{\gamma_H}{\phi}$. The planner therefore merely *warns* alleged period 1-shirkers ($f_1 = 0$).

I now consider outcome $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$:

Lemma 2. Outcome $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ is obtained at minimal punishment costs by using a punishment scheme $\boldsymbol{f}$ satisfying $\phi f_2 = \gamma_L$ and $\phi f_1 + \epsilon_{II}\phi\hat{f}_2 = (\epsilon_I + \epsilon_{II})\gamma_H + (1-\epsilon_I)\gamma_L$ with $\phi f_1 \geq 0$ and $\phi\hat{f}_2 \geq \gamma_H$.

I again have that $f_1 < \hat{f}_2$ and $f_2 < \hat{f}_2$. However, whether $f_1 \leq f_2$ or $f_1 > f_2$ now depends on the choice for $\hat{f}_2$. Whereas the low types determine the extent of the TSE for outcome $\{\hat{\mathcal{H}}_2\}$, it is now determined by the high types. If a high type is punished in period 1, then he contributes in period 2 and his expected period 2-costs are $\gamma_H + \epsilon_{II}\hat{f}_2$. If he is not punished in period 1, then he does not contribute in period 2 and his expected period 2-costs are $(1 - \epsilon_I)f_2$. The planner exploits the TSE by subtracting the difference between these two expected costs from the *high* uniform punishment. Indeed, we have that $f_1 = \frac{\gamma_H}{\phi} - \left(\gamma_H + \epsilon_{II}\hat{f}_2 - (1-\epsilon_I)f_2\right)$. In contrast to outcome $\{\hat{\mathcal{H}}_2\}$, in which the TSE must be subtracted from the *low* uniform punishment, the TSE can now always be fully exploited.

I now compare the total welfare for the outcomes $\{\hat{\mathcal{H}}_2\}$ and $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ and the optimal uniform punishment, starting with the case $\gamma_L \geq \epsilon_{II}(\gamma_H - \gamma_L)$:

Proposition 2. Suppose $\gamma_L \geq \epsilon_{II}(\gamma_H - \gamma_L)$. Then there exists a $\check{\rho} \in (0, \overline{\rho})$ such that the planner maximizes total welfare by

i. *using the high uniform punishment if* $\rho \leq \check{\rho}$;

ii. *inducing outcome* $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ *if* $\rho \in (\check{\rho}, \overline{\rho}]$*;*
iii. *inducing outcome* $\{\hat{\mathcal{H}}_2\}$ *if* $\rho > \overline{\rho}$*.*

If the population consists mainly of high types, then the planner opts for the uniform punishment. As ρ increases the advantages of using graduated punishments start playing a role. If $\rho = \check{\rho}$, then the increase in welfare caused by the TSE geared to the high types exactly offsets the loss in the contributions from group $\mathcal{H}_2$. As long as $\rho \leq \overline{\rho}$ the fraction of high types is relatively large and the planner does not want to lose the contributions from group $\mathcal{H}_1$.

If $\rho > \overline{\rho}$, the planner succumbs to the other rationale for using graduated punishments: the screening effect. By using a period 1-punishment that is such that only low types contribute in period 1 the planner is able to (imperfectly) sort agents by type, enabling her to tailor period 2-punishment to types. Given that only high types shirk in period 1 most of the agents who are punished in period 1 are high types. As they can only be deterred from shirking by a punishment of at least $\frac{\gamma_H}{\phi}$, the planner uses a punishment $\hat{f}_2 \geq \frac{\gamma_H}{\phi}$ for repeat offenders. On the other hand, an agent who is not found guilty of shirking in period 1 is most likely a low type and the punishment $f_2 = \frac{\gamma_L}{\phi}$ therefore suffices to deter most of the agents who were not punished in period 1 from shirking in period 2.

By exploiting the screening effect the planner obtains the period 2-contributions at relatively low costs, but she has to forgo the contributions from group $\mathcal{H}_1$. Outcome $\{\hat{\mathcal{H}}_2\}$ is therefore only optimal if $\rho \geq \overline{\rho}$. Of course, any screening is vacuous if the population is homogeneous. The welfare difference between outcome $\{\hat{\mathcal{H}}_2\}$ and the low uniform punishment therefore vanishes as ρ approaches 1. Yet, as any punishment scheme with the property $\phi f_1 + \epsilon_{II}\phi\hat{f}_2 = (1+\epsilon_{II})\gamma_L$, including the low uniform punishment, results in the same *total* punishment costs attributable to (law-abiding) low types, the planner never strictly prefers the low uniform punishment.

The welfare comparisons are less straightforward if $\gamma_L < \epsilon_{II}(\gamma_H - \gamma_L)$:

PROPOSITION 3. Suppose $\gamma_L < \epsilon_{II}(\gamma_H - \gamma_L)$. Then there exists a possibly empty interval $(\tilde{\rho}_\ell, \tilde{\rho}_r) \subset (\overline{\rho}, 1)$ such that the planner maximizes total welfare by[13]

i. *using the high uniform punishment if* $\rho \leq \check{\rho}$*;*
ii. *inducing outcome* $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ *if* $\rho \in (\check{\rho}, \tilde{\rho}_\ell]$*;*
iii. *inducing outcome* $\{\hat{\mathcal{H}}_2\}$ *if* $\rho \in (\tilde{\rho}_\ell, \tilde{\rho}_r)$*;*
iv. *using the low uniform punishment if* $\rho \geq \tilde{\rho}_r$*,*

where $\check{\rho}$ has the same value as in Proposition 2.

Compared to the case $\gamma_L \geq \epsilon_{II}(\gamma_H - \gamma_L)$, the planner induces outcome $\{\hat{\mathcal{H}}_2\}$ less often. Since the TSE geared to low types cannot be fully exploited if $\gamma_L < \epsilon_{II}(\gamma_H - \gamma_L)$ and the planner therefore resorts to warning alleged period 1-shirkers if she wants to induce outcome $\{\hat{\mathcal{H}}_2\}$, this outcome loses ground to both outcome $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ and the low uniform punishment. Because the planner is indifferent between outcome $\{\hat{\mathcal{H}}_2\}$ and the low uniform punishment for $\rho = 1$ *only if* the TSE could be fully exploited, the low uniform punishment now outperforms outcome $\{\hat{\mathcal{H}}_2\}$ for ρ sufficiently

[13] If the interval is empty, then $\tilde{\rho}_\ell = \tilde{\rho}_r$.

large. Since the TSE geared to the high types can always be fully exploited, the fraction of low types $\widetilde{\rho}_\ell$ for which the planner is indifferent between outcome $\{\mathcal{H}_1, \hat{\mathcal{H}}_2\}$ and outcome $\{\hat{\mathcal{H}}_2\}$ starts increasing as $\gamma_L - \epsilon_{II}(\gamma_H - \gamma_L)$ becomes negative.

Propositions 2 and 3 reveal that it is often optimal to use graduated punishments. These punishments are such that repeat offenders receive a harsher punishment than first-time offenders. Furthermore, first-time offenders in general face a different punishment across periods. This is an artifact of the finite horizon of the two-period setting: as the game ends after period 2, the planner cannot exploit a TSE to reduce f_2. There is also only one opportunity to sort agents by type and the screening effect can only be exploited during a single period. Let's now turn attention to a setting which does not have these features.

