Production and Abatement Distortions under Noisy Green Taxes

Hongli Feng and David A. Hennessy

Working Paper 05-WP 409
October 2005

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Hongli Feng and David Hennessy are with the Department of Economics and the Center for Agricultural and Rural Development at Iowa State University. Senior authorship is not assigned.

Questions or comments about the contents of this paper should be directed to Hongli Feng, 560D Heady Hall, Iowa State University, Ames, IA 50011-1070; Ph: (515) 294-6307; Fax: (515) 233-6336; E-mail: hfeng@iastate.edu.

Abstract

Pigouvian taxes are typically imposed in situations where there is imperfect knowledge on the extent of damage caused by a producing firm. A regulator imposing imperfectly informed Pigouvian taxes may cause the firms that should (should not) produce to shut down (produce). In this paper we use a Bayesian information framework to identify optimal signal-conditioned taxes in the presence of such losses. The tax system involves reducing (increasing) taxes on firms identified as causing high (low) damage. Unfortunately, when an abatement decision has to be made, the tax system that minimizes production distortions also dampens the incentive to abate. In the absence of wrong-firm concerns, a regulator can solve the problem by not adjusting taxes for signal noise. When wrong-firm losses are a concern, the regulator has to trade off losses from distorted production incentives with losses from distorted abatement incentives. The most appropriate policy may involve a combination of instruments.

Keywords: conditioning; heterogeneity; informativeness; Pigouvian tax; signaling

JEL classification: D62, H20

Production and Abatement Distortions under Noisy Green Taxes

1. Introduction

Nonpoint source pollution of air and water is an important and particularly vexing economic problem. When accurate monitoring is prohibitively expensive, the corrective regulation will inevitably involve errors in assessing contributions to damage. A large literature, much of which is surveyed in Shortle and Horan [13], has emerged with intent to identify optimal policies. Some studies, as in Xepapadeas [14], and Millock, Sunding, and Zilberman [11], consider the act of monitoring as a policy choice variable. Most of them, such as the ambient taxes modeled in Segerson [12] and Cabe and Herriges [3], use proxies for damage when assessing an imperfect Pigouvian tax. However, as far as we can ascertain, no studies have been done on how taxation in the presence of noisy signals concerning damage might affect incentives to abate. That is the matter of our article.

We pose a Bayesian model in which external damage can take two levels and polluters can be of two types with regard to pollution activities. The types differ in the probability of causing low damage. Otherwise, types differ by production costs. A regulator can monitor a firm to better discern the firm's type, and then use this information to guide firm activities (produce or not). The costs of monitoring are not relevant to the argument we make, and we ignore them. The signal received, though unbiased and not manipulable by the firm, is imperfect or noisy. In the absence of any opportunity to abate on the part of the firm, we show it is optimal for the regulator to use the signal to assess a tax equal to the firm's expected damage conditional upon the signal. In this way, losses from what we call the wrong-firm problem (WFP) are minimized. The WFP is the sum of two concerns: that a noisy Pigouvian tax induces i) a firm that would on average generate positive social surplus under production to not produce, and ii) a firm that would on average generate negative social surplus under production to produce.

We then introduce an abatement opportunity, the cost of which varies across firms. In this setting, specifying the tax equal to expected damage may not be optimal. This is because the averaging involved in forming an expectation on damage reduces the incentive to incur the abatement cost. A further issue is that in the Bayesian game there may be a unique equilibrium set of abaters, or multiple equilibrium sets of abaters. In all cases, however, too little abatement will occur. The adverse consequences for abatement are shown to be most severe when the WFP is a significant problem, i.e., when one might otherwise have benefited most from setting taxes equal to expected damage. This presents a policy dilemma in that, without further information, production distortions may have to be traded off against abatement distortions. We discuss the judicious use of policy instrument combinations.

2. Model framework

The model we develop concerns regulation over a set of firms sufficiently small that production consequences do not affect output market prices. A consumption good has market value B and can be produced by a continuum of firms with each supplying one unit of output. These firms differ by the cost $c \geq 0, c \in [\underline{c}, \overline{c}]$, required to produce that single unit of output, and by the level of negative externality $\theta \geq 0, \theta \in [\underline{\theta}, \overline{\theta}]$, produced as a by-product.[1] Entries and exits occur at no cost. A signal $x \in [\underline{x}, \overline{x}]$ can be obtained on θ at no cost to the viewer. The signal may be thought of as a technology indicator, field slope (nitrogen run-off), or proximity to a city (hog odor). Or the signal may be the result of a science-based test on air or water emissions.[2]

---

[1] Our model can also be interpreted as one of heterogeneity within a firm whereby different outputs are tagged with different c and θ values. Then the firm's decision is not whether to produce a single output but rather how much to produce.

