MEMORANDUM

No 32/2001

Should the standard of evidence be lowered to reduce crime?

By
Tone Ognedal

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the **University of Oslo** **Department of Economics** P. O.Box 1095 Blindern N-0317 OSLO Norway Telephone: + 47 22855127 Fax: + 47 22855035 Internet: http://www.oekonomi.uio.no/ e-mail: econdep@econ.uio.no	In co-operation with **The Frisch Centre for Economic Research** Gaustadalleén 21 N-0371 OSLO Norway Telephone: +47 22 95 88 20 Fax: +47 22 95 88 25 Internet: http://www.frisch.uio.no/ e-mail: frisch@frisch.uio.no

List of the latest 10 Memoranda:

No 31	Jon Strand: Public- and private-good values of statistical lives: Results from a combined choice-experiment and contingent-valuation survey. 55 p.
No 30	Ove Wolfgang: Cost-effective Abatement of Ground-level Ozone in Cities and for Larger Regions: Implications of Non-monotonicity. 28 pp.
No 29	Svenn-Erik Mamelund: The Spanish Influenza among Norwegian ethnic minorities 1918-1919. 35 pp.
No 28	Erik Biørn: The Efficiency of Panel Data Estimators: GLS Versus Estimators Which Do Not Depend on Variance Components. 15 pp.
No 27	Qaisar Farooq Akram and Ragnar Nymoen: Employment behaviour in slack and tight labour markets. 34 pp.
No 26	Snorre Kverndokk, Knut Einar Rosendahl and Thomas F. Rutherford: Climate policies and induced technological change: Which to choose the carrot or the stick? 25 pp.
No 25	Olav Bjerkholt: Tracing Haavelmo’s steps from Confluence Analysis to the Probability Approach. 33 pp.
No 24	Sverre A.C. Kittelsen, Guri Galtung Kjæserud and Odd Jarle Kvamme: Errors in Survey Based Quality Evaluation Variables in Efficiency Models of Primary Care Physicians. 23 pp.
No 23	Ove Wolfgang: Eco-Correlation in Acidification Scenarios. 52 pp.
No 22	Øystein Kravdal: The High Fertility of College EducatWomen in Norway: An Artefact of the ’Piecemeal Approach’. 14 pp.

A complete list of this memo-series is available in a PDF® format at:
http://www.oekonomi.uio.no/memo/

Should the standard of evidence be lowered to reduce crime?

by

Tone Ognedal (*)

Department of Economics

University of Oslo

JEL-classification G34

Keywords: standard of evidence, crime, economic crime

Abstract

As crime becomes increasingly widespread it may be optimal not to lower but to increase the standard of evidence. Even though a higher standard of evidence results in a lower expected penalty for all levels of crime, it increases the expected penalty for high levels of crime relative to the expected penalty for lower levels of crime. Consequently, a high standard of evidence is more effective in deterring high levels of crime. As the proportion of agents with a low opportunity cost of committing crime increases, the standard of evidence should be increased.

(*) Tone Ognedal, Department of Economics, Box 1095, 0317 Oslo, Norway. Tel.: + 47 22 85 51 31; fax: + 47 22 85 50 35

E-mail: tone.ognedal@econ.uio.no

1. Introduction

Economic crime is often difficult to prove because the main evidence is circumstantial, and might not meet the criminal standard of proof. For instance, incorrect accounting in a firm might suggest tax evasion or a misreading of the accounting law, accidental spills can be caused by neglect of safety or by bad luck. Often the authorities do not even try a case in court because the evidence is obviously too weak to result in a conviction. Consequently, the probability of conviction for economic crime is low.

In theory, a low probability of conviction could be compensated for by the imposition of a high penalty. In practice, however, the penalty is limited by the legal principle that "the penalty should fit the crime", or by limited liability when the defendant is a firm. Consequently, to maintain a suffiently high expected penalty, many countries resort to disguised reductions in the standard of proof, and thereby in the standard of evidence[(1)] . One reform that lowers the standard of evidence is the introduction of further civil sanctions in addition to criminal ones. While guilt must be proven "beyond reasonable doubt" in a criminal case, "preponderance of evidence" is sufficient in a civil case. Hence, if civil sanctions can be used when the evidence is too weak for criminal conviction, the standard of evidence is effectively lower.[(2)]

The idea behind these reforms is that a lower standard of evidence increases the probability of conviction and thereby the expected penalty. In turn this leads to lower crime rates. However, this reasoning may be too simplistic. A reduced standard of evidence raises the expected penalty for minor violations relative to both obeying the law and committing more serious violations. While the marginal cost of minor violations goes up, the marginal cost of serious violations goes down. Therefore, when agents can choose the magnitude of the violation, the net effect on overall crime is ambiguous. When the standard of evidence is reduced, agents contemplating minor violations have stronger incentives to obey the law. However, agents contemplating serious crimes have less incentive to be moderate. Paradoxically, therefore, a reduced standard of evidence may increase the cost of crime, because it reduces the deterrence for serious violations of the law. Even though the number of violations may go down, the total cost of crime may well increase because the violations become more serious. Moreover, since a reduced standard of evidence increases the probability of wrongful convictions, it may increase costly over-compliance by law-abiding agents who will try to reduce the chance of unfavorable evidence.

The effects of policy changes on the magnitude of crime are not captured in many studies because each agent's decision is modeled as a discrete choice between crime and obedience to the law. One example is the seminal paper by Schrag and Scotchmer (1994), where the agents face exogenous opportunities for crimes of a given magnitude. When the magnitude of the crime is exogenous to each agent, changes in the number of crimes can be used to measure the effect of policy changes. However, if the agents can choose the magnitude of the crime, as in the model developed in this paper, the number of crimes can be a misleading measure. The effect of an incremental increase in the magnitude of an already serious violation may do much more harm than the effect from an agent committing a minor violation instead of obeying the law. However, while the latter scenario increases the number of crimes, the first does not. Hence, a policy that aims at minimizing the number of crimes may not be optimal, since it may imply that the standard of evidence is set too low to discourage violators from choosing serious crimes instead of more moderate ones.

In models where crimes are determined exogenously, a lower standard of evidence reduces crime if the expected penalty increases for crimes relative to obeying the law. When agents choose the magnitude of the crime, however, there is no such simple relationship between the expected penalty and the crime. In my model, agents decide the magnitude of the crime by comparing the marginal gain from the crime to the increase in the expected penalty for different levels of crime. Hence, it is not only the expected penalty for crimes compared to obeying the law that matters, but also the differences in the expected penalty between different magnitudes of crime. Even though a lower standard of evidence increases the expected penalty for all crimes relative to obeying the law, it may increase crime because the marginal costs of the most serious crimes become lower.

