The Optimal Standard of Proof in Criminal Law When Both Fairness and Deterrence Matter

Lando, Henrik

Document Version
Final published version

Publication date:
2000

License
CC BY-NC-ND

Citation for published version (APA):
Lando, H. (2000). *The Optimal Standard of Proof in Criminal Law When Both Fairness and Deterrence Matter*. Institut for Finansiering, Copenhagen Business School. Working Papers / Department of Finance. Copenhagen Business School No. 2000-7

Link to publication in CBS Research Portal

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 22. Mar. 2025

WP 2000-7

The Optimal Standard of Proof in Criminal Law When Both Fairness and Deterrence are Social Aims

af

Henrik Lando

INSTITUT FOR FINANSIERING, Handelshøjskolen i København
Solbjerg Plads 3, 2000 Frederiksberg C
tlf.: 38 15 36 15   fax: 38 15 36 00

DEPARTMENT OF FINANCE, Copenhagen Business School
Solbjerg Plads 3, DK - 2000 Frederiksberg C, Denmark
Phone (+45)38153615, Fax (+45)38153600
www.cbs.dk/departments/finance

ISBN 87-90705-41-6
ISSN 0903-0352

The Optimal Standard of Proof in Criminal Law When Both Fairness and Deterrence Matter

Henrik Lando,
Department of Finance,
Copenhagen Business School,
e-mail: hl....@cbs.dk

August 8, 2000

Abstract

This paper addresses the issue of the optimal standard of proof in criminal law. It is assumed that people in society care about both fairness and deterrence. It is important to punish those who are guilty and only those. However, error is unavoidable and hence a trade-o¤ emerges between the three aims of punishing the guilty, not punishing the innocent and deterring potential criminals. It is shown that when only deterrence matters the optimal standard of proof is a preponderance-of-the-evidence standard (given some other assumptions) while if fairness is an issue the standard will generally be stricter and involve Bayesian up-dating. When both fairness and deterrence matter the standard of proof will (generally) lie in between the two standards. An example illustrates how the model might be applied in practice to determine the optimal standard of proof for a given crime.

1. Introduction

The standard of proof is of central importance in almost any legal regulation. In general, a high standard of proof creates fewer false 'convictions', but more false acquittals and hence often less deterrence. In this paper we focus on criminal law. The question analyzed is the following: If a society holds certain aggregate preferences on the seriousness of a false acquittal, on the seriousness of a false conviction and on how these 'costs' compare with the 'cost' of the crime itself, then how much certainty would that society want before convicting one of its members of a crime[1]? The starting point of the analysis is the realization that full certainty concerning guilt can only rarely be obtained. This may be an understatement; recent application of DNA-tests to criminal cases from the past suggests that erroneous convictions are not as rare and unique as sometimes believed[2].

We will assume that people in society hold views concerning what is a 'proper' sanction for a given o¤ence, and that people will dislike if the sanction is either (much) lower or (much) higher than this. In extension, we will assume that society su¤ers a welfare loss by false convictions. The falsely convicted will know he is innocent and is likely to su¤er a loss due to being sanctioned which is higher than he would feel if he were guilty. The rest of society will not know whether a given person is guilty or not (though people may form beliefs in this regard) but we assume that people dislike, from a fairness perspective, the idea that false convictions occur[3].

In some legal cases[4], fairness may not be very important[5]. Hence, it seems worthwile establishing as a benchmark the optimal standard of proof when deterrence is the only concern. We also establish the optimal standard of proof when fairness is the only concern. However, the question which we are mainly interested in is what the standard of proof should be when both deterrence and fairness matter. We show that the optimal standard of proof will except in an exceptional case lie between the two benchmarks just mentioned when both deterrence and

[1] This question is sometimes ignored or even denied in legal discourse where 'legal security' is sometimes treated as an absolute good that should not be traded-o¤ against other goods such as deterrence.

[2] For American evidence involving capital cases, see e.g. Liebman et. al. [9].

[3] The same assumptions are made by e.g. Miceli [10]; see also Andreoni [1].

[4] especially when we include civil law sanctions such as remedies for breach of contract. The focus of this article is on criminal law but some of the analysis applies to civil law as well.

[5] This may e.g. be the case in civil law disputes, such as in a tort case, where both parties may be insured.

fairness are social aims, and we attempt to establish how the trade-o¤ should be made in an actual case.

To determine the optimal standard of proof it will be necessary to simultaneously determine other variables such as the level of sanctions and the level of enforcement. For example, if punishment is very harsh this may call for a higher standard of proof than if punishment is more lenient since it may be felt to be more serious to falsely convict a person when the sanction is heavy. We will hence address the question of the optimal standard of proof within a model that simultaneously determines also the optimal sanction and optimal levels of enforcement[6] in the presence of uncertainty concerning guilt.

Analyzing what is the optimal standard of proof leads one to consider the meaning of the concept itself, since it is not as clear as intuition may perhaps suggest. One may associate the notion of 'standard of proof' with the notion of 'probability of guilt' but that may be misleading. Probability of guilt involves Bayesian updating of prior probabilities with the advent of new information (evidence), but it may not be optimal to do Bayesian updating. It may be optimal for society that the judge looks only at the evidence of the case before him and not at the prior probability that the defendant would commit such an act. To illustrate the importance of this distinction, imagine two societies, A and B, where in A very few people even consider committing theft while almost everyone in B lacks moral notions and commit theft if it is to their (amoral) personal advantage. Imagine further that two people are apprehended by the police, one in each society, and that the evidence raised against them turns out to be identical. Ignoring the issue of equality before the law, should the courts then treat the two cases in the same way (if we abstract from the fact that preferences concerning proper punishment is likely to di¤er between the two societies)? As a contribution of this paper, it will be shown to depend on the aims of society (deterrence or fairness) and on whether sanctions are socially costly.

The article follows the approach to criminal law initiated by Bentham, revived by Becker[2] and further developed by Shavell and Polinsky, among others (for an overview see [17], and Garoupa [5]).

More speci…cally, we shall address the issue within a model by Shavell and Polinsky [18] that incorporates fairness concerns. The model of the present article di¤ers from the Shavell and Polinsky model by assuming that a judge (or jurors) must base verdicts on imperfect information about how people have acted. This

[6] Since these variables are of interest in themselves we shall at points digress brie‡y to interpret the results.

entails the risk that someone will be falsely convicted; as mentioned above, we shall assume that people (including the falsely convicted) will thereby suffer a welfare loss due to a feeling of unfairness over and above the loss due to the sanction itself.

Two important papers by Miceli: [10] and [11] are closely related to the present one. In [11], Miceli analyzes the question of the optimal standard of proof assuming that only fairness, not deterrence, matters, and in [10] he analyzes the optimal levels of enforcement and punishment when both fairness and deterrence matter. In the latter paper, uncertainty is present but the standard of proof is taken as given. In contrast, the present paper analyzes the optimal standard of proof when both fairness and deterrence matter[7].

Concerning the specification of the model, two distinctions are worth making. The first concerns an effect which has been discussed by Png [12], namely that the risk of being falsely convicted of a crime lowers the incentive to obey the law since it lowers the expected sanction for committing the illegal act. For certain kinds of crimes, however, this effect is not present. It can be excluded for those kinds of crime, e.g. murder, where not committing the illegal act implies that there will be no harm, hence no investigation and no possibility of a false conviction for this particular (non-existing) crime. For other kinds of illegal behavior, it may play a role particularly if the probability of false conviction is high. The example analyzed by Png [12] concerns careless driving. If a motorcyclist risks being falsely convicted of careless driving in the event of an accident, this may lower his incentive to drive carefully, Png maintains. Our model concerns the latter kind of behavior where the effect is, at least potentially, present, but the model only needs slight modification to be useful for analyzing the former kind of behavior. We will do so to analyze the optimal standard of proof when only deterrence matters.

The second distinction concerns whether the police is faced with the question: 'who among several suspects did it?' rather than with the question: 'did person A do it, or was it an accident/has there been observational error'? Again, murder sometimes falls in the first category. When someone has been murdered, the problem can be to identify the offender among several potential offenders. Tax-evasion and speeding, on the other hand, fall in the second category where investigations

[7] There are some other differences. Miceli does not consider the possibility of over-deterrence (rightly so for such crimes as murder); he does not make use of the concept of the likelihood ratio and the way in which he models the legal system is in minor ways different from that of the present paper.

usually only concern the actions of one person. Murder-cases can also be of this kind, as when it is uncertain whether someone was shot, as an act of will or as an accident ('the gun went o¤'). In terms of modeling, the former kind of crime is more like a moral-hazard-in-teams problem while the latter is more like a simple, one-agent moral hazard problem which is simpler to analyze. In this article we will analyze the case where only one person's actions are involved. The question is whether a harm can be attributed to the unlawful, reckless or negligent actions of somebody or whether the harm has been caused by (pure) accident. One example would be the following: …sh are found dead in a lake and suspicion falls on a …rm for having emitted a larger amount of some substance into the lake than permitted. While the evidence points in this direction it does so with less than full certainty. The …sh may be dead due to other, e.g. natural, causes[8].

The article is organized as follows: In the following section the model is introduced. In the two -following sections we analyze two polar cases: First the case where only deterrence matters and fairness concerns are absent and next the case where only fairness matters and deterrence is absent. Then we bring results together in an analysis of the realistic case where both deterrence and fairness matter. Discussion of an extension and a conclusion ends the article.

