

Prevention of Crime and the Optimal Standard of Proof in Criminal Law

Lando, Henrik

Document Version
Final published version

Publication date:
2003

License
CC BY-NC-ND

Citation for published version (APA):
Lando, H. (2003). *Prevention of Crime and the Optimal Standard of Proof in Criminal Law*.

[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.

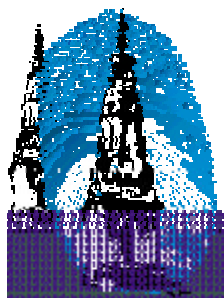
Take down policy

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: 18. Jun. 2025

LEFIC

Copenhagen Business School
Solbjerg Plads 3
DK-2000 Frederiksberg



Center for Law, Economics and Financial Institutions at CBS

LEFIC WORKING PAPER 2003-04

Prevention of Crime and the Optimal Standard of Proof in Criminal Law

Henrik Lando

www.cbs.dk/LEFIC

Prevention of Crime and the Optimal Standard of Proof in Criminal Law

Henrik Lando*

IVS and

Center for Law, Economics and Financial Institutions
at Copenhagen Business School

(LEFIC)

e-mail: hl.ivs@cbs.dk

September 23, 2003

Abstract

The standard of proof in criminal law affects retributive justice through the number of wrong convictions and wrong acquittals. It also affects the level of crime, since a higher standard of proof implies less deterrence and less incapacitation. This article derives an expression for the optimal standard from a trade-off between these effects, and applies the expression to the crime of sexual violation against women. For this crime, social preferences for justice versus prevention of crime are elicited through a survey and inserted into the expression for the optimal standard. The result indicates that the concern for prevention of crime may have a significant effect on the optimal standard of proof.

I wish to thank Steven Levitt for providing empirical estimates of the preventive effect of sanctions, and Clas Wihlborg for extensive comments. Also, I thank participants at the American Law and Economics Association Conference at Harvard Law School, May 2002. Peter Schmidt and Jens Dick-Nielsen provided able research assistance.

Introduction

According to a prevalent view, the standard of proof in criminal law should balance the number of wrong acquittals with the number of wrong convictions so as to minimize the sum of ‘injustice costs’ from these two types of error. This article investigates the consequences of also including the concern for prevention of crime as a determinant of the optimal standard of proof. Naturally, a higher standard will in general increase the level of crime, since it will lead to less deterrence and less incapacitation of criminals. Therefore, including the concern for preventing crime will tend to lower the socially optimal standard of proof. This article inquires into the potential size of this effect. I first derive - within a Becker-type [2] model of crime - a general expression for the optimal standard of proof as it depends on social preferences concerning justice and prevention of crime and on the elasticity of crime with respect to the expected sanction. Then I apply the expression to the particular crime of sexual violation of women. For this crime, the standard of proof is well-known to play an important role for the outcome of legal trials, so in this respect the crime seems well suited for illustrating the potential importance of including the concern for preventing crime. In effect, I demonstrate that for this crime, the optimal standard falls from 92% to 67% when prevention of crime is taken into account. Thus, the analysis suggests that the concern for prevention of crime may have a substantial effect on the standard of proof. Indeed, if it is correct to associate the standard ‘beyond a reasonable doubt’ with 95% (subjective) certainty and ‘clear and convincing evidence’ with 75% certainty, as some do (e.g. Schauer and Zeckhauser [24]), the analysis suggests that for the crime of sexual violation of women including the concern for prevention lowers the standard from ‘beyond a reasonable doubt’ to ‘clear and convincing evidence’.

For the crime of sexual violation, social preferences for justice and prevention will be derived from survey answers to the two following questions: ‘How many guilty people are you willing to let go free to avoid that someone is wrongly convicted of sexual violation?’ and ‘How many sexual violations can you/are you willing to accept in order to avoid that someone is wrongly convicted of sexual violation?’. The optimal standard is found by inserting the survey population’s median responses to these two questions, together with an estimate of the preventive effect of sanctions, into the expression for the optimal standard.

Structure

Before introducing the model, I shall review the law and economics literature on the subject of the standard of proof, and discuss whether the risk of being innocently convicted should be seen as affecting the incentive to commit crime. This issue will be discussed at some length, since it will be assumed in the model that there is no effect on incentives of wrong convictions, which may appear counterintuitive. After the discussion of this point, the model will be presented, and the main result will be derived and applied to the crime of sexual violation against women. A discussion and a conclusion end the paper.

The Literature

The articles by Schrag and Scotchmer [25] and Farmer and Terrell [9] are closely related to this.

Like this paper, Schrag and Scotchmer weigh justice and prevention concerns, but their focus is primarily on which kind of evidence should be allowed in court, rather than on how much evidence should suffice for conviction. In particular, they analyze whether character evidence should be allowed in a criminal trial, and show that allowing character evidence may, under certain circumstances, have a detrimental effect on deterrence of crime, since it may provide criminals who have committed crime in the past too little incentive to abandon their criminal career.

Farmer and Terrell analyze the following choice involving justice and prevention of crime: when the propensity for crime is higher for one (race or gender) group than for another, allowing members of the former group to be convicted on relatively less evidence will lead to more incapacitation at a given cost in terms of unfair conviction, but will violate the principle of equality before the law. There is hence a trade-off between justice and prevention of crime, and they explore the likely significance of this trade-off through simulations of a simple model. According to their simulations, adhering to the principle of equality before the law may be costly in terms of less incapacitation and hence more crime.

Three papers by, respectively, Rubinfeld and Sappington [22], Miceli [14] and Yilankaya [28] are worth mentioning because they address an issue that will not be included in the present analysis but which may be important when setting the standard of proof. The three papers analyze the effect of the standard of proof and of the size of the sanction on both the prosecutor's effort to establish

guilt and the defendant's effort to establish innocence. Thus, Rubinfeld and Sappington investigate how the standard of proof and the size of the sanction affect the defendant's effort to signal innocence, Miceli analyzes how the two variables affect prosecutorial effort, and Yilankaya explores how the two variables affect both the defendant's and the prosecutor's efforts, which are seen as strategically interdependent. In contrast, the present paper abstracts from ('black-boxes') this and other interactions of the standard of proof with the judicial process¹. The question analyzed in this paper is what level of certainty society should optimally strive for, and this can be analyzed without going into the process through which this level of certainty is attained.

Another important article by Miceli [13] should also be mentioned, although it is not mainly concerned with the standard of proof. He analyzes how much effort the legal system should spend to investigate some given crime and how high the sanction for the crime should be, as choices involving trade-offs between justice (fairness) and efficiency.

