Crime and Punishment: On the Optimality of Imprisonment although Fines Are Feasible[#]

Ingolf Dittmann[§]

November 1999

Abstract:

A general result of the literature on crime and punishment is that imprisonment is not optimal if fines can be used instead. This paper presents a positive model which predicts the opposite for serious crimes, namely that imprisonment will be used, even if offenders could pay a correspondingly high fine. Hence, this model can explain mandatory prison sentencing, which is often found in practice for serious crimes.

In contrast to the standard normative model in which a social planner chooses the detection probability and the type of punishment, this model separates these two decisions. In the first stage of the model, the individuals determine the type of punishment in a referendum. Given this decision, the government chooses the detection probability in the second stage. The main result is that individuals vote unanimously for imprisonment if the harm caused by the considered crime – and therefore the size of the penalty – is large.

JEL Classification Codes: K42, D72

[#] I would like to thank Cay Folkers, Clive Fraser, Wolfgang Leininger, Alistair Munro, Dillip Mookherjee, Wolfram Richter, and the seminar participants at the Universities of Leicester and Dortmund for helpful discussions and comments. The paper was written while the author enjoyed the hospitality and stimulating atmosphere of the Public Sector Economic Research Centre at the University of Leicester. Financial support from the Rudolf Chaudoire Stiftung is gratefully acknowledged.

[§] University of Dortmund, Department of Economics, Graduiertenkolleg, 44221 Dortmund, Germany, currently: University of California at San Diego, Department of Economics, 9500 Gilman Drive, La Jolla, CA 92093-0508. E-mail: i.dittman@weber.ucsd.edu

1. Introduction

The literature on the economics of crime, pioneered by Becker (1968) and Polinsky and Shavell (1979, 1984), applies the economic approach of decisions under uncertainty to criminal behaviour, assuming that the behaviour of criminals does not differ in principle from the behaviour of other economic agents. The main issue of this literature is to derive the optimal prosecution policy, including the optimal probability that an offender is detected and punished, the optimal size of punishment and the optimal type of punishment (i.e., imprisonment or fine). The standard approach is a two-stage game. In the first stage, a social planner chooses the control variables, usually the detection probability, the size of punishment and the type of punishment. After the planner's decisions are made public, each individual decides whether or not to commit the crime in the second stage of the game. In the basic set-up, individuals only differ in the gain they expect from committing the crime.

One prominent result of this literature, which is very robust to modifications of the standard model, is that imprisonment is not optimal if the offender is able to pay a fine instead. The economic intuition is straightforward: Imprisonment is costly for society, as prisons must be maintained and prisoners are hindered to take up legal employment. On the other hand, fines can be used to finance the police or to compensate victims. Therefore, Polinsky and Shavell (1984) conclude that "it is desirable to use the fine to its maximum feasible extent before possibly supplementing it with an imprisonment term." This result is somewhat at odds with reality, however. Typically, offenders are punished with either a fine or a prison term but rarely with a fine and a prison term. Moreover, many crimes – and especially serious ones – are punished with mandatory imprisonment, so that the judge has no discretion to impose a fine instead, no matter how rich the offender is.

The objective of the present paper is to offer an explanation for the existence of mandatory prison sentencing even if individuals are able to pay a fine instead. In contrast to most of the literature on crime and punishment which is concerned with normative issues, my approach is positive. Starting out from the standard model as described above, I substitute the social planner with a government which is in charge of criminal prosecution in practice. An important difference between the government and a social planner is that the government typically cannot choose all variables of the criminal justice system. While it can determine the expenditure on police and thereby the detection probability for a specific crime, it usually cannot change the type and the size of the punishment – at least not during one parliamentary term. Typically, the type and the size of the punishment is laid down as a law by the parliament (which in turn is elected by the individuals) so that these two variables are less flexible than the detection probability.

In order to model this complicated decision process, I consider a three-stage game in this paper. In the first stage, all individuals decide in a referendum whether the considered crime should be punished by a prison term or a fine. In the second stage, the government chooses the detection probability by determining the police budget, and, in the third stage, each individual decides whether or not to commit the crime. Finally, criminals are prosecuted and punished as announced. The referendum in the first stage of the game is a typical constitutional mechanism for the in-period choice of social institutions (see, e.g., Switzerland). It should be regarded as a convenient approximation to the complex decision process in reality. The analysis focuses on the detection probability and the *type* of punishment, whereas the *size* of punishment is treated as an exogenous variable. A number of other papers have shown that the size of the punishment should be related to the harm caused by the considered crime (see, e.g., Polinsky and Shavell, 1979 or Mookherjee and Png, 1994). As the present model considers only one specific crime, its harm and the size of punishment are assumed to be exogenous. An important and driving assumption of the model is that the objectives of the government differ from those of the individuals. In particular, I assume that the government does not fully take into account the individuals' gains and losses from criminal activity or that individuals do not fully take into account the government's budget considerations.[1]

[1] The different objectives of government and citizens are reminiscent of principal-agent-models, where the interests of principal and agent differ by assumption. Indeed, Dittmann (1999) considers a principal-agent model and determines the behaviour of the principal contingent on the type of punishment (prison or fine) which she can inflict on the agent if he is convicted. However, principal-agent models cannot be used to describe a political economy, because there are no contracts (or, more formally, no participation constraints) between the government as principal and the citizens as agents. The relationship between government and citizens is more like a mutual principal-agent relationship in which the government is the principal (when it comes to prosecuting criminals) and the agent (when it comes to elections).

The analysis of the second stage of this model reveals that the government's choice of the detection probability strongly depends on the type of the punishment. Consider a short prison term and an equivalent fine, so that individuals are indifferent between these two punishments. In this case, the government chooses a higher detection probability if the punishment is a fine than if it is a prison term, because the revenues from fines make prosecution more worthwhile. Now consider a very long prison term with an equivalent high fine, so that deterrence is much cheaper. In the case of mandatory prison sentencing, the government now chooses a detection probability that deters all individuals from committing the crime, so that social harm from crime is zero. On the other hand, if the punishment is a fine, the government chooses a *smaller* detection probability, in order to encourage some crime. Thereby, the government receives revenues from fines which more than offset the disutility it attaches to the corresponding level of crime.

Moreover, the analysis of the first stage of the model shows that most individuals are interested in a high detection probability if the harm caused by the considered crime is large. Taking the government’s behaviour in the second stage into account, individuals therefore vote for mandatory imprisonment if the punishment is large and for a fine if the punishment is small. As more serious crimes are usually associated with larger penalties, this model predicts that the law will prescribe mandatory imprisonment for serious crimes.

To my knowledge, there are only two other papers, Chu and Jiang (1993) and Levitt (1997), which try to explain why imprisonment is employed before fines are used to their maximum feasible extent. Both papers consider variants of the standard two-stage model with a social planner. Chu and Jiang (1993) assume that individuals can choose between different crimes and that they differ in their wealth, which is perfectly observable. Ideally, the fine should now be conditioned on the committed crime and the offender’s wealth. In particular, rich individuals should face higher fines than poor individuals for less severe crimes, so that poor individuals are deterred from serious crimes and rich individuals from any crime. Chu and Jiang (1993) assume, however, that the fine for a particular crime cannot depend on the offender’s wealth. If, additionally, rich individuals are more prison averse than poor individuals, imprisonment becomes an alternative means to punish rich individuals more severely than poor individuals. Consequently, a combination of imprisonment and a fine, which is smaller than the individual’s wealth, can be optimal for less severe crimes. While this model can explain prison terms accompanied by fines, it cannot explain the use of imprisonment without additional fines.