The purpose of this paper is to analyze how the standard of evidence should be determined when the aim is to minimize the cost of crime, under conditions where penalties are limited. I do not discuss the effect of changes in the standard of evidence upon the cost of legal errors. Since the court's decisions are based on uncertain evidence, such errors will occur. For a given crime rate, a lower standard of evidence will increase the number of wrongful convictions (type I errors) and reduce the number of wrongful acquittals (type II errors). Hence, there is a trade-off between the two types of error for a given crime rate. The arguments in favor of reduced standards of evidence seem to assume that there is also a trade-off between wrongful convictions and the cost of crime, when the level of crime may vary with the standard of evidence. Proponents of a lower standard of evidence argue that the price of a reduction in crime is a greater number of wrongful convictions. This paper demonstrates that there may not be a trade-off between wrongful convictions and crime. On the contrary, they may move in tandem, which means a lower standard of evidence leads to both higher costs of crime and more wrongful convictions.

The framework developed below applies to many types of crime where evidence is uncertain, although the characteristics may differ in many other respects. For example, the framework covers both "intentional crimes" and "neglectful crimes". For intentional crimes, such as fraud and insider trading, the harmful outcome is intended and the violator gains from the loss he or she inflicts upon others. With "neglectful crimes", such as failure to comply with safety standards, the violator does not gain from the harmful outcome per se, but saves costs from his or her low effort to abide by the rules. Although the harmful outcome is not intended, the neglect of the regulations is intended. Consequently, the problem of efficient deterrence when evidence is uncertain applies to both types of crime.

Section 2 presents the model and derives the crime-minimizing standard of evidence. It analyzes how this standard varies with the penalty level and with a firm's opportunity cost of crime and the moral cost of crime. When the opportunity costs or the moral costs of crime differ between firms, ideally the standard of evidence should also differ between them. However, equality before the law dictates that there must be a common standard for all. Section 3 derives the optimal common standard of evidence when the cost of crime between firms differs, and examines how this standard should change as the composition of those committing crimes changes. Section 4 discusses the case where some firms over-comply, i.e. use too many resources to avoid wrongful convictions. Section 5 concludes the paper.

2. A model of optimal penalty design

The model focuses on the example of environmental crime. Firms in an industry may accidentally emit substances that are environmentally harmful. Monitoring and other costly safety procedures reduce the probability of high levels of harm and increase the probability of small levels of harm. However, a firm with excellent safety procedures may experience a severe accident, while a firm with no safety procedures may avoid accidents. While the level of harm caused by a firm is observable and verifiable by the authorities, the firm’s effort to reduce harm is not.[(3)] Hence, the punishment must be based on harm rather than effort.

Let h be the accidental harm and e the firm's effort to reduce harm: h can take any value between 0 and H, and e can take any value between 0 and E. F(h, e) is the firm's perceived cumulative probability distribution function for h, given effort level e. The corresponding density function is f(h,e). The distribution of h satisfies the following four assumptions:

(a1) The court cannot exclude any level of effort from their observation of h, that is, the support of f (h, e) is [0, H] for all e.

(a2) Effort reduces the chance of high levels of harm, but at a decreasing rate. Formally, $F_e \geq 0$ for all h, with strict inequality for some h and $F_{ee} \leq 0$ for all h, with strict inequality for some h.

(a3) Higher levels of h signify that effort is lower. Formally, f_e / f is strictly decreasing in h, which is equivalent to $f_{eh} - f_e f_h < 0$, sometimes called the monotone likelihood ratio condition.

(a4) For harm levels that become less or equally likely as effort increases, the rate of change in the likelihood of harm as effort increases is negative. Formally, $f_{ee}(h, e) < 0$ for h and e such that $f_e(h,e) \leq 0$. The assumption is illustrated in Figure 1. As effort is increased, high levels of harm become less likely and low levels become more likely. Assumptions (a3) and (a4) together ensure that as effort increases, the harm level that is equally likely for higher values of e (h* in Figure 1) is lower the higher the value of e.

In addition, the following assumptions are made about the cost of effort and the social cost of harm:

(a5) The social cost of harm d(h) is increasing in h and does so at an increasing rate, i.e., $d'(h) > 0$ and $d''(h) > 0$. Since the firms' effort affects the probability of different levels of harm, the expected social cost of harm is a function of the effort level e, as represented below:

$$D(e) = \int_0^H d(h) f(h,e) dh$$

Here, D(e) is a decreasing, convex function of effort, since $d'(h) > 0$ and $F_{ee} \leq 0$ [(4)], i.e., $D'(e) < 0$ and $D''(e) > 0$.

(a6) The firm's cost of reducing harm c(e) is an increasing, convex function of effort, i.e., $c'(e) > 0$ and $c''(e) > 0$.

Assumptions (a5) and (a6) together imply that total social costs c(e) + D(e) are convex in e. If there is an inner solution to the cost minimizing problem, the cost minimizing effort e^S is therefore given by the first order condition:

$$c'(e^S) + D'(e^S) = 0 \tag{1}$$

If there were no restrictions on the penalty, and it was set to equal the social costs, the firm could be induced to choose the optimal effort e^S. However, when the penalties are restricted, it may be not be feasible to induce the socially optimal effort level. How the penalty function should be designed in this case is examined below.

Firms are assumed to be risk neutral, and maximizes net expected income. Let P(h) be the penalty function. The expected penalty of the firm is then $\int_0^H P(h) f(h,e) dh$. If the penalty is a non-decreasing function of harm, i.e., if $P'(h) \geq 0$, it follows from assumptions (a1) and (a2) that the expected penalty is a decreasing, convex function of e.[(5)] The penalty may be limited by the firm's ability to pay or by a legal maximum penalty. Here, both limits are represented by a maximum penalty P. This paper analyzes the case where the penalty is a fine with negligible administrative costs. Since the penalty is then simply a transfer from the firm to the government, social welfare is not affected.[(6)]

The firm may internalize part of the expected social costs directly. There may be several reasons for such internalization. For example, owners may take a genuine interest in the environment, they may be guided by social or moral norms, or they may wish to avoid negative publicity. Let m be the share of the expected social costs internalized by the firm, hereafter called moral concern, where $0 \leq m < 1$. The firm's total (expected) cost is then:

$$T(e) = c(e) + \int_0^H P(h) f(h,e) dh + mD(e) \tag{2}$$

For each feasible penalty function, P(h), which the authorities might choose, there is an optimal choice of effort by the firm. The firm chooses the effort that minimizes total costs T(e). Since all three cost terms in (2) are convex in e, T(e) is convex. Hence, the firm's choice of effort is a unique solution to the firm's optimization problem given by the first order condition:

$$c'(e) + mD'(e) = -\int_0^H P(h) f_e(h,e) dh \tag{3}$$

The condition simply states that the marginal cost of effort plus the reduction in the internalized social cost should equal the marginal gain from reduced expected penalties. The interesting case is when the maximum penalty P is too low to induce the socially optimal effort level e^S. The aim of the enforcement policy is then to induce the highest possible effort. The authorities should therefore pick the penalty function that induces maximum effort, when effort is given by equation (3), as summarized in the following proposition.

Proposition 1

To maximize the firm's effort, the penalty should be zero for observed harm levels below a threshold h and equal to its maximal level P for observed harm levels above the threshold. The threshold h* is a decreasing function of the moral concern m and of the maximum penalty P.*

Since firms are not punished if a harm level below h* is observed, h* is the standard of evidence that induces maximum effort.