2. The Model

Agents are atomistic and the whole population is normalized to one. Each agent may commit an illegal act which grants utility equal to g. g di¤ers between agents; it has density function $z(g)$ and cumulative distribution function $Z(g)$: The di¤erent gains may re‡ect di¤erent moral costs of committing a crime. Harm to society of a criminal act is h: An agent will choose to commit a criminal act when his expected utility is higher from this act than from other acts he can choose. For the purpose of tractability we assume agents are risk-neutral. The critical level of gain for which the agent will be just indi¤erent between committing the criminal act and not committing it will be termed $\underline{g}$:

We assume that the police decides on e^1 which signi…es the number of cases it chooses to go into or the level of monitoring undertaken. $p(e^1)$ will signify the probability of being investigated or monitored. e^2 will signify the resources spent

[8] Taking this example, our assumption in the model will be that there are many …rms facing a similar situation, and that …rms di¤er in their utility from breaking the law. Or, alternatively, that there is only one …rm whose type is unknown

on accumulating evidence in a case. e^2 affects the precision of the information on which a verdict is based[9].

We can express the evidence gathered by the stochastic variable (or vector) x: The criminal act will induce one density function over x : $v_c(x; e^2)$ while non-criminal acts induce (another) density-function: $v_n(x; e^2)$: We assume that the likelihood ratio $\frac{v_n(x;e^2)}{v_c(x;e^2)}$ is decreasing in x: Thus, a higher realized value of x implies a higher likelihood that the person has committed the harmful act. The standard of evidence set by the court can then be expressed as a level $\overline{x}$, where a person will only be convicted in court if x is larger than $\overline{x}$. Naturally, in the absence of uncertainty, the police will in our simplified world only take those cases to court where $x > \overline{x}$[10].

The judge must decide whether or not to impose a sanction f when he observes x: The sanction may be interpreted as a fine or as imprisonment (or as the death penalty). When the sanction is interpreted as imprisonment, it will involve a social cost. When fairness concerns are part of the analysis, we can include the social economic cost of sanctions in the analysis by subtracting it in the benefit of sanctioning the guilty and including it in the cost of sanctioning the innocent, see below[11]. When only deterrence matters, we need to include the cost explicitly. Stigma is often an important factor; if it is a linear function of the sanction our model can incorporate it without modification, otherwise we need to make minor adjustments when the optimal level of sanctions is determined.

The system thus determines a function $f(x)$ and we assume in line with what is at present realistic that $f(x)$ takes the value f for $x > \overline{x}$ and the value 0 for $x \le \overline{x}$[12]. For the case where the sanction is monetary we further assume that

---

[9] Miceli [10] models this in a slightly different way. He assumes that the police spends effort δ in apprehending a person after a crime and that higher effort by the police increases the probability of apprehending the right person. δ hence corresponds to our e^2 which determines the amount of information collected in a case.

[10] We abstract from the existence of plea bargaining and from issues concerning the prosecutor's incentives that may differ from society's, see Grossman and Katz [6], Miceli [11] and Reinganum [14].

[11] However, in the first section where deterrence is the only motive, these costs do not appear and hence we need to introduce the social cost of imposing sanctions explicitly in this section if we wish to interpret the sanction as imprisonment. This will not be done since the implications are obvious.

[12] It will hence be assumed that there is either full conviction or full acquittal; the possibility that the level of certainty may determine the level of the sanction will not be explored in this article, though it is a very important possibility. It is investigated in a tort context by Shavell [19].

there is a maximal fine f_{max} due to limited wealth of offenders.

3. The Case With No Fairness-Dependent Utility

3.1. Optimal Sanctions, Enforcement and Standard of Proof

Given risk-neutrality, the maximization problem can be stated as follows:

$$\begin{aligned} Max_{A;f;e^1;e^2} W &= \int_{\bar{g}}^{\infty} g z(g) dg - (1 - Z(\bar{g}))h - e^1 - e^2 \\ \text{where } \bar{g} &= p(e^1) f \int_{\underline{x}}^{\infty} v_c(x; e^2) - v_n(x; e^2) dx \text{ if } f \leq f_{max} \\ \bar{g} &= p(e^1) f_{max} \int_{\underline{x}}^{\infty} v_c(x; e^2) - v_n(x; e^2) dx \text{ if } f > f_{max} \end{aligned}$$

The first equation expresses that social welfare amounts to the welfare of those who commit the act (which we include but which we could also omit) minus the loss to society from the acts committed and the costs of detecting and investigating. The second equation expresses that an agent will commit the illegal act if his benefit of committing the criminal act exceeds the difference in expected sanction from doing rather than not doing the act[13]. The third equation expresses that only the part of a fine that a person can pay has a deterrent effect. If the sanction is non-monetary, such as incarceration, f_{max} would be the life-time of the individual and the cost of incarcerating a person, $c(f)$; to society should be included in the $W-$function which becomes: $W = \int_{\bar{g}}^{\infty} g z(g) dg - (1 - Z(\bar{g}))h - e^1 - e^2 - c(f) * p(e^1)((1 - Z(\bar{g})) \int_{\underline{x}}^{\infty} v_c(x; e^2) dx + Z(\bar{g}) \int_{\underline{x}}^{\infty} v_n(x; e^2) dx)$. We first analyze the case where sanctions are monetary.

3.1.1. Monetary Sanctions

The maximization problem is in many respects analogous to that of Shavell-Polinsky [16].

[13] If we choose the specification where the non-occurrence of harm precludes the existence of a crime, this amounts to the case where $v_n(x) = 0$ for all x, since the only x's that call for an investigation involve harm. In this case we obtain: $\bar{g} = p(e^1) f \int_{\underline{x}}^{\infty} v_c(x; e^2) dx$: This is the specification chosen by Miceli [10]. We shall return to an implication of this alternative specification.

We first state a solution to the maximization problem and then interpret the result concerning the standard of proof. Denote a solution by $(f_d; e_d^1; e_d^2; \bar{x}_d)$:

Proposition 3.1. Under the given assumptions, a solution is:

$f_d = f_{max}$,

$\bar{x}_d$ is given by the equation $v_c(\bar{x}_d; e_d^2) = v_n(\bar{x}_d; e_d^2)$,

e_d^1 is either zero or given by:

$p'(e_d^1) = 1/(h - \theta)z(\theta)f_{max}S$;

e_d^2 is either zero (if $e_d^1 = 0$ or $p(e_d^1)f_{max}\int_{\bar{x}_d}^{\infty} v_c(\bar{x}_d; 0) - v_n(\bar{x}_d; 0)dx \geq h$) or given by the equation: $(h - \theta)z(\theta)p(e_d^1)f_{max}(\partial(\int_{\bar{x}_d}^{\infty} v_c(\bar{x}_d; e^2) - v_n(\bar{x}_d; e^2)dx)/\partial e^2) = 1$

Proof of the proposition: The proof of the proposition lies essentially in the following idea: Assume that it is worthwhile deterring all types below some θ: Then θ must equal $p(e^1)f\int_{\bar{x}}^{\infty} v_c(x; e^2) - v_n(x; e^2)dx$: If we set f and $\int_{\bar{x}^*}^{\infty} v_c(x; e^2) - v_n(x; e^2)dx$ at their maximal levels then this will minimize the need for enforcement e^1:

The results concerning e_d^1 and e_d^2 are quite straightforward to derive; while that concerning $\bar{x}_d$ may need a proof: The claim is that under the assumption of a monotone likelihood ratio, $\int_{\bar{x}}^{\infty} v_c(x; e_d^2) - v_n(x; e_d^2)dx$ is maximized for that $\bar{x}$ which solves $v_c(\bar{x}; e_d^2) = v_n(\bar{x}; e_d^2)$: To prove this denote this level $\bar{x}_d$: $\frac{v_c(\bar{x}; e_d^2)}{v_n(\bar{x}; e_d^2)}$ increasing in $\bar{x}$ implies that $v_c(\bar{x}; e_d^2) > v_n(\bar{x}; e_d^2)$ for $\bar{x} > \bar{x}_d$ and $v_c(\bar{x}; e_d^2) < v_n(\bar{x}; e_d^2)$ for $\bar{x} < \bar{x}_d$: Hence the integral $\int_{\bar{x}}^{\infty} v_c(x; e_d^2) - v_n(x; e_d^2)dx$ must be maximized for $\bar{x} = \bar{x}_d$:

Another way of obtaining the result is through the first-order condition which reads: $\frac{dW}{d\bar{x}} = \frac{\partial W}{\partial \theta} \cdot \frac{\partial \theta}{\partial \bar{x}} = (h - \theta) \cdot z(\theta) \cdot p(e^1) \cdot f \cdot (v_n(\bar{x}; e_d^2) - v_c(\bar{x}; e_d^2))$: This equals zero when $v_n(\bar{x}; e_d^2) = v_c(\bar{x}; e_d^2)$:

Concerning e^2, increasing accuracy of information will deter more people in the measure $z(\theta)p(e_d^1)f_{max}(\partial(\int_{\bar{x}_d}^{\infty} v_c(\bar{x}_d; e^2) - v_n(\bar{x}_d; e^2)dx)/\partial e^2)$ and the benefit of so doing is $(h - \theta)$: The term $\partial(\int_{\bar{x}_d}^{\infty} v_c(\bar{x}_d; e^2) - v_n(\bar{x}_d; e^2)dx)/\partial e^2$ expresses how much an increase in e^2 increases the difference of likelihood of conviction when one has versus one has not committed the act and hence expresses how important greater effort in finding the truth is for deterrence.

We state the result concerning $\bar{x}_d$ in a separate proposition:

Proposition 3.2. In the absence of fairness concerns, if sanctions are monetary and people are risk-neutral, fines should be imposed on a person when, given the received signal x about his behavior, this signal is more likely to be received if the act has been committed than if it has not.

This result establishes the preponderance-of-the-evidence-standard as a benchmark.

As mentioned above, the modeling employed here does not apply directly to such cases, e.g. certain murder-cases, where investigations are contingent on the act. We shall briefly mention the result which applies to this kind of crime as far as the standard of proof is concerned. When $g = p(e^1) f \int_{\bar{x}}^{\infty} v_c(x; e^2) dx$; deterrence is maximized when $\int_{\bar{x}}^{\infty} v_c(x; e^2) dx$ is maximized. Since sanction by assumption does not involve social cost (and agents are risk-neutral), the optimal standard of proof in this case would involve $\bar{x} = -\infty$; i.e. when harm occurs sanction should always follow.