In contrast, Levitt (1997) assumes that the enforcement agency cannot observe the individuals' wealth, so that fines can only be enforced under the threat of imprisonment. As criminals typically are poor and hence have a low disutility of jail, the fine which can be enforced under threat of a given prison term is quite small. Yet, a small fine would make the crime more attractive for rich individuals. Therefore, it might be optimal not to use fines but only prison terms in many situations. Since, in reality, at least some information about the offender's wealth is available, Levitt (1997) argues that courts should decide whether the offender is punished with a fine (accompanied by a threat of imprisonment) or with imprisonment alone, i.e., he argues against mandatory prison sentencing.[2]

The rest of this paper is organised as follows. Section 2 introduces the model in detail. Section 3 analyses the government's choice of the detection probability given the type of punishment. Section 4 derives the outcome of the referendum in the first stage of the model, and Section 5 contains conclusions and further notes.

2. The Model

I consider a specific crime that can be committed by any of an infinite number of risk-neutral individuals. By committing the crime, individual i gains an extra utility A_i and another randomly chosen individual j suffers the harm H. Alternatively, the harm H, which does not depend on i, can be regarded as damage to public property, i.e., as a public "bad". The utility from crime, A_i, is individual i's private knowledge, whereas the distribution of the A_i's is public knowledge. Let $f(A)$ denote the corresponding density function with support [0, 1], i.e., the highest utility from crime is normalised to 1. For computational simplicity, I further assume that the A_i’s are uniformly distributed, so that

---

[2] Dittmann (2000) provides a more detailed comparison between Chu and Jiang (1993), Levitt (1997) and the present paper.

f(A) = 1 for $0 \le A \le 1$. Individuals only differ in their utility A_i they can derive from committing the considered crime. In all other respects, they are identical.[3]

An individual who has committed the crime is convicted with probability *r*. If convicted, he incurs a loss *P* either because he must pay the fine *P* or because he is imprisoned for an equivalent period. A fine and a prison term of a certain length are "equivalent" if individuals are indifferent between these two punishments. All individuals have enough personal funds to pay the fine *P*. The detection probability *r* is set by the government which thereby incurs the costs *c(r)* with $c(0) = 0$, $c'(r) > 0$, $c''(r) > 0$ and $\lim_{r \to 1} c(r) = \infty$. Accordingly, individual *i* chooses to commit the crime if $A_i > rP$, so that the proportion of criminals is $q = q(r,P) = \int_{rP}^{\infty} f(A)dA = \min\{1 - rP, 0\}$.

If the punishment is a fine, the government's utility function is the sum of the utility loss from crime ($-sqH$), the expected revenues from fines (qrP), and the costs of policing ($-c(r)$):

$$U_G(r) = -sqH + qrP - c(r)\,. \tag{1}$$

Here, the total population is normalised to one and $s \in [0, 1]$ denotes the government's "self-interest" parameter with which those elements of social welfare are discounted that do not directly influence the government's revenues. The underlying assumption is that governments have a strong preference for a discretionary budget which can be used to increase the probability of re-election. If the punishment is a prison term, the government's utility function reduces to

$$U_G(r) = -sqH - c(r)\,. \tag{1'}$$

In order to keep the model simple, costs of imprisonment are not included in (1'). This restriction does not affect the model's qualitative results, as we will see in the next section. Note that both utility functions contain neither the gain from crime nor the expected disutility of criminals due to punishments. If *s = 1*, (1) and (1') become Benthamite social welfare functions with zero weight given to criminals.

---

[3] In Section 4, I will briefly consider the effect of relaxing the two assumptions (i) that individuals are equally wealthy and (ii) that the individuals' utilities from crime are uniformly distributed. The main results of this paper continue to hold if either assumption is dropped. The chief purpose of these two assumptions is to make the model solvable with standard calculus and to arrive at clear-cut results.

The utility function of individual i is given by

$$U_i(r) = \max\{A_i - rP, 0\} - qH + mqrP - mc(r) \tag{2}$$

if the penalty is a fine and by

$$U_i(r) = \max\{A_i - rP, 0\} - qH - mc(r) \tag{2'}$$

in the case of mandatory imprisonment. Here, $\max\{A_i - rP, 0\}$ is the net gain from criminal activity, $-qH$ is the expected loss from other individuals' criminal activity and $m(qrP - c(r))$ or $-mc(r)$, respectively, is the government's budget discounted with the "myopia" parameter $m \in [0, 1]$. The smaller m, the more myopic individuals are, in the sense that they attach a smaller weight to the costs of the police system and to potential revenues from fines paid by other individuals. A possible explanation for $m < 1$ is that the government will use the remaining budget for pet projects which are only of limited value to the average individual. Note that, as long as $m < 1$ or $s < 1$, the government's objectives differ from those of the (non-criminal) individuals. This difference in objectives will turn out to be the driving force of the model.

The game considered consists of three stages. In stage 1, individuals decide in a referendum on the type of the punishment. In stage 2, the government chooses the detection probability r and, in stage 3, each individual decides whether or not to commit the crime. In stage 1, each individual can choose between three options: (a) the crime is punished with a prison term, (b) the crime is punished with a fine and (c) the crime is not prosecuted or punished at all. The option which receives the largest number of votes is implemented before the government chooses the detection probability r in stage 2 of the game.

Note that the size of the punishment, P, is fixed. Obviously, if government or individuals could choose P, they would choose P as high as possible, in order to deter all crime with a very small detection probability at a very low cost. This is the so-called "maximum-punishment" result (see Becker, 1968). In more detailed models, however, this result often does not hold. Stigler (1970) argues, for instance, that infinite punishments are generally not optimal if individuals can choose between different crimes which differ in the amount of harm they cause (see also Mookherjee and Png, 1994, and Friedman and Sjostrom, 1993). He argues that the most harmful crime should be punished with the most severe punishment whereas less severe crimes should receive smaller punishments in order to give notorious criminals an incentive not to commit serious crimes. Polinsky and Shavell (1979, 1991) show that, if individuals are risk averse or if individuals differ in their wealth, optimal penalties are finite, i.e., smaller than the wealth of most individuals. Moreover, Andreoni (1991) demonstrates that finite fines may become optimal if there is a positive error probability with which the government "convicts" non-criminals.

3. The government's choice of the detection probability

We first consider the case that individuals have voted for a mandatory prison sentence in the first stage of the game. Proposition 1 below describes which detection probability the government chooses.

Proposition 1 (optimal detection probability under mandatory prison sentencing):

a) If $P < \frac{c'(0)}{sH}$, the government does not prosecute criminals, i.e., $r_P^* = 0$.

b) If $\frac{c'(0)}{sH} \le P < \tilde{P}$, where $\tilde{P}$ is the solution of $c'(1/\tilde{P}) = sH\tilde{P}$, the government chooses r_P^* such that $c'(r_P^*) = sHP$.

c) If $P \ge \tilde{P}$, the government chooses $r_P^* = \frac{1}{P}$ and all criminals are deterred.

Proof: See Appendix.

If the punishment is very small, a given level of deterrence $r{\cdot}P$ is very expensive to achieve so that the government does not prosecute criminals. Therefore, I call this area the *no-prosecution* region. If the size of the punishment, *P*, is larger than the threshold $c'(0)/(sH)$, the detection probability r_P^* is positive and increases with increasing *P*, *s* and *H*. As long as $P < \tilde{P}$, there is always some crime, wherefore I call this area the *crime-and-prosecution* region. If $P \ge \tilde{P}$, i.e., in the *full-deterrence* region, all crime is deterred because $r_P^* P = 1$. Here, r_P^* decreases with further increasing *P*. Note that the *full-deterrence* region gets larger if *s* or *H* become larger, whereas the *no-prosecution* region shrinks.

