The Use of Warnings in the Presence of Errors [a]

Sandra Rousseau

CEDON, KULeuven, Warmoesberg 26, B-1000 Brussel, Belgium

E-mail: sandra.rousseau@kuleuven.be

CITE AS: Rousseau, S. (2009). The use of warnings in the presence of errors. *International Review of Law and Economics,* 29, 191-201 (doi:10.1016/j.irle.2009.01.001)

Abstract

This article studies the effects of warnings, an enforcement instrument which is often used by environmental inspection agencies. Due to regulatory errors, measured emissions are uncertain and some firms are unjustly penalized. Warnings can then be used as a means to reduce the consequences of these errors. Even though the presence of warnings creates some underdeterrence of medium-cost firms, such a system reduces the overcompliance of low-cost firms caused by the uncertainty surrounding measured emissions. Further, warnings reduce the number of incorrect prosecutions in the case of measurement errors, which is also welfare enhancing, albeit at the cost of increasing the number of violators that go unpunished. For small error sizes, the use of warnings is shown to be welfare improving compared to only using fines.

JEL: K42 Illegal behaviour and the Enforcement of Law

Keywords: Enforcement; non-monetary instruments; warnings; measurement errors

[a] I gratefully acknowledge the financial support of the FWO project 'An economic approach to modelling the enforcement of environmental regulation' and the SBO project 060034 (www.environmental-lawforce.be). I would also like to thank Eef Delhaye, Ronald Rousseau, Ellen Moons and two anonymous referees for their useful suggestions.

1. Introduction

Firms are typically subject to public monitoring and enforcement in a variety of areas. This contribution describes firms' compliance decisions concerning environmental regulations, but the insights provided extend beyond that specific policy area. We focus on the use of warnings as an enforcement instrument and investigate what implications this entails on firms' pollution levels, on the prevalence of regulatory errors and on social welfare levels. To this end, we define a warning as the act of not issuing a fine to a firm measured to be in violation (i.e. a fine-amnesty). In the model, a previously detected violator faces a slightly increased fine by the withdrawing of this fine-amnesty and it takes this effect into account when making its compliance decisions.

The economic literature[1] on the deterrence of crime, which starts with Becker (1968), has focused mainly on monetary sanctions, and more specifically on fines. When non-monetary sanctions are studied in the literature, authors typically discuss prison sentences (e.g. Shavell, 1987a and Garoupa and Klerman, 2004). Informal enforcement instruments, such as warnings and advices, are only rarely studied. Warnings are, nonetheless, often used as an enforcement instrument by the environmental agency and they instruct firms to end a situation of non-compliance and abide with all appropriate laws, decrees and permits. For example, Rousseau (2007) describes the enforcement actions taken after audits that found one or more breaches of environmental regulations by Flemish textile companies. This study finds that in almost ninety percent of the cases where an enforcement action was taken a warning was given to the violator.[2] Further evidence of the frequent use of warnings can be found in Harrington (1988), Eckert (2004) and Nyborg and Telle (2006).

A recent study focusing on warnings is Nyborg and Telle (2004), in which the authors investigate the potential of warnings to help regulators keep control. In their model a warning provides firms with the opportunity to move into compliance before the end of the period and in exchange these firms will not be prosecuted. Thus a warning does not increase the expected penalty of the warned firm, contrary to the way we define a warning. Using a game theoretic approach, Nyborg and Telle argue that *'warnings reduce substantially the probability of (...) accidental switches from the full-compliance to the no-compliance equilibrium'*. Our approach differs from theirs in that we look at the welfare implications of using warnings while they focus on the compliance rate among firms. Moreover, we allow for continuous violations, for cost heterogeneity between firms and consider measurement errors rather than managerial errors. A more empirical approach is taken by Eckert (2004), who shows that warnings are used by the Canadian inspection

[1] For an overview of the literature on enforcement of environmental regulation, see Cohen (2000).

[2] In concrete figures; 178 warnings were issued out of the 199 cases where an enforcement action was taken after the observation of an environmental violation during an inspection. This warning was either the only instrument used (19 %) or accompanied by a notice of violation (70%). A notice of violation formally documents a violation and a copy is always sent to the Public Prosecutor. For more detail on this case study we refer to Rousseau (2007).

agency to target firms for inspection, with past warnings increasing the probability of an inspection relative to a past finding of compliance.

Furthermore, the basic dynamic enforcement model of Harrington (1988)[3] can also be reinterpreted as clarifying the role of warnings. In his model, firms are divided into two groups: 'good' (group 1) and 'bad' (group 2) firms. When 'good' firms are found in violation, they are not fined. They are, however, sent to group 2, where they face a higher expected fine. This act of not fining firms and sending them to group 2 could also be translated as a warning. Indeed, in Harrington's model it is clear that being moved to group 2 with a higher fine and an increased inspection frequency is in fact an enforcement action that increases the expected penalty.

In our model, the agency increases the expected fine of a previously detected violator by withdrawing the fine-amnesty. In agreement with traditional rational crime theory, the expected fine of a warned firm therefore (weakly) increases. Indeed it is common practice to punish repeat offenders more severely for the same offense than first-time offenders. Escalating penalties have therefore been studied in various settings: see, for example, Chu et al. (2000), Baik and Kim (2001), Polinsky and Rubinfeld (1991), Emons (2003, 2004 and 2007) and Delhaye (2007). However, escalating penalty schemes often provide mixed results with respect to deterring potential offenders. In Emons (2007), for instance, sanctions increasing with offense history are optimal when the benefit of the offense is high; however, when the benefit is low, decreasing sanctions will minimize enforcement costs. Indeed also in our set-up, increasing penalties for repeat offenders (i.e. the use of warnings combined with fines) are not unambiguously better than uniform penalties.

Furthermore, we take into account that sometimes firms' emissions are incorrectly measured due to equipment failure, human errors while taking samples, or because of inaccuracies in laboratory tests. If such measurement errors are present, some compliant firms are unjustly sanctioned and some violating firms remain unpunished. Bose (1995) has already shown that regulatory errors can cause the optimal penalty to be non-maximal. Looking at only one enforcement instrument, he has optimized the level of the fine, which is independent of the seriousness of violation. Similarly, Chu et al. (2000) argue that optimal sanctions should be mild for first time offenders but severe for repeat offenders if erroneous conviction of innocent offenders is possible. Polinsky and Shavell (2000), too, state that '*introducing the possibility of mistakes may increase the desirability of lowering the fine because, due to type II errors*' (i.e. mistakenly punishing compliant individuals), *'individuals who do not violate the law are subject to the risk of having to pay a fine'*. The presence of errors implies that firms are confronted with uncertainty concerning the type and size of punishment they can expect.

[3] The analyses of Harford (1991), Russell (1990) and Russell et al. (1986) have extended this model in order to allow for measurement errors. However, these studies concentrate on maximizing deterrence for a given regulatory budget, and hence they do not consider the welfare implications of regulatory errors.

As Craswell and Calfee (1986) point out, uncertainty changes the deterrent impact of legal rules by creating two opposing effects. Firstly, it creates a positive chance that a violator will not be punished, thus reducing the incentives to comply. We call this effect 'underdeterrence' of firms in the remainder of the text. However, uncertainty often allows an actor to reduce the probability of punishment even further by modifying his behavior by more than the law requires. This second effect can give even risk-neutral parties an incentive to 'overcomply'. The relative strength of these two effects will depend on various features of the legal environment, such as the benefits and costs generated by the regulated activity and the penalty structure for violations of the legal rule. The findings of Craswell and Calfee (1986) on the effects of errors on compliance are also confirmed by Shavell (1987b) and Cooter and Ulen (1997).

Finally, the model proposed here allows for fines increasing in the size of the violation. We show that informal enforcement instruments should be considered as a complement to formal, monetary sanctions. In the presence of regulatory errors, three different regulatory systems are compared. Hence, we investigate the effect of warnings on firms' emissions, the impact on the prevalence of errors and the welfare comparisons of regimes with and without warnings. Regulatory errors are typically associated with social costs: incorrect sanctioning decreases the trust people put into the legal system and it is perceived to be unjust and arbitrary. In this instance, warnings can reduce the welfare cost of falsely accused firms. Also, warnings can reduce the overcompliance of low-cost firms that is caused by including measurement errors. These errors motivate some low-cost firms to reduce their emissions below the emissions standard in order to avoid the risk of being fined. However, warnings can worsen the underdeterrence of medium-cost firms caused by the measurement errors. Those firms take advantage of the fact that, under certain conditions, small violations lead to warnings and not to monetary punishments. We find that, if the social costs of the conviction of compliant firms are sufficiently large, the use of warnings can indeed increase welfare.

The remainder of the paper is organized as follows. In section 2, we present the model and the assumptions made. In section 3, we analyze the basic model without measurement errors. In section 4, we deal with the model including measurement errors. In section 5 we compare the welfare effects of the different policy scenarios. We conclude in section 6. All the proofs are in the Appendix.

2. Assumptions

The model considers firms each of which face an identical emission standard $\overline{e}$. In each period *t,* firms are inspected with an exogenously given unconditional probability *p*, with $t > 0$ representing the number of periods the enforcement system has been in place. Inspections are assumed to be independent of the compliance history. In the concluding section, the implications of increased inspections for repeat offenders are briefly discussed.