In Appendix A we show that when a criminal case involves more than one suspect, the specification of the model nees to be changed and a similar result as that just obtained concerning the optimal standard of proof when only deterrence matters does not apply.

We will now provide some understanding of the proposition above and discuss a particular implication.

3.1.2. Why Does the Preponderance-of-the-Evidence Standard Maximize Deterrence?

Before committing a criminal act, a person is likely to contemplate the information or signals that alternative acts, criminal or non-criminal, will produce. Naturally, some of the signals are more likely to be received by the court-system if the criminal act will be committed than if it will not be. If for each such signal received by the court there will be conviction, the contemplation of this signal by the potentially criminal person will act as an inducement not to perform the criminal act, since the expected pay-off of the criminal act will hereby be diminished in relation to the expected pay-o¤ of non-criminal acts. The remaining signals are more likely to be received by the judge if no criminal act has been performed. If the receipt of each such signal leads to non-conviction, again this will act as an inducement not to perform the criminal act: The person will think that this favorable outcome is more likely if he does not commit the criminal act than if he does. Thus, to maximize the di¤erence between the expected utility from not committing and committing a crime, the court should convict whenever it receives signals that are more likely to be received when the criminal act has been committed than when it has not[14].

3.1.3. On the Non-use of Ex-ante Probability of Guilt

The result that the optimal standard of proof is a preponderance of the evidence-standard tells us what information to use in reaching verdicts but, equally important, also what kind not to use. A preponderance-of-the-evidence standard should be distinguished from a standard which passes a verdict of guilty when it is more likely than not that the person is guilty. The notion of a probability of guilt is problematic since to de…ne it we must set a priori probabilities of guilt which are updated in a Bayesian fashion by the informativeness of the signal x: Only if the a priori probability of a given apprehended person being criminal is assumed to be equal to 1/2, can we say that when the standard of proof is $\overline{x}_d$ people are sanctioned when it is more likely than not that they are guilty. Rather, the proposition should be interpreted as saying that under the given assumptions (no fairness-utility, no risk-aversion, no social cost of imposing penalties) à priori probabilities of guilt or innocence play no role in the optimal system of sanctions. This relates to the example in the introduction where society A and B were to decide on cases that were from the viewpoint of the concrete evidence identical but where the ex-ante probability of guilt was much higher in society B than in society A. The only di¤erence between the two societies lies in $Z(g)$; the distribution function for the types, which a¤ects the à priori likelihood of guilt. Hence, according to the proposition above there should be no di¤erence in the standard of proof in the two societies. The density-functions v_c and v_n do not depend on the distribution of types in the population but only on the accuracy of information. Thus, the verdict should in both cases be 'guilty' if the police measurement is more likely to be realized when the evidence put before the judge is more likely

[14] Note that this is the universal principle in the case of risk-neutrality also for the kind of crime where not committing the act implies no risk of conviction.

to be the outcome of a theft than not. It may seem surprising that the optimal system in this way totally disregards ex ante probabilities. But note that imposing a fine is costless to society and it may hence be worthwhile to impose a fine even when it is unlikely in a Bayesian sense that a person is guilty. It constitutes an efficient way for society A of deterring its few potential offenders.

3.1.4. Non-monetary Sanctions

The result that the preponderance of the evidence standard is optimal and ex ante probabilities irrelevant does not carry over to non-monetary sanctions that are costly to apply. When

$$W = \int_{\bar{g}}^{\infty} g z(g) dg - (1 - Z(\bar{g}))h - e^1 - e^2 -$$

$$c(f) \cdot p(e^1)((1 - Z(\bar{g})) \int_{\bar{x}}^{\infty} v_c(x; e^2)dx + Z(\bar{g}) \int_{\bar{x}}^{\infty} v_n(x; e^2)dx)$$

and

$\bar{g} = p(e^1) f \int_{\bar{x}}^{\infty} v_c(x; e^2) - v_n(x; e^2)dx$; the first-order condition becomes:

$$\frac{dW}{d\bar{x}} = \frac{\partial W}{\partial \bar{g}} \cdot \frac{\partial \bar{g}}{\partial \bar{x}} = p(e^1) \cdot f \cdot (v_n(\bar{x}; e_d^2) - v_c(\bar{x}; e_d^2)) \cdot$$

$$z(\bar{g}) \left[(h - \bar{g}) + c(f) \cdot p(e^1) \cdot \int_{\bar{x}}^{\infty} v_c(x; e^2) - v_n(x; e^2)dx \right] +$$

$$c(f) \cdot p(e^1) \cdot \left[Z(\bar{g})v_n(\bar{x}; e^2) + (1 - Z(\bar{g}))v_c(\bar{x}; e^2) \right] = 0$$

Deterrence is more important than when sanction is costless as can be seen from the expression

$$z(\bar{g})c(f) \int_{\bar{x}}^{\infty} v_c(x; e^2) - v_n(x; e^2)dx$$

but increasing $\bar{x}$ confers the benefit of punishing both guilty and people innocent people less, since:

$$c(f)\cdot p(e^1)\cdot\left[Z(\underline{g})v_n(\bar{x};e^2)+(1-Z(\underline{g}))v_c(\bar{x};e^2)\right]>0$$

the standard of proof is higher than the preponderance of the evidence standard in this case. If $\bar{x}=\bar{x}_d$ then $v_n(\bar{x};e_d^2)-v_c(\bar{x};e_d^2)=0$; hence the derivative equals the latter part $c(f)\cdot p(e^1)\cdot[Z(\underline{g})v_n(\bar{x};e^2)+(1-Z(\underline{g}))v_c(\bar{x};e^2)]$ which as stated is positive. Hence, it adds to welfare to increase $\bar{x}$ beyond $\bar{x}_d$.

It is also clear that the irrelevance of ex-ante probabilities no longer holds, since $Z(\underline{g})$ and $(1-Z(\underline{g}))$ now enter the first-order condition but exactly how ex ante probabilities enter the choice of standard of proof will not be explored further here[15].

4. The Case Including Fairness-utility

We shall follow Miceli [10],[11], Shavell-Polinsky[17] and Diamond [3] in the way we formulate preferences for fairness. If a person is sanctioned at the level f although the person is innocent, there will be a loss of utility in the population (including his personal loss) of $\mu(f)$ while if a guilty person goes free, there will be a loss of utility of $Q(0)$. If a guilty person is fined f there will be a utility-gain equal to $Q(f)$. Thus, $Q(f)$ expresses the aggregate utility attached to alternative levels of sanctions for guilty people while $W(f)$ expresses the aggregate utility attached to punishing the innocent. $\mu(f)$, $Q(f)$ and $Q(0)$ can include the social cost of punishing at the level f when the sanction is non-monetary. Hence, we can exclude the cost of imprisonment explicitly in the analysis. If we express the standard of proof in terms of the cut-off signal $\bar{x}$, the maximization problem can now be expressed as:

$$\begin{aligned}\max_{\bar{x};e^1;e^2;f} W = {} & \int_{\underline{g}}^{1} gz(g)dg - (1-Z(\underline{g}))h \\ & -\mu(f)Z(\underline{g})p(e^1)\int_{\bar{x}}^{1} v_n(x;e^2)dx \\ & -Q(0))(1-Z(\underline{g}))(1-p(e^1)) \\ & -Q(0)(1-Z(\underline{g}))p(e^1)(1-\int_{\bar{x}}^{1} v_c(x;e^2)dx)\end{aligned}$$

[15] Note that it should be analyzed together with the variable e^1.

$$+Q(f)(1 - Z(\tilde{e}))p(e^1)\int_{\bar{x}}^{\infty} v_c(x; e^2)dx$$

$$- e^1 - e^2$$

$$\text{where } \tilde{e} = p(e^1)f \int_A v_c(x; e^2) - v_n(x; e^2)dx$$

Some notation will be useful: W can be written as $W_d + W_f - e^1 - e^2$ where $W_d = \int_{\tilde{e}}^{\infty} gz(g)dg - (1 - Z(\tilde{e}))h$ and

$$\begin{aligned}
W_f = & -\theta(f)Z(\tilde{e})p(e^1)\int_{\bar{x}}^{\infty} v_n(x; e^2)dx \\
& - Q(0))(1 - Z(\tilde{e}))(1 - p(e^1)) \\
& - Q(0)(1 - Z(\tilde{e}))p(e^1)(1 - \int_{\bar{x}}^{\infty} v_c(x; e^2)dx) \\
& + Q(f)(1 - Z(\tilde{e}))p(e^1)\int_{\bar{x}}^{\infty} v_c(x; e^2)dx
\end{aligned}$$

Also, the W-function can be written: $W(\tilde{e}(f; e^1; e^2; \bar{x}); f; e^1; e^2; \bar{x})$: Hence,

$$\begin{aligned}
\frac{dW}{df} &= \frac{\partial W}{\partial \tilde{e}}\frac{\partial \tilde{e}}{\partial f} + \frac{\partial W}{\partial f} \\
\frac{dW}{de^1} &= \frac{\partial W}{\partial \tilde{e}}\frac{\partial \tilde{e}}{\partial e^1} + \frac{\partial W}{\partial e^1} \\
\frac{dW}{de^2} &= \frac{\partial W}{\partial \tilde{e}}\frac{\partial \tilde{e}}{\partial e^2} + \frac{\partial W}{\partial e^2} \\
\frac{dW}{d\bar{x}} &= \frac{\partial W}{\partial \tilde{e}}\frac{\partial \tilde{e}}{\partial \bar{x}} + \frac{\partial W}{\partial \bar{x}}
\end{aligned}$$

Note that the partial derivatives $\frac{\partial W}{\partial f}$ and $\frac{\partial W}{\partial \bar{x}}$ can be written as $\frac{\partial W_f}{\partial f}$ and $\frac{\partial W_f}{\partial \bar{x}}$ while $\frac{\partial W}{\partial e^1}$ and $\frac{\partial W}{\partial e^2}$ can be written as $\frac{\partial W_f}{\partial e^1} - 1$ and $\frac{\partial W}{\partial e^2} - 1$: The reason is that the variables only affect W_d through the $\tilde{e}(f; e^1; e^2; \bar{x})$-function. Hence, we can write the first-order conditions as:

$$\begin{aligned}
\frac{dW}{df} &= \frac{\partial W}{\partial \tilde{e}}\frac{\partial \tilde{e}}{\partial f} + \frac{\partial W_f}{\partial f} = 0 \\
\frac{dW}{de^1} &= \frac{\partial W}{\partial \tilde{e}}\frac{\partial \tilde{e}}{\partial e^1} + \frac{\partial W_f}{\partial e^1} - 1 = 0
\end{aligned}$$

$$\frac{dW}{de^2} = \frac{\partial W}{\partial \tilde{g}}\frac{\partial \tilde{g}}{\partial e^2} + \frac{\partial W_f}{\partial e^2} - 1 = 0$$
$$\frac{dW}{d\bar{x}} = \frac{\partial W}{\partial \tilde{g}}\frac{\partial \tilde{g}}{\partial \bar{x}} + \frac{\partial W_f}{\partial \bar{x}} = 0$$

In this way we separate the effects that depend on people being deterred and the pure fairness-effects (where the number of people deterred is held constant).