The effect of the government's self-interest parameter *s* on the optimal detection probability *r* can best be seen in an example. Let *H* = 1 and $c(r) = -0.1\ln(1-r)$. Figure 1 shows r_P^* for two values of *s*, namely *s* = 1 (solid line) and *s* = 0.5 (broken line). It demonstrates that a drop in *s*, i.e., an increase in the government's self-interest leads to a smaller detection probability for small punishments. For large punishments, the government still chooses to deter all criminals.

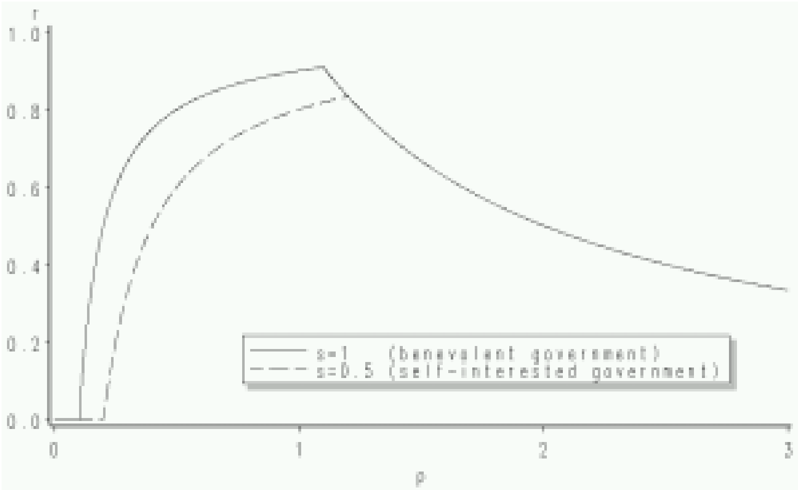

Figure 1: The government's choice of the detection probability r if the punishment is a prison term

Note that imprisonment actually occurs only in the *crime-and-prosecution* region. In the *no-prosecution* region, criminals are not punished, and, in the *full-deterrence* region, nobody commits the crime. At this point, it is straight-forward to see why neglecting costs of imprisonment does not affect the qualitative results of my model: Assume for a moment that there are positive costs of imprisonment. Then the *no-prosecution* and the *full-deterrence* regions would become larger, whereas the *crime-and-prosecution* region would shrink and the rise in $r_P^*(P)$ would become steeper. (These findings can be easily verified algebraically.) By changing the prosecution policy in this way, the government effectively reduces the number of individuals imprisoned.

Next, we consider the government's behaviour, if, in the first stage of the game, individuals have decided that the penalty is a fine. In this case, the government maximises (1) instead of (1') and we obtain:

Proposition 2 (optimal detection probability if fines are allowed):

a) If $P < \dfrac{c'(0)}{1+sH}$, the government does not prosecute criminals, i.e. $r_F^* = 0$.

b) If $sH > 1$ and $P \geq \hat{P}$, where $\hat{P}$ is the solution of $c'(1/\hat{P}) = (sH-1)\hat{P}$, the government chooses $r_F^* = \dfrac{1}{P}$ and all criminals are deterred.

c) In all other cases, the government chooses r_F^* such that $c'(r_F^*) + 2P^2 r_F^* = (sH+1)P$.

In particular, if $sH \leq 1$, there is never full deterrence for any finite punishment *P*.

Proof: See Appendix.

If *sH* > 1, i.e., if the harm from the government's point of view is larger than the largest utility an individual can obtain from committing the crime, we again have the three regions *no-prosecution*, *crime-and-prosecution* and *full-deterrence*. If *sH < 1*, however, the *full-deterrence* region disappears. In order to obtain an intuition for this important result of Proposition 2, assume that *sH* is small, *P* is large and *r = 1/P*, i.e., that the government chooses full deterrence. If the government now reduces *r* slightly, a small proportion of individuals will commit the crime and cause a harm that is quite small from the government's point of view. On the other hand, the government now not only saves expenditures on the police but also receives revenues from fines paid by criminals. If *sH* is sufficiently small, these benefits from allowing some crime outweigh the costs and the government decides not to deter all criminals.

The optimal detection probability r_F^* increases in the *crime-and-prosecution* region only for small fines. For large fines, r_F^* decreases with increasing *P*. Moreover, r_F^* decreases if *s* decreases. Hence, a more self-interested government always chooses a smaller detection probability. Figure 2 illustrates these findings for *H* = 1.5 and $c(r) = \text{-}0.1\ln(1-r)$.

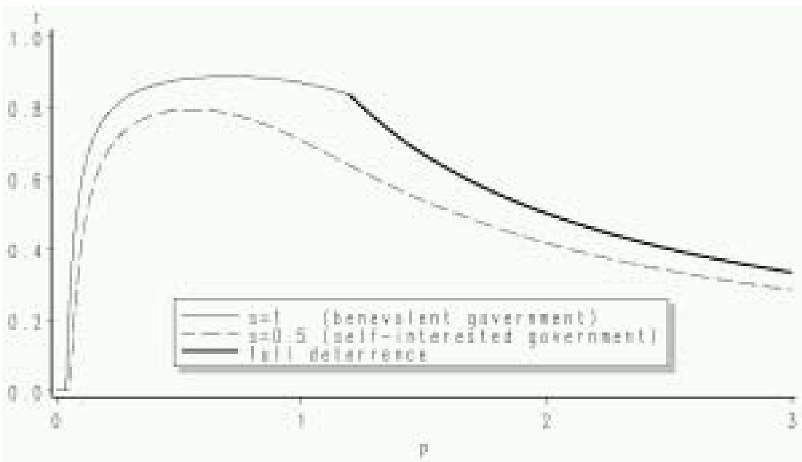

Figure 2: The government's choice of the detection probability r if the punishment is a fine

In order to better understand the voting behaviour of the individuals in the first stage of the game, we now compare the government’s choice of the detection probability between the two regimes *imprisonment* and *fine*. If we compare Propositions 1 and 2, we immediately see that the *no-prosecution* region is smaller if the punishment is a fine (Proposition 2) than if it is a prison term (Proposition 1). The additional revenues from fines induce the government to prosecute the crime even if the fine is quite small. Moreover, we obtain:

Corollary 1 (Comparison of the chosen detection probabilities given the type of punishment):

Let $\overline{P}$ be the solution of $c'\left(\frac{1}{2\overline{P}}\right) = sH\overline{P}$.

a) If $P < \overline{P}$, $r_F^* \geq r_P^*$ with strict inequality if $P > \frac{c'(0)}{1+sH}$.

b) If $P > \overline{P}$, $r_F^* \leq r_P^*$ with strict inequality if $sH \leq 1$ or if $sH > 1$ and $P < \hat{P}$.

Proof: See Appendix.

Figures 3 and 4 illustrate the statements of Corollary 1 for H = 1 and $c(r) = \text{-}0.1\ln(1-r)$. Figure 3 considers the case of a benevolent government with s = 1, whereas Figure 4 illustrates the behaviour of a strongly self-interested government with s = 0.3. For small P (i.e. $P < \overline{P}$, where $\overline{P}$ is the intersection of the two curves $r_F^*(P)$ and $r_P^*(P)$), the government chooses a higher detection probability if the penalty is a fine. For large P, it chooses a higher detection probability if the penalty is a prison term.

