

DISCUSSION PAPER SERIES

No. 10249

**OVERDETERRENCE OF REPEAT
OFFENDERS WHEN PENALTIES FOR
FIRST-TIME OFFENDERS ARE RESTRICTED**

Daniel Müller and Patrick W Schmitz

INDUSTRIAL ORGANIZATION



Centre for Economic Policy Research

OVERDETERRENCE OF REPEAT OFFENDERS WHEN PENALTIES FOR FIRST-TIME OFFENDERS ARE RESTRICTED

Daniel Müller and Patrick W Schmitz

Discussion Paper No. 10249
November 2014
Submitted 08 November 2014

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Daniel Müller and Patrick W Schmitz

OVERDETERRENCE OF REPEAT OFFENDERS WHEN PENALTIES FOR FIRST-TIME OFFENDERS ARE RESTRICTED

Abstract

When penalties for first-time offenders are restricted, it is typically optimal for the lawmaker to overdeter repeat offenders. First-time offenders are then deterred not only by the (restricted) fine for a first offense, but also by the prospect of a large fine for a subsequent offense. Now suppose the restriction on penalties for first-time offenders is relaxed; i.e., larger fines for a first offense become enforceable. Should overdeterrence of repeat offenders now be reduced? We show that this is the case only if the original restriction was not very strong. Otherwise, overdeterrence of repeat offenders should actually be further amplified.

JEL Classification: D82, H23, K14, K42 and L51

Keywords: incentives, law enforcement, limited liability, penalties and repeat offenders

Daniel Müller daniel.mueller@uni-bonn.de
University of Bonn

Patrick W Schmitz patrick.schmitz@uni-koeln.de
University of Cologne and CEPR

1. Introduction

The law often sanctions repeat offenders more severely than first-time offenders.¹ By now, the literature has provided various justifications for the fact that the sanction imposed on an offender depends on whether he was convicted previously.² Some authors have argued that a record of prior offenses provides information about characteristics of the offender (e.g., a higher-than-average propensity to commit crimes).³ Yet, making sanctions depend on offense history may be advantageous even when individuals are ex ante identical such that there are no characteristics to be learned about. As has been emphasized by Shavell (2004, p. 529), when “detection of a violation implies not only an immediate sanction, but also a higher sanction for a future violation, an individual will be deterred more from committing a violation presently.” In this paper, we follow Shavell’s (2004) insight and further explore how penalties for repeat offenders should be designed when we take their effect on the deterrence of first-time offenders into account.

Specifically, suppose that there is an exogenously given restriction on the penalties for first-time offenders; i.e., there is an upper limit l which a fine for a first-time offender must not exceed, while there is no (binding) restriction on the fine that a repeat offender has to pay.⁴ Suppose that in each of two periods a potential offender engages in an activity that may cause a harm h . For simplicity, assume that whenever a harm is caused, the offender is convicted to pay a fine. Of course, if the harm is smaller than the maximum fine l , then

¹For example, with regard to civil penalties in the U.S.A., Shavell (2004, ch. 22) points out that for certain violations of the Occupational Safety and Health Act there is a maximum fine of \$7,000 for a first offense, while a repeat offender may be fined \$70,000.

²See Polinsky and Shavell (2007) and Miceli (2013) for recent literature reviews.

³See e.g. Rubinstein (1980), Polinsky and Rubinfeld (1991), and Chu et al. (2000).

⁴There may be various reasons why society does not accept larger penalties for first-time offenders. For instance, Stigler (1970, p. 528) has pointed out that a “first-time offender may have committed the offense almost accidentally” and Polinsky and Shavell (1998, p. 313) argue that “considerations of fairness might constrain the sanction imposed on first-time offenders but not on repeat offenders.”

by setting the fine equal to the harm h in both periods the negative externality of the activity is internalized and the first-best (i.e., socially optimal) activity level is implemented. Yet, if $l < h$, then first-time offenders in the first period are underdeterred, given that the fine for repeat offenders is set equal to the harm h . As a consequence, in general it will be optimal for the lawmaker to set the fine for repeat offenders larger than h . While in the second period overdeterrence of repeat offenders is ex post inefficient, the advantage of such a policy is that the large fine for a second offense has a spillover effect on the first period.⁵ Individuals in the first period are deterred not only by the (restricted) fine that they have to pay when they cause harm as a first-time offender, but also by the prospect of having to pay a large fine as a repeat offender in the second period.