When firm i is inspected and its actual emissions are e_{ti}, its measured emissions equal:

$$\begin{aligned} e_{ti}^{m} &= e_{ti}\left(1-\gamma\right) && \textit{with probability } q \\ &= e_{ti} && \textit{with probability} \left(1-2q\right) \\ &= e_{ti}\left(1+\gamma\right) && \textit{with probability } q \end{aligned} \tag{1}$$

This formulation implies that there is uncertainty with respect to the discharge level that is actually measured. This uncertainty is represented by the parameter γ measuring the size of the measurement error $\left(0 \le \gamma \le 1\right)$ and the parameter q that measures the probability of erroneous measurement $\left(0 \le q < \frac{1}{2}\right)$. This implies that type I (some violators are mistakenly found to be complying) and type II (some compliant firms are erroneously found liable) measurement errors are present and equally likely. The sampling and measuring equipment used by the inspection agency can be inexact or the analysis of samples can be inaccurate, such that with probability q the measured result is higher than the actual emissions and with probability q measured emissions are below actual discharges[4]. The accuracy of the monitoring technology is higher, the lower is q and the lower is γ. If the technology is perfect, $q = \gamma = 0$. These parameters are assumed to be known to both the regulator and the firm.

Furthermore, the measurement errors are assumed to be independent over time. The uncertainty in measured emissions implies that the emission level on which the sanction is based is not exactly known. The firm will respond to this uncertainty by adjusting its behavior (see Craswell and Calfee, 1986). Note that re-measurement is not sufficient to remove all errors even though the probability of having two consecutive errors is small. In this model we assume exogenous errors; nonetheless the probability of making a mistake as well as its size may be functions of both firm and regulatory actions. Indeed, firms can invest in avoidance activities (see, for instance, Malik, 1990 and Innes, 2005) and the regulator might invest in inspection technologies of better quality leading to fewer and/or smaller mistakes (see Malik, 1993).[5]

The fine function is based on measured emissions and is defined as:

$$\begin{aligned} f\left(e_{ti}\right) &= F\left(e_{ti}^{m}\right) && \textit{if } e_{ti}^{m} > \overline{e} \\ &= 0 && \textit{if } e_{ti}^{m} \le \overline{e} \end{aligned}$$

If a firm i is found in violation at time t $\left(e_{ti}^{m} > \overline{e}\right)$, then a fine $F\left(e_{ti}^{m}\right)$ is levied with $F' > 0$ and $F'' > 0$. Also we assume that the right-hand side derivative of the fine function in point

[4] The assumption of symmetric error probabilities is a simplification. However, incorporating asymmetric errors would not significantly affect the results of the analysis.

[5] As shown by Malik (1990) one implication of incorporating avoidance behavior by firms would be that penalties need not always be set as high as possible. Furthermore, allowing the agency to invest in more accurate monitoring technology would reduce the social costs associated with regulatory errors; whether this is also welfare enhancing will depend on the size of the investment required.

$\overline{e}$ is strictly positive: $F'^{+}\left(\overline{e}\right) > 0$. This implies that even an infinitely small violation would results in a non-negligent positive fine and it reflects the practice that fines are generally constrained by a minimum (and maximum) amount defined in the corresponding legislation or sentencing guidelines. Reasons for imposing such bounds on penalties include fairness considerations and the observation that the administrative costs associated with imposing penalties are at least partly independent of the actual size of that penalty.

Firms are assumed to be risk neutral and identical except for the continuous cost parameter $\theta_i \in \left[\underline{\theta}, \overline{\theta}\,\right]$. Importantly, these cost parameters are completely unobservable by the agency. Firms initially emit $e_o\left(> \overline{e}\right)$ units of the regulated pollutant. In order to reduce emissions, firm *i* incurs a cost $\theta_i\, C\left(e_{ti}\right)$ in period *t* with $C' < 0,\ C'' > 0$.[6] Actual emissions e_{ti} are equal to baseline emissions e_o subtracted by abated emissions.

3. Basic model without measurement errors

In order to compare the impact of different regulatory schemes, it is first necessary to define what happens if there is no uncertainty with respect to the measured emissions and measurements carried out by the inspection agency correctly determine the firm's emissions $\left(\gamma = 0\right)$. Managers know that, if they choose to exceed the emission standard and they are inspected, the violation will be detected with certainty (no type I errors). Likewise, managers can also be certain that, if they decide to comply with the regulation, they will not accidentally be found disobeying (no type II errors).

Because nothing in this model depends on past or future behavior, the problem can be solved as if it were a static cost minimization problem. The firm chooses its emissions e_{ti} in order to minimize the costs associated with the emission standard:

$$\min_{e_{ti}} \left\{ \theta_i C\left(e_{ti}\right) + pf\left(e_{ti}\right) \right\} \tag{2}$$

Thus the optimal strategy is to comply when the marginal expected penalty for non-compliance is larger than the marginal abatement costs savings of exceeding the standard; that is, when $pF'^{+}\left(\overline{e}\right) \geq -\theta_i C'\left(\overline{e}\right)$. In that case, the optimal strategy is $e_{ti} = \overline{e}$.[7] However, the optimal strategy is to exceed the standard if the marginal expected penalty is below the marginal abatement cost savings at the standard. In that case, we have $e_{ti} > \overline{e}$ and

$$-\theta_i C'\left(e_{ti}\right) = pF'\left(e_{ti}\right) \tag{3}$$

---

[6] Turning to the second order conditions, it is assumed that $pF'' > -\theta_i C''$ for all values of abatement. Thus firms face a strictly convex total cost function.

[7] In this perfect measurement model, the firm never chooses an emission level strictly below the standard: it just increases abatement costs, but there are no penalty savings.

This is the familiar result that, for an interior solution, marginal abatement costs equal the marginal expected fine.

Next, we define the cost parameter $\tilde{\theta}$ as the value for which a firm is indifferent between complying with the standard or exceeding it. The firm's abatement decisions are then described by:

$$\begin{array}{llll} If \;\; \theta_i \leq \tilde{\theta} & then & e_{ti} = \overline{e} & \\ if \;\; \theta_i > \tilde{\theta} & then & e_{ti} \;\; solves & -\theta_i C'(e_{ti}) = pF'(e_{ti}) \end{array} \tag{4}$$

Firms will obey the rules and emit exactly $\overline{e}$ if their abatement costs are sufficiently low. Otherwise they will violate the emission standard and the seriousness of the infraction depends on the level of their emission reduction costs. This is illustrated for two firms in figure 1. Firm 1 is a low-cost firm $\left(\theta_1 < \tilde{\theta}\right)$ and can relatively cheaply reduce its emissions. Consequently this firm will decide to comply with the emission standard. Firm 2 is however confronted with high abatement costs $\left(\theta_2 > \tilde{\theta}\right)$ and will find it beneficial to exceed the standard. Thus firm 2 weighs the marginal expected fine with its marginal abatement costs and selects emission level e_{t2}.

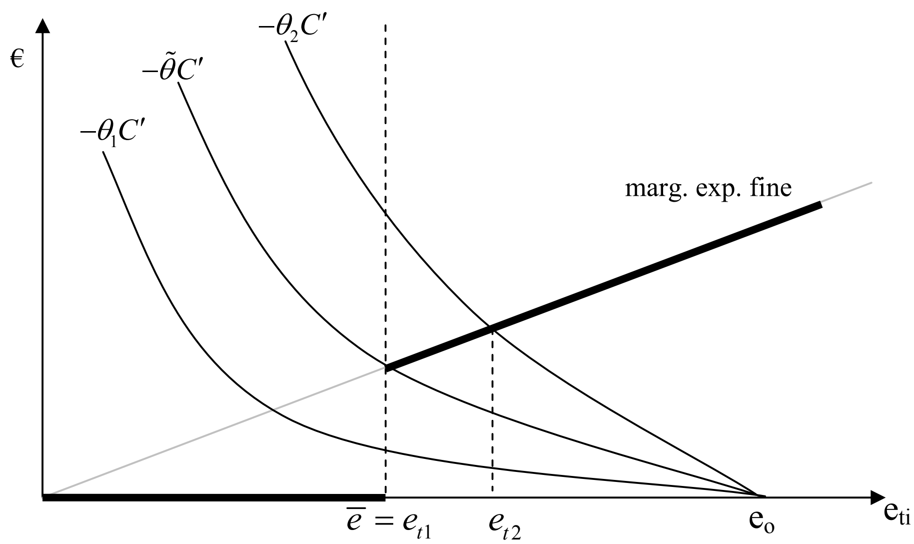


Figure 1: Abatement decision in period *t* with fines and under certainty

4. Model with measurement errors

Measurement errors are now introduced into the model. Firms again choose their actual emissions e_{ti} in order to minimize the total expected costs $E\left[TC_{ti}\right]$ associated with the environmental standard:

$$\min_{e_{ti}} E\left[TC_{ti}\right] = \min_{e_{ti}} \left\{\theta_i C\left(e_{ti}\right) + p\, E\left[f\left(e_{ti}\right)\right]\right\} \tag{5}$$

First, a regulatory scheme with fines as the only enforcement instrument available is discussed; the next section looks at what happens if both warnings and fines can be used.

