Graduated Punishments in Public Good Games

Allard van der Made *

May, 2011

Abstract

A host of social situations feature graduated punishments. We explain this phenomenon 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. A low punishment today imperfectly sorts agents by type: only low-cost agents contribute. The planner uses this information optimally, punishing tomorrow's (alleged) repeat shirkers harsher than first-time shirkers. The threat of becoming branded as repeat offender allows the planner to use a very mild punishment for first-time shirkers, attenuating the costs associated with administering punishments. Graduated punishments are consequently socially optimal as long as the population is not too homogeneous.

Keywords: graduated punishments, imperfect monitoring, collective action, reputation.

JEL classification codes: D82, H41, K49

---

*CEREC, Facultés universitaires Saint-Louis, 43 Boulevard du Jardin Botanique, 1000 Brussels, Belgium, e-mail: `vandermade@fusl.ac.be.nl`. Financial support from the Belgian Federal Government through the IAP Project (contract 6/09) is gratefully acknowledged. This paper owes its existence to numerous discussions with Wouter Vergote. I would like to thank Remco van Eijkel, attendees at PET10, SMYE2011, and workshop participants at CEREC and at CORE (Université Catholique de Louvain).

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 are inclined to choose the collectively preferred action is the presence of a monitoring institution that is able to punish (alleged) wrongdoers. Many scholars studying collective action problems have observed that successful punishment schemes often exhibit graduated sanctions, in which repeat offenders are punished more severely than first-time offenders (e.g. Agrawal, 2003, Ellickson, 1991, Ostrom, 1990, 2000, Wade, 1994). Graduated sanctions also appear in many judiciary systems, stipulating that habitual offenders can or must be punished more severely than first-time offenders.[1] In its most extreme form graduated punishments are such that first-time offenders receive a mere warning. Given its widespread use, it is surprising that this phenomenon has received limited theoretical attention.

We present a theory that explains the prevalence of punishment schemes with graduated punishments. We 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.

More precisely, in our 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 social planner monitors the behaviour of individual agents, but monitoring is imperfect: some non-contributors (shirkers) escape being detected and some contributors are found guilty of something they did not do. The planner can administer punishments to alleged shirkers, but this entails a social cost.[2] The individually borne cost of contributing to the public good differs among agents and is either high or low. An agent's cost type is private information. The planner maximizes total welfare, i.e the total social benefits of 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 such a punishment is only optimal if the number of high-cost agents is sufficiently large. If this number is not sufficiently large, then the social costs of (erroneously) administering severe punishments to a large group of low-cost agents outweigh the benefits of deterring a small group of high-cost agents from shirking. The planner then sets a low punishment and only low-cost agents contribute to the public good. If agents are not only supposed to contribute

[1]For example, various state governments in the United States have enacted so called 'Three Strikes Laws'. Such laws require the 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. See also http://en.wikipedia.org/wiki/Three_strikes_law.

[2]These costs include the administrative and legal costs associated with punishing someone. They could also include the cost of keeping someone in jail for some time.

today, but also in future periods, then the planner can often improve upon this outcome by employing graduated punishments.

Using graduated punishments instead of a uniform punishment improves welfare for two reasons. Firstly, by imposing a mild sanction today the planner is able to (imperfectly) 'sort' agents by cost type: only low-cost agents contribute if punishments are low. As a consequence, the planner can in the future (again imperfectly) tailor punishments to types by imposing a harsh punishment on repeat offenders and a moderate one on first-time offenders. Such tailoring enables the planner to induce a given amount of contributions in a more cost-efficient way.

Secondly, the mere threat of becoming 'branded' as shirker and thereby moving from the low punishment regime to the high punishment regime makes an agent reluctant to shirk today: since monitoring is imperfect, being caught shirking today increases expected future punishments, even if the agent plans to contribute in future periods. In other words, an agent fears losing his reputation of being a contributor. This fear enables the planner to reduce the punishment for first-time shirkers below the low punishment of the one-shot setting (i.e. below the punishment that prevails if the number of high-cost agents is small). The reputation effect is particularly strong if agents are patient, i.e. if agents have a long horizon. In fact, for all cost parameters one can find a discount rate above which it suffices to issue a mere warning to first-time offenders.

Using graduated punishments is not always optimal. If the society consists mainly of high-cost agents, then using this punishment scheme would yield a very low level of public good provision. To increase contributions, the planner then opts for a uniform punishment that deters all agents from shirking. Nonetheless, because monitoring is imperfect, some agents are still punished on the equilibrium path.[3] On the other hand, if the vast majority of agents incur the low cost when contributing, then most agents who end up in the high punishment regime are low-cost agents. It can then be optimal to use the (uniform) low punishment of the one-shot setting, as this leads to considerably lower punishment costs without significantly reducing the level of aggregate contributions.

Our main results hinge crucially on the presence of type II errors, i.e. the possibility that the planner falsely judges someone guilty of shirking. If type II errors were completely absent, then only shirkers are punished. The presence of type II errors has two effects. Firstly, if the planner would never erroneously punish contributors, then agents would not mind losing their reputation and the planner would consequently be unable to reduce the punishment to first-time shirkers below the low punishment of the one-shot setting. The reason that only type II errors matter in the determination of the reputation effect is that all agents contribute in each future period as soon as they move to the high punishment regime. So, only type II errors lead to punishments being administered to agents who are branded as shirkers. Secondly, absent type II errors the planner always ensures that high-cost agents contribute in the one-shot setting. The only advantage of setting a punishment that does not suffice to deter high-cost agents from shirking is that such a low punishment

[3]This is a common feature of equilibria of games with private information and imperfect monitoring. See e.g. Green and Porter (1984).

entails low social costs of administering punishments to low-cost agents. Yet, since low-cost agents are only punished if a type II error occurs, this advantage does not play a role if type II errors are never made.[4]

If the planner knew each agent's cost of contributing to the public good, then sorting agents by type would be redundant and the planner would never resort to graduated punishments. She would then be able to perfectly deter shirking: it suffices to 'promise' a given agent an expected punishment at least as large as this agent's cost of contributing.[5]

Note that our framework not only applies to classic public good situations, but also to law enforcement problems. The cost of contributing to the public good becomes the opportunity cost of not committing some crime in the latter category. Furthermore, 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 stemming from cybercrimes. Not engaging in criminal activities thence increases welfare at the aggregate level in a similar fashion as contributing to a public good does.

Most situations of collective action are plagued by informational problems. Consider for instance a groundwater basin shared by hundreds of farmers. Such basins can be destroyed by overextraction.[6] Whether a particular farmer extracts more water than he is entitled to is difficult to determine: a sudden drop in the water level could equally well be caused by overextraction by one of his neighbours. So, both type I and type II errors are bound to occur. How 'tempted' a farmer is to overextract water depends on unobservable psychological traits as well as the finer details of the microclimate and the soil composition he faces. His cost type is consequently private information.

This paper is organized as follows. Section 2 introduces the main ingredients of the model and presents the optimal punishment scheme of the one-shot setting. In Section 3 we study a two-period version of our model. The full-fledged dynamic model can be found in Section 4. In Section 5 we consider some extensions and we relate our work to the literature. Section 6 offers concluding remarks. All proofs are relegated to various appendices.

2 The Environment

A social planner faces a public good problem. If a fraction π of the population contributes to the public good, then the total social benefits of the public good amount to $\pi \times v$. However, contributing is costly for individuals. A fraction $1-\rho$ of the population consists of agents who incur the high cost of γv when contributing, where $\gamma < 1$. The remaining fraction ρ consists of agents who incur the low cost of $(\gamma - \alpha)v$ when contributing, for some

[4]Type I errors reduce the probability that shirking is detected. Making a type I error with some probability p has the same impact on the optimal punishment(s) as only monitoring a sample containing a fraction $1-p$ of the population: a larger p leads to higher punishments, but the expected punishments remain constant.

[5]Since monitoring is imperfect, expected punishments are not equal to actual punishments.

[6]Ostrom (1990), chapter 4, provides analytical narratives of the collective action problems surrounding the groundwater basins near Los Angeles.

$\alpha \in (0, \gamma)$. Importantly, an agent's type (low-cost or high-cost) is private information. All agents are risk-neutral. We normalize v to 1. We use the subscript L to refer to low-cost agents and the subscript H to refer to high-cost agents.

Without any further rules (with associated sanctions) imposed by the social planner no agent would contribute. Yet, because $\gamma < 1$, it is socially efficient for all agents to contribute. The planner can monitor agents' behaviour, enabling her to punish alleged non-contributors. However, the planner's monitoring technology is flawed: with probability ϵ the planner draws the wrong conclusion when investigating a given agent's behaviour, where $\epsilon \in (0, \frac{1}{2})$. So, only a fraction $1 - \epsilon$ of the non-contributors are caught, whereas a fraction ϵ of the contributors are found guilty of something they did not do.[7] We assume that monitoring agents comes at no cost, but that administering punishments to agents is costly. Specifically, if the planner administers a punishment to an agent that reduces this agent's utility by f, then society bears a cost of βf, where $\beta < 1$. One can think of βf as the administrative and legal costs associated with fining an agent or the cost of keeping someone in jail for some time.

The planner tries to maximize total welfare by choosing the level of the punishments. We assume that she can commit to this choice. Punishments are made public before agents advance to the contribution stage. Total welfare consists of the social benefits of the public good, the individually borne costs of contributing to the public good, and the costs of administering punishments.[8] In a one shot-setting total welfare $W = W(f_0)$ therefore reads

$$W = \rho(1 - \gamma + \alpha)\delta_L + (1 - \rho)(1 - \gamma)\delta_H - \Phi(\delta_L, \delta_H)\beta f_0, \qquad (1)$$

where $\delta_L = 1$ ($\delta_L = 0$) if low-cost agents (do not) contribute, $\delta_H = 1$ ($\delta_H = 0$) if high-cost agents (do not) contribute, and $\Phi(\delta_L, \delta_H)$ is the fraction of the population that is punished. Of course, the agents' actions δ_L and δ_H depend on the level of the punishment f_0.

The planner can ensure that each agent contributes to the public good by choosing sufficiently harsh punishments. However, since she would still punish a fraction ϵ of the agents should everybody contribute, this outcome need not be optimal. We do assume, however, that the planner prefers the 'everybody-contributes' outcome to the laissez-faire outcome in which there is no public good provision at all. To ensure that the planner never opts for laissez-faire, we maintain the following condition throughout the paper:

Condition 1 *Laissez-faire is never optimal:* $1 - \gamma > \frac{1-\phi}{2\phi}\beta\gamma$,

where $\phi := 1 - 2\epsilon > 0$ signifies the quality of the planner's monitoring technology. In Appendix A we show that Condition 1 suffices to ensure that the planner always prefers harsh punishments to laissez-faire, even if all agents would incur the high cost when contributing ($\rho = 0$). Condition 1 states that the gain in total welfare associated with one high-cost

[7]Note that the probability that the planner makes a type I error is equal to the probability that she makes a type II error. In section 5 we relax this assumption.

[8]From a normative point of view, it seems natural not to include the actual punishments in the welfare measure. In Section 5 we consider an extension of the model in which punishing agents who did in fact contribute is detrimental to welfare.

agent contributing (i.e. $1-\gamma$) exceeds the expected cost society bears when the planner chooses a punishment level that is such that high-cost agents do indeed contribute. This expected cost equals the marginal cost of administering punishments (β) times the probability that a high-cost agent who contributes is erroneously found guilty of shirking ($\frac{1-\phi}{2}$) times the minimal punishment that is required to deter high-cost agents from shirking ($\frac{\gamma}{\phi}$).

Before we move to settings with more than one period, we derive the optimal strategy of the planner in the one-shot setting. This static equilibrium outcome serves as a benchmark for the two-period setting and the infinite-horizon setting that we discuss in the next sections: in both settings the planner can always simply replicate the static outcome in each period. However, in doing so the planner discards any information regarding agents' types that can be inferred from past actions. So, our analysis of the one-shot setting provides a lower bound on the per-period welfare that can be attained in the other settings.

An agent contributes to the public good if and only if the associated expected costs do not exceed the expected costs the agent faces when shirking. A low-cost agent consequently contributes as long as $\gamma-\alpha+\epsilon f_0 \leq (1-\epsilon)f_0$, i.e. as long as $f_0 \geq \frac{\gamma-\alpha}{\phi}$. A high-cost agent contributes if and only if $\gamma+\epsilon f_0 \leq (1-\epsilon)f_0$. This inequality reduces to $f_0 \geq \frac{\gamma}{\phi}$. So, the planner must choose between a low punishment ($\phi f_0 = \gamma-\alpha$) which only induces low-cost agents to contribute and a high punishment ($\phi f_0 = \gamma$) which ensures that high-cost agents also contribute. Of course, the latter option becomes more attractive as ρ decreases. This observation is formalized as follows:

Proposition 1 *In the one-shot setting the social planner commits to*

$$\phi f_0^* = \begin{cases} \gamma & \text{if } \rho \leq \bar{\rho} \\ \gamma-\alpha & \text{if } \rho > \bar{\rho}, \end{cases} \tag{2}$$

where

$$\bar{\rho} = 1 - \frac{\frac{1-\phi}{2\phi}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)} \in (0,1). \tag{3}$$

Proof. See Appendix A. ■

Even though the cost of contributing to the public good incurred by a high-cost agent is less than the social welfare generated by this contribution, the planner does not always opt to incentivize the high-cost agents to contribute. The reason is that administering punishments is socially costly. If the number of high-cost agents is small (ρ large), the increase in the provision of the public good brought about by moving from the low punishment ($\phi f_0 = \gamma-\alpha$) to the high punishment ($\phi f_0 = \gamma$) is small. Such a move would also entail administering the high punishment instead of the low one to a fraction ϵ of the low-cost agents. If the population consists mainly of low-cost agents the detrimental effect on welfare of the higher punishments dominates the positive effect of more contributions. Note that the planner is particularly inclined to simply ignore high-cost agents (i.e. $\bar{\rho}$ is relatively small) if there is a large cost difference between low-cost and high-cost agents

(α large or γ small), if the cost of administering punishments is large (β large), or if the planner makes a lot of mistakes (ϕ small).

The trade-off between higher contributions and lower social costs also plays a prominent role in settings with multiple periods. Yet, it turns out that the planner can often use information regarding agents' types obtained in past periods to alleviate the social costs of punishments.

3 The Two-Period Model

In this section we consider a setting in which agents are supposed to contribute to the public good in two periods. An agent's type is again private information. However, the social planner does remember in the second period whether or not she has punished a given agent in the first period. Just like in the one-shot setting the planner draws the wrong conclusion with probability ϵ when investigating an agent's behaviour. Drawing the wrong conclusion with respect to a particular agent's behaviour in period 1 does not affect the probability with which she misjudges that agent's behaviour in period 2.

Recalling who has been punished in the first period enables the planner to use differentiated punishments in the second period, one for agents who have not been punished in the first period (f_2) and one for agents who *have* been punished ($\hat{f}_2$). Of course, in the first period the planner can only use one punishment level (f_1). We assume that the planner announces all punishments at the start of the game and that she is able to commit to these levels. Each agent employs backward induction to arrive at his optimal strategy. The precise timing of the game is thus as follows:

0. The monitor announces f_1, f_2, and $\hat{f}_2$.

1a. Each agent decides whether or not to contribute (δ_L and δ_H are chosen).

1b. The monitor carries out investigations and administers punishments.

2a. Each agent decides whether or not to contribute.

2b. The monitor carries out investigations and administers punishments.

Payoffs are realized at the end of each period. Each agent minimizes his total expected costs. The planner maximizes aggregate welfare $\mathcal{W}$, the sum of total welfare in period 1 (W_1) and total welfare in period 2 (W_2).

Using two different punishments in period 2 need not be optimal. It might be optimal to either incentivize all agents to contribute in both periods or to simply ignore high-cost agents altogether. Obviously, if the planner decides to use only one punishment level, then society is best off if she uses the one described in Proposition 1. Potential improvements on this strategy are the subject of the next subsection.

3.1 Analysis

Clearly, using two different punishments in the second period can only be optimal if low-cost agents and high-cost agents behaved differently in the first period. The reason is that the planner cannot distill any information regarding a given agent's type from his behaviour in period 1 should the two types employ the same strategy in that period. So, using differentiated punishments requires agents choosing $\delta_L = 1$ and $\delta_H = 0$.[9]

Notice that, since the game ends after the second period, strategic motives do not play a role in stage 2a: agents simply compare the expected costs associated with the two possible choices and pick the one with the lowest expected costs. Therefore, if the planner contemplates using two different punishments in period 2, she can confine attention to $f_2^* = \frac{\gamma-\alpha}{\phi}$ (for those who have not been punished in the first period) and $\hat{f}_2^* = \frac{\gamma}{\phi}$ (for those who have been punished in the first period). This pair of punishments induces high-cost agents who were punished in the first period to contribute in period 2. Moreover, each low-cost agent, whether or not punished in period 1, decides to contribute in period 2 when faced with the pair $(f_2^*, \hat{f}_2^*)$. Lastly, high-cost agents who dodged being punished in period 1 shirk again in period 2, as the expected punishment ϕf_2^* is less than their cost of contributing γ.[10]

The planner induces the choices $\delta_L = 1$ and $\delta_H = 0$ if she picks a moderate punishment level f_1 in period 1: it must be such that only low-cost agents opt to contribute in the first period. Formally, the following incentive compatibility constraints must hold:

- Low-cost agents prefer to contribute in period 1 if and only if:

  $$\gamma-\alpha+\epsilon(f_1+\gamma-\alpha+\epsilon\hat{f}_2^*)+(1-\epsilon)(\gamma-\alpha+\epsilon f_2^*) \leq (1-\epsilon)(f_1+\gamma-\alpha+\epsilon\hat{f}_2^*)+\epsilon(\gamma-\alpha+\epsilon f_2^*).$$

  The left-hand side of this constraint consists of the expected costs a low-cost agent faces when contributing in period 1. It equals the actual cost of contributing $\gamma-\alpha$ plus expected costs associated with being erroneously punished in period 1 and/or period 2. The right-hand side consists of the expected costs a low-cost agent faces when shirking in period 1. Note that we have used the fact that low-cost agents always contribute in period 2 when faced with the pair $(f_2^*, \hat{f}_2^*)$. Using the expressions for f_2^* and $\hat{f}_2^*$ and the definition of ϕ reduces the constraint to

  $$\phi f_1 \geq \gamma - \alpha - \tfrac{1-\phi}{2}\alpha. \quad (4)$$

- High-cost agents prefer to shirk in period 1 if and only if:

  $$\gamma + \epsilon(f_1 + \gamma + \epsilon\hat{f}_2^*) + (1-\epsilon)(1-\epsilon)f_2^* > (1-\epsilon)(f_1 + \gamma + \epsilon\hat{f}_2^*) + \epsilon(1-\epsilon)f_2^*.$$

[9] Because low-cost agents are more inclined to contribute than high-cost agent are (given any first period punishment f_1), we can immediately discard the possibility that only high-cost agents contribute in period 1.

[10] Formally, expressions like ϕf_2^* denote a *difference* in expected punishment: $\phi f_2^* = (1-\epsilon)f_2^* - \epsilon f_2^*$ is the expected punishment faced when shirking minus the expected punishment faced when contributing. We frequently omit the "difference in" for ease of exposition.

This constraint is derived in a similar fashion as the one pertaining to low-cost agents. Observe that a high-cost agent only contributes in the second period if he has been punished in the first period. The constraint for high-cost agents is equivalent to

$$\phi f_1 < \gamma - \tfrac{1+\phi}{2}\alpha. \tag{5}$$

It is not difficult to see that one can always find an f_1 such that the two incentive compatibility constraints (4)-(5) are met simultaneously. In fact, if the planner opts for differentiated punishments in period 2, then she sets

$$\phi f_1^* = \max\{\gamma - \alpha - \tfrac{1-\phi}{2}\alpha, 0\}.$$

It remains to determine when differentiated punishments are optimal. The *menu of punishments* $\boldsymbol{f}^* = (f_1^*, f_2^*, \hat{f}_2^*)$ is optimal if the associated aggregate welfare $\mathcal{W}(\boldsymbol{f}^*)$ exceeds the aggregate welfare $\mathcal{W}(f_0^*)$ generated when the planner uses the single punishment given in (2) in both periods. In Appendix A we compare these two welfare expressions. This comparison leads to

Proposition 2 *There exist $\check{\rho} = \check{\rho}(\alpha, \beta, \gamma, \phi) < \bar{\rho}$ and $\hat{\rho}(\alpha, \beta, \gamma, \phi) > \bar{\rho}$ such that the social planner maximizes aggregate welfare $\mathcal{W}$ by committing to the menu of punishments*

$$f_1^* = \max\{\tfrac{\gamma-\alpha}{\phi} - \tfrac{1-\phi}{2\phi}\alpha, 0\}, \quad f_2^* = \tfrac{\gamma-\alpha}{\phi}, \quad \hat{f}_2^* = \tfrac{\gamma}{\phi} \tag{6}$$

as long as $\rho \in [\check{\rho}, \hat{\rho}]$. If $\rho < \check{\rho}$ or $\rho > \hat{\rho}$, then the social planner commits in both periods to the single punishment given in (2). Furthermore, $\hat{\rho} < 1$ if and only if $\gamma - \alpha < \frac{1-\phi}{2}\alpha$, whereas $\check{\rho} > 0$ for all parameter configurations.

Proof. See Appendix A. ■

If $\boldsymbol{f}^*$ used, then a low-cost agent always contribute, whereas a high-cost agent shirks in period 1 and contributes in period 2 only if he is punished in period 1.