Let us now explore what happens when the restriction l that society has put on the admissible fines for first-time offenders is relaxed. At first sight, one might guess that the lawmaker should reduce the ex post inefficient overdeterrence of repeat offenders, because the deterrence of first-time offenders can now be improved by a larger penalty for first offenses. Yet, it turns out that this is the case only if initially the upper limit l was not very restrictive. If l was very small, then an increase in l will actually prompt the lawmaker to further increase the fine for repeat offenders; i.e., overdeterrence of second offenses will be further aggravated. Intuitively, when a first-time offender in the second period can be more severely punished and the fine for a repeat offender does not go up, then the indirect deterrence effect on the first period is reduced, which the lawmaker may prefer to counterbalance by further increasing the fine for repeat offenders (provided that the direct punishment in the first period is still rather small,

⁵Note that related spillover-of-incentives effects also occur in sequential moral hazard models with limited liability. See e.g. Schmitz (2005) and Ohlendorf and Schmitz (2012), who show how second-period rents may act as carrot and stick for an agent's first-period effort choice, such that optimal contracts exhibit memory even though the periods are technologically independent. Recent papers that exploit related effects include e.g. Kräkel and Schöttner (2010), Tsai and Kung (2011), Chen and Chiu (2013), and Pi (2014).

which is the case for relatively small values of the upper limit l).

To the best of our knowledge, in this paper we present the first analysis of how fines for repeat offenders should be optimally adapted when an exogenous restriction on fines for first-time offenders is relaxed. By considering a two-period framework, we build on Polinsky and Rubinfeld's (1991) seminal work on repeat offenders. In contrast to their paper and following Polinsky and Shavell (1998), we make the standard assumption that social welfare is given by the private benefit generated by the risky activity minus the harm caused (i.e., we do not consider illicit gains that are not included in social welfare). Throughout, we assume that it is not possible to directly regulate the activity level (i.e., the activity is a hidden action), such that the lawmaker who wants to maximize the expected social welfare can implement the desired activity level only by setting suitable penalties which have to be paid ex post when a harm has been caused.⁶

The remainder of this paper is organized as follows. In Section 2, the formal framework is presented and in Section 3, the model is analyzed. Concluding remarks follow in Section 4. Finally, the proofs of our results have been relegated to the Appendix.

2. The model

In each of two consecutive periods, $t = 1, 2$, a risk-neutral individual (he) chooses the level $a_t \in [0, 1]$ of a potentially harmful activity. With probability a_t the individual causes a harm $h > 0$ in period t . If a harm is caused, the individual is convicted to pay a fine. The individual's private benefit from pursuing the activity is $b(a) : [0, 1] \rightarrow \mathbb{R}$ with $b'(a) > 0$, $b''(a) < 0$, $\lim_{a \rightarrow 0} b'(a) = \infty$ and $\lim_{a \rightarrow 1} b'(a) = 0$.

Let $y \in \{0, 1\}$ denote the individual's offense history at the beginning of period $t = 2$. If $y = 1$, then the individual is pre-convicted because he caused

⁶See Shavell (1984) and Schmitz (2000) for models in which both ex ante and ex post interventions are feasible in order to regulate activities that are potentially risky for the environment or third parties.

a harm h in period $t = 1$. If $y = 0$, the individual has a clean slate. In period $t = 0$ the lawmaker (she) commits to a constitution, in particular stipulating the (finite) fine $F_t \geq 0$ to be paid by an individual in period t if he causes a harm. While the fine in period $t = 2$ may condition on the individual's offense history, $F_2 = F_2(y)$, the lawmaker is not allowed to discriminate according to whether a first offense was committed in $t = 1$ or $t = 2$.⁷ Thus, $F_1 = F_2(0) \equiv F^0$ and $F_2(1) \equiv F^1$. Moreover, while there is no (binding) restriction regarding the punishment F^1 of a repeat offender, by social convention punishment of a first-time offender must not be overly drastic, $F^0 \leq l$, where $l \geq 0$.

If the activity level in each period were directly enforceable, then the lawmaker would implement the activity levels that maximize the expected social surplus $S(a_1) + S(a_2)$, where

$$S(a_t) = b(a_t) - h \cdot a_t. \quad (1)$$

Thus, the first-best solution is given by $a_1 = a_2 = a^{FB} > 0$, where a^{FB} is implicitly characterized by $b'(a^{FB}) = h$. In the remainder of this paper, we assume that the activity level is a hidden action, so the lawmaker can regulate the activity level only indirectly by setting suitable fines that have to be paid ex post if a harm has been caused.