4.1 Regulatory scheme with fines (scenario F)

Given the fact that inspections and sanctions are independent of past behavior, the problem in period *t* is independent of the compliance history and is quite straightforward. Firm *i* chooses the emission level e_{ti} in order to minimize its expected costs. An important modeling assumption is that abatement efforts are only valid for one period and thus abatement costs return each period. The types of abatement costs under consideration are operating and maintenance costs that are necessary to minimize the emissions associated with the production process.[8]

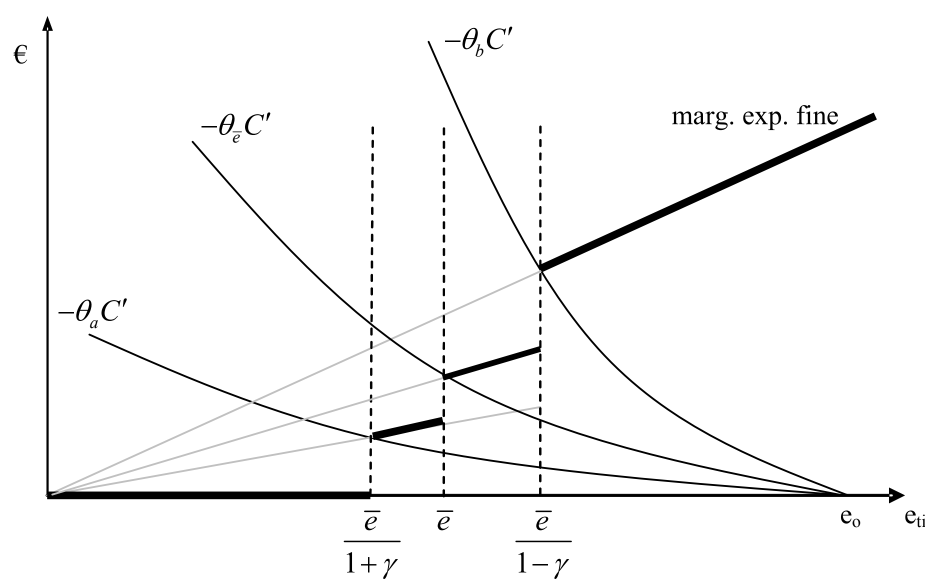


Figure 2: Abatement decision in period *t* with fines

The marginal expected fine faced by the firms as well as three possible marginal abatement costs curves are depicted in figure 2, where e_o are the baseline emissions if the firms do not abate their emissions. As is shown in figure 2, the marginal expected fine

[8] The implications of incorporating irreversible long-term abatement investments in the model are not directly clear (see, for instance, Zhao, 2002). On the one hand, firms would loose flexibility since they would not be able to easily revoke previously made abatement decisions. This would lead to higher compliance rates. On the other, abatement costs will be higher and the benefit of postponing the investment will also be higher.

curve can be divided into four regions.[9] If the firms' true emissions are below $\frac{\overline{e}}{1+\gamma}$, the expected fine is zero since the firms will never be fined. If the firms' emissions e_{ti} are between $\frac{\overline{e}}{1+\gamma}$ and $\overline{e}$, firms may be incorrectly penalized if a type II error occurs and the marginal expected fine then equals $a_1(e) \equiv pq\,F'(e(1+\gamma))$. If the firms' true emissions are between $\overline{e}$ and $\frac{\overline{e}}{1-\gamma}$, firms have a positive probability of being (correctly) fined and the marginal expected fine is $a_2(e) \equiv a_1(e) + p(1-2q)F'(e)$. Fourthly, if $e_{ti} > \frac{\overline{e}}{1-\gamma}$, the firms truly violate the standard and are always fined if inspected. The marginal expected fine in this region is equal to $a_3(e) = a_2(e) + pqF'(e(1-\gamma))$.

Due to the shape of the marginal expected fine curve, firms can be divided into four groups with similar behavior depending on their cost parameter. A firm with cost parameter θ_a finds it optimal to emit exactly $\frac{\overline{e}}{1+\gamma}$ since its marginal abatement costs are exactly equal to the marginal expected fine in that point. Analogously, it is optimal to emit exactly $\overline{e}$ for a firm with cost parameter $\theta_{\overline{e}}$ and for a firm with cost parameter θ_b to emit exactly $\frac{\overline{e}}{1-\gamma}$. Thus for firms with cost parameters below θ_a, it is optimal to emit $\frac{\overline{e}}{1+\gamma}$, since they will never be fined and additional emission reductions only generate extra costs and no benefits. Next firms with cost parameters above θ_a and below $\theta_{\overline{e}}$ will anticipate to be fined only when a type II error occurs $\left(e_{ti}^m = e_{ti}[1+\gamma]\right)$ and will then discharge e_i^a, defined by $-\theta_i C'(e_i^a) = a_1(e_i^a)$. Thirdly firms with cost parameters above $\theta_{\overline{e}}$ and below θ_b will anticipate to be fined when no error occurs and when a type II error occurs. Thus they will discharge $e_i^{\overline{e}}$, defined by $-\theta_i C'(e_i^{\overline{e}}) = a_2(e_i^{\overline{e}})$. Finally, firms will emit e_i^b, defined by $-\theta_i C'(e_i^b) = a_3(e_i^b)$, if they expect a positive fine no matter which type of measurement error occurs (i.e. firms with $\theta_i > \theta_\beta$). The mathematical definitions of these parameters can be found in appendix A. The firms' emission decisions when they are only confronted with a fine system are summarized in proposition 1:

[9] Note that we have $\overline{e}/1+\gamma \le \overline{e} \le \overline{e}/1-\gamma$ and that the difference between $\overline{e}$ and $\overline{e}/1+\gamma$ will always be smaller than the difference between $\overline{e}$ and $\overline{e}/1-\gamma$.

Proposition 1*: When firm i only faces fines (scenario F), its emissions e_{ti} in period t are determined as follows*

$$
\begin{array}{lll}
If\ \theta_i \le \theta_a & then & e_{ti} = \dfrac{\overline{e}}{1+\gamma} \\
If\ \theta_a < \theta_i \le \theta_{\overline{e}} & then & e_{ti} = min\left[e_i^a , \overline{e} \right] \\
If\ \theta_{\overline{e}} < \theta_i \le \theta_b & then & e_{ti} = min\left[e_i^{\overline{e}} , \dfrac{\overline{e}}{1-\gamma} \right] \\
If\ \theta_b < \theta_i & then & e_{ti} = e_i^b
\end{array}
$$

As mentioned in the introduction, Craswell and Calfee (1986) have shown that the presence of errors can lead to two effects: overcompliance and underdeterrence. On the one hand, overcompliance is observed when firms take excessive precautions in order to avoid mistakenly being found in fault. In our model low-cost firms $\left(\theta_i \le \theta_a\right)$ overcomply with the emission standard as well as some medium-cost firms $\left(\theta_a < \theta_i \le \theta_{\overline{e}}\right)$. These medium-cost firms are sometimes incorrectly accused of violating the emission standard (type II errors). On the other hand, underdeterrence takes place when violators' face a positive chance of not being punished thus reducing their incentives to comply. This is the case for the other medium-cost firms $\left(\theta_{\overline{e}} < \theta_i \le \theta_b\right)$ since these firms are sometimes incorrectly assumed to be in compliance when inspected due to type I errors. High-cost firms $\left(\theta_i > \theta_b\right)$ never comply. Finally, we can also note that more accurate monitoring technology (γ and q smaller) reduces both the level of overcompliance and the level of underdeterrence.

4.2 Regulatory scheme with fines and warnings (scenarios W1 and W2)

In the two schemes discussed in this section, the regulator can use warnings as well as fines in order to enforce the emission standard. A warning notifies the firm about a detected (small) violation but does not impose a punishment and is thus not a penalizing enforcement instrument. It does, however, slightly increase the expected fine in the next period for repeat offenders. Two specifications are considered: firstly, enforcement is independent of past compliance (scenario *W1*) and secondly, enforcement depends on past findings of non-compliance (scenario *W2*). We explicitly distinguish these two scenarios in order to clearly separate the effect of using warnings from the effect of having increasing penalties for repeat offenders. As we show later, scenario *W2* is nothing but a particular combination of scenarios *F* and *W1*.

In the first specification of the model, a warning is issued when the firm's measured emissions e_{ti}^{m} are between $\overline{e}$ and $\overline{e}\left[1+\gamma\right]$. If the measured emissions exceed $\overline{e}\left[1+\gamma\right]$, a fine is imposed[10]. So, only large offenders are fined[11] while small offenders get a warning. In the second specification, a warning is issued when the firm's measured emissions e_{ti}^{m} are between $\overline{e}$ and $\overline{e}\left[1+\gamma\right]$, *and* the firm was not given a warning or a fine in the previous period. A fine $F\left(e_{ti}^{m}\right)$ is directly applied for large detected violations, $e_{ti}^{m}>\overline{e}\left[1+\gamma\right]$, or for small violations of previously sanctioned firms, $\overline{e}<e_{ti}^{m}\leq\overline{e}\left[1+\gamma\right]$ and $e_{t-1,i}^{m}>\overline{e}$. The regulator implicitly assumes that a firm is truly violating the emission limit if two consecutive minor violations are detected. In this scenario, large and repeated violators are fined and only first-time small offenders are warned.[12]

4.2.1 State independent warnings (W1)

Under the first regulatory scheme *W1*, none of the low-cost firms $\left(\theta_i\leq\theta_{\overline{e}}\right)$ will overcomply due to the presence of warnings and they strictly implement the emission standard. After all, even if a type II error occurs, their measured emissions will never exceed $\overline{e}\left[1+\gamma\right]$ and thus there is no cost associated with their compliance decision since these firms will never be fined. Moreover, the use of warnings will lead to an increase in the number of medium-costs firms that will be underdeterred, since the expected penalty is reduced compared to the fine system due to the presence of type I errors. Firms emitting pollution levels up to $\dfrac{\overline{e}\left[1+\gamma\right]}{1-\gamma}$ are not fined – with a positive probability q – even if they are inspected and this decreases their incentives for reducing pollution levels. The cost parameter θ_β is defined in analogy to θ_a, $\theta_{\overline{e}}$ and θ_b (see figure 3). A firm with cost parameter θ_β will emit exactly $\dfrac{\overline{e}\left[1+\gamma\right]}{1-\gamma}$. The different parameters can be ranked as follows: $\theta_a<\theta_{\overline{e}}<\theta_b<\theta_\beta$. High-cost firms $\left(\theta_i>\theta_\beta\right)$ are certain to be fined and the expected fine is equal to the one imposed in the fine system without warnings. Thus these firms select an emission level e_i^b, which we previously defined as solving the equation $-\theta_i C'\left(e_i^b\right)=a_3\left(e_i^b\right)$. Again we see that a more accurate monitoring technology (with lower

[10] This warning-fine system *W1* is thus identical to a fine system *(F)* with an emission standard equal to $\overline{e}\left[1+\gamma\right]$ rather than $\overline{e}$. It is, however, still interesting to discuss this case as it presents an extreme case when warnings are used next to fines.