5. The Infinite-Horizon Setting

The public good game is now repeated *ad infinitum.* Each period consists of three stages. In the first stage each agent chooses between contributing to the public good and shirking. In the second stage the social planner carries out investigations and punishes agents who have been found guilty of shirking. Payoffs are realized directly after this punishment stage. In the last stage a fraction $1 - \beta \in (0, 1)$ of the population dies and is replaced by new agents. Each agent advances to the next period with probability β, that is, this probability does not depend on an agent's type, how often he has shirked, or his punishment history. The population is again characterized by the parameters ρ, γ_L, and γ_H. In particular, a fraction ρ of each generation and thus of the total population in any period consists of low types. An agent's type remains constant over time. Since $\gamma_H < 1$, it is socially optimal that every agent contributes in each period that he lives.

The planner keeps track of whether or not a given agent has been punished in the past. The quality of her monitoring technology is again given by the probabilities ϵ_I and ϵ_{II}. She announces all punishments that could be administered during a given period at the start of that period, before agents decide whether or not to contribute. She can use two different punishments, one for agents who have never been punished before (agents with a *clean record*) and one for agents who have been punished at least once, or she can use the same punishment for all alleged shirkers.

The population can be divided in four groups: low types with a clean record, low types who have been punished at least once, high types with a clean record, and high types who have been punished at least once. I call the set of all agents with a clean record q and the set of all agents who have been punished before $\hat{q}$. I focus on the *stationary equilibria* of the repeated game, that is, the equilibria that can prevail if the composition of the population with respect to the above categorization remains unaltered as the economy moves from some period to the next one. A stationary equilibrium is supported by a punishment f for agents in q, a punishment $\hat{f}$ for agents in $\hat{q}$, and a contribution rule for each group.

Each period the planner maximizes the welfare generated in that period.[14] An agent minimizes his expected current and discounted future costs. The only difference between the planner and the

[14] Note that a strategy that supports a stationary equilibrium of the present game also supports the corresponding stationary equilibrium of the game in which the planner maximizes current welfare plus discounted future welfare, irrespective of the discount rate.

agents regarding their attitude toward the future stems from the fact that the planner is immortal whereas agents die with probability $1-\beta$ at the end of a period. I therefore use β as the agents' discount factor. Note that at any time the expected number of future periods that an agent stays in the economy, that is, an agent's (remaining) life expectancy, is equal to $\frac{\beta}{1-\beta}$.

In contrast to the two-period setting, I do not need a condition like Condition 2 to ensure that graduated punishments are never optimal if the population is homogeneous. The reason is that in a stationary equilibrium of the infinite-horizon setting agents cannot be treated differently across periods. The planner can only make punishments contingent on whether an agent has received a punishment in the past, not on the period itself. So, the punishments f_1 and f_2 of section 4 are now replaced by the single punishment f and the two groups $\mathcal{H}_1$ and $\mathcal{H}_2$ are now replaced by a single group $\mathcal{H}$, the group of high types with a clean record. Condition 1 now also suffices to ensure that low types always contribute in equilibrium. To see this suppose that low types in q contribute and that those in $\hat{q}$ shirk. Then high types in $\hat{q}$ definitely shirk and we end up with either contributions from all agents in q or with only contributions from the low types in q. The first outcome results in less welfare than with the high uniform punishment and the second one results in less welfare than with the low uniform punishment. A similar reasoning reveals that incentivizing only low types in $\hat{q}$ to contribute cannot be optimal. So, if the planner does not opt for a uniform punishment, then she uses a punishment scheme that induces all low types to contribute and on top of that either group $\mathcal{H}$ (*outcome* $\{\mathcal{H}\}$) or group $\hat{\mathcal{H}}$ (*outcome* $\{\hat{\mathcal{H}}\}$), where group $\hat{\mathcal{H}}$ consists of the high types in $\hat{q}$.

Inducing outcome $\{\mathcal{H}\}$ always yields lower per-period welfare than the high uniform punishment.[15] The reason is that the punishments supporting outcome $\{\mathcal{H}\}$ cause a negative TSE: to ensure that group $\hat{\mathcal{H}}$ shirks whereas group $\mathcal{H}$ contributes $\hat{f}$ must be lower than f. This creates an incentive to shirk for agents in q. The planner is consequently best off setting both punishments as close to the high uniform punishment as possible. By instead using the high uniform punishment the planner also obtains contributions from group $\hat{\mathcal{H}}$ without significantly increasing the punishment costs.

The optimal punishment pairs for outcome $\{\hat{\mathcal{H}}\}$ are as follows:

LEMMA 3. Outcome $\{\hat{\mathcal{H}}\}$ is obtained at minimal punishment costs by using a punishment pair $(f,\hat{f})$ satisfying

i. $(1-\beta)\phi f+\beta\epsilon_{II}\phi\hat{f}=(1-\beta+\beta\epsilon_{II})\gamma_L$ with $\phi f\geq 0$ and $\phi\hat{f}\geq\gamma_H$ if $(1-\beta)\gamma_L\geq\beta\epsilon_{II}(\gamma_H-\gamma_L)$;
ii. $\phi f=0$ and $\phi\hat{f}=\gamma_H$ if $(1-\beta)\gamma_L<\beta\epsilon_{II}(\gamma_H-\gamma_L)$.

Any optimal punishment pair supporting outcome $\{\hat{\mathcal{H}}\}$ is such that $f<\hat{f}$. These pairs are reminiscent of the pairs $(f_1,\hat{f}_2)$ given in Lemma 1. In fact, if $\beta=\frac{1}{2}$, then these two collections of pairs coincide. This is not very surprising: if $\beta=\frac{1}{2}$, then an agent's life expectancy is 1, which equals the number of future periods of an agent in the two-period setting.

If $(1-\beta)\gamma_L\geq\beta\epsilon_{II}(\gamma_H-\gamma_L)$, then $f=\frac{\gamma_L}{\phi}-\frac{\beta}{1-\beta}\epsilon_{II}\left(\hat{f}-\frac{\gamma_L}{\phi}\right)$. So, agents in q face a punishment that equals the low uniform punishment reduced by $\frac{\beta}{1-\beta}\epsilon_{II}\left(\hat{f}-\frac{\gamma_L}{\phi}\right)$. This reduction is equal to an agent's

[15] Detailed derivations regarding the infinite-horizon setting can be found in the Supporting Information Appendix.

life expectancy times the probability that an agent is erroneously found guilty of shirking times the difference between $\hat{f}$ and the low uniform punishment. This is the increase in a low type's total expected future punishments if he would move to $\hat{q}$ and the low uniform punishment would be used for agents in q. If $(1-\beta)\gamma_L < \beta\epsilon_{II}(\gamma_H - \gamma_L)$, then the planner cannot fully exploit the TSE and she therefore merely warns alleged shirkers in q. This happens in particular if β is sufficiently large, that is, an agent's life expectancy is sufficiently high. This is intuitive: the larger β is, the more important future costs are relative to costs incurred in the current period and the more agents fear moving to $\hat{q}$ and the lower f consequently can be.

Comparing the per-period welfare associated with the optimal uniform punishment f_0^* given in Equation 1 with that generated if outcome $\{\hat{\mathcal{H}}\}$ is induced yields:

Proposition 4. There are two cases:

- If $(1-\beta)\gamma_L \geq \beta\epsilon_{II}(\gamma_H - \gamma_L)$, then there exists a $\breve{\rho} \in (0,\overline{\rho})$ such that the planner maximizes total welfare by
  i. using the high uniform punishment if $\rho \leq \breve{\rho}$;
  ii. inducing outcome $\{\hat{\mathcal{H}}\}$ *if* $\rho > \breve{\rho}$.
- If $(1-\beta)\gamma_L < \beta\epsilon_{II}(\gamma_H - \gamma_L)$, then there exists a $\breve{\rho} \in (0,\overline{\rho})$ and a $\widetilde{\rho} \in (\overline{\rho},1)$ such that the planner maximizes total welfare by
  i. using the high uniform punishment if $\rho \leq \breve{\rho}$;
  ii. inducing outcome $\{\hat{\mathcal{H}}\}$ *if* $\rho \in (\breve{\rho},\widetilde{\rho}]$;
  iii. using the low uniform punishment if $\rho > \widetilde{\rho}$.