If the population is (all but) homogeneous (ρ close to 0 or 1), then the planner opts for the single punishment given in (2) instead of graduated punishments.[11] Of course, when faced with a homogeneous population, the best the planner can do is to commit in both periods to the smallest punishment that deters each agent from shirking.

The positive effects associated with using graduated punishments (i.e. using $\boldsymbol{f}^*$) instead of a single punishment start playing a role as ρ departs from 0 or 1. If the population is heterogeneous, then using graduated punishments allows the planner to (imperfectly) sort agents by type. The reason is that in period 1 the punishment level f_1^* is too low to incentivize high-cost agents to contribute and as a result only low-cost agents contribute. This implies that an agent who is caught shirking in period 2 for the second time is likely to be a high-type. Because the planner does occasionally draw the wrong conclusion when investigating agents' behaviour, this mechanism only imperfectly sorts agents by type.

[11] If $\hat{\rho} = 1$, then the difference in aggregate welfare $\mathcal{W}(\boldsymbol{f}^*) - \mathcal{W}(f_0^*)$ vanishes as ρ approaches 1.

This sorting enables the planner to tailor second-period punishments to types to a large extent. Since the vast majority of those who have been found guilty of shirking in period 1 are high-cost agents and high-cost agents can only be deterred from shirking by 'promising' them an expected punishment of at least γ, the planner announces a punishment of $\frac{\gamma}{\phi}$ for repeat offenders. On the other hand, an agent who is found guilty of shirking for the first time in period 2 is probably not a high-cost agent. A punishment of $\frac{\gamma-\alpha}{\phi}$ therefore suffices to deter most of the agents who were not punished in period 1 from shirking in period 2.

Recall that the choice between $\phi f_0 = \gamma - \alpha$ and $\phi f_0 = \gamma$ in the one-shot setting is a 'choice between two evils': choosing $\phi f_0 = \gamma - \alpha$ leads to a suboptimal contribution level (only a fraction ρ of the population contributes), whereas choosing $\phi f_0 = \gamma$ leads to large costs of administering punishments. However, if the planner can (imperfectly) tailor punishments to types, then this choice can be avoided. This has two consequences. Firstly, with graduated punishments only high-cost agents who escaped being punished in period 1 shirk in period 2 and the contribution level in period 2 therefore exceeds ρ. Secondly, by administering the severe punishment only to repeat offenders, the planner moderates the aggregate costs of administering punishments.

If the population consists mainly of high-cost agents (ρ small), then using graduated punishments would result in a very low level of public good provision in the first period. At the same time a considerable part of the population would be punished in that period (recall that a fraction $\frac{1+\phi}{2}$ of the high-cost agents ends up being punished if $\boldsymbol{f}^*$ is used). These costs associated with using graduated punishments become smaller as ρ increases. This explains why in general $\check{\rho}$ departs considerably from 0 whereas $\hat{\rho}$ often equals 1.

Observe that f_1^* is less than $\frac{\gamma-\alpha}{\phi}$, the smallest punishment that deters low-cost agents from shirking in the one-shot setting. The reason that the planner is able to incentivize low-cost agents to contribute in period 1 with an expected punishment below $\gamma - \alpha$ is that an agent found guilty of shirking in that period receives part of his 'effective punishment' indirectly. Such an agent not only faces the direct punishment f_1^*, but he will also receive the high punishment $\hat{f}_2^*$ instead of the lower punishment f_2^* should he be found guilty of shirking a second time. So, the threat of becoming known as a repeat offender, i.e. the fear of losing one's reputation, allows the planner to reduce the expected punishment used in period 1 below the level that is required in a one-shot setting. Notice that the size of this *reputation effect* simply equals the loss in expected utility associated with losing one's reputation: $\frac{1-\phi}{2\phi}\alpha$ is the difference between $\hat{f}_2^*$ and f_2^* times the probability that a contributing agent is erroneously found guilty of shirking.[12]

Whether the planner opts for graduated punishments depends crucially on the fraction of low-cost agents ρ: using graduated punishments is only optimal if $\rho \in [\check{\rho}, \hat{\rho}]$. The following comparative statics results can be obtained:

[12] The size of the reputation effect is smaller if the corner solution $f_1^* = 0$ prevails. If $\phi f_1^* = \gamma - \alpha - \frac{1-\phi}{2}\alpha$ and $\rho = 1$, then the reduction in aggregate punishments in period 1 due to the reputation effect exactly equals the increase in aggregate punishments in period 2 caused by the fact that some agents receive the punishment $\hat{f}_2^*$ instead of f_2^*. The planner therefore becomes indifferent between using f_0^* and using $\boldsymbol{f}^*$ as $\rho \to 1$ if the size of the reputation effect is $\frac{1-\phi}{2\phi}\alpha$. If the reputation effect is smaller, then it cannot offset the detrimental effect $\hat{f}_2^*$ has on welfare and $\hat{\rho}$ is consequently less than 1.

Corollary 1 *An increase in ϕ leads to an increase in $\check{\rho}$, whereas an increase in α, β, or γ leads to a decrease in $\check{\rho}$.*

Proof. See Appendix A. ■

To understand these results, recall that if the planner would have a perfect monitoring technology at her disposal, then she would never use graduated punishments. In such an ideal world the planner would announce a single punishment that is sufficiently severe to deter both low-cost and high-cost agents from shirking. Since no mistakes would be made during the monitoring stages, aggregate welfare would then amount to $2(1-\gamma+\rho\alpha)$, the social benefits of the public good minus the costs of contributing if everybody contributes in both periods. Such 'perfect deterrence' is not possible if ϕ is less than one. In fact, as the monitoring technology deteriorates (ϕ decreases), the planner becomes increasingly keen on moderating its punishment level(s). Using graduated punishments instead of a single one does that. The threshold $\check{\rho}$ above which the planner opts for the menu of punishments $\boldsymbol{f}^*$ consequently increases in ϕ. A similar reasoning explains why $\check{\rho}$ decreases in both β and γ. Lastly, an increase in α decreases the social costs associated with administering the punishments f_1^* and f_2^*. Using the menu $\boldsymbol{f}^*$ instead of $f_0^* = \frac{\gamma}{\phi}$ consequently becomes more attractive as α increases. This explains why $\check{\rho}$ decreases in α.

The aggregate welfare that society enjoys when graduated punishments are used (as well as the optimality of this option) depends crucially on the extent of the reputation effect. In the two-period model an agent suffers from a bad reputation for at most one period. Moreover, agents' behaviour is bound to be prone to 'endgame effects': losing one's reputation in period 2 has no effect on expected payoffs. Because of these reasons it is not a priori clear whether our results carry over to settings with more than two periods. We therefore investigate in the next section the relation between the agents' horizon and the viability of graduated punishments using a full-fledged dynamic model.

4 The Dynamic Model

Time $t = 1, 2, 3, \ldots$ is discrete. Each period t 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 those agents who have been found guilty of shirking. The last stage is a renewal stage: a fraction $1-\zeta \in (0,1)$ of the population dies and is replaced by new agents. The probability that a certain agent dies does not depend on his type (low-cost or high-cost), his behaviour with respect to the public good, or the number of times he has been punished. So, each agent advances to the next period with probability ζ. The population is again characterized by the parameters α, γ, and ρ. In particular, a fraction ρ of each generation and thus of the population in any period consists of low-cost agents. Note that since $\gamma < 1$, it is socially optimal if an agent contributes in *each* period that he lives.

The social planner is immortal. Moreover, she keeps track of whether or not a given agent has been punished in the past.[13] The quality of her monitoring technology is again given by ϕ. She announces all punishment levels that might be applicable in period t at the start of that period. She can choose to use two different punishments, one for agents who have never been punished before (f_t) and one for agents who have been punished at least once ($\hat{f}_t$). Alternatively, she can administer the same punishment to all agents found guilty of shirking, irrespective of their punishment record. Since the fraction low-cost agents is always ρ, the planner chooses the punishment given in (2) in the latter case. After the punishment levels for period t have been announced each agent decides whether or not to contribute in that period.

Note that the population can be divided into four categories: low-cost agents who have never been punished, low-cost agents who have been punished at least once, high-cost agents who have never been punished, and high-cost agents who have been punished at least once. We focus on the *stationary equilibria* of the model, i.e. equilibria that 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. We thus focus on the very long run ($t \to \infty$).

In each period the planner aims to maximize the total welfare generated in that period. Observe that a strategy of the planner that supports a stationary equilibrium in the present setting would also support the corresponding stationary equilibrium of the game in which the planner instead maximizes current welfare *plus* discounted future welfare, irrespective of the discount rate. An agent tries to minimize his expected discounted current and future costs. The only difference between the planner and an agent regarding their attitude towards the future lies in the fact that the former is immortal whereas an agent only has a probability of ζ to live on for another period. It is therefore natural to use this survival rate as agents' discount factor between periods. Note that ζ determines an agent's expected horizon: the expected horizon equals $\frac{\zeta}{1-\zeta}$, which is strictly increasing in ζ.[14] Importantly, payoffs are realized after the punishment stage, but before the renewal stage.

A stationary equilibrium is supported by a pair of punishment levels f^* and $\hat{f}^*$ and four contribution rules: δ_L^*, $\hat{\delta}_L^*$, δ_H^*, and $\hat{\delta}_H^*$. Here, $\delta_j^* = 1$ ($\delta_j^* = 0$) if a type j-agent who has never been punished decides (not) to contribute, $j = L, H$. Likewise, $\hat{\delta}_j^* = 1$ ($\hat{\delta}_j^* = 0$) if a type j-agent who *has* been punished at least once decides (not) to contribute, $j = L, H$. Of course, the equilibrium contribution rules must be best responses to the equilibrium punishment levels and vice versa.

The strategy of the planner depends crucially on the precise composition of the population. We therefore first derive the composition of the population in the limit in Subsection 4.1 before we derive the stationary equilibria in Subsection 4.2. We skip any reference to taking limits in these subsections if there is no risk of confusion.

[13]In Section 5 we discuss the possibility that the planner has imperfect recall.

[14]An agent stays alive for exactly k periods after the current period with probability $\zeta^k(1-\zeta)$. His expected horizon thus equals $\sum_{k\in\mathbb{N}} k\zeta^k(1-\zeta) = \frac{\zeta}{(1-\zeta)^2}(1-\zeta)$.

4.1 The Composition of the Population

Whether or not a given agent has been punished in the past is of course irrelevant if the planner opts for the uniform punishment given in (2). Just like in the two-period model graduated punishments can only be optimal if these punishments are such that low-cost agents always contribute whereas high-cost agents only contribute if they have been punished in past periods at least once. We can thus confine attention to the case $(\delta_L, \hat{\delta}_L, \delta_H, \hat{\delta}_H) = (1, 1, 0, 1)$.

Let $\hat{q}$ (q) be the fraction of the population that has (never) been punished before. Denote the fraction of the population that consists of low-cost agents who are in q ($\hat{q}$) by μ ($\hat{\mu}$).[15] By construction $\hat{q} = 1 - q$ and $\hat{\mu} = \rho - \mu$. From the above definitions and the fact that a fraction $1 - \zeta$ of the old population is replaced by pristine agents at the end of each period one infers that in the stationary equilibrium q abides by the following 'flow equation':

$$\begin{aligned} q =& 1 - \zeta + \zeta q \times \left(\tfrac{1+\phi}{2}\delta_L \tfrac{\mu}{q} + \tfrac{1-\phi}{2}(1 - \delta_L)\tfrac{\mu}{q} + \tfrac{1+\phi}{2}\delta_H(1 - \tfrac{\mu}{q}) + \tfrac{1-\phi}{2}(1 - \delta_H)(1 - \tfrac{\mu}{q})\right) \\ =& 1 - \zeta + \zeta\phi\mu + \zeta\tfrac{1-\phi}{2}q. \end{aligned} \tag{7}$$

The right-hand side consists of the inflow of new agents (which equals $1 - \zeta$) and of the agents who stay in q because they have not been punished in the previous period (which equals $\frac{1+\phi}{2}$ times the fraction of contributors in q plus $\frac{1-\phi}{2}$ times the fraction of shirkers in q). The flow equation for μ reads

$$\mu = (1 - \zeta)\rho + \zeta q \times \left(\tfrac{1+\phi}{2}\delta_L \tfrac{\mu}{q} + \tfrac{1-\phi}{2}(1 - \delta_L)\tfrac{\mu}{q}\right) = (1 - \zeta)\rho + \zeta\tfrac{1+\phi}{2}\mu. \tag{8}$$

Straightforward algebra yields the following:

Lemma 1 *Suppose* $(\delta_L, \delta_H) = (1, 0)$. *Then in the very long run one has:*

$$q = \frac{(1 - \zeta)(1 - \zeta\frac{1+\phi}{2} + \rho\zeta\phi)}{(1 - \zeta\frac{1-\phi}{2})(1 - \zeta\frac{1+\phi}{2})}, \quad \mu = \frac{\rho(1 - \zeta)}{1 - \zeta\frac{1+\phi}{2}}. \tag{9}$$

Proof. See Appendix A. ■

Observe that

$$\frac{\mu}{q} = \rho \times \frac{1 - \zeta\frac{1-\phi}{2}}{1 - \zeta\frac{1-\phi}{2} - (1 - \rho)\zeta\phi} > \rho,$$

i.e. in q the fraction of low-cost agents exceeds ρ. This is intuitive: because low-cost agents in q do contribute, most of them stay in q. On the other hand, the majority of the high-cost agents, being found guilty of shirking, move to $\hat{q}$ and hence $1 - \frac{\hat{\mu}}{\hat{q}} > 1 - \rho$.

The above results regarding the composition of the population enable us to determine the stationary equilibria of the dynamic model. This is done in the next subsection.

[15] We allow ourselves some abuse of notation by using q for the fraction of the population that has never been punished before as well as for the *set* of agents with this feature. A similar remark applies to $\hat{q}$.

4.2 Results

An agent minimizes his expected discounted costs by choosing between contributing and shirking.[16] Denote the *continuation cost* of a type j-agent who is in q ($\hat{q}$) by C_j ($\hat{C}_j$), $j = L, H$. Then agents' equilibrium behaviour is governed by the following four Bellman equations:

- Bellman equation for low-cost agents who have never been punished:

$$C_L = \min_{\delta \in \{0,1\}} \Big[\delta\big(\gamma - \alpha + \tfrac{1-\phi}{2}(f + \zeta \hat{C}_L) + \tfrac{1+\phi}{2}\zeta C_L\big) + (1-\delta)\big(\tfrac{1+\phi}{2}(f + \zeta \hat{C}_L) + \tfrac{1-\phi}{2}\zeta C_L\big)\Big]. \quad (10)$$

  The expected discounted costs a low-cost agent who has never been punished incurs should he contribute ($\delta = 1$) is given by the δ-part of the right-hand side of this equation. If such an agent contributes, then he incurs the cost of contributing $\gamma - \alpha$. With probability $\frac{1-\phi}{2}$ he is erroneously found guilty of shirking. If so, he is punished by an amount f and moves to $\hat{q}$. If the planner does not make a mistake (which happens with probability $\frac{1+\phi}{2}$), then he is not punished and stays in q. If the agent decides to shirk (the $(1-\delta)$-part), then the planner detects him misbehaving with probability $\frac{1+\phi}{2}$, in which case the agent is punished by an amount f and moves to $\hat{q}$. With probability $\frac{1-\phi}{2}$ he escapes being punished and thus stays in q. In all cases the agent continues to the next period with probability ζ.

- Bellman equation for low-cost agents who have been punished before:

$$\hat{C}_L = \min_{\delta \in \{0,1\}} \Big[\delta\big(\gamma - \alpha + \tfrac{1-\phi}{2}\hat{f}\big) + (1-\delta)\tfrac{1+\phi}{2}\hat{f} + \zeta \hat{C}_L\Big]. \quad (11)$$

  This equation (and the ones pertaining to high-cost agents displayed below) is constructed in a similar fashion as (10) is. Notice that an agent cannot escape from $\hat{q}$. So, C_L is absent from (11).[17] Furthermore, agents in $\hat{q}$ receive the punishment $\hat{f}$ when found guilty of shirking.

- Bellman equation for high-cost agents who have never been punished:

$$C_H = \min_{\delta \in \{0,1\}} \Big[\delta\big(\gamma + \tfrac{1-\phi}{2}(f + \zeta \hat{C}_H) + \tfrac{1+\phi}{2}\zeta C_H\big) + (1-\delta)\big(\tfrac{1+\phi}{2}(f + \zeta \hat{C}_H) + \tfrac{1-\phi}{2}\zeta C_H\big)\Big]. \quad (12)$$

- Bellman equation for high-cost agents who have been punished before:

$$\hat{C}_H = \min_{\delta \in \{0,1\}} \Big[\delta(\gamma + \tfrac{1-\phi}{2}\hat{f}) + (1-\delta)\tfrac{1+\phi}{2}\hat{f} + \zeta \hat{C}_H\Big]. \quad (13)$$

[16] Clearly, if low-cost or high-cost agents would employ a mixed strategy, then an infinitesimal increase in (one of) the punishment(s) would lead to an upward jump in contributions. This effect renders mixed strategy equilibria impossible. We can thus confine attention to pure strategies.

[17] If the planner would have limited recall, a possibility that is explored in Section 5, then an agent who is currently in $\hat{q}$ can return to q. The continuation cost $\hat{C}_L$ would in that case depend on C_L.

Just like in the two-period model, any equilibrium in which two different punishments ($f \neq \hat{f}$) are used, must be such that low-cost agents always contribute, whereas a high-cost agent only contributes when in $\hat{q}$. One easily verifies that $\hat{\delta}_L = 1$ is optimal as long as $\phi\hat{f} \geq \gamma - \alpha$ and that $\hat{\delta}_H = 1$ is optimal as long as $\phi\hat{f} \geq \gamma$. Consequently, if the planner does use differentiated punishments, then $\hat{f} = \frac{\gamma}{\phi}$. Combining this observation with the strategies $\hat{\delta}_L = \hat{\delta}_H = 1$ yields

$$\hat{C}_L\Big|_{f\neq\hat{f}} = \frac{\gamma - \alpha + \frac{1-\phi}{2\phi}\gamma}{1-\zeta}, \quad \hat{C}_H\Big|_{f\neq\hat{f}} = \frac{\gamma + \frac{1-\phi}{2\phi}\gamma}{1-\zeta}. \tag{14}$$

These continuation costs are simply the discounted costs of contributing in each period plus the discounted expected costs of being (erroneously) punished.

Low-cost agents in q decide to contribute if $\gamma - \alpha \leq \phi(f + \zeta(\hat{C}_L - C_L))$, as can be gathered from (10). On the other hand, by rewriting (12) one sees that high-cost agents in q do *not* contribute as long as $\gamma > \phi(f + \zeta(\hat{C}_H - C_H))$. In Appendix A we show that with differentiated punishments these incentive compatibility constraints are equivalent to

$$\phi f \geq \gamma - \alpha - \tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2}\alpha, \quad \phi f < \gamma. \tag{15}$$

Observe that the incentive compatibility constraint of low-cost agents resembles its counterpart in the two-period setting (see (4)). The reduction in punishment $\frac{1-\phi}{2\phi}\alpha$ (i.e. the size of the reputation effect) that the planner uses if she employs differentiated punishments in the two-period setting is now multiplied by $\frac{\zeta}{1-\zeta}$, an agent's expected horizon. If $\zeta = \frac{1}{2}$, then this expectation equals 1 and the constraint reduces to the one of the two-period setting (in which the expected number of future periods is also 1). Since $\frac{\zeta}{1-\zeta}$ increases in ζ, the smallest punishment that ensures that low-cost agents in q contribute decreases in the discount factor ζ. In fact, if ζ is sufficiently large, it suffices to merely 'warn' agents in q (i.e. $f^* = 0$). This is intuitive: the larger ζ is, the more important expected 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.

The incentive compatibility constraint of high-cost agents does differ dramatically from its counterpart (5). The reason is that the planner uses at most two punishment levels in the current setting, whereas she uses three levels (the menu $\boldsymbol{f}^*$) in the two-period setting if using differentiated punishments is optimal. Since both the expected punishment for agents in $\hat{q}$ and a high-cost agent's cost of contributing are equal to γ, a high-cost agent simply refrains from contributing when in q as long as the expected punishment for agents in q is less than γ.

Clearly, since $\gamma - \alpha - \frac{\zeta}{1-\zeta}\frac{1-\phi}{2}\alpha < \gamma$, the planner can always find an f such that agents opt for $(\delta_L, \delta_H) = (1, 0)$. In Appendix A we show that inducing these choices can be optimal. In fact:

Proposition 3 *There exist $\check{\rho} = \check{\rho}(\alpha, \beta, \gamma, \phi, \zeta) < \bar{\rho}$ and $\hat{\rho} = \hat{\rho}(\alpha, \beta, \gamma, \phi, \zeta) > \bar{\rho}$ such that the social planner maximizes per-period welfare by committing to the pair of punishments*

$$f^* = \max\{\tfrac{\gamma-\alpha}{\phi} - \tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2\phi}\alpha, 0\}, \quad \hat{f}^* = \tfrac{\gamma}{\phi} \tag{16}$$

as long as $\rho \in [\check{\rho}, \hat{\rho}]$. If $\rho < \check{\rho}$ or $\rho > \hat{\rho}$, then the social planner commits to the single punishment given in (2). Furthermore, $\hat{\rho} < 1$ if and only if $\gamma - \alpha < \frac{\zeta}{1-\zeta}\frac{1-\phi}{2}\alpha$, whereas $\check{\rho} > 0$ for all parameter configurations.