[11] Note that this fine can be corrected for the presence of errors by subtracting a constant reflecting the expected size of the error from the calculated fine. This would not influence our results.

[12] The model could easily be extended to allow warnings to be used during *n* periods (rather than only one period) as long as the violations are small. However, in the end, if the (small) violation continues period after period, it becomes a serious violation and the inspection agency will fine the violator. Including this would complicate the mathematics without adding extra insight into the problem.

γ) would lead to fewer firms that are underdeterred, since the expression $\frac{\bar{e}[1+\gamma]}{1-\gamma}$ would decrease and the gap with the emission standard would decline.

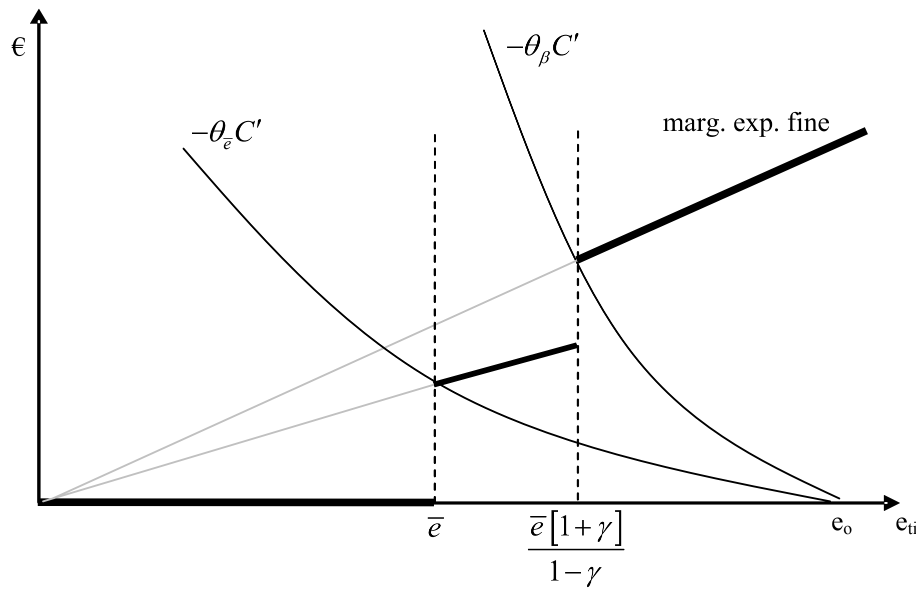


Figure 3: Abatement decisions in period *t* under warning-fine system W1

The emission levels chosen by firms facing a warning-fine system defined as W1 are summarized in proposition 2.

Proposition 2: *When firm i faces fines and state independent warnings (scenario W1), its emissions e_{ti} in period t are determined as follows*

$$
\begin{array}{lll}
\text{If } \theta_i \le \theta_{\bar{e}} & \text{then} & e_{ti} = \bar{e} \\
\text{If } \theta_{\bar{e}} < \theta_i \le \theta_\beta & \text{then} & e_{ti} = \min\left[e_i^{\bar{e}}, \frac{\bar{e}[1+\gamma]}{1-\gamma}\right] \\
\text{If } \theta_\beta < \theta_i & \text{then} & e_{ti} = e_i^b
\end{array} \qquad (6)
$$

If the firm has low abatement costs, $\theta_i \le \theta_{\bar{e}}$, it complies exactly with the emission standard. Medium-cost firms, $\theta_{\bar{e}} < \theta_i \le \theta_\beta$, choose their emission levels such that they have a lower probability $(q<1)$ to incur a fine than high-cost firms with $\theta_i > \theta_\beta$. The use of warnings in this setting implies that the overcompliance effect is now completely absent, while the underdeterrence effect increases. This implies that no compliant firms will ever be erroneously fined, but several small violators will only be warned and not fined even if they are repeat violators.

4.2.2 State dependent warnings (W2)

The second specification of a warning-fine enforcement scheme *W2* implies that firms receive a warning (or a fine-amnesty) when the measured emissions are between $\overline{e}$ and $\overline{e}\left[1+\gamma\right]$, *and* the firm was not given a warning or a fine in the previous period. A fine is directly applied for large detected violations, $e_{ti}^{m} > \overline{e}\left[1+\gamma\right]$, or for small violations of previously sanctioned firms (i.e. the fine-amnesty is revoked). The scenario *W2* is hence a combination of the fine system *F* and the warning-fine system *W1*. Under this system, the abatement decision of the firm depends on its compliance history and we distinguish two different possibilities: 1) the firm was not fined or warned in the previous period and 2) the firm was fined or warned in the preceding period. Next these two situations are discussed.

Firstly, when a firm was not fined or warned in period *t-1*, its emission decision in period *t* can be described by proposition 2. These firms have not lost the right of receiving a fine-amnesty if a small violation should be detected. Thus for these firms there is no difference between system *W1* and system *W2* in period *t*.

Secondly, when a firm was warned or fined in the previous period, it will make the same emission decision in the current period irrespective of whether they face fines (system *F*) or fines and warnings combined (system *W2*). Its emissions decisions can therefore be described by proposition 1. After all, if the firm exceeded the emission limit in the previous period, the environmental inspection agency will not use warnings in the current period, not even for small abuses.

Finally, in order to determine total expected emission decisions in each period, we need to know the distribution of firms over the two possible scenarios and we have to determine the probability that a firm was fined or warned in the previous period. For this reason, we derive the steady state probabilities (also called equilibrium probabilities) for each firm. These probabilities can be used to describe the long-run[13] behavior of our model. In steady state, the 'flow' of probability into each state must equal to the flow of probability out of each state (Winston, 1994). For example, in steady state exactly the same number of firms enters the state 'received a warning in the previous period' as the number of firms that leave this state (and enter a different state).

Next, we use the following notation: the equilibrium probability that a firm received a warning in the previous period equals $P^{W}\left(\theta_{i}\right)$; the equilibrium probability that the firm paid a fine in the previous period is denoted by $P^{F}\left(\theta_{i}\right)$ and the equilibrium probability that the firm was not fined or warned in the previous period is denoted by $P^{N}\left(\theta_{i}\right)$. Note that

[13] Behavior before the steady state is reached is called short-run or transient behavior. For large *t*, the steady-state probabilities accurately describe the probability of being in any state.

each firm is in one and only one state in each period and thus $P^W(\theta_i)+P^F(\theta_i)+P^N(\theta_i)=1$. Further, the conditions under which the different types of firms are in each particular state are described in table 1.

Table 1: Conditions determining the probability of being in each state in period t

	Firm received a warning in period t-1	*Firm received a fine in period t-1*	*Firm was not fined or warned in period t-1*
1. $\theta_i > \theta_\beta$	Never	Inspected in t-1	Not inspected in t-1
2. $\theta_b < \theta_i \le \theta_\beta$	Inspected in t-1 with type I error and no fine or warning in t-2	a. Inspected in t-1 without type I error b. Inspected in t-1 with type I error and fine or warning in t-2	Not inspected in t-1
3. $\theta_{\bar{e}} < \theta_i \le \theta_b$	Inspected in t-1 without error and no fine or warning in t-2	a. Inspected in t-1 with type II error b. Inspected in t-1 without error and fine or warning in t-2	a. Not inspected in t-1 b. Inspected in t-1 with type I error
4. $\theta_a < \theta_i \le \theta_{\bar{e}}$	Inspected in t-1 with type II error and no fine or warning in t-2	Inspected in t-1 with type II error and fine or warning in t-2	a. Not inspected in t-1 b. Inspected in t-1 without type II error
5. $\theta_i \le \theta_a$	Inspected in t-1 with type II error and no warning received in period t-2	Never	a. Not inspected in t-1 b. Inspected in t-1 without type II error c. Inspected in t-1 with type II error and warning received in period t-2

The expected emissions for each firm in scenario *W2* can then described in the following proposition:

Proposition 3: *When firm i faces fines and state dependent warnings (scenario W2), its expected emissions* $E(e_{ti})$ *in period t are determined as followed*

(i) *If* $\theta_i > \theta_\beta$, *then* $E(e_{ti}) = e_i^b$

(ii) If $\theta_b < \theta_i \le \theta_\beta$, then $E(e_{ti}) = \left[1 - P^N(\theta_i)\right] e_i^b + P^N(\theta) \min\left[e_i^{\bar{e}}, \frac{\bar{e}\left[1+\gamma\right]}{1-\gamma}\right]$

(iii) If $\theta_{\bar{e}} < \theta_i \le \theta_b$, then $E(e_{ti}) = \left[1 - P^N(\theta_i)\right] \min\left[e_i^{\bar{e}}, \frac{\bar{e}}{1-\gamma}\right] + P^N(\theta_i) \min\left[e_i^{\bar{e}}, \frac{\bar{e}\left[1+\gamma\right]}{1-\gamma}\right]$

(iv) If $\theta_a < \theta_i \le \theta_{\bar{e}}$, then $E(e_{ti}) = \left[1 - P^N(\theta_i)\right] \min\left[e_i^a, \bar{e}\right] + P^N(\theta_i)\, \bar{e}$

(v) If $\theta_i \le \theta_a$, then $E(e_{ti}) = P^W(\theta_i)\left[\frac{\bar{e}}{1+\gamma}\right] + \left[1 - P^W(\theta_i)\right]\bar{e}$.