In the infinite-horizon setting the planner makes similar choices as in the two-period setting. If ρ is small, then the planner uses the high uniform punishment to ensure that all agents contribute in each period. As long as the planner can fully exploit the TSE, then for all $\rho > \breve{\rho}$ she uses graduated punishments. By using graduated punishments the planner (imperfectly) sorts agents in q by type, allowing her to tailor future punishments to type and hence obtain contributions from high types at relatively low costs. As the total benefits of this strategy increase with the agents' life expectancy $\frac{\beta}{1-\beta}$ the threshold $\breve{\rho}$ decreases in β. If $(1-\beta)\gamma_L < \beta\epsilon_{II}(\gamma_H - \gamma_L)$, then the planner cannot fully exploit the TSE. This makes the low uniform punishment relatively more attractive vis-à-vis outcome $\{\hat{\mathcal{H}}\}$ and the planner consequently opts for the low uniform punishment if ρ is sufficiently large.

Note that it is very likely that a low type spends a large part of his life in $\hat{q}$ if β is large: as monitoring is imperfect, the probability that a law-abiding agent is found guilty of shirking at least once in t periods goes to 1 as $t \to \infty$. Administering the high punishment $\hat{f}$ to these low types is clearly suboptimal: the punishment $\frac{\gamma_L}{\phi}$ suffices to deter them from shirking. This observation suggests that it might be a bad idea to keep an agent who has been found guilty of shirking once in the high-punishment regime $\hat{q}$ for the rest of his life.[16] The planner could alternatively reward those in $\hat{q}$ who (allegedly) contributed in the last period with a clean slate, that is, move them back to q and treat them in the next period as if they had never shirked. In a working paper version of this article (van der Made, 2015) I show that the planner prefers rewarding good behavior in such a fashion to

[16] It could also be argued that it is rather harsh to do this given that the planner occasionally judges an innocent person guilty of shirking.

using the optimal uniform punishment as long as the population is relatively heterogeneous. However, rewarding good behavior has also disadvantages: it leads to more shirking by high types and a weakening of the TSE. It is therefore ambiguous if and when the planner should combine graduated punishments with rewarding good behavior.

6. Concluding Remarks

I have determined the welfare-maximizing punishment strategies for a social planner who is confronted with a repeated public good problem. Because monitoring is imperfect and administering punishments is costly, a uniform punishment is often suboptimal. To alleviate the detrimental effects on welfare of monitoring mistakes and costly punishments the planner can use a punishment scheme featuring graduated punishments: repeat offenders are punished harsher than first-time offenders. Such a punishment scheme allows the planner to (imperfectly) sort agents by cost type, enabling her to tailor future punishments to type. Moreover, the threat of facing harsher punishments in the future makes agents more reluctant to shirk today. This temporal spillover effect enables the planner to sanction first-time offenders very mildly. In fact, merely warning first-time offenders sometimes suffices.

This framework not only applies to classic public good situations, but also to law enforcement problems. The cost of contributing to the public good is then replaced by the opportunity cost of not committing a crime. Most crimes bestow a negative externality upon society at large. This ranges from commonly felt disgust following a gruesome murder to a reduction in the safety of online services caused by cyber crimes. Not engaging in criminal activities therefore increases aggregate welfare in a similar fashion as contributing to a public good does.

Obviously, one can envision more elaborate punishment strategies than the ones I have analyzed. For instance, in most judiciary systems the punishment a convicted criminal receives does not simply depend on whether this person has a criminal record, but also on the precise content of such a record. Furthermore, we have only looked at the stationary equilibria of the infinite-horizon setting. I have consequently left an important question unanswered: under what conditions do societies reach steady states in which graduated punishments are used? These issues might prove fruitful avenues for future research.

Appendix A

Details Regarding Condition 1

Suppose that $\rho = 0$. To induce agents to contribute the planner has to set a punishment f such that $\gamma_H + \epsilon_{II} f \leq (1 - \epsilon_I) f$ and hence the planner opts for $f^* = \frac{\gamma_H}{\phi}$. The associated welfare reads $W(f^*) = 1 - \gamma_H - \epsilon_{II} c \frac{\gamma_H}{\phi}$, which is positive if $1 - \gamma_H > \epsilon_{II} c \frac{\gamma_H}{\phi}$ holds.

Proof of Proposition 1

Welfare with the low uniform punishment equals

$$W\left(\frac{\gamma_L}{\phi}\right) = \rho(1-\gamma_L) - \rho\epsilon_{II} c \frac{\gamma_L}{\phi} - (1-\rho)(1-\epsilon_I) c \frac{\gamma_L}{\phi} = \rho(1-\gamma_L) - \epsilon_{II} c \frac{\gamma_L}{\phi} - (1-\rho) c \gamma_L. \tag{A.1}$$

If the planner uses the high uniform punishment, then welfare becomes

$$W\left(\frac{\gamma_H}{\phi}\right)=\rho(1-\gamma_L)+(1-\rho)(1-\gamma_H)-\epsilon_{II}c\frac{\gamma_H}{\phi}. \tag{A.2}$$

The welfare difference $\Delta=\Delta(\rho)=W\left(\frac{\gamma_H}{\phi}\right)-W\left(\frac{\gamma_L}{\phi}\right)$ reads

$$\Delta=(1-\rho)(1-\gamma_H)-\epsilon_{II}c\frac{\gamma_H}{\phi}+\epsilon_{II}c\frac{\gamma_L}{\phi}+(1-\rho)c\gamma_L=(1-\rho)(1-\gamma_H+c\gamma_L)-\epsilon_{II}c\frac{\gamma_H-\gamma_L}{\phi}.$$

Solving $\Delta(\rho)=0$ yields $\rho=\overline{\rho}$. The claims now follow from the facts that $\Delta^{'}(\rho)<0$ and $\Delta(1)<0<\Delta(0)$.

Details Regarding Condition 2

I first show that graduated punishments are never optimal if the population is homogeneous and Condition 2 holds. Let $\xi\in\{\gamma_L,\gamma_H\}$ be the extant type's cost of contributing and suppose that $\hat{f}_2$ is such that agents who are punished in period 1 do contribute in period 2.

Assume first that f_2 is such that an agent who is not punished in period 1 does contribute in period 2, that is, there is no TSE. Then an agent contributes in period 1 if:

$$\xi+\epsilon_{II}\left(f_1+\xi+\epsilon_{II}\hat{f}_2\right)+(1-\epsilon_{II})(\xi+\epsilon_{II}f_2)\le(1-\epsilon_I)\left(f_1+\xi+\epsilon_{II}\hat{f}_2\right)+\epsilon_I(\xi+\epsilon_{II}f_2).$$

The left-hand side of this constraint consists of the expected costs an agent faces when contributing in period 1. It equals the costs of contributing twice plus the expected costs associated with being erroneously punished in one or both periods. The right-hand side consists of the expected costs of an agent if he shirks in period 1, given that he plans to always contribute in period 2. This constraint simplifies to

$$\phi f_1-\epsilon_{II}\phi f_2+\epsilon_{II}\phi\hat{f}_2\ge\xi. \tag{A.3}$$

Since an agent who is not punished in period 1 only contributes in period 2 if $\phi f_2\ge\xi$, Inequality A.3 has to be evaluated at $\phi f_2=\xi$. This yields $\phi f_1\ge(1+\epsilon_{II})\xi-\epsilon_{II}\phi\hat{f}_2$.