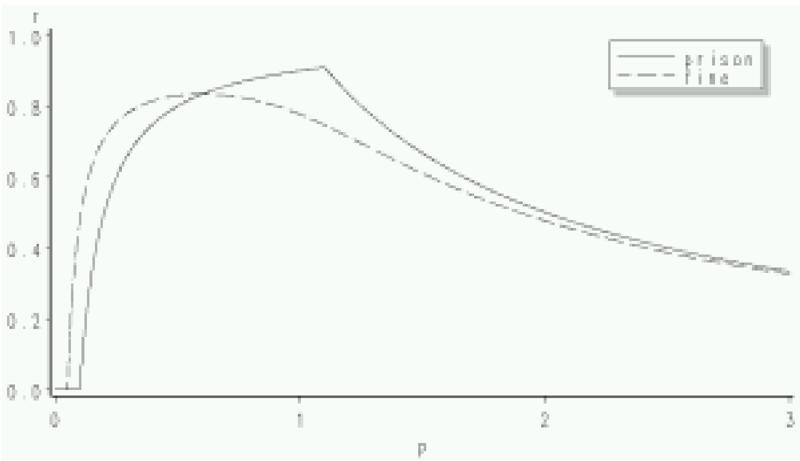

Figure 3: The benevolent government's choice of the detection probability r (s = 1)

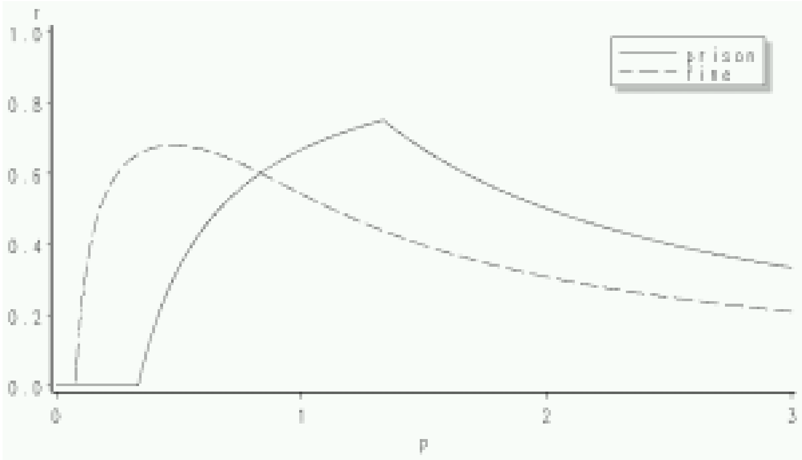

Figure 4: The self-interested government's choice of the detection probability r (s = 0.3)

In order to understand this apparently counter-intuitive result, look at the fine as a monetary reward for catching a criminal. If the punishment is small, the government chooses – without such a reward - a quite low or even zero detection probability, which leads to a high number of criminals. Hence, the introduction of a reward induces the government to increase the detection probability in order to raise the probability of catching a criminal; the incentives work in the "right" direction. If the punishment is very large, however, the government chooses – without such a reward – a quite high detection probability, so that the number of criminals is quite low or even zero. Now the introduction of monetary incentives induces the government to *reduce* the detection probability in order to increase the number of criminals and thereby the probability of catching a criminal. Here, monetary incentives work in the "wrong" direction.

In Figure 3, i.e., for a benevolent government with $s = 1$, the two detection probabilities $r_F^*(P)$ and $r_P^*(P)$ already differ significantly. Figure 4 demonstrates that, if the government is strongly self-interested, i.e., if s is considerably smaller than 1, the difference between the two curves is much larger and the intersection $\overline{P}$ lies further to the right. Therefore, if a government strongly appreciates a discretionary budget, its expenditure on the police will critically depend on the type of punishment.

4. The individuals' choice of the type of punishment

This section analyses the first stage of the game in which the individuals determine the type of punishment in a referendum.

Proposition 3 (The optimality of imprisonment):

Suppose that $m < 1$ or $s < 1$. Then for every $H \in \left]\max\left\{\frac{m}{2-sm}, \frac{1}{2}\right\}, \frac{1}{s}\right[$ there exists a P_0 such that for all $P > P_0$ more than 50% of the individuals strictly prefer imprisonment to a fine and to no prosecution. If $m = 1$ and $s = 1$, no individual strictly prefers imprisonment for any combination of H and P.

Proof: See Appendix.

Basically, Proposition 3 states that imprisonment is optimal for some (H, P) unless the objectives of government and individuals are exactly the same (i.e., $s = 1$ and $m = 1$). The more the objectives of government and individuals differ the larger is the interval of harms H for which individuals will vote for imprisonment. Note that for $H \geq 1/s$, imprisonment is still weakly optimal for sufficiently large P. The reason is that both types of punishment, imprisonment and fine, lead to full deterrence, so that individuals and government are indifferent between prison and fines (cf. Proposition 1(c) and Proposition 2(b)).

Much stronger results than those in Proposition 3 can be obtained, if we consider the special case $m = 0$, i.e., that individuals are completely myopic and judge the law enforcement system only by their personal expected harm and gain and do not take the government's budget into account. If we suppose that all savings or revenues from running the criminal justice system are not paid back to the individuals but spent on some public good at the government's discretion, the assumption that $m = 0$ is reasonable and corresponds to an established approach in optimal taxation theory. Hence, this is an important special case of the model.

Mathematically, assuming $m = 0$ greatly simplifies the individuals' utility function, which becomes piecewise linear. The utility of an individual who does not commit the crime then is $U_i^N(r) = -qH = -(1-rP)H$ and, thus, increases monotonically in the detection probability r. On the other hand, the utility of a "criminal" is $U_i^C(r) =$

$A_i - rP - (1 - rP)H = A_i + (H - 1)rP - H$, which also increases monotonically in r if $H > 1$. Therefore, if the harm H is larger than the largest gain an individual can obtain from committing the crime, all individuals vote for the type of punishment which leads to the largest detection probability r. Combined with the results from Corollary 1, this implies that individuals choose a prison term if the size of the punishment is large whereas they choose a fine if it is small. If $H < 1$, individuals are no longer unanimous and the derivation of the referendum's outcome becomes more complicated.

Proposition 4 (The voting behaviour of the individuals if *m* = 0):

Suppose that $m = 0$ and that all individuals hold a referendum to decide whether the considered crime should be punished with (i) a fine, (ii) a prison term or (iii) not at all.

(a) If $H \geq 1$, the individuals choose unanimously that type of punishment which leads to the higher detection probability r, i.e., they vote for a prison term if $P > \overline{P}$ and for a fine if $P \leq \overline{P}$.

(b) If $H < 0.5$, the majority (i.e., at least 50% of all individuals) choose not to prosecute the crime, i.e., no sanction at all.

(c) If $0.5 \leq H < 1$, there are two cases which need to be distinguished:

(1) If H is large enough, so that $sHP > c'\left(\frac{1}{2HP}\right)$ holds, the majority of individuals choose a prison term.

(2) Otherwise, the majority choose not to prosecute the crime.

Proof: See Appendix.

Figure 5 illustrates the findings of Proposition 4 for a benevolent government with $s = 1$, using the same example as in the previous section ($H = 1$; $c(r) = -0.1\ln(1-r)$); it depicts the majority's decision contingent on the punishment P and the harm H. Note that the *prison* region contains a sub-region for large H in which individuals are indifferent between fine and prison, because $r_P^*(P) = r_F^*(P)$ (cf. Proposition 1(c) and Proposition 2(b)).