The proof of this proposition and the values of the steady state probabilities can be found in appendix B.

Next we discuss the compliance behavior of the firms as described in proposition 3. Firstly, since the high-cost firms (case *i* of proposition 3) are always in violation, they are always found in violation when inspected and consequently they always have to pay a fine. This implies that these firms will always emit the same amount of pollution irrespective of whether they had to pay a fine the previous period or not. Secondly, in case *ii* of proposition 3, all firms are exceeding the emission standard. Still, some of them will not be fined and receive only a warning if they are inspected and a type I error occurs. However, none of these firms will ever be found to be compliant, not even erroneously. Thirdly, all medium-cost firms falling under case *iii* of proposition 3 are in violation but they have a positive probability of being found in compliance. When these violators are caught, they get warned unless a type II error occurs during audit or no errors occurs but they were repeat offenders. Further the medium-cost firms falling under case *iv* of proposition 3 are all compliant. Nonetheless, these firms face a positive chance of incorrectly receiving a warning or even a fine. A fine is imposed only after the occurrence of two consecutive type II errors, while a warning follows a type II error that was not preceded by a type II error in the previous period. Finally, the low-cost firms (case *v* of proposition 3) will always comply since their emissions will never exceed the emission standard. Even so they are sometimes incorrectly warned that they have committed a minor violation. However, there will never be any wrongly fined firms in this group. After all, a firm that received a warning will emit $\frac{\bar{e}}{1+\gamma}$ in the next period and will thus do better than legally required, in order to avoid the possibility of being fined for a small violation should a type II error occur in the next period.

5. Comparing the different policy scenarios

This section compares three regulatory scenarios: 1) Scenario *F*: every detected violation is fined, 2) Scenario *W1*: detected large violations are fined and detected small violators are warned, and 3) Scenario *W2*: large and repeated violators are fined and small first-time offenders are warned.[14] We discuss the impact of the penalty schemes in the presence of errors on firms' emission levels, on the prevalence of incorrect convictions and erroneously acquittals as well as the effect on social welfare. In order to compare the total impact on social welfare imposed by each regulatory scheme, different counteracting effects need to be weighed against each other. Social welfare takes the net social benefits from the changes in emissions levels into account as well as fairness and justice considerations. The social welfare function[15] V thus depends on the level of emissions $\left(E=\sum_i e_i\right)$, the number of incorrectly fined firms *(I)* and the amount of acquitted violators *(II)* associated with each enforcement scenario*:* $V(E,I,II)$. Welfare levels also depend on the distribution of the firms with respect to the cost parameter: the impact of the different scenarios will depend on the proportion of high-cost, low-cost and medium-cost firms in the industry. In the remainder of the analysis we assume that all cost types are represented in the industry.[16]

If we adopt a classical utilitarian approach, we can sum the different effects on social welfare. A slight generalization is allowing for a weighted sum of the social welfare effects. The impact of an enforcement policy in social welfare (ΔV) is then equal to the weighted sum of the social cost of the change in emissions $SC(\Delta E)$, the social cost associated with the change in the number of type I errors $SC(\Delta I)$ and the social cost associated with the change in type II errors $SC(\Delta II)$:

$$\Delta V=\lambda_1 SC(\Delta E)+\lambda_2 SC(\Delta I)+\lambda_3 SC(\Delta II) \tag{7}$$

The weights λ depend on the importance attached to each welfare aspect by society.

In order to better understand the impact of warnings on these three components, we first look at the impact on emissions levels and the extent of the regulatory errors. Next we discuss the social valuation of these effects and finally we compare the fine system with the state-dependent warning-fine system.

5.1 Impact on emission levels and errors

The three policy options each have a different impact on the total emission levels and the amount of erroneous acquittals (type I) and erroneous convictions (type II). These different effects are summarized in table 2. The formal derivation of this table can be found in appendix C.

[14] The results are derived for the steady state. For the scenarios *F* and *W1*, these steady state results are equal to the static results. Still, modeling the state-dependent warning-fine system (*W2*) needs a dynamic setting.
[15] Note that fines are modeled as costless transfers.
[16] For example, if there are only high-cost firms in the industry, the impact of the three systems will be identical.

Table 2: Effects of the enforcement schemes

	Fine ***F***	*State-independent warning-fine* ***W1***	*State-dependent warning-fine* ***W2***
Emission levels	Lowest	Highest	Medium
Number of erroneous acquittals	Lowest	Highest	Medium
Number of erroneous convictions	Highest	Lowest	Medium

First, the effect of the three enforcement strategies on total emission levels can be derived from the propositions 1, 2 and 3. The fine system always leads to the lowest emission levels since it imposes the highest sanction on firms found to be in violation. The state-independent warning system *(W1)*, on the other hand, implies the highest emission levels because it is less stringent for suspected violators and small violators are not sanctioned. Also, since the state-dependent warning system *(W2)* is a combination of two other systems, it brings about intermediate emission levels.

Secondly, the influence of the enforcement scheme on the occurrence of type I errors is such that, under a fine system, fewer violators incorrectly escape a sanction than under a warning-fine system. The number of erroneous acquittals is thus lowest under the fine system since all detected violators are fined, while under the warning-fine systems some of the detected violators are warned rather than fined. Hence, the W1-system leads to the highest level of erroneous acquittals while the state-dependent warning-fine system *W2* leads to intermediate levels.

Finally, we discuss the number of erroneous convictions. These will be highest under the fine system since all detected violators are fined, while under the warning-fine systems some of the detected violators are warned rather than fined. The W1-system thus leads to the lowest level (i.e. zero) of erroneous convictions and again the W2-system leads to an intermediate level. Hence, the use of state-independent warnings implies fewer erroneous convictions than using state-dependent warnings.

5.2 Social valuation

In this section we discuss the social valuation of emissions and emission reductions as well as the social valuation of regulatory errors.

5.2.1 Social valuation of emissions

We can value emissions at their net marginal social costs, which equal the difference between the marginal cost of an extra emission reduction and the marginal social benefit of reduced environmental harm. Thus this marginal social cost of emissions is positive if the additional abatement required for an extra unit of emission reduction is socially costly and does not outweigh the associated environmental benefit. The net marginal social cost is negative[17] if the reduction in emissions is socially beneficial and the marginal abatement costs are lower than the related environmental gains.

If we assume that the emission standard $\overline{e}$ is set such that the marginal cost of an extra unit emission reduction equals the marginal social benefit of the increase in the environmental quality, then the marginal social costs will be positive for emission reductions below the standard since these extra reductions are socially detrimental, and the marginal social cost will be negative for emission levels exceeding the emission limit. For 'small' values of the error sizes (see appendix D), the net social value of increasing emission levels by using warnings will be positive due to the convexity of the individual and aggregate abatement cost functions and the assumption of constant marginal benefits[18] (see figure 4) . Thus, under these circumstances, the increase in emissions associated with using a warning-fine system *W2* is welfare improving compared to only imposing fines (see appendix D). In the presence of large errors sizes, the regulator might consider investing in improving monitoring accuracy rather than trying to correct the effects associated with the errors by using warnings. However, this falls outside the scope of the current analysis.

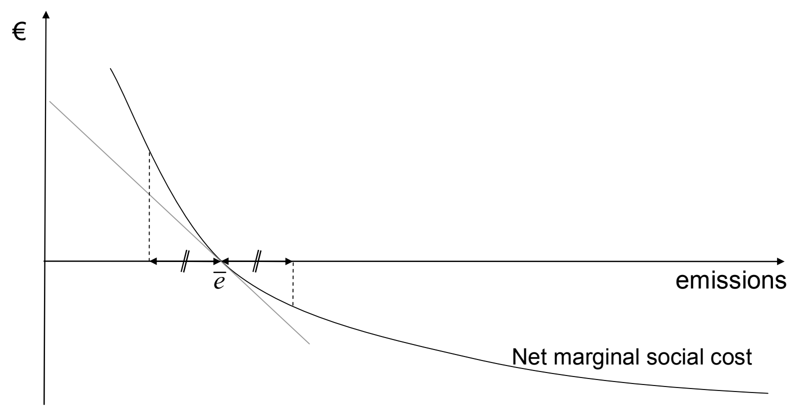


Figure 4: Net marginal social cost curve

5.2.2 Social valuation of errors

Fairness is a social concern and this implies that there is a social cost associated with sanctioning the innocent as well as letting the guilty go unpunished. Generally, it might be

[17] A negative cost thus equals a benefit.
[18] The assumption of constant marginal benefits is only used for the derivation of the threshold level of error sizes. The other findings in the paper do not depend on this assumption. For non-constant marginal benefits the threshold is less clear-cut to define. However, the main finding, that for 'sufficiently small' error sizes warnings will be socially beneficial, continues to hold.

better to acquit guilty individuals than convicting an innocent person, since society places a higher weight on the latter and also because type II errors seem to have substantially higher costs. Indeed, society considers conviction of the innocent to be much worse than acquittal of the guilty (Dworkin, 2002, Chu et al., 2000, Miceli and Segerson, 2007 and Persson and Siven, 2007). As a case in point, the standard of proof used in criminal cases reflects the importance that society attaches to the danger of convicting the innocent. Scholars have long contended that it is better to acquit five, twenty or even a thousand guilty persons than to convict one innocent person (e.g. Blackstone, 1769 and Kitai, 2003). Thus the social weight attached to erroneous convictions seems to be higher than that associated with erroneous acquittals (i.e. $\lambda_3 > \lambda_2$ in expression (7) of the social welfare function).