Assume now that f_2 is such that an agent who is not punished in period 1 shirks in period 2. Then an agent contributes in period 1 if:

$$\xi+\epsilon_{II}\left(f_1+\xi+\epsilon_{II}\hat{f}_2\right)+(1-\epsilon_{II})(1-\epsilon_I)f_2\le(1-\epsilon_I)\left(f_1+\xi+\epsilon_{II}\hat{f}_2\right)+\epsilon_I(1-\epsilon_I)f_2.$$

Since agents shirk in period 2 if they have not been punished in period 1, ξ is not included in the f_2–parts of this constraint. This constraint reduces to

$$\phi f_1-(1-\epsilon_I)\phi f_2+\epsilon_{II}\phi\hat{f}_2\ge(\epsilon_I+\epsilon_{II})\xi. \tag{A.4}$$

Evaluating it at $f_2=0$ (any $f_2>0$ would merely increase costs) yields $\phi f_1\ge(\epsilon_I+\epsilon_{II})\xi-\epsilon_{II}\phi\hat{f}_2$. This is the constraint the planner has to take into account if she opts to exploit the TSE.

Compare the two situations. Since the behavior of agents who are punished in period 1 is the same in both cases, so is $\hat{f}_2$. The reduction in the optimal period 1-punishment caused by the TSE is therefore $(1+\epsilon_{II})\xi-(\epsilon_I+\epsilon_{II})\xi=(1-\epsilon_I)\xi$. The TSE also reduces f_2 from $\frac{\xi}{\phi}$ to 0. On the other hand, the agents who were not punished in period 1 (a fraction $1-\epsilon_{II}$ of the population) no longer contribute in period 2 if the TSE is being used. Since a fraction ϵ_{II} of the population is punished in period 1 and of those who are not a fraction ϵ_{II} receives the punishment f_2 in period 2, the increase in total welfare stemming from the TSE is

$$T(\xi)=\epsilon_{II}c(1-\epsilon_I)\frac{\xi}{\phi}+(1-\epsilon_{II})\epsilon_{II}c\frac{\xi}{\phi}-(1-\epsilon_{II})(1-\xi)=(1+\phi)\epsilon_{II}c\frac{\xi}{\phi}-(1-\epsilon_{II})(1-\xi),$$

which is negative if Condition 2 holds.

I next show that low types always contribute if Condition 2 holds. First, any punishment scheme which is such that all low types shirk in one of the two periods boils down to laissez-faire in that period and is therefore suboptimal. Since $T(\gamma_L)<0$, it is suboptimal to "allow" low types who were not punished in period 1 to shirk in period 2 if $\rho=1$. If $\rho\in(0,1)$ and one "allows" non-punished low types to shirk in period 2, then non-punished high types also shirk in period 2. The change in total welfare is then thus either $\rho T(\gamma_L) < 0$ (if nonpunished high types were going to shirk anyway) or $\rho T(\gamma_L) + (1-\rho)T(\gamma_H) < 0$.

It remains to prove that "allowing" low types who have been punished in period 1 to shirk in period 2 cannot be optimal. Suppose first that low types shirk in period 2 if punished in period 1. Then they contribute in period 1 if:

$$\gamma_L + \epsilon_{II}\left(f_1 + (1-\epsilon_I)\hat{f}_2\right) + (1-\epsilon_{II})(\gamma_L + \epsilon_{II}f_2) \le (1-\epsilon_I)\left(f_1 + (1-\epsilon_I)\hat{f}_2\right) + \epsilon_I(\gamma_L + \epsilon_{II}f_2).$$

This constraint reduces to $\phi f_1 \ge (1+\phi)\gamma_L + \epsilon_{II}\phi f_2 - (1-\epsilon_I)\phi\hat{f}_2$. On the other hand, if low types do contribute in period 2 if punished in period 1, then they contribute in period 1 if Inequality A.3 with ξ replaced by γ_L holds, that is, if $\phi f_1 \ge \gamma_L + \epsilon_{II}\phi f_2 - \epsilon_{II}\phi\hat{f}_2$. Because punished low types (do not) contribute in period 2 if $\phi\hat{f}_2 \ge \gamma_L$ ($\phi\hat{f}_2 < \gamma_L$), the change in the period 1-punishment that is minimally required to induce low types who have been punished in period 1 to contribute in period 2 is negative. So, the impact on period 1-welfare of "allowing" punished low types to shirk in period 2 is negative. Furthermore, Condition 1 implies that shirking by those low types reduces period 2-welfare. We conclude that "allowing" low types who have been punished in period 1 to shirk in period 2 is never optimal.

Proof of Lemma 1

Group $\hat{\mathcal{H}}_2$ contributes if $\hat{f}_2 \ge \frac{\gamma_H}{\phi}$ and group $\mathcal{H}_2$ shirks if $f_2 < \frac{\gamma_H}{\phi}$. Furthermore, all low types contribute in period 2 if both f_2 and $\hat{f}_2$ are at least $\frac{\gamma_L}{\phi}$. Given that low types always contribute in period 2 they also contribute in period 1 if Equation A.3 with ξ replaced by γ_L holds, that is, if

$$\phi f_1 + \epsilon_{II}\phi\hat{f}_2 - \epsilon_{II}\phi f_2 \ge \gamma_L. \tag{A.5}$$

Given that high types only contribute in period 2 when punished in period 1 they shirk in period 1 if Equation A.4 with ξ replaced by γ_H does not hold, that is, if

$$\phi f_1 + \epsilon_{II}\phi\hat{f}_2 - (1-\epsilon_I)\phi f_2 < (\epsilon_I + \epsilon_{II})\gamma_H. \tag{A.6}$$

Total welfare with outcome $\{\hat{\mathcal{H}}_2\}$ equals

$$\begin{aligned}\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2) = {} & 2\rho(1-\gamma_L) + (1-\rho)(1-\epsilon_I)(1-\gamma_H) - \rho\epsilon_{II}cf_1 - (1-\rho)(1-\epsilon_I)cf_1 \\ & -\rho\epsilon_{II}^2 c\hat{f}_2 - \rho(1-\epsilon_{II})\epsilon_{II}cf_2 - (1-\rho)(1-\epsilon_I)\epsilon_{II}c\hat{f}_2 - (1-\rho)\epsilon_I(1-\epsilon_I)cf_2.\end{aligned}$$

The first two terms of $\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2)$ are the net benefits of the contributions. The next two terms are the punishment costs of period 1. Since only low types contribute in period 1, a fraction ϵ_{II} of the low types and a fraction $1-\epsilon_I$ of the high types receive the punishment f_1. The last four terms make up the punishment costs of period 2. Since low types always contribute, a fraction ϵ_{II} of them receive a punishment (either $\hat{f}_2$ or f_2) in period 2. Only the high types who are punished in period 1 contribute in period 2, implying that a fraction $(1-\epsilon_I)\epsilon_{II}$ of the high types receive the punishment $\hat{f}_2$ and a fraction $\epsilon_I(1-\epsilon_I)$ of them receive the punishment f_2.

