Faculty of Economics

Psychic Punishment Costs and Deterrence

by
Barna Bakó, Péter Isztin

Psychic Punishment Costs and Deterrence

Barna Bakó[*]   Péter Isztin[†]

August 1, 2016

Abstract

In this paper we analyze criminal deterrence in the presence of specific psychic costs of punishments. We consider a dynamic model with three players, analyzing the choices of a representative lawmaker, potential criminal and judge. In our setting the lawmaker decides whether to introduce a fixed punishment enhancement above a chosen threshold of crime level, depending on its popularity among the voters. In reaction, the judge, who is influenced by her own preferences as well as the opinion of her peer group, might change the probability of punishment, through affecting the standard of reasonable doubt. Our results suggest that large discontinuous and mandatory increases in punishment can have unintended effects that are contrary to the stated goal of such punishment enhancements. In equilibrium, when either the judge or her peer group is "anti-punishment" enough, the level of criminal activity might increase in response to the punishment enhancement. This perverse effect is less likely to occur if there is a higher number of peer groups within the "elite", so that a greater extent of self-selection by judges can occur. Our results have relevance for a number of areas outside the traditional criminal justice system as well, such as special courts (such as ecclesiastical or military courts), or the strictness and enforcement of regulations.

JEL codes: K12, K42, D81

Keywords: crime, deterrence, punishment, peer effects

---

[*]MTA-BCE 'Lendület' Strategic Interactions Research Group, Department of Microeconomics, Corvinus University of Budapest, Fővám tér 8, E225-A, Budapest, 1093, Hungary, e-mail: barna.bako@uni-corvinus.hu.

[†]Department of Microeconomics, Corvinus University of Budapest, Fővám tér 8, E225, Budapest, 1093, Hungary, e-mail: isztin.peter@uni-corvinus.hu.

1 Introduction

The economic analysis of crime and punishment goes back to Becker (1968)'s seminal paper, which forms the basis of a rational choice account of criminal behavior. Since then there has been a large empirical literature testing the basic predictions of the economic approach to crime, most of which have been validated (see e.g. Levitt, 2004; Durlauf and Nagin, 2011). For instance, it has been shown that criminals respond to both increases in the amount of punishment as well as the probability of punishment, although the elasticity is ususaly larger for the latter than for the former. While the literature on deterrence is wide-ranging, there has been little research on the determinants of punishment, as opposed to crime. Among the few forays into this area we can list Andreoni (1991), Glaeser et al (2000) or Glaeser and Sacerdote (2000).

In this paper we attempt to merge two strands of literature: we analyze criminal deterrence, but in the context of social and political incentives both in the case of lawmakers and judges. In doing that we build on the literature on social interactions (see e.g. Becker and Stigler, 1974; Glaeser and Sacerdole, 2003; Brock and Durlauf, 2001; Cicala et al, 2015), and in particular, the literature on peer pressure and esteem and disesteem (Brennan and Pettit, 2005; Cowen, 2002; Cowen and Glazer, 2007). Moreover, we also build on the political economy and public choice literatures in assuming that lawmakers do not maximize social welfare, but their private utility. In this paper we aim to extend economic analysis to the behavior of law enforcers as well as lawmakers in a way that does not assume them to maximize some social welfare function. The social welfare maximizing approach is useful when analyzing the welfare implications of crime and punishment, but it is less suitable when one wishes to understand the behavior of actual decision makers in criminal policy and law enforcement.

The quest for the approval of others and other social interactions have featured prominently in the works of early economists, such as Smith (1790). In the later economic literature, however, such topics have been overlooked, up until the past few decades. The political economy of punishment also has a long history, but again, modern economic literature has been slow to get started on this topic. One motivation of our paper has been the intuition provided by Mill (1868) in his speech in defense of capital punishment, where he notes that the deterrent effect of capital punishment can be compromised when judges and juries are too reluctant to apply it. A similar, but more formal statement is found in Andreoni (1991), in which the author claims that higher penalties may encourage criminal activity. This perverse result is due to the fact that judges are interested in being right or "just" in their judgement. When the severity of punishment increases the cost of wrongful conviction also increases, therefore the judge raises the threshold of evidence for a guilty verdict, thus lowering the probability of conviction. This can lead to the surprising result that stricter punishment can induce more crime. Yet, in our view, Andreoni's model is inadequate in two respects: first, adjusting the probability of conviction in response to higher probability of a wrongful conviction can actually enhance deterrence as wrongful convictions may weaken deterrence. Second, judges might have an incentive to lower the probability of conviction even apart from caring about wrongful convictions.

Our paper also relates to the literature of credible threats and commitments, starting with Schelling (1960) which highlights the role of time inconsistency within agents. Closer to our approach in this regard is the research by Acemoglu and Robinson (2009) who emphasize the commitment problems in politics, arising from the fact that agreements between decision-makers often does not result in perfect Coasian solutions as these agreements suffer from a time inconsistency problem. Typically, credible third-party enforcements are rare in politics, thus, promises and threats are hard to make credible. In our view similar problems can arise in judicial decision-making, especially when the judicial branch of the government is independent from the executive branch, and therefore cannot be directly incentivized by the latter. In our model there are multiple agents involved in deterrence, and it is the differences in incentives faced by these agents that cause the potential problems in deterrence credibility.

The paper is organized as follows. First, we introduce the problem of punishment costs and other private incentives faced by lawmakers and judges in relation to punishment, followed by a model of criminal deterrence with three period. In period 1, the lawmaker decides whether to introduce an enhanced form of punishment, and she chooses the threshold above which the enhanced punishment must be applied by the judge. In the second period, the potential criminal chooses a crime level, her utility depending on the gross benefit of crime as well as on the expected punishment. In period 3, the judge makes a decision regarding the probability of punishment through choosing a threshold of evidence. We assume that the judge faces positive psychic (social) punishment costs, stemming largely from the potential disesteem by other members of her peer group. One of our main result suggests that the level of criminal activity may increase in response to the punishment enhancement. In section 5 we provide some examples which our framework can be applied for and finally we conclude.