We will be interested in finding the optimal standard of proof; to do so we must to find interior solutions solve this simultaneous set of equations. We start by finding the right-hand sides of the derivatives just given, i.e. $\frac{\partial W_f}{\partial f}; \frac{\partial W_f}{\partial e^1}; \frac{\partial W_f}{\partial e^2}$ and $\frac{\partial W_f}{\partial \bar{x}}$:

4.1. The Case Without a Deterrence Motive

Let us assume that the number of people who violate the law is unaffected by the standard of proof and by the other variables: $\tilde{g} = \bar{g}$: Then the problem reduces to:

$$\max_{\bar{x};e^1;e^2;f} W_f - e^1 - e^2$$

Denote optimal levels by $e^1_f; e^2_f; f_f$ and $\bar{x}_f$. We introduce the following notation:

$$T_1 \equiv \text{probability of sanctioning the innocent} = p(e^1)\int_{\bar{x}}^{1} v_n(x;e^2)dx$$
$$T_2 \equiv \text{probability of not sanctioning the guilty} = (1 - p(e^1)\int_{\bar{x}}^{1} v_c(x;e^2)dx)$$
$$C \equiv \text{probability of sanctioning the guilty} = p(e^1)\int_{\bar{x}}^{1} v_c(x;e^2)dx$$

In the following section, we analyze the optimal levels of sanctions and enforcement. The section can be skipped by the reader who is only interested in the results concerning the optimal standard of proof.

4.1.1. Optimal Sanctions and Enforcement

e^1_f : $\frac{\partial W_f}{\partial e^1}$ is given by

$$p'(e^1)\left[(Q(0) + Q(f))(1 - Z(\bar{g}))\int_{\bar{x}}^{1} v_c(x;e^2)dx - Z(\bar{g})\int_{\bar{x}}^{1} v_n(x;e^2)dx\right] =$$

$$\frac{\partial C}{\partial e^1}(Q(0) + Q(f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^1} Z(\overline{g})\mu(f):$$

Thus, in an interior maximum $(e^1_f; e^2_f; f_f, \overline{x}_f)$ it must be the case that:

$$\frac{\partial C}{\partial e^1}(Q(0) + Q(f_f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^1} Z(\overline{g})\mu(f_f) = 1$$

where $\frac{\partial C}{\partial e^1}$ and $\frac{\partial T_1}{\partial e^1}$ are evaluated in $(e^1_f; e^2_f; f_f, \overline{x}_f)$. This and some of the other results of this section are similar to the results obtained by Miceli [11].

Increased enforcement effort increases the likelihood both of convicting a guilty person and of convicting an innocent person. If he is guilty and convicted there is a double gain $Q(0) + Q(f)$; the person who would unfairly go free is now fairly punished. If he is innocent and convicted there is a cost of $\mu(f_f)$ as expressed in the last term.

It is worth noting that if accuracy is high enforcement creates fairness and that in this case the idea of setting f at a very high level in order to be able to set enforcement at a low level may not be very attractive from a fairness perspective. When fairness is an issue and accuracy is not too low, an interior solution ($f < f_{max}$) hence becomes more likely. Still, the possibility that e^1_f should be zero exists. A necessary condition for this to be optimal is that $1/p'(0) > -Z(\overline{g})\mu(f_f)\int_{\overline{x}_f}^{1} v_n(x; e^2_f)dx) + (1 - Z(\overline{g}))Q(0)\int_{\overline{x}_f}^{1} v_c(x; e^2_f)dx +$
$(1 - Z(\overline{g}))Q(f_f)\int_{\overline{x}_f}^{1} v_c(x; e^2_f)dx$

The interpretation is that enforcement should take place if it is not too ineffective in apprehending people and if apprehension is beneficial in the sense that it creates more fairness than unfairness.

e^2_f : $\frac{\partial W_f}{\partial e^2}$ is given by

$$(1 - Z(\overline{g}))p(e^1)\frac{\partial(\int_{\overline{x}}^{1} v_c(x; e^2)dx)}{\partial e^2}(Q(0) + Q(f)) - Z(\overline{g}))p(e^1)\frac{\partial(\int_{\overline{x}}^{1} v_n(x; e^2)dx)}{\partial e^2}\mu(f)$$

$$= \frac{\partial C}{\partial e^2}(Q(0) + Q(f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^2} Z(\overline{g})\mu(f):$$

Note that $\frac{\partial(\int_{\overline{x}}^{1} v_n(x;e^2)dx)}{\partial e^2} = \frac{\partial T_1}{\partial e^2}$ is negative. Precision creates the benefit that guilty persons may be more adequately punished, which is expressed by $\frac{\partial C}{\partial e^2}(Q(0) +$

$Q(f))(1 - Z(\bar{g}))$ as well as the benefit that innocent people will less often be convicted, which is expressed by the positive term $-\frac{\partial T_1}{\partial e^2} Z(\bar{g})\mu(f)$:

In an interior maximum $(e_f^1; e_f^2; f_f; \bar{x}_f)$ it must be the case that:

$$\frac{\partial C}{\partial e^2}(Q(0) + Q(f_f))(1 - Z(\bar{g})) - \frac{\partial T_1}{\partial e^2} Z(\bar{g})\mu(f_f) = 1$$

If

$$(1 - Z(\bar{g}))p(e_f^1)\frac{\partial(\int_{\bar{x}_f}^{\infty} v_c(x;0)dx)}{\partial e^2}(Q(0)+Q(f_f)) - Z(\bar{g}))p(e_f^1)\frac{\partial(\int_{\bar{x}_f}^{\infty} v_n(x;0)dx)}{\partial e^2}\mu(f_f)$$

$$< 1$$

and we assume that informativeness is a concave function of e^2[16]; then $e_f^2 = 0$:

f_f :

$$\frac{\partial W_f}{\partial f} = -Z(\bar{g})p(e^1)\int_{\bar{x}}^{\infty} v_n(x;e^2)dx\mu'(f) + (1 - Z(\bar{g}))p(e^1)\int_{\bar{x}_f}^{\infty} v_c(x;e^2)dxQ'(f)$$

Hence: $\frac{\partial W_f}{\partial f} = C(1 - Z(\bar{g}))Q'(f) - T_1 Z(\bar{g})\mu'(f)$

In an interior maximum $(e_f^1; e_f^2; f_f, \bar{x}_f)$ it must hence be the case that:

$$\frac{\mu'(f_f)}{Q'(f_f)} = \frac{C}{T_1}\frac{(1 - Z(\bar{g}))}{Z(\bar{g})}$$

where T_1 and C are, respectively, the probabilities of sanctioning the innocent and sanctioning the guilty given $(e_f^1; e_f^2; f_f, \bar{x}_f)$[17]. Increasing the fine imposes a cost on those who are convicted innocently while it may confer a benefit in punishing the guilty in a more adequate fashion. Assuming that μ is convex and Q is concave, the optimal sanction will be higher the greater the accuracy of trial $(\frac{C}{T_1})$ and the higher the crime-rate $\frac{(1 - Z(\bar{g}))}{Z(\bar{g})}$. The optimal sanction may well be lower but may

[16] I.e. $\frac{\partial(\int_{\bar{x}}^{\infty} v_c(x;e_2)dx)}{\partial e_2}$ is decreasing in e_2 and $\frac{\partial(\int_{\bar{x}}^{\infty} v_n(x;e_2)dx)}{\partial e_2}$ is increasing in e_2 (from one negative number to a numerically smaller negative number).

[17] Sufficient conditions for an interior maximum, $0 < f < \infty$; are: $Z(\bar{g})\int_{\bar{x}}^{\infty} v_n(x;e_2)dx\mu'(0) < (1 - Z(\bar{g}))\int_{\bar{x}}^{\infty} v_c(x;e_2)dxQ'(0)$

and $Z(\bar{g})\int_{\bar{x}}^{\infty} v_n(x;e_2)dx\mu'(f) > (1 - Z(\bar{g}))\int_{\bar{x}}^{\infty} v_c(x;e_2)dxQ'(f)$ for $f > K$ where K is some high number. The second condition wecan assume to be fulfilled, but the first may not be, and we will keep the possibility in mind that fairness considerations may suggest not sanctioning at all.

also be higher than the maximal. (If it is higher than the maximal sanction we assume that it will be set at the maximal level f_{max} since this is the maximum amount people can effectively be sanctioned. The assumption here is that people are concerned with the level of sanctions actually carried out and not with the symbolic value of imposing a high sanction).