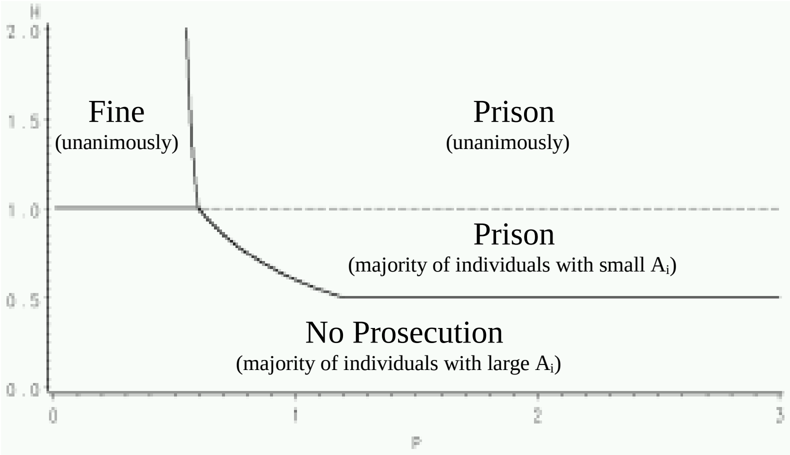

Figure 5: The result of the referendum on the type of punishment

For other values of the government's self-interest parameter *s*, the corresponding picture is quite similar. The two curved thresholds move somewhat towards the right if *s* decreases, so that the *prison* region shrinks while the *fine* and the *no-prosecution* regions expand. This implies that *less* crimes are punished by imprisonment if the government is *more* self-interested.

Note that Proposition 4 and Figure 5 only display the preferences of the majority. They do not give any information as to how strong these preferences are, i.e., how large the differences in the individuals' utilities between prison and fine are. As the detection probabilities $r_P^*(P)$ and $r_F^*(P)$ stay close together if the government is benevolent (s = 1, cf. Figure 3), individuals may well vote for imprisonment, but their utility is likely not to differ much between the two types of punishment. In contrast, if the government is strongly self-interested, i.e., if s << 1, the detection probabilities $r_P^*(P)$ and $r_F^*(P)$ lie far apart (cf. Figure 4), so that individuals' preferences between the two options are much stronger in general. Therefore, individuals will be more interested in fixing the type of punishment if the government is strongly self-interested.

Throughout the paper, I have assumed that each individual is able to pay the punishment *P* if it is a fine. (Note that this does not necessarily mean that all individuals are equally wealthy.) If we drop this assumption and assume instead that an individual with wealth w_i smaller than *P* pays a fine of w_i and is imprisoned for a period that corresponds to a monetary value of $P - w_i$ (if there is no mandatory prison sentencing), then all qualitative results of my model continue to hold. The government's choice of the detection probability if the punishment is a fine, $r_F^*(P)$, will move somewhat towards the corresponding choice if the punishment is a prison term, $r_P^*(P)$.

Another assumption I would like to discuss is that the utilities from crime A_i are uniformly distributed. For more general distributions *f(A)*, only partial results can be obtained. In particular, the government still chooses no prosecution for small punishments and full deterrence for large prison terms. For large fines, full deterrence is still established if and only if $sH > 1$. Hence, $r_F^* \leq r_P^*$ for large P and $r_F^* \geq r_P^*$ for small P, and, as a consequence, individuals will vote unanimously for imprisonment for large P and for a fine for small P if the harm is large enough – especially if the government is self-interested
($s < 1$). Moreover, the majority of individuals vote for no prosecution if the harm is smaller than the median of the A_i's. For intermediate values of harm or punishment, however, no general results of the referendum can be obtained.

5. Conclusions and further notes

This paper offers an economic explanation of why serious crimes are punished by mandatory prison sentencing even if offenders could pay an equivalent fine. In contrast to standard models of crime and punishment, in which a social planner chooses all the control variables, this paper considers a political economy model in which the decisions on the type of punishment and on the expenditure on policing are separated. In the first stage of the model, the citizens (or individuals) determine the type of punishment (prison or fine) in a referendum. Given this decision, the government chooses the expenditure on police and thereby the detection probability in the second stage of the model. The size of the punishment is treated as exogenous, in order to limit the model's complexity.

The model's main result is that individuals vote unanimously for imprisonment if the size of the punishment is large, whereas they vote for a fine if the size of the punishment is small (provided that the harm of the considered crime is sufficiently large). If the harm of the considered crime is small, the majority of individuals do not want the crime to be prosecuted at all. As more harmful crimes are usually punished with larger penalties (see, e.g., Polinsky and Shavell, 1979, 1991, and Stigler, 1970), this model can explain why, in practice, serious crimes are punished with prison terms whereas less serious crimes are fined.

In order to understand the result of the referendum in the first stage of the model, consider the government's choice of the detection probability given the type of punishment in the second stage. If the punishment is small, the government chooses a higher detection probability if the punishment is a fine than if it is a prison term. This is due to the fact that the number of criminals is quite high if the punishment is small. Therefore, increasing the detection probability increases the number of convictions and thereby the expected revenues from fines. Hence, fines serve as a monetary incentive to improve policing if the punishment is small. If the punishment is large, on the other hand, the government chooses a higher detection probability if the punishment is a prison term. The intuitive reason is that only few individuals commit the crime if the punishment is a large prison term. If the punishment is a large fine instead, the government has an incentive to choose a *lower* detection probability in order to *increase* the number of criminals. A higher number of criminals leads to more convictions and thereby to higher revenues from fines. Consequently, fines are an incentive to *reduce* policing if the punishment is large.

Now consider the first stage of the model. If the harm of the crime is large enough, the utility of every individual monotonically increases in the detection probability. Hence, they vote unanimously for that type of punishment which leads to the higher detection probability in the second stage: for imprisonment if the punishment is high and for a fine if the punishment is small. This result already holds if the government is benevolent, i.e., maximises social welfare. If the government places a higher weight on its own budget than on the individuals' utility, the government becomes more responsive to monetary incentives and the above described effects are much stronger. As a result, individuals will be strongly interested in fixing the type of punishment if the government is self-interested.

This discussion raises the question why citizens only choose the type of punishment and not the detection probability as well. The main reason is that it is quite easy to enforce the type and the size of punishment but nearly impossible to enforce a given detection probability. Moreover, citizens would have to know the cost function, i.e., the prosecution technology, in order to make a good decision. And finally, the technologies of crime and prosecution are likely to change over time, so that a fixed detection probability might not be optimal.

Note that the present model does not directly take into account that the government is elected regularly by the citizens. Admittedly, general elections are the chief means by which citizens exercise power over the government, but individuals must decide on a complex bundle of different issues in a single election. Even if crime prevention and prosecution usually are important subjects during election campaigns, general elections have only limited influence on this issue. Moreover, the present model indirectly takes general elections into consideration, as the government's objective function includes social welfare. In this context, the assumption that the government is self-interested can be restated as follows: The government can increase the probability of re-election by diverting some money from the criminal justice system to some other policy area, which is "more productive" (in terms of votes for a re-election) than police.

Finally, note that the present paper only investigates the deterrence effect of imprisonment. It does not take into account that imprisonment also prevents crime, because notorious criminals cannot commit any further crimes while they are imprisoned. However, Levitt (1998a) presents empirical evidence that deterrence is more important than incapacitation in reducing crime, particularly in the case of property crime. Furthermore, Levitt (1998b) and Kessler and Levitt (1999) confirm that imprisonment has a strong deterrence effect.