Proof. See Appendix A. ■

If graduated punishments are used, then low-cost agents always contribute, whereas high-cost agents shirk as long as they have never been punished. On average, a high-cost agent shirks $\frac{1}{1-\zeta\frac{1-\phi}{2}}$ times.[18]

The structure of the optimal graduated punishments scheme in the dynamic setting departs in several respects from the one we obtained in Proposition 2 for the two-period setting. In the latter setting agents only fear losing their good reputation in the first period. The incentives of agents consequently differ across periods and the planner therefore has to use three punishment levels when opting for graduated punishments: one for those found guilty of shirking in period 1, one for first time offenders in period 2, and one for repeat offenders in that period. Only two punishment levels are used in the stationary equilibrium of the dynamic setting. The punishment level that is not used in the dynamic setting is $\frac{\gamma-\alpha}{\phi}$, the punishment that is administered in period 2 of the two-period setting to first-time offenders. These agents do not mind losing their good reputation. A loss of reputation always increases an agent's expected discounted future costs in the dynamic setting. In the dynamic setting the planner can therefore allow herself to always administer the cost-efficient punishment f^* to (alleged) first time offenders instead of the more costly punishment $\frac{\gamma-\alpha}{\phi}$.

The cost reduction that the planner (maximally) achieves by using graduated punishments instead of a uniform punishment scheme (i.e. $\frac{\zeta}{1-\zeta}\frac{1-\phi}{2\phi}\alpha$) is the size of the reputation effect ($\frac{1-\phi}{2\phi}\alpha$) times the expected horizon ($\frac{\zeta}{1-\zeta}$). As ζ becomes sufficiently large, the planner arrives at a corner solution in which she merely issues warnings to first time offenders ($f^* = 0$). Even though such a solution leads to zero costs of punishing first-time offenders, it is suboptimal if the fraction of low-cost agents ρ is very large. The explanation of this result has already been alluded to in footnote 13: If $\rho = 1$ and $f^* = \frac{\gamma-\alpha}{\phi} - \frac{\zeta}{1-\zeta}\frac{1-\phi}{2\phi}\alpha$, then the reduction in aggregate punishments to alleged first-time offenders due to the reputation effect exactly equals the increase in aggregate punishments to alleged repeat offenders due to the fact that they receive the punishment $\frac{\gamma}{\phi}$ instead of $\frac{\gamma-\alpha}{\phi}$. However, if $\frac{\gamma-\alpha}{\phi} < \frac{\zeta}{1-\zeta}\frac{1-\phi}{2\phi}\alpha$, then the planner, being forced to set $f^* = 0$, cannot fully exploit the reputation effect and the reduction in aggregate punishments to alleged first-time offenders consequently drops below the increase in aggregate punishments to alleged repeat offenders.

Intuitively, if both ρ and ζ are close to 1, then the population consists mainly of (long lived) low-cost agents. Because ζ is large, it is very likely that such a low-cost agent spends

[18] With probability $\frac{1+\phi}{2} + \frac{1-\phi}{2}(1-\zeta) = 1 - \zeta\frac{1-\phi}{2}$ a shirking high-cost agent is caught shirking or fails to advance to the next period. In both cases he stops shirking. With the complementary probability $\zeta\frac{1-\phi}{2}$ he advances to the next period and shirks in that period. So, the expected number of times a high-cost agent shirks is $\sum_{k=0}^{\infty}(k+1)(1-\zeta\frac{1-\phi}{2})(\zeta\frac{1-\phi}{2})^k = \frac{1}{1-\zeta\frac{1-\phi}{2}}$.

a large part of his life in $\hat{q}$: because monitoring is imperfect, the probability that a (law-abiding) low-cost agent is found guilty of shirking at least once in τ periods goes to 1 as $\tau \to \infty$. In fact, $\hat{q} = 1 - q$, the fraction of the population that has been punished at least once, converges to 1 as $\zeta \uparrow 1$, as can be inferred from (9). So, if ρ and ζ are both large, then the vast majority of agent in $\hat{q}$ are low-cost-agents. Administering the punishment $\hat{f}^* = \frac{\gamma}{\phi}$ to such agents is clearly suboptimal: it suffices to commit to a punishment of $\frac{\gamma-\alpha}{\phi}$ to deter these agents from shirking. Since the number of high-cost agents in $\hat{q}$ is negligible if ρ is close to 1, administering the punishment $\hat{f}^*$ to repeat offenders is dominated by administering the more cost-efficient punishment $\frac{\gamma-\alpha}{\phi}$. The planner therefore does not use graduated punishments if ρ and ζ are both close to 1.

The comparative statics with respect to $\check{\rho}$ do not differ qualitatively from those obtained for the two-period model. Summarizing:

Corollary 2 *An increase in ϕ leads to an increase in $\check{\rho}$, whereas an increase in α, β, γ, or ζ leads to a decrease in $\check{\rho}$.*

Proof. See Appendix A. ■

5 Discussion

In this section we investigate the robustness of our results by considering some extensions and discuss our work's relation to the literature.

5.1 Extensions

Differing Type I and Type II Errors: We have assumed in the previous sections that the planner's monitoring technology is such that the probability that she fails to detect a shirker (a type I error) equals the probability that she falsely judges someone guilty of shirking (a type II error). It is very likely that monitoring technologies are in general prone to both types of mistakes. Yet, the assumption that the associated probabilities are equal is rather strong. Our calculations reveal that relaxing this assumption does not affect the nature of the prevailing equilibrium. It does reveal that type II errors are essential to the optimality of graduated punishments: if the planner would never make type II errors, then she would never opt for graduated punishments. By contrast, graduated punishments can be optimal if the planner never makes type I errors.

Consider the two-period model of Section 3, but with the probability ϵ that an error occurs replaced by two probabilities: a probability $\epsilon_I \in (0, \frac{1}{2})$ that a type I error occurs and a probability $\epsilon_{II} \in (0, \frac{1}{2})$ that a type II error occurs. The quality of a monitoring technology with this more general specification reads $\tilde{\phi} := 1 - \epsilon_I - \epsilon_{II} > 0$. The following results pertaining to this generalization hold:

Proposition 4 *Suppose laissez-faire is never optimal.*[19] *Then:*

- *In the one-shot setting the planner commits to*

$$\tilde{\phi}\tilde{f}_0^* = \begin{cases} \gamma & \text{if } \rho \leq \tilde{\tilde{\rho}} \\ \gamma - \alpha & \text{if } \rho > \tilde{\tilde{\rho}}, \end{cases} \tag{17}$$

  where

$$\tilde{\tilde{\rho}} = 1 - \frac{\epsilon_{II}\beta\alpha}{\tilde{\phi}\big(1-\gamma+\beta(\gamma-\alpha)\big)}.$$

- *In the two-period setting there exist $\tilde{\rho} < \tilde{\tilde{\rho}}$ and $\tilde{\hat{\rho}} > \tilde{\tilde{\rho}}$ such that the planner maximizes aggregate welfare by committing to the menu of punishments*

$$\tilde{f}_1^* = \max\{\tfrac{\gamma-\alpha-\epsilon_{II}\alpha}{\tilde{\phi}}, 0\}, \quad \tilde{f}_2^* = \tfrac{\gamma-\alpha}{\tilde{\phi}}, \quad \tilde{\hat{f}}_2^* = \tfrac{\gamma}{\tilde{\phi}} \tag{18}$$

  as long as $\rho \in [\tilde{\rho}, \tilde{\hat{\rho}}]$. If $\rho < \tilde{\rho}$ or $\rho > \tilde{\hat{\rho}}$, then the planner commits in both periods to the single punishment given in (17). Furthermore, $\tilde{\hat{\rho}} < 1$ if and only if $\gamma - \alpha < \epsilon_{II}\alpha$, whereas $\tilde{\rho} > 0$ for all parameter configurations.

Proof. See Appendix B. ■

Comparing Proposition 4 with Propositions 1-2 reveals that in the one-shot setting as well as in the two-period model the probability ϵ_{II} that the planner erroneously punishes someone who did contribute plays a more prominent role than the probability ϵ_I that the planner fails to detect a shirker. To understand the relation between $\tilde{\tilde{\rho}}$ and ϵ_{II}, recall that a fraction ϵ_{II} is punished if $\tilde{\phi}\tilde{f}_0^* = \gamma$ is used, whereas a fraction $\rho\epsilon_{II} + (1-\rho)(1-\epsilon_I)$ is punished if $\tilde{\phi}\tilde{f}_0^* = \gamma - \alpha$. In words, if the planner moves from $\tilde{\phi}\tilde{f}_0^* = \gamma$ to $\tilde{\phi}\tilde{f}_0^* = \gamma - \alpha$, then she stops making type II errors related to high-cost agents, but she starts making type I errors related to those agents. This move therefore leads to an increase of $(1-\rho)\tilde{\phi}$ in the fraction of the population that is punished. Such an increase makes using $\tilde{\phi}\tilde{f}_0^* = \gamma - \alpha$ instead of $\tilde{\phi}\tilde{f}_0^* = \gamma$ less attractive. The second disadvantage of the low punishment stems from the fact that high-cost agents do not contribute if the punishment is low. These two disadvantages are the building blocks of the denominator of $1 - \tilde{\tilde{\rho}}$. The sole advantage of using $\tilde{\phi}\tilde{f}_0^* = \gamma - \alpha$ instead of $\tilde{\phi}\tilde{f}_0^* = \gamma$ is that the punishment $\frac{\gamma-\alpha}{\tilde{\phi}}$ is, of course, less costly than the punishment $\frac{\gamma}{\tilde{\phi}}$. This reduction in punishment costs only applies to the fraction ϵ_{II} of the population that would be punished if the high punishment were used, explaining the fact that the low punishment is never used if type II errors do not occur ($\tilde{\tilde{\rho}} = 1$ if $\epsilon_{II} = 0$). Because a low-cost agent contributes in both periods, the probability that such an agent is found guilty of shirking in period 2 is simply the probability ϵ_{II} that a type II error occurs. So, the reputation effect equals $\tilde{\hat{f}}_2^* - \tilde{f}_2^*$ times ϵ_{II} and hence $\tilde{f}_1^* = \max\{\frac{\gamma-\alpha-\epsilon_{II}\alpha}{\tilde{\phi}}, 0\}$. Note that the reputation effect vanishes as ϵ_{II} approaches 0.

[19] This condition boils down to $1 - \gamma > \frac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma$.

Punishing the Innocent: Making a type II error results in an injustice: an innocent person is punished for an offense he did not commit. In general societies appear to be more worried about such injustices than about possible type I errors. Indeed, several authors have pointed out that there may be social costs associated with the conviction of law-abiding citizens above and beyond the costs associated with administering punishment.[20] This idea can be incorporated in the extension that allows for differing type I and type II errors by attaching an additional weight $\delta > 0$ to punishments stemming from type II errors. In other words, the marginal social cost of punishing someone who did contribute becomes $\beta + \delta$, whereas the marginal social cost of punishing a shirker remains equal to β. Our main results continue to hold in this more general setting:

Proposition 5 *Suppose punishing the innocent entails an additional social cost. Specifically, let the marginal social cost of punishing a contributor be $\beta + \delta$, $\delta > 0$, instead of β (the marginal social cost of punishing a shirker). Then for any $\delta > 0$ such that laissez-faire is never optimal one has:*[21]

- *In the one-shot setting the planner commits to*

$$\tilde{\phi}\tilde{f}_0^*(\delta) = \begin{cases} \gamma & \text{if } \rho \le \tilde{\tilde{\rho}}(\delta) \\ \gamma - \alpha & \text{if } \rho > \tilde{\tilde{\rho}}(\delta), \end{cases} \tag{19}$$

  where

$$\tilde{\tilde{\rho}}(\delta) = 1 - \frac{\frac{\epsilon_{II}}{\tilde{\phi}}(\beta + \delta)\alpha}{1 - \gamma + \beta(\gamma - \alpha) - \frac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma - \alpha)}$$

  decreases in δ.

- *In the two-period setting there exist $\tilde{\check{\rho}}(\delta) < \tilde{\tilde{\rho}}(\delta)$ and $\tilde{\hat{\rho}}(\delta) > \tilde{\tilde{\rho}}(\delta)$ such that the planner maximizes aggregate welfare by committing to the menu of punishments $\tilde{\boldsymbol{f}}^*$ given in (18) as long as $\rho \in [\tilde{\check{\rho}}(\delta), \tilde{\hat{\rho}}(\delta)]$. If $\rho < \tilde{\check{\rho}}(\delta)$ or $\rho > \tilde{\hat{\rho}}(\delta)$, then the planner commits in both periods to the single punishment given in (19). Furthermore, $\tilde{\check{\rho}}(\delta)$ is strictly decreasing in δ, whereas $\tilde{\hat{\rho}}(\delta)$ is weakly decreasing in δ.*

Proof. See Appendix C. ∎

If punishing the innocent entails an additional social cost ($\delta > 0$), then using the high punishment $\frac{\gamma}{\tilde{\phi}}$ is particularly unattractive: as this punishment ensures that all agents contribute, the number of type II errors is maximal should the planner use this punishment. The parameter range for which the optimal punishment of the one-shot setting is $\frac{\gamma}{\tilde{\phi}}$ therefore shrinks as δ increases. In other words, $\tilde{\tilde{\rho}}(\delta)$ decreases in δ.

[20] See for instance the discussion in Chu et al. (2000, p. 130).

[21] Laissez-faire is never optimal if and only if $1 - \gamma > \frac{\epsilon_{II}}{\tilde{\phi}}(\beta + \delta)\gamma$. This condition obviously becomes more difficult to meet as δ increases.

The same reasoning explains why $\tilde{\tilde{\rho}}(\delta)$ decreases in δ. If $\rho \leq \tilde{\tilde{\rho}}(\delta)$, then in the two-period setting the planner chooses between the high uniform punishment and graduated punishments. The larger δ, the larger the social costs associated with the first option relative to the social costs associated with the second option, implying that indeed $\tilde{\tilde{\rho}}'(\delta) < 0$. If $\rho > \tilde{\tilde{\rho}}(\delta)$, then the planner chooses between the low uniform punishment and graduated punishments. Using the low uniform punishment minimizes the number of type II errors that occur. The planner is consequently more inclined to use the low uniform punishment instead of graduated punishments the larger δ is. Yet, if δ is such that laissez-faire is never optimal, then the effect of δ on aggregate welfare is small compared to the advantages of using graduated punishments. The upper bound $\tilde{\tilde{\rho}}(\delta)$ therefore only decreases in δ if the upper bound is already less than 1 in the absence of δ, i.e. if the planner administers warnings to first-time shirkers.

Limited Recall: In the dynamic model of Section 4 the planner has a perfect memory. This is not only a strong assumption, but it might also be suboptimal. The reason is that an agent, once he is in the high punishment regime $\hat{q}$, cannot return to q if the planner 'refuses to forget' that the agent has been found guilty of shirking in some past period. The planner is consequently forced to apply the high punishment $\hat{f}^*$ each time this agent is erroneously found guilty of shirking. Because monitoring is imperfect, this happens with positive probability in equilibrium. With limited recall, however, the planner enables herself to reduce the cost of administering punishments by 'moving agents back to q'. On the other hand, such 'forgetting' comes at a cost. Compared to the perfect memory setting agents spend on average more time in q if the planner has limited recall. High-cost agents consequently shirk more frequently in the latter setting. Furthermore, losing one's reputation is less frightening if a stay in $\hat{q}$ is only temporary. The size of the reputation effect is thus bound to be smaller if the planner has limited recall.

Our calculations indicate that in general the costs of forgetting outweigh its benefits. Yet, using graduated punishments in the presence of limited recall often improves welfare compared to using a uniform punishment in the same way as using graduated punishments with perfect recall does. Formally:

Proposition 6 *Suppose the planner can only remember whether or not an agent has received a punishment in the previous period. Then in the stationary equilibrium she uses graduated punishments instead of a uniform punishment if* $\rho \in [\check{\rho}_{lim}, \hat{\rho}_{lim}]$, *for some* $\check{\rho}_{lim} \in (0, \bar{\rho})$ *and* $\hat{\rho}_{lim} \in (\bar{\rho}, 1]$.[22] *The associated punishments abide by*

$$\phi f^*_{lim} = \max\{\gamma - \alpha - \frac{\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha, 0\}, \quad \phi \hat{f}^*_{lim} = \max\{\gamma - \frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}\alpha, \frac{1}{1+\zeta\frac{1+\phi}{2}}\gamma\}.$$

Furthermore, if using graduated punishments welfare-dominates the uniform punishment

[22]The thresholds $\check{\rho}_{\text{lim}}$ and $\hat{\rho}_{\text{lim}}$ exhibit the same behaviour as their counterparts $\check{\rho}$ and $\hat{\rho}$ of the perfect recall model. In particular, $\hat{\rho}_{\text{lim}} = 1$ if and only if the punishment that alleged first-time offenders receive equals $\gamma - \alpha$ minus the reputation effect.

scheme, then the planner prefers perfect recall to limited recall as long as she would be able to fully exploit the reputation effect (i.e. $\phi f^* = \gamma - \alpha - \frac{\zeta}{1-\zeta}\frac{1-\phi}{2}\alpha \geq 0$).[23]

Proof. See Appendix D. ■

The advantage of limited recall that we already mentioned, i.e. the reduced frequency with which the planner administers the high punishment associated with repeat offenders, is especially important if the population consists mainly of long-lived low-cost agents (both ρ and ζ are large). If the planner opts for graduated punishments with perfect recall in such a case, then the high punishment $\hat{f}^*$ will quite often be used to punish law-abiding (low-cost) agents who happen to reside in $\hat{q}$ by mistake. As is explained in Section 4, the planner prefers the uniform punishment scheme to the graduated punishments with perfect recall should this problem be severe. Using limited recall instead of perfect recall certainly alleviates the problem, but it does not take it away completely. The planner therefore resorts to the uniform punishment scheme.[24]

The expression for $\hat{f}^*_{\text{lim}}$ reveals that limited recall has a second advantage vis-à-vis perfect recall. Since agents in $\hat{q}$ are able to return to q, the punishment such agents receive when found guilty of shirking is in general lower than $\hat{f}^*$, the punishment they would receive if the planner had perfect recall. The reason is that agents in $\hat{q}$ have an incentive to 'regain a good reputation': if such an agent contributes, then he does not only dodge (with large probability) being punished, but he also moves back to the low punishment regime q (with the same probability). So, contributing when in $\hat{q}$ bestows a benefit (in expected terms) on the contributing agent, thereby reducing the net expected cost of contributing. The planner can consequently set the punishment for repeat offenders below $\hat{f}^*$ if she has limited recall. Despite this second advantage of limited recall, the downsides of limited recall - more shirking by high-cost agents, less fear of losing one's reputation when in q - outweigh its merits, rendering graduated punishments with limited recall inutile should the planner possess a perfect memory. Nevertheless, if the planner can only choose between graduated punishments with limited recall and the uniform punishment, then she does choose graduated punishment with limited recall.

5.2 Relation to the Literature

Graduated punishments have received quite some theoretical attention, most notably from law and economics scholars. Various explanations for this phenomenon have been proposed. Miceli and Bucci (2005) argue that the dire labour market prospects of convicted criminals makes committing crimes relatively more attractive for those who already have a criminal record. This effect can be negated by punishing repeat offenders harsher than first-time

[23]If the planner is unable to fully exploit the reputation effect, then one cannot assess analytically the sign of the per-period welfare difference. However, numerical computations indicate that the planner also prefers perfect recall to limited recall in those cases.

[24]The fact that using limited recall alleviates the problem does imply that graduated punishments with limited recall welfare-dominates graduated punishments with perfect recall if ρ and ζ are sufficiently large.

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 the actual punishment for first-time offenders. Of course, law enforcers could also learn from past offenses, yielding an increase in the probability that repeat offenses are detected. If law enforcers learn more than offenders, then the optimal punishment for repeat offenders is lower than the one for first-time offenders.[25]

Stigler (1970) argued informally that heavy penalties are unnecessary for first-time offenders if they are likely to have committed the offense accidentally and the probability of repetition is negligible. In Rubinstein (1979) offenses may also have been committed by accident. Convicting innocent offenders is detrimental to welfare. Rubinstein shows that it is then optimal to be lenient towards individuals with a 'reasonable' criminal record, i.e. those individuals are not administered the exogenously given punishment. Erroneous convictions also play a central role in Chu et al. (2000). Their planner tries to minimize total social costs, which consists of the harm imposed on society by criminal conduct and the cost of erroneous convictions. Chu et al. (2000) establish that in a two-period setting society is always best off if alleged repeat offenders are punished more severely than alleged first-time offenders. Such a solution is optimal, because the probability of convicting an innocent offender twice is much lower than convicting an innocent offender only once. Since punishing those who did commit crimes is costless, Chu et al.'s planner does not face a trade-off between crime prevention and cost minimization comparable to our trade-off between public good provision and cost minimization. Furthermore, they do not allow the punishment for first-time offenders in period 1 to differ from its counterpart in period 2. Their solution consequently fails to appreciate any reputation effects.