Next we discuss the determinants of the social costs associated with convictions. According to Dworkin (2002) and Kitai (2003), every punishment, whether justified or not, brings about 'a bare harm' to the convicted individual. This bare harm cost encompasses suffering, frustration, pain and an inability to fulfill one's aspirations due to the loss of freedom. However, when an innocent party is convicted, the injustice brings an additional suffering, which Dworkin (2002) termed 'the injustice factor'. This injustice factor is an objective notion of the inherent evil of a mistaken conviction and is independent of the degree of bare harm and of the question whether anyone is conscious of the injustice. Also, erroneous convictions can spread uncertainty or fear of being wrongly convicted among the law-abiding population. Both types of errors, in general, have a corrosive effect on the legitimacy of the criminal justice system and undermine the public's confidence in police, government, prosecutors and the courts.

5.3 Comparison of system F with W2

We now compare the size of these impacts for the fine policy and the state-dependent warning-fine enforcement strategy *(F versus W2)*. The latter warning-fine system is, after all, the prevailing one in real-world enforcement policy situations. We have already mentioned that if the measurement errors are small, the net social value of increasing emission levels by using warnings will be positive. Thus, the increase in emissions associated with using a warning-fine system *W2* is welfare improving compared to only imposing fines. From the discussion on the social valuation of errors, it follows that there are compelling reasons to assume that the social benefit of decreasing erroneous convictions by using warnings is much higher than the social cost associated with the rise in erroneous acquittals. This implies that warnings are socially beneficial to use.

Comparing the two different specifications of the warning-fine system shows that no policy scenario is strictly better than the other. However, the warning-fine scheme without influence of past compliance behavior completely removes the likelihood of accidentally sanctioning innocent firms. Thus, the policy scenario *W1* is more likely to outperform the state-dependent warning-fine system *W2,* as the weight of erroneous convictions on social welfare is likely to be higher than that of erroneous acquittals.

6. Conclusion

This article has shown that the prevalent role of warnings in the daily policy of an inspection agency can be explained as a corrective measure when there is an imperfect match between a firm's emission decision and the agency's measured emissions. Introducing the use of warnings as an enforcement instrument can thus increase social welfare when the results of emission measurements are uncertain. This uncertainty implies that abatement decisions taken by the firm are not fully reflected in the level of emissions measured by the inspection agency. Measured emissions can be biased if, for example, the agency's measuring equipment does not work perfectly or as a result of human errors. In such a situation, warnings can, firstly, be used as a means to reduce the environmental and social consequences of these errors. Even though the use of warnings creates some underdeterrence of medium-cost firms, such a system reduces the overcompliance of low-cost firms caused by the difference between actual and measured emissions. Secondly, warnings reduce the number of incorrect prosecutions in the case of measurement errors, which is also welfare enhancing.

In this contribution we use a constant inspection frequency. However, theoretical (see Harrington, 1988) and empirical literature indicates that an increase in inspection probability after the environmental inspection agency has found firms in violation is very plausible. As Eckert (2004) and Rousseau (2007) show, warnings as well as fines are used by environmental agencies to target firms for inspections. Here this would imply that, if a firm was previously caught violating the emission standard, it would be more frequently inspected and thus their expected penalty increases. Hence the fine system will also be dependent on the firm's compliance history. If firms have a past record of non-compliance, they will lower their emissions and more firms will comply because they will be inspected more frequently. This is the classical result established by Becker (1968); that an increase in expected penalty will reduce the level of crime. In this model, higher inspections of repeat offenders will lead to more overcompliance by firms and since warnings reduce this overcompliance effect, their case is strengthened. However, if we assume that the monitoring resources are fixed, the increase in inspections for one group of firms has to be compensated by a lower inspection frequency for the other firms. Thus the firms with a good compliance history would face a lower expected penalty and these firms would therefore have an incentive to increase their emission levels. The total effect of targeting repeat offenders through increased inspections is thus ambiguous and will depend on the specific assumptions made.

The results of this study show theoretically that it is important to consider informal as well as formal enforcement instruments. Both types of instruments can not only be used as substitutes but also, and more importantly, as complements. Moreover, the results presented here cast some doubts on the appropriateness of the monitoring strategy used in many western countries that employs warnings only for first-time small offenders. It might be worthwhile to consider switching to a system that uses warnings for *all* small offenders irrespective of their compliance history as long as the size of the existing measurement errors is within reasonable bounds. After all, such a system – if correctly designed – removes all erroneous convictions and the small violations will only have small associated environmental costs.

APPENDIX

A. Definition of parameters used in proposition 1

Due to the shape of the marginal expected fine curve, firms can be divided into four groups with similar behavior depending on their cost parameter. The definitions of e_i^a and θ_a are subsequently given, where $^+$ denotes the right hand derivative:

$$-\theta_a C'\left(\frac{\overline{e}}{1+\gamma}\right) = pq\,F'^{+}\left(\overline{e}\right) \text{ and } -\theta_i C'\left(e_i^a\right) = a_1\left(e_i^a\right)$$

The threshold θ_a is the cost parameter of the firm for which it is optimal to emit exactly $\frac{\overline{e}}{1+\gamma}$. Further, firms will discharge e_i^a, if they anticipate to be fined only when a type II error occurs and $e_{ti}^m = e_{ti}\left(1+\gamma\right)$. Further we also define $e_i^{\overline{e}}$ and $\theta_{\overline{e}}$, with $\theta_a < \theta_{\overline{e}}$:

$$-\theta_{\overline{e}} C'\left(\overline{e}\right) = p\left(1-2q\right)F'^{+}\left(\overline{e}\right) + pq\,F'\left(\overline{e}\left(1+\gamma\right)\right) \text{ and } -\theta_i C'\left(e_i^{\overline{e}}\right) = a_2\left(e_i^{\overline{e}}\right)$$

The values e_i^b and θ_b, with $\theta_a < \theta_{\overline{e}} < \theta_b$, are defined by:

$$-\theta_b C'\left(\frac{\overline{e}}{1-\gamma}\right) = pq\,F'^{+}\left(\overline{e}\right) + p\left(1-2q\right)F'\left(\frac{\overline{e}}{1-\gamma}\right) + pq\,F'\left(\frac{\overline{e}\left(1+\gamma\right)}{1-\gamma}\right) \text{ and}$$

$$-\theta_i C'\left(e_i^b\right) = a_3\left(e_i^b\right)$$

Firms will emit e_i^b, if they expect a positive fine no matter which type of error occurs.

B. Proof proposition 3

We use the following notation: $P^W\left(\theta_i\right)$ = equilibrium probability that firm *i* received a warning in the previous period; $P^F\left(\theta_i\right)$ = equilibrium probability that firm *i* received a fine in the previous period and $P^N\left(\theta_i\right)$ = equilibrium probability that firm *i* received no fine or warning in the previous period. So we categorize firms into three groups (three states): 1) received a warning in previous period, 2) received a fine in previous period and 3) received no warning or fine in previous period. Since this division is exhaustive, the following condition is always satisfied: $P^W\left(\theta_i\right) + P^F\left(\theta_i\right) + P^N\left(\theta_i\right) = 1$. The expected emissions and the steady state probabilities of firm *i* in period *t* (with *t* sufficiently large such that the steady state is reached) can be described for each category depending on the parameter θ_i.

<u>*1. For* $\theta_i > \theta_\beta$</u>

In equilibrium, the expected emissions of these high-cost firms equal: $E(e_{ti}) = e_i^b$. If these high-costs firms are inspected, they are always found in violation and their emissions always exceed $\frac{\bar{e}}{1-\gamma}$. Consequently, if inspected, these firms always have to pay a fine. Thus: $P^F = p$ and $P^N = 1-p$.

2. For $\theta_b < \theta_i \leq \theta_\beta$

Firms in this group received no warning or fine in period *t-1* if they were not inspected. Thus: $P^N(\theta_i) = 1-p$. Firms in this group received a warning in period *t-1,* if they were inspected in period *t-1* and a type I error occurred (pq), on condition that they did not receive a warning or fine in period *t-2*. Thus the steady-state probability that such a firm received a warning in the previous period equals $P^W(\theta_i) = pq\,P^N(\theta_i) = pq[1-p]$. Thirdly, firms were fined in period *t-1,* a) if they were inspected in that period and no type I error occurred, or b) if a type I error occurred during measurement and they were warned or fined in the previous period *t-2*. The equilibrium probability that the firm paid a fine in the previous period is equal to: $P^F(\theta_i) = p[1-q] + pq\left[1-P^N(\theta_i)\right] = p[1-q] + pqp$.

The expected emissions of these firms are therefore equal to:

$$E(e_{ti}) = \left[P^W(\theta_i) + P^F(\theta_i)\right]e_i^b + \left[1-P^W(\theta_i) - P^F(\theta_i)\right]\min\left[e_i^{\bar{e}}, \frac{\bar{e}[1+\gamma]}{1-\gamma}\right]$$

All these firms are exceeding the emission standard. Still, some of them will escape being fined, if they are inspected, and will receive only a warning if a type I error occurs. However, none of these firms will ever be found to be compliant, not even erroneously.