3. The Analysis

In period $t = 2$, an individual with offense history $y \in \{0, 1\}$ chooses the activity level $a_2(F^y) = \arg \max_{a_2 \in [0, 1]} U(a_2; F^y)$, where

$$U(a; F) = b(a) - F \cdot a \quad (2)$$

denotes an individual's expected utility from activity level a when facing fine F in case of a harm. The second-period activity level that is optimal for the individual satisfies

$$b'(a_2(F^y)) = F^y \quad (3)$$

⁷Qualitatively similar results hold in the case in which a first-time offender may face different fines in the two periods and there is an exogenous restriction on fines in the first period only.

with $da_2(F^y)/dF^y = 1/b''(a_2(F^y)) < 0$; i.e., the higher the fine, the lower the individual's optimal activity level. Note that $U(a; h) = S(a)$, hence $a_2(F^y) \geq a^{FB}$ if and only if $F^y \leq h$. Application of the envelope theorem reveals that the individual's expected second-period utility is decreasing in the second-period fine, $dU(a_2(F^y); F^y)/dF^y = -a_2(F^y)$.

In period $t = 1$, the individual chooses his activity level $a_1 \in [0, 1]$ in order to maximize his overall expected utility,

$$EU(a_1) = U(a_1; F^0) + a_1 \cdot U(a_2(F^1); F^1) + (1 - a_1) \cdot U(a_2(F^0); F^0), \quad (4)$$

which is strictly concave, $d^2EU(a_1)/d(a_1)^2 = b''(a_1) < 0$. In consequence, if $dEU(a_1)/da_1|_{a_1=1} = -F^0 + U(a_2(F^1); F^1) - U(a_2(F^0); F^0) > 0$, then the optimal first-period activity level is $a_1(F^0, F^1) = 1$. If $dEU(a_1)/da_1|_{a_1=1} \leq 0$, the optimal first-period activity level is characterized by the first-order condition

$$b'(a_1(F^0, F^1)) = F^0 + U(a_2(F^0); F^0) - U(a_2(F^1); F^1). \quad (5)$$

In the latter case, application of the envelope theorem yields that the individual's activity level is strictly decreasing in both the fine for first-time offenders and the fine for repeat offenders, $\partial a_1(F^0, F^1)/\partial F^0 = (1 - a_2(F^0))/b''(a_1(F^0, F^1)) < 0$ and $\partial a_1(F^0, F^1)/\partial F^1 = a_2(F^1)/b''(a_1(F^0, F^1)) < 0$.

Anticipating the individual's behavior, the lawmaker's problem at date $t = 0$ amounts to setting fines $F^0 \leq l$ and F^1 in order to maximize the expected welfare

$$\begin{aligned} W(F^0, F^1) := & S(a_1(F^0, F^1)) + a_1(F^0, F^1) \cdot S(a_2(F^1)) \\ & + (1 - a_1(F^0, F^1)) \cdot S(a_2(F^0)). \end{aligned} \quad (6)$$

Making use of (3) and (5), the partial derivatives of the expected welfare are given by

$$\begin{aligned} \frac{\partial W(F^0, F^1)}{\partial F^0} = & \frac{\partial a_1(F^0, F^1)}{\partial F^0} [(F^0 - h)(1 - a_2(F^0)) + (F^1 - h)a_2(F^1)] \\ & + (1 - a_1(F^0, F^1)) \cdot [F^0 - h] \cdot \frac{da_2(F^0)}{dF^0} \end{aligned} \quad (7)$$

and

$$\begin{aligned} \frac{\partial W(F^0, F^1)}{\partial F^1} &= \frac{\partial a_1(F^0, F^1)}{\partial F^1} [(F^0 - h)(1 - a_2(F^0)) + (F^1 - h)a_2(F^1)] \\ &\quad + a_1(F^0, F^1) \cdot [F^1 - h] \cdot \frac{da_2(F^1)}{dF^1}. \end{aligned} \quad (8)$$

We can now characterize the penalties F^{0*} and F^{1*} that the lawmaker will stipulate at date $t = 0$.

Proposition 1. *If $l \geq h$, the lawmaker sets $F^{0*} = F^{1*} = h$.*

When the lawmaker sets $F^0 = F^1 = h$, then in the second period the individual will exert the socially desirable activity level $a_2(h) = a^{FB}$ irrespective of his offense history. Since his second-period utility does not depend on his offense history, according to (5) the individual's first-period activity level is purely determined by the fine $F^0 = h$ for first-time offenders, so the individual chooses the socially desirable activity level also in the first period, $a_1(h, h) = a^{FB}$.

Let us now consider the case in which the restriction on penalties for first-time offenders becomes relevant. We first establish that repeat offenders are punished at least as hard as first-time offenders; i.e., decreasing punishment schemes are never optimal.