Proposition 1 implies that when the crime level goes up because firms' moral concern (m) go down, it may be optimal to raise the standard of evidence even though this would lower the expected penalty for all levels of effort. The reason for this paradoxical result is that even though the expected penalty goes down for all levels of effort when the standard of evidence is raised, the expected penalty for serious violations of the law increases relative to more moderate violations. Hence, the marginal penalty for committing a serious violation of the law has increased. As long as firms are relatively law abiding (m is high), a low standard of evidence is optimal, because it creates a high marginal penalty even for moderate violations, thus discouraging firms from anything more than minor violations. When firms become less law abiding (m becomes lower) and therefore choose to commit more serious violations, the low standard of evidence is no longer optimal. The reason is that when a firm is punished even for low levels of harm, there is no extra penalty for causing high levels.(7) Consequently, the increase in the expected penalty for committing a serious violation instead of a more moderate one is small. By raising the standard of evidence, the marginal penalty is raised for serious violations.

Proposition 1 is proved in two steps. First, I explain why the effort maximizing penalty function jumps from 0 to P at the threshold h^*. The formal proof is given in Appendix A. Next, I derive the threshold h^* as a function of the parameters P and m, and prove the second part of the proposition.

The right-hand side of (3) is the marginal reduction in the expected penalty from higher effort, hereafter called the marginal penalty. The firm chooses higher effort the higher the marginal penalty. Let e^* denote the maximum effort that can be induced. This implies that the marginal penalty must be maximized at e^*. It follows from (3) that the marginal penalty is maximized when the penalty is as large as possible for harm levels that become less likely if effort is increased and as low as possible for harm levels that become more likely as effort is increased. As illustrated in Figure 1, the threshold harm h* between the more likely and the less likely outcomes is given by:

$$f_e(h^*, e^*) = 0 \tag{4}$$

Consequently, there should be no penalty when harm levels below h* are observed, and the penalty should be the maximal level P for harm levels above h*. A formal proof is given in Appendix A.

The next step of the proof is to determine the maximum induceable effort e^*, such that h^* can be determined by (4). With the optimal penalty function described in Proposition 1, the firm's total costs can be written as:

$$T(e) = c(e) + mD(e) + P\left[1 - F(h^*, e)\right] \tag{5}$$

The first order condition from (3) becomes:

$$c'(e) + mD'(e) = PF_e(h^*, e) \tag{6}$$

Equation (6) determines the maximum effort that can be induced (e^*) as a function of the optimal threshold (h^*) the moral concern m, and the maximum penalty (P), i.e.

$$e^* = e(h^*; m, P) \tag{7}$$

It can easily be verified by differentiating (6) that $de^*/dm > 0$ and $de^*/dP > 0$. Inserting (7) into (4) gives us:

$$f_e(h^*, e(h^*; m, P)) = 0 \tag{8}$$

Equation (8) determines h^* as a function of the firm's moral concern m and the maximum penalty P. Implicit differentiation with respect to m and P, and using that $de^*/dh = 0$ in optimum, yields the following equations:

$$\frac{dh^*}{dm} = -\frac{f_{ee}(h^*, e^*)}{f_{eh}(h^*, e^*)} \frac{de^*}{dm} \tag{9}$$

$$\frac{dh^*}{dP} = -\frac{f_{ee}(h^*, e^*)}{f_{eh}(h^*, e^*)} \frac{de^*}{dP} \tag{10}$$

From (6) recall that $de^*/dm > 0$ and $de^*/dP > 0$. It follows from assumptions (a3) and (a4) that $f_{eh}(h^*,e^*) < 0$ and $f_{ee}(h^*,e^*) < 0$.[8] This implies that dh^*/dm and dh^*/dP are negative, which completes the proof.

The result of Proposition 1 is illustrated in Figure 2. The two upward sloping curves depict the left-hand side of equation (6), i.e., the marginal cost of effort minus moral gain, for two levels of moral concern, $m_A > m_B$. The two downward sloping curves depict the right-hand side of Equation (6), i.e., the marginal penalty for two different standards of evidence. The firm's optimal choice of effort is where the marginal penalty equals the firm's marginal cost of effort minus moral gain. When firms have moral concern m_A, the effort maximizing standard of evidence is h^*_A. The corresponding marginal penalty is given by the solid downward sloping curve. The firm chooses effort e^*_A. If moral concern is reduced to m_B, while the standard of evidence is unchanged, the firm's optimal effort is reduced to e_B. However, if the standard of evidence is increased in response to reduced moral concern the marginal penalty curve becomes steeper. The dotted line in figure 2 shows the marginal penalty curve for a higher standard of evidence. As a consequence, the negative effect on effort from lower moral concern is alleviated. While effort was reduced to e_B when the standard of evidence remained at h^*_A, it stops at e_B' when the standard of evidence is increased.

The marginal penalty curve becomes steeper when h* is increased. This can be verified by differentiating the marginal penalty with respect to the standard of evidence:

$$\frac{d(PF_e(h^*,e)}{dh^*} = Pf_e(h^*,e)$$

The marginal penalty does not change at $e = e^*_A$ since $f_e(h^*_A, e^*_A) = 0$. However, since $f_{ee}(h^*_A, e^*_A) < 0$, f_e must be positive for effort levels below e_A^*, and so the marginal penalty PF_e has increased.

3. The optimal standard of evidence when firms' moral costs differ

To minimize the cost of crime, the standard of evidence should ideally differ for firms with different moral costs, as shown above. However, the principle of equal legal treatment implies that the standard of evidence must be the same for all firms. As a consequence, it is too low for some firms and too high for others, and their reactions to changes in the standard of evidence differ. In this section, I derive the optimal standard of evidence when firms' moral concern differ, and analyze how it should be changed as moral cost changes.

Let us consider the case where there are only two types of firms. Type A firms have higher moral concerns than type B firms, i.e., $m_A > m_B$. It follows from Proposition 1 that the effort maximizing standard of evidence is lower for A-firms than for B-firms, i.e., $h_A^* < h_B^*$. I focus on the case where the maximum penalty is too low to induce effort above the socially optimal even for type A firms, and therefore too low to induce the socially optimal effort in type B firms. (The case where the penalties may cause overcompliance in some types, is studied in section 4.) Appendix B(i) demonstrates that with two types of firms, it is still optimal to use a threshold penalty function where the penalty is zero for harm levels below a threshold h* and equal to the maximum level P for harm levels above h*. Hence, the threshold h* is the optimal common standard of evidence.

The optimal common standard of evidence lies between the effort maximizing standards of evidence for the two types of firms, i.e., $h_A^* < h^* < h_B^*$. This is shown in Figure 3. The two bell shaped curves, derived in Appendix A(ii), depict the optimal choice of effort for different standards of evidence, for each of the two types of firms. Since $m_A > m_B$, the effort of type A firms is higher than for type B firms for all standards of evidence. A standard of evidence below or equal to h_A^* is not optimal, since a higher standard of evidence would lead to higher effort in type B firms with higher or equal effort in type A firms. Similarly, a standard of evidence above or equal to h_B^* is not optimal for either type of firm. However, between h_A^* and h_B^* there is a trade-off between inducing high effort in type A firms and type B firms. By moving a standard of evidence closer to the ideal standard for type A firms, h_A^*, the A-firms choose higher effort but the B-firms choose lower effort.