Appendix

Proof of Proposition 1

The government solves the maximisation problem

$$\max_r\{-sqH - c(r)\} \text{ with } q = \begin{cases} 0, & \text{if } rP > 1 \\ 1 - rP, & \text{if } rP \le 1 \end{cases} \text{ subject to } 0 \le r \le 1 \tag{A1}$$

Obviously, $r > \frac{1}{P}$ cannot be optimal because increasing *r* beyond 1/*P* increases costs without affecting the benefit of prosecution. Hence, problem (A1) can be transformed by setting $q = 1 - rP$ and introducing a new constraint:

$$\max_r\{-sH + sHrP - c(r)\} \text{ subject to } 0 \le r \le \frac{1}{P}$$

The condition $r \le 1$ has been dropped, because it always holds due to the assumption $\lim_{r\to 1} c(r) = \infty$.

Assume that r fulfils the two constraints. The derivation of the objective function with respect to *r* and equating it to zero gives

$$c'(\hat{r}) = sHP. \tag{A2}$$

Since the cost function *c(·)* is convex, condition $r \ge 0$ is fulfilled if $c'(0) \le c'(\hat{r}) = sHP \Leftrightarrow P \ge \frac{c'(0)}{sH}$.

Hence, $r_P^* = 0$ is optimal if $P < \frac{c'(0)}{sH}$.

Condition $r \le \frac{1}{P}$ is satisfied if $c'\left(\frac{1}{P}\right) \ge c'(\hat{r}) = sHP$, i.e., if $P \le \tilde{P}$ with $c'\left(\frac{1}{\tilde{P}}\right) = sH\tilde{P}$. Therefore, $r_P^* = 1/P$, if $P \ge \tilde{P}$. Moreover, if $\frac{c'(0)}{sH} \le P < \tilde{P}$, r_P^* is given by (A2). ■

Proof of Proposition 2

The government solves the maximisation problem

$$\max_r\{-sqH + qrP - c(r)\} \text{ with } q = \begin{cases} 0, & \text{if } rP > 1 \\ 1 - rP, & \text{if } rP \le 1 \end{cases} \text{ subject to } 0 \le r \le 1$$

As in the proof of Proposition 1, this problem can be transformed to:

$$\max_r\{-sH + (sH+1)rP - r^2P^2 - c(r)\} \text{ subject to } 0 \le r \le \frac{1}{P}$$

Assume that r fulfils the two constraints. The derivation of the objective function with respect to *r* and equating it to zero gives

$$c'(\tilde{r}) + 2P^2\tilde{r} = (sH+1)P. \tag{A3}$$

Consider $P_0 = \frac{c'(0)}{1+sH}$. If $P > P_0$, (A3) can only hold if $\tilde{r} > 0$, i.e. $\tilde{r}(P) > 0$ if $P > P_0$. On the other hand, if $P < P_0$, (A3) can only hold if $\tilde{r} < 0$, i.e. $\tilde{r}(P) < 0$ if $P < P_0$. Hence, due to the constraint $r \ge 0$, $r_F^* = 0$ if $P < P_0$. Otherwise, (A3) does not violate this constraint.

Consider $\hat{P}$ which solves

$$c'\left(\frac{1}{\hat{P}}\right) + 2\hat{P} = (sH+1)\hat{P} \Leftrightarrow c'\left(\frac{1}{\hat{P}}\right) = (sH-1)\hat{P} \tag{A4}$$

If $sH \leq 1$, $\hat{P}$ does not exist. Since $\tilde{r}(\cdot)$, as implicitly defined by (A3) is a continuous function on $[0, \infty)$ and $\tilde{r}(0) < \infty$, $\tilde{r}(P) < 1/P$ for all P in this case. Hence, in this case, r_F^* is given by (A3) if $P \geq P_0$.

If $sH > 1$, $\hat{P}$ is uniquely defined by (A4). As $\tilde{r}(\cdot)$ is continuous and $\tilde{r}(P) < 1/P$ for small P, $\tilde{r}(P) < 1/P$ for all $P < \hat{P}$ and $\tilde{r}(P) \geq 1/P$ for all $P \geq \hat{P}$. Hence, r_F^* is given by (A3) if $P_0 \leq P < \hat{P}$ and $r_F^* = 1/P$ if $P \geq \hat{P}$. ■

Proof of Corollary 1:

I first derive the $\overline{P}$ for which $\hat{r}(\overline{P}) = \tilde{r}(\overline{P})$. $\hat{r}(P)$ and $\tilde{r}(\overline{P})$ are implicitly defined by (A2) and (A3), respectively. (A2) and (A3) induce $sH\overline{P} + 2\overline{P}^2\hat{r} = (sH+1)\overline{P} \;\Leftrightarrow\; 2\overline{P}^2\hat{r} = \overline{P} \;\Leftrightarrow\; \hat{r} = 1/(2\overline{P})$

$$\Leftrightarrow \quad (c')^{-1}\{sH\overline{P}\} = 1/(2\overline{P}) \quad \Leftrightarrow \quad c'(1/(2\overline{P})) = sH\overline{P} \tag{A5}$$

Note that $\overline{P}$ is uniquely defined. Moreover, one immediately obtains (due to $c''(\cdot) > 0$): $\overline{P} < \tilde{P} < \hat{P}$, i.e., $\overline{P}$ does not lie in the *full-deterrence* region. Furthermore, $\overline{P}$ does not fall into the *no-prosecution* region, as $\overline{P} = \dfrac{c'(1/(2\overline{P}))}{sH} > \dfrac{c'(0)}{sH}$. Hence, $\overline{P}$ is the unique intersection of the functions $r_P^*(P)$ and $r_F^*(P)$ in the *crime-and-prosecution* region. ■

Proof of Proposition 3

Assume that $s < 1$ or $m < 1$ and choose $H \in \left]\max\left\{\dfrac{m}{2-sm}, \dfrac{1}{2}\right\}, \dfrac{1}{s}\right[$ and (for the time being) $P > \tilde{P}$. Then, $r_P^*(P) = \dfrac{1}{P}$ and $r_F^*(P) < \dfrac{1}{P}$ (due to Propositions 1(c) and 2(b)), i.e., we have full deterrence if the punishment is a prison term and less than full deterrence if the punishment is a fine. The utility of individual i in the cases of no prosecution (0), imprisonment (P) and fine (F) is given by $U_{i0}(P) = A_i - H$, $U_{iP}(P) = -mc(1/P)$, and $U_{iF}(P) = \max\{A_i - r_F^*(P)P, 0\} - (1 - r_F^*(P)P)H + m\left[(1 - r_F^*(P)P)r_F^*(P)P - c(r_F^*(P))\right]$.

First observe, that $\lim_{P\to\infty} U_{iP}(P) = 0$ whereas $U_{i0}(P)$ is independent of P and (due to H > ½) negative for more than 50% of individuals. Hence, there is a P_1 such that $U_{iP}(P) > U_{i0}(P)$ for more than 50% of all individuals for all $P > P_1$.