Andreoni [1] argues that higher sanctions may not lead to less crime, since jurors may adjust the standard of proof to the level of sanctions in order not to risk sanctioning an innocent person harshly. Andreoni hence points to a conflict of interest between jurors and the rest of society; jurors must live with the thought of perhaps having sanctioned an innocent person, and may therefore have an incentive to apply a higher standard of proof than that which is in society's overall interest. This and other agency issues will however not be addressed in the present paper, since the main aim of this paper is to establish the socially optimal standard of proof as a benchmark.

Other important law and economics references on the standard of proof in criminal law are Kaplow [11], Posner [19], Reinganum [21], Kaplow and Shavell [12], and Polinsky and Shavell [17].

Does the risk of unfair conviction affect the incentive to commit crime?

In the model below, it will be assumed that the incentive to commit crime

¹Most importantly, how the standard of proof interacts with the outcome of plea bargaining will not be analyzed in this paper. For an analysis of plea bargaining, see e.g. Reinganum [21]. Note that the vast majority of felony convictions in the US (approx. 95%) are through plea bargaining, see Bureau of Justice Statistics Bulletin, *Felony Convictions in State Courts, 2000*: <http://www.ojp.usdoj.gov/bjs/pub/pdf/fssc00.pdf>. In contrast, plea bargaining does not exist as such in many European countries, including Denmark.

is unaffected by the risk of innocent conviction. This goes against the intuition that the risk of being wrongly convicted must lower the reward to not committing crime, and that it must hence act as an inducement to commit crime. However, this logic does not apply in all circumstances. To take a concrete example, if Smith decides not to steal Jones's car, Smith need generally not fear being falsely convicted of having stolen Jones's car for the simple reason that Jones' car has not been stolen. Naturally, Smith may fear being innocently convicted for another theft or another crime, but this risk exists whether or not he steals Jones' car². Naturally, if Smith does not steal it, he may fear being innocently convicted of stealing Jones' car, if somebody else steals it. This is the case analyzed by Schrag and Scotchmer. However, this seems a somewhat special case. First, criminals do not always compete for a limited set of crime opportunities. There are e.g. many cars to steal and many banks to rob. Second, for each individual the probability of being falsely convicted tends to be very low, since many innocent individuals share the risk³.

The risk of unfair conviction may, however, affect the incentive for crime in other ways than that analyzed by Schrag and Scotchmer. Png [16] cites the example of a motorcyclist, who even though he or she drives with due care, may risk being falsely accused of negligent driving in the case of an accident. This, Png argues, may reduce his or her incentive for exercising due care, by reducing the expected pay-off from doing so⁴. Another and perhaps better example is given by R. Posner [19], who points out that if the sanction for speeding is applied at random, it provides no incentive not to speed.

The question then arises what fundamental feature determines when there is an effect on incentives. I submit that the important point lies in defining the criminal act as a *specific act* committed at a given time and place. The question is whether there can be evidence indicating that a person has committed the specific act, when he or she has not. When Jones decides not to steal Jones' car, Smith need not fear being falsely convicted of it, while when Smith decides not to speed

²Of course, if the sanction for one theft is felt less when applied on top of a sanction for another, the risk of unfair conviction may affect the incentive for theft. For several reasons, e.g. that the sanction is usually increasing in the number of offenses, this effect will not be included in the analysis.

³Schrag and Scotchmer must have in mind that if Smith does not steal Jones' car, Smith will be near the place of the car-theft, if somebody else steals it, and will therefore be more likely to be innocently convicted than the average person. This, again, is a somewhat special case, although of course it can occur.

⁴Png ascribes the basic insight to R. Posner in [20].

at a given point in time, he may nevertheless fear being sanctioned for speeding at that point in time. This article pertains to such crimes as theft, assault and murder for which there is almost zero probability of being falsely convicted for that specific act which one chooses not to commit⁵.

The Model

The analysis concerns one particular kind of criminal act. Each such act causes harm, h , to society, and is assumed for simplicity to be undertaken by one person alone. Potential criminals are assumed to be risk neutral, and, following the Becker [2] approach to crime, each will commit the crime if the benefit, v , is greater than the expected sanction⁶. The total population is normalized to one, and the fraction of the population with benefit v is given by the density function $z(v) > 0$, and the cumulative distribution function $Z(v)$, both defined on the interval $[0; \infty[$.

Whenever the criminal act has been committed, somebody will report it to the police, who in turn will undertake an investigation and take the case to court if there is sufficient evidence against the prime suspect⁷. The evidence generated through investigations and court proceedings will be stochastic and may be more or less incriminating. It will be assumed that the judge/juror can rank the evidence in this respect, and that the ranking can be expressed by a continuous variable $y \in [0, \infty[$, where the number is higher the more conclusive the evidence.

The ratio of the probability of the prime suspect's guilt (g), and his or her innocence (i), can be found from the Bayesian formula $P = A \times L$, where P is the ratio:

$$P = \frac{\text{probability of guilt given the evidence}}{\text{probability of innocence given the evidence}}$$

and A is the ratio:

$$\frac{\text{ex-ante probability of guilt}}{\text{ex-ante probability of innocence}}$$

⁵Such cases are usually about the identity of the wrongdoer and not about the action of the wrongdoer, since the existence of a victim is often transparent, and there would not be a victim if a wrongful act had not been performed.

⁶Many people will not even think of committing the crime or will have great disutility from it in terms of loss of esteem from others, etc. This is consistent with the model.

⁷We are interested in such cases where there is a prime suspect, since if this is not the case, a criminal conviction will be out of the question (as will be clear).

where ex-ante refers to the information existing before any evidence is presented. L is the likelihood ratio, which is:

$$\frac{\text{the probability of observing the evidence given guilt}}{\text{the probability of observing the evidence given innocence}}$$

It will be convenient to split the evidence y into two parts; that which is sufficient to generate a probability of guilt equal to $\frac{1}{2}$, which we will term $y_{\frac{1}{2}}$, and the amount of evidence beyond that⁸, which we will term x ⁹. Note that the posterior beliefs can be expressed as $P = A \times L(y_{\frac{1}{2}}) \times L(x | y_{\frac{1}{2}})$, where A is the ex-ante ratio of beliefs before any evidence is presented, $L(y_{\frac{1}{2}})$ is the likelihood ratio for the level of evidence $y_{\frac{1}{2}}$ and $L(x | y_{\frac{1}{2}})$ is the ratio of the conditional probabilities of observing x when y has been observed. This follows from the fact that the probability of observing both $y_{\frac{1}{2}}$ and x equals the probability of observing $y_{\frac{1}{2}}$ multiplied by the probability of observing x given that one has observed $y_{\frac{1}{2}}$. Note here that $A \times L(y_{\frac{1}{2}})$ equals 1 by the definition of $y_{\frac{1}{2}}$; this simplifies expressions.