4.1.2. The Optimal Standard of Proof

$$\frac{\partial W_f}{\partial \overline{x}} = p(e^1)\left[Z(\overline{g})v_n(\overline{x};e^2)\mu(f) - ((1 - Z(\overline{g}))v_c(\overline{x};e^2)(Q(0) + Q(f_f))\right]$$

In an interior maximum $(e_f^1; e_f^2; f_f, \overline{x}_f)$, it must hence be the case that:

$$\frac{v_n(\overline{x}_f;e_f^2)}{v_c(\overline{x}_f;e_f^2)} = \frac{1 - Z(\overline{g})}{Z(\overline{g})} \cdot \frac{(Q(0) + Q(f_f))}{\mu(f_f)}:$$

For proof, see Appendix B. A similar result is obtained by Miceli [11].

The question arises, however, whether there exists an interior optimal level of $\overline{x}_f$: There may not exist a solution to the equation above or the solution may conceivably represent a global minimum. We can distinguish three cases:

Case 1: $\lim_{\overline{x} \to -\infty} \frac{v_n(\overline{x};e_f^2)}{v_c(\overline{x};e_f^2)} < \frac{1-Z(\overline{g})}{Z(\overline{g})} \cdot \frac{(Q(0)+Q(f_f))}{\mu(f_f)}$. In this case there is no solution to the equation; it never adds to welfare to increase the standard of proof and hence $\overline{x}_f = -\infty$: It is the case where it is overwhelmingly important to punish the guilty and unimportant not to punish the innocent.

Case 2: $\lim_{\overline{x} \to \infty} \frac{v_n(\overline{x};e_f^2)}{v_c(\overline{x};;e_f^2)} > \frac{1-Z(\overline{g})}{Z(\overline{g})} \cdot \frac{(Q(0)+Q(f_f))}{\mu(f_f)}$. In this case it always adds welfare to increase the standard of proof since it is all-important not to punish the innocent, hence $\overline{x}_f = \infty$: (In this partial analysis we take as given the level of e_f^1: In the general analysis $\overline{x}_f = \infty$ will be ruled out as an equilibrium since a better way of achieving the result that nobody is convicted is to let e^1 equal zero).

Case 3: If neither of the two conditions above obtains, then there exists a solution to the equation (by continuity and by the fact that $\frac{v_n(\overline{x};e_f^2)}{v_c(\overline{x};e_f^2)}$ is decreasing). Denote this solution $\overline{x}_f$: The second order condition for a maximum becomes: If $v_c'(\overline{x}_f;e_f^2) < 0$:

$$\frac{v_n'(\overline{x}_f;e_f^2)}{v_c'(\overline{x}_f;e_f^2)} > \frac{1 - Z(\overline{g})}{Z(\overline{g})} \cdot \frac{(Q(0) + Q(f_f))}{\mu(f_f)}:$$

If $v_c'(\bar{x}_f; e^2) > 0$:

$$\frac{v_n'(\bar{x}_f; e_f^2)}{v_c'(\bar{x}_f; e_f^2)} < \frac{1 - Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0) + Q(f_f))}{\mu(f_f)}.$$

Both of these conditions are fulfilled when the likelihood ratio is monotone: $\frac{v_n(\bar{x}; e_f^2)}{v_c(\bar{x}; e_f^2)}$ decreasing in $\bar{x}$ implies that

$$v_n'(\bar{x}; e_f^2) v_c(\bar{x}; e_f^2) - v_c'(\bar{x}; e_f^2) v_n(\bar{x}; e_f^2) < 0$$

which implies that $\frac{v_n'(\bar{x}; e_f^2)}{v_c'(\bar{x}; e_f^2)} > \frac{v_n(\bar{x}; e_f^2)}{v_c(\bar{x}; e_f^2)}$ when $v_c'(\bar{x}; e_f^2) < 0$ and $\frac{v_n'(\bar{x}; e_f^2)}{v_c'(\bar{x}; e_f^2)} < \frac{v_n(\bar{x}; e_f^2)}{v_c(\bar{x}; e_f^2)}$ when $v_c'(\bar{x}; e_f^2) > 0$. The claim then follows from: $\frac{v_n(\bar{x}_f; e_f^2)}{v_c(\bar{x}_f; e_f^2)} = \frac{1 - Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0)+Q(f_f))}{\mu(f_f)}$.

We can thus state the following proposition:

Proposition 4.1. If only fairness matters, and if the likelihood ratio is monotone, the optimal standard of proof $\bar{x}_f$ is given by:

$$\frac{v_n(\bar{x}_f; e_f^2)}{v_c(\bar{x}_f; e_f^2)} = \frac{1 - Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0) + Q(f_f))}{\mu(f_f)}.$$

unless $\lim_{\bar{x} \to -\infty} \frac{v_n(\bar{x}; e_f^2)}{v_c(\bar{x}; e_f^2)} < \frac{1 - Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0)+Q(f_f))}{\mu(f_f)}$ or $\lim_{\bar{x} \to \infty} \frac{v_n(\bar{x}; e_f^2)}{v_c(\bar{x}; e_f^2)} > \frac{1 - Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0)+Q(f_f))}{\mu(f_f)}$. The exceptions represent the extreme outcomes of either 'always' or 'never' convicting.

To interpret this standard of evidence, note that $\frac{(Q(0)+Q(f_f))}{\mu(f_f)}$ expresses the gravity of false acquittals relative to false convictions. This ratio determines how many offenders we are willing to let go unpunished in order to avoid that one person is falsely convicted. Assume that we are willing to let exactly ten guilty people go free each time we convict an innocent person. Consider a situation where one in a thousand people are criminals. Then the critical level of evidence is that which is ten thousand times more likely to be forthcoming when the person has in fact violated the law than when he has not. When this kind of evidence is presented in court, we should expect that only in one out of ten cases the person will not be guilty, which is the ratio we are willing to accept.

Note that:

1. The ratio of falsely convicted to falsely acquitted may be very di¤erent from the ratio $\frac{(Q(0)+Q(f_f))}{\mu(f_f)}$: The latter ratio expresses the relative costs of false acquittals and false convictions that are relevant when considering marginal changes in the standard of proof. So it is only at the margin that the ratio of falsely acquitted and falsely convicted equals $\frac{(Q(0)+Q(f_f))}{\mu(f_f)}$: This distinction will be illustrated in the example below. It follows that if we estimate that the actual number of falsely acquitted is e.g. twenty times higher than the number of falsely convicted while we think the cost of falsely convicting is only ten times higher than the cost of falsely convicting, we should not jump to the conclusion that the standard of proof should be lowered.

2. In the pure deterrence case the optimal standard of proof was a preponderance of the evidence-standard: $v_n(\overline{x}; e_f^2)=v_c(\overline{x}; e_f^2) = 1$: When only fairness matters, the standard of proof is stricter than the preponderance of the evidence-standard when there are more innocent than guilty people and when the 'cost' of punishing the innocent is greater than that of letting a guilty person o¤ without punishment. Since these assumptions are often met, we can say that the standard of proof is in most normal circumstances stricter when fairness is the issue than when deterrence is the issue. However, we cannot exclude that in rare situations fairness concerns may call for a more lenient standard of proof than $\overline{x}_d$.

These results seem broadly in line with empirical observation, e.g. that the standard of evidence is stricter for criminal o¤ences than for civil o¤ences.

On the Non-use of Ex-ante Probability of Guilt It is worth noting that the the standard of proof based on the evidence will be stricter the greater the fraction of law abiding citizens. This reveals an important di¤erence between the pure deterrence and the pure fairness case. When only deterrence matters, and sanctions are monetary, the optimal standard of proof does not depend on à priori probabilities of guilt or innocence, as noted above. In contrast, as the result just stated reveals, these play an important role when the aim is justice. The reason behind this result is that probabilities of making type-1 and type-2 errors when using a given standard based only on the evidence depends on à priori probabilities of a person being criminal. The probabilities of making errors are in themselves not important when the aim is to control behavior; important is then how to construct the system such that a person can change the probability of sanction through his action. In the example mentioned in the introduction, if society A cares for justice they will acquit the alleged criminal on the grounds that the probability that it is all a mistake is high while society B should pass a verdict of guilty on the grounds that doing so is likely to create more fairness than unfairness. By following this policy, society A will rarely let a criminal off and often let an innocent man go, while society B will also find their policy optimal since they will often sanction a guilty and rarely convict an innocent person[18].

We shall now turn to the realistic case where both fairness and deterrence matter.

4.2. The Case With Both a Deterrence and a Fairness Motive

In the section above with only a concern for fairness the partial derivatives $\frac{\partial W}{\partial f}; \frac{\partial W}{\partial e^1}; \frac{\partial W}{\partial e^2}$ and $\frac{\partial W}{\partial \bar{x}}$ were found:

We now turn to that part of the derivative of social welfare that depends on a deterrent effect.

Derivation and rewriting yields the following result:

$$\frac{\partial W}{\partial \bar{e}} = [(h - \bar{e}) + Q(0)T_2 - T_1\mu(f) - cQ(f)]$$

The positive effects of deterrence are: the smaller harm to society $(h - \bar{e})$, fewer false convictions $Q(0)T_2$ since fewer guilty people ($Q(0)$ is likely to be positive, it is defined as the disutility of not punishing a guilty person); the negative effects are: more falsely convicted, $T_1\mu(f)$; since more innocent people, and fewer people justly punished; $cQ(f)$; since fewer guilty people ($Q(f)$ is likely to be positive, it is defined as the utility of punishing a criminal). Following Shavell and Polinsky we can define the full deterrence ideal as that level of $\bar{e}$ where $\frac{\partial W}{\partial \bar{e}} = 0$ hence where $[(h - \bar{e}) + Q(0)T_2 - T_1\mu(f) - cQ(f)] = 0$:

In the following section, we analyze the optimal levels of sanctions and enforcement. The section can be skipped by the reader who is only interested in the results concerning the optimal standard of proof.