The planner maximizes $\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2)$ subject to $\phi f_1 \ge 0$, $\gamma_L \le \phi f_2 < \gamma_H$, $\phi\hat{f}_2 \ge \gamma_H$, Inequality A.5, and Inequality A.6. The Lagrangian of the program in which the strict inequalities $\phi f_2 < \gamma_H$ and Inequality A.6 are replaced by weak inequalities is

$$\begin{aligned}\mathcal{L}(\boldsymbol{f},\boldsymbol{\lambda},\boldsymbol{\mu};\hat{\mathcal{H}}_2) = {} & \mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2) - \lambda_1\left(\gamma_L - \phi f_1 - \epsilon_{II}\phi\hat{f}_2 + \epsilon_{II}\phi f_2\right) - \lambda_2\left(\phi f_1 + \epsilon_{II}\phi\hat{f}_2 - (1-\epsilon_I)\phi f_2 - (\epsilon_I + \epsilon_{II})\gamma_H\right) + \mu_1\phi f_1 \\ & -\mu_2(\gamma_L - \phi f_2) - \mu_3(\phi f_2 - \gamma_H) - \mu_4\left(\gamma_H - \phi\hat{f}_2\right),\end{aligned}$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ *and* $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ are shadow prices. The first order conditions (FOCs) of the maximization problem consequently include:

$$-(1-\rho)c\phi - \epsilon_{II}c + \phi\lambda_1 - \phi\lambda_2 + \phi\mu_1 = 0, \tag{A.7}$$

$$-\rho(1-\epsilon_{II})\epsilon_{II}c - (1-\rho)\epsilon_I(1-\epsilon_I)c - \epsilon_{II}\phi\lambda_1 + (1-\epsilon_I)\phi\lambda_2 + \phi\mu_2 - \phi\mu_3 = 0, \tag{A.8}$$

$$-(1-\rho)\epsilon_{II}c\phi - \epsilon_{II}^2 c + \epsilon_{II}\phi\lambda_1 - \epsilon_{II}\phi\lambda_2 + \phi\mu_4 = 0. \tag{A.9}$$

On top of these there are six complementary slackness conditions (one for each inequality constraint), the original constraints must hold, and the shadow prices must be nonnegative. Condition A.8 can only hold if either $\lambda_2 > 0$ or $\mu_2 > 0$. Furthermore, Condition A.7 and Condition A.9 together imply that μ_1 and μ_4 must have the same sign.

Suppose $\mu_1 > 0$ and $\mu_4 > 0$. Then, $\phi f_1 = 0$ and $\phi \hat{f}_2 = \gamma_H$. If $\lambda_2 > 0$, then the complementary slackness condition pertaining to (A.6) implies that $\phi f_2 = -\frac{\epsilon_I \gamma_H}{1-\epsilon_I} < 0$, violating the constraint $\phi f_2 \geq \gamma_L$. So, $\mu_2 > 0$ and hence $\phi f_2 = \gamma_L$. One can verify that all constraints are now satisfied as long as $\gamma_L \leq \epsilon_{II}(\gamma_H - \gamma_L)$.

Suppose $\mu_1 = \mu_4 = 0$. Then Condition A.7 can only hold if $\lambda_1 > 0$, implying that $\phi f_1 + \epsilon_{II} \phi \hat{f}_2 = \epsilon_{II} \phi f_2 + \gamma_L$. Because Inequality A.5 and Inequality A.6 cannot both hold with equality without violating the constraint $\phi f_2 \geq \gamma_L$, I have that $\lambda_2 = 0$ and therefore $\mu_2 > 0$, which gives $\phi f_2 = \gamma_L$. It follows that $\phi f_1 + \epsilon_{II} \phi \hat{f}_2 = (1+\epsilon_{II})\gamma_L$. One can only find $f_1 \geq 0$ and $\hat{f}_2 \geq \frac{\gamma_H}{\phi}$ such that this equality holds as long as $\gamma_L \geq \epsilon_{II}(\gamma_H - \gamma_L)$.

Since the two constraints that are strict inequalities in the planner's problem are never binding in the optimum, I have indeed obtained the optimal punishment schemes.

Proof of Lemma 2

The only constraint that differs from its counterpart for outcome $\{\hat{\mathcal{H}}_2\}$ is the one pertaining to high types in period 1. Since high types now do contribute in period 1, Inequality A.6 is replaced by

$$\phi f_1 + \epsilon_{II} \phi \hat{f}_2 - (1-\epsilon_I)\phi f_2 \geq (\epsilon_I + \epsilon_{II})\gamma_H. \tag{A.10}$$

Total welfare now reads

$$\mathcal{W}(\boldsymbol{f}; \mathcal{H}_1, \hat{\mathcal{H}}_2) = 2\rho(1-\gamma_L) + (1-\rho)(1+\epsilon_{II})(1-\gamma_H) - \epsilon_{II} c f_1 - \epsilon_{II}^2 c \hat{f}_2 - \rho(1-\epsilon_{II})\epsilon_{II} c f_2 - (1-\rho)(1-\epsilon_{II})(1-\epsilon_I) c f_2.$$

Because low types contribute *a forteriori* if high types do, Inequality A.5 is redundant. I therefore maximize $\mathcal{W}(\boldsymbol{f}; \mathcal{H}_1, \hat{\mathcal{H}}_2)$ subject to $\phi f_1 \geq 0$, $\gamma_L \leq \phi f_2 \leq \gamma_H$, $\phi \hat{f}_2 \geq \gamma_H$, and Inequality A.10. As $\phi f_2 \leq \gamma_H$ turns out to be non-binding at the optimum, the solution of this program coincides with the solution of the planner's problem. The Lagrangian of the program is

$$\begin{aligned}\mathcal{L}(\boldsymbol{f}, \lambda, \boldsymbol{\mu}; \mathcal{H}_1, \hat{\mathcal{H}}_2) = \mathcal{W}(\boldsymbol{f}; \mathcal{H}_1, \hat{\mathcal{H}}_2) - \lambda\Big((\epsilon_I + \epsilon_{II})\gamma_H - \phi f_1 - \epsilon_{II}\phi \hat{f}_2 + (1-\epsilon_I)\phi f_2\Big) \\ + \mu_1 \phi f_1 - \mu_2(\gamma_L - \phi f_2) - \mu_3(\phi f_2 - \gamma_H) - \mu_4\Big(\gamma_H - \phi \hat{f}_2\Big),\end{aligned}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$. Besides the complementary slackness conditions, the original constraints, and the fact that shadow prices must be nonnegative, I have the following FOCs:

$$\begin{gathered}-\epsilon_{II} c + \phi\lambda + \phi\mu_1 = 0, \\ -\rho(1-\epsilon_{II})\epsilon_{II} c - (1-\rho)(1-\epsilon_{II})(1-\epsilon_I) c - (1-\epsilon_I)\phi\lambda + \phi\mu_2 - \phi\mu_3 = 0, \\ -\epsilon_{II}^2 c + \epsilon_{II}\phi\lambda + \phi\mu_4 = 0.\end{gathered}$$

The second FOC can only be satisfied if $\mu_2 > 0$, implying that $\phi f_2 = \gamma_L$. The other two FOCs reveal that μ_1 and μ_4 must have the same sign. If $\mu_1 > 0$ and $\mu_4 > 0$, then $\phi f_1 = 0$ and $\phi \hat{f}_2 = \gamma_H$. However, Inequality A.10 does not hold for $\boldsymbol{f} = \left(0, \frac{\gamma_L}{\phi}, \frac{\gamma_H}{\phi}\right)$. So, $\mu_1 = \mu_4 = 0$. It follows that $\lambda > 0$ and hence $\phi f_1 + \epsilon_{II}\phi \hat{f}_2 = (\epsilon_I + \epsilon_{II})\gamma_H + (1-\epsilon_I)\phi f_2 = \gamma_L + \epsilon_I(\gamma_H - \gamma_L) + \epsilon_{II}\gamma_H$. One can always find $f_1 \geq 0$ and $\hat{f}_2 \geq \frac{\gamma_H}{\phi}$ such that this equality holds.