$L(x | y_{\frac{1}{2}})$ can be expressed as $\frac{f_g(x)}{f_i(x)}$ where the differentiable density function $f_g(x)$ expresses the probability of x forthcoming against the prime suspect, conditional on $y_{\frac{1}{2}}$, if it is assumed that the prime suspect did commit the crime, and $f_i(x)$ is the probability of x forthcoming against the prime suspect, conditional on $y_{\frac{1}{2}}$, if it is assumed that some other person than the prime suspect committed the crime¹⁰. F_g and F_i denote the corresponding distribution functions.

It follows from Bayes' formula that:

$$\frac{\text{prob}(g | x)}{\text{prob}(i | x)} = \frac{\frac{1}{2}}{\frac{1}{2}} * \frac{f_g(x)}{f_i(x)}$$

⁸Note that we are only interested in such cases where P exceeds 1, i.e. where the probability of the prime suspect's guilt is at least equal to $\frac{1}{2}$.

⁹This way of representing the evidence, by x instead of by y both simplifies calculations and creates a link between our intuition and the analysis. To illustrate, when we say that there is more than a preponderance of the evidence, we mean that the evidence would be more likely to come forth if the suspect did commit the crime than if he or she did not. But in a Bayesian sense this is only correct if we consider the prior belief that the suspect is guilty to be $\frac{1}{2}$. So we are really saying that there is more evidence than that which would create a probability of $\frac{1}{2}$, and we are not talking about the likelihood ratio which may be much higher than 1 if the prior probability of the suspect's guilt is very low.

¹⁰For consistency, it will be assumed that a potential offender and the judge/juror agree on the density functions $f_g(x)$ and $f_i(x)$, and there will also be assumed to be agreement on $L(y_{\frac{1}{2}})$.

from which it follows that

$$prob(g \mid x) = \frac{f_g(x)}{f_g(x) + f_i(x)}$$

since $prob(g \mid x) + prob(i \mid x) = 1$.

If the threshold level of x above which there will be conviction is denoted by \tilde{x} , the conditional probabilities of the main events can now be stated in terms of the F_g and the F_i distribution-functions. Thus, the probability that the guilty will be convicted, $prob\{(x \geq \tilde{x} \mid y_{\frac{1}{2}}) \cap g\}$, and the probability of somebody being falsely convicted, $prob\{(x \geq \tilde{x} \mid y_{\frac{1}{2}}) \cap i\}$, are:

$$\begin{aligned} prob\{(x \geq \tilde{x} \mid y_{\frac{1}{2}}) \cap g\} &= \int_{\tilde{x}}^{\infty} f_g(x) dx = 1 - F_g(\tilde{x}) \\ prob\{(x \geq \tilde{x} \mid y_{\frac{1}{2}}) \cap i\} &= 1 - F_i(\tilde{x}) \end{aligned}$$

Let q be the probability that $y \geq y_{\frac{1}{2}}$, i.e. the probability that the evidence against the prime suspect is such that it is more probable than not that the prime suspect is guilty. If the sanction is s and the level of utility for which a person will be exactly deterred is \tilde{v} , then

$$\tilde{v} = s \times q \times \int_{\tilde{x}}^{\infty} f_g(x) dx$$

$Z(\tilde{v})$ is then the law-abiding part of the population and $1 - Z(\tilde{v})$ is the criminal part.

I will incorporate justice concerns into the social welfare function in the same way as Diamond [6], Miceli [13] and Polinsky and Shavell [18]. When a crime has been committed, not sanctioning the guilty comes at a social cost Q in terms of ‘justice disutility’, as compared with the situation where the offender is caught and sanctioned. The social justice cost from not sanctioning the guilty will hence be $(1 - Z(\tilde{v}))((1 - q)Q + qF_g(\tilde{x})Q)$.

The injustice cost incurred by society when a person is innocently convicted will be denoted θ . The expected cost associated with unfair convictions is hence $(1 - Z(\tilde{v}))q(1 - F_i(\tilde{x}))\theta$.

Let c denote the cost to society of applying the sanction, i.e. the cost of incarceration. It will be applied both when the innocent and when the guilty

are sanctioned. The total expected sanctioning costs are then: $(1 - Z(\tilde{v}))q(1 - F_i(\tilde{x}) + 1 - F_g(\tilde{x}))c$.

Social welfare is maximized when total social costs SC are minimized with respect \tilde{x} . SC can be written¹¹:

$$(1 - Z(\tilde{v})) [h + (1 - q)Q + qF_g(\tilde{x})Q + q(1 - F_i(\tilde{x}))\theta + q(1 - F_i(\tilde{x}) + 1 - F_g(\tilde{x}))c]$$

where

$$\tilde{v} = s \times q \times \int_{\tilde{x}}^{\infty} f_g(x)dx$$

The Optimal Standard of Proof

The optimal standard of proof can be found by inserting the expression for \tilde{v} into the social cost function and differentiating with respect to \tilde{x} . It is convenient first to express social costs as $SC = (1 - Z(\tilde{v}))(h + w)$, where $(1 - Z(\tilde{v}))$ is the number of crimes, and where $h + w$ is the total harm per crime, including not only the direct harm, h , but also the injustice and sanctioning cost per crime, w . w is given by:

$$w = (1 - q)Q + F_g(\tilde{x})qQ + (1 - F_i(\tilde{x}))q\theta + qc(1 - F_i(\tilde{x}) + 1 - F_g(\tilde{x}))$$

where $(1 - q)Q + F_g(\tilde{x})qQ$ is the expected injustice cost per crime associated with the offender escaping sanction, $(1 - F_i(\tilde{x}))q\theta$ is the expected injustice cost per crime due to false conviction, and $qc(1 - F_i(\tilde{x}) + 1 - F_g(\tilde{x}))$ is the expected sanctioning cost per crime.