[18] This difference reveals a conflict between two different notions of fairness. If the two societies are two neighborhoods of the same country (e.g. a poor and a rich) using á priori probabilities will conflict with the notion that there should be equality before the law. Note that in the model above this trade-off does not occur since there is only one group of people; people within this group are assumed indistinguishable from the point of view of the judge.

4.2.1. Optimal Sanctions and Enforcement

The full derivatives can now be stated:

f : The condition which must be fulfilled in an interior maximum, is:

$$\alpha z(\underline{g})\frac{\partial W}{\partial \underline{g}} + Q'(f)T_1 Z(\overline{g}) - \theta'(f)C(1 - Z(\overline{g})) = 0$$

where $\alpha = \frac{\partial \underline{g}}{\partial f}$:

The possibility that $f = f_{max}$ is worth examining since W is not differentiable in f_{max}. Two conditions must be met for f_{max} to be a optimum. When $f = f_{max}$ the right-derivative $\frac{\partial \underline{g}}{\partial f}_{(+)}$ is zero, thus $\alpha_{(+)} = 0$; and hence $cQ'(f_{max})(1 - Z(\underline{g})) - T_1\theta'(f_{max})Z(\underline{g})$ must be negative (if $cQ'(f_{max})(1 - Z(\underline{g})) - T_1\theta'(f_{max})Z(\underline{g}) > 0$ increasing f will be worthwhile since there is no deterrent effect but the impact on fairness is positive). And, as the second condition,

$$\alpha_{(-)}z(\underline{g})[(h - \underline{g}) + Q(0)T_2 - T_1\theta(f_{max}) - cQ(f_{max})] +$$
$$cQ'(f_{max})(1 - Z(\underline{g})) - T_1\theta'(f_{max})Z(\underline{g}) \geq 0$$

where $\alpha_{(-)}$ is the left-derivative of $\frac{\partial \underline{g}}{\partial f}$ at f_{max}. We see that if $Q'(f_{max})$ is sufficiently negative (the optimal penalty for the guilty is thought to be less than the maximal penalty) and $\theta'(f_{max})$ sufficiently positive (lowering the size of the sanction from the maximal level will decrease the conceived degree of unfairness, cet. par.), then the optimal sanction will be lower than the maximal. On the other hand, we also see that if people want to punish criminals harshly for retributive reasons it may be that fairness considerations (despite fear of punishing the innocent) overall lead to higher sanctions than what is warranted from the viewpoint of prevention.

e^1 : Define $\frac{\partial \underline{g}}{\partial e^1}$ which is equal to $p'(e^1)f\int_A v_c(x; e^2) - v_n(x; e^2)dx$ as β:

The first-order condition becomes:

$$\beta z(\underline{g})\frac{\partial W}{\partial \underline{g}} + \frac{\partial C}{\partial e^1}(Q(0) + Q(f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^1}Z(\overline{g})\theta(f) = 1$$

The possibility that e^1 should be set at its minimal level exists in this situation as well as the possibility of not enforcing at all ($e^1 = 0$); deriving when these cases arise is quite straightforward but will not be undertaken here.

e^2 : Define $\frac{\partial \underline{g}}{\partial e^2}$ as γ: The first-order condition becomes: $\gamma z(\underline{g})\frac{\partial W}{\partial \underline{g}} + \frac{\partial C}{\partial e^2}(Q(0) + Q(f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^2}Z(\overline{g})\theta(f) = 1$:

4.2.2. The Standard of Proof

Denote $\frac{\partial q}{\partial \overline{x}} = -f \cdot p(e^1)(v_c(\overline{x}, e^2) - v_n(\overline{x}, e^2))$ by γ.

We obtain the first-order condition:

$$\gamma z(\mathbf{e})\frac{\partial W}{\partial \mathbf{e}} + p(e^1)\left[Z(\overline{g})v_n(\overline{x}, e^2)\theta(f) - ((1 - Z(\overline{g}))v_c(\overline{x}, e^2)(Q(0) + Q(f_f))\right] = 0$$

$\Longrightarrow$

$$\gamma z(\mathbf{e})\frac{\partial W}{\partial \mathbf{e}} + \frac{\partial T_1}{\partial \overline{x}}Z(\overline{g})\theta(f) - \frac{\partial C}{\partial \overline{x}}(Q(0) + Q(f))(1 - Z(\overline{g})) = 0$$

Note that γ is negative if $\overline{x} > \overline{x}_d$ ($\overline{x}_d$ maximizes deterrence) and positive if $\overline{x} < \overline{x}_d$.

The simultaneous set of equations which must be fulfilled in an interior optimum ($0 < f < f_{max}$) is hence:

$$\alpha z(\mathbf{e})\frac{\partial W}{\partial \mathbf{e}} + Q'(f)T_1 Z(\overline{g}) - \theta'(f)C(1 - Z(\overline{g})) \quad \text{condition for } f$$

$$\beta z(\mathbf{e})\frac{\partial W}{\partial \mathbf{e}} + \frac{\partial C}{\partial e^1}(Q(0) + Q(f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^1}Z(\overline{g})\theta(f) = 1 \quad \text{cond. for } e^1$$

$$\delta z(\mathbf{e})\frac{\partial W}{\partial \mathbf{e}} + \frac{\partial C}{\partial e^2}(Q(0) + Q(f))(1 - Z(\overline{g})) - \frac{\partial T_1}{\partial e^2}Z(\overline{g})\theta(f) = 1 \quad \text{cond. for } e^2$$

$$\gamma z(\mathbf{e})\frac{\partial W}{\partial \mathbf{e}} + \frac{\partial T_1}{\partial \overline{x}}Z(\overline{g})\theta(f) - \frac{\partial C}{\partial \overline{x}}(Q(0) + Q(f))(1 - Z(\overline{g})) = 0 \quad \text{cond. for } \overline{x}$$

This is the simultaneous set of four equations which determines the four variables $(f, e^1, e^2, \overline{x})$ when the solution is interior. It should be noted that interior solutions are more likely when fairness matters since the outcome where a very high penalty is imposed while very few people are apprehended will tend to lack appeal from a fairness viewpoint, as argued above. From these equations we can do comparative statics analysis to derive the consequences for criminal policy of e.g. an increase in the μ or Q-functions or of a decrease in the perceived level of accuracy, as has followed from the recent use of DNA-tests. However, the main concern lies in establishing the optimal level of the standard of proof. It can now be established that the optimal standard of proof will usually lie in between the level which is optimal from a deterrence perspective ($\overline{x}_d$) and that which is optimal from a pure deterrence perspective $\overline{x}_f$. Recall that $\overline{x}_d$ is given by:

$$\frac{v_n(\overline{x}_d; e^2)}{v_c(\overline{x}_d; e^2)} = 1$$

while $\overline{x}_f$ is given by

$$\frac{v_n(\overline{x}_f; e^2)}{v_c(\overline{x}_f; e^2)} = \frac{1 - Z(\hat{e})}{Z(\hat{e})} \cdot \frac{(Q(0) + Q(f))}{\mu(f)}.$$

The claim is hence that the optimal likelihood ratio for conviction will usually lie in the interval $\left\langle 1; \frac{1-Z(\hat{e})}{Z(\hat{e})} \cdot \frac{(Q(0)+Q(f))}{\mu(f)} \right\rangle$ when both deterrence and fairness matter.

There are two basic cases to be distinguished: One where fairness concerns constrains deterrence, which we will define here as the case where the full deterrence ideal (defined above) is higher than the fairness ideal, and the other where fairness concerns stretch deterrence. More accurately, we will define the first case as that where it would be optimal to set punishment at a higher level were it not for pure fairness reasons, and the second as the case where it would be optimal to set punishment at a lower level absent (pure) fairness concerns.

Proposition 4.2. If fairness constrains deterrence, the optimal standard of proof will lie between $\overline{x}_d$ and $\overline{x}_f$.

Proof: When fairness constrains deterrence, from a full deterrence perspective it would increase welfare to impose higher f. Hence,

$$\alpha z(\hat{e})\left[(h - \hat{e}) + Q(0)T_2 - T_1\mu(f) - cQ(f)\right] > 0$$

This implies that

$$z(\hat{e})\left[(h - \hat{e}) + Q(0)T_2 - T_1\mu(f) - cQ(f)\right] > 0;$$

since $\alpha > 0$. There are two possibilities to consider.

a) $\overline{x}_d < \overline{x}_f$:

b) $\overline{x}_d \geq \overline{x}_f$

In a), the typical case, $\overline{x}^*$ will be higher than $\overline{x}_d$ but lower than $\overline{x}_f$: This can be proven by contradiction:

1) If $\overline{x}^* < \overline{x}_d$ then $\gamma > 0$ by the fact that $\overline{x}_d$ is the deterrence ideal ($\frac{\partial \underline{e}}{\partial \overline{x}} = 0$); hence $\gamma z(\underline{e})\,[(h - \underline{e}) + Q(0)T_2 - T_1\theta(f) - cQ(f)] > 0$ from above. Then $\frac{\partial W_f}{\partial \overline{x}}$ must be negative which by the above-proven concavity of $\frac{\partial W_f}{\partial \overline{x}}$ implies that $\overline{x}^* > \overline{x}_f$. But this is a contradiction since $\overline{x}^*$ cannot be smaller than $\overline{x}_d$ and larger than $\overline{x}_f$ when $\overline{x}_d < \overline{x}_f$.

2) If $\overline{x}^* > \overline{x}_f > \overline{x}_d$; then $\gamma < 0$; hence $\gamma z(\underline{e})\,[(h - \underline{e}) + Q(0)T_2 - T_1\theta(f) - cQ(f)] < 0$ and $\frac{\partial W_f}{\partial \overline{x}}$ must be positive. This implies $\overline{x}^* < \overline{x}_f$; a contradiction.

In b) which is the atypical case, $\overline{x}^*$ will be higher than $\overline{x}_f$ but lower than $\overline{x}_d$: The same proof through contradiction can be applied here.