From Proposition 2(c) we know that

$$r_F^*(P)P = \frac{sH+1}{2} - \frac{c'(r_F^*(P))}{2P} \overset{P\to\infty}{\rightarrow} \frac{sH+1}{2}, \tag{A8}$$

because $r_F^*(P) \overset{P\to\infty}{\rightarrow} 0$. I first consider the case that $s > 0$. Then $(sH + 1)/2 >$ ½, so that there is a P_2 such that more than 50% of all individuals decide not to commit the crime if the punishment is a fine and $P > P_2$. Now assume that $P > P_2$ and consider the difference in utilities between imprisonment and fine for these 50%+ non-criminals:

$$\begin{aligned} U_{iP}^N(P) - U_{iF}^N(P) &= -mc(1/P) + (1 - r_F^*(P)P)H - m\left[(1 - r_F^*(P)P)r_F^*(P)P - c(r_F^*(P))\right] \\ &= m(1 - r_F^*(P)P)(H/m - r_F^*(P)P) - m\left[c(1/P) - c(r_F^*(P))\right] \end{aligned} \tag{A9}$$

Note that $H > m/(2 - sm)$ induces $(sH + 1)/2 < H/m$. Together with (A8) this implies that there is a P_3, such that $r_F^*(P)P < H/m$ for all $P > P_3$. If we assume that $P > P_3$, we therefore get

$$U_{iP}^N(P) - U_{iF}^N(P) > m\left(\frac{1-sH}{2}\right)\left(\frac{H}{m} - r_F^*(P)P\right) - m\left[c\left(\frac{1}{P}\right) - c(r_F^*(P))\right].$$

Note that $m\left[c(1/P) - c(r_F^*(P))\right] \underset{P\to\infty}{\to} 0$, whereas $m\left(\frac{1-sH}{2}\right)\left(\frac{H}{m} - r_F^*(P)P\right) \underset{P\to\infty}{\to} m\left(\frac{1-sH}{2}\right)\left(\frac{H}{m} - \frac{sH+1}{2}\right) > 0$ if $m > 0$. Hence, there is a P_4, such that $U_{iP}^N(P) - U_{iF}^N(P) > 0$ for all $P > P_4$. If $m = 0$, $U_{iP}^N(P) - U_{iF}^N(P) > 0$ follows directly from (A9).

This proves that for every $H \in \left]\max\left\{\frac{m}{2-sm}, \frac{1}{2}\right\}, \frac{1}{s}\right[$ and $P > \max\left\{\tilde{P}, P_1, P_2, P_3, P_4\right\}$ at least 50% of all individuals strictly prefer imprisonment if $s > 0$. If $s = 0$, we know from the above argument that asymptotically 50% of all individuals are non-criminals and will vote for imprisonment. Now choose $\varepsilon < m\left(\frac{1-sH}{2}\right)\left(\frac{H}{m} - \frac{sH+1}{2}\right)$ and consider the individuals with $A_i \in [0.5 - \varepsilon, r_F^*(P)P]$. These individuals will commit the crime if the punishment is a fine and will asymptotically receive a net utility smaller or equal to ε from their criminal activity. Due to the construction of ε, there is a P_5 such that these individuals will prefer imprisonment to a fine if $P > P_5$. Hence, if $P > P_0 \equiv \max\left\{\tilde{P}, P_1, P_2, P_3, P_4, P_5\right\}$, 50% + ε strictly prefer imprisonment to a fine or to no prosecution, which proves the first statement of Proposition 3.

Now consider the case that $s = 1$ and $m = 1$. Then the utility function of non-criminals is exactly the same as the government's utility function, so that non-criminals would never strictly prefer imprisonment to a fine.

For criminals we have: $U_{i0}^C(P) = A_i - H$, $U_{iP}^C(P) = A_i - r_P^* P - qH - c(r_P^*)$ and $U_{iF}^C(P) = A_i - (1-q)r_F^* P - qH - c(r_F^*)$. If $r_P^* \geq r_F^*$, then $U_{iF}^C(P) > U_{iP}^C(P)$, i.e., criminals prefer fines to imprisonment. If $r_P^* < r_F^*$, on the other hand, then either $r_P^* = 0$, in which case $U_{i0}^C(P) = U_{iP}^C(P)$, or $c'(r_P^*) = HP$ (cf. Proposition 1(b)). Integrating up this last equation with respect to r over $[0, r_P^*]$ gives: $c(r_P^*) = HPr_P^*$. Consequently,

$$U_{i0}^C(P) - U_{iP}^C(P) = r_P^* P - r_P^* PH + c(r_P^*) = r_P^* P + (1 - r_P^* P)H + c(r_P^*) - H > 0.$$

Thus, no criminal strictly prefers imprisonment to no prosecution. This proves the second statement of Proposition 3. ■

Proof of Proposition 4

Assume that individuals can choose between r_1, r_2 and 0 with $0 < r_1 < r_2$. I first show that no individual votes for r_1. To this end, I consider 3 cases:

(1) An individual does not commit the crime for $r = r_1$ (and, consequently, for $r = r_2$). Since $U_i(r) = -(1-rP)H = rPH - H$ for $r \geq r_1$, such an individual will prefer r_2 to r_1.

(2) An individual does commit the crime for all three values of *r*. Since $U_i(r) = A_i - rP - (1-rP)H$ $= rP(H-1) + A_i - H$ for $r \leq r_2$, such an individual will prefer r_2 to r_1 if $H \geq 1$ and 0 over r_1 if $H < 1$.

(3) An individual does not commit the crime for $r = r_2$ but he commits the crime for $r = r_1$. Since $U_i(r) = rP(H-1) + A_i - H$ for $r \leq r_1$, such an individual will prefer 0 to r_1 if $H < 1$. If $H \geq 1$, he will prefer r_2 over r_1, because $U_i(r_2) - U_i(r_1) = r_2PH - r_1P(H-1) + A_i > (r_2 - r_1)P(H-1) \geq 0$, as $A_i < r_2P$.

Therefore, r_1 is never voted for and we can restrict our analysis to $r = 0$ and $r = \max\{r_P^*, r_F^*\} \equiv \tilde{r}$.

(a) Suppose $H \geq 1$. The gain in utility of a "non-criminal" from $\tilde{r}$ over 0 is $\Delta U_i^N = U_i^N(\tilde{r}) - U_i^N(0)$ $= -(1-\tilde{r}P)H - A_i + H \geq \tilde{r}HP - \tilde{r}P \geq 0$, as $A_i \leq \tilde{r}P$. The gain in utility of a "criminal" is $\Delta U_i^C = U_i^C(\tilde{r}) - U_i^C(0) = A_i - \tilde{r}P - (1-\tilde{r}P)H - A_i + H = (H-1)\tilde{r}P \geq 0$, so all individuals prefer $\tilde{r}$ to 0.

(b) Assume $H < 0.5$. As $\Delta U_i^C = (H-1)\tilde{r}P < -\dfrac{\tilde{r}P}{2}$, criminals vote for $r = 0$. For non-criminals, we obtain: $\Delta U_i^N = \tilde{r}PH - A_i < \tfrac{1}{2}\tilde{r}P - A_i \leq \tfrac{1}{2} - A_i$. Hence, $\Delta U_i^N < 0$ if $A_i \geq \tfrac{1}{2}$, which means that at least 50% of all individuals vote for $r = 0$. (Only those individuals with $A_i < \tfrac{1}{2}$ can possibly prefer $\tilde{r}$.)