We can simplify the derivative of social cost with respect to the standard of proof by noting that $\frac{\partial \tilde{v}}{\partial \tilde{x}} = -f_g(\tilde{x})sq$, $\frac{\partial Z(\tilde{v})}{\partial \tilde{v}} = z(\tilde{v})$, and that therefore $\frac{\partial (1 - Z(\tilde{v}))}{\partial \tilde{x}} = z(\tilde{v})f_g(\tilde{x})sq$. It follows that :

$$\frac{\partial SC}{\partial \tilde{x}} = (h + w)z(\tilde{v})f_g(\tilde{x})sq + (1 - Z(\tilde{v}))\frac{\partial w}{\partial \tilde{x}}$$

In this expression $\frac{\partial w}{\partial \tilde{x}}$ equals $f_g(\tilde{x})qQ - q\theta f_i(\tilde{x}) - (f_i(\tilde{x}) + f_g(\tilde{x}))qc$. Denote the fraction of criminals deterred by a one-unit increase in \tilde{v} by α , i.e. $\alpha = z(\tilde{v})/(1 - Z(\tilde{v}))$. Then $\frac{\partial SC}{\partial \tilde{x}} = (h + w)z(\tilde{v})f_g(\tilde{x})sq + (1 - Z(\tilde{v}))\frac{\partial w}{\partial \tilde{x}}$, and:

¹¹It is assumed here that the crime is of the kind where the utility derived by the offender by most people would be considered ‘socially illicit’, and so should not be included in the social welfare function. However, whether or not the utility of the offender is included, plays no important role in what follows.

$$\frac{\partial SC}{\partial \tilde{x}} = (h + w)z(\tilde{v})f_g(\tilde{x})sq + (1 - Z(\tilde{v}))\frac{\partial w}{\partial \tilde{x}} = 0 \Rightarrow$$

$$(h + w)z(\tilde{v})f_g(\tilde{x})sq + (1 - Z(\tilde{v})) [f_g(\tilde{x})qQ - q\theta f_i(\tilde{x}) - (f_i(\tilde{x}) + f_g(\tilde{x}))qc] = 0 \Rightarrow$$

$$\frac{f_i(\tilde{x})}{f_g(\tilde{x})} = \frac{[(h + w)z(\tilde{v})sq + (1 - Z(\tilde{v}))(qQ - qc)]}{(1 - Z(\tilde{v}))(q\theta + qc)} = \frac{(h + w)\alpha sq + (qQ - qc)}{(q\theta + qc)} \Rightarrow$$

$$\frac{f_i(\tilde{x}) + f_g(\tilde{x})}{f_g(\tilde{x})} = \frac{(h + w)\alpha sq + (qQ - qc) + (q\theta + qc)}{(q\theta + qc)} \Rightarrow$$

$$\frac{f_g(\tilde{x})}{f_i(\tilde{x}) + f_g(\tilde{x})} = \frac{(q\theta + qc)}{(h + w)\alpha sq + (qQ - qc) + (q\theta + qc)} = \frac{\theta + c}{(h + w)\alpha s + (Q - c) + (\theta + c)}$$

which proves the following proposition:

Proposition: The optimal standard of proof, $prob(g \mid \tilde{x}^*)$, is given by:

$$prob(g \mid \tilde{x}^*) = \frac{f_g(\tilde{x})}{f_i(\tilde{x}) + f_g(\tilde{x})} = \frac{\theta + c}{(h + w)\alpha s + (Q - c) + (\theta + c)}$$

Two comments may be made. First, note that in the absence of a deterrent effect, i.e. if $z(\tilde{v}) = 0$, the result is

$$prob(g \mid \tilde{x}^*) = \frac{\theta + c}{(Q - c) + (\theta + c)}$$

This condition equates the marginal relative number of falsely acquitted and falsely convicted with their relative costs. It should be stressed that the condition is marginal. Thus, if e.g. $prob(g \mid \tilde{x}^*)$ equals 90%, this does not mean that one in every ten verdicts will be an unjust conviction. Rather, it means that whenever the evidence is at the critical level, the defendant is innocent with a probability of 10%. A majority of defendants will be found guilty at a higher level of certainty.

Second, note that the extent to which the optimal standard is affected by the concern for prevention of crime depends on two factors: how harmful the crime is ($h + w$), where harm includes the expected injustice and sanctioning cost per committed crime, and how large a fraction of the criminal population will be deterred by a lowering of the standard (αs).

An Application to the Crime of Sexual Violation Against Women

Trials of alleged sexual violation must often be decided by the court on the basis of evidence that leaves room for doubt.

To insert into the formula:

$$prob(g \mid \tilde{x}^*) = \frac{\theta + c}{(h + w)\alpha s + (Q - c) + (\theta + c)}$$

we need estimates for θ , Q , h , c , w , s , α and q . In the following, the numbers apply to Denmark, but it should be obvious that a similar analysis can be carried out for any other jurisdiction¹².

One variable can be chosen freely, let us set $Q = 1$.

To establish the social welfare function, I conducted a survey among a group of graduate students and faculty at Copenhagen Business School, whom I asked for their preferences concerning the ratios of θ/Q and θ/h .

The exact wording of the survey is in the appendix. Below, figure 1 and figure 2 present the answers to the two questions, respectively.

¹²In the appendix, a similar analysis is conducted for the US.

Figure 1: How many guilty people are you willing to let go free in order to avoid that someone is wrongly convicted of sexual violation against women?