Lemma 1. *If $l < h$, the lawmaker sets $F^{0*} \leq F^{1*}$.*

Note that harsher punishment of repeat offenders implies that $U(a_2(F^0); F^0) \geq U(a_2(F^1); F^1)$. Hence, the individual's first-period activity is characterized by (5), such that $\partial a_1(F^0, F^1)/\partial F^y < 0$ for $y \in \{0, 1\}$. Moreover, with $\lim_{a \rightarrow 1} b'(a) = 0$, $U(a; F)$ is bounded from above, such that $\lim_{a \rightarrow 0} b'(a) = \infty$ implies that $a_1(F^0, F^1) > 0$ must hold.

Next, Lemma 1, (7), and (8) imply that whenever $0 < l < h$, optimal sentencing requires $F^{1*} > h$ and thus overdeterrence of repeat offenders, i.e., $a_2(F^{1*}) < a^{FB}$. The reason for implementing this overdeterrence is the fact that the prospect of a larger fine that the individual may have to pay in the future helps to deter him from choosing an overly large activity level in the present.

Proposition 2. (i) If $l \in (0, h)$, the lawmaker sets $F^{1*} > h$.
(ii) If $l = 0$, she sets $F^{1*} = h$.

Note that if $l = 0$, then the activity level of a first-time offender in the second period will be equal to one, since he cannot be deterred at all. In this case, it is very attractive for the individual not to cause a harm in the first period, because then in the second period he may choose the largest possible activity level. As a result, in this special case the indirect effect of the second-period prospects on the first-period incentives is already strong enough such that the lawmaker prefers not to make use of overdeterrence of repeat offenders.

Finally, observe that the lawmaker optimally imposes maximum punishment for first-time offenders.

Proposition 3. If $l < h$, the lawmaker sets $F^{0*} = l$.

In what follows, assume that there is a unique interior solution regarding the choice of F^1 . Under the optimal punishment scheme, there is a non-monotonic relationship between the maximum punishment of first-time offenders and the optimal punishment of repeat offenders.

Corollary 1. There exist \underline{l} and \bar{l} with $0 < \underline{l} < \bar{l} < h$, such that $\left. \frac{dF^{1*}}{dl} \right|_{l < \underline{l}} > 0$ and $\left. \frac{dF^{1*}}{dl} \right|_{l > \bar{l}} < 0$.

Thus, for sufficiently large values of l , a relaxation of the restriction on penalties for first-time offenders prompts the lawmaker to reduce the overdeterrence of repeat offenders, since satisfactory incentives in the first period can now be provided by relatively large fines for first-time offenders. However, for small values of l , when society accepts an increase in the penalties for first-time offenders, then the lawmaker optimally reacts by also increasing the penalties for repeat offenders, thus aggravating overdeterrence. This policy allows the lawmaker to uphold the desirable indirect effect that large fines for repeat offenders have for the incentives of first-time offenders in the first period.

Our main results are illustrated in Figures 1 and 2.⁸ In particular, Figure 1 shows that when the restriction on the fine for first-time offenders is relaxed,

⁸Specifically, in the figures $b(a) = \sqrt{a} - a/2$ and $h = 1$. Hence, $a^{FB} \approx 0.11$.

then initially the optimal fine for repeat offenders increases. Hence, as depicted in Figure 2, the optimal activity level of a repeat offender is initially decreasing in the maximum penalty l that may be inflicted on a first-time offender. The optimal fine for repeat offenders is decreasing and thus the activity level of repeat offenders is increasing in l only if the maximum penalty is sufficiently large.

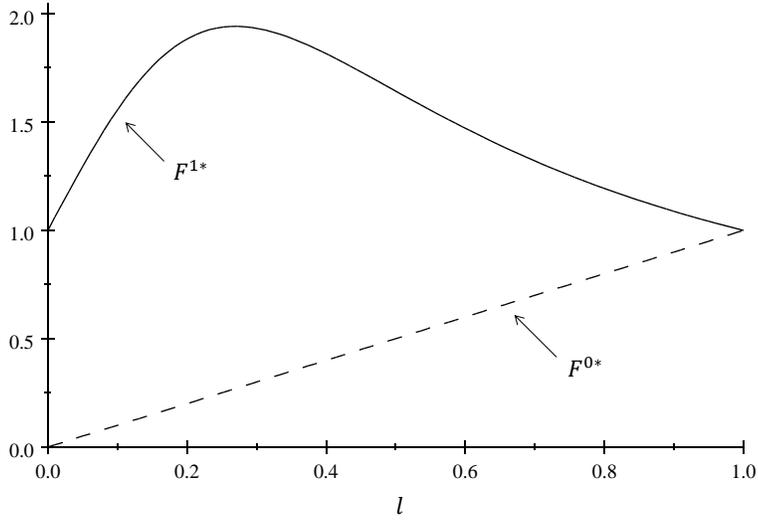


Figure 1. The optimal fine for repeat offenders, F^{1*} , and the optimal fine for first-time offenders, F^{0*} , as functions of the maximum penalty for first-time offenders, l .