2 Incentives in punishment

The standard economic model of crime and punishment (see Becker, 1968) focuses on the behavioral responses a criminal might have to the changing costs and benefits of the criminal activity. However, it does not extends the analysis to the incentives faced by the actors in the penal system. Therefore, the literature that followed Becker has typically assumed, explicitly or implicitly, that lawmakers, judges and prosecutors maximize a utilitarian social welfare function and aim to determine optimal punishments. Much of these literature has been concerned with what the optimal punishment should be in a given situations and on how criminals respond to different punishment regimes, thus helping policy-makers choose the right approach to criminal policy. However, accounting for the incentives of judicial and political decision-makers may be important for a number of reasons. For example, the interaction of these incentives might not produce criminal sentences that are consistent with optimal punishment theory. This question has been studied by Glaeser and Sacerdote (2000), who conclude that in a number of cases the criminal justice system does not work according to basic optimal punishment criteria. The au-

thors show, for example, that black criminals typically receive higher sentences while murders of black victims on average are followed by lower sentences than murders of white ones. Moreover, the murder of a woman is typically punished with a longer sentence than the murder of a man (see Glaeser and Sacerdote, 2000). Such results indicate that the incentives of participants in the criminal justice system might have an important impact on the criminal sentences and such should be taken seriously.

Since Becker (1968)'s paper there has been a decades old debate about the deterrent effect of capital punishment. For example, Ehrlich (1973) in one of the first modern econometric study of capital punishment claims that each execution deters 8 murders. Among else, this result has lead to a reconsideration of criminal policy and many states in the United States reinstated capital punishment into their books. In a more recent paper, however, Donohue and Wolfers (2005) argue that much of the empirical work in the topic is highly flawed, as the results are very sensitive to changes in specifications. Yet, this might be driven by the fact that presently there are very few executions in the United States, and so researchers simply have too little data to separate "normal" fluctuation in murder rates from changes due to deterrence from executions. So while there are studies indicating significant deterrent effects, there are several researchers as well as public figures who are not persuaded by this evidence.

There are other cases as well where there is some skepticism toward the efficacy of tough sentencing laws. Often, the skeptics provide explanations based on short-sighted or outrightly irrational behavior on the part of criminals. In this paper we offer an explanation where criminals respond perversely to the introduction of a (mandatory) punishment enhancement precisely because they are rational and forward-looking enough. The mechanism we are about to elaborate and model is roughly the following. Suppose that over a specific threshold of evidence (of the nature of crime) a defendant has to receive an enhanced punishment. We assume that judges have some power over determining this threshold of evidence, that is, the judge has some influence over the probability that defendants are found guilty. Among else, we will concentrate on the judge's decision of this probability. It is traditionally assumed that the probability of punishment and the severity of punishment are independent of each other. In this paper, however, we will show that under certain circumstances this might not be the case. If the punisher faces some, either psychic or reputational costs from a very strict sentence, she might have an incentive to lower the probability of conviction. However, if would-be criminals foresee this behavior, than they will also adjust their behavior accordingly. While the introduction of a "tough" punishment enhancement increases the severity of punishment, lowering the probability of conviction can offset this effect. Depending on its magnitude, it can even lead to a decrease in the expected cost of committing a crime that is followed by that particular punishment enhancement. If this is so, then the would-be criminal has a greater incentive to commit a corresponding level of crime. The reason such a scenario might come about is that criminal policy-makers and judges face different incentives: policymakers might face a strongly pro law and order public opinion, so they want to ensure that the most serious crimes are met with a serious punishment. On the

other hand, judges may be more influenced by the opinion of their peer group, whose members might disapprove of punishments which are considered by them too strict. Therefore, applying the penalty is personally costly for them. Note, that we do not describe a relationship between strictness and probability of punishment so that probability always changes as the applicable punishment changes. Instead, we assume that the change in probability occurs when there is a jump in the severity of punishment after a given threshold. Continuous increases in punishment are, in our model, proportional to increases in the crime level, while the discontinuous change after the established threshold is not. The latter change is therefore both more salient and possibly more outrageous or delighting to the audience (voters, to the lawmaker and peers, to the judge).

Lawmakers of course usually do not restrict the judiciary as strongly as we have assumed in the above thought experiment. They typically determine an interval within which judges are usually expected to apply the middle-way punishment. However, if policy-makers know that, and therefore expect judges to apply the penalty only very rarely, they might still tie the judges' hands by specifying certain cases where capital punishment needs to be applied.

3 Attitudes in crime and punishment

A motivating observation behind our model is that attitudes toward crime and punishment can affect both lawmaking as well as the justice system by electoral incentives in the former and approbation or social incentives in the latter case. Such attitudes often strongly differ between different parts of the population and can be thought of as ideological groups (Haidt, 2012), or, as we assume in this article, there can be a difference in the attitudes of different peer groups. For example, in many European countries, a majority of the population supports certain forms of strict punishments such as capital punishment or life imprisonment without parole as well as generally having a "tough on crime" attitude, while members of the intellectual "elite" often hold diametrically opposite views on these matters (see e.g. Whitehead et al, 1999). Moreover, Baum (2007) provides compelling evidence that judges are influenced by the opinions of their social groups, and these opinions often reflect "elite" preferences.

It would be interesting to examine how these different views form, although this paper cannot embark on this task. However, it can still be useful to consider some possible mechanisms. There are ample evidence that political and societal views are partly determined by personality traits, many of which are likely to be innate (see e.g. Haidt, 2012). Personality traits also correlate with education and income, thus establishing correlations between belonging to some "elite" group and having certain beliefs. However, it is also possible that societal beliefs form endogenously. For instance, certain views which are salient enough can serve as signals about the personality, or general political inclination of an individual. Members of particular groups might wish to coordinate along the same or similar beliefs in order to signal their membership in that group. In order for such a coordination to be possible, the said view should be salient or focal

enough, following Schelling (1960)'s concept of focal points. A possible further consequence can be that beliefs can at times be especially stark, and possibly starker than it would be optimal. Within certain social groups there can also occur some social learning (see e.g. Glaeser and Sunstein, 2009) where group members learn new information primarily from their fellow group members and this leads to large between-group differences and smaller within-group differences in beliefs. Social influence therefore can be purely informational but it can also affect attitudes and political preferences, or simply the expression of such preferences. Kuran (1997) documents the phenomenon of "preference falsification", the practice abundant across human societies, where members of a group are not willing to express views different from the perceived majority opinion, leading to a "pluralistic ignorance", where the majority might not hold a particular view anyone, yet group members still believe it does so they do not challenge it, as public deviation from a "sacred" opinion is costly.