It follows from the arguments above that a policy that minimizes the number of crimes may not be optimal. Let us assume that if the standard of evidence is h_A^*, type A firms choose the socially optimal effort e^S, i.e., they obey the law. Since the effort of type B firms is lower than of type A firms, type B firms violate the law for any standard of evidence. Hence, if the aim is to minimize the number of crimes, h_A^* should be chosen as the common standard of evidence. However, as shown above, h_A^* is not an optimal common standard of evidence if the aim is to minimize the total cost of crime. The problem with the standard of evidence that minimizes the number of crimes is that it makes the marginal penalty too low for the serious violations committed by type B firms.

To determine the optimal standard of evidence more precisely and to find how it varies with changes in the distribution of types of firms, this paper undertakes a more formal analysis of how the social cost of crime varies with the standard of evidence. The social cost s_i from a firm of type i can be written as

$$s(h,m_i) \equiv c\big(e_i(h;m_i,P)\big) + D\big(e_i(h;m_i,P)\big) \qquad (10)$$

The superscript on h is omitted where it is clear what the symbol refers to the standard of evidence. The total social cost is the weighted sum of the costs for the two types of firms, with the share of each type as weights. Hence, if α is the share of the firms with the highest moral concern m_A, the total social cost S is:

$$S = \alpha s(h,m_A) + (1-\alpha)s(h,m_B) \qquad (11)$$

There is a trade-off between the social costs of violations from the two groups. A standard of evidence closer to h_B^* results in lower effort and therefore higher social costs from type A firms, but higher effort and therefore lower social costs from type B firms. When S is convex (see Appendix B), the optimal standard of evidence is where the marginal cost of crime from A-firms equals the marginal cost of crime from B-firms, i.e., where

$$\frac{dS}{dh} = \alpha s_h(h, m_A) + (1-\alpha) s_h(h, m_B) = 0 \qquad (12)$$

If the share of type B firms (with low moral concern) increases, the cost of deviating from the ideal standard of evidence for B-firms becomes higher. As a consequence, the standard of evidence should be closer to h_B^*, as summarized in the following proposition.

Proposition 2

The optimal common standard of evidence is higher when the proportion of firms with low moral concern rises.

The proof is given in Appendix B.

The main results from the discussion above are easily generalized to many types of firms. A reduction in the standard of evidence, say from h^1 to $h^2 < h^1$, increases the penalty for harm levels between h^1 and h^2 relative to both higher and lower levels of harm. As a consequence, the expected penalty for effort levels with a high probability of harm between h^1 and h^2 increases relative to both higher and lower effort levels. Hence, the marginal penalty becomes higher for firms with higher effort and lower for firms with lower effort. In other words, the marginal penalty as a function of effort becomes less steep as the standard of evidence is reduced. As a consequence, the firms with high effort (high moral concern) choose even higher effort and the firms with low effort (low moral concern) choose even lower effort. The optimal common standard of evidence is determined as the optimal tradeoff between these opposing effects. If the proportion of the firms with low moral concern is increased, it becomes more important to raise the marginal penalty for these firms, and as a consequence the standard of evidence should be increased.

While Proposition 2 implies that the standard of evidence should be lowered when there are more high moral firms, it does not imply that it should be lowered when high moral firms become even more morally concerned. Even if the ideal standard of evidence for A-firms is lower as their moral concern (m_A) becomes higher, the optimal common standard of evidence may actually be higher when A-firms become more morally concerned.

The reason for this paradoxical result is that higher moral concerns for either of the types of firms have two effects. First, firms that become more morally concerned increase their effort for all standards of evidence. This implies that the marginal gain from increased effort in these firms goes down, and more weight should be given to increasing effort in the firms that have not become more morally concerned. There is less need to deter the firms as they internalize more of the social cost of crime. As a consequence, the standard of evidence should be closer to the optimal level for the firms where moral concern has not increased. The second effect is that higher moral concern may make effort more or less responsive to changes in the standard of evidence. If effort becomes more responsive, the standard of evidence should be closer to the ideal level in the firms that have become more morally concerned. If effort becomes less responsive, the standard of evidence should be closer to the ideal level for the other group. Combining these two effects gives us the following result:

Proposition 3

If higher moral concern in firms with high moral makes their effort less responsive to changes in the standard of evidence, the standard of evidence should be set closer to the ideal level for firms with low moral , and vice versa.

A formal proof is given in Appendix B.

4. Over-compliance

In the analysis above I assumed that the firms' moral concerns are so low that no firms choose effort above the socially optimal level. However, when firms differ, an enforcement policy that induces optimal effort in firms with low moral concern may cause too much effort in firms with high moral concern. Figure 4 illustrates one such case. Type A firms have higher moral concern than do type B firms. At the standard of evidence that would induce maximum effort e_B^* in type B firms (with low moral concern), type A firms (with high moral concern) would choose an effort level that is higher than the social optimal effort e^S. As seen from the figure, both h^I or h^{II} would induce the social effort level e^S in type A firms. This implies that $s(h, m_A)$ is not concave, but has local minima at h^I and h^{II}. As a consequence, total social costs S may not be concave in h and we cannot determine the optimal standard of evidence from the first order condition only.

In Figure 4, the highest standard of evidence that induces the socially optimal effort in type A firms is too high to induce the maximal effort in type B firms. Appendix C demonstrates that in this case the high moral firms always over-comply when the standard of evidence is set optimally. The optimal standard of evidence, h^*, is either between h^I and h_A^* or between h_B^* and h^{II}. In cases where $h_B^* < h^* < h^{II}$, h^* is too high for the low moral firms and too low for the high moral firms. As a consequence, a reduced standard of evidence (below h^*) would reduce crime, but also increase over-compliance among the law-abiding firms. The loss from increased over-compliance would exceed the gain from less crime.

5. Concluding remarks

Illegal insider trading is an example of an economic crime to which my model may be applied.[(9)] A reduced standard of evidence for insider trading has been discussed in many countries. To prove that a trader had access to precise and price-affecting confidential information, the prosecution usually has to make uncertain inferences from telephone calls, meetings, trading patterns, relationships between people and other circumstantial evidence. Such inferences seldom meet the standard of evidence required in a criminal court. Hence, in countries that have only provided for criminal enforcement of insider trading, like the Netherlands and Norway, few people are convicted. The Netherlands had only one conviction in the period 1988-98 (Newkirk and Robertson, 1998). Norway has had only one conviction since 1985, and has recently revised its law on securities exchange (Verdipapirhandelsloven) to provide for civil sanctions of insider trading.

Before the revision, sanctions were only imposed after a criminal charge, and the criminal standard of evidence applied. After the revision, some violations of insider rules can be sanctioned administratively, and the civil standard of evidence applies (Ot.prp. 80, 2001). Using my model, we can examine whether these reductions in the standard of evidence will reduce the cost of illegal insider trading.