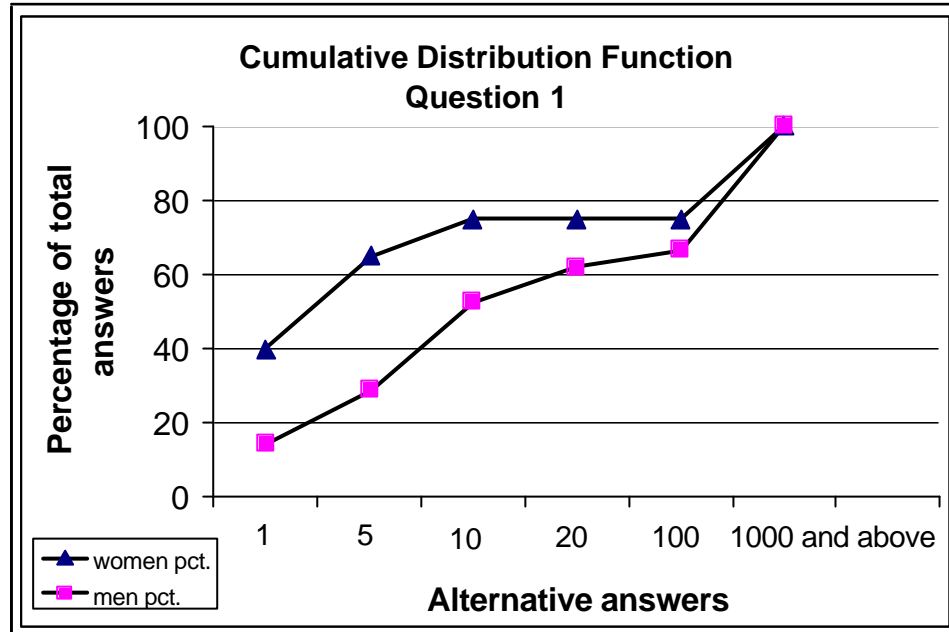
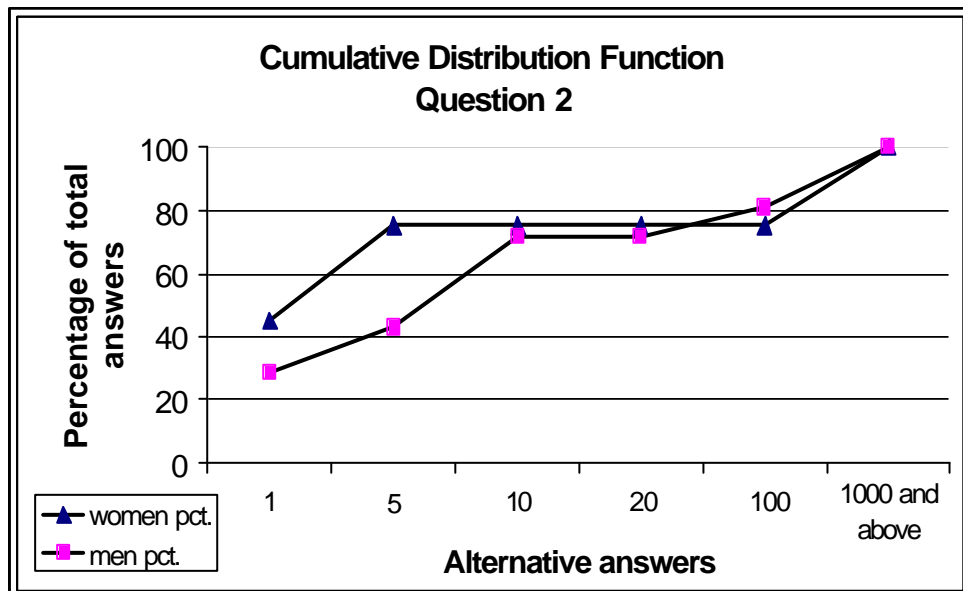


Figure 2: How many sexual violations can you/are you willing to accept in order to avoid that someone is wrongly convicted of sexual violation against women?



Note that women attach a higher cost to false acquittals and to the crime, relative to the cost of a false conviction, than men do, as one might perhaps expect.

Society's preferences will be taken to be equal to the median answer in the total population. In the group of 49 respondents, the median preference for θ/Q was a ratio of 10:1, while the median ratio for θ/h was 5:1¹³.

The respondents were not asked to take into account the cost, c , to society of applying the sanction, and presumably did not do so. The average sanction is a year and a half and the offender is released after serving $2/3$ of the time, so the effective average sanction is one year. The cost, c , has been estimated at 35,000\$ per year and the question is what this amounts to in terms of utility. It may be compared with the value of avoiding one violation which in utility terms has been set at 2 (by the fact that $Q = 1$, $\theta/Q = 10$ and $\theta/h = 5$). The monetary value of avoiding one violation is of course controversial; if we assume it is in the interval between 200,000\$ and 300,000\$¹⁴, then, since the utility 2 equals between 200,000\$ and 300,000\$, the amount of 35,000\$ is approximately in the interval between $1/6$ and $1/9$. Since it will not be very important to our results where in the interval it lies, this will not be pursued any further.

In the expression for w : $w = (1-q)Q + F_g(\tilde{x})qQ + (1-F_i(\tilde{x}))q\theta + qc(1-F_i(\tilde{x}) + 1 - F_g(\tilde{x}))$, the first two terms, $(1-q)Q + F_g(\tilde{x})qQ$, dominate the others. This is due to the fact that only 60 of 1000 criminal acts lead to conviction. Since 940 of 1000 offenders are not sanctioned, the expected injustice cost of wrong acquittals is $\frac{940}{1000} \times Q = .94$. Thus, $w \simeq 1$, so $h + w \simeq 3$.

The variable s is the level of the sanction in terms of disutility. As noted, it is approximately one years' imprisonment¹⁵. The disutility of this sanction to the criminal is difficult to estimate. Given that $h \simeq 2$, let us assume as a very rough approximation that $s = 1/2$, meaning that one year in prison corresponds

¹³It should be mentioned that some respondents expressed the view that nothing could compensate for a wrong conviction; when calculating median preferences, I treated this view the same as the view that the cost is very high (1000 and above). Likewise, some people gave the answer 0, which was treated as the same as a very low answer when calculating the median. Both the answer ∞ and the answer 0 may be interpreted as claiming that costs are incommensurable.

¹⁴Cohen et al [4] estimate society's willingness to pay for one less rape or sexual assault to be 237,000 \$.

¹⁵This level may be lower than optimal which may have some effect on the results concerning the optimal standard of proof. If the sanction were increased, crime would fall, and it would then be more difficult to achieve a further decline in the level of crime. This suggests that lowering the standard of proof and increasing the sanction are (strategic) substitutes, but it should be added that there are other cross-effects. The issue will not be pursued further here.

to one fourth the utility loss of a violation. This utility loss should be understood to include social sanctions following a guilty verdict. Still the number may be controversial; the reader can explore alternative assumptions.

An estimate of α , i.e. of $z(\tilde{v})/(1-Z(\tilde{v}))$, can be obtained as follows: $z(v)$ equals $\frac{dZ(v)}{dv}$, i.e. $z(v^*)$ equals the decrease in the fraction of the population which follows one unit's increase in v^* , i.e. one unit's increase in the expected sanction. If we substitute the actual value of \tilde{x}^* for the optimal, then since $v^* = s \times q \times \int_{\tilde{x}^*}^{\infty} f_g(x)dx$,

and $q \times \int_{\tilde{x}^*}^{\infty} f_g(x)dx$ approximately equals $\frac{60}{1000} = 6\%$, the expected sanction equals around 0.03 since s equals $1/2$. To calculate $\frac{dZ(v)}{dv}$, assume that the expected sanction is increased from 0.03 to 0.04. If it is assumed that the elasticity of crime with respect to the expected sanction is -0.1 ¹⁶, an increase in the sanction of $33\frac{1}{3}\%$ creates a 2.8% decrease in crime, which corresponds to 28 violations per year¹⁷.