4 The Model

In this section we analyze a simple sequential-move game with 3 periods. First the lawmaker decides whether to introduce a particular form of punishment and sets a threshold of crime level above which the punishment has to be applied by the judge. In the second period, the potential criminal (from now on we simply refer to her as criminal) chooses a level of crime and in the third period the judge determines the probability of conviction, which is her only choice variable, as the applicable punishment is determined by the lawmaker.

4.1 The lawmaker's problem

The lawmaker chooses the threshold $\bar{x}$, depending on the following considerations: we assume that the lawmaker is indifferent toward social welfare and will not choose $\bar{x}$ so to maximize a social welfare function, instead, she derives utility (or rather: she evades disutility) by choosing according to the preferences of the median voter. Moreover, we assume that the lawmaker is also influenced by the preferences of the median voter of her party which may differ from the former one because of ideological commitments.

The lawmaker makes two choices: first decides whether to introduce a punishment enhancement α, and after a threshold, $\bar{x}$, for which the enhancement needs to be applied.

The lawmakers's utility function is given with

$$U(\bar{x}) = \psi - t_l(x^v - \bar{x})^2 - z_l(x^p - \bar{x})^2 \tag{1}$$

where ψ is the utility she derives from the introduction of the enhancement, x^v is the preferred threshold of the median voter and x^p is the preferred threshold of the median party voter (from now on we will refer to it as the party median). We assume that $0 \le \bar{x} \le 1, 0 \le x^v \le 1, 0 \le x^p \le 1$. We do not make any restrictions on either the magnitude of the sign of ψ, so it could be

positive as well as negative. We also note that ψ captures the preferences of both the general as well as the party median.

The intuition behind the above formalization is that both voters among the general population and loyal party voters care about the lawmaker's decision in criminal affairs. We do not specify the preferences of the party voters, and neither do we fully specify other voters' preferences. We assume that voting is the simple majority voting. This allows us to concentrate on the effect of the median voter's preferences. Furthermore, for the party voters' preferences we also take the median. If $\psi > 0$, merely introducing the punishment enhancement α, the lawmaker gets rewarded by the median voter. However, if she does not choose the threshold level favored by the median voter or the median party voter, she receives some punishment, which depends on the distance between her choice and the median voter's preference and takes a strictly convex form. We denote with t_l and z_l the parameters determining how much the lawmaker cares about deviating from the median voter's position x^v and the party median x^p, respectively.

As for the first choice, the lawmaker introduces the enhancement if and only if

$$\psi > t_l(x^v - \bar{x})^2 + z_l(x^p - \bar{x})^2 \tag{2}$$

Obviously, if ψ is zero or negative, so that the punishment enhancement is generally not popular or outright unpopular among the voters, the lawmaker never introduces the enhancement. From the first order condition we have that:

$$2t_l x^v - 2t_l \bar{x} + 2z_l x^p - 2z_l \bar{x} = 0$$

which yields:

$$\bar{x} = \frac{t_l \bar{x}^v + z_l \bar{x}^p}{t_l + z_l} \tag{3}$$

Remark 1 *Note that if $t_l = z_l$ then $\bar{x} = \frac{\bar{x}^v + \bar{x}^p}{2}$. In this special case, the lawmaker's chosen $\bar{x}$ will be an average of the preferred thresholds of the voters and her fellow party members.*

4.2   The criminal's problem

The potential criminal chooses a level of crime denoted by x. The level of crime can have various interpretations: it can mean either the number of crimes committed, or, perhaps most appropriately for our model, the seriousness of the committed crime. If $x = 0$, the criminal does not commit any crime (therefore, he is not a criminal, in this case). The potential criminal's utility from crime has the following properties: the criminal's utility consists of a gross benefits from the crime, $B(x)$, and the expected cost of the crime, $F(x)$, and we assume that both the benefits and the costs are increasing in x. More formally, we assume that $\partial B(x)/\partial x > 0$ and $\partial F(x)/\partial x > 0$. However, we assume that if a punishment enhancement is introduced by the lawmaker then there is a jump in the criminal's cost function: after a certain level of x, (denoted by $\bar{x}$), the criminal faces an enhanced punishment, $F(x) + \alpha$, where $\alpha > 0$. Beside, there can

also be a kink in the function, as the probability of punishment can be higher or lower below and above the threshold. We denote the probability below the threshold as p and the probability above the threshold as p'. We allow for the possibility that $p = p'$.

The criminal's utility function takes the following form:

$$EU(x) = B(x) - \begin{cases} pF(x) & \text{if } x \leq \bar{x} \\ p'(F(x) + \alpha), & \text{if } x > \bar{x} \end{cases} \tag{4}$$

Besides the monotonicity of $B(x)$ and $F(x)$, we assume that both functions are continuous, twice-differentiable and concave. The first two assumptions are imposed for tractability reasons, while the latter assumption is made in order to approximate real life conditions, where the marginal increase in punishment is rarely equal and usually lower than the marginal increase in the crime level. In order to guarantee the existence of an interior solution, we assume that $pF''(x) > B''(x)$ and $B''(x) < pF''(x)$ for every level of x. That is, the criminal's cost function is steeper than his benefit function at any crime level below the threshold, so that the two functions intersect giving an interior solution.[1]

From the criminal's problem we get the following first order conditions:

i) In Case I:
$$\frac{\partial B(x)}{\partial x} = p\frac{\partial F(x)}{\partial x} \tag{5}$$

ii) In Case II:
$$\frac{\partial B(x)}{\partial x} = p'\frac{\partial F(x)}{\partial x} \tag{6}$$

If $p' < p$, the optimal crime level is higher above the punishment threshold. However, above the threshold the criminal has to pay the additional α punishment enhancement which is a fixed cost, so it does not influence the marginal decision. It does however influence whether the criminal will go above the threshold. The condition for this is $p'(F(x) + \alpha) > pF(x)$, leading to:

$$\frac{p'}{p} > \frac{F(x)}{F(x) + \alpha} \tag{7}$$

Notice, that as p changes above the threshold level $\bar{x}$, the marginal expected punishment curve becomes either steeper of less steep and we did not make any assumption regarding these cases. Therefore, first of all, it is completely possible to have no interior solution if a kink occurs before the marginal expected punishment curve would intersect the marginal benefit curve. In this case,