Remarks Regarding Welfare Expressions

To prove Propositions 2–3 one has to compare $\mathcal{W}(\hat{\mathcal{H}}_2)$, $\mathcal{W}(\mathcal{H}_1, \hat{\mathcal{H}}_2)$, and $\mathcal{W}^0$, where $\mathcal{W}(\hat{\mathcal{H}}_2)$ ($\mathcal{W}(\mathcal{H}_1, \hat{\mathcal{H}}_2)$) is $\mathcal{W}(\boldsymbol{f}; \hat{\mathcal{H}}_2)$ ($\mathcal{W}(\boldsymbol{f}; \mathcal{H}_1, \hat{\mathcal{H}}_2)$) evaluated at an optimal punishment scheme, and $\mathcal{W}^0$ is the total welfare if the optimal uniform punishment is used, that is:

$$\mathcal{W}^0 = \begin{cases} 2W\left(\dfrac{\gamma_H}{\phi}\right) & \text{if } \rho \leq \overline{\rho} \\ 2W\left(\dfrac{\gamma_L}{\phi}\right) & \text{if } \rho > \overline{\rho} \end{cases},$$

where $W\left(\frac{\gamma_L}{\phi}\right)$ is given in Equation A.1 and $W\left(\frac{\gamma_H}{\phi}\right)$ is given in Equation A.2. Let $\Delta_2 = \Delta_2(\rho) = \mathcal{W}(\hat{\mathcal{H}}_2) - \mathcal{W}^0$, $\Delta_{12} = \Delta_{12}(\rho) = \mathcal{W}(\mathcal{H}_1, \hat{\mathcal{H}}_2) - \mathcal{W}^0$, and $\hat{\Delta} = \hat{\Delta}(\rho) = \mathcal{W}(\mathcal{H}_1, \hat{\mathcal{H}}_2) - \mathcal{W}(\hat{\mathcal{H}}_2)$. To simplify the comparisons, rewrite $\mathcal{W}(\boldsymbol{f}; \hat{\mathcal{H}}_2)$ as follows:

$$\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2)=2\rho(1-\gamma_L)+(1-\rho)(1-\epsilon_I)(1-\gamma_H)-\epsilon_{II}cf_1-(1-\rho)c\phi f_1-\epsilon_{II}^2c\hat{f}_2$$
$$-(1-\rho)\epsilon_{II}c\phi\hat{f}_2-\rho(1-\epsilon_{II})\epsilon_{II}cf_2-(1-\rho)\epsilon_I(1-\epsilon_I)cf_2.$$

Furthermore, for every parameter configuration:

$$\mathcal{W}(\mathcal{H}_1,\hat{\mathcal{H}}_2)=2\rho(1-\gamma_L)+(1-\rho)(1+\epsilon_{II})(1-\gamma_H)-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H}{\phi}-\epsilon_{II}(1-\epsilon_I)c\frac{\gamma_L}{\phi}$$
$$-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)(1-\epsilon_{II})(1-\epsilon_I)c\frac{\gamma_L}{\phi}.$$

Proof of Proposition 2

The welfare difference between outcomes $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ *and* $\{\hat{\mathcal{H}}_2\}$ is given by:

$$\begin{aligned}\hat{\Delta}&=(1-\rho)(\epsilon_I+\epsilon_{II})(1-\gamma_H)+\epsilon_{II}(1+\epsilon_{II})c\frac{\gamma_L}{\phi}+(1-\rho)(1+\epsilon_{II})c\gamma_L\\&\quad+\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}+(1-\rho)\epsilon_I(1-\epsilon_I)c\frac{\gamma_L}{\phi}-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H}{\phi}-\epsilon_{II}(1-\epsilon_I)c\frac{\gamma_L}{\phi}\\&\quad-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)(1-\epsilon_{II})(1-\epsilon_I)c\frac{\gamma_L}{\phi}\\&=(1-\rho)(\epsilon_I+\epsilon_{II})(1-\gamma_H+c\gamma_L)-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H-\gamma_L}{\phi},\end{aligned}$$

where I used that $\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2)$ must be evaluated at $\phi f_2=\gamma_L$ and $\phi f_1+\epsilon_{II}\phi\hat{f}_2=(1+\epsilon_{II})\gamma_L$. Note that $\hat{\Delta}(\overline{\rho})=0$ and that $\hat{\Delta}$ is strictly decreasing in ρ. It follows that outcome $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ results in higher (lower) total welfare than outcome $\{\hat{\mathcal{H}}_2\}$ *if* $\rho\leq\overline{\rho}$ ($\rho>\overline{\rho}$).

For $\rho\leq\overline{\rho}$, I thus have to compare $\mathcal{W}(\mathcal{H}_1,\hat{\mathcal{H}}_2)$ with $\mathcal{W}^0=2W\left(\frac{\gamma_H}{\phi}\right)$:

$$\begin{aligned}\Delta_{12}&=\mathcal{W}(\mathcal{H}_1,\hat{\mathcal{H}}_2)-2W\left(\frac{\gamma_H}{\phi}\right)=-(1-\rho)(1-\epsilon_{II})(1-\gamma_H)-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H}{\phi}\\&\quad-\epsilon_{II}(1-\epsilon_I)c\frac{\gamma_L}{\phi}-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)(1-\epsilon_{II})(1-\epsilon_I)c\frac{\gamma_L}{\phi}+2\epsilon_{II}c\frac{\gamma_H}{\phi}\\&=-(1-\rho)(1-\epsilon_{II})(1-\gamma_H+c\gamma_L)+\epsilon_{II}(1+\phi)c\frac{\gamma_H-\gamma_L}{\phi}.\end{aligned}$$

Note that

$$\Delta_{12}(\overline{\rho})=-(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_H-\gamma_L}{\phi}+\epsilon_{II}(1+\phi)c\frac{\gamma_H-\gamma_L}{\phi}>0$$

and that Δ_{12} is strictly increasing in ρ. Furthermore:

$$\Delta_{12}(0)=-(1-\epsilon_{II})(1-\gamma_H+c\gamma_L)+\epsilon_{II}(1+\phi)c\frac{\gamma_H-\gamma_L}{\phi}<-\epsilon_{II}(1+\phi)c\frac{\gamma_H}{\phi}-(1-\epsilon_{II})c\gamma_L+\epsilon_{II}(1+\phi)c\frac{\gamma_H-\gamma_L}{\phi}<0,$$

where the first inequality stems from Condition 2. There consequently exists a $\check{\rho}\in(0,\overline{\rho})$ such that outcome $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ results in higher (lower) welfare than the high uniform punishment if $\rho>\check{\rho}$ ($\rho<\check{\rho}$).

For $\rho>\overline{\rho}$, I have to compare $\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2)$ evaluated at $\phi f_2=\gamma_L$ and $\phi f_1+\epsilon_{II}\phi\hat{f}_2=(1+\epsilon_{II})\gamma_L$ with $\mathcal{W}^0=2W\left(\frac{\gamma_L}{\phi}\right)$:

$$\begin{aligned}\Delta_2&=\mathcal{W}(\hat{\mathcal{H}}_2)-2W\left(\frac{\gamma_L}{\phi}\right)=(1-\rho)(1-\epsilon_I)(1-\gamma_H)-\epsilon_{II}(1+\epsilon_{II})c\frac{\gamma_L}{\phi}\\&\quad-(1-\rho)(1+\epsilon_{II})c\gamma_L-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)\epsilon_I(1-\epsilon_I)c\frac{\gamma_L}{\phi}+2\epsilon_{II}c\frac{\gamma_L}{\phi}+2(1-\rho)c\gamma_L\\&=(1-\rho)(1-\epsilon_I)(1-\gamma_H+c\gamma_L)\geq 0.\end{aligned}$$

So, outcome $\{\hat{\mathcal{H}}_2\}$ outperforms the optimal uniform punishment for all $\rho>\overline{\rho}$.