(c) Let $\tfrac{1}{2} \leq H < 1$. Criminals never vote for $\tilde{r}$. Therefore, $\tilde{r}$ can only be chosen if more than 50% of all individuals are non-criminals and better off with $\tilde{r}$ compared with $r = 0$. At least 50% of all individuals are non-criminals if $\tilde{r}P \geq \tfrac{1}{2}$. Then, at least 50% of all individuals are better off with $\tilde{r}$, if $\Delta U_i^N = \tilde{r}PH - A_i \geq 0$ for 50% of all individuals *i*, i.e., if $\tilde{r}PH \geq \tfrac{1}{2}$. As $\tilde{r}P > \tilde{r}PH$, at least 50% of all individuals vote for $\tilde{r}$ if

$$\tilde{r}PH \geq \tfrac{1}{2}. \tag{A6}$$

(1) Let $P > \tilde{P}$, then $\tilde{r} = r_P^*$ and $\tilde{r}P = 1$. Together with $H \geq \tfrac{1}{2}$ this induces (A6). Hence, for $\tfrac{1}{2} \leq H < 1$ and $P > \tilde{P}$ the majority will vote for $\tilde{r}$.

(2) Let $\overline{P} \leq P \leq \tilde{P}$, then $\tilde{r} = r_P^*$ and $\tilde{r}$ is given by (A2), i.e., $\tilde{r} = (c')^{-1}\{sHP\}$. Substituting in (A6) gives

$$c'\left(\frac{1}{2HP}\right) \leq sHP, \tag{A7}$$

which holds if *H* is "large enough".

(3) Let $\tilde{H}$ be the smallest *H* which satisfies (A7), i.e., $c'\left(\dfrac{1}{2\tilde{H}P}\right) = s\tilde{H}P$ and consider $P = \tilde{P}$, i.e., $c'\left(\dfrac{1}{P}\right) = s\tilde{H}P$. We obtain: $c'\left(\dfrac{1}{P}\right) = c'\left(\dfrac{1}{2\tilde{H}P}\right) \Leftrightarrow \dfrac{1}{P} = \dfrac{1}{2\tilde{H}P} \Leftrightarrow \tilde{H} = \tfrac{1}{2}$. Hence, at $P = \tilde{P}$, (A7) holds if and only if $H \geq \tfrac{1}{2}$. If *P* increases, inequality (A7) becomes less tight. Therefore, (A7) is satisfied for all $H \in [\tfrac{1}{2}, 1)$ and $P > \tilde{P}$. Consequently, cases (1) and (2) can be merged.

(4) Let $P < \overline{P}$, then $\tilde{r} = r_F^*$ and $\tilde{r}$ is given by (A3). I first show that the deterrence $\tilde{r}P$ increases monotonically in P over this region. Deriving $c'(\tilde{r}(P)) + 2P^2\tilde{r}(P) - (sH+1)P = 0$ with respect to

P and equating it to zero yields $c''(\tilde{r})\tilde{r}'(P) + 4P\tilde{r} + 2P^2\tilde{r}'(P) - (sH+1) = 0$ $\Leftrightarrow$ $\tilde{r}'(P) = \dfrac{sH + 1 - 4P\tilde{r}}{c''(\tilde{r}) + 2P^2}$.

Therefore, $\dfrac{\partial\tilde{r}(P)P}{\partial P} = \tilde{r}'(P)P + \tilde{r}(P) = \dfrac{(sH + 1 - 4P\tilde{r})P}{c''(\tilde{r}) + 2P^2} + \tilde{r}$. Now assume, $\tilde{r}P$ has a maximum:

$\dfrac{\partial\tilde{r}(P)P}{\partial P} = 0 \quad \Leftrightarrow \quad (sH+1)P - 4P^2\tilde{r} = -\tilde{r}c''(\tilde{r}) - 2P^2\tilde{r} \quad \Leftrightarrow \quad \tilde{r}c''(\tilde{r}) + (sH+1)P = 2P^2\tilde{r}$. As $c''(\cdot) > 0$, it follows $2P^2\tilde{r} > (sH+1)P \quad \Leftrightarrow \quad \tilde{r}P > \dfrac{sH+1}{2}$. However, (A3) and $c'(\cdot) > 0$ induce that $\tilde{r}P < \dfrac{sH+1}{2}$. Hence, $\tilde{r}P$ does not have any extremum and increases monotonically. Therefore, $\tilde{r}P$ is maximal at $P = \overline{P}$ in the region $P \leq \overline{P}$.

Now consider $P = \overline{P}$. At this point, $r_F^* = r_P^*$ and the minimal H which leads to a majority for $\tilde{r}$ is given by $c'\left(\dfrac{1}{2HP}\right) = sHP$ (see (A7)). Together with (A5), the equation which defines $\overline{P}$, we get $c'\left(\dfrac{1}{2HP}\right) = c'\left(\dfrac{1}{2P}\right) \quad \Leftrightarrow \quad \dfrac{1}{2HP} = \dfrac{1}{2P} \quad \Leftrightarrow \quad H = 1$. Hence, at $P = \overline{P}$, condition (A6) is only satisfied if $H \geq 1$. If $P < \overline{P}$, $\tilde{r}P < \tilde{r}\overline{P}$ so that H must be even larger to fulfil (A6). Altogether, this proves that for $\frac{1}{2} \leq H < 1$ and $P < \overline{P}$ the majority of individuals vote for $r = 0$. ■

References

ANDREONI, JAMES (1991): "Reasonable doubt and the optimal magnitude of fines: should the penalty fit the crime?", *RAND Journal of Economics*, 22, pp. 385-395

BECKER, GARY S. (1968): "Crime and punishment: An economic approach", *Journal of Political Economy*, 76, pp. 169-217

CHU, C. Y. CYRUS AND NEVILLE JIANG (1993): "Are fines more efficient than imprisonment?", *Journal of Public Economics*, 51, pp. 391-413

DITTMANN, INGOLF (1999): "How reliable should auditors be?: optimal monitoring in principal-agent relationships", *European Journal of Political Economy*, 15, pp. 523-546

DITTMANN, INGOLF (2000): "Imprisonment versus fines: A theoretical perspective", in: Ziggy MacDonald and David Pyle: *Illicit Activity: An Economic Analysis of Crime, Drugs and Fraud*, Ashgate, Aldershot, forthcoming

FRIEDMAN, DAVID AND WILLIAM SJOSTROM (1993): "Hanged for a sheep - the economics of marginal deterrence", *Journal of Legal Studies*, 22, pp. 345-366

KESSLER, DANIEL AND STEVEN D. LEVITT (1999): " Using sentence enhancements to distinguish between deterrence and incapacitation", *Journal of Law and Economics*, 42, pp. 343-363

LEVITT, STEVEN D. (1997): "Incentive compatibility constraints as an explanation for the use of prison sentences instead of fines", *International Review of Law and Economics*, 17, pp. 179-192

LEVITT, STEVEN D. (1998a): "Why do increased arrest rates appear to reduce crime: deterrence, incapacitation, or measurement error?", *Economic Inquiry*, 36, pp. 353-372

LEVITT, STEVEN D. (1998b): "Juvenile crime and punishment", *Journal of Political Economy*, 106, pp. 1156-1185

MOOKHERJEE, DILIP AND I. P. L. PNG (1994): "Marginal deterrence in enforcement of law", *Journal of Political Economy*, 102, pp. 1039-1066

POLINSKY, A. MITCHELL AND STEVEN SHAVELL (1979): "The optimal tradeoff between the probability and magnitude of fines", *American Economic Review*, 69, pp. 880-891

POLINSKY, A. MITCHELL AND STEVEN SHAVELL (1984): "The optimal use of fines and imprisonment", *Journal of Public Economics*, 24, pp. 89-99

POLINSKY, A. MITCHELL AND STEVEN SHAVELL (1991): "A note on optimal fines when wealth varies among individuals", *American Economic Review*, 81, pp. 618-621

STIGLER, GEORGE J. (1970): "The optimum enforcement of laws", *Journal of Political Economy*, 78, pp. 526-536