---

[1]Alternatively, we could assume that $B(x)$ is concave, while $pF(x)$ is convex. This assumption would guarantee not only the existence but the unicity of a interior solution. However, in real life the punishment function seldom if ever can be described as increasing and convex. Furthermore, it is also worth noting that both in the concave as well as the convex case we can imagine a situation where there is only one intersection and $\bar{x}$ is set above the optimal criminal choice x^*. The reason for this is that as long as voters are largely uninformed about criminal statistics, which could be assumed under rational ignorance, and the lawmaker does not care about social welfare, the lawmaker might choose an $\bar{x}$ which is above the crime level the criminal would choose, however that would be irrelevant from a deterrence perspective.

if p decreases and α is not high enough to offset this change, the criminal will always choose $x = 1$. However, if p decreases and the kink occurs after $B'(x)$ and $pF'(x)$ intersect, another problem arises: uniqueness of the optimum is not assured as if the marginal expected punishment curve becomes steep enough after the kink it can have two intersections with the marginal benefit curve. However, only one of these two possible optima will be chosen by the criminal. As we have shown above, the optimum selection will be guided by taking into account the value of α. If $p'(F(x) + \alpha) > pF(x)$, so that the condition $\frac{p'}{p} > \frac{F(x)}{F(x)+\alpha}$ holds, the criminal will choose the crime level at the intersection that occurs before the kink. Examining further the case when the marginal benefit and the marginal expected punishment functions have two intersections, we also note that the second intersection is not a stable optimum as in that point and its environment the $p'F'(x)$ function is steeper than the $B'(x)$ function, therefore, by increasing the crime level above the second possible optimum, the criminal increases her utility. It is easy to see that provided the conditions under which there exists two intersections, the higher potential optimum will not actually be at the point of the second intersection, but at $x = 1$. Formally:

Remark 2 *If there exist x^* and $x^{*\prime}$ such that $x^* < x^{*\prime}$ and $B'(x^*) = pF'(x^*)$, $B'(x^{*\prime}) = p'F'(x^{*\prime})$ both hold, then $x^{*\prime}$ is not a stable optimum. By choosing an arbitrary $x^{*\prime} + \epsilon$ crime level, where $\epsilon > 0$, the criminal's utility is strictly higher in the latter case: $B(x^{*\prime} + \epsilon) - p'F(x^{*\prime} + \epsilon) > B(x^{*\prime}) - p'F(x^{*\prime})$, therefore the criminal is better off increasing the crime level beyond $x^{*\prime}$. Also, as $B''(x) > p'F''(x)$ when $\bar{x} \leq x \leq 1$ (the condition that needs to hold when there are two intersections), the criminal increases the crime level until it reaches the maximum. Therefore, if the criminal chooses a crime level that is above $\bar{x}$ and the marginal benefit and expected punishment functions intersect twice, the criminal will choose $x = 1$, the highest possible crime level.*

4.3 The judge's problem

In the last period, the judge chooses the probability of a guilty verdict, p. We assume that the judge can effect the probability through adjusting the threshold above which the defendant is considered guilty. This assumption can be supported by the observation that this threshold is generally not spelled out in any of the sentencing guidelines, so judges enjoy some degree of freedom over it. We assume that the judge's utility depends on two variables: her preferred probability, p^o, and the preferred probability of her peer group, p^p. If she chooses a p that differs from her preferred value, the judge receives some disutility. Since the judge is also influence by the prevailing attitude, she derives a disutility by moving away from p^p as well. Our preferred interpretation of this peer effect is that the judge receives some punishment from any member of the "elite" for deviating from their position. Such punishment can take many forms: fewer invitations to certain social or public events, fewer friendship offers or simply a general disapproval. Different members of this peer group might administer different levels of punishment, depending on their own preferred value of p. We assume that the judge by making her decision will consider the average of all these punishments.

In the followings we first briefly outline a background model for the formation of peer groups and derive some conclusions about how much peer group membership will influence judicial behavior in different conditions. Next, we consider the choice problem of a representative judge in our punishment game.

4.3.1 Formation of peer groups

Suppose that judges receive their compensation in three forms: they receive a monetary wage, w_m, a general non-monetary wage, w_p, that largely depends on working conditions and is independent of work effort and productivity, and a social wage, w_s, which comes from being a member of a complementary social group. In the later assumption we follow Cicala et al (2015) and we assume a judicial labour market and a social market such that $w_m' \left(\frac{\partial f(j)}{\partial l_j} \right) > 0$ and $w_s' \left(\frac{\partial f(e)}{\partial s_j} \right) > 0$. I.e., the judge's wages are positively related to her marginal product in both the labor market and the social market, respectively. We do not make any assumption about the market structures, leaving our results applicable for various different cases. We assume that making judicial decisions that go against the preferences of the judge's peer group lowers her marginal product in the production of the peer group's social goods, for various reasons: it might indicate a worse complementarity between the judge and other members of the group, so their interactions will be less productive in the future; it might cause distress among group members, again lowering her productivity. The judge can avoid this by making choices and expressing views that are consistent with the average position in the group. We assume therefore that the social wage is a function of the difference between the peer group members' average position and the judge's revealed position, and can be written in the following general form:

$$w_s = \bar{w}_s - \mu |p - o|$$

where $\bar{w}_s$ denotes the social wage the member would get if her revealed position would be equal to the group position, p stands for the group's position, o is the peer group member's own position and μ is a positive parameter expressing the effect of a given distance from the peer group position on the social wage. We assume that a country (or other administrative unit) with greater population has more opportunity to form new peer groups. The intuition behind this assumption is that investing in group specific human capital has some fixed cost, and the bigger the market, the more groups can be supported. This is, in essential an application of the well-known "the division of labor is limited by the extent of the market": in our setting, the extent of the division of labor means specialization into different peer groups. There can be for example a "tough" and a "soft" peer group with respect to punishment, often combined with "left" and "right wing" views. In the long run equilibrium, judges can sort into different peer groups, from all of which they get a given social wage.