One implication of my model is that there is a trade-off between minor and large crimes. If one chooses a high standard of evidence to deter serious crimes, the number of small violations is higher than if the standard of evidence was reduced. Hence, if there are few convictions but many small violations, as is the case for illegal insider trading, this does not imply that the choice of the standard of evidence is wrong. On the contrary, it may be that the agent with low moral costs (or a low opportunity cost of crime) is successfully induced to choose small crimes instead of serious crimes because the high standard of evidence makes minor crimes relatively profitable. If one reduces the standard of evidence, the number of crimes will most likely decrease, but the total cost of insider crimes may well increase because the marginal cost of serious crimes is reduced.

When agents can choose the severity of the violation, the optimal standard of evidence depends on the composition of the agents. For example, if agents with high opportunity costs of crime dominate, such that there are few serious criminals to deter, the standard of evidence should be low to make the marginal costs for minor crimes high. Hence, if the main problem of insider trading is that a large proportion of the agents may be tempted to commit minor violations, access to further civil sanctions may help. However, if deterrence of serious violations is most important, the standard of evidence should remain high.

One crucial assumption in my analysis is that a reduction in the standard of evidence increases the expected penalty even for those who obey the law. An agent may reduce the chance of wrongful conviction, but there is no "safe" behavior. It might be argued , however, that for some types of economic crime law obedient agents are safe, and proper business conduct will not be convicted. If there exist such a safe alternative, a reduction in the standard of evidence may be more effective to reduce crime than the model predicts. If law obedience is a safe alternative, a reduction in the standard of evidence has two effects: On the one hand, it reduces the penalty for serious violations relative to minor ones. On the other hand, it increases the expected penalty for all violations relative to law obedience. Yet, the main message of my analysis is still relevant: Although a lower standard of evidence may induce more agents to obey the law , it causes more serious law violations among those who do not.

Appendix A:

(I) Proof of Proposition 1: The optimal penalty function

The firms' choice of effort is determined by Equation (3). Since c''(e) + mD''(e) > 0 the penalty function P(h) that induces maximum effort is the one that maximizes the right hand side of (3). Let e* denote the maximum effort. The optimal penalty function is then:

P(h) = P for h-values where $f_e(h,e^*) \leq 0$ and P(h) = 0 for h-values where $f_e(h,e^*) > 0$.

The threshold h* that triggers the maximum penalty is given by Equation (4), i.e., by $f_e(h^*,e^*) = 0$. Since $f_{eh}-f_ef_h < 0$ for all h (from (a3)), it follows that $f_{eh}(h^*,e^*) < 0$. As a consequence, h* is the unique h-value that satisfies (4). Since $f_e(h,e^*) > 0$ for h < h* and $f_e(h,e^*) < 0$ for h > h*, it follows that the optimal penalty is zero for h < h* and P for $h \geq h^*$, which completes the proof.

(II) The optimal effort as a function of the standard of evidence

For a standard of evidence h^t, the firm's optimal choice of effort is given by Equation (6) when h^* is substituted with h^t, i.e.

$$c'(e) + mD'(e) = PF_e(h^t, e) \tag{A.1}$$

Hence, if e^0 denotes the firm's optimal effort choice, e^0 is given by

$$e^0 = e(h^t; m, P) \tag{A.2}$$

Since h^* is the threshold that induces the maximum effort e^*, it follows that that $e^0 < e^*$ for $h^t \neq h^*$. Differentiating (A.1) with respect to h^t we obtain:

$$\frac{de}{dh^t} = \frac{Pf_e(h^t, e)}{T''(e)} \tag{A.3}$$

where $T''(e^0) > 0$. Since h* is the only h-value that satisfies $f_e(h, e(h;m,P)) = 0$, and $f_{eh} < 0$ for $f_e = 0$, it follows that $f_e(h^t,e^0) > 0$ for $h^t < h*$ and $f_e(h^t,e^0) < 0$ for $h^t < h*$. Hence,

$$de^0/dh^t \begin{cases} > 0 \text{ for } h^t < h^* \\ = 0 \text{ for } h^t = h^* \\ < 0 \text{ for } h^t > h^* \end{cases} \qquad \text{(A.4)}$$

Differentiating de/dh^t twice with respect to h^t yields

$$\frac{d^2e}{dh^2} = p\left\{f_{eh} + f_{ee}\frac{de}{dh} - f_e\frac{T'''}{T''}\right\} \qquad \text{(A.5)}$$

For $h^t = h^*$, $f_e=0$ and $de/dh^t=0$ such that $\frac{d^2e}{dh^2} = pf_{eh}<0$. Hence e^0 is concave in h^t at the optimum, but it may not be concave for other values of h^t .

Appendix B:

(I) The optimal penalty function with two types of firms.

The total social costs is a function of e_A and e_B, given by

$$S = \alpha\big(c(e_A) + D(e_A)\big) + (1-\alpha)\big(c(e_B) + D(e_B)\big) \qquad \text{(B.1)}$$

When both e_A and e_B are below the socially optimal level, $c'(e_i) + D'(e_i) < 0$ such that S is decreasing in e_A and e_B. The optimal effort for a firm of type i is given by (3), i.e.

$$T'(e_i) = c'(e_i) + m_i D'(e_i) + \int_0^H P(h) f_e(h, e_i)dh = 0 \qquad \text{(B.2)}$$

The optimal penalty function P(h), where $0 \leq P(h) \leq P$, is the one that maximizes S when e_A and e_B are determined by (B.2).

Let us assume that we have found the optimal P(h), and thereby determined the optimal effort in the two types of firms, ε_A and ε_B. I proceed in two steps. (i) First, I show that P(h) = 0 is optimal for $h \leq h_1$, where h_1 is given by $f_e(h_1, \varepsilon_A) = 0$ and P(h) = P is optimal for $h \geq h_2$, where h_2 is given by $f_e(h_2, \varepsilon_B) = 0$. (ii) Second, I show that for $h_1 < h < h_2$, the penalty must either be zero or maximal (P). Moreover, P(h) jumps from zero to the maximal value P at a threshold level h* between h_1 and h_2.

(i) It follows from (B.2) that the optimal P(h) is zero, for h values where $f_e > 0$ for both types of firms. If this were not the case, a decrease in the penalty for these h-values would increase the marginal gain from effort for both types of firms, and thereby increase the effort of both types. Since h_1 is a unique h-value that satisfies $f_e(h_1,\varepsilon_A) = 0$, and $f_{eh}(h_1, \varepsilon_A) < 0$ (from (a.3)) it follows that $f_e(h,\varepsilon_A) > 0$ for $h < h_1$. Using the same arguments, $f_e(h, \varepsilon_B) > 0$ for $h < h_1$ since $f_e(h_2, \varepsilon_B) = 0$ and $h_1 < h_2$. Hence, P(h) = 0 is optimal for $h < h_1$. From (B.2), the optimal P(h) equals P for h values where $f_e < 0$ for both types of firms. By the same reasoning as in the case where $h < h_1$, both $f_e(h, \varepsilon_B) < 0$ and $f_e(h,\varepsilon_A) < 0$ for $h > h_2$, i.e P(h) = P is optimal.