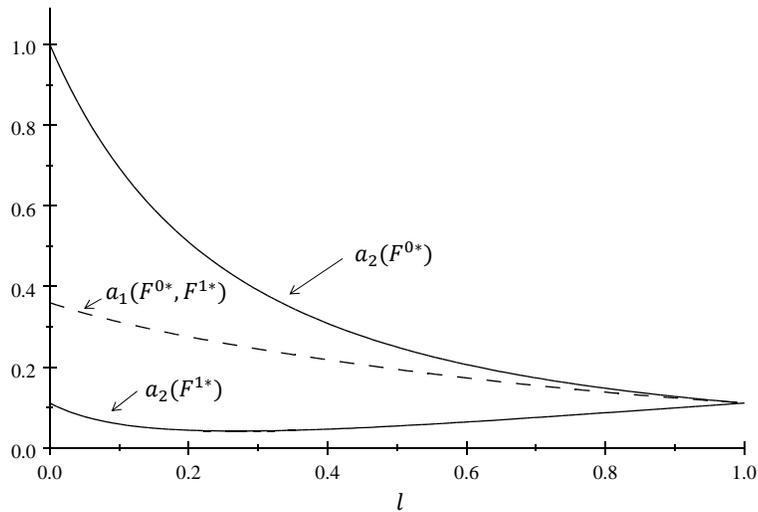


Figure 2. The first-period activity level is $a_1(F^{0*}, F^{1*})$, while the second-period activity level is $a_2(F^{1*})$ or $a_2(F^{0*})$, depending on whether or not the individual was previously convicted.

4. Conclusion

The optimal magnitude of sanctions is a major issue in the theory of public enforcement of law.⁹ In this paper, we have taken as given that society may put restrictions on admissible fines for first-time offenders and we have explored the implications that such restrictions may have for the optimal penalties of repeat offenders. Even though the socially optimal activity level of an individual does not depend upon whether or not he has been convicted before, overdeterrence of repeat offenders can be desirable because it improves the deterrence of first-time offenders. At first sight, one might be tempted to think that a relaxation of the restriction on fines for first-time offenders should lead the lawmaker to reduce the ex post inefficient overdeterrence of repeat offenders. Yet, we have shown that this is the case only if the restriction on the penalties of first-time offenders is relatively weak. Otherwise, when the restriction on penalties for first-time offenders is relaxed, then overdeterrence of repeat offenders should be further aggravated in order to uphold the incentive spillovers from the second period to the first period.

⁹See e.g. the literature survey by Kaplow and Shavell (2002).

Appendix

Proof of Proposition 1.

Given in the text. ■

Proof of Lemma 1.

We proceed by contradiction. First, suppose that $F^1 < F^0$ and $U(a_2(F^1)) > U(a_2(F^0)) + F^0$ is optimal. Then $a_1(F^0, F^1) = 1$ and $a_2(F^1) > a_2(F^0) > a^{FB}$ because $F^1 < F^0 \leq l < h$. Now consider an increase in the fine for repeat offenders to $\tilde{F}^1 \in [F^1, F^0]$. Clearly $a_2(F^0)$ is left unchanged and, as long as \tilde{F}^1 is sufficiently close to F^1 , also $a_1(F^0, \tilde{F}^1) = a_1(F^0, F^1) = 1$. Since $a_2(F^1) > a_2(\tilde{F}^1) > a^{FB}$, $S(a_2(\tilde{F}^1)) > S(a_2(F^1))$, thereby strictly increasing the value of the lawmaker's objective. In consequence, $F^1 < F^0$ and $U(a_2(F^1)) > U(a_2(F^0)) + F^0$ cannot be optimal.

Next, suppose that $F^1 < F^0$ and $U(a_2(F^1)) \leq U(a_2(F^0)) + F^0$ is optimal. Since $F^1 < F^0 \leq l < h$, we have $a^{FB} < a_2(F^0) < a_2(F^1)$, $S(a_2(F^1)) < S(a_2(F^0))$, and $U(a_2(F^0), F^0) < U(a_2(F^1), F^1)$. Moreover, $a_1(F^0, F^1)$ is characterized by (5), $\partial a_1(F^0, F^1)/\partial F^1 < 0$, and $a_1(F^0, F^1) > a^{FB}$ because $b'(a_1(F^0, F^1)) = F^0 + U(a_2(F^0), F^0) - U(a_2(F^1), F^1) < F^0 < h$. In consequence, a slight increase in F^1 increases overall expected total surplus,