4.3.2 Specialization and the extent of the market

One corollary of our assumptions is that an increase in the number of peer groups should increase the reservation wage $\bar{w}_s$ of a random peer group member, and in turn, as a greater population can support a greater number of peer groups. In other words, specialization is limited by the extent of the market, countries with greater populations should have more "elite" peer groups and membership in these peer groups should have a smaller effect on judicial decision-making. More specifically, consider an "elite" population of N_e, and its n partition. Let us assume that these peer groups are distributed along a policy space $[0, 1]$, where a 0 position means an extreme anti-punishment or "left-wing" position, while a position 1 means an extreme pro-punishment or "right-wing" position. The peer groups have a distribution that is not specified here. We assume that the number of members in a peer group is greater than one. Assume that the positions these peer groups inhabit along the policy space is the average of the positions of their members. Individuals can decide which peer group they join or quit and we assume that they cannot themselves form a "singleton coalition", so they must belong to a peer group.

Suppose that individuals, following the previous outline, are rewarded in their peer groups according to their marginal product, and assume that the marginal product depends on the difference between the group position (the group average) and their respective revealed individual position. Thus, the closest a member is to a group position, the higher social wage she receives from that group. Therefore, any individual will join the peer group whose position is the closest to her own position.

Consider a marginal individual i in a market with two peer groups, with half way between the two peer group's positions, a and b (i.e. $i = \frac{a+b}{2}$). She is indifferent between joining the two groups, so she randomly joins one of the groups in that case, as $|a - \frac{a+b}{2}| = |b - \frac{a+b}{2}|$. Now consider what happens when a new group enters the market, with a position c, such that $a < c < b$. Regardless of whether c is smaller or greater than $\frac{a+b}{2}$, c will be i's new chosen group, as the social wage she receives there, $\bar{w}_s - \mu|c - \frac{a+b}{2}| > \bar{w}_s - \mu|a - \frac{a+b}{2}| = \bar{w}_s - \mu|b - \frac{a+b}{2}|$. Notice that all three peer groups now have to pay i more than before in order for i to join any of them, as the reservation wage is $\bar{w}' = \bar{w}_s - \mu|c - \frac{a+b}{2}|$, whereas previously it was equal to $\bar{w}_s - \mu|a - \frac{a+b}{2}| = \bar{w}_s - \mu|b - \frac{a+b}{2}|$. It follows, that an increase in the number of peer groups increases the reservation wage and therefore lowers the marginal social wage received from any peer group.

So far we assumed peer group position as exogenous, assuming that any individual member does not influence the position of the peer group, as there are a large enough number of individuals in any group. Now suppose that this condition does not hold. To show that our results hold in this case as well consider the following. Suppose that there is a preexisting group position a and b and individual i with a position $\frac{a+b}{2}$. However, now if an individual decides to join either group (and we assume that she basically *has* to join one), than simply by joining and thus influencing the group's position in her own direction, her social wage will always be higher than $\bar{w}_s - \mu|a - \frac{a+b}{2}|$. So if she joins group a, she will receive $\bar{w}_s - \mu|a' - \frac{a+b}{2}|$, where $\bar{w}_s - \mu|a' - \frac{a+b}{2}| > \bar{w}_s - \mu|a - \frac{a+b}{2}|$,

since $a' > a$. Now consider what happens when a group with a position c (with the previous conditions applying) enters the market. It is easy to see that i can only receive a higher social wage after the entry, the only difference being is that there will be a new group position for the entrant c' such that $|c' - \frac{a+b}{2}| < |c - \frac{a+b}{2}|$. That is, it can be shown that when peer group members have some market power, that is, they can partially set the position of a group, the increase in the reservation wage is even greater, and therefore the effect on judicial decision-making of an increase in the number of peer groups is also greater.

4.3.3 The judge's choice

In the following we add the following specifications to our outline above: we assume that the judge's disutility that comes from deviation from the peer group position takes a quadratic form. This assumption does not influence the comparative statics of our results but greatly simplifies our calculations. We further assume that being a judge gives the judge a baseline utility that depends on her monetary wage, her working conditions as well as her social wage received by her respective peer group. This baseline utility is written as $\bar{U}(w_m, w_s, w_p)$ and we assume that it is increasing, continuous, twice differentiable and concave for all the variables. We do not make any restrictions concerning either the judge's or her peer group's average preferences, yet we assume that $0 \le p^o \le 1$, $0 \le p^{o'} \le 1$, $0 \le p^p \le 1$, $0 \le p^{p'} \le 1$. Therefore, we allow both the judge's own ideal position and the peer group's ideal position to change in any direction above $\bar{x}$.

The judge's utility function can be written as:

$$U_j(p, w_m, w_s, w_p) = \bar{U}(w_m, w_s, w_p) - \begin{cases} t_j(p^o - p)^2 + z_j(p^p - p)^2, & \text{if } x \le \bar{x}, \\ t_j(p^{o'} - p)^2 + z_j(p^{p'} - p)^2, & \text{if } x > \bar{x} \end{cases} \tag{8}$$

If $x > \bar{x}$, the judge needs to administer higher punishments for any crime level in this interval. The judge determines the probability of punishment p, and she has a preferred or ideal p in each cases (p^o and $p^{o'}$). She derives disutility from deviating from these ideal values. She also faces some form of punishment by deviating from the preferred p of the median member of her peer group. From the first order conditions we have that:

i) In Case I:

$$2t_j p^o - 2t_j p + 2z_j p^p - 2z_j p = 0$$

    or

$$p = \frac{t_j p^o + z_j p^p}{t_j + z_j} \tag{9}$$

ii) In Case II:

$$2t_j p^{o'} - 2t_j p + 2z_j p^{p'} - 2z_j p = 0$$

    or

$$p = \frac{t_j p^{o'} + z_j p^{p'}}{t_j + z_j} \tag{10}$$

12

The p chosen by the judge can be higher or lower depending on whether the criminal chooses x below or above the threshold. p decreases above the threshold if the judge herself is lenient enough, that is, it gives her disutility to administer especially strict punishments, or her peer group is lenient or anti-punishment enough. In the converse cases, the judge is more likely to choose a higher p above the threshold. Note that, as we have seen in the previous subsection, the larger the number of peer groups the judge can join, the lower the influence of the peer group will be, as the judge can more easily self-select into a group whose position is close enough to her own position.