QED

Proposition 4.3. If fairness concerns lead to overdeterrence compared with the full deterrence ideal, the optimal standard of proof will not lie between $\overline{x}_d$ and $\overline{x}_f$:

Proof: This follows from the proof just given when the fact that

$$[(h - \underline{e}) + Q(0)T_2 - T_1\theta(f) - cQ(f)]$$

is now negative is taken into account. This fact changes the sign of

$$\gamma z(\underline{e})\,[(h - \underline{e}) + Q(0)T_2 - T_1\theta(f) - cQ(f)]$$

and hence leads to the opposite result of the above. QED

In the case where $\overline{x}_d < \overline{x}_f$, consider the point $\overline{x} = \overline{x}_f$: This point maximizes fairness in a sense (keeping deterrence constant). If $\overline{x}$ is decreased there will be two effects: More deterrence (by the fact that $\overline{x}$ will be closer to $\overline{x}_d$) and less fairness. Both of these effects are bad since we start from a situation of overdeterrence. If, on the other hand $\overline{x}$ is increased, deterrence will suffer which is good and the negative effect on fairness will only be second-order when the increase is small. An increase beyond $\overline{x}_f$ will hence improve welfare[19].

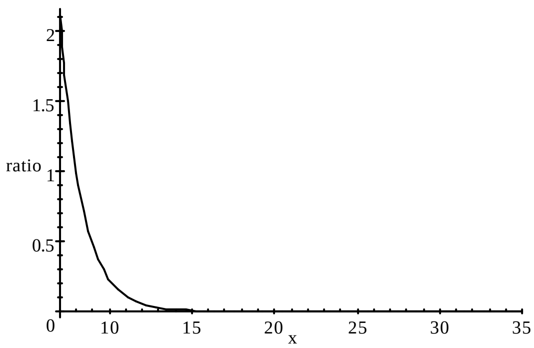
When fairness concerns lead to overdeterrence it may hence be optimal to stretch the standard of proof beyond what is optimal from a pure fairness perspective in order to mitigate over-deterrence.

[19] By a similar reasoning it cna be shown that it may be optimal to decrease $\overline{x}$ below $\overline{x}_d$: We shall not analyze when it is optimal to sacrifice deterrence by sanctioning even when it is more

4.3. Examples

The following example is intended to illustrate the kind of trade-off between deterrence and fairness described above. In reality, the numbers will be uncertain. Still, it is important to know how the standard of proof should be determined if the variables were all known. Consider the crime of sexual offense (such as violation). Assume that the number of offenders to non-offenders is in the range of 1:10000 (a number which is not fixed but may to some extent be affected by sanctions). Let there be 10 mill people in the population out of which 1000 people are likely to commit a serious sexual offense at some point. Assume further that if the crime is committed, the evidence x that will come up in a thorough police investigation has a random component; we assume the evidence is normally distributed with mean equal to 14 and standard deviation equal to 4. Its density is $\frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-14)^2}{32}}$. If the crime is not committed, the mean will equal 2 and the standard deviation will be 4. The density function will be $\frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-2)^2}{32}}$.

The likelihood ratio equals $\frac{\frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-2)^2}{32}}}{\frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-14)^2}{32}}}$. It is decreasing as can be seen graphically:



Let the fraction $\frac{(Q(0)+Q(f))}{\mu(f)}$ equal 1/10 such that $\frac{1-Z(e)}{Z(e)} * \frac{(Q(0)+Q(f))}{\mu(f)}$ is approximately 1:100000. Set $\mu(f) = 10$.

---

likely that a crime has not been committed. Note, however, that in the case where fairness dominates deterrence concerns, in the sense that it is optimal to create overdeterrence, it seems unlikely that this should be optimal.

Likewise, we will not explore the case where $\overline{x}_d > \overline{x}_f$ further; as mentioned the more realistic case is $\overline{x}_d < \overline{x}_f$.

Assume that $p(e^1) = 1$, investigations will take place. Assume further that for each year's increase in expected sanction the number of potential offenders who will be deterred by the increase equals z. Part of the deterrence effect is likely to be due to stigma. Assume that f is five years' imprisonment.

The question is now how the size of harm h and z affect the optimal standard of proof.

If fairness is the only motive to punish, then given the accuracy of the evidence, punishment would virtually never occur. Thus,

$$\frac{\frac{1}{4\sqrt{2\pi}}e^{-\frac{(x-2)^2}{32}}}{\frac{1}{4\sqrt{2\pi}}e^{-\frac{(x-14)^2}{32}}} = 1/100000 \Rightarrow x = 23.351$$

To find the probability of conviction if one has committed the act, this number is inserted in the cumulative normal distribution. It equals

$$1 - \text{NormalDist}(8 + \frac{4}{3}\ln 100\,000; 14; 4) = .0097$$

If one sets the standard of proof at this level, out of 1000 offenders, only $1000 \times .0097 = 9.7$ will be punished. Thus, a vast majority of the guilty will go free.

If one has not committed the act, there is almost zero probability that one will be convicted:

$(1 - \frac{1}{4\sqrt{2\pi}}\int_{-\infty}^{8+\frac{4}{3}\ln 100\,000} e^{-\frac{(x-2)^2}{32}}dx) \times 10^7 = .471$. Thus, only half a person will be innocently convicted.

While the number of unfairly punished is low, the number of people who are guilty but nevertheless go free is very high. Yet from a fairness perspective this is the best one can do; the problem is that the evidence is not sufficiently conclusive to allow a satisfactory number of convictions without an unsatisfactory number of false convictions. (In cases of sexual offense, a higher level of certainty can no doubt be obtained). Thus, if the likelihood ratio is set at 1/1000, the same calculations as above yield the outcome that 700.0 will be punished unfairly while 788.87 will go free although they are guilty. If set at 1/10000 942 people will go free and 24 will be punished unfairly. Both of these alternatives ar worse in terms of overall unfair costs.

Note how misleading it is to assume that the ratio $\frac{(Q(0)+Q(f))}{\mu(f)}$ will in optimum equal the actual ratio of falsely acquitted to falsely convicted.

Let us analyze how strong the deterrence effect has to be for it to be optimal to set the standard of proof at 1/10000 rather than at 1/100000. The unfairness cost of a standard of proof equal to 1=10000 equals 942 + 240 = 1182 The unfairness cost of a standard of proof equal to 1=100000 equals 990; 3 + 4:71 t 995:

The expected sanction when the standard is 1=100000 equals approximately two weeks (…ve years times the probability of charges being levelled of less than one per cent). Naturally, stigma may follow and increase the cost substantially for some types of o¤enders.

The expected sanction when the standard is 1=10000 equals approximately a month and a half since the probability of conviction is now 2,4%. If we think that the loss of utility su¤ered by society, including of course mainly the victim, from three sexual violations equals the loss from falsely convicting one person, then the number of deterred o¤enders by an increase from 1% to 2,4% in conviction probability needs to be $\frac{1182 ¡ 995}{3:33}$ = 56: 156: Out of 1000 o¤enders we then need to deter 56 for it to be worthwhile. If we consider the disutility of a sexual violation as serious as the disutility of a false conviction, we need to deter around 19 people in order for it to be optimal to lower the standard of proof.

To illustrate in a slightly di¤erent way how our results may be applied in a real case, we may imagine ourselves to be a judge who must decide whether or not to convict an alleged o¤ender in a given case where the evidence is given as $\mathbf{x}$. Let us imagine that we can establish with some degree of con…dence that the probability of the given evidence forthcoming if the person has committed the crime (v_c ($\mathbf{x}$)) would be, say, around 25 % while the probability that the same evidence would be forthcoming if the person has not committed the crime (v_n ($\mathbf{x}$)) would be something like 2 %: We can also imagine that the à priori probability of a person committing the crime (1 ¡ $Z(\mathbf{e})$) is 5 %: If we know the relative seriousness of false convictions vs. false acquittals, i.e. if we know $\frac{(Q(0)+Q(f))}{\mu(f)}$; we would be able to make a …rst test as to whether the person should be found guilty: Given $\mathbf{x}$; we can calculate the fairness related costs and bene…ts of convicting the person. The probability of $\mathbf{x}$ is given by

$$prob(x) = prob(\mathbf{x}\ j\ crime) ¤ prob(crime) + prob(\mathbf{x}\ j\ no\text{-}crime) ¤ prob(no\ crime):$$

The probability that he is guilty is calculated by Bayesian updating as follows: prob(crimej $\mathbf{x}$) = $\frac{prob(\mathbf{x}\ j\ crime)¤prob(crime)}{prob(\mathbf{x})}$ = $\frac{:25¤0:05}{:0315}$ = :4 while prob(no-crimej $\mathbf{x}$) = :6

Thus, with this evidence and this à priori probability, the person should not be convicted on grounds of fairness unless one considers it more (one and a half times more) important to sanction a guilty person than to not sanction an innocent.

If numbers were different, we may discover at this point that whether or not there is a deterrent effect we should convict the person on fairness grounds alone. If, e.g. we consider it ten times more important not to sanction than an innocent than to sanction a guilty, and if the numbers above were $v_c(\mathbf{e}) = 90\ \%$; $v_n(\mathbf{e}) = 0.1\ \%$ and $1 - Z(\mathbf{e}) = 50\%$; then prob(crime given $\mathbf{e}$) $= \frac{.9 \times 0.5}{.4505} = .99889$ and we would sanction the person unless the ratio of seriousness is around 1:1000.