(ii) Differentiating (B.1) and (B.2) gives

$$dS = \alpha[c'(e_A) + D'(e_A)]de_A + (1-\alpha)[c'(e_B) + D'(e_B)]de_B \tag{B.3}$$

$$T''(e_i)de_i = -\int_{h1}^{h2} dP_h f_e(h, e_i)dh \tag{B.4}$$

where dP_h denotes a change in the penalty for observed harm h. Inserting de_A and de_B from (B.4) into (B.3), the following must hold for h-values between h_1 and h_2 :

$$dS = \int_{h1}^{h2} \{\alpha\psi(\varepsilon_A) f_e(h, \varepsilon_A) + (1-\alpha)\psi(\varepsilon_B) f_e(h, \varepsilon_B)\} dP_h dh \tag{B.5}$$

where $\psi(\varepsilon_i) = [c'(\varepsilon_i)+D'(\varepsilon_i)]/T''(\varepsilon_i)$. Since $s'(\varepsilon_i) < 0$ and $T''(\varepsilon_i) > 0$, $\psi(\varepsilon_i) < 0$. Moreover, $f_e(h,\varepsilon_A) < 0$ and $f_e(h,\varepsilon_B) > 0$ between h_1 and h_2, such that the first term inside the brackets is positive and the second is negative. For h-values such that the bracket is positive, $dS/dP_h < 0$ which implies that $P(h) = 0$ is optimal. For h-values such that the bracket is negative, $dS/dP_h < 0$ which implies that $P(h) = P$ is optimal. Hence, $P(h) = 0$ is optimal for h-values where $\alpha\psi_A f_e(h,\varepsilon_A) > -(1-\alpha)\psi_B f_e(h,\varepsilon_B)$, and $P(h) = P$ otherwise. When h is close to h_1, $dS/dP_h < 0$ since $f_e(h_1,\varepsilon_A) = 0$ and $f_e(h_1,\varepsilon_B)>0$. When h is close to h_2, $dS/dP_h > 0$ is positive since $f_e(h, \varepsilon_B) = 0$. Hence, if $dS/dP_h = 0$ for one h-value only, given by $\alpha\psi_A f_e(h^*,\varepsilon_A) = -(1-\alpha)\psi_B f_e(h^*,\varepsilon_B)$, it follows that $dS/dP_h < 0$ for $h_1<h<h^*$ and $dS/dP_h > 0$ for $h^*<h<h_2$, i.e. the optimal penalty function changes from 0 to P at h*.

(II) Is S convex in h?

Total social cost S is a weighted sum of social cost from the two types of firms, as given by Equation (11). A sufficient, but not necessary, condition for S to be convex in h is that $s(h,m_i)$ is convex in h, where $s(h,m_i)$ is given by(10). Differentiating $s(h,m_i)$ twice with respect to h gives us

$$s_i''(h,m_i) = \left(c'(e_i) + D'(e_i)\right)\frac{d^2e_i}{dh^2} + \left(c''(e_i) + D''(e_i)\right)\left(\frac{de_i}{dh}\right)^2 \tag{B.6}$$

The last term is positive since $c'' > 0$ and $D'' > 0$. Moreover, we know that $c'(e) + D'(e) < 0$, when the maximum feasible effort is below the socially optimal effort. Hence, a sufficient, but not necessary, condition for s_i to be convex is that $d^2e_i/dh^2 < 0$, i.e., $e(h;m,P)$ is concave in h . As shown in Appendix A, $e(h;m_i)$ is concave in h for h_i^*, but not necessarily for other h-values.

(III) Proof of Proposition 2

Differentiating Equation (12) with respect to α, and reorganizing, we obtain:

$$\frac{dh}{d\alpha} = \frac{s_h(h, m_A) - s_h(h, m_B)}{S_{hh}} \tag{B.7}$$

If S is convex in h, the denominator is positive. Moreover, it follows from Equation (12) that the nominator is negative. This implies that $dh/d\alpha < 0$, which completes the proof.

Proof of Proposition 3

Differentiating Equation (12) with respect to m_i we obtain:

$$\frac{dh}{dm_A} = \frac{-\alpha s_{hm}(h, m_A)}{S_{hh}} \tag{B.8}$$

$$\frac{dh}{dm_B} = \frac{-(1-\alpha) s_{hm}(h, m_B)}{S_{hh}} \tag{B.9}$$

When $S_{hh} > 0$, the sign of dh/dm_B has the opposite sign of $s_{hm}(h,m_B)$. Differentiating Equation (10) with respect to h and then with respect to m_i we obtain:

$$s_{sm}(h, m_i) = (c''(e_i) + D''(e_i))\frac{de_i}{dh}\frac{de_i}{dm_i} + (c'(e_i) + D'(e_i))\frac{d^2 e_i}{dhdm_i} \tag{B.10}$$

Since $c'' + D'' > 0$ and $de_i/dm_i > 0$, the first term has the same sign as de_i/dh. Hence, the term is positive for type B firms, and negative for type A firms. Since $c' + D' < 0$, the second term has the same sign as $d^2e_i/dhdm_i$, which illustrates how the responsiveness of effort to h changes with m.

If the effort becomes less responsive to h as m is higher, i.e., if $d^2e_A/dhdm_A > 0$ and $d^2e_B/dhdm_B < 0$, we obtain $s_{hm}(h,m_A) < 0$ and $s_{hm}(h,m_B) > 0$. Hence, (B.3) and (B.4) implies that: $dh_B/dm_B < 0$ and $dh_A/dm_A > 0$.

Appendix C: The optimal standard of evidence when one type of firm over-complies.

Let h^I denote the lowest standards of evidence that induce optimal effort in type A firms (with high moral concern), while h^{II} denotes the highest standard of evidence. h^I will be too low to induce the maximum feasible effort e_B^* in type B firms (with low moral concern) and only the case where h^{II} is too high to induce e_B^*, shown on Figure 4, is discussed. This implies that the optimal standard of evidence, h^*, either lies between h^I and h_A^* or between h_B^* and h^{II}, i.e., $h^I < h^* < h_A^*$ *or* $h_B^* < h^* < h^{II}$. To prove this, I show that the other possible choices of h^* cannot be optimal. These possible choices are:

(i) $h^* < h^I$; (ii) $h_A^* < h^* < h_B^*$; or (iii) $h^* > h_B^*$.

(i) First, it follows from an inspection of Figure 4 that $h^* < h^I$ cannot be optimal, since an increase in h^* would induce both type A and type B firms to choose effort levels closer to their respective optimal levels. *(ii)* Second, by similar reasoning, $h^* > h^{II}$ cannot be optimal. *(iii)* Finally, if $h_A^* < h^* < h_B^*$, e_A is higher than the optimal level, while e_B is lower. Hence, $c'(e_A) + D'(e_A) > 0$ and $c'(e_B) + D'(e_B) < 0$. Moreover, we see from Figure 4 that $de_A/dh < 0$ and $de_B/dh > 0$. This yields:

$$\frac{dS}{dh} = \alpha[c'(e_A) + D'(e_B)]\frac{de_A}{dh} + (1-\alpha)[c'(e_B) + D'(e_B)]\frac{de_B}{dh} < 0$$

i.e., the total social cost is reduced if h^* is increased. It follows that $h_A^* < h^* < h_B^*$ cannot be optimal.