4.4 Equilibrium and comparative statics

To solve the game for its subgame-perfect Nash equilibrium we use backward induction. To do so, we substitute the results of the judge's maximization problem to the choice problem of the criminal. Thus, the criminal's problem can be written as:

$$EU(x) = B(x) - \begin{cases} \frac{t_j p^o + z_j p^p}{t_j + z_j} F(x) & \text{if } x \le \bar{x} \\ \frac{t_j p^{o\prime} + z_j p^{p\prime}}{t_j + z_j} (F(x) + \alpha), & \text{if } x > \bar{x} \end{cases} \tag{11}$$

First order conditions lead to:

i) In Case I:
$$\frac{\partial B(x)}{\partial x} = \frac{t_j p^o + z_j p^p}{t_j + z_j} \frac{\partial F(x)}{\partial x} \tag{12}$$

ii) In Case II:
$$\frac{\partial B(x)}{\partial x} = \frac{t_j p^{o\prime} + z_j p^{p\prime}}{t_j + z_j} \frac{\partial F(x)}{\partial x} \tag{13}$$

Intuitively, increasing the severity of punishment above a threshold level $\bar{x}$ can have two possible effects: the severity of punishment increases, but, if the judge is not "tough" enough, the probability of receiving any punishment decreases, so the final effect on expected punishment, and therefore, the level of crime chosen is ambiguous, depending on the values of the exogenous variables in the model. Examining the equation we get from the criminal's problem through backward induction, we can see that the more lenient the judge's peer group and the judge herself is, the higher the level of crime the criminal will choose in both cases. Also, the more $p^{o\prime}$ or $p^{p\prime}$ is larger than p^o or p^p, respectively, the more likely is that the criminal will choose a crime level above the threshold.

As we have seen from the lawmaker's problem above, the lawmaker's choice does not depend on either the judge's or the criminal's decision, so the statements made on her individual maximization problem apply. From that, the solution for $\bar{x}$ is given by:

$$\bar{x} = \frac{t_l \bar{x}^v + z_l \bar{x}^p}{t_l + z_l} \tag{14}$$

Our model implies the following results:

Proposition 1 *Criminal activity x decreases in p and $F(x)$. A sufficient increase in α leads the criminal to choose a crime level that is no higher than $\bar{x}$.*

Proof: see Appendix A.

Intuitively, if only the severity of punishment $F(x)$ changes above the threshold levels, x should decrease, but p is not independent of $F(x)$. When judges are "strict" enough, the probability of conviction can increase above $\bar{x}$.

Proposition 2 *There is an α^*, such that for an $\alpha < \alpha^*$ a decrease in $\bar{x}$ will always increase x if $p' < p$ and will always decrease it if $p' > p$.*

Proof: see Appendix B.

A decrease in the probability of conviction decreases the expected marginal cost of crime and *ceteris paribus* increases the crime level. The opposite holds in the case of increasing probability. However, for a big enough α, $p' < p$ is not enough for this to hold. As we have seen in the discussion of the criminal's problem, if the condition $p'(F(x) + \alpha) > pF(x)$ holds, decreasing $\bar{x}$ *ceteris paribus* decreases x.

Proposition 3 *$x^* \leq \bar{x}$ if and only if $p'(F(x) + \alpha) > pF(x)$*

Proof: see Appendix C.

That is, the crime level will never be higher than the threshold chosen by the lawmaker if either the probability increases after the threshold or the punishment enhancement is big enough to offset the effect of a decrease in p. If either of these is not the case, the criminal will choose an x such that $x > \bar{x}$.

5 Examples and extensions

5.1 Examples

Our model offers a number of possible applications and extensions. First of all, our predictions can be applied in the context of the standard criminal justice system. As we have noted in Section 2, there is evidence that judges have some room influencing the probability of conviction and their decisions are often influenced by their own preferences as well as social incentives like peer opinion, apart from material incentives. Obviously, the maneuvering room for judges is not the same in every context. We expect greater applicability of our model in cases where judges have a greater influence over the probability of conviction. This is more likely in cases where the decision of the judge is more hidden from the public eye, but can be monitored by some select group of individuals, described as the peer group in our model. Our results also have more

relevance in situations when there is some "schism" between public and elite opinion, or, in other settings, a difference in general between the reputational incentives of rule-setters or lawmakers, and those who apply the rules. The perverse result predicted under certain conditions in our model is also more likely to occur when there are not too many peer groups that judges can choose from. In the following we consider some settings where our results are especially relevant, both from a descriptive and from a policy perspective. Also, the more personal the relationship might be between investigators and the investigated, the more we should expect this relationship to hold, and the more we should worry about the perverse effects of strict punishments.

We begin with one of our motivating examples, the current schism in most European countries between general public opinion and the opinion of intellectual groups. Majorities in many European countries support capital punishment. This support is much weaker among more educated and higher income individuals, especially among "progressives" (see Whitehead et al, 1999). This seems to be the case in the case of many other forms of punishment, as well. Considering this, we expect the perverse result to be more likely in smaller European countries, where the schism between the general population and the "elite" on the matter of crime and punishment is coupled with a relatively small and homogeneous "elite". On the other hand, in the United States, while progressive views are increasingly common among intellectuals, there still seems to be many ideological groups and subgroups.

Our predictions may have relevance in the area of economic crimes, as well. Two examples in particular might be rules related to tax evasion and tax fraud as well as collusion among firms. In the latter case, it has been traditionally the case that only the firm itself gets punished for violations of competition law. A crucial problem with this traditional approach that it does not provide sufficient incentives against collusion, as the punishment is distributed among shareholders and the executive is not held responsible for the collusion. Considering this problem, many countries have enacted changes in the law according to which the executive or the executive directors can also be held liable in collusion cases. However, one problem with this approach, is that competition authorities and courts typically apply a higher threshold of reasonable doubt when deciding in a case where the punishment must be applied to one or more individuals as opposed to a firm. Therefore, increasing the punishment for collusion for the executive (from zero to non-zero level) might be coupled with a decrease in the probability that a firm gets any punishment for collusion. The final effect on the expected punishment might be unclear.