Hence, $\frac{dZ(v)}{dv}$ equals¹⁸ $\frac{15}{0.01} = 2800$, and α then equals $2800/1000 = 2.8$.

To calculate q , which is the probability that the evidence is beyond y_2 , i.e. the probability that the evidence points to the prime suspect with a probability at or higher than 50%, note that the total number of reported violations is approximately 500 per year of which approximately 225 end up in court and 60 lead to a prison sentence¹⁹. If we assume that only one in two sexual violations is reported, which is the official estimate, the total number of violations is 1000²⁰. If we can assume that for crimes that end up in court, the probability is higher than .5, clearly q must exceed $\frac{225}{1000}$, since 225 cases do end up in court out of the 1000 yearly crimes. On the other hand, 500 crimes go unreported, so no evidence is forthcoming at all. q cannot therefore exceed $\frac{1}{2}$. So q lies between 22.5% and 50%. Note furthermore that of the 500 reported there must be a number of cases where the offender is not found. On the other hand, of the 275 cases that are

¹⁶This is a low figure. Donohue and Seigelman [7] use the number -0.15 and denote it the consensus estimate.

¹⁷If the elasticity is -0.1 , the percentage decrease in the level of crime as a consequence of increasing the sanction by 33.3%, would be $1 - 1.333^{-0.1} = 2.8\%$.

¹⁸Both the numerator and the denominator should be divided by the size of the population which then cancels out.

¹⁹Source: Statistical Yearbook 2001, Table 217 [27].

²⁰An organisation protecting the rights of victims of sexual violation, the Joan-Sisters, estimates that the number is much higher, between 5000 and 10000 whereas the official estimate is 1000. The official estimate is closer to the number that victimization studies find in the US; the estimated percentage of reported rapes was 53,7% in 2002, according to Bureau of Justice Statistics, see p. 11 in <http://www.rainn.org/ncvs2002.pdf>.

reported but not taken to court, there must be some for which the probability of guilt exceeds $\frac{1}{2}$. Based on these considerations q will be taken to be approximately .35.

Given these assumptions, and assuming that the cost of sanctioning is $1/6$, the optimal standard of proof becomes:

$$p_g = \frac{\theta + c}{(h + w)\alpha s + (Q - c) + (\theta + c)} = \frac{10 + \frac{1}{6}}{(3 \times 2.8 \times 0.5) + (1 - \frac{1}{6}) + (10 + \frac{1}{6})} = .67$$

In comparison, the optimal standard becomes:

$$p_g = \frac{10 + \frac{1}{6}}{10 + \frac{1}{6} + 1 - \frac{1}{6}} = 92\%$$

in the absence of a preventive effect, if the cost of sanctioning is $1/6$.

To see what is behind the result that .67 is the optimal standard when the preventive effect is included, consider the consequences of lowering the standard of proof from 75% to 65%. If, say, 100 offenders are convicted at 75% certainty²¹, and 10 more defendants are convicted when the standard is lowered (this number is irrelevant since all effects are scaled up or down) to 65%, of these approximately 7 will be guilty while 3 will be innocent, assuming that defendants' probabilities of guilt are uniformly distributed on the interval from 75% to 65%. The added cost in terms of unfair convictions is then $10\frac{1}{6} \times 3$, and the added benefit in terms of corrections of wrongful acquittals is $7 * (1 - \frac{1}{6})$. From a justice viewpoint, costs exceed benefits by approximately 24-25 utils. However, the increase in the expected sanction will be approximately $10/100 = 10\%$, so violations will decrease by $1 - 1.1^{-0.1} = .9\%$ which amount to 9 fewer violations. This adds 27 in utility ($h + w = 3$). Given our assumptions, marginal benefits are higher than marginal costs when prevention is taken into account; this reflects that the standard is too high at 75%.

Applying the proposition of this paper to the case of sexual violation hence leads to the conclusion mentioned in the introduction: If 'beyond a reasonable standard' can be taken to mean approximately 95%, and 'clear and convincing evidence' to mean approximately 75% certainty, as has been suggested by Schauer and Zeckhauser [24] and others, the optimal standard is closer to 'clear and convincing evidence' than to 'beyond a reasonable standard' when the effect of the standard of proof on prevention of crime is taken into account.

²¹compared with 60 at the existing, higher standard. The number 100 is chosen for illustration.

Discussion

This section discusses some general issues and some objections that may be raised against the theory and the application.

A basic issue is the choice of social welfare function. It was assumed in this paper that when people express a strong preference for not sanctioning the innocent, this may be taken at face value as expressing a strong preference for avoiding an unjust outcome. Against this view, it has been argued that people are more concerned with the risk that they themselves will be wrongly convicted than with justice for others, and that the analysis can therefore be framed in terms of efficiency in a narrow sense. However, it may be replied that the view that people are only concerned about their own risk of false conviction is not confirmed by the survey. Both men and women report strong preferences for not sanctioning the innocent (although women are in general somewhat less concerned than men, as shown in figure 1), and women's concern is hard to reconcile with the view that they should be simply risk averse, since women need not fear being falsely convicted for sexual violation.

The theory presented in this paper can also be criticized on the basis that it is morally wrong for society to (purposely) risk convicting an innocent person in order to avoid crime, i.e. in order to avoid that *some other person* commits a wrong²². However, stated in this way the criticism is too strong, for if it is generally wrong to risk convicting an innocent person, then the very existence of a legal system must also be wrong, for we know that any legal system will inevitably make mistakes²³.

In a similar vein, it may be held that for ethical reasons, society should rather change other policy variables, e.g. increase the level of police effort, the size of sanctions, or establish social programs that help potential offenders, rather than to compromise on the fundamental value of 'legal security'. However, to this criticism it can be replied that crime will persist also when all of these other variables are set at their optimal level, and that the expression for the optimal standard derived in this paper can be understood as taking these variables as

²²This moral dilemma is mentioned also by Miceli [13] (p. 6) who gives references to discussions by Nozick and Elster.

²³So the point needs to be put in terms of degree: it can more reasonably be argued that there is a limit to the *degree* to which society can in a utilitarian fashion sacrifice the interests of possibly innocent defendants in order to benefit other members of society. This cannot be entered further into here.

given *at their optimal level*.