FOOTNOTES

Acknowledgements: I am grateful to Steinar Holden, Aanund Hylland, Tor Jakob Klette, Karl Ove Moene and Atle Seierstad for useful comments

(1) The standard of evidence is derived from the standard of proof .The defendant is convicted if the probability that he is guilty given the evidence exceeds the "standard of proof". If for example the probability of guilt is increasing in the observed variable x, the standard of proof determines the standard of evidence as the lowest value of x that leads to conviction .

(2) The argument that civil sanctions should be an option because they lower the standard of evidence is used explicitly in debates on crime enforcement. For example, a speech by staff of the US Securities Exchange Commission at the16th International Symposium on Economic Crime argued that, "The burden of proving a purely circumstantial case is less onerous in the civil context, where guilt need be shown only by a preponderance of evidence, rather than beyond reasonable doubt, and where the use of presumption may shift the burden of proof to the defendant under certain circumstances." (Newkirk and Robertson, 1998).

(3) In this paper I do not address the question of how the evidence can be improved by the judicial process, as discussed in Rubinfeld and Sappington (1987). Moreover, I ignore purely legal mistakes, and only discuss legal errors that are inevitable because evidence is uncertain. The probabilities of the two errors are therefore endogenously determined by the distribution of the evidence and the judgement rules followed by the court. By contrast, Polinsky and Shavell (1989) discuss the effects of legal errors that include pure mistakes by the court such as misinterpretation of the evidence or misinterpretation of the law. As a consequence, they treat the probabilities of the two types of error as exogenous variables that may vary independently of one another.

(4) By taking the derivative of the expected harm $\int_0^H d(h) f(h,e)dh$ twice and using partial integration, we obtain $\int_0^H d(h) f_{ee}(h,e)dh = -\int_0^H d'(h) F_{ee}(h,e)dh \geq 0$. Since d'(h) > 0 and $F_{ee} \geq 0$, it follows that D(e) is convex in effort.

(5) Since P'(h) ≥ 0, the proof that expected penalty is convex in effort follows the same steps as the proof that D(e) is convex, as in footnote 1.

(6) Including socially costly penalties, such as the withdrawal of production licenses, complicates the analysis without adding significant insights.

(7) In the extreme case where $h^* = 0$, i.e., where any observed harm is punished, the punishment has no effect, since expected punishment is independent of the effort level chosen.

(8) In optimum, $f_e(h^*,e^*) = 0$. Hence, since assumption (3) implies that f_{eh}-$f_e f_h < 0$, it follows that $f_{eh} < 0$ in optimum. Moreover, following assumption (4), $f_e(h^*,e^*) = 0$ further implies that $f_{ee}(h^*,e^*) < 0$ in optimum.

(9) I define insider trading as trading on precise and confidential information, not only illegal trading by those defined as "primary insiders", whose trading is explicitly regulated by law. While the trading of the primary insiders is relatively easy to monitor, it is difficult to prove that other traders had access to inside information. One may therefore expect that the main problem is insider trading by other than those defined as insiders by law (see e.g., Eckbo and Smith (1998)).

REFERENCES

Eckbo, E.B. and D.C. Smith (1998), "The conditional performance of insider trades", *The Journal of Finance*, vol. LIII, no.2, pp 467- 498.

Newkirk, T.C. and M.A. Robertson (1998), "Insider trading - A U.S. perspective", Speech at the 16th International Symposium on Economic Crime, Jesus College, Cambridge, England.

Ot.prp.nr.42 (1997-98), Lov om årsregnskap.

Ot.Prp.nr.80 (2000-20001), "Om lov om endringer i lov 9.juni 1997 nr.79 om verdipapirhandel" (tiltak mot ulovlig innsidehandel).

Polinsky, A.M. and S. Shavell (1989), Legal error, litigation and the incentive to obey the law", *Journal of Law, Economics and Organization*, vol.81, no.2, pp 99- 108.

Rubinfeld, D.L. and D.E.M. Sappington (1987), "Efficient awards and standards of proof in judicial proceedings", *RAND Journal of Economics*, vol.81, no.2, pp 308-15.

Schrag, J. and S. Scotchmer (1994), "Crime and prejudice: The use of character evidence in criminal trials", *Journal of Law, Economics and Organization*, vol.10, no.2, pp 319-42.