Another area where our model could be applied, with some modifications, is the investigation and punishment of crimes and other rule-breaking behavior (including breaking certain specifically religious or job-specific proscriptions) among clergy. For example, the Catholic Church has an autonomous body that adjudicates claims of crimes that is independent of secular authorities and it can issue punishments against its members. An especially stern form of ecclesiastical punishment is excommunication, which basically means an end of a priestly career, and could mean even more for a religious individual. Obviously, certain crimes such as the sexual abuse of

children often incite demand for very strict punishment within the Church as well as the general criminal justice system. Here the incentives faced by the Church leadership are ambiguous. First, the Church leadership might receive reputational benefits from a tough approach toward crimes committed by clergymen. The reputational benefit, however, is received by the Church as a whole, so even Church leaders might be unlikely to internalize the full reputational benefits of action against crimes allegedly committed by priests. However, leaders certainly derive some part of this benefit, therefore they still have some incentives for stricter punishment. Church leaders who take part in the investigation might have other incentives as well, however. First, there might be criminals among their peers, or their peers might disapprove of the possible convictions. While most of the positive reputational benefits might go to the highest Church leaders, such as the Pope, who devise the rules, individual investigators might face higher cost from offending their peers, especially if they still have career ambitions.

Another area where there is a system of punishment that is distinct from the general criminal justice system is the army. There are martial courts where traditionally the strictest punishment has been the death penalty (that is not the case any more in most of the Western world). Similarly to the case of the Church, the incentives of the army leaders and those involved in investigating allegations against army personnel can be very different. Army leaders arguably internalize more of the reputational costs and benefits that are conferred to the army as a whole, while investigators, as in the case of ecclesiastical investigations, might either have personal stakes in cases or are subject to peer pressure. For instance, war crimes and sexual crimes committed by army personnel might not be properly investigated due to a sense of solidarity among army members. Note however, that reputational concerns can incentivize both the furthering as well as stifling of investigations, depending on the probability that a particular allegation becomes known to the public. Given the above incentives in the army, it is possible that increasing the severity of punishment for military crimes and misdemeanors can decrease the probability that an investigation will begin or end with a conviction.

The intuition of our model can also be applied to regulatory enforcement. One of the central problems facing European labor markets are labor market rigidities that are often the result of inefficient regulations (see e.g. Blanchard and Wolfers, 2000). However, new empirical results by Kanbur and Ronconi (2016) suggest a counterveiling effect provided by the enforcement of these regulations: according to the results, the strictness of employment regulations correlate with weak levels of enforcement. This relationship makes the possible effect of stricter or more rigid labor regulations, or any type of regulations, ambiguous. Regulatory agencies might weaken enforcement in response to the stricter rules, thus the effective regulatory burden can increase as well as decrease, depending on the magnitude of changes in the rules versus the enforcement intensity.

5.2 Extensions

One limitation of our approach taken in this paper is that we treat the punishment enhancement α as exogenously given, whereas it could be assumed that α can be chosen by the lawmaker. In real life, both of these intuitions can find some support. If we take the example of classic punishments like life in prison, the absence of the possibility of parole or capital punishment, these punishments are not made up by the current lawmaker, rather, their application is what gets decided by her. However, when the law prescribes punishment enhancements that, for instance, add a certain number of years to the punishment, this number is actually determined by the lawmaking body. The reason we have taken α as given is that we wanted to show the implications of it, including the possible perverse effect of a mandatory punishment enhancement. However, if the lawmaker can choose the value of α, se can avoid such perverse results. In our model it is assumed that the lawmaker does not care about social welfare, including deterrence, so even if we allowed a choice of α, the lawmaker would not be interested in minimizing the perverse effect of a punishment enhancement. However, if she wished to, she could achieve undoing the perverse effect of introducing α by setting α so that $\alpha > \frac{p'}{p} - 1$, which comes directly from equation (7). The more lenient the judge is, the higher α needs to be in order to offset the effects of leniency. However, this is only one way the lawmaker could outweigh the potentially perverse effect of punishment enhancement. One other way to do that is to directly incentivize judges not to lower the probability of conviction after the $\bar{x}$ threshold. This can be done more easily in states where the judicial branch is, either formally or informally, not fully independent from executive power. Therefore, we expect liberal democracies to be more prone to the problem we described in this paper. However, as we have not assumed that lawmakers care about deterrence, our analysis suggests that lawmakers in democratic countries probably face greater incentives to introduce criminal legislations that have perverse effects on deterrence. While non-democratic leaders might be expected to internalize more of the benefit from criminal deterrence, democratic leaders are more responsive to public opinion.

6 Conclusions

In this paper we have examined the relationship between the severity and the probability of punishment when there are psychic costs judges derive from giving out certain punishments. The type of cost we consider specifically is the disutility the judge gets from deviating from the views of her peers. Due to this effect, we have found that increasing the severity of punishment, in certain cases, can decrease the expected punishment faced by the criminal and, therefore, can both increase and decrease the level of criminal behavior. Such an effect is less likely the more peer groups judges can join, which can be assumed to be correlated with country size. One consequence of this is that we expect the perverse result of punishment enhancement most likely to occur in countries with sufficiently pro-punishment general population and with relatively homogeneous, sufficiently anti-punishment "elite". We have also considered a few possible applications and

extensions of our model, including child abuse allegations in the catholic Church or decision making in martial courts. However, quantitative empirical studies are needed in order to test the predictions provided by our paper.

Appendix A

Proof of Proposition 1: See equations (5) and (6) for the first order conditions under p and p'. When either p or p' increases, the expected marginal punishment increases, which implies that the optimal criminal level x needs to be chosen at a higher level of marginal benefit. The above combined with the concavity of $B(x)$ implies a lower x. The same holds for changes in F (which is exogenously given in the model). A change in α increases the total cost of punishment in Case II. As α is a fixed cost, it cannot influence the marginal decision, only the optimum selection. Consider an initial optimal choice $x^* > \bar{x}$ and $x^{*\prime} \leq \bar{x}$ as an other possible optimum. Suppose that initially, $p'(F(x^*) + \alpha) < pF(x^{*\prime})$. There is a large enough increase in α such that $p'(F(x^*) + \alpha) > pF(x^{*\prime})$, so the criminal chooses the lower crime level.