$$\begin{aligned} \frac{\partial W(F^0, F^1)}{\partial F^1} &= \overbrace{\frac{\partial S(a_1(F^0, F^1))}{\partial a_1}}^{<0 \text{ for } a_1(F^0, F^1) > a^{FB}} \cdot \frac{\partial a_1(F^0, F^1)}{\partial F^1} \\ &\quad + \frac{\partial a_1(F^0, F^1)}{\partial F^1} [S(a_2(F^1)) - S(a_2(F^0))] \\ &\quad + a_1(F^0, F^1) \cdot \underbrace{\frac{dS(a_2(F^1))}{da_2}}_{<0 \text{ for } a_2(F^1) > a^{FB}} \cdot \frac{da_2(F^1)}{dF^1} > 0, \end{aligned} \quad (9)$$

contradicting the original choice of fines to be optimal. ■

Proof of Proposition 2.

(i) From Lemma 1, in the optimum we must have $F^0 \leq F^1$. Distinguishing four different cases, we proceed by contradiction.

First, suppose that $F^0 \in (0, l]$ and $F^1 \in [F^0, h]$ is optimal. With $0 < F^0 \leq F^1$, we have $a_1(F^0, F^1) \in (0, 1)$, $\partial a_1(F^0, F^1)/\partial F^1 < 0$, $a_2(F^0) \in (0, 1)$, and

$a_2(F^1) \in (0, 1)$. Then

$$\begin{aligned} \frac{\partial W(F^0, F^1)}{\partial F^1} &= \frac{\overbrace{\partial a_1(F^0, F^1)}^{<0}}{\partial F^1} \left[\overbrace{(F^0 - h)}^{<0} \overbrace{(1 - a_2(F^0))}^{>0} + \overbrace{(F^1 - h)}^{<0} \overbrace{a_2(F^1)}^{>0} \right] \\ &\quad + \underbrace{a_1(F^0, F^1)}_{>0} \cdot \underbrace{[F^1 - h]}_{\leq 0} \cdot \underbrace{\frac{da_2(F^1)}{dF^1}}_{<0} > 0, \end{aligned} \quad (10)$$

which contradicts $F^0 \in (0, l]$ and $F^1 \in [F^0, h]$ to be optimal.

Second, suppose that $F^0 = 0$ and $F^1 \in (0, h)$ is optimal. With $0 = F^0 < F^1$, we have $a_1(F^0, F^1) \in (0, 1)$, $\partial a_1(F^0, F^1)/\partial F^1 < 0$, $a_2(F^0) = 1$, and $a_2(F^1) \in (0, 1)$. Then

$$\frac{\partial W(F^0, F^1)}{\partial F^1} = \overbrace{(F^1 - h)}^{<0} \left[\overbrace{\frac{\partial a_1(F^0, F^1)}{\partial F^1}}^{<0} \cdot \overbrace{a_2(F^1)}^{>0} + \overbrace{a_1(F^0, F^1)}^{>0} \cdot \overbrace{\frac{da_2(F^1)}{dF^1}}^{<0} \right] > 0, \quad (11)$$

which contradicts $F^0 = 0$ and $F^1 \in (0, h]$ to be optimal.

Third, suppose that $F^0 = 0$ and $F^1 = 0$ is optimal. In this case, $a_1(F^0, F^1) = a_2(F^0) = a_2(F^1) = 1$ and $\partial a_1(F^0, F^1)/\partial F^1 < 0$. Then

$$\frac{\partial W(F^0, F^1)}{\partial F^1} = \overbrace{-h}^{<0} \left[\overbrace{\frac{\partial a_1(F^0, F^1)}{\partial F^1}}^{<0} + \overbrace{\frac{da_2(F^1)}{dF^1}}^{<0} \right] > 0, \quad (12)$$

which contradicts $F^0 = F^1 = 0$ to be optimal.

Finally, suppose that $F^0 = 0$ and $F^1 = h$ is optimal. With $0 = F^0 < F^1 = h$, we have $a_1(F^0, F^1) \in (0, 1)$, $\partial a_1(F^0, F^1)/\partial F^1 < 0$, $a_2(F^0) = 1$, and $a_2(F^1) \in (0, 1)$. While $\partial W(F^0, F^1)/\partial F^1 = 0$, in this case

$$\frac{\partial W(F^0, F^1)}{\partial F^0} = \overbrace{-h}^{<0} \cdot \overbrace{(1 - a_1(F^0, F^1))}^{>0} \cdot \overbrace{\frac{da_2(F^0)}{dF^0}}^{<0} > 0, \quad (13)$$

which contradicts $F^0 = 0$ and $F^1 = h$ to be optimal.