Polinsky and Rubinfeld (1991) study a setting with perfect monitoring. They assume that an individual's gain from committing some crime has two components: a socially acceptable gain and an illicit gain. The latter is a fixed trait. By contrast, an individual's acceptable gain is drawn from some distribution at the start of each period. Both components are private information. The planner maximizes aggregate acceptable gains minus harms stemming from criminal activities by choosing fines for first and second offenses. Since some crimes are socially efficient, the planner never opts for full deterrence. Individuals who commit crimes in the first period are likely to enjoy high illicit gains, especially if the fine for first offenses is low. This allows the planner to sort agents by 'illicit type'. Using higher fines for second offenses reduces underdeterrence vis-à-vis low uniform fines and reduces overdeterrence vis-à-vis high uniform fines, making such graduated fines socially optimal for some parameter values.[26]

Unlike Polinsky and Rubinfeld (1991), Polinsky and Shavell (1998) only consider acceptable gains in their two-period model with perfect monitoring. Polinsky and Shavell's

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

[26] If acceptable gains were fixed and illicit gains were drawn at the start of each period, then it can be optimal to use *lower* fines for second offenses.

planner has to expend resources to apprehend offenders and punishments cannot exceed some upper bound. Because administering punishments itself is costless, the planner uses this maximal punishment should using a uniform punishment be optimal. Since employing graduated punishments creates a reputation effect (a difference in tomorrow's punishments for first-time and repeat offenders makes agents more reluctant to commit a crime today), it can be optimal to set the punishment for first-time offenders in the second period below the maximal punishment. This reputation effect increases crime deterrence in period one, but reduces deterrence in period two. Whether the positive period-one effect outweighs the negative period-two effect depends on the distribution from which acceptable gains are drawn. In contrast to our reputation effect, Polinsky and Shavell's reputation effect has no impact on the punishment that prevails in period one.

In Rubinstein (1980) an agent's income should he abide the law is stochastic. Because the probability that the agent is caught when committing a crime is less than one, his income from criminal activities is also stochastic. Whether a uniform punishment scheme (in which punishments for first-time offenders and repeat offenders equal the maximal punishment) or a graduated punishment scheme is best at minimizing the number of offenses depends on the agent's risk attitude.

Warnings play a prominent role in Harrington (1988). Harrington, studying the enforcement of compliance with environmental regulations, shows that a planner who knows each firm's cost of compliance can achieve a higher compliance rate (compared to a system with a uniform punishment) by resorting to a system in which firms with relatively good compliance records are merely warned. Just like in our model, firms do not want to lose their good reputation, i.e. move to a high punishment regime. Yet, since Harrington (1988) assumes perfect monitoring, this result hinges on the presence of an upper bound to punishments. In a more recent paper, Rousseau (2009) argues that the use of warnings reduces the number of erroneous convictions and at the same time mitigates overcompliance to regulations by low types. Importantly, Rousseau assumes that the structure of punishments is exogenously given and that the planner can only choose between administering the appropriate punishment and warning the alleged violator.

Our approach is related to the model developed by Abreu et al. (2005). They study ongoing relationships between two players in which one player is tempted to depart from jointly efficient behaviour. How tempted that player is is private information. The other player receives signals regarding the tempted player's behaviour and can administer punishments to that player. In equilibrium punishments can go in either direction after perceived bad behaviour. The sign of the change in punishment depends crucially on the distribution from which the level of temptation is drawn. Although Abreu et al. (2005) stress that both asymmetric information and imperfect monitoring are a prerequisite for graduated punishments to occur, the setting they consider differs considerably from ours. They investigate a one-sided prisoner's dilemma with players who try to maximize their own payoff. In our public good game only the agents are selfish, the planner is benevolent. More importantly, the player who is tempted to depart from jointly efficient behaviour is infinitely impatient. As a consequence, reputation effects do not play a role in Abreu et al. (2005).

6 Conclusions

We have investigated the optimal punishment scheme a benign social planner uses when confronted with a repeated public good problem. Because monitoring is imperfect and administering punishments is costly, a uniform punishment that deters all agents from shirking is often suboptimal. To alleviate the detrimental effects on total welfare of monitoring mistakes and costly punishments, the planner employs 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, because agents fear losing their reputation and becoming branded as shirkers, the planner can allow herself to sanction first time offenders very mildly. In fact, mere warnings are often optimal.

Obviously, one can envision more elaborate punishment schemes. For instance, in most judiciary systems punishments do not simply depend on whether a convicted criminal has a criminal record, but on the precise content of such a record. Furthermore, we have only looked at the stationary equilibria of the dynamic model. We have consequently left an important question unanswered: when do groups or societies reach steady states in which graduated punishments are employed? To answer this question one needs to analyze the short run of the dynamic setting. These two issues might prove fruitful avenues for future research.

Appendix A: Proofs

Details regarding Condition 1

High-cost agents contribute if and only if the expected costs they face when contributing ($\gamma + \epsilon f_0$) does not exceed the expected costs they face when shirking ($(1-\epsilon)f_0$). So, the lowest punishment that ensures that all agents contribute is $f_0 = \frac{\gamma}{\phi}$. Using (1) and the definition of ϕ one obtains $W(\frac{\gamma}{\phi}) = 1 - \gamma + \rho\alpha - \beta\frac{1-\phi}{2\phi}\gamma$, a decreasing function of ρ. Since laissez-faire yields zero welfare, it follows that the planner does not opt for laissez-faire if Condition 1 holds.

Proof of Proposition 1

Straightforward calculations reveal that

$$W(\tfrac{\gamma-\alpha}{\phi}) = \rho(1-\gamma+\alpha) - (\tfrac{1+\phi}{2\phi} - \rho)\beta(\gamma-\alpha), \quad W(\tfrac{\gamma}{\phi}) = 1 - \gamma + \rho\alpha - \beta\tfrac{1-\phi}{2\phi}\gamma. \tag{A.1}$$

The difference in welfare between the two options (as a function of ρ) reads

$$\Delta(\rho) := W(\tfrac{\gamma}{\phi}) - W(\tfrac{\gamma-\alpha}{\phi}) = (1-\rho)(1-\gamma) + (1-\rho)\beta\gamma - (\tfrac{1+\phi}{2\phi} - \rho)\beta\alpha.$$

Because $\frac{\partial\Delta}{\partial\rho} = -(1-\gamma) - \beta(\gamma-\alpha) < 0$ and $\Delta(1) = -\frac{1-\phi}{2\phi}\beta\alpha < 0$, it is optimal to set $\phi f_0 = \gamma - \alpha$ if ρ is sufficiently large. Specifically, $f_0^* = \frac{\gamma-\alpha}{\phi}$ if $\rho > \bar{\rho}$, where

$$\bar{\rho} = 1 - \frac{\frac{1-\phi}{2\phi}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}$$

solves $\Delta(\rho) = 0$. Observe that:

$$\Delta(0) = 1 - \gamma + \beta(\gamma - \tfrac{1+\phi}{2\phi}\alpha) > 1 - \gamma + \beta(\gamma - \tfrac{1+\phi}{2\phi}\gamma) = 1 - (1 + \tfrac{1-\phi}{2\phi}\beta)\gamma > 0,$$

where we have used the fact that $\gamma > \alpha$ and Condition 1. The inequality $\Delta(0) > 0$ implies that $\bar{\rho}$ must be strictly positive. ■

Proof of Proposition 2

The aggregate welfare $\mathcal{W}(f_0^*)$ generated when the planner uses only one punishment follows from Proposition 1:

$$\mathcal{W}(f_0^*) = \begin{cases} 2W(\frac{\gamma}{\phi}) & \text{if } \rho \leq \bar{\rho} \\ 2W(\frac{\gamma-\alpha}{\phi}) & \text{if } \rho > \bar{\rho}. \end{cases}$$

We have to compare this figure with $\boldsymbol{f}^*$. Let us first derive the total welfare $W_2(\boldsymbol{f}^*)$ generated in period 2 if the menu of punishments $\boldsymbol{f}^*$ is used.

In period 2 all low-cost agents as well as those high-cost agents who have been caught shirking in the first period contribute. This yields, after taking into account agents' costs of contributing, an aggregate payoff of

$$\rho(1 - \gamma + \alpha) + (1 - \rho)\tfrac{1+\phi}{2}(1 - \gamma), \tag{A.2}$$

where we have used the fact that a fraction $1 - \epsilon = \frac{1+\phi}{2}$ of the high-cost agents are punished in period 1. We have to deduct the social costs of administering punishments from this figure. The aggregate punishment in period 2 reads:

$$\rho\epsilon^2\hat{f}_2^* + \rho(1-\epsilon)\epsilon f_2^* + (1-\rho)(1-\epsilon)\epsilon\hat{f}_2^* + (1-\rho)\epsilon(1-\epsilon)f_2^* = \tfrac{1-\phi}{2\phi} \times \big((1 + \phi - \rho\phi)\gamma - \tfrac{1+\phi}{2}\alpha\big).$$

To see this, recall that in each period the planner misjudges a particular agent's behaviour with probability ϵ. So, in period 1 a fraction ϵ of the low-cost agents are erroneously punished, whereas a fraction ϵ of the high-cost agents escape being befittingly punished. A fraction ϵ of the low-cost agents who have been punished in the first period are again punished in the second period. Since they are considered repeat offenders the high punishment $\hat{f}_2^*$ is applied. This reasoning explains the ρ-part of the left-hand side. The $(1 - \rho)$-part is constructed along similar lines. Combining the aggregate punishment with (A.2) results in

$$W_2(\boldsymbol{f}^*) = \rho(1 - \gamma + \alpha) + (1 - \rho)\tfrac{1+\phi}{2}(1 - \gamma) - \beta\tfrac{1-\phi}{2\phi} \times \big((1 + \phi - \rho\phi)\gamma - \tfrac{1+\phi}{2}\alpha\big).$$

We now determine the total welfare $W_1(\boldsymbol{f}^*)$ generated in period 1. This equals

$$W_1(\boldsymbol{f}^*) = \rho(1 - \gamma + \alpha) - \big(\rho\tfrac{1-\phi}{2} + (1 - \rho)\tfrac{1+\phi}{2}\big)\beta f_1^*.$$

This figure depends crucially on whether $\gamma - \alpha - \frac{1-\phi}{2}\alpha$ is positive or negative, i.e. on whether $\phi f_1^* = \gamma - \alpha - \frac{1-\phi}{2}\alpha$ or $\phi f_1^* = 0$. We consider these two possibilities in turn:

Suppose $\gamma \geq (1 + \frac{1-\phi}{2})\alpha$. Then $\phi f_1^* = \gamma - \alpha - \frac{1-\phi}{2}\alpha$. Adding $W_1(\boldsymbol{f}^*)$ and $W_2(\boldsymbol{f}^*)$ now yields:

$$\mathcal{W}(\boldsymbol{f}^*) = 2\rho(1 - \gamma + \alpha) + (1 - \rho)\tfrac{1+\phi}{2}(1 - \gamma) - \big(\tfrac{1+\phi}{2\phi} + \tfrac{1-\phi^2}{2\phi} - \rho(1 + \tfrac{1-\phi}{2})\big)\beta(\gamma - \alpha).$$

We have to analyze the cases $\rho \leq \bar{\rho}$ and $\rho > \bar{\rho}$ separately:

The case $\rho \leq \bar{\rho}$: In this case $\mathcal{W}(f_0^*) = 2W(\frac{\gamma}{\phi})$ and therefore:

$$\begin{aligned}\Psi :=& \mathcal{W}(\boldsymbol{f}^*) - \mathcal{W}(f_0^*) \\ =& -2(1-\rho)(1-\gamma) + (1-\rho)\tfrac{1+\phi}{2}(1-\gamma) - \big(\tfrac{1+\phi}{2\phi} + \tfrac{1-\phi^2}{2\phi} - \rho(1+\tfrac{1-\phi}{2})\big)\beta(\gamma-\alpha) + \tfrac{1-\phi}{\phi}\beta\gamma \\ =& -(1-\rho)\big(1+\tfrac{1-\phi}{2}\big)\big((1-\gamma)+\beta(\gamma-\alpha)\big) + \tfrac{1-\phi}{\phi}\beta\alpha.\end{aligned} \tag{A.3}$$

Observe that Ψ is strictly increasing in ρ. Evaluating $\Psi = \Psi(\rho)$ at $\rho = \bar{\rho}$ yields

$$\begin{aligned}\Psi(\bar{\rho}) =& -\frac{\frac{1-\phi}{2\phi}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)} \times \big(1+\tfrac{1-\phi}{2}\big)\big((1-\gamma)+\beta(\gamma-\alpha)\big) + \tfrac{1-\phi}{\phi}\beta\alpha \\ =& \tfrac{1-\phi}{2\phi}\beta\alpha \times \big(-1-\tfrac{1-\phi}{2}+2\big) > 0.\end{aligned}$$

The last two observations imply that $\Psi(\rho) \geq 0$ if $\rho \in [\check{\rho}, \bar{\rho}]$, where $\check{\rho} = \check{\rho}(\alpha,\beta,\gamma,\phi) < \bar{\rho}$ equals $\max\{\Psi^{-1}(0), 0\}$.

To prove that $\check{\rho} = \Psi^{-1}(0) > 0$ if $\gamma \geq (1+\frac{1-\phi}{2})\alpha$, it suffices to show that $\Psi(0) < 0$:

$$\begin{aligned}\Psi(0) =& -\big(1+\tfrac{1-\phi}{2}\big)\big((1-\gamma)+\beta(\gamma-\alpha)\big) + \tfrac{1-\phi}{\phi}\beta\alpha < -\big(1+\tfrac{1-\phi}{2}\big)\tfrac{1-\phi}{2\phi}\beta\gamma + \tfrac{1-\phi}{\phi}\beta\alpha \\ =& \tfrac{1-\phi}{2\phi}\beta \times \Big(-\big(1+\tfrac{1-\phi}{2}\big)\gamma + 2\alpha\Big) \leq \tfrac{1-\phi}{2\phi}\beta \times \Big(-\big(1+\tfrac{1-\phi}{2}\big)^2\alpha + 2\alpha\Big) < 0,\end{aligned}$$

where the first inequality follows from Condition 1 and the fact that $\gamma > \alpha$ and the second one holds as long as $\gamma \geq (1+\frac{1-\phi}{2})\alpha$.

The case $\rho > \bar{\rho}$: The difference in aggregate welfare now reads:

$$\begin{aligned}\mathcal{W}(\boldsymbol{f}^*) - \mathcal{W}(f_0^*) =& (1-\rho)\tfrac{1+\phi}{2}(1-\gamma) - \big(\tfrac{1+\phi}{2\phi} + \tfrac{1-\phi^2}{2\phi} - \rho(1+\tfrac{1-\phi}{2})\big)\beta(\gamma-\alpha) \\ &+ \big(\tfrac{1+\phi}{\phi} - 2\rho\big)\beta(\gamma-\alpha) \\ =& (1-\rho)\tfrac{1+\phi}{2}(1-\gamma+\beta(\gamma-\alpha)) > 0.\end{aligned}$$

We conclude that the planner opts for the menu $\boldsymbol{f}^*$ if $\rho > \check{\rho}$ and $\gamma - \alpha - \frac{1-\phi}{2}\alpha \geq 0$.

Suppose next $\gamma < (1+\frac{1-\phi}{2})\alpha$. Then $\phi f_1^* = 0$ and aggregate welfare using the menu $\boldsymbol{f}^*$ equals

$$\mathcal{W}(\boldsymbol{f}^*) = 2\rho(1-\gamma+\alpha) + (1-\rho)\tfrac{1+\phi}{2}(1-\gamma) - \beta\tfrac{1-\phi}{2\phi}\big((1+\phi-\rho\phi)\gamma - \tfrac{1+\phi}{2}\alpha\big).$$

Again, we have to distinguish between $\rho \leq \bar{\rho}$ and $\rho > \bar{\rho}$:

The case $\rho \leq \bar{\rho}$: The difference in aggregate welfare $\Psi(\rho)$ is:

$$\begin{aligned}\Psi(\rho) =& -(1-\rho)(2-\tfrac{1+\phi}{2})(1-\gamma) - \beta\tfrac{1-\phi}{2\phi}\big((1+\phi-\rho\phi)\gamma - \tfrac{1+\phi}{2}\alpha - 2\gamma\big) \\ =& -(1-\rho)(1+\tfrac{1-\phi}{2})(1-\gamma) + \tfrac{1-\phi}{2\phi}(1-(1-\rho)\phi)\beta\gamma + \tfrac{1-\phi}{2\phi}\tfrac{1+\phi}{2}\beta\alpha.\end{aligned} \tag{A.4}$$

Evaluating this expression at $\rho = \bar{\rho}$ results in:

$$\begin{aligned}
\Psi(\bar{\rho}) =& \frac{\frac{1-\phi}{2\phi}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)} \times \Big[-(1-\gamma)(1+\tfrac{1-\phi}{2}) - \tfrac{1-\phi}{2}\beta\gamma \Big] + \tfrac{1-\phi}{2\phi}\beta\gamma + \tfrac{1-\phi}{2\phi}\tfrac{1+\phi}{2}\beta\alpha \\
=& \tfrac{1-\phi}{2\phi}\beta\alpha \times \Big[\tfrac{1+\phi}{2} - \frac{(1+\frac{1-\phi}{2})(1-\gamma) + \frac{1-\phi}{2}\beta\gamma}{1-\gamma+\beta(\gamma-\alpha)} \Big] + \tfrac{1-\phi}{2\phi}\beta\gamma \\
>& \tfrac{1-\phi}{2\phi}\beta\alpha \times \Big[\frac{-(1-\phi)(1-\gamma) + \frac{1+\phi}{2}\beta(\gamma-\alpha) - \frac{1-\phi}{2}\beta\gamma}{1-\gamma+\beta(\gamma-\alpha)} + 1 \Big] \\
=& \tfrac{1-\phi}{2\phi}\beta\alpha \times \frac{(1+\frac{1+\phi}{2})\beta(\gamma-\alpha) + \phi(1-\gamma) - \frac{1-\phi}{2}\beta\gamma}{1-\gamma+\beta(\gamma-\alpha)} \\
>& \tfrac{1-\phi}{2\phi}\beta\alpha \times \frac{(1+\frac{1+\phi}{2})\beta(\gamma-\alpha)}{1-\gamma+\beta(\gamma-\alpha)} > 0,
\end{aligned} \tag{A.5}$$

where the first inequality follows from the fact that $\gamma > \alpha$ and the second one from Condition 1. Since $\frac{\partial\Psi}{\partial\rho} = (1+\frac{1-\phi}{2})(1-\gamma) + \frac{1-\phi}{2}\beta\gamma > 0$, we again conclude that $\Psi(\rho) \geq 0$ if $\rho \in [\check{\rho}, \bar{\rho}]$, where $\check{\rho} = \check{\rho}(\alpha, \beta, \gamma, \phi) < \bar{\rho}$ equals $\max\{\Psi^{-1}(0), 0\}$.

Again, $\check{\rho} > 0$:

$$\begin{aligned}
\Psi(0) =& -\big(1+\tfrac{1-\phi}{2}\big)(1-\gamma) + \tfrac{1-\phi}{2\phi}(1-\phi)\beta\gamma + \tfrac{1-\phi}{2\phi}\tfrac{1+\phi}{2}\beta\alpha \\
<& -\big(1+\tfrac{1-\phi}{2}\big)\tfrac{1-\phi}{2\phi}\beta\gamma + \tfrac{1-\phi}{2\phi}(1-\phi)\beta\gamma + \tfrac{1-\phi}{2\phi}\tfrac{1+\phi}{2}\beta\alpha = \tfrac{1-\phi}{2\phi}\beta \times \Big[\tfrac{1+\phi}{2}\alpha - \tfrac{1+\phi}{2}\gamma\Big] < 0,
\end{aligned}$$

The case $\rho > \bar{\rho}$: The relevant welfare difference $\Psi(\rho) = \mathcal{W}(\boldsymbol{f}^*) - \mathcal{W}(f_0^*)$ now reads

$$\begin{aligned}
\Psi(\rho) =& (1-\rho)\tfrac{1+\phi}{2}(1-\gamma) - \beta\tfrac{1-\phi}{2\phi}\big((1+\phi-\rho\phi)\gamma - \tfrac{1+\phi}{2}\alpha\big) + (\tfrac{1+\phi}{2\phi} - \rho)\beta(\gamma-\alpha) \\
=& (1-\rho)\Big[\tfrac{1+\phi}{2}(1-\gamma) + \tfrac{1+\phi}{2}\beta\gamma - \beta\alpha\Big] - \tfrac{1-\phi}{2}\tfrac{1-\phi}{2\phi}\beta\alpha.
\end{aligned}$$

Clearly, this expression is negative if ρ is sufficiently large. On the other hand, by continuity of Ψ, we infer from (A.5) that $\Psi(\rho) > 0$ for $\rho > \bar{\rho}$ not too large. Since, using Condition 1,

$$\tfrac{1+\phi}{2}(1-\gamma) + \tfrac{1+\phi}{2}\beta\gamma - \beta\alpha = \tfrac{1+\phi}{2}(1-\gamma) - \tfrac{1-\phi}{2}\beta\gamma + \beta(\gamma-\alpha) > \tfrac{1+\phi}{2}(1-\gamma) - \phi(1-\gamma) + \beta(\gamma-\alpha) > 0,$$

we know that Ψ decreases monotonically in ρ for $\rho \in (\bar{\rho}, 1)$. We conclude that there exists a $\hat{\rho} = \hat{\rho}(\alpha, \beta, \gamma, \phi) \in (\bar{\rho}, 1)$ such that $\Psi(\rho) \geq 0$ if $\rho \in (\bar{\rho}, \hat{\rho}]$, whereas $\Psi(\rho) < 0$ if $\rho \in (\hat{\rho}, 1]$. This observation completes the proof. ∎

Proof of Corollary 1

Suppose first that $\gamma - \alpha - \frac{1-\phi}{2}\alpha \geq 0$. Then from (A.3) we infer that $\check{\rho}$ solves

$$-(1-\rho)\big(1+\tfrac{1-\phi}{2}\big)\big((1-\gamma) + \beta(\gamma-\alpha)\big) + \tfrac{1-\phi}{\phi}\beta\alpha = 0. \tag{A.6}$$

Totally differentiating this expression yields after some rearranging

$$\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi} = \frac{-\frac{1-\check{\rho}}{2}(1-\gamma+\beta(\gamma-\alpha)) + \frac{1}{\phi^2}\beta\alpha}{(1+\frac{1-\phi}{2})(1-\gamma+\beta(\gamma-\alpha))}.$$

The denominator of $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi}$ is clearly positive. The equality (A.6) implies that

$$\tfrac{1}{\phi^2}\beta\alpha = \tfrac{1}{\phi(1-\phi)}(1-\check{\rho})\big(1+\tfrac{1-\phi}{2}\big)\big((1-\gamma)+\beta(\gamma-\alpha)\big).$$

The sign of the numerator of $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi}$ is consequently the same as the sign of

$$-\frac{1}{2}+\frac{1+\frac{1-\phi}{2}}{\phi(1-\phi)}=\frac{2+(1-\phi)^2}{2\phi(1-\phi)}.$$

We conclude that $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi}>0$.