Proof of Proposition 3

The welfare difference between outcomes $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}\{\hat{\mathcal{H}}_2\}$ is now given by:

$$\begin{aligned}\hat{\Delta} &= (1-\rho)(\epsilon_I+\epsilon_{II})(1-\gamma_H)-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H}{\phi}-\epsilon_{II}(1-\epsilon_I)c\frac{\gamma_L}{\phi}-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}\\ &\quad-(1-\rho)(1-\epsilon_{II})(1-\epsilon_I)c\frac{\gamma_L}{\phi}+\epsilon_{II}^2c\frac{\gamma_H}{\phi}+(1-\rho)\epsilon_{II}c\gamma_H+\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}+(1-\rho)\epsilon_I(1-\epsilon_I)c\frac{\gamma_L}{\phi}\\ &= (1-\rho)(\epsilon_I+\epsilon_{II})(1-\gamma_H)-\epsilon_I\epsilon_{II}c\frac{\gamma_H-\gamma_L}{\phi}-\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)(1-\epsilon_I)c\gamma_L+(1-\rho)\epsilon_{II}c\gamma_H,\end{aligned}$$

where I used that $\mathcal{W}(\boldsymbol{f};\hat{\mathcal{H}}_2)$ must be evaluated at $\boldsymbol{f}=\left(0,\frac{\gamma_L}{\phi},\frac{\gamma_H}{\phi}\right)$. The inequality $\gamma_L<\epsilon_{II}(\gamma_H-\gamma_L)$ implies that:

$$\hat{\Delta}'(\rho)=-(\epsilon_I+\epsilon_{II})(1-\gamma_H)+(1-\epsilon_I)c\gamma_L-\epsilon_{II}c\gamma_H<-(\epsilon_I+\epsilon_{II})(1-\gamma_H)+(1-\epsilon_I)c\gamma_L-(1+\epsilon_{II})c\gamma_L<0$$

and

$$\begin{aligned}\hat{\Delta}(\overline{\rho}) &= (1-\overline{\rho})(\epsilon_I+\epsilon_{II})(1-\gamma_H+c\gamma_L)+(1-\overline{\rho})\epsilon_{II}c(\gamma_H-\gamma_L)-(1-\overline{\rho})c\gamma_L-\epsilon_I\epsilon_{II}c\frac{\gamma_H-\gamma_L}{\phi}-\epsilon_{II}c\frac{\gamma_L}{\phi}\\ &= (1-\overline{\rho})\epsilon_{II}c(\gamma_H-\gamma_L)-(1-\overline{\rho})c\gamma_L+\epsilon_{II}^2c\frac{\gamma_H-\gamma_L}{\phi}-\epsilon_{II}c\frac{\gamma_L}{\phi}>0.\end{aligned}$$

Since $\hat{\Delta}(1)<0$, there exists a $\hat{\rho}\in(\overline{\rho},1)$ such that $\hat{\Delta}(\hat{\rho})=0$, that is, outcome $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ results in higher (lower) welfare than outcome $\{\hat{\mathcal{H}}_2\}$ if $\rho<\hat{\rho}$ $(\rho>\hat{\rho})$.

For $\rho\leq\overline{\rho}$ I thus have to make the exact same comparison as in the proof of Proposition 2 and I hence conclude that outcome $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ results in higher (lower) welfare than the high uniform punishment if $\rho>\check{\rho}$ $(\rho<\check{\rho})$.

For $\rho\in(\overline{\rho},\hat{\rho}]$ I have to compare $\mathcal{W}(\mathcal{H}_1,\hat{\mathcal{H}}_2)$ with $\mathcal{W}^0=2W\left(\frac{\gamma_L}{\phi}\right)$:

$$\begin{aligned}\Delta_{12} &= (1-\rho)(1+\epsilon_{II})(1-\gamma_H)-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H}{\phi}-\epsilon_{II}(1-\epsilon_I)c\frac{\gamma_L}{\phi}\\ &\quad-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)(1-\epsilon_{II})(1-\epsilon_I)c\frac{\gamma_L}{\phi}+2\epsilon_{II}c\frac{\gamma_L}{\phi}+2(1-\rho)c\gamma_L\\ &= (1-\rho)(1+\epsilon_{II})(1-\gamma_H+c\gamma_L)-\epsilon_{II}(\epsilon_I+\epsilon_{II})c\frac{\gamma_H-\gamma_L}{\phi}.\end{aligned}$$

Observe that $\Delta'_{12}(\rho)<0$ and that $\Delta_{12}(\overline{\rho})=(1-\epsilon_I)\epsilon_{II}c\frac{\gamma_H-\gamma_L}{\phi}>0$. So, outcome $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ yields higher welfare than the low uniform punishment for $\rho>\overline{\rho}$ sufficiently close to $\overline{\rho}$.

For $\rho>\hat{\rho}$ I have to compare $\mathcal{W}(\hat{\mathcal{H}}_2)$ $\mathcal{W}^0=2W\left(\frac{\gamma_L}{\phi}\right)$:

$$\begin{aligned}\Delta_2 &= (1-\rho)(1-\epsilon_I)(1-\gamma_H)-\epsilon_{II}^2c\frac{\gamma_H}{\phi}-(1-\rho)\epsilon_{II}c\gamma_H-\rho(1-\epsilon_{II})\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)\epsilon_I(1-\epsilon_I)c\frac{\gamma_L}{\phi}+2\epsilon_{II}c\frac{\gamma_L}{\phi}+2(1-\rho)c\gamma_L\\ &= (1-\rho)(1-\epsilon_I)(1-\gamma_H+c\gamma_L)-\epsilon_{II}^2c\frac{\gamma_H-\gamma_L}{\phi}+\epsilon_{II}c\frac{\gamma_L}{\phi}-(1-\rho)\epsilon_{II}c(\gamma_H-\gamma_L)+(1-\rho)c\gamma_L.\end{aligned}$$

Condition 2 implies that:

$$\begin{aligned}\Delta'_2(\rho) &= -(1-\epsilon_I)(1-\gamma_H+c\gamma_L)+\epsilon_{II}c(\gamma_H-\gamma_L)-c\gamma_L<-(1-\epsilon_I)\frac{1+\phi}{1-\epsilon_{II}}\epsilon_{II}c\frac{\gamma_H}{\phi}+\epsilon_{II}c\gamma_H-(2-\epsilon_I+\epsilon_{II})c\gamma_L\\ &= -\epsilon_{II}^2c\frac{\gamma_H}{\phi}-(1-\epsilon_I)\frac{1-\epsilon_I}{1-\epsilon_{II}}\epsilon_{II}c\frac{\gamma_H}{\phi}-(2-\epsilon_I+\epsilon_{II})c\gamma_L<0.\end{aligned}$$

The inequality $\gamma_L<\epsilon_{II}(\gamma_H-\gamma_L)$ ensures that $\Delta_2(1)<0$. So, outcome $\{\hat{\mathcal{H}}_2\}$ yields lower welfare than the low uniform punishment for $\rho>\hat{\rho}$ sufficiently close to 1.