3. For $\theta_{\bar{e}} < \theta_i \leq \theta_b$

All these medium-cost firms are in violation but they have a positive probability of being found in compliance. Within this group of firms, the expected emissions equal:

$$E(e_{ti}) = \left[P^W(\theta_i) + P^F(\theta_i)\right]\min\left[e_i^{\bar{e}}, \frac{\bar{e}}{1-\gamma}\right] + \left[1-P^W(\theta_i) - P^F(\theta_i)\right]\min\left(e_i^{\bar{e}}, \frac{\bar{e}[1+\gamma]}{1-\gamma}\right)$$

These firms were not fined or warned in period t-1 a) if they were not inspected in t-1 or b) if they were inspected but a type I error occurred: $P^N(\theta_i) = 1-p+pq$. Secondly, these firms were sent a warning in period *t-1,* if they were inspected and no error occurred, as long as they were not fined nor warned in period *t-2*: $P^W(\theta_i) = p[1-2q]P^N(\theta_i)$ $= p[1-2q][1-p+pq]$. Thirdly, firms are fined in period *t-1* a) if a type II error occurs during audit or b) if their emissions are measured accurately and they were fined or warned in period *t-2*: $P^F(\theta_i) = pq + p[1-2q]\left[1-P^N\right] = pq + p[1-2q]p[1-q]$.

4. For $\theta_a < \theta_i \leq \theta_{\bar{e}}$

All these firms are compliant and have a positive chance of incorrectly receiving a warning or even a fine. For these firms, expected emissions are:

$$E(e_{ti}) = \left[P^W(\theta_i) + P^F(\theta_i)\right] \min\left[e_i^a, \overline{e}\right] + \left[1 - P^W(\theta_i) - P^F(\theta_i)\right] \overline{e}$$

Firstly, these firms did not receive a fine or warning in period t-1, a) if they were not inspected or b) if they were inspected but no type II error occurred: $P^N(\theta_i) = 1 - p + p[1-q] = 1 - pq$. Secondly, these medium-cost firms were warned in period *t-1*, if a type II error occurred during inspection and they did not receive a warning or fine in period *t-2*: $P^W(\theta_i) = pq\, P^N(\theta_i) = pq[1-pq]$. Finally, the firm received a fine in period *t-1*, if a type II error occurred in t-1 and the firm received a fine or warning in *t-2*: $P^F(\theta_i) = pq\left[1 - P^N(\theta_i)\right] = pqpq$.

<u>5. For $\theta_i \le \theta_a$</u>

Expected emissions of these low-cost firms are: $E(e_{ti}) = P^W(\theta_i)\left[\frac{\overline{e}}{1+\gamma}\right] + \left[1 - P^W(\theta_i)\right]\left[\overline{e}\right]$.

These low-cost firms will always comply since their emissions will never exceed the emission standard. There will never be any wrongly fined firms in this group: $P^F(\theta_i) = 0$.

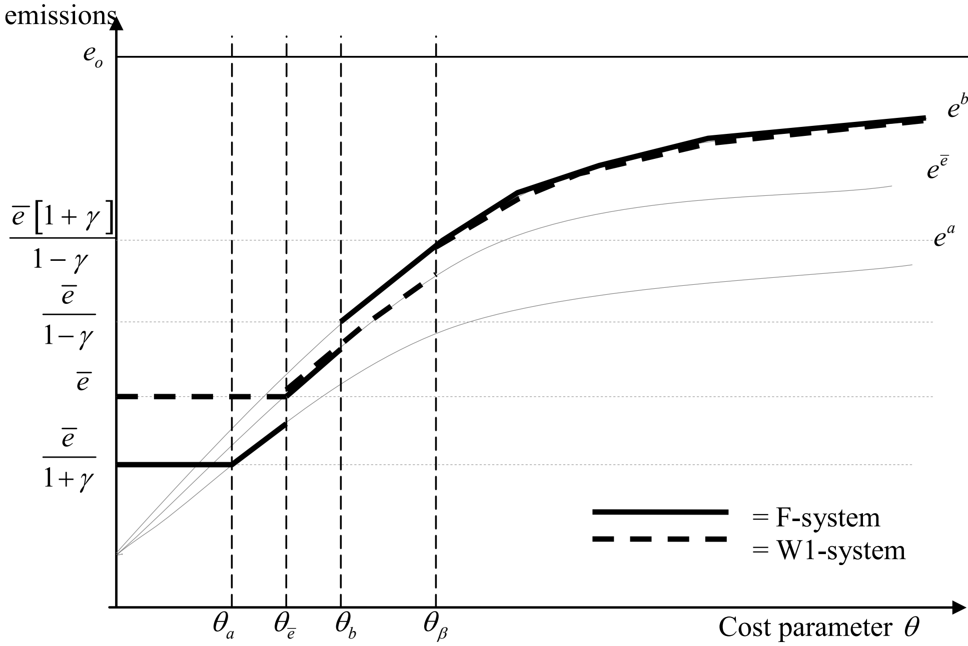
However, these firms were warned in period *t-1*, if a type II error occurred during inspection and they were not warned in period *t-2*: $P^W = pq\, P^N$. The firms were not warned (nor fined) in period *t-1*, if a) no inspection took place, b) they were inspected and no type II error occurred or c) they were inspected, a type II error occurred and they were warned in period *t-2*: $P^N(\theta_i) = 1 - p + p[1-q] + pq\, P^W(\theta_i)$. Using $P^W(\theta_i) + P^F(\theta_i) + P^N(\theta_i) = 1$, we know that $P^W(\theta_i) + 1 - p + p[1-q] + pq\, P^W(\theta_i) = 1$ and thus $P^W(\theta_i) = \frac{pq}{1+pq}$ and $P^N(\theta_i) = 1 - p + p[1-q] + pq\frac{pq}{1+pq}$.

In summary, we have:

	$P^W(\theta_i)$	$P^F(\theta_i)$	$P^N(\theta_i)$
1. $\theta_i > \theta_\beta$	0	p	1-p
2. $\theta_b < \theta_i \le \theta_\beta$	pq (1-p)	pq p + p(1-q)	1-p
3. $\theta_{\overline{e}} < \theta_i \le \theta_b$	p(1-2q) (1-p) + p(1-2q) pq	p(1-2q) p(1-q) + pq	1-p + pq
4. $\theta_a < \theta_i \le \theta_{\overline{e}}$	pq p(1-q) + pq (1-p)	pq pq	1-p + p(1-q)
5. $\theta_i \le \theta_a$	pq / (1 + pq)	0	1-p + p(1-q) + pq pq / (1 + pq)

C. Derivation of table 2

The emissions levels associated with the three different systems follow directly from propositions 1, 2 and 3. A graphical representation of the emissions associated with the fine system and the warning-fine system W1 is given below. Since the W2-system is a combination of W1 and F, its expected emission levels are weighted averages of the levels associated with the two other systems.



The W1-system will always imply higher emission levels than the F-system for the firms with cost parameters below $\theta_{\bar{e}}$ and lower emission levels than for the firms with cost parameters between θ_b *and* θ_β. Due to the definition of the θ-parameters, overall emission levels will always be higher under the W1-system than under the F-system.

We now look at the number of erroneous acquittals. These are lowest under the F-system since all detected violators are fined, while under the warning-fine system some of the detected violators are warned rather than fined. The W1-system will thus lead to the highest level of erroneous acquittals. The number of erroneous convictions is discussed next. These will be highest under the F-system since all detected violators are fined, while under the warning-fine systems some of the detected violators are warned rather than fined. The W1-system will thus lead to the lowest (i.e. zero) level of erroneous convictions and the W2-system leads to intermediate levels.

We now compare the W2-system versus F-system more formally. The difference in emissions induced by the warning-fine system W2 and those brought about by the fine system can be expressed as (the emissions in scenario *F* are subtracted from those associated with scenario *W2*):

$$
\begin{aligned}
\Delta E(\theta) \equiv & \int_{0}^{\theta_a} \left[1-P^W(\theta)\right]\left[\bar{e}-\frac{\bar{e}}{1+\gamma}\right]d\theta \\
& +\int_{\theta_a}^{\theta_{\bar{e}}} \left[1-P^W(\theta)-P^F(\theta)\right]\left[\bar{e}-\min\left(e_i^a(\theta),\bar{e}\right)\right]d\theta \\
& +\int_{\theta_{\bar{e}}}^{\theta_b} \left[1-P^W(\theta)-P^F(\theta)\right]\left[\min\left(e_i^{\bar{e}}(\theta),\frac{\bar{e}[1+\gamma]}{1-\gamma}\right)-\min\left(e_i^{\bar{e}}(\theta),\frac{\bar{e}}{1-\gamma}\right)\right]d\theta \\
& +\int_{\theta_b}^{\theta_\beta} \left[1-P^W(\theta)-P^F(\theta)\right]\left[\min\left(e_i^{\bar{e}}(\theta),\frac{\bar{e}[1+\gamma]}{1-\gamma}\right)-e_i^b(\theta)\right]d\theta
\end{aligned} \tag{8}
$$

This expression is always positive and total emissions levels are thus lower under scenario *F* than under scenario *W2*. The increase in the number of type I errors associated with the use of warnings equals: $\Delta I(\theta) \equiv \int_{\theta_b}^{\theta_\beta} pq_1\left[1-P_{t-1}^W(\theta)-P_{t-1}^F(\theta)\right]d\theta$. Then again, using warnings implies fewer type II errors and this reduction in errors is: $\Delta II(\theta) \equiv \int_{\theta_a}^{\theta_{\bar{e}}} pq_2\left[1-P_{t-1}^W(\theta)-P_{t-1}^F(\theta)\right]d\theta + \int_{\theta_{\bar{e}}}^{\theta_b} pq_2\, d\theta$.