Using (A.6) one obtains

$$\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\beta}=\frac{(1-\check{\rho})(1+\frac{1-\phi}{2})(\gamma-\alpha)-\frac{1-\phi}{\phi}\alpha}{(1+\frac{1-\phi}{2})(1-\gamma+\beta(\gamma-\alpha))}.$$

The denominator of this expression is clearly positive. Furthermore:

$$(1-\check{\rho})(1+\tfrac{1-\phi}{2})(\gamma-\alpha)-\tfrac{1-\phi}{\phi}\alpha=\beta^{-1}\times-(1-\check{\rho})\big(1+\tfrac{1-\phi}{2}\big)(1-\gamma)<0,$$

where the equality follows from (A.6). So, $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\beta}<0$.

It is straightforward to verify that both $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\gamma}$ and $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\alpha}$ are negative.

Next, suppose $\gamma-\alpha-\frac{1-\phi}{2}\alpha<0$. Equation (A.4) informs us that $\check{\rho}$ now solves

$$-(1-\rho)\big(1+\tfrac{1-\phi}{2}\big)(1-\gamma)-(1-\rho)\tfrac{1-\phi}{2}\beta\gamma+\tfrac{1-\phi}{2\phi}\beta\gamma+\tfrac{1-\phi}{2\phi}\tfrac{1+\phi}{2}\beta\alpha=0. \qquad \text{(A.7)}$$

The derivative of $\check{\rho}$ with respect to ϕ therefore abides by

$$\Big[\big(1+\tfrac{1-\phi}{2}\big)(1-\gamma)+\tfrac{1-\phi}{2}\beta\gamma\Big]\mathrm{d}\check{\rho}+\tfrac{1}{2}\Big[(1-\check{\rho})(1-\gamma)+(1-\check{\rho})\beta\gamma-\tfrac{1}{\phi^2}\beta\gamma-\tfrac{1+\phi^2}{2\phi^2}\beta\alpha\Big]\mathrm{d}\phi=0.$$

Since (A.7) is equivalent to

$$(1-\rho)(1-\gamma)+(1-\rho)\beta\gamma=\tfrac{1}{\phi}\beta\gamma+\tfrac{1+\phi}{2\phi}\beta\alpha-\tfrac{2}{1-\phi}(1-\rho)(1-\gamma),$$

we conclude that

$$(1-\check{\rho})(1-\gamma)+(1-\check{\rho})\beta\gamma-\tfrac{1}{\phi^2}\beta\gamma-\tfrac{1+\phi^2}{2\phi^2}\beta\alpha=-\tfrac{1-\phi}{\phi^2}\beta\gamma-\tfrac{1-\phi}{2\phi^2}\beta\alpha-\tfrac{2}{1-\phi}(1-\check{\rho})(1-\gamma)<0$$

and thus that $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi}>0$.

The derivative of $\check{\rho}$ with respect to β reads

$$\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\beta}=\frac{(1-\check{\rho})\frac{1-\phi}{2}\gamma-\frac{1-\phi}{2\phi}\gamma-\frac{1-\phi}{2\phi}\frac{1+\phi}{2}\alpha}{(1+\frac{1-\phi}{2})(1-\gamma)+\frac{1-\phi}{2}\beta\gamma}.$$

Because (A.7) implies that

$$(1-\check{\rho})\tfrac{1-\phi}{2}\gamma-\tfrac{1-\phi}{2\phi}\gamma-\tfrac{1-\phi}{2\phi}\tfrac{1+\phi}{2}\alpha=\beta^{-1}\times-(1-\check{\rho})\big(1+\tfrac{1-\phi}{2}\big)(1-\gamma)<0,$$

this derivative is negative.

Again, one easily verifies that both $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\gamma}$ and $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\alpha}$ are negative. ■

Proof of Lemma 1
Combining (7) and (8) results in the following system of equalities:

$$\begin{pmatrix} q \\ \mu \end{pmatrix} = \begin{bmatrix} \zeta\frac{1-\phi}{2} & \zeta\phi \\ 0 & \zeta\frac{1+\phi}{2} \end{bmatrix} \begin{pmatrix} q \\ \mu \end{pmatrix} + \begin{pmatrix} 1-\zeta \\ \rho(1-\zeta) \end{pmatrix} \Longleftrightarrow \begin{pmatrix} q \\ \mu \end{pmatrix} = \begin{bmatrix} 1-\zeta\frac{1-\phi}{2} & -\zeta\phi \\ 0 & 1-\zeta\frac{1+\phi}{2} \end{bmatrix}^{-1} \begin{pmatrix} 1-\zeta \\ \rho(1-\zeta) \end{pmatrix}$$

The expressions given in (9) now follow immediately from the fact that

$$\begin{bmatrix} 1-\zeta\frac{1-\phi}{2} & -\zeta\phi \\ 0 & 1-\zeta\frac{1+\phi}{2} \end{bmatrix}^{-1} = \frac{1}{(1-\zeta\frac{1-\phi}{2})(1-\zeta\frac{1+\phi}{2})} \begin{bmatrix} 1-\zeta\frac{1+\phi}{2} & \zeta\phi \\ 0 & 1-\zeta\frac{1-\phi}{2} \end{bmatrix}.$$ ■

Details regarding (15)
If $\delta_L = 1$, then (10) reduces to

$$C_L = \gamma - \alpha + \tfrac{1-\phi}{2}(f + \zeta\hat{C}_L) + \tfrac{1+\phi}{2}\zeta C_L \Leftrightarrow C_L = \frac{\gamma - \alpha + \frac{1-\phi}{2}(f+\zeta\hat{C}_L)}{1-\frac{1+\phi}{2}\zeta},$$

from which one infers using (14) that

$$\hat{C}_L - C_L = \frac{(1-\zeta)\hat{C}_L - (\gamma-\alpha) - \frac{1-\phi}{2}f}{1-\zeta\frac{1+\phi}{2}} = \frac{\frac{1-\phi}{2}(\frac{\gamma}{\phi} - f)}{1-\zeta\frac{1+\phi}{2}}.$$

Consequently:

$$\gamma - \alpha \leq \phi(f + \zeta(\hat{C}_L - C_L)) \Leftrightarrow \phi f \geq (\gamma-\alpha) - \frac{\zeta\frac{1-\phi}{2}(\gamma - \phi f)}{1-\zeta\frac{1+\phi}{2}} \Leftrightarrow \phi f \geq \gamma - \alpha - \tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2}\alpha.$$

Substituting $\delta_H = 0$ in (12) yields

$$C_H = \tfrac{1+\phi}{2}(f + \zeta\hat{C}_H) + \tfrac{1-\phi}{2}\zeta C_H \Leftrightarrow C_H = \frac{\frac{1+\phi}{2}(f+\zeta\hat{C}_H)}{1-\zeta\frac{1-\phi}{2}}.$$

Combining the latter expression with (14) results in

$$\hat{C}_H - C_H = \frac{(1-\zeta)\hat{C}_H - \frac{1+\phi}{2}f}{1-\zeta\frac{1-\phi}{2}} = \frac{\frac{1+\phi}{2}(\frac{\gamma}{\phi} - f)}{1-\zeta\frac{1-\phi}{2}}.$$

Therefore:

$$\gamma > \phi(f + \zeta(\hat{C}_H - C_H)) \Leftrightarrow \phi f < \gamma - \frac{\zeta\frac{1+\phi}{2}(\gamma - \phi f)}{1-\zeta\frac{1-\phi}{2}} \Leftrightarrow \phi f < \gamma.$$

Proof of Proposition 3

We have to compare the total welfare $\Omega = \Omega(f^*, \hat{f}^*)$ generated in a period if the pair $(f^*, \hat{f}^*)$ is used with $W(f_0^*)$ (given in (A.1)), the per-period welfare generated if the planner uses only one punishment. We calculate Ω step by step. The per-period welfare without the costs of administering punishments reads

$$\rho(1-\gamma+\alpha) + (\hat{q}-\hat{\mu})(1-\gamma) = \rho(1-\gamma+\alpha) + \big[(1-\rho)-(q-\mu)\big](1-\gamma),$$

where we have used the identities $\hat{q} = 1-q$, $\hat{\mu} = \rho - \mu$ and the fact that low-cost agents always contribute whereas high-cost agents only contribute when in $\hat{q}$. Lemma 1 informs us that

$$q-\mu = \frac{(1-\zeta)(1-\zeta\frac{1+\phi}{2}+\rho\zeta\phi)}{(1-\zeta\frac{1-\phi}{2})(1-\zeta\frac{1+\phi}{2})} - \frac{\rho(1-\zeta)}{1-\zeta\frac{1+\phi}{2}} = \frac{(1-\zeta)(1-\rho)}{1-\zeta\frac{1-\phi}{2}}. \tag{A.8}$$

We conclude that, ignoring punishments, per-period welfare equals

$$\rho(1-\gamma+\alpha) + (1-\rho)\times\frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}(1-\gamma). \tag{A.9}$$

The aggregate per-period punishment reads

$$\big(\tfrac{1-\phi}{2}\tfrac{\mu}{q} + \tfrac{1+\phi}{2}(1-\tfrac{\mu}{q})\big)qf^* + \tfrac{1-\phi}{2}\hat{q}\hat{f}^* = (\tfrac{1+\phi}{2}q-\phi\mu)f^* + \tfrac{1-\phi}{2}(1-q)\hat{f}^*.$$

Notice that $f^* > 0$ only if $\gamma > \check{\gamma}$, where $\check{\gamma} := \alpha + \frac{\zeta}{1-\zeta}\frac{1-\phi}{2}\alpha$. We have to treat the cases $\gamma \geq \check{\gamma}$ (i.e. $\phi f^* = \gamma - \alpha - \frac{\zeta}{1-\zeta}\frac{1-\phi}{2}\alpha$) and $\gamma < \check{\gamma}$ (i.e. $\phi f^* = 0$) separately:

The case $\gamma \geq \check{\gamma}$: Then aggregate per-period punishment is:

$$(\tfrac{1+\phi}{2\phi}q-\mu)(\gamma-\alpha-\tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2}\alpha) + \tfrac{1-\phi}{2\phi}(1-q)\gamma.$$

In this expression γ is multiplied by

$$\tfrac{1+\phi}{2\phi}q-\mu+\tfrac{1-\phi}{2\phi}(1-q) = \frac{(1-\zeta)(1-\rho)}{1-\zeta\frac{1-\phi}{2}} + \tfrac{1-\phi}{2\phi},$$

where we have used (A.8) to arrive at the right-hand side of this equation. Furthermore, $-\alpha$ is multiplied by

$$\begin{aligned}(\tfrac{1+\phi}{2\phi}q-\mu)(1+\tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2}) &= \frac{1-\zeta}{1-\zeta\frac{1+\phi}{2}}\times\Big[\tfrac{1+\phi}{2\phi}\times\frac{1-\zeta\frac{1+\phi}{2}+\rho\zeta\phi}{1-\zeta\frac{1-\phi}{2}}-\rho\Big]\times\frac{1-\zeta\frac{1+\phi}{2}}{1-\zeta}\\ &= \frac{\frac{1+\phi}{2\phi}(1-\zeta\frac{1+\phi}{2})-\rho(1-\zeta)}{1-\zeta\frac{1-\phi}{2}} = \frac{(1-\zeta)(1-\rho)}{1-\zeta\frac{1-\phi}{2}}+\tfrac{1-\phi}{2\phi}.\end{aligned}$$

It follows that

$$(\tfrac{1+\phi}{2}q-\phi\mu)f^* + \tfrac{1-\phi}{2}(1-q)\hat{f}^* = \frac{(1-\zeta)(1-\rho)}{1-\zeta\frac{1-\phi}{2}}(\gamma-\alpha) + \tfrac{1-\phi}{2\phi}(\gamma-\alpha). \tag{A.10}$$

Combining (A.9) and (A.10) yields

$$\Omega(f^*,\hat{f}^*) = \rho(1-\gamma+\alpha) + (1-\rho)\times\frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}(1-\gamma) - \frac{(1-\zeta)(1-\rho)}{1-\zeta^{\frac{1-\phi}{2}}}\beta(\gamma-\alpha) - \tfrac{1-\phi}{2\phi}\beta(\gamma-\alpha) \quad \text{if } \gamma-\alpha-\tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2}\alpha \geq 0.$$

The subcase $\rho \leq \bar{\rho}$: The difference in welfare between the two options now reads

$$\begin{aligned}\Omega(f^*,\hat{f}^*) - W(f_0^*) =& (1-\rho)\frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}(1-\gamma) - \frac{(1-\zeta)(1-\rho)}{1-\zeta^{\frac{1-\phi}{2}}}\beta(\gamma-\alpha) - \tfrac{1-\phi}{2\phi}\beta(\gamma-\alpha)\\ &- (1-\rho)(1-\gamma) + \tfrac{1-\phi}{2\phi}\beta\gamma\\ =& -\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}(1-\rho)\big((1-\gamma)+\beta(\gamma-\alpha)\big) + \tfrac{1-\phi}{2\phi}\beta\alpha.\end{aligned}$$

Evaluating this expression at $\rho = \bar{\rho}$ results in

$$\Omega(f^*,\hat{f}^*) - W(f_0^*)\big|_{\rho=\bar{\rho}} = -\frac{(1-\zeta)}{1-\zeta^{\frac{1-\phi}{2}}}\tfrac{1-\phi}{2\phi}\beta\alpha + \tfrac{1-\phi}{2\phi}\beta\alpha = \Big(1-\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}\Big)\tfrac{1-\phi}{2\phi}\beta\alpha > 0.$$

Because $\Omega(f^*,\hat{f}^*) - W(f_0^*)$ increases in ρ, we conclude that the pair $(f^*,\hat{f}^*)$ maximizes per-period welfare if $\rho \geq \check{\rho}$ for some $\check{\rho} = \check{\rho}(\alpha,\beta,\gamma,\phi,\zeta) < \bar{\rho}$. Since

$$\begin{aligned}\Omega(f^*,\hat{f}^*) - W(f_0^*)\Big|_{\rho=0} &= -\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}\big((1-\gamma)+\beta(\gamma-\alpha)\big) + \tfrac{1-\phi}{2\phi}\beta\alpha\\ &\leq -\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}(1-\gamma) + \frac{1-\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\tfrac{1-\phi}{2\phi}\beta\alpha\\ &\leq -\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}(1-\gamma) + \frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}\tfrac{1-\phi}{2\phi}\beta\gamma < 0,\end{aligned}$$

where the fist two inequalities follow from $\gamma \geq \check{\gamma}$ and the last one from Condition 1, we infer that $\check{\rho} > 0$.

The subcase $\rho > \bar{\rho}$: In this parameter range the difference in welfare is

$$\begin{aligned}\Omega(f^*,\hat{f}^*) - W(f_0^*) =& \frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}(1-\rho)(1-\gamma) - \frac{(1-\zeta)(1-\rho)}{1-\zeta^{\frac{1-\phi}{2}}}\beta(\gamma-\alpha) - \tfrac{1-\phi}{2\phi}\beta(\gamma-\alpha)\\ &+ \big(\tfrac{1+\phi}{2\phi}-\rho\big)\beta(\gamma-\alpha)\\ =& \frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}(1-\rho)\Big((1-\gamma)+\beta(\gamma-\alpha)\Big) > 0.\end{aligned}$$

So, using the pair $(f^*,\hat{f}^*)$ instead of the single punishment f_0^* is also optimal if $\rho > \bar{\rho}$. This result implies that $\hat{\rho} = 1$ if $\gamma \geq \check{\gamma}$.

The case $\gamma < \check{\gamma}$: The aggregate punishment now reduces to $\frac{1-\phi}{2}(1-q)\hat{f}^*$. With the aid of (9) one can calculate this figure:

$$\tfrac{1-\phi}{2}(1-q)\hat{f}^* = \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1-\phi}{2}} \times \Big[\tfrac{1+\phi}{2\phi} - \rho\frac{(1-\zeta)}{1-\zeta\frac{1+\phi}{2}}\Big]\gamma.$$

Consequently:

$$\begin{aligned}\Omega(f^*,\hat{f}^*) =& \rho(1-\gamma+\alpha) + \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}(1-\rho)(1-\gamma)\\ &- \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1-\phi}{2}} \times \Big[\tfrac{1+\phi}{2\phi} - \rho\frac{(1-\zeta)}{1-\zeta\frac{1+\phi}{2}}\Big]\beta\gamma \quad \text{if } \gamma-\alpha-\tfrac{\zeta}{1-\zeta}\tfrac{1-\phi}{2}\alpha < 0.\end{aligned} \tag{A.11}$$

We again have to distinguish between $\rho \leq \bar{\rho}$ and $\rho > \bar{\rho}$:

The subcase $\rho \leq \bar{\rho}$: Subtracting $W(\frac{\gamma}{\phi})$ from (A.11) yields

$$\begin{aligned}\Omega(f^*,\hat{f}^*) - W(f_0^*) =& \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}(1-\rho)(1-\gamma) - \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1-\phi}{2}}\Big[\tfrac{1+\phi}{2\phi} - \rho\frac{(1-\zeta)}{1-\zeta\frac{1+\phi}{2}}\Big]\beta\gamma\\ &-(1-\rho)(1-\gamma) + \tfrac{1-\phi}{2\phi}\beta\gamma\\ =& \frac{1-\zeta}{1-\zeta\frac{1-\phi}{2}}\Big[-(1-\rho)(1-\gamma) + \tfrac{1-\phi}{2\phi}\beta\gamma + \frac{\rho\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1+\phi}{2}}\beta\gamma\Big]\\ =& \frac{1-\zeta}{1-\zeta\frac{1-\phi}{2}}\Big[-(1-\rho)(1-\gamma) + \frac{1-\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1+\phi}{2}}\tfrac{1-\phi}{2\phi}\beta\gamma - \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1+\phi}{2}}(1-\rho)\beta\gamma\Big].\end{aligned} \tag{A.12}$$

We again show that the difference $\Psi(\rho) := \Omega(f^*,\hat{f}^*) - W(f_0^*)$ is positive if $\rho = \bar{\rho}$:

$$\begin{aligned}\Psi(\bar{\rho}) =& \frac{\beta(1-\zeta)\frac{1-\phi}{2\phi}}{1-\zeta\frac{1-\phi}{2\phi}}\Big[-\frac{\alpha(1-\gamma)}{1-\gamma+\beta(\gamma-\alpha)} + \frac{1-\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1+\phi}{2}}\gamma - \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1+\phi}{2}} \times \frac{\alpha\beta\gamma}{1-\gamma+\beta(\gamma-\alpha)}\Big]\\ =& \frac{\beta(1-\zeta)\frac{1-\phi}{2\phi}}{1-\zeta\frac{1-\phi}{2\phi}}\Big[\frac{-\alpha(1-\gamma)(1-\zeta\frac{1+\phi}{2}) + (1-\gamma+\beta(\gamma-\alpha))(1-\zeta\frac{1-\phi}{2})\gamma - \zeta\frac{1-\phi}{2}\alpha\beta\gamma}{(1-\gamma+\beta(\gamma-\alpha))(1-\zeta\frac{1+\phi}{2})}\Big]\\ >& \frac{\beta(1-\zeta)\frac{1-\phi}{2\phi}}{1-\zeta\frac{1-\phi}{2\phi}}\Big[\frac{\zeta\phi(1-\gamma)\gamma - \zeta\frac{1-\phi}{2}\beta\gamma^2}{(1-\gamma+\beta(\gamma-\alpha))(1-\zeta\frac{1+\phi}{2})}\Big] > 0,\end{aligned} \tag{A.13}$$

where the first inequality follows from the fact that $\gamma > \alpha$ and the second one from Condition 1. Because $\Psi(\rho)$ increases in ρ, we conclude that $\Omega(f^*,\hat{f}^*) > W(f_0^*)$ if $\rho > \check{\rho}$, for some $\check{\rho} = \check{\rho}(\alpha,\beta,\gamma,\phi,\zeta) < \bar{\rho}$. Condition 1 implies that the threshold $\check{\rho}$ is again strictly positive:

$$\Omega(f^*,\hat{f}^*)-W(f_0^*)\Big|_{\rho=0}=\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}\Big[-(1-\gamma)+\frac{1-\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1+\phi}{2}}}\tfrac{1-\phi}{2\phi}\beta\gamma-\frac{\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1+\phi}{2}}}\beta\gamma\Big]$$
$$=\frac{1-\zeta}{1-\zeta^{\frac{1-\phi}{2}}}\Big[-(1-\gamma)+\tfrac{1-\phi}{2\phi}\beta\gamma\Big]<0.$$

The subcase $\rho>\bar{\rho}$: In this case we need to subtract $W(\frac{\gamma-\alpha}{\phi})$ from (A.11) to arrive at the relevant welfare difference:

$$\Psi(\rho):=\Omega(f^*,\hat{f}^*)-W(f_0^*)=\frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}(1-\rho)(1-\gamma)$$
$$-\frac{\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\Big(\tfrac{1+\phi}{2\phi}-\rho\frac{(1-\zeta)}{1-\zeta^{\frac{1+\phi}{2}}}\Big)\beta\gamma+(\tfrac{1+\phi}{2\phi}-\rho)\beta(\gamma-\alpha). \tag{A.14}$$

Observe that:

$$\Psi'(\rho)=-\frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}(1-\gamma)+\frac{\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\times\frac{1-\zeta}{1-\zeta^{\frac{1+\phi}{2}}}\beta\gamma-\beta(\gamma-\alpha)$$
$$<-\frac{\zeta^{\frac{1+\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\times\tfrac{1-\phi}{2\phi}\beta\gamma+\frac{\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\times\frac{1-\zeta}{1-\zeta^{\frac{1+\phi}{2}}}\beta\gamma$$
$$=\Big(-\tfrac{1+\phi}{2\phi}+\frac{1-\zeta}{1-\zeta^{\frac{1+\phi}{2}}}\Big)\times\frac{\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\beta\gamma<0,$$

where we have used Condition 1 and the fact that $\gamma>\alpha$ to establish the first inequality. The second inequality follows from the facts that $-\frac{1+\phi}{2\phi}<-1$ and $\frac{1-\zeta}{1-\zeta^{\frac{1+\phi}{2}}}<1$. We infer from the above inequality combined with (A.13) that $\Psi(\rho)$ is positive for $\rho>\bar{\rho}$ not 'too large'.[27] As ρ approaches 1, one obtains

$$\Psi(1)=-\frac{\zeta^{\frac{1-\phi}{2}}}{1-\zeta^{\frac{1-\phi}{2}}}\Big(\tfrac{1+\phi}{2\phi}-\frac{1-\zeta}{1-\zeta^{\frac{1+\phi}{2}}}\Big)\beta\gamma+\tfrac{1-\phi}{2\phi}\beta(\gamma-\alpha)$$
$$=-\frac{\zeta^{\frac{1-\phi}{2}}\frac{1-\phi}{2\phi}}{1-\zeta^{\frac{1+\phi}{2}}}\beta\gamma+\tfrac{1-\phi}{2\phi}\beta(\gamma-\alpha)=\tfrac{1-\phi}{2\phi}\beta\times\Big[\frac{1-\zeta}{1-\zeta^{\frac{1+\phi}{2}}}\gamma-\alpha\Big].$$

Because $\gamma<\check{\gamma}=\frac{1-\zeta^{\frac{1+\phi}{2}}}{1-\zeta}\alpha$, $\Psi(1)$ is always negative, implying that $\hat{\rho}<1$ if $\gamma<\check{\gamma}$. ■

Proof of Corollary 2

The claims regarding $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\alpha}$, $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\beta}$, and $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\gamma}$ as well as those pertaining to $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi}$ and $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\zeta}$ for $\gamma\geq\hat{\gamma}$ can be straightforwardly obtained. Their proofs are omitted.