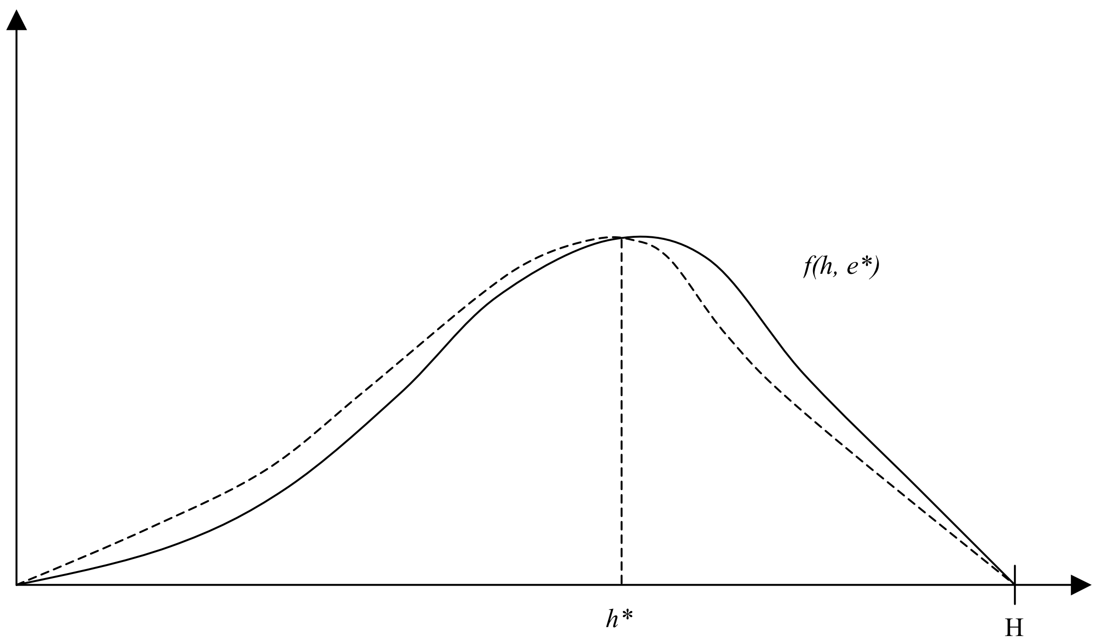

Figure 1

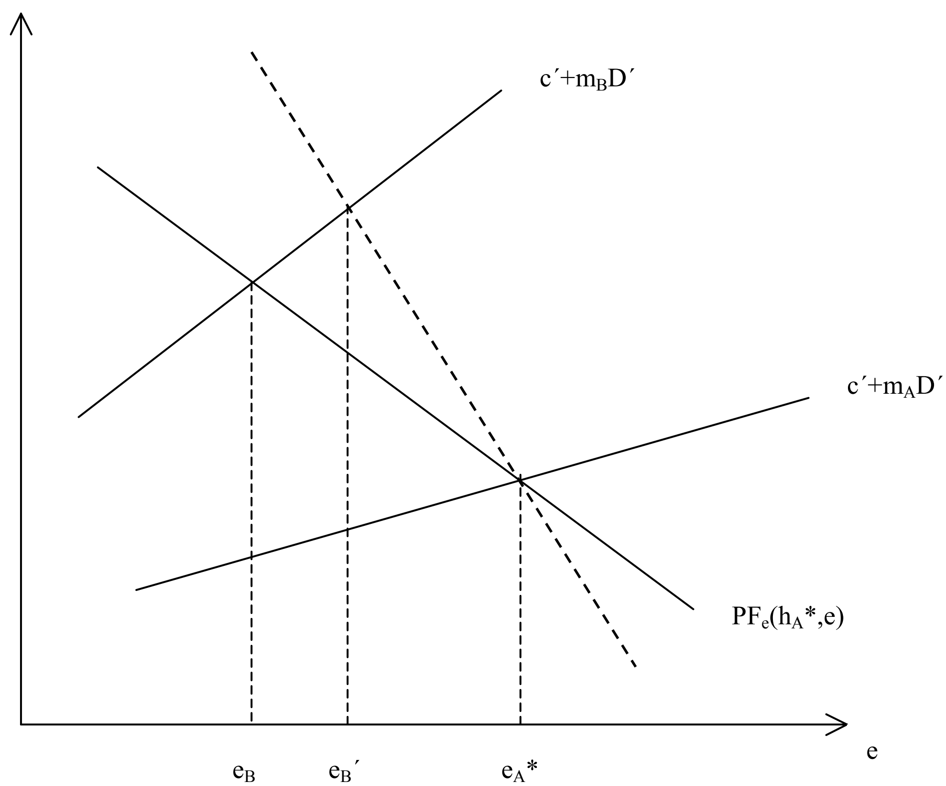

Figure 2

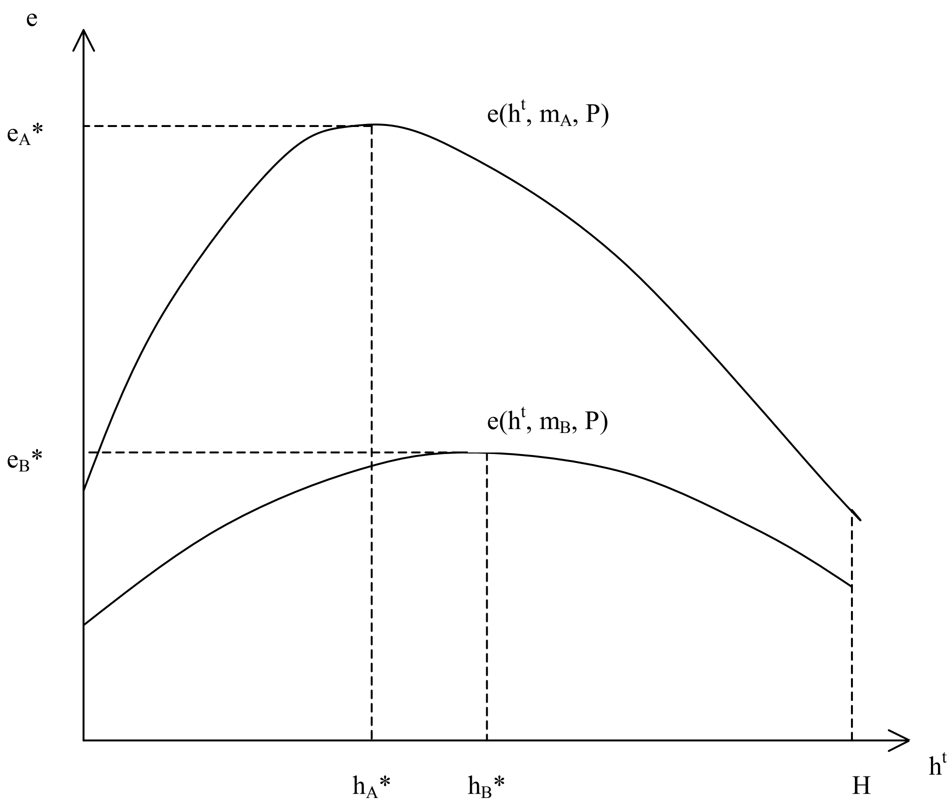

Figure 3

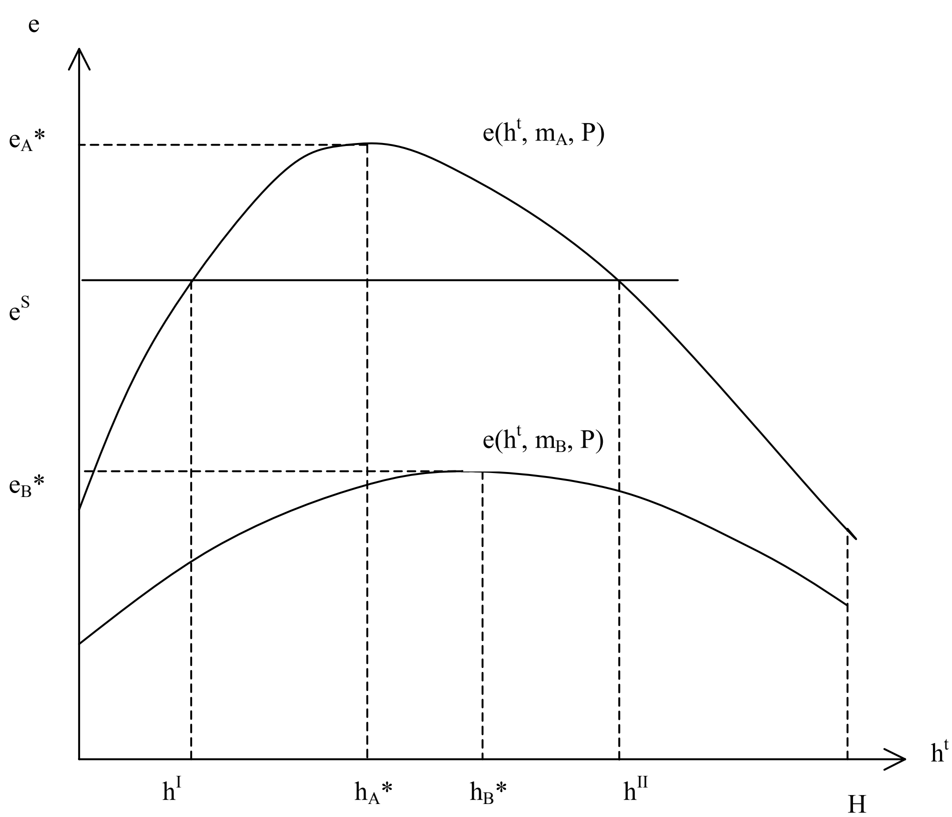

Figure 4