(ii) To see that $F^1 = h$ is optimal for $l = 0$, note that $a_2(F^0) = a_2(0) = 1$. From (8), we have

$$\frac{\partial W(0, F^1)}{\partial F^1} = (F^1 - h) \left[\underbrace{\frac{\partial a_1(0, F^1)}{\partial F^1}}_{<0} \cdot \underbrace{a_2(F^1)}_{>0} + \underbrace{a_1(0, F^1)}_{>0} \cdot \underbrace{\frac{da_2(F^1)}{dF^1}}_{<0} \right]. \quad (14)$$

The desired result then follows from the fact that $\partial W(0, F^1)/\partial F^1 > 0$ for $F^1 < h$ and $W(0, F^1)/\partial F^1 < 0$ for $F^1 > h$. ■

Proof of Proposition 3.

If $l = 0$, then $F^0 = l = 0$. In the remainder of the proof, we consider $l > 0$.

First, with $F^1 > h$ by Proposition 2, $F^0 = 0$ cannot be optimal because in this case $a_2(F^0) = 1$ and $a_1(F^0, F^1) < 1$ such that

$$\frac{\partial W(F^0, F^1)}{\partial F^1} = (F^1 - h) \left[\frac{\partial a_1(F^0, F^1)}{\partial F^1} \cdot a_2(F^1) + a_1(F^0, F^1) \cdot \frac{da_2(F^1)}{dF^1} \right] < 0. \quad (15)$$

Hence, overall expected total surplus could be increased by reducing F^1 , which makes $F^0 = 0$ incompatible with optimality.

Next, suppose that $0 < F^0 < l$ is optimal, i.e., the constraint $F^0 \leq l$ is not binding in the optimum. Then the optimal values of F^0 and F^1 have to be characterized by the corresponding system of first-order conditions, i.e., $\frac{\partial W(F^0, F^1)}{\partial F^0} = 0$ and $\frac{\partial W(F^0, F^1)}{\partial F^1} = 0$ have to be jointly satisfied. From (7) and (8), this requires

$$\frac{(1 - a_1(F^0, F^1)) \cdot [F^0 - h] \cdot \frac{da_2(F^0)}{dF^0}}{\left(\frac{\partial a_1(F^0, F^1)}{\partial F^0} \right)} = \frac{a_1(F^0, F^1) \cdot [F^1 - h] \cdot \frac{da_2(F^1)}{dF^1}}{\left(\frac{\partial a_1(F^0, F^1)}{\partial F^1} \right)}, \quad (16)$$

where $a_1(F^0, F^1) \in (0, 1)$, and $da_2(F^y)/dF^y < 0$ and $\partial a_1(F^0, F^1)/\partial F^y < 0$ for $F^y \in \{F^0, F^1\}$. With $F^0 < l < h$ by hypothesis, a necessary condition for (16) to hold is that $F^1 < h$. According to Proposition 2, however, $F^1 < h$ cannot be optimal, which contradicts $F^0 < l$ to be optimal. ■

Proof of Corollary 1.

With $F^{0*} = l$ the optimal fine for repeat offenders is characterized by

$$\begin{aligned} \frac{\partial W(l, F^{1*})}{\partial F^1} &= \frac{\partial a_1(l, F^{1*})}{\partial F^1} [(l - h)(1 - a_2(l)) + (F^{1*} - h)a_2(F^{1*})] \\ &\quad + a_1(l, F^{1*}) \cdot [F^{1*} - h] \cdot \frac{da_2(F^{1*})}{dF^1} = 0. \end{aligned} \quad (17)$$

Implicit differentiation of (17) with respect to l yields

$$\begin{aligned} & \frac{\partial^2 W(l, F^{1*})}{\partial (F^1)^2} \cdot \frac{dF^{1*}}{dl} + \frac{\partial^2 a_1(l, F^{1*})}{\partial F^1 \partial F^0} [(l-h)(1-a_2(l)) + (F^{1*}-h)a_2(F^{1*})] \\ & + \frac{\partial a_1(l, F^{1*})}{\partial F^1} \left[(1-a_2(l)) - \frac{da_2(l)}{dF^0} (l-h) \right] \\ & + \frac{\partial a_1(l, F^{1*})}{\partial F^0} (F^{1*}-h) \frac{da_2(F^{1*})}{dF^1} = 0. \end{aligned} \quad (18)$$

For $l=0$ we have $a_2(0) = 1$ and $F^{1*} = h$ such that (18) becomes

$$\frac{\partial^2 W(0, F^{1*})}{\partial (F^1)^2} \cdot \frac{dF^{1*}}{dl} + \frac{\partial a_1(0, F^{1*})}{\partial F^1} \cdot \frac{da_2(0)}{dF^0} \cdot h = 0. \quad (19)$$