[27] Notice that, by construction, $\Psi(\cdot)$ is continuous.

Suppose $\gamma < \hat{\gamma}$. Then from (A.12) we infer that $\check{\rho}$ solves

$$\tilde{\Psi} = -\big(1 - \zeta\tfrac{1+\phi}{2}\big)(1-\rho)(1-\gamma) + \big(1 - \zeta\tfrac{1-\phi}{2}\big)\tfrac{1-\phi}{2\phi}\beta\gamma - \zeta\tfrac{1-\phi}{2}(1-\rho)\beta\gamma = 0.$$

It is not difficult to see that $\frac{\partial\tilde{\Psi}}{\partial\check{\rho}} > 0$. Furthermore:

$$\frac{\partial\tilde{\Psi}}{\partial\phi} = \tfrac{\zeta}{2}(1-\rho)(1-\gamma) + \tfrac{\zeta}{2}\tfrac{1-\phi}{2\phi}\beta\gamma - \big(1 - \zeta\tfrac{1-\phi}{2}\big)\tfrac{1}{2\phi^2}\beta\gamma + \tfrac{\zeta}{2}(1-\rho)\beta\gamma.$$

Note that $\tilde{\Psi} = 0$ implies that

$$\tfrac{\zeta}{2}(1-\check{\rho})(1-\gamma) + \tfrac{\zeta}{2}(1-\check{\rho})\beta\gamma = \tfrac{1}{1-\phi} \times \Big(- (1-\zeta)(1-\check{\rho})(1-\gamma) + \big(1 - \zeta\tfrac{1-\phi}{2}\big)\tfrac{1-\phi}{2\phi}\beta\gamma\Big).$$

Therefore:

$$\begin{aligned}
\frac{\partial\tilde{\Psi}}{\partial\phi}\Big|_{\rho=\check{\rho}} &= -\tfrac{1}{1-\phi}(1-\zeta)(1-\check{\rho})(1-\gamma) + \tfrac{1}{2\phi}\beta\gamma - \big(1 - \zeta\tfrac{1-\phi}{2}\big)\tfrac{1}{2\phi^2}\beta\gamma \\
&< \Big(- (1-\zeta)(1-\check{\rho}) + 1 - \big(1 - \zeta\tfrac{1-\phi}{2}\big)\tfrac{1}{\phi}\Big) \times \tfrac{\beta\gamma}{2\phi} \\
&= \Big(\check{\rho}(1-\zeta) - \tfrac{1}{\phi} + \zeta\tfrac{1+\phi}{2\phi}\Big) \times \tfrac{\beta\gamma}{2\phi} < \Big((1-\zeta) - \tfrac{1}{\phi} + \zeta\tfrac{1+\phi}{2\phi}\Big) \times \tfrac{\beta\gamma}{2\phi} \\
&= \tfrac{1-\phi}{\phi}\big(-1 + \tfrac{\zeta}{2}\big) \times \tfrac{\beta\gamma}{2\phi} < 0,
\end{aligned}$$

where the first inequality follows from Condition 1. We conclude that $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\phi} > 0$. Differentiating $\tilde{\Psi}$ with respect to ζ yields

$$\frac{\partial\tilde{\Psi}}{\partial\zeta} = \tfrac{1+\phi}{2}(1-\rho)(1-\gamma) - \tfrac{1-\phi}{2}\tfrac{1-\phi}{2\phi}\beta\gamma - \tfrac{1-\phi}{2}(1-\rho)\beta\gamma.$$

Note that $\tilde{\Psi} = 0$ implies that

$$1 - \check{\rho} = \frac{(1 - \zeta\frac{1-\phi}{2})\frac{1-\phi}{2\phi}\beta\gamma}{(1 - \zeta\frac{1+\phi}{2})(1-\gamma) + \zeta\frac{1-\phi}{2}\beta\gamma}.$$

Combining this result with the above expression for $\frac{\partial\tilde{\Psi}}{\partial\zeta}$ results in

$$\begin{aligned}
\frac{\partial\tilde{\Psi}}{\partial\zeta}\Big|_{\rho=\check{\rho}} &= \frac{\frac{1+\phi}{2}(1-\gamma)(1 - \zeta\frac{1-\phi}{2})\frac{1-\phi}{2\phi}\beta\gamma}{(1 - \zeta\frac{1+\phi}{2})(1-\gamma) + \zeta\frac{1-\phi}{2}\beta\gamma} - \tfrac{1-\phi}{2}\tfrac{1-\phi}{2\phi}\beta\gamma - \frac{\frac{1-\phi}{2}\beta\gamma(1 - \zeta\frac{1-\phi}{2})\frac{1-\phi}{2\phi}\beta\gamma}{(1 - \zeta\frac{1+\phi}{2})(1-\gamma) + \zeta\frac{1-\phi}{2}\beta\gamma} \\
&= \tfrac{1-\phi}{2\phi}\beta\gamma \times \frac{\phi(1-\gamma) - \frac{1-\phi}{2}\beta\gamma}{(1 - \zeta\frac{1+\phi}{2})(1-\gamma) + \zeta\frac{1-\phi}{2}\beta\gamma} > 0,
\end{aligned}$$

where the inequality is courtesy of Condition 1. So, $\frac{\mathrm{d}\check{\rho}}{\mathrm{d}\zeta} < 0$. ■

Appendix B: Differing Type I and Type II Errors

Proof of Proposition 4
Consider the one-shot setting. A low-cost agent contributes as long as $\gamma - \alpha + \epsilon_{II}\tilde{f}_0 \leq (1-\epsilon_I)\tilde{f}_0 \Leftrightarrow \tilde{\phi}\tilde{f}_0 \geq \gamma - \alpha$, whereas a high-cost agent contributes as long as $\gamma + \epsilon_{II}\tilde{f}_0 \leq (1-\epsilon_I)\tilde{f}_0 \Leftrightarrow \tilde{\phi}\tilde{f}_0 \geq \gamma$. If the planner sets the expected punishment $\tilde{\phi}\tilde{f}_0$ equal to γ, then low-cost as well as high-cost agents contribute. Welfare with this punishment equals

$$\tilde{W}(\tfrac{\gamma}{\tilde{\phi}}) = \rho(1-\gamma+\alpha) + (1-\rho)(1-\gamma) - \tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma.$$

Setting $\rho = 0$ results in a welfare of $1 - \gamma - \frac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma$. So, laissez-faire is never optimal if $1 - \gamma > \frac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma$. If the planner opts for an expected punishment of $\gamma - \alpha$ only low-cost agents contribute and welfare amounts to

$$\tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}}) = \rho(1-\gamma+\alpha) - \rho\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta(\gamma-\alpha) - (1-\rho)\tfrac{1-\epsilon_I}{\tilde{\phi}}\beta(\gamma-\alpha).$$

Observe that:

$$\tilde{W}(\tfrac{\gamma}{\tilde{\phi}}) \geq \tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}}) \Longleftrightarrow 1-\rho \geq \frac{\frac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}.$$

This observation proves the optimality of (17).

We move on to the two-period setting. If the planner employs two different punishments in period 2, then she optimally sets $\tilde{\phi}\tilde{f}_2^* = \gamma - \alpha$ for first-time shirkers and $\tilde{\phi}\tilde{\tilde{f}}_2^* = \gamma$ for second-time shirkers. Denoting the punishment in period 1 by $\tilde{f}_1$, a low-cost agent contributes in that period if and only if

$$\gamma - \alpha + \epsilon_{II}(\tilde{f}_1 + \gamma - \alpha + \epsilon_{II}\tilde{\tilde{f}}_2^*) + (1-\epsilon_{II})(\gamma - \alpha + \epsilon_{II}\tilde{f}_2^*) \leq$$
$$(1-\epsilon_I)(\tilde{f}_1 + \gamma - \alpha + \epsilon_{II}\tilde{\tilde{f}}_2^*) + \epsilon_I(\gamma - \alpha + \epsilon_{II}\tilde{f}_2^*),$$

which, using the expressions for $\tilde{f}_2^*$ and $\tilde{\tilde{f}}_2^*$, reduces to $\tilde{\phi}\tilde{f}_1 \geq \gamma - \alpha - \epsilon_{II}\alpha$. A high-cost agent does not contribute in period 1 as long as

$$\gamma + \epsilon_{II}(\tilde{f}_1 + \gamma + \epsilon_{II}\tilde{\tilde{f}}_2^*) + (1-\epsilon_{II})(1-\epsilon_I)\tilde{f}_2^* > (1-\epsilon_I)(\tilde{f}_1 + \gamma + \epsilon_{II}\tilde{\tilde{f}}_2^*) + \epsilon_I(1-\epsilon_I)\tilde{f}_2^*,$$

that is as long as $\tilde{\phi}\tilde{f}_1 < \gamma - \alpha + \epsilon_I\alpha$. The two incentive compatibility constraints thus imply that the planner sets $\tilde{\phi}\tilde{f}_1^* = \max\{\gamma - \alpha - \epsilon_{II}\alpha, 0\}$ should she opt for graduated punishments. Let $\tilde{\boldsymbol{f}}^* = (\tilde{f}_1^*, \tilde{f}_2^*, \tilde{\tilde{f}}_2^*)$. If the planner employs graduated punishments, then in period 1 welfare amounts to

$$\tilde{W}_1(\tilde{\boldsymbol{f}}^*) = \rho(1-\gamma+\alpha) - \beta\big(\rho\epsilon_{II}\tilde{f}_1^* + (1-\rho)(1-\epsilon_I)\tilde{f}_1^*\big) \tag{B.1}$$

and in period 2 welfare amounts to

$$\tilde{W}_2(\tilde{\boldsymbol{f}}^*) = \rho(1-\gamma+\alpha) + (1-\epsilon_I)(1-\rho)(1-\gamma) - \beta\Big(\rho\epsilon_{II}^2\tilde{\tilde{f}}_2^* + \rho(1-\epsilon_{II})\epsilon_{II}\tilde{f}_2^* + (1-\rho)(1-\epsilon_I)\epsilon_{II}\tilde{\tilde{f}}_2^* + (1-\rho)\epsilon_I(1-\epsilon_I)\tilde{f}_2^*\Big). \tag{B.2}$$

We have to treat the two cases ($\gamma - \alpha - \epsilon_{II}\alpha \geq 0$, $\gamma - \alpha - \epsilon_{II}\alpha < 0$) separately:

The case $\gamma - \alpha - \epsilon_{II}\alpha \geq 0$: Total welfare $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*) = \tilde{W}_1(\tilde{\boldsymbol{f}}^*) + \tilde{W}_2(\tilde{\boldsymbol{f}}^*)$ now equals

$$\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*) = 2\rho(1-\gamma+\alpha) + (1-\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha)\big) - 2\beta\big(\rho\epsilon_{II} + (1-\rho)(1-\epsilon_I)\big)\tfrac{\gamma-\alpha}{\tilde{\phi}}.$$

Let us compare $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*)$ with $2\tilde{W}(\tilde{f}_0^*)$, the total welfare generated if the planner uses the uniform punishment. Two possibilities require attention:

The subcase $\rho \leq \tilde{\tilde{\rho}}$: The welfare difference $\tilde{\Psi} = \tilde{\Psi}(\rho) := \tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*) - 2\tilde{W}(\frac{\gamma}{\tilde{\phi}})$ now reads

$$\begin{aligned}\tilde{\Psi} =&(1-\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha)\big) - 2\beta\big(\rho\epsilon_{II} + (1-\rho)(1-\epsilon_I)\big)\tfrac{\gamma-\alpha}{\tilde{\phi}} \\ &- 2(1-\rho)(1-\gamma) + 2\beta\epsilon_{II}\tfrac{\gamma}{\tilde{\phi}} \\ =&-(1+\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha)\big) + 2\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha,\end{aligned} \tag{B.3}$$

where we have used the fact that $\tilde{\phi} = 1 - \epsilon_I - \epsilon_{II}$ to establish the second equality. Since

$$\tilde{\Psi}(\tilde{\tilde{\rho}}) = -(1+\epsilon_I)\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha + 2\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha = (1-\epsilon_I)\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha > 0$$

and $\tilde{\Psi}'(\rho) > 0$, we conclude that there exists a $\tilde{\tilde{\rho}} \in (0, \tilde{\tilde{\rho}})$ such that $\tilde{\Psi}(\rho) \geq 0$ if $\rho \in [\tilde{\tilde{\rho}}, \tilde{\tilde{\rho}}]$.[28]

The subcase $\rho > \tilde{\tilde{\rho}}$: Subtracting $2\tilde{W}(\frac{\gamma-\alpha}{\tilde{\phi}})$ from $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*)$ results in $(1-\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha)\big)$, which is clearly positive. So, $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*) \geq 2\tilde{W}(\frac{\gamma-\alpha}{\tilde{\phi}})$ for all $\rho \in (\tilde{\tilde{\rho}}, 1]$.

The case $\gamma - \alpha - \epsilon_{II}\alpha < 0$: With $\tilde{f}_1^*$ equal to 0, total welfare $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*)$ reduces to

$$\begin{aligned}\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*) =&2\rho(1-\gamma+\alpha) + (1-\rho)(1-\epsilon_I)(1-\gamma) - (1-\rho)\epsilon_I\beta\gamma - \tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma \\ &+ (1-\rho)(\epsilon_I - \epsilon_{II})\beta\alpha + \tfrac{(1-\epsilon_{II})\epsilon_{II}}{\tilde{\phi}}\beta\alpha.\end{aligned}$$

We again need to treat the two subcases separately:

The subcase $\rho \leq \tilde{\tilde{\rho}}$: Subtracting $2\tilde{W}(\frac{\gamma}{\tilde{\phi}})$ from $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*)$ evaluated at $\tilde{\boldsymbol{f}}^* = (0, \frac{\gamma-\alpha}{\tilde{\phi}}, \frac{\gamma}{\tilde{\phi}})$ yields

$$\begin{aligned}\tilde{\Psi}(\rho) =&-(1+\epsilon_I)(1-\rho)(1-\gamma) - (1-\rho)\epsilon_I\beta(\gamma-\alpha) + \tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma \\ &- (1-\rho)\epsilon_{II}\beta\alpha + (1-\epsilon_{II})\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha.\end{aligned} \tag{B.4}$$

[28]The laissez-faire condition together with the assumption that $\gamma \geq (1+\epsilon_{II})\alpha$ ensure that $\tilde{\tilde{\rho}} > 0$.

In particular:

$$
\begin{aligned}
\tilde{\Psi}(\tilde{\tilde{\rho}}) =& -\big((1+\epsilon_I)(1-\gamma)+\epsilon_I\beta(\gamma-\alpha)+\epsilon_{II}\beta\alpha\big)\frac{\frac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}+\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma+(1-\epsilon_{II})\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha \\
=& -\epsilon_I\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha-(1-\gamma+\epsilon_{II}\beta\alpha)\frac{\frac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}+\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma+(1-\epsilon_{II})\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha \\
=&\epsilon_{II}\beta\alpha-\frac{1-\gamma+\epsilon_{II}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha+\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma>\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha\times\Big[\tilde{\phi}-\frac{1-\gamma+\epsilon_{II}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}+1\Big] \\
=&\frac{\frac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}\times\Big[\tilde{\phi}\big(1-\gamma+\beta(\gamma-\alpha)\big)+\beta(\gamma-\alpha-\epsilon_{II}\alpha)\Big] \\
>&\frac{\frac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha}{1-\gamma+\beta(\gamma-\alpha)}\times\Big[\epsilon_{II}\beta\gamma+\tilde{\phi}\beta(\gamma-\alpha)+\beta(\gamma-\alpha)-\epsilon_{II}\beta\alpha\Big]>0,
\end{aligned}
$$

where the first and the last inequality stem from the fact that $\gamma > \alpha$ and the second one is a consequence of the laissez-faire condition ($\tilde{\phi}(1-\gamma) > \epsilon_{II}\beta\gamma$). Furthermore:

$$
\begin{aligned}
\tilde{\Psi}(0) =& -(1+\epsilon_I)(1-\gamma)-\epsilon_I\beta(\gamma-\alpha)+\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma-\epsilon_{II}\beta\alpha+(1-\epsilon_{II})\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha \\
<&\big(-(1+\epsilon_I)\tfrac{\epsilon_{II}}{\tilde{\phi}}-\epsilon_I+\tfrac{\epsilon_{II}}{\tilde{\phi}}\big)\beta\gamma+\big(\epsilon_I-\epsilon_{II}+(1-\epsilon_{II})\tfrac{\epsilon_{II}}{\tilde{\phi}}\big)\beta\alpha \\
=& -(1-\epsilon_I)\tfrac{\epsilon_I}{\tilde{\phi}}\beta\gamma+(1-\epsilon_I)\tfrac{\epsilon_I}{\tilde{\phi}}\beta\alpha<0,
\end{aligned}
$$

where the first inequality follows from the laissez-faire condition. Since $\tilde{\Psi}$ increases in ρ, the facts that $\tilde{\Psi}(0) < 0$ and $\tilde{\Psi}(\tilde{\tilde{\rho}}) > 0$ imply that there exists a $\tilde{\rho} \in (0, \tilde{\tilde{\rho}})$ such that $\tilde{\Psi}(\rho) \geq 0$ if $\rho \in [\tilde{\rho}, \tilde{\tilde{\rho}}]$.