D. Sufficient condition for warnings to be socially beneficial

We use the following notation for the regulatory scheme: $m \in \{F, W1, W2\}$. Thus we have the level of emissions $\left(\sum_i e_i(m)\right)$, the number of erroneous acquittals $\left(I(m)\right)$ and the amount of erroneous convictions $\left(II(m)\right)$ associated with each enforcement scenario *m*. The net marginal social costs (MSC), which equal the difference between the marginal cost of an extra emission reduction and the marginal social benefit:

$$MSC(e) = MAC(e) - MB(e)$$

with *MAC* the aggregate marginal abatement cost function derived from the individual abatement cost functions $\theta_i C$ and *MB* representing society's (constant) willingness-to-pay for an additional reduction in emissions. Now we can derive sufficient conditions to have a positive net social benefit - i.e. a negative net social cost - of the change in emissions associated with using warnings:

$$\sum_i \left[MSC\left(e_i(F)\right)e_i(F) - MSC\left(e_i(W1)\right)e_i(W1)\right] < 0 .$$

We show under which conditions the net social benefit of reducing overcompliance is larger than the net social cost associated with increasing underdeterrence.

First, we assume that the emission standard $\bar{e}$ is set such that $MAC(\bar{e}) = MB$. This implies that $MSC(e)$ will be positive for emission reductions below the standard since these extra reductions are socially detrimental, and $MSC(e)$ will be negative for emission levels exceeding the emission limit. Moreover due to the convexity of the individual and aggregate abatement cost functions and the constant marginal benefits *MB,* we have $\left|MSC(\overline{e}-\varepsilon)\right| > \left|MSC(\overline{e}+\varepsilon)\right|$ for $0 < \varepsilon \le \overline{e}$ with ε representing the size of the deviation from the emission standard.

Secondly, since the change in emissions will be highest when we compare system *F* and system *W1*, we evaluate these two policy scenarios and do not consider system *W2.* Under system *W1*, all probabilities $P^{W}(\theta)$ and $P^{F}(\theta)$ in equation (8) become zero.

Thirdly, we investigate under which conditions the following inequality holds true:

$$\int_{0}^{\theta_a}\left[\overline{e}-\frac{\overline{e}}{1+\gamma}\right]d\theta+\int_{\theta_a}^{\theta_{\overline{e}}}\left[\overline{e}-\min\left(e_i^a(\theta),\overline{e}\right)\right]d\theta \ge$$

$$\int_{\theta_{\overline{e}}}^{\theta_b}\left[\min\left(e_i^{\overline{e}}(\theta),\frac{\overline{e}[1+\gamma]}{1-\gamma}\right)-\min\left(e_i^{\overline{e}}(\theta),\frac{\overline{e}}{1-\gamma}\right)\right]d\theta+\int_{\theta_b}^{\theta_\beta}\left[\min\left(e_i^{\overline{e}}(\theta),\frac{\overline{e}[1+\gamma]}{1-\gamma}\right)-e_i^b(\theta)\right]d\theta$$

Replacing the functions e_i^a, $e_i^{\overline{e}}$ *and* e_i^b by their lower or upper bounds, allows us to derive a sufficient condition for this inequality to hold:

$$\overline{e}\frac{\gamma}{1-\gamma}\theta_a(\gamma) \ge \overline{e}\frac{\gamma}{1-\gamma}\left[\theta_\beta(\gamma)-\theta_{\overline{e}}(\gamma)\right] \Leftrightarrow \frac{1-\gamma}{1+\gamma} \ge \frac{\theta_\beta(\gamma)-\theta_{\overline{e}}(\gamma)}{\theta_a(\gamma)}$$

We define $\tilde{\gamma}$ as the solution of the following set of conditions:

$$\frac{\tilde{\gamma}}{1-\tilde{\gamma}}\frac{d\theta_a(\tilde{\gamma})}{d\gamma}+\theta_a(\tilde{\gamma})\frac{1}{(1-\tilde{\gamma})^2}-\frac{\tilde{\gamma}}{1-\tilde{\gamma}}\left[\frac{d\theta_\beta(\tilde{\gamma})}{d\gamma}-\frac{d\theta_{\overline{e}}(\tilde{\gamma})}{d\gamma}\right]-\left[\theta_\beta(\tilde{\gamma})-\theta_{\overline{e}}(\tilde{\gamma})\right]\frac{1}{(1-\tilde{\gamma})^2}=0$$

$$0 \le \tilde{\gamma} \le 1$$

Note also that from the definitions of the thresholds follows that:

$$\frac{d\theta_a(\gamma)}{d\gamma}<0, \frac{d\theta_{\overline{e}}(\gamma)}{d\gamma}>0 \; and \; \frac{d\theta_{\overline{e}}(\gamma)}{d\gamma}>0$$

We can now state that: for $\gamma \le \tilde{\gamma}$, $\sum_i\left[MSC\left(e_i(F)\right)e_i(F)-MSC\left(e_i(W1)\right)e_i(W1)\right]<0$.

References

Becker, G.S., 1968. Crime and punishment: An economic approach. *Journal of Political Economy* 76 (2), 169-217.

Blackstone, W., 1769. *Commentaries on the Laws of England*. Clarendon Press.

Bose, P., 1995. Regulatory errors, optimal fines and the level of compliance. *Journal of Public Economics* 56, 475-484.

Chu, C.Y.C., Sheng-cheng, H., Huang, T., 2000. Punishing repeat offenders more severely. *International Review of Law and Economics* 20 (1), 127-140.

Cohen, M.A., 2000. *Monitoring and enforcement of environmental policy*, in Tietenberg, T. and Folmer, S. (eds.), International Yearbook of Environmental and Resource Economics, Volume III. Edward Elgar Publishers.

Cooter, R., Ulen, T., 1997. *Law and Economics*. Addison-Wesley.

Craswell R., Calfee, J.E., 1986. Deterrence and uncertain legal standards. *Journal of Law, Economics and Organization* 2, 279-303.

Delhaye, E., 2007. The enforcement of speeding: Should fines be higher for repeated offences? *Transportation Planning and Technology* 30(3), 355-375.

Dworkin, R., 2002. *Taking rights seriously*. Harvard University Press.

Eckert, H., 2004. Inspections, warnings, and compliance: The case of petroleum storage regulation. *Journal of Environmental Economics and Management* 47, 232-259.

Emons, W., 2003. A note on the optimal punishment for repeat offenders. *International Review of Law and Economics* 23, 253-259.

Emons, W., 2004. Subgame perfect punishment for repeat offenders. *Economic Inquiry* 42(3), 496-502.

Emons, W., 2007. Escalating penalties for repeat offenders. *International Review of Law and Economics* 27, 170-178.

Garoupa, N., Klerman, D., 2004. Corruption and the optimal use of nonmonetary sanctions. *International Review of Law and Economics* 24, 219-225.

Harford, J.D., 1991. Measurement error and state-dependent pollution control enforcement. *Journal of Environmental Economics and Management* 21, 67-81

Harrington, W., 1988. Enforcement leverage when penalties are restricted. *Journal of Public Economics* 37, 29-53.

Innes, R., 2005. Violators avoidance activities and self-reporting in optimal law enforcement. *Journal of Law, Economics, and Organization* 17(1), 239-256.

Kitai, R., 2003. Protecting the guilty. *Buffalo Criminal Law Review* 6, 1163-1187.

Malik, A.S., 1990. Avoidance, screening and optimum enforcement. *RAND Journal of Economics* 21(3), 341-353.

Malik, A.S., 1993. Self-reporting and the design of policies for regulating stochastic pollution. *Journal of Environmental Economics and Management* 24, 241-257.

Miceli, T.J., Segerson, K., 2007. Punishing the innocent along with the guilty: the economics of individual versus group punishment. *Journal of Legal Studies* 36, 81-106.

Nyborg, K., Telle, K., 2004. The role of warnings in regulation: Keeping control with less punishment. *Journal of Public Economics* 88 (12), 2801-2816.

Nyborg, K., Telle, K., 2006. Firms' compliance to environmental regulation: Is there really a paradox? *Environmental and Resource Economics* 35(1), 1-18.

Persson, M., Siven, C., 2007. The Becker paradox and type I vs. type II errors in the economics of crime. *International Economic Review* 48(1), 211-233.

Polinsky, A.M., Shavell, S., 2000. The economic theory of public enforcement of law. *Journal of Economic Literature* 37, 45-76.

Polinsky, A.M., Rubinfeld, D.L., 1991. A model of optimal fines for repeat offenders. *Journal of Public Economics* 46, 291-309.

Rousseau, S., 2007. Timing of environmental inspections: Survival of the compliant. *Journal of Regulatory Economics* 32 (1), 17-36.

Russell, C.S., 1990. Game models for structuring monitoring and enforcement systems. *Natural Resource Modeling* 4, 143-173.

Russell, C.S., Harrington, W., Vaughan, W.J., 1986. *Enforcing Pollution Control Laws*. Resources for the Future.

Shavell, S., 1987a. The optimal use of nonmonetary sanctions as a deterrent. *American Economic Review* 77 (4), 584-592.

Shavell, S., 1987b. *Economic Analysis of Accident Law*. Harvard University Press.

Winston, W.L., 1994. *Operations Research. Applications and algorithms*. Duxbury Press.

Zhao, J., 2002. Irreversible abatement investment under cost uncertainties: tradable emission permits and emissions charges. *Journal of Public economics* 87, 2765– 2789.