Concerning the application to sexual violation, it may be objected that this crime does not fall into the category for which the risk of unfair conviction does not affect the incentive for committing the crime, and that the crime therefore cannot be used as an application of the model. Arguably, a man who has chosen not to commit a particular act of sexual violation may nevertheless be falsely accused of it. However, there is reason to be skeptical of the idea that the risk of unfair conviction will significantly affect incentives for committing sexual violation, since either the woman consents or she does not; if she does not, the possibility of having voluntary intercourse is not present, and if the man then abstains, the risk of being accused of sexual violation afterwards is very small, since there has been no intercourse. Thus, it seems likely that false accusations do not increase incentives for committing sexual violation to any significant extent²⁴.

The question may be raised what quantitative expressions for the standard of proof really mean. For example, what does it mean in terms of jury-instructions to say that the standard is 67%? This issue cannot be fully dealt with here, except by noting that it should not be taken for granted that standards of proof cannot be expressed quantitatively in terms of probabilities of guilt. Thus, in an overview of the literature on how jurors make decisions, Michael Saks [23] (section 6) notes that: ... ‘several studies have found that the applicable standard of proof is best comprehended by jurors when it is presented to them in the one way that the courts refuse to use: by way of quantitative metaphors’.

However, while quantitative expressions or metaphors may be better grasped by jurors than verbal statements, one may fear that if a jury is told to apply a given quantified standard, it will do so with error, and that this will increase the likelihood of innocent conviction. How exactly the potential for error affects that standard which one would want the jury and the judge to follow is a question that will be left to future research.

Finally, it may be objected that the survey is unlikely to yield reliable information concerning people’s true preferences for justice versus efficiency. It is well-known that answers may depend on the way questions are phrased. Whether such is the case here remains to be seen. It may also be objected that several respondents seem not to share the utilitarian logic of this paper. In the first question, 15 of the 49 answers and in the second, 13 of the 49 answers can be interpreted as denying that logic (including both the answer infinite and the answer

²⁴Although it may be argued, theoretically, that the incentive to *achieve consent* will be lower when consent does not preclude conviction.

zero). On the other hand, a strong majority, more than 2/3 of the total group of respondents, were in fact willing to quantify their preferences.

Conclusion

This paper has explored the consequences of viewing the standard of proof as determined by justice concerns as well as by the concern for prevention of crime. It has demonstrated that the optimal standard can be derived from two properties of the social welfare function together with an estimate of the elasticity of the number of crimes with respect to the expected sanction. The two properties are: the marginal rate of substitution between a false conviction and a false acquittal, and the marginal rate of substitution between a false conviction and a criminal act. Thus, for any crime, given an estimate of the size of the preventive effect, the optimal standard can be derived from the answer to two questions: how many offenders is society willing to let go to avoid sanctioning one person who is innocent?, and how many criminal acts is society willing to accept to avoid that one person is innocently convicted? In the paper, the social welfare function was derived from median responses to the two questions among a group of 50 people (mainly graduate students), for the crime of sexual violation. For this crime, the aim of preventing crime was shown to have a significant effect on the socially desirable standard of proof.

Let me finally add that the spirit of this article has not been to suggest that wrong convictions should be taken lightly; as revealed by the answers to the survey, social ‘costs’ of wrongful convictions are high. Rather, the idea was that while the wrongful convictions are ‘costly’, so are criminal offenses to the victims, and that it may therefore be misleading to only consider injustice costs when setting the standard of proof.

Appendix 1

The Survey

Dear Colleague/Student,

I am writing to request your participation in a short questionnaire. All participants will remain anonymous. The answers will be used in an article I am writing on the burden of proof in criminal law. The issue to be looked at in the article deals with the degree of certainty required before a man is convicted of sexual violation.

In the first question, please ignore the fact that the punishment can act as a deterrent. Instead, focus exclusively on your sense of justice. There are, however, two things to consider in contrast to one another. On the one hand, with regard to justice, it is important to avoid convicting someone who is innocent, while on the other hand it can seem unjust (also for the victim) if someone who is guilty is not convicted. If high demands are made on the strength of the evidence, fewer innocent people will be convicted while more guilty people will go free. Fewer innocent people will be convicted while more guilty people will go free, the stronger the evidence required.

Question 1) How many guilty people are you willing to let go free in order to avoid that someone is wrongly convicted of sexual violation against women?

- a) 1
- b) 5
- c) 10
- d) 20
- e) 100
- f) 1000
- g) an infinite number, impossible to compensate
- h) other, please specify the amount

In the second question, please take into consideration that the punishment can act as a deterrent. Please assume that fewer innocent people will be convicted, but that more people will be in danger of being violated because the deterrent affect will be small/will not be as great. There is no logical connection between this and the above question, which means that one can give any answer whatsoever without contradicting ones self.

Question 2) How many sexual violations can you/are you willing to accept in order to avoid that someone is wrongly convicted of sexual violation against women?

- a) 1
- b) 5
- c) 10
- d) 20
- e) 100
- f) 1000
- g) an infinite number, impossible to compensate
- h) other, please specify the amount.

Please indicate whether you are male or female.

Appendix 2

The optimal standard in the US

The following calculations of the optimal standard of proof in the US for the crime of sexual violation should be treated with caution. For one thing, the estimates of the the injustice costs are derived from a group of Danish students and faculty and the difference in the length of the prison sentence in Denmark and the US may be due to different preferences for justice.

In the US, the total number of sexual violations is approximately 147.000 per year²⁵ of which approximately 90186 are reported to the police and filed by the police, 22964 lead to arrests and 10600 are convicted²⁶. The mean expected time served in prison is 88 months²⁷; let us assume that the utility loss is at least three times larger than that in Denmark (where the sentence served is more than five times shorter, and prison conditions are perhaps not as rough). Thus, assume that $s = 3$. If the cost to society of sanctioning was between 1/3 and 2/3 in the Danish case, and the prison sentence served is nearly five times longer in the US, we can assume that the cost of the sanction to society is something in the order of 7/3 in utility terms. Assume further that $(h + w) \simeq 3$, as in Denmark. α can be calculated as follows: $z(v)$ equals $\frac{dZ(v)}{dv}$, i.e. $z(v^*)$ equals the decrease in the fraction of the population which follows one unit's increase in v^* , i.e. one unit's increase in the expected sanction. If we substitute the actual value of \tilde{x}^* for the optimal, then since $v^* = s \times q \times \int_{\tilde{x}^*}^{\infty} f_g(x)dx$, and $q \times \int_{\tilde{x}^*}^{\infty} f_g(x)dx$ approximately equals $\frac{10600}{147160} = 7.2\%$, the expected sanction equals $\frac{10600}{147160} \times 2 = 14.4\%$. If the sanction is increased by 1 percentage point to 15.4 %, i.e. increased by $\frac{0.01}{.144} = 7\%$, and if the elasticity is -0.1 , the percentage decrease in the level of crime as a consequence of increasing the sanction by 7%, would be $1 - 1.06944^{-0.1} = 0.67\%$. That would amount to $147160 \times .67\% = 986$ fewer violations. Hence, $\frac{dZ(v)}{dv}$ equals $\frac{986}{0.01} = 98600$ and α then equals $98600/147160 = .67$.