The subcase $\rho > \tilde{\tilde{\rho}}$: The welfare difference $\tilde{\mathcal{W}}(0, \frac{\gamma-\alpha}{\tilde{\phi}}, \frac{\gamma}{\tilde{\phi}}) - 2\tilde{W}(\frac{\gamma-\alpha}{\tilde{\phi}})$ equals

$$
\begin{aligned}
\tilde{\Psi}(\rho) =&(1-\epsilon_I)(1-\rho)(1-\gamma)-(1-\rho)\epsilon_I\beta\gamma-\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma+(1-\rho)(\epsilon_I-\epsilon_{II})\beta\alpha \\
&+(1-\epsilon_{II})\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta\alpha+2(1-\rho)\beta(\gamma-\alpha)+2\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta(\gamma-\alpha) \\
=&(1-\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha)\big)+(1-\rho)\beta(\gamma-\alpha-\epsilon_{II}\alpha)+\tfrac{\epsilon_{II}}{\tilde{\phi}}\beta(\gamma-\alpha-\epsilon_{II}\alpha).
\end{aligned}
$$

Because $\gamma - \alpha - \epsilon_{II}\alpha < 0$, $\tilde{\Psi}(\rho) < 0$ for ρ sufficiently large. On the other hand, by continuity of $\tilde{\Psi}$ we know that $\tilde{\Psi}(\rho) > 0$ for $\rho > \tilde{\tilde{\rho}}$ close to $\tilde{\tilde{\rho}}$. Invoking the monotonicity of $\tilde{\Psi}(\rho)$ we conclude that there exists a $\hat{\tilde{\rho}} \in (\tilde{\tilde{\rho}}, 1)$ such that $\tilde{\Psi}(\rho) \geq 0$ if $\rho \in [\tilde{\tilde{\rho}}, \hat{\tilde{\rho}}]$. ■

Appendix C: Punishing the Innocent

Proof of Proposition 5

Notice that since the incentives of the agents do not depend on δ, provided laissez-faire is not optimal the planner sets either the low punishment $\frac{\gamma-\alpha}{\tilde{\phi}}$ or the high punishment $\frac{\gamma}{\tilde{\phi}}$ in the one-shot setting. Welfare with the low punishment equals

$$\tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}};\delta) = \tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}};0)-\rho\tfrac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha) = \rho(1-\gamma+\alpha)-\rho\tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)(\gamma-\alpha)-(1-\rho)\tfrac{1-\epsilon_{I}}{\tilde{\phi}}\beta(\gamma-\alpha)$$

and welfare with the high punishment equals

$$\tilde{W}(\tfrac{\gamma}{\tilde{\phi}};\delta) = \tilde{W}(\tfrac{\gamma}{\tilde{\phi}};0) - \tfrac{\epsilon_{II}}{\tilde{\phi}}\delta\gamma = \rho(1-\gamma+\alpha) + (1-\rho)(1-\gamma) - \tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\gamma.$$

Laissez-faire is never optimal as long as

$$\max\{\tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}};\delta), \tilde{W}(\tfrac{\gamma}{\tilde{\phi}};\delta)\} > 0.$$

Since $\tilde{W}(\frac{\gamma}{\tilde{\phi}};\delta)$ increases in ρ and $\tilde{W}(\frac{\gamma}{\tilde{\phi}};\delta)\big|_{\rho=0} = 1-\gamma-\frac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\gamma$, a sufficient condition for laissez-faire to never be optimal is

$$1-\gamma > \tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\gamma. \tag{C.1}$$

Because $\tilde{W}(\frac{\gamma-\alpha}{\tilde{\phi}};\delta)\big|_{\rho=0} < 0$ for all $\delta > 0$, (C.1) is also a necessary condition.

Straightforward calculations reveal that[29]

$$\tilde{W}(\tfrac{\gamma}{\tilde{\phi}};\delta) \geq \tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}};\delta) \iff 1-\rho \geq 1-\tilde{\tilde{\rho}}(\delta) := \frac{\frac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\alpha}{1-\gamma+\beta(\gamma-\alpha)-\frac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)},$$

proving the optimality of (19). Since

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\left(\frac{\frac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\alpha}{1-\gamma+\beta(\gamma-\alpha)-\frac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)}\right) > 0,$$

the threshold $\tilde{\tilde{\rho}}(\delta)$ is decreasing in δ.

We now look at the two-period setting. The planner then has to choose between using the uniform punishment $\tilde{f}_0^*$ in both periods and using the menu of punishments $\tilde{\boldsymbol{f}}^*$. The latter option yields a total welfare of

$$\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*;\delta) = \tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*;0) - \rho\epsilon_{II}\delta\tilde{f}_1^* - \big(\rho\epsilon_{II}^2\delta\tilde{\tilde{f}}_2^* + \rho(1-\epsilon_{II})\epsilon_{II}\delta\tilde{f}_2^* + (1-\rho)(1-\epsilon_I)\epsilon_{II}\delta\tilde{\tilde{f}}_2^*\big),$$

where $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*;0) = \tilde{W}_1(\tilde{\boldsymbol{f}}) + \tilde{W}_2(\tilde{\boldsymbol{f}})$ (i.e. (B.1) plus (B.2)). Two cases ($\gamma-\alpha-\epsilon_{II}\alpha \geq 0$ and $\gamma-\alpha-\epsilon_{II}\alpha < 0$) with each two subcases ($\rho \leq \tilde{\tilde{\rho}}(\delta)$ and $\rho > \tilde{\tilde{\rho}}(\delta)$) require attention.

The case $\gamma-\alpha-\epsilon_{II}\alpha \geq 0$: In this case the planner sets $\tilde{\phi}\tilde{f}_1^* = \gamma-\alpha-\epsilon_{II}\alpha$.

[29]The condition (C.1) ensures that $\tilde{\tilde{\rho}}(\delta) \in (0,1)$.

The subcase $\rho \leq \tilde{\tilde{\rho}}(\delta)$: The relevant welfare difference (welfare with graduated punishments minus two times $\tilde{W}(\frac{\gamma}{\tilde{\phi}};\delta)$) then reads

$$\begin{aligned}\tilde{\Psi}(\rho;\delta) =& \tilde{\Psi}(\rho;0) - \rho\epsilon_{II}\delta\tfrac{\gamma-\alpha-\epsilon_{II}\alpha}{\tilde{\phi}} - \big(\rho\epsilon_{II}^2\delta\tfrac{\gamma}{\tilde{\phi}} + \rho(1-\epsilon_{II})\epsilon_{II}\delta\tfrac{\gamma-\alpha}{\tilde{\phi}} + (1-\rho)(1-\epsilon_I)\epsilon_{II}\delta\tfrac{\gamma}{\tilde{\phi}}\big)\\ &+ 2\tfrac{\epsilon_{II}}{\tilde{\phi}}\delta\gamma = \tilde{\Psi}(\rho;0) + \tfrac{\epsilon_{II}}{\tilde{\phi}}\delta\big((1+\epsilon_I)(1-\rho)\gamma + 2\rho\alpha\big)\\ =& -(1+\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha) - \tfrac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)\big) + 2\tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\alpha\\ &-(1-\epsilon_I)\tfrac{\epsilon_{II}}{\tilde{\phi}}\delta(1-\rho)\alpha,\end{aligned}$$

where $\tilde{\Psi}(\rho;0)$ can be found in (B.3). One has consequently:

$$\tilde{\Psi}(\tilde{\tilde{\rho}}(\delta);\delta) = -(1+\epsilon_I)\tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\alpha + 2\tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\alpha - (1-\epsilon_I)\tfrac{\epsilon_{II}}{\tilde{\phi}}\delta(1-\tilde{\tilde{\rho}}(\delta))\alpha > 0.$$

This observation together with the fact that $\tilde{\Psi}(\rho;\delta)$ increases in ρ proves that $\tilde{\rho}(\delta) < \tilde{\tilde{\rho}}(\delta)$. Totally differentiating $\tilde{\Psi}(\rho;\delta) = 0$ reveals that $\tilde{\rho}'(\delta) < 0$.

The subcase $\rho > \tilde{\tilde{\rho}}(\delta)$: The difference in total welfare between the two possibilities now equals

$$\begin{aligned}&\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*;0) - 2\tilde{W}(\tfrac{\gamma-\alpha}{\tilde{\phi}};0) - \rho\epsilon_{II}\delta\tfrac{\gamma-\alpha-\epsilon_{II}\alpha}{\tilde{\phi}}\\ &- \big(\rho\epsilon_{II}^2\delta\tfrac{\gamma}{\tilde{\phi}} + \rho(1-\epsilon_{II})\epsilon_{II}\delta\tfrac{\gamma-\alpha}{\tilde{\phi}} + (1-\rho)(1-\epsilon_I)\epsilon_{II}\delta\tfrac{\gamma}{\tilde{\phi}}\big) + 2\rho\tfrac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)\\ =&(1-\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha) - \tfrac{\epsilon_{II}}{\tilde{\phi}}\delta\gamma\big),\end{aligned}$$

where we have used the fact that $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*;0) - 2\tilde{W}(\frac{\gamma-\alpha}{\tilde{\phi}};0) = (1-\epsilon_I)(1-\rho)\big(1-\gamma+\beta(\gamma-\alpha)\big)$. The laissez-faire condition (C.1) implies that $1-\gamma+\beta(\gamma-\alpha)-\frac{\epsilon_{II}}{\tilde{\phi}}\delta\gamma > \frac{\epsilon_{II}}{\tilde{\phi}}\beta\gamma+\beta(\gamma-\alpha) > 0$, proving that using graduated punishments is optimal if $\rho > \tilde{\tilde{\rho}}(\delta)$. So, $\hat{\tilde{\rho}}(\delta) = 1$.

The subcase $\rho \leq \tilde{\rho}(\delta)$: Subtracting $2\tilde{W}(\frac{\gamma}{\tilde{\phi}};\delta)$ from $\tilde{\mathcal{W}}(\tilde{\boldsymbol{f}}^*;\delta)$ evaluated at $\tilde{\boldsymbol{f}}^* = (0, \frac{\gamma-\alpha}{\tilde{\phi}}, \frac{\gamma}{\tilde{\phi}})$ yields

$$\tilde{\Psi}(\rho;\delta) = \tilde{\Psi}(\rho;0) + \tfrac{\epsilon_{II}}{\tilde{\phi}}\delta\big(\gamma + (1-\rho)\epsilon_I\gamma + \rho(1-\epsilon_{II})\alpha\big),$$

where $\tilde{\Psi}(\rho;0)$ is given in (B.4). Evaluating $\tilde{\Psi}(\rho;\delta)$ at $\rho = \tilde{\tilde{\rho}}(\delta)$ yields after some manipulations

$$\begin{aligned}\tilde{\Psi}(\tilde{\tilde{\rho}}(\delta);\delta) =& \epsilon_{II}(\beta+\delta)\alpha + \tfrac{\epsilon_{II}}{\tilde{\phi}}(\beta+\delta)\gamma - (1-\tilde{\tilde{\rho}}(\delta))(1-\gamma) - (1-\tilde{\tilde{\rho}}(\delta))\epsilon_{II}(\beta+\delta)\alpha\\ =& \epsilon_{II}(\beta+\delta)\big(\alpha+\tfrac{\gamma}{\tilde{\phi}}\big) - \epsilon_{II}(\beta+\delta)\alpha\frac{1-\gamma+\epsilon_{II}(\beta+\delta)\alpha}{1-\gamma+\beta(\gamma-\alpha) - \frac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)}.\end{aligned}$$

This expression is positive precisely if

$$\frac{\gamma}{\tilde{\phi}\alpha} > \frac{\epsilon_{II}(\beta+\delta)\alpha - [\beta(\gamma-\alpha) - \frac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)]}{1-\gamma+[\beta(\gamma-\alpha) - \frac{\epsilon_{II}}{\tilde{\phi}}\delta(\gamma-\alpha)]}. \tag{C.2}$$

Observe that the RHS of this inequality decreases in the term between square brackets. Note that:

$$\beta(\gamma-\alpha) - \tfrac{\epsilon_{II}}{\tilde\phi}\delta(\gamma-\alpha) > -\tfrac{\epsilon_{II}}{\tilde\phi}\delta(\gamma-\alpha) > -\tfrac{1-\gamma}{\gamma}(\gamma-\alpha),$$

where we have used (C.1) to establish the second inequality. Furthermore, (C.1) also implies that $\epsilon_{II}(\beta+\delta)\alpha < \tilde\phi\,\frac{1-\gamma}{\gamma}\alpha$. The inequality (C.2) consequently holds *a forteriori* if

$$\frac{\gamma}{\tilde\phi\alpha} > \frac{\tilde\phi\frac{1-\gamma}{\gamma}\alpha + \frac{1-\gamma}{\gamma}(\gamma-\alpha)}{(1-\gamma) - \frac{1-\gamma}{\gamma}(\gamma-\alpha)} = \frac{\tilde\phi\alpha + (\gamma-\alpha)}{\gamma - (\gamma-\alpha)} \Leftrightarrow \tfrac{\gamma}{\alpha} > \tilde\phi^2 + \tilde\phi\tfrac{\gamma}{\alpha} - \tilde\phi \Leftrightarrow \tfrac{\gamma}{\alpha} > -\tilde\phi,$$

an inequality which trivially holds. This proves that $\tilde\Psi(\tilde{\tilde\rho}(\delta);\delta) > 0$ and hence that $\tilde\rho(\delta) < \tilde{\tilde\rho}(\delta)$. Totally differentiating $\tilde\Psi(\rho;\delta) = 0$ again reveals that $\tilde\rho'(\delta) < 0$.

The subcase $\rho > \tilde{\tilde\rho}(\delta)$: The relevant welfare difference is

$$\begin{aligned}
&\tilde{\mathcal{W}}(0,\tfrac{\gamma-\alpha}{\tilde\phi},\tfrac{\gamma}{\tilde\phi};0) - 2\tilde W(\tfrac{\gamma-\alpha}{\tilde\phi};0) - \tfrac{\epsilon_{II}}{\tilde\phi}\delta\big(\rho\epsilon_{II}\gamma + \rho(1-\epsilon_{II})(\gamma-\alpha) + (1-\rho)(1-\epsilon_I)\gamma\big)\\
&+ 2\rho\tfrac{\epsilon_{II}}{\tilde\phi}\delta(\gamma-\alpha)\\
=&\tilde{\mathcal{W}}(0,\tfrac{\gamma-\alpha}{\tilde\phi},\tfrac{\gamma}{\tilde\phi};0) - 2\tilde W(\tfrac{\gamma-\alpha}{\tilde\phi};0) + \tfrac{\epsilon_{II}}{\tilde\phi}\delta\big(\rho(\gamma-\alpha-\epsilon_{II}\alpha) - (1-\rho)(1-\epsilon_I)\gamma\big).
\end{aligned}$$

Because $\gamma-\alpha-\epsilon_{II}\alpha < 0$, this difference is less than $\tilde{\mathcal{W}}(0,\frac{\gamma-\alpha}{\tilde\phi},\frac{\gamma}{\tilde\phi};0) - 2\tilde W(\frac{\gamma-\alpha}{\tilde\phi};0)$. Since $\tilde\Psi(\tilde{\tilde\rho}(\delta);\delta) > 0$, the planner does employ graduated punishments as long as ρ does not exceed $\tilde{\tilde\rho}(\delta)$ by 'too much'. However, because the welfare difference decreases in δ, $\hat{\tilde\rho}(\delta)$ also decreases in δ. ■

Appendix D: Limited Recall

Proof of Proposition 6

To be able to determine whether using graduated punishments is optimal, we need to know the long run ($t \to \infty$) composition of the population, $q = 1 - \hat q$ as well as μ and $\hat\mu = \rho - \mu$, if graduated punishments are indeed used.

As before, when looking at graduated punishment schemes it suffices to consider situations in which low-cost agents always contribute ($\delta_L = \hat\delta_L = 1$), whereas high-cost agents only contribute when in $\hat q$ ($\delta_H = 0$, $\hat\delta_H = 1$).

The 'flow equation' for the fraction q with limited recall reads

$$q = 1 - \zeta + \zeta q(\tfrac{1+\phi}{2}\tfrac{\mu}{q} + \tfrac{1-\phi}{2}(1-\tfrac{\mu}{q})) + \zeta(1-q)\tfrac{1+\phi}{2} = 1 - \zeta\tfrac{1-\phi}{2} - \zeta\phi q + \zeta\phi\mu. \qquad \text{(D.1)}$$

The difference between (D.1) and its counterpart of the perfect recall setting (see (7)) is that the latter does not contain a term pertaining to agents returning from $\hat q$ to q (the term $\zeta(1-q)\frac{1+\phi}{2}$ in (D.1)).

If the planner has only limited recall, then the 'flow equation' for the fraction of the population that consists of low-cost agents residing in q is:

$$\mu = (1-\zeta)\rho + \zeta q \tfrac{1+\phi}{2}\tfrac{\mu}{q} + \zeta(1-q)\tfrac{1+\phi}{2}\tfrac{\rho-\mu}{1-q} = \rho(1-\zeta\tfrac{1-\phi}{2}). \tag{D.2}$$

Substituting this expression into (D.1) results in

$$q = \frac{(1-\zeta\frac{1-\phi}{2})(1+\rho\zeta\phi)}{1+\zeta\phi}. \tag{D.3}$$

In order to derive the optimal punishments f and $\hat{f}$ that prevail should the planner opt for graduated punishments, we need to know the incentive compatibility constraints that ensure that $(\delta_L, \hat{\delta}_L) = (1,1)$ ($(\delta_H, \hat{\delta}_H) = (0,1)$) is indeed optimal from the point of view of the low-cost agents (high-cost agents). These constraints follow from the following four Bellman equations that govern agents' behaviour:

- Bellman equation for low-cost agents in q:

$$C_L = \min_{\delta\in\{0,1\}} \Big[\delta\big(\gamma-\alpha+\tfrac{1-\phi}{2}(f+\zeta\hat{C}_L)+\tfrac{1+\phi}{2}\zeta C_L\big)+(1-\delta)\big(\tfrac{1+\phi}{2}(f+\zeta\hat{C}_L)+\tfrac{1-\phi}{2}\zeta C_L\big)\Big].$$

- Bellman equation for low-cost agents in $\hat{q}$:

$$\hat{C}_L = \min_{\delta\in\{0,1\}} \Big[\delta\big(\gamma-\alpha+\tfrac{1-\phi}{2}(\hat{f}+\zeta\hat{C}_L)+\tfrac{1+\phi}{2}\zeta C_L\big)+(1-\delta)\big(\tfrac{1+\phi}{2}(\hat{f}+\zeta\hat{C}_L)+\tfrac{1-\phi}{2}\zeta C_L\big)\Big].$$

- Bellman equation for high-cost agents in q:

$$C_H = \min_{\delta\in\{0,1\}} \Big[\delta\big(\gamma+\tfrac{1-\phi}{2}(f+\zeta\hat{C}_H)+\tfrac{1+\phi}{2}\zeta C_H\big)+(1-\delta)\big(\tfrac{1+\phi}{2}(f+\zeta\hat{C}_H)+\tfrac{1-\phi}{2}\zeta C_H\big)\Big].$$

- Bellman equation for high-cost agents in $\hat{q}$:

$$\hat{C}_H = \min_{\delta\in\{0,1\}} \Big[\delta\big(\gamma+\tfrac{1-\phi}{2}(\hat{f}+\zeta\hat{C}_H)+\tfrac{1+\phi}{2}\zeta C_H\big)+(1-\delta)\big(\tfrac{1+\phi}{2}(\hat{f}+\zeta\hat{C}_H)+\tfrac{1-\phi}{2}\zeta C_H\big)\Big].$$

Optimality of $(\delta_L, \hat{\delta}_L, \delta_H, \hat{\delta}_H) = (1,1,0,1)$ requires:

$$\begin{aligned}
&\gamma-\alpha \le \phi f + \phi\zeta(\hat{C}_L - C_L), \quad \gamma-\alpha \le \phi\hat{f} + \phi\zeta(\hat{C}_L - C_L),\\
&\gamma > \phi f + \phi\zeta(\hat{C}_H - C_H), \quad \gamma \le \phi\hat{f} + \phi\zeta(\hat{C}_H - C_H).
\end{aligned}$$

Clearly, as long as $\hat{f} > f$ the second and third inequality contain slack and the planner opts for punishments that solve

$$\phi f \ge \gamma - \alpha - \phi\zeta(\hat{C}_L - C_L), \quad \phi\hat{f} \ge \gamma - \phi\zeta(\hat{C}_H - C_H). \tag{D.4}$$

Combining the Bellman equations with the fact that $(\delta_L^*, \hat{\delta}_L^*, \delta_H^*, \hat{\delta}_H^*) = (1,1,0,1)$ yields after some rearranging

$$\hat{C}_L - C_L = \tfrac{1-\phi}{2}(\hat{f} - f), \quad \hat{C}_H - C_H = \frac{\gamma + \frac{1-\phi}{2}\hat{f} - \frac{1+\phi}{2} f}{1+\zeta\phi}. \tag{D.5}$$

Plugging these expressions into (D.4) with equality signs replacing the inequality signs results after some straightforward algebra in:

$$\phi f = \gamma - \alpha - \frac{\zeta \frac{1-\phi}{2}}{1+\zeta\phi}\alpha, \quad \phi\hat{f} = \gamma - \frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}\alpha = \gamma - \alpha + \frac{1-\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha. \tag{D.6}$$

Of course, just like in the model with perfect recall, the expression for the punishment for first time offenders can be negative. However, unlike in the model with perfect recall, simply setting $\phi f = 0$ and using the expression for $\phi\hat{f}$ given above is not an option. The reason is that the resulting pair of punishments violates the incentive compatibility constraint for high-cost agents in $\hat{q}$. This is caused by the fact that, in contrast to the optimal punishment for repeat offenders in the perfect recall setting, $\hat{f}$ as given in (D.6) depends negatively on α. Because $\hat{f}$ decreases in α, the difference between $\hat{f}$ and f becomes smaller as α increases beyond the value above which $\gamma - \alpha - \frac{\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha < 0$, i.e. above which $f = 0$. The incentives for high-cost agents to contribute when in $\hat{q}$ are consequently weakened via two channels. Clearly, a lower punishment $\hat{f}$ weakens the incentives to contribute. On top of that, a smaller difference between $\hat{f}$ and f decreases the relative attractiveness of being in q, making agents less inclined to try to move back to q by contributing.