Appendix B

Proof of Proposition 2: Consider the first order conditions in the two branches of the criminal's utility function. In Case I, the first order condition yields $\frac{\partial B(x)}{\partial x} = p\frac{\partial F(x)}{\partial x}$. In Case II it yields $\frac{\partial B(x)}{\partial x} = p'\frac{\partial F(x)}{\partial x}$. Suppose $p' < p$. Suppose there is an optimal solution x^* that satisfies the first order condition in Case I. Then if the lawmaker lowers the threshold $\bar{x}$, there exists a sufficient decrease in the threshold such that x^* is no longer optimal. There is a new downward kink in the expected punishment function that is below the original optimal choice. The new optimal choice in Case I is the same, but in Case II the optimal choice is either the same or lower than previously. It is the same only if Case II does not yield an interior solution ($x = 1$) under either the old or the new threshold.

Appendix C

Proof of Proposition 3 We consider two cases. In the first case, $p' > p$. It is easy to see that in Case II the criminal faces a higher marginal expected punishment. If the condition in the proposition also holds (and in this case it always does as $\alpha > 0$ by definition), x^* cannot exceed $\bar{x}$, as the criminal can only make herself worse off by choosing a crime level above the threshold. Now consider the case when $p' < p$. Here the marginal expected punishment decreases above $\bar{x}$. This implies an $x^{*\prime} > \bar{x}$. Suppose the optimal choice in Case II is x^*. We can modify the condition in the proposition as $p'(F(x^{*\prime}) + \alpha) > pF(x^*)$ The criminal is therefore better off

choosing $x^* < x^{*\prime}$. x^* must be lower or equal to $\bar{x}$ as above the kink $x^{*\prime}$ is the optimal choice, therefore x^* cannot be the optimal choice in Case II, so it must locate either at the kink or below it. Therefore, if α is great enough to offset a decrease in p, the chosen crime level will never exceed $\bar{x}$.

References

Acemoglu, Daron and Robinson, James A. (2009): *Economic Origins of Dictatorship and Democracy*, Cambridge University Press

Andreoni, James (1991), 'Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?', *The RAND Journal of Economics*, Vol. 22, No. 3, pp. 385-395

Baum, Lawrence (2007), *Judges and Their Audiences: A Perspective on Judicial Behavior*, Princeton University Press

Becker, Gary S. (1974), 'A Theory of Social Interactions', *NBER Working Paper No. 42*

Becker, Gary S. (1968), 'Crime and Punishment: An Economic Approach', *Journal of Political Economy*, Vol. 76, No. 2, pp. 169-217.

Becker, Gary S. and Stigler, George J. (1974), 'Law Enforcement, Malfeasance, and Compensation of Enforcers', *The Journal of Legal Studies*, Vol. 3, No. 1, pp. 1-18.

Blanchard, Olivier and Wolfers, Justin (2000), 'The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence', *The Economic Journal*, Vol. 110, Issue 462, pp. 1-33.

Brennan, Geoffrey and Pettit, Philip (2005), 'The Economy of Esteem. An Essay on Civil and Political Society', *Journal of Economics*, Vol. 86, No. 2, pp. 183-191.

Brock, William A. and Durlauf, Steven N. (2001), 'Discrete Choice with Social Interactions', *The Review of Economic Studies*, Vol. 68, No. 2, pp. 235-260.

Cicala, Steve; Fryer, Roland J., Jr. and Spenkuch, Jörg L. (2015), 'Comparative Advantage in Social Interactions', *NBER Working Paper No. 16880*

Cowen, Tyler (2002), 'The Esteem Theory of Norms', *Public Choice*, Vol. 113, No. 1, pp. 211-224.

Cowen, Tyler and Glazer, Amihai (2007), 'Esteem and Ignorance', *Journal of Economic Behavior and Organization*, Vol. 63, No. 3, pp. 373-383.

Donohue, John J. and Wolfers, Justin (2005), 'Uses and Abuses of Empirical Evidence in the Death Penalty Debate', *Stanford Law Review*, Vol. 58, No. 3, pp. 791-841.

Durlauf, Steven N. and Nagin, Daniel S. (2011), 'Imprisonment and Crime: Can Both Be Reduced?', *Criminology and Public Policy*, Vol. 10, No. 1, pp. 13-54.

Ehrlich, Isaac (1973), 'The Deterrent Effect of Capital Punishment: A Question of Life and Death', *NBER Working Paper No. 18*

Glaeser, Edward L.; Kessler, Daniel P. and Piehl, A. Morrison (2000), 'What Do Prosecutors Maximize? An Analysis of The Federalization of Drug Crimes', *American Law and Economics Review*, Vol. 2, No 2. pp. 259-290.

Glaeser, Edward L. and Sacerdote, Bruce L. (2000), 'The Determinants of Punishment: Deterrence, Incapacitation and Vengeance', *NBER Working Paper No. 7676*

Glaeser, Edward L. and Sacerdote, Bruce L. (2003), 'The Social Multiplier', *Journal of the European Economic Association*, Vol. 1, No. 2-3, pp. 345-353.

Glaeser, Edward L. and Sunstein, Cass R. (2009), 'Extremism and Social Learning', *Journal of Legal Analysis*, Vol. 1, Issue 1, pp. 263-324.

Haidt, Jonathan (2012), *The Righteous Mind: Why Good People are Divided by Politics and Religion*, Vintage

Kanbur, R and Ronconi, L. (2016), 'Enforcement Matters: The Effective Regulation of Labor', *CEPR Discussion Paper No. 11098*

Kuran, Timur (1997), *Private Truths, Public Lies: The Social Consequences of Preference Falsification*, Harvard University Press

Levitt, Steven D. (2004), 'Understanding Why Crime Fell in the 1990s: Four Factors That Explain the Decline and Six That Do Not', *The Journal of Economic Perspectives*, Vol. 18, No. 1, pp. 163-190.

Mill, John S. (1868), 'Speech in Favor of Capital Punishment', in *The Collected Works of John Stuart Mill, Volume XXVIII - Public and Parliamentary Speeches Part I November 1850 - November 1868*, ed. John M. Robson and Bruce L. Kinzer, Toronto: University of Toronto Press, London: Routledge and Kegan Paul

Schelling, Thomas C. (1960), *The Strategy of Conflict*, Harvard University Press

Smith, Adam (1790): *The Theory of Moral Sentiments*, 6th edition, London: A. Millar

Whitehead, John T., Blankenship, Michael B., Wright, John Paul (1999), 'Elite versus Citizen Attitudes on Capital Punishment: Incongruity Between the Public and Policymakers', *Journal of Criminal Justice*, Vol. 27, Issue 3, pp. 249-258.