Likewise, for $l=h$ we have $F^{1*} = h$ such that (18) becomes

$$\frac{\partial^2 W(h, F^{1*})}{\partial (F^1)^2} \cdot \frac{dF^{1*}}{dl} + (1-a_2(h)) \cdot \frac{\partial a_1(h, F^{1*})}{\partial F^1} = 0. \quad (20)$$

Finally, note that

$$\begin{aligned} \frac{\partial^2 W(l, F^{1*})}{\partial (F^1)^2} &= \frac{\partial^2 a_1(l, F^{1*})}{\partial F^1 \partial F^1} [(l-h)(1-a_2(l)) + (F^{1*}-h)a_2(F^{1*})] \\ &+ \frac{\partial a_1(l, F^{1*})}{\partial F^1} \left[\frac{\partial a_2(F^{1*})}{\partial F^1} (F^{1*}-h) + a_2(F^{1*}) \right] \\ &+ \frac{\partial a_1(l, F^{1*})}{\partial F^1} (F^{1*}-h) \frac{\partial a_2(F^{1*})}{\partial F^1} \\ &+ a_1(l, F^{1*}) \frac{\partial a_2(F^{1*})}{\partial F^1} + a_1(l, F^{1*}) (F^{1*}-h) \frac{d^2 a_2(F^{1*})}{dF^1 dF^1} \end{aligned} \quad (21)$$

such that

$$\frac{\partial^2 W(0, F^{1*})}{\partial (F^1)^2} = \frac{\partial a_1(0, F^{1*})}{\partial F^1} a_2(F^{1*}) + a_1(0, F^{1*}) \frac{da_2(F^{1*})}{dF^1} < 0 \quad (22)$$

and

$$\frac{\partial^2 W(h, F^{1*})}{\partial (F^1)^2} = \frac{\partial a_1(h, F^{1*})}{\partial F^1} a_2(F^{1*}) + a_1(h, F^{1*}) \frac{da_2(F^{1*})}{dF^1} < 0. \quad (23)$$

Combining (19) and (22) reveals that $dF^{1*}/dl|_{l=0} > 0$. Likewise, combining (20) and (23) yields $dF^{1*}/dl|_{l=h} < 0$. The results stated in the corollary then follow from continuity. ■

References

- Chen, B.R., Chiu, Y.S., 2013. Interim performance evaluation in contract design. *Economic Journal* 123, 665–698.
- Chu, C.Y., Hu, S.-C., Huang, T.-Y., 2000. Punishing repeat offenders more severely. *International Review of Law and Economics* 20, 127–140.
- Kaplow, L., Shavell, S., 2002. Economic analysis of law. *Handbook of Public Economics* 3, 1661–1784.
- Kräkel, M., Schöttner, A., 2010. Minimum wages and excessive effort supply. *Economics Letters* 108, 341–344.
- Miceli, T.J., 2013. Escalating penalties for repeat offenders: why are they so hard to explain? *Journal of Institutional and Theoretical Economics* 169, 587–604.
- Ohlendorf, S., Schmitz, P.W., 2012. Repeated moral hazard and contracts with memory: the case of risk-neutrality. *International Economic Review* 53, 433–452.
- Pi, J., 2014. Job design with sequential tasks and outcome externalities revisited. *Economics Letters* 125, 274–277.
- Polinsky, A.M., Rubinfeld, D.L., 1991. A model of optimal fines for repeat offenders. *Journal of Public Economics* 46, 291–306.
- Polinsky, A.M., Shavell, S., 1998. On offense history and the theory of deterrence. *International Review of Law and Economics* 18, 305–324.
- Polinsky, A.M., Shavell, S., 2007. The theory of public enforcement of law. *Handbook of Law and Economics* 1, 403–454.
- Rubinstein, A., 1980. On an anomaly of the deterrent effect of punishment. *Economics Letters* 6, 89–94.
- Schmitz, P.W., 2000. On the joint use of liability and safety regulation. *International Review of Law and Economics* 20, 371–382.
- Schmitz, P.W., 2005. Should contractual clauses that forbid renegotiation always be enforced? *Journal of Law, Economics, and Organization* 21, 315–329.

- Shavell, S., 1984. A model of the optimal use of liability and safety regulation. *RAND Journal of Economics* 15, 271–280.
- Shavell, S., 2004. *Foundations of Economic Analysis of Law*. Harvard University Press.
- Stigler, G.J., 1970. The optimum enforcement of laws. *Journal of Political Economy* 78, 526–536.
- Tsai, T.-S., Kung, S.-C., 2011. Sequential investments, know-how transmission, and optimal organization. *BE Journal of Theoretical Economics* 11.