²⁵According to the Bureau of Justice's Crime Victimization Survey for 2000, see <http://www.ojp.usdoj.gov/bjs/pub/pdf/cvus0002.pdf>, table 26.

²⁶These numbers are for the year 2000, see Bureau of Justice Statistics Bulletin, Felony Convictions in State Courts, 2000: <http://www.ojp.usdoj.gov/bjs/pub/pdf/fssc00.pdf>. By far the largest part of convictions are in State courts.

²⁷See p.5 in Bureau of Justice Statistics Bulletin, Felony Convictions in State Courts, 2000: <http://www.ojp.usdoj.gov/bjs/pub/pdf/fssc00.pdf>

Given all of this, the optimal standard of proof becomes: $p_g = \frac{\theta+c}{\theta+c+\alpha s(h+w)+Q-c} = \frac{10+\frac{7}{3}}{10+\frac{7}{3}+0.67*2*3+1-\frac{7}{3}} = 82\%$

It may be worth emphasizing that this calculation should be considered to be preliminary. The purpose here was to do a rough calculation.

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, pp 385-95.
- [2] Becker, Gary: ‘Crime and Punishment: An Economic Approach’, Journal of Political Economy; 1968, 76 (2), pp 169-217.
- [3] Bentham, Jeremy: ‘Theory of Legislation’, London, 1876.
- [4] Cohen, M. et al: ‘Willingness-To-Pay for Crime Control Programs’, Working Pape, September 2003, Vanderbilt University, Nashville, TN.
- [5] Davis, M: ‘The Value of the Truth and the Optimal Standard of Proof in Legal Disputes’, Journal of Law, Economics and Organization; October 1994, vol. 10, no. 2.
- [6] Diamond, Peter: ‘Integrating Punishment and Efficiency Concerns in Punitive Damages for Reckless Disregard of Risks to Others’, Journal of Law, Economics and Organization; vol. 18, no. 1, 2002.
- [7] Donohue, J and Seigelman, P: ‘Allocating Resources Among Prisons and Social Programs in the Battle Against Crime’, 27 J. Legal Studies 1; January, 1998.
- [8] The Economist: ‘United States: A Pandora’s box; DNA evidence’, Dec 14, 2002. vol. 365, Iss. 8303; p 47
- [9] Farmer, Amy, Terrell, Dek: ‘Crime versus Justice: Is There a Trade-Off?’, 44 Journal of Law and Economics; October 2001.
- [10] Grossman G. and Katz M: ‘Plea Bargaining and Social Welfare’, American Economic Review; 73, 749-57.

- [11] Kaplow, Louis: 'The Value of Accuracy in Adjudication: An Economic Analysis', *Journal of Legal Studies*; 23(1), Part 2, 1994, pp 307-401.
- [12] Kaplow, Louis and Shavell, Steven: 'Accuracy in the Determination of Liability', *Journal of Law and Economics*; 37(1), April 1994, pp 1-15.
- [13] Miceli, Thomas J: 'Optimal Criminal Procedure: Fairness and Deterrence', *International Review of Law and Economics*; 11(1), May 1991, pp 3-10.
- [14] Miceli, Thomas: 'Optimal Prosecution of Defendants Whose Guilt is Uncertain', *Journal of Law, Economics and Organization*; 6(1), Spring 1990, pp 189-202.
- [15] New York Times, editorial, September 1, 2003: www.nytimes.com/2003/09/01/opinion/01MON5.html.
- [16] 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.
- [17] Polinsky, A. Mitchell and Shavell, Steven: 'Legal Error, Litigation, and the Incentive to Obey the Law', *Journal of Law, Economics and Organization*; 5(1), Spring 1989, pp 99-108.
- [18] Polinsky, A. Mitchell and Shavell, Steven: 'The Fairness of Sanctions: Some Implications for Optimal Enforcement Policy', *American Law and Economics Review*; 2, 2000, pp 223-37.
- [19] Posner, R: 'An Economic Approach to the Law of Evidence', 51 *Stanford Law Review* 1477, July 1999, and John M. Olin *Law and Economics working Papers* No. 66, 2000, downloadable from www.law.uchicago.edu/Publications/Working/index.html.
- [20] Posner, R: 'An Economic Approach to Legal Procedure and Judicial Administration', *Journal of Legal Studies*; vol. 39, no. 2, 1973.
- [21] Reinganum, J.: 'Plea Bargaining and Prosecutorial Discretion', *American Economic Review*; 78, pp 713-28.
- [22] Rubinfeld, D, and Sappington D: 'Efficient Awards and Standards of Proof in Judicial Proceedings', *Rand Journal of Economics*; vol. 18, no. 2, 1987, pp 308-315.

- [23] Saks, M: ‘What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions’, 6 Southern California Interdisciplinary Law Journal, vol. 1, Fall 1997.
- [24] Schauer F. and Zeckhauser R.: ‘On the Degree of Confidence for Adverse Decisions’, Journal of Legal Studies; vol. 25 (January 1996).
- [25] Schrag, J. and Scotchmer S.: ‘Crime and Prejudice: The Use of Character Evidence in Criminal Trials’, Journal of Law, Economics and Organization; October 1994, vol. 10, no. 2.
- [26] Schrag, J. and Scotchmer S.: ‘The Self-Reinforcing Nature of Crime’, International Review of Law and Economics; September 1997, vol. 17, iss. 3, pp 325-35.
- [27] : Statistisk Årbog, 2001. Danmarks Statistik.
- [28] Yilankaya, Okan: ‘A Model of Evidence Production and Optimal Standard of Proof and Penalty in Criminal Trials’, Canadian Journal of Economics; May 2002, vol. 35, iss. 2, pp 385-409.