Returning to the former case where the probability of unlawful behavior was only 40%, the question is how we should consider the deterrent effect of convicting given $\mathbf{e}$: Let us assume $f = 1$ and $p(e^1) = 1$. Calculations above have given us the expression: $f \times (v_c(\mathbf{e}) - v_n(\mathbf{e})) \times z(\mathbf{e}) \times h$ as the size of the deterrent welfare-effect (ignoring cross-effects, i.e. mainly that the deterred people risk being falsely convicted). Thus, when considering the deterrent effect we should calculate the change in expected sanction from convicting on the given evidence: $(0.25 - 0.02) \times f = 0.23$. We then need to estimate how many people are deterred by one unit of extra deterrence $z(\mathbf{e})$, in other words we must ascertain the number of people who will be deterred by the increase in sanction just calculated and multiply this number by h. The relevant comparison hence becomes:

Cost of conviction: $\mu(1) \times 0.6$

Cost of acquittal: $(Q(0) + Q(f)) \times 0.4 + 0.23 \times z(\mathbf{e}) \times h$:

Note that apart from preferences (μ and Q and h) what the judge needs to know is $v_c(\mathbf{e})$, $v_n(\mathbf{e})$, $Z(\mathbf{e})$ and $z(\mathbf{e})$.

5. An Extension

Several issues have not been touched upon in this paper. We shall only mention one: risk-aversion, which is likely to be an important factor. It enters both into the incentive constraint and into the social welfare consequences of applying sanctions. Applying sanctions with some probability incurs a social loss when people are risk-averse. If people are strongly risk-averse, they may also be strongly averse to the possibility of being falsely convicted, and this factor seems likely to call for a stricter standard of proof. On he other hand, they may also be very averse to the possibility of becoming a victim of crime. How exactly risk-aversion affects the optimal standard of proof will be left to future research.

6. Conclusion

The main points are the following:

- When only deterrence matters and sanctions are monetary, the optimal standard of proof is the preponderance-of-the-evidence standard. This means that the likelihood ratio $\frac{v_n(\bar{x}_d; e_d^2)}{v_c(\bar{x}_d; e_d^2)}$ equals one.

- When only deterrence matters and sanctions impose a cost on society, the standard of proof is higher than the the preponderance-of-the-evidence standard.

-When only fairness matters, the optimal likelihood ratio $\frac{v_n(\bar{x}_f; e_f^2)}{v_c(\bar{x}_f; e_f^2)}$ equals $\frac{1-Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0)+Q(f_f))}{\mu(f_f)}$. It hence reflects the ratio of criminals in the population and the relative severity of punishing the innocent vs. acquitting the guilty.

When both deterrence and fairness matter the optimal likelihood ratio, except when there is overdeterrence, lies in the interval $\left[1; \frac{1-Z(\bar{g})}{Z(\bar{g})} \times \frac{(Q(0)+Q(f_f))}{\mu(f_f)}\right]$. Where in the interval it lies depends on the weight of fairness vs. deterrence motives.

-When only deterrence matters and sanctions are monetary, ex-ante probabilities (the fraction of criminals in the population) hence do not matter whereas ex-ante probabilities are important for the standard of proof when the issue is fairness.

-It is important to view the optimal standard of proof in terms of marginal changes. The question is what the effect is of a marginal change in the standard of proof. Hence, if the general sentiment in society is roughly that the cost of punishing an innocent person is ten times higher than that of falsely acquitting a guilty person ($\frac{(Q(0)+Q(f_f))}{\mu(f_f)}$ is $1/10$), while it is believed that twenty people go free every time an innocent is sanctioned, this does not in itself tell us that the standard of proof should be lowered. The relevant question is whether a marginal change in the standard of proof creates a ratio of extra false convictions to extra false acquittals that corresponds to their relative costliness.

-Apart from preferences (μ and Q and h) what the judge needs to know to decide a case where the evidence is $\bar{\mathbf{x}}$ is $v_c(\bar{\mathbf{x}})$, $v_n(\bar{\mathbf{x}})$, $Z(\bar{\mathbf{g}})$ and $z(\bar{\mathbf{g}})$.

7. Appendix A

We will show that when the actions of more than one person are involved the analysis becomes more complicated and a result such as that obtained above about the preponderance of the evidence being optimal when only deterrence matters does not apply. We do not model the case in generality but take a simplified model to show a general point. Assume that two people may commit a murder. If A does it, there will be a density-function $f_A(x; e^2)$ of the evidence vector x while if B does it there will be a density-function $f_B(x;e^2)$. Assume also that other assumptions of the model in the text are maintained, e.g. that sanctions involve zero cost to society. If only deterrence matters, it may be thought that it would be optimal to convict A if

$$f_A(x;e^2) > f_B(x;e^2)$$

and vice-versa. However, this is not generally so. Imagine e.g. that there are three states of nature: $x_1 = (1,0,0)$;

$x_2 = (0,1,0)$ and $x_3 = (0,0,1)$. If B commits the crime the three states obtain with the probabilities $(0.9, 0.1, 0)$ while if A commits the crime the probabilities of the three states are $(0.8, 0.09, 0.11)$. If x_1 occurs, B should be punished if the rule above applies ($f_A(x;e^2) < f_B(x;e^2)$) . But this may be enough to deter B and it may therefore be optimal to sanction A in the other two states to deter him also, although in one of them $f_A(x;e^2) < f_B(x;e^2)$. Thus, what corresponds to a preponderance of the evidence standard is not necessarily optimal when more than one person is involved.

8. Appendix B

Finding $\overline{x}_f^*$: The criterion function can be written as follows when we only include terms that incorporate $\overline{x}$, and reduce terms:

$$
\begin{aligned}
&-Z(\overline{g})p(e^1)\int_{\overline{x}}^{1} v_n(x;e^2)dx\mu(f^*) \\
&-(1-Z(\overline{g}))(1-p(e^1)\int_{\overline{x}}^{1} v_c(x;e^2)dx))Q(0) \\
&+(1-Z(\overline{g}))p(e^1)\int_{\overline{x}}^{1} v_c(x;e^2)dxQ(f^*)
\end{aligned}
$$

Differentiating with respect to $\overline{x}$ yields:

$$Z(\overline{g})p(e^1)v_n(x;e^2)\mu(f) - (1-Z(\overline{g}))p(e^1)v_c(x;e^2)Q(0) -$$

$$(1-Z(\overline{g}))p(e^1)v_c(x;e^2)Q(f)$$

which must be zero in the optimum. Thus, an interior maximum, if it exists, must fulfill the condition:

$$v_n(\overline{x}_f;e^2)p(e^1)\mu(f) - (1-Z(\overline{g}))p(e^1)v_c(\overline{x}_f;e^2)Q(0) - (1-Z(\overline{g}))p(e^1)v_c(\overline{x}_f;e^2)Q(f)$$

$$= 0$$

which yields the stated condition:

$v_n(\overline{x}_f; e^2) = v_c(\overline{x}_f; e^2) = \frac{(1 - Z(\overline{g}))}{Z(\overline{g})}(Q(0) + Q(f))/\mu(f).$

QED.

References

[1] Andreoni,-James: "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?" Rand-Journal-of-Economics; 22(3), Autumn 1991, pages 385-95.

[2] Becker, Gary: "Crime and Punishment: An Economic Approach", Journal of Political Economy, 1968, 76 (2), pp 169-217.

[3] Diamond, Peter: "Integrating Punishment and E¢ciency Concerns in Punitive Damages for Reckless Disregard of Risks to Others", Working Paper, No-97-19, Department of Economics, MIT, October 1997.

[4] Ehrlich, Isaach: "The Optimum Enforcement of Laws and the Concept of Justice: A Positive Analysis", International Review of Law and Economics, vol 2, No. 1, (June 1982), pp 3-27.

[5] Garoupa, Nuno: "The Theory of Optimal Law Enforcement", Journal of Economic Surveys, vol 11, No. 3, 1997, pp 267-295.

[6] Grossman G. and Katz M: "Plea Bargaining and Social Welfare", Am. Economic Review 73, 749-57.

[7] Kaplow, Louis: "The Value of Accuracy in Adjudication", : An Economic Analysis", Journal of Legal Studies; 23(1), Part 2 Jan. 1994, pages 307-401.

[8] Lando, Henrik: "An Attempt to Incorporate Fairness in an Economic Model of Tort Law", International Review of Law and Economics, vol. 17, No. 4, (December 1997), pp. 575-587.

[9] James Liebman et al: "A Broken System: Error Rates in Capital Cases, 1973-1995", New York University Columbia Law School, Public Law Research Paper No. 15.

[10] Miceli,-Thomas-J: "Optimal Criminal Procedure: Fairness and Deterrence", International-Review-of-Law-and-Economics; 11(1), May 1991, pages 3-10.

[11] Miceli,-Thomas: "Optimal Prosecution of Defendants Whose Guilt Is Uncertain", Journal-of-Law,-Economics-and-Organization; 6(1), Spring 1990, pages 189-202.

[12] Png, Ivan: "Optimal Subsidies and Damages in the Presence of Judicial Error", International Review of Law and Economics, 1986, Vol 6, No. 1, pp 101-105.

[13] Posner, R: "An Economic Approach to the Law of Evidence", John M. Olin Law and Economics working Papers No. 66, 2000, downloadable from www.law.uchicago.edu/Publications/Working/index.html

[14] Reinganum, J.: "Plea Bargaining and Prosecutorial Discretion", Am. Economic Review 78, 713-28.

[15] Rubinfeld, D, and Sappington D: "E¢cient awards and standards of proof in judicial proceedings", Rand Journal of Economics, vol 18, No. 2, 1987, pp. 308-315.

[16] Polinsky,-A.-Mitchell; Shavell,-Steven: "Legal Error, Litigation, and the Incentive to Obey the Law", Journal-of-Law,-Economics-and-Organization; 5(1), Spring 1989, pages 99-108.

[17] Polinsky,-A.-Mitchell; Shavell,-Steven: "The economic theory of public enforcement of law", Journal of Economic Literature, Nashville; Mar 2000, Vol. 38, Iss. 1; pg. 45, 32 pgs.

[18] Polinsky,-A.-Mitchell; Shavell,-Steven; "The Fairness of Sanctions: Some Implications for Optimal Enforcement Policy", Working Paper, Harvard University, Olin Center, 12/98. (Downloadable from http://www.law.harvard.edu/programs/olin_center/papers/abstract247.html).

[19] Shavell,-Steven: Economic Analysis of Accident Law, Harvard University Press, 1987.