Evaluating $\hat{C}_H - C_H$ (see (D.5)) at $f = 0$ and combining the result with the constraint on $\hat{f}$ given in (D.4) yields $(1 + \zeta\frac{1+\phi}{2})\phi\hat{f} \geq \gamma$. We conclude that

$$\phi f^* = \begin{cases} \gamma - \alpha - \frac{\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha & \text{if } \gamma \geq \bar{\gamma} \\ 0 & \text{if } \gamma < \bar{\gamma} \end{cases}, \quad \phi\hat{f}^* = \begin{cases} \gamma - \alpha + \frac{1-\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha & \text{if } \gamma \geq \bar{\gamma} \\ \frac{1}{1+\zeta\frac{1+\phi}{2}}\gamma & \text{if } \gamma < \bar{\gamma} \end{cases}, \tag{D.7}$$

where $\bar{\gamma} := \alpha + \frac{\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha$ increases in ζ.

We can now calculate the total welfare $\Omega(f^*, \hat{f}^*)$ generated in one period if the planner, having limited recall, employs graduated punishments. The per-period welfare ignoring the costs of administering punishments reads

$$\rho(1-\gamma+\alpha) + (1 - \tfrac{\mu}{\hat{q}})\hat{q}(1-\gamma) = \rho(1-\gamma+\alpha) + \big[(1-q) - (\rho-\mu)\big](1-\gamma).$$

Note that:

$$(1-q) - (\rho - \mu) = \frac{(1-\rho)\zeta\phi + (1+\rho\zeta\phi)\zeta\frac{1-\phi}{2}}{1+\zeta\phi} - \frac{\rho\zeta\frac{1-\phi}{2}(1+\zeta\phi)}{1+\zeta\phi} = \frac{(1-\rho)\zeta\frac{1+\phi}{2}}{1+\zeta\phi}.$$

So, per-period welfare gross of punishment costs equals

$$\rho(1-\gamma+\alpha)+\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}(1-\rho)(1-\gamma). \tag{D.8}$$

Let us now determine the per-period costs of administering punishments. These costs amount to β times

$$q\big(\tfrac{1-\phi}{2}\tfrac{\mu}{q}+\tfrac{1+\phi}{2}(1-\tfrac{\mu}{q})\big)f^*+(1-q)\tfrac{1-\phi}{2}\hat{f}^*=(\tfrac{1+\phi}{2}q-\phi\mu)f^*+(1-q)\tfrac{1-\phi}{2}\hat{f}^*. \tag{D.9}$$

where we have used that in an equilibrium with graduated punishments a fraction $\frac{1-\phi}{2}$ of the low-cost agents in q, a fraction $\frac{1+\phi}{2}$ of the high-cost agents in q, and a fraction $\frac{1-\phi}{2}$ of the agents in $\hat{q}$ are punished. We have to treat the cases $\gamma\geq\bar{\gamma}$ and $\gamma<\bar{\gamma}$ separately:

The case $\gamma\geq\bar{\gamma}$: In this case the aggregate punishment equals

$$\begin{aligned}(\tfrac{1+\phi}{2}q-\phi\mu)f^*+(1-q)\tfrac{1-\phi}{2}\hat{f}^*=&(q-\mu)(\gamma-\alpha)+\tfrac{1-\phi}{2\phi}(\gamma-\alpha)-(\tfrac{1+\phi}{2\phi}q-\mu)\frac{\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha\\&+\tfrac{1-\phi}{2\phi}(1-q)\frac{1-\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\alpha\\=:&(q-\mu+\tfrac{1-\phi}{2\phi})(\gamma-\alpha)+\chi\alpha.\end{aligned}$$

Observe that:

$$\begin{aligned}\chi=&\frac{-\zeta\frac{1-\phi}{2}\frac{1+\phi}{2\phi}-\frac{1-\phi}{2\phi}+\zeta\frac{1-\phi}{2}\frac{1-\phi}{2\phi}}{1+\zeta\phi}q+\frac{\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\rho(1-\zeta\tfrac{1-\phi}{2})+\frac{1-\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\tfrac{1-\phi}{2\phi}\\=&-\tfrac{1-\phi}{2\phi}q+\tfrac{1-\phi}{2\phi}(1-\zeta\tfrac{1-\phi}{2})\frac{1+\rho\zeta\phi}{1+\zeta\phi}=0.\end{aligned}$$

Furthermore:

$$q-\mu+\tfrac{1-\phi}{2\phi}=\big(1-\zeta\tfrac{1-\phi}{2}\big)\tfrac{1-\rho}{1+\zeta\phi}+\tfrac{1-\phi}{2\phi}=\frac{\frac{1+\phi}{2\phi}-\rho(1-\zeta\frac{1-\phi}{2})}{1+\zeta\phi}.$$

The social costs associated with administering punishments thus amount to

$$\beta\frac{\frac{1+\phi}{2\phi}-\rho+\zeta\rho\frac{1-\phi}{2}}{1+\zeta\phi}(\gamma-\alpha).$$

Subtracting this from (D.8) yields the total per-period welfare:

$$\Omega(f^*,\hat{f}^*)=\rho(1-\gamma+\alpha)+\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}(1-\rho)(1-\gamma)-\beta\frac{\frac{1+\phi}{2\phi}-\rho(1-\zeta\frac{1-\phi}{2})}{1+\zeta\phi}(\gamma-\alpha),\ \ \gamma\geq\bar{\gamma}. \tag{D.10}$$

To assess the difference in per-period welfare between using the single punishment f_0^* and using graduated punishments, we consider the subcases $\rho\leq\bar{\rho}$ and $\rho>\bar{\rho}$ in turn.

The subcase $\rho \leq \bar{\rho}$: Subtracting $W(\frac{\gamma}{\phi})$ (see (A.1)) from $\Omega(f^*, \hat{f}^*)$ yields

$$\begin{aligned}\Omega(f^*, \hat{f}^*) - W(\tfrac{\gamma}{\phi}) =& \Big(\frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi} - 1\Big)(1-\rho)(1-\gamma) - \beta\frac{\frac{1+\phi}{2\phi} - \rho(1-\zeta^{\frac{1-\phi}{2}})}{1+\zeta\phi}(\gamma-\alpha) + \beta\tfrac{1-\phi}{2\phi}\gamma \\ =& -(1-\rho)\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi} \times \big[(1-\gamma) + \beta(\gamma-\alpha)\big] + \beta\tfrac{1-\phi}{2\phi}\alpha.\end{aligned}$$

Evaluating this expression at $\rho = \bar{\rho}$ results in

$$\Omega(f^*, \hat{f}^*) - W(\tfrac{\gamma}{\phi})\big|_{\rho=\bar{\rho}} = -\tfrac{1-\phi}{2\phi}\beta\alpha\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi} + \tfrac{1-\phi}{2\phi}\beta\alpha = \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi} \times \tfrac{1-\phi}{2\phi}\beta\alpha > 0,$$

where we have used the fact that $1 - \bar{\rho} = \big((1-\gamma) + \beta(\gamma-\alpha)\big)^{-1}\frac{1-\phi}{2\phi}\beta\alpha$. One easily verifies that $\Omega(f^*, \hat{f}^*) - W(\frac{\gamma}{\phi})$ is strictly increasing in ρ and we consequently conclude that $\Omega(f^*, \hat{f}^*) > W(\frac{\gamma}{\phi})$ if $\rho \in (\check{\rho}_1, \bar{\rho}]$ for some $\check{\rho}_1 < \bar{\rho}$.

The subcase $\rho > \bar{\rho}$: Subtracting $W(\frac{\gamma-\alpha}{\phi})$ (see (A.1)) from $\Omega(f^*, \hat{f}^*)$ and some rearranging suffices to show that using graduated punishments always dominates using a single punishment in this case:

$$\begin{aligned}\Omega(f^*, \hat{f}^*) - W(\tfrac{\gamma-\alpha}{\phi}) =& \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi}(1-\rho)(1-\gamma) - \beta\frac{\frac{1+\phi}{2\phi} - \rho(1-\zeta^{\frac{1-\phi}{2}})}{1+\zeta\phi}(\gamma-\alpha) \\ &+ (\tfrac{1+\phi}{2\phi} - \rho)\beta(\gamma-\alpha) \\ =& \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi}(1-\rho)(1-\gamma) + \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi}(1-\rho)\beta(\gamma-\alpha) > 0.\end{aligned}$$

The case $\gamma < \bar{\gamma}$: Since $f^* = 0$ and $\hat{f}^* = \frac{\gamma}{\phi(1+\zeta^{\frac{1+\phi}{2}})}$, the aggregate punishment now equals

$$(1-q)\frac{\frac{1-\phi}{2\phi}}{1+\zeta^{\frac{1+\phi}{2}}}\gamma = \frac{(1-\rho)(1-\zeta^{\frac{1-\phi}{2}}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\gamma.$$

The per-period welfare $\Omega(f^*, \hat{f}^*)$ consequently reads

$$\begin{aligned}\Omega(f^*, \hat{f}^*) =& \rho(1-\gamma+\alpha) + \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi}(1-\rho)(1-\gamma) \\ &- \frac{(1-\rho)(1-\zeta^{\frac{1-\phi}{2}}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\beta\gamma, \quad \gamma < \bar{\gamma}.\end{aligned} \tag{D.11}$$

The subcase $\rho \le \bar{\rho}$: The relevant welfare difference is

$$\begin{aligned}\Omega(f^*,\hat{f}^*) - W(\tfrac{\gamma}{\phi}) =& \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi}(1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta^{\frac{1-\phi}{2}}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\beta\gamma \\ & - (1-\rho)(1-\gamma) + \tfrac{1-\phi}{2\phi}\beta\gamma \\ =& -\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi}(1-\rho)(1-\gamma) - (1-\rho)\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi}\times\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\beta\gamma \\ & + \frac{1+\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}\times\tfrac{1-\phi}{2\phi}\beta\gamma.\end{aligned}$$

Since this welfare difference is increasing in ρ, it suffices to show that this difference evaluated at $\rho = \bar{\rho}$ is positive, i.e. that $\Delta :=$

$$\tfrac{1-\phi}{2\phi}\beta\times\left[-\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi}\times\frac{(1-\gamma)\alpha}{1-\gamma+\beta(\gamma-\alpha)} - \frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi}\times\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\frac{\beta\gamma\alpha}{1-\gamma+\beta(\gamma-\alpha)} + \frac{1+\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}\gamma\right]$$

is positive. Condition 1 implies that

$$\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\beta\gamma < \frac{\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}(1-\gamma).$$

Therefore:

$$\begin{aligned}\big(\tfrac{1-\phi}{2\phi}\beta\big)^{-1}\Delta >& -\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi}\times\Big(1+\frac{\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}\Big)\frac{(1-\gamma)\alpha}{1-\gamma+\beta(\gamma-\alpha)} + \frac{1+\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}\gamma \\ >& -\frac{1-\zeta^{\frac{1-\phi}{2}}}{1+\zeta\phi}\times\Big(1+\frac{\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}\Big)\gamma + \frac{1+\zeta\phi}{1+\zeta^{\frac{1+\phi}{2}}}\gamma \\ =& \frac{(\zeta^{\frac{1+\phi}{2}})^2}{(1+\zeta\phi)(1+\zeta^{\frac{1+\phi}{2}})}\gamma > 0.\end{aligned}$$

We conclude that $\Omega(f^*,\hat{f}^*) > W(\frac{\gamma}{\phi})$ if $\rho \in (\check{\rho}_1, \bar{\rho}]$ for some $\check{\rho}_1 < \bar{\rho}$.

The subcase $\rho > \bar{\rho}$: Subtracting $W(\frac{\gamma-\alpha}{\phi})$ from (D.11) yields the relevant welfare difference:

$$\Psi(\rho) = \frac{\zeta^{\frac{1+\phi}{2}}}{1+\zeta\phi}(1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta^{\frac{1-\phi}{2}}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\frac{\zeta^{\frac{1-\phi}{2}}}{1+\zeta^{\frac{1+\phi}{2}}}\beta\gamma + \big(\tfrac{1+\phi}{2\phi} - \rho\big)\beta(\gamma-\alpha).$$

Analysis of $\Psi'(\rho)$ conveys that this difference decreases monotonically in ρ:

$$\begin{aligned}\Psi'(\rho) =& -\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}(1-\gamma) + \frac{1-\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\frac{\zeta\frac{1-\phi}{2}}{1+\zeta\frac{1+\phi}{2}}\beta\gamma - \beta(\gamma-\alpha)\\ <& -\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}(1-\gamma) + \frac{1-\zeta\frac{1-\phi}{2}}{1+\zeta\phi}\frac{\zeta\phi}{1+\zeta\frac{1+\phi}{2}}(1-\gamma) - \beta(\gamma-\alpha)\\ =& -\frac{\zeta\frac{1-\phi}{2} + \zeta^2(\frac{1+\phi}{2})^2 + \zeta^2\phi\frac{1-\phi}{2}}{(1+\zeta\phi)(1+\zeta\frac{1+\phi}{2})} - \beta(\gamma-\alpha) < 0,\end{aligned}$$

where the first inequality follows from Condition 1. Because $\Psi(\bar{\rho}) = \Delta > 0$, we conclude that using graduated punishments is optimal for $\rho \in (\bar{\rho}, \hat{\rho}_1]$ for some $\hat{\rho}_1 > \bar{\rho}$. Observe that

$$\begin{aligned}\Psi(1) =& -\frac{\zeta\frac{1-\phi}{2}}{1+\zeta\frac{1+\phi}{2}}\tfrac{1-\phi}{2\phi}\beta\gamma + \tfrac{1-\phi}{2\phi}\beta(\gamma-\alpha) = \tfrac{1-\phi}{2\phi}\beta\Big[\frac{1+\zeta\phi}{1+\zeta\frac{1+\phi}{2}}\gamma - \alpha\Big]\\ <& \tfrac{1-\phi}{2\phi}\beta\Big[\frac{1+\zeta\phi}{1+\zeta\frac{1+\phi}{2}}\bar{\gamma} - \alpha\Big] = 0.\end{aligned}$$

So, $\hat{\rho}_1 < 1$ if $\gamma < \bar{\gamma}$. Since $\bar{\gamma}$ increases in ζ, this observation implies that the planner resorts to the uniform punishment scheme if ρ and ζ are both sufficiently large.

We have now shown that using graduated punishments improves welfare compared to using a uniform punishment if the planner has limited recall as long as $\rho \in [\check{\rho}_1, \hat{\rho}_1]$. We next compare the per-period welfare generated if the planner has perfect recall with the per-period welfare generated if the planner has only limited recall. To that end it is useful to denote the per-period welfare under perfect recall by Ω_∞ and the per-period welfare under limited recall by Ω_1. For the sake of convenience we first repeat these welfare expressions below:

$$\Omega_\infty = \begin{cases}\rho(1-\gamma+\alpha) + \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}(1-\rho)(1-\gamma) - \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1-\phi}{2}}\frac{(1-\rho)(1-\zeta)+\frac{1-\phi}{2\phi}(1-\zeta\frac{1-\phi}{2})}{1-\zeta\frac{1+\phi}{2}}\beta\gamma & \text{if } \gamma < \check{\gamma}\\ \rho(1-\gamma+\alpha) + \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}(1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta)+\frac{1-\phi}{2\phi}(1-\zeta\frac{1-\phi}{2})}{1-\zeta\frac{1-\phi}{2}}\beta(\gamma-\alpha) & \text{if } \gamma \geq \check{\gamma},\end{cases} \tag{D.12}$$

$$\Omega_1 = \begin{cases}\rho(1-\gamma+\alpha) + \frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}(1-\rho)(1-\gamma) - \frac{\zeta\frac{1-\phi}{2}}{1+\zeta\frac{1+\phi}{2}}\frac{(1-\rho)(1-\zeta\frac{1-\phi}{2})+\frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\beta\gamma & \text{if } \gamma < \bar{\gamma}\\ \rho(1-\gamma+\alpha) + \frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi}(1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta\frac{1-\phi}{2})+\frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\beta(\gamma-\alpha) & \text{if } \gamma \geq \bar{\gamma}.\end{cases} \tag{D.13}$$

Straightforward calculations reveal that $\bar{\gamma} < \check{\gamma}$. We thus have to consider three (disjoint) cases: $\gamma \in (\alpha, \bar{\gamma})$, $\gamma \in [\bar{\gamma}, \check{\gamma})$, and $\gamma \in [\check{\gamma}, 1)$.

The case $\gamma \in (\alpha, \bar{\gamma})$: The welfare difference now reads:

$$\begin{aligned}\Omega_1 - \Omega_\infty =& \Big[\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi} - \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}\Big](1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta\frac{1-\phi}{2}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\frac{\zeta\frac{1-\phi}{2}}{1+\zeta\frac{1+\phi}{2}}\beta\gamma\\ &+ \frac{(1-\rho)(1-\zeta) + \frac{1-\phi}{2\phi}(1-\zeta\frac{1-\phi}{2})}{1-\zeta\frac{1-\phi}{2}}\frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1+\phi}{2}}\beta\gamma.\end{aligned}$$

The sign of this expression cannot be assessed analytically. Numerical computations indicate that $\Omega_1 - \Omega_\infty$ is negative if $\gamma \in (\alpha, \bar{\gamma})$.

The case $\gamma \in [\bar{\gamma}, \check{\gamma})$: The welfare difference is:

$$\Omega_1 - \Omega_\infty = \Big[\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi} - \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}\Big](1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta\frac{1-\phi}{2}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\beta(\gamma-\alpha) + \frac{\zeta\frac{1-\phi}{2}}{1-\zeta\frac{1-\phi}{2}} \frac{(1-\rho)(1-\zeta) + \frac{1-\phi}{2\phi}(1-\zeta\frac{1-\phi}{2})}{1-\zeta\frac{1+\phi}{2}}\beta\gamma. \quad \text{(D.14)}$$

Again, one has to resort to numerical methods to assess the sign of the welfare difference. It appears that $\Omega_1 - \Omega_\infty$ is also negative if $\gamma \in [\bar{\gamma}, \check{\gamma})$.

The case $\gamma \in [\check{\gamma}, 1)$: With γ large the welfare difference equals

$$\begin{aligned}\Omega_1 - \Omega_\infty =& \Big[\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi} - \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}\Big](1-\rho)(1-\gamma) - \frac{(1-\rho)(1-\zeta\frac{1-\phi}{2}) + \frac{1-\phi}{2\phi}(1+\zeta\phi)}{1+\zeta\phi}\beta(\gamma-\alpha) \\ &+ \frac{(1-\rho)(1-\zeta) + \frac{1-\phi}{2\phi}(1-\zeta\frac{1-\phi}{2})}{1-\zeta\frac{1-\phi}{2}}\beta(\gamma-\alpha) \\ =& \Big[\frac{\zeta\frac{1+\phi}{2}}{1+\zeta\phi} - \frac{\zeta\frac{1+\phi}{2}}{1-\zeta\frac{1-\phi}{2}}\Big](1-\rho)(1-\gamma) - \frac{\zeta^2(\frac{1+\phi}{2})^2}{(1+\zeta\phi)(1-\zeta\frac{1-\phi}{2})}(1-\rho)\beta(\gamma-\alpha) < 0,\end{aligned}$$

implying that there exists a $\tilde{\gamma} < \check{\gamma}$ such that $\Omega_\infty > \Omega_1$ as long as $\gamma > \tilde{\gamma}$. This observation completes the proof. ∎

References

Abreu, D., D. Bernheim, and A. Dixit (2005): "Self-Enforcing Cooperation with Graduated Punishments," Working paper, Princeton University.

Agrawal, A. (2003): "Sustainable Governance of Common-Pool Resources: Context, Methods, and Politics," *Annual Review of Anthropology*, 32, 243–262.

Chu, C. Y. C., S.-C. Hu, and T.-Y. Huang (2000): "Punishing Repeat Offenders more Severely," *International Review of Law and Economics*, 20, 127–140.

Dana, D. A. (2001): "Rethinking the Puzzle of Escalating Penalties for Repeat Offenders," *Yale Law Journal*, 110, 733–783.

Ellickson, R. (1991): *Order without Law: How Neighbors Settle Disputes.* Harvard University Press, Cambridge, MA.

Green, E., and R. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87–100.

Harrington, W. (1988): "Enforcement Leverage when Penalties are Restricted," *Journal of Public Economics*, 37, 29–53.

Miceli, T. J., and C. Bucci (2005): "A Simple Theory of Increasing Penalties for Repeat Offenders," *Review of Law and Economics*, 1.

Mungan, M. C. (2010): "Repeat Offenders: If they learn, we punish them more severely," *International Review of Law and Economics*, 30, 173–177.

Ostrom, E. (1990): *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge, UK.

——— (2000): "Collective Action and the Evolution of Social Norms," *Journal of Economic Perspectives*, 14, 137–158.

Polinsky, A. M., and D. L. Rubinfeld (1991): "A Model of Optimal Fines for Repeat Offenders," *Journal of Public Economics*, 46, 291–306.

Polinsky, A. M., and S. Shavell (1998): "On Offense History and the Theory of Deterrence," *International Review of Law and Economics*, 18, 305–324.

Rousseau, S. (2009): "The Use of Warnings in the Presence of Errors," *International Review of Law and Economics*, 29, 191–201.

Rubinstein, A. (1979): "An Optimal Policy for Offenses that May Have Been Committed by Accident," in *Applied Game Theory*, ed. by S. Brams, A. Schotter, and G. Schwodiauer, pp. 406–413. Physica-Verlag, Würzburg.

——— (1980): "On an Anomaly of the Deterrent Effect of Punishment," *Economic Letters*, 6, 89–94.

Stigler, G. J. (1970): "The Optimum Enforcement of Laws," *Journal of Political Economy*, 78, 526–536.

Wade, R. (1994): *Village Republics: Economic Conditions for Collective Action in South India*. ICS Press, Oakland.