[2] Various issues that lead to noise in a field level test for phosphorus runoff are provided in Mallarino *et al.* [9]. Klatt *et al.* [8] use five soil phosphorus tests on 332 field soil samples in an Iowa watershed. Result correlations range from 0.97 down to 0.88. Bearing in mind that none of the tests measures true soil phosphorus, this attests to the significance of noise in real-world signals that might be used.

Cost is distributed independently of the signal and damage so that firms are distributed according to density $f_c(c)f_{\theta,x}^X(\theta,x)$ on support $S=[\underline{c},\overline{c}]\times[\underline{\theta},\overline{\theta}]\times[\underline{x},\overline{x}]\subset\mathbb{R}^3$, a box in the space of three-dimensional reals. The superscripted X on $f_{\theta,x}^X(\theta,x)$ indicates the signal structure. If a different set of signals, labeled X', is available then we write the new joint density between signals and damage as $f_{\theta,x}^{X'}(\theta,x)$. For convenience, we normalize such that $\int_S f_c(c)f_{\theta,x}^X(\theta,x)dcd\theta dx=1$. In addition, and again for convenience, we assume that $f_c(c)f_{\theta,x}^X(\theta,x)>0$ over the entire support. We do not consider at this point one additional source of heterogeneity, namely, the cost of an abatement action.

Write the mass density function for social cost $z=c+\theta$ as $f_z(z)$,[3] the cumulative mass function as $F_z(z)=\int_0^z f_z(s)ds$, the cumulative marginal for c as $F_c(c)=\int_0^c f_c(s)ds$, and the signal-conditioned mass density for θ as $f_{\theta|x}^X(\theta\,|\,x)=f_{\theta,x}^X(\theta,x)/\int_\theta f_{\theta,x}^X(\theta,x)d\theta$. Consistency requires $\int_\theta f_{\theta,x}^X(\theta,x)dx\equiv\int_\theta f_{\theta,x}^{X'}(\theta,x)dx\equiv f_\theta(\theta)\ \forall X,X'$, so that we may omit the signal structure superscript in this case. Other marginals, cumulatives, and conditionals are represented in like manner. Describe means by the average, or expectation, operator $E[\cdot]$. In particular, the signal-conditioned mean of θ under signal structure X is written as $E_X[\theta\,|\,x]=\int\theta f_{\theta|x}^X(\theta\,|\,x)d\theta$.

We first consider a perfectly informed (i.e., knowing θ) social planner who wishes to implement a tax to maximize social surplus. Another way of saying this is that the tax should optimally sort firms according to production status. With social cost per unit output written as $z=c+\theta$, the social planner's problem is to choose the value of $z=\hat{z}$ so that when only firms with z values at or below $\hat{z}$ produce then social surplus is maximized, i.e., find the argument

[3] Variable $c+\theta$ has nothing to do with the signals, so we omit the superscript indicator of the signal structure when writing $f_z(z)$.

that solves[4]

$$\max_{\hat{z}} \quad BF_z(\hat{z}) - \int_{z\le\hat{z}} zf_z(z)dz. \tag{1}$$

The optimality condition is

$$B - z^*_{fb} = 0, \tag{2}$$

where z^*_{fb} is the marginal firm's value of $c+\theta$. A firm with $c+\theta \le z^*_{fb}$ produces while a firm with $c+\theta > z^*_{fb}$ does not.