Now consider $\rho\approx\hat{\rho}$. If $\Delta_{12}(\hat{\rho})>0$, then outcome $\{\mathcal{H}_1,\hat{\mathcal{H}}_2\}$ yields higher welfare than the low uniform punishment for all $\rho\in(\overline{\rho},\hat{\rho}]$. Since $\Delta_2(\hat{\rho})=\Delta_{12}(\hat{\rho})$, once I know the sign of $\Delta_{12}(\hat{\rho})$ I also know whether outcome $\{\hat{\mathcal{H}}_2\}$ is preferred to the low uniform punishment for $\rho\geq\hat{\rho}$ close to $\hat{\rho}$. In particular, if $\Delta_{12}(\hat{\rho})>0$, then there exists a $\widetilde{\rho}_r\in(\hat{\rho},1)$ such that for $\rho\in(\widetilde{\rho}_\ell,\widetilde{\rho}_r)$ total welfare is maximized by inducing outcome $\{\hat{\mathcal{H}}_2\}$, where $\widetilde{\rho}_\ell=\hat{\rho}$. By contrast, if $\Delta_{12}(\hat{\rho})<0$, then inducing outcome $\{\hat{\mathcal{H}}_2\}$ is never optimal. In that case $\widetilde{\rho}_\ell=\widetilde{\rho}_r=\check{\rho}$.

I now show that $\Delta_{12}(\hat{\rho})$ can have either sign. If $\gamma_L = \epsilon_{II}(\gamma_H - \gamma_L) \Leftrightarrow \frac{\gamma_H}{\gamma_L} = 1 + \frac{1}{\epsilon_{II}}$, then

$$\hat{\Delta} = (1-\rho)(\epsilon_I + \epsilon_{II})(1-\gamma_H + c\gamma_L) - (\epsilon_I + \epsilon_{II})c\frac{\gamma_L}{\phi},$$
$$\Delta_{12} = (1-\rho)(1+\epsilon_{II})(1-\gamma_H + c\gamma_L) - (\epsilon_I + \epsilon_{II})c\frac{\gamma_L}{\phi}.$$

I now have that $1-\hat{\rho} = \frac{1}{1-\gamma_H + c\gamma_L} c\frac{\gamma_L}{\phi}$ and hence $\Delta_{12}(\hat{\rho}) = (1-\epsilon_I)c\frac{\gamma_L}{\phi} > 0$. Since Δ_{12} depends continuously on all parameters, I conclude that $\Delta_{12}(\hat{\rho}) > 0$ if $\frac{\gamma_H}{\gamma_L}$ is sufficiently close to $1 + \frac{1}{\epsilon_{II}}$. On the other hand, for $\gamma_L \to 0$ I obtain

$$\hat{\Delta} = (1-\rho)(\epsilon_I + \epsilon_{II})(1-\gamma_H) - \epsilon_I\epsilon_{II}c\frac{\gamma_H}{\phi} + (1-\rho)\epsilon_{II}c\gamma_H,$$
$$\Delta_{12} = (1-\rho)(1+\epsilon_{II})(1-\gamma_H) - \epsilon_{II}(\epsilon_I + \epsilon_{II})c\frac{\gamma_H}{\phi}.$$

In this case, I have that

$$1-\hat{\rho} = \frac{\epsilon_I\epsilon_{II}c\frac{\gamma_H}{\phi}}{(\epsilon_I + \epsilon_{II})(1-\gamma_H) + \epsilon_{II}c\gamma_H}.$$

Evaluating Δ_{12} at $\hat{\rho}$ results in

$$\Delta_{12}(\hat{\rho}) = \frac{(1+\epsilon_{II})(1-\gamma_H)}{(\epsilon_I + \epsilon_{II})(1-\gamma_H) + \epsilon_{II}c\gamma_H}\epsilon_I\epsilon_{II}c\frac{\gamma_H}{\phi} - \epsilon_{II}(\epsilon_I + \epsilon_{II})c\frac{\gamma_H}{\phi}.$$

This expression is negative for ϵ_I sufficiently close to 0.

Acknowledgments

Financial support from the Belgian Federal Government through the Interuniversity Attraction Poles Project (contract 6/09) is gratefully acknowledged. This article owes its existence to numerous discussions with Wouter Vergote. The author would like to thank Remco van Eijkel, Marco Haan, attendees at PET10, SMYE 2011, EEA-ESEM 2011, PEJ 2014, workshop participants at CEREC (Facultés universitaires Saint-Louis), CORE (Université Catholique de Louvain), and IEEF as well as two anonymous referees and the editor for useful comments.

References

Agrawal, Arun. 2003. Sustainable governance of common-pool resources: context, methods, and politics. *Annual Review of Anthropology* 32:243–62.

Chu, C. Y. Cyrus, Sheng-cheng Hu, and Ting-yuan Huang. 2000. Punishing repeat offenders more severely. *International Review of Law and Economics* 20:127–40.

Dana, David A. 2001. Rethinking the puzzle of escalating penalties for repeat offenders. *Yale Law Journal* 110:733–83.

Ellickson, Robert C. 1991. *Order without law: how neighbors settle disputes*. Cambridge, MA: Harvard University Press.

Emons, Winand. 2007. Escalating penalties for repeat offenders. *International Review of Law and Economics* 27:170–8.

Endres, Alfred, and Bianca Rundshagen. 2012. Escalating penalties: a supergame approach. *Economics of Governance* 13:29–49.

Funk, Patricia. 2004. On the effective use of stigma as a crime-deterrent. *European Economic Review* 48:715–28.

Harrington, Winston. 1988. Enforcement leverage when penalties are restricted. *Journal of Public Economics* 37:29–53.

Landsberger, Michael, and Isaac Meilijson. 1982. Incentive generating state dependent penalty system. *Journal of Public Economics* 19:333–52.

Miceli, Thomas J. 2013. Escalating penalties for repeat offenders: why are they so hard to explain? *Journal of Institutional and Theoretical Economics* 169:587–604.

Miceli, Thomas J., and Catherine Bucci. 2005. A simple theory of increasing penalties for repeat offenders. *Review of Law and Economics* 1:71–80.

Miles, Stan, and Derek Pyne. 2015. Deterring repeat offenders with escalating penalty schedules: a Bayesian approach. *Economics of Governance* 16:229–50.

Mungan, Murat C. 2010. Repeat offenders: if they learn, we punish them more severely. *International Review of Law and Economics* 30:173–7.

Mungan, Murat C. 2013. Optimal warning strategies: punishment ought not to be inflicted where the penal provision is not properly conveyed. *Review of Law and Economics* 9:303–39.

Mungan, Murat C. 2014. A behavioral justification for escalating punishment schemes. *International Review of Law and Economics* 37:189–97.

Nyborg, Karine, and Kjetil Telle. 2004. The role of warnings in regulation: keeping control with less punishment. *Journal of Public Economics* 88:2801–16.

Ostrom, Elinor. 1990. *Governing the commons: the evolution of institutions for collective action*. Cambridge, UK: Cambridge University Press.

Ostrom, Elinor. 2000. Collective action and the evolution of social norms. *Journal of Economic Perspectives* 14:137–58.

Polinsky, A. Mitchell, and Daniel L. Rubinfeld. 1991. A model of optimal fines for repeat offenders. *Journal of Public Economics* 46:291–306.

Polinsky, A. Mitchell, and Steven Shavell. 1998. On offense history and the theory of deterrence. *International Review of Law and Economics* 18:305–24.

Rousseau, Sandra. 2009. The use of warnings in the presence of errors. *International Review of Law and Economics* 29: 191–201.

Ariel Rubinstein. 1979. An optimal policy for offenses that may have been committed by accident. In *Applied game theory*, edited by S. J. Brams, A. Schotter, and G. Schwödiauer. Würzburg: Physica-Verlag. pp. 406–413.

Stigler, George J. 1970. The optimum enforcement of laws. *Journal of Political Economy* 78:526–36.

van der Made, Allard. 2015. *Graduated Punishments in Public Good Games*. University of Groningen: Unpublished paper.

Wade, Robert. 1994. *Village republics: economic conditions for collective action in South India*. Oakland: ICS Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article.