It is shown in the appendix that a tax at rate $E[\theta] = \int \theta f_\theta(\theta)d\theta$, i.e., average damage over the entire firm set, is optimal when the regulator has available no conditioning information. Under this tax, firms produce if $c \le B - E[\theta]$. Different firms are active under the first-best tax when compared with under the uninformed tax. Partition parameter space S into four sets. Set A_d^{sn} (A_{dn}^{sn}) comprises firms that do (do not) produce under the average tax solution and should not produce under no conditioning information. Set A_d^s (A_{dn}^s) comprises firms that do (do not) produce under the average tax solution and should produce under no conditioning information. The set measures are

$$\begin{aligned} A_d^{sn} &= \int_{\substack{c\le B-E[\theta],\\ c+\theta > z^*_{fb}}} f_c(c)f_\theta(\theta)dcd\theta; & A_{dn}^{sn} &= \int_{\substack{c > B-E[\theta],\\ c+\theta > z^*_{fb}}} f_c(c)f_\theta(\theta)dcd\theta; \\ A_d^{s} &= \int_{\substack{c\le B-E[\theta],\\ c+\theta \le z^*_{fb}}} f_c(c)f_\theta(\theta)dcd\theta; & A_{dn}^{s} &= \int_{\substack{c > B-E[\theta],\\ c+\theta \le z^*_{fb}}} f_c(c)f_\theta(\theta)dcd\theta. \end{aligned} \tag{3}$$

Fig. 1 illustrates, where the shaded areas represent the two different types of errors. The problem with the uninformed tax solution is that, in general, the measure of firms in each of sets A_d^{sn} and A_{dn}^s is likely to be nontrivial. Pigouvian taxation may improve performance, but the extent of success will depend upon the statistical attributes of parameters c and θ, and in particular on how well a tax matches the damage done. Fig. 1 provides a benchmark for our

[4] We could introduce a downward-sloping demand function and include consumer surplus in the calculation. All the insights to follow still apply, but the demonstrations are longer.

analysis. It is, however, only a limiting case for the problem we will address. In Fig. 1 it is assumed that all information on damage is available to the regulator. More generally, we will study the case where a regulator receives a signal and uses it to design signal-conditioned optimal taxes.

3. Pigouvian taxation

The model in this section does not address the abatement issue. It only asks how a regulator should tax in the presence of noisy signals on damage when the single decision a firm can make is whether to produce. Since a firm's value of θ is not always available to the regulator, it is more realistic to condition the tax on some other attribute. Let the timeline be that 0) nature draws c and type for a firm, 1) then the regulator receives a signal about a firm's production environment and levies a tax on the unit in order to optimally deter damage, and 2) then the producer makes the production decision. The signal received at time 1 is random; it is not manipulated in any way by the sender. Fig. 2 provides a timeline.

Our focus is on economic losses due to policy issues. In order to be concrete about how policy problems can arise and can be mitigated, we place more structure on the firm set and on signal structures. There are two damage levels, H with value θ_H and L with value $\theta_L < \theta_H$. There are two types of firms, type α and type $\alpha+\delta$, $\delta > 0$. For type α (type $\alpha+\delta$), the regulator observing the producer's operation will receive an L signal with probability $\alpha \in [0,1)$ (respectively, $\alpha+\delta \in (0,1]$) and an H signal with probability $1-\alpha$ (respectively, $1-\alpha-\delta$). The way to view α is as a lottery for the regulator whereby the true damage will turn out to be θ_L with probability α and θ_H with probability $1-\alpha$. The signal emitted by a firm is unbiased in that if the firm is an α then L is emitted with probability α and damage is θ_L with probability α. But the signal realization is *independent* of the damage realization. Note that even if the regulator were to observe the firm type, she would still be exposed to this lottery risk. When we assert that a firm should produce we mean that the firm provides a lottery that makes the firm, in expectation, a (positive) surplus generator.

Type-conditioned expected damage levels are

$$\begin{aligned} &E_{\alpha}[\theta] = \alpha\theta_L + (1-\alpha)\theta_H; \\ &E_{\alpha+\delta}[\theta] = (\alpha+\delta)\theta_L + (1-\alpha-\delta)\theta_H = E_{\alpha}[\theta] - \delta(\theta_H - \theta_L) < E_{\alpha}[\theta]; \end{aligned} \tag{4}$$

on each of the respective types. These would be appealing tax assessments were firm type known. But firm type is not known, and the only information available to the regulator is signal $i, i \in \{L, H\}$. The share of firms that are $\alpha+\delta$ is, for this section, exogenous at $\phi \in (0,1)$. Later we will endogenize ϕ when we introduce an abatement choice and identify incentives problems concerning abatement investments. In this section the intent is to show that there are efficiency problems that have nothing to do with abatement inefficiencies.

When the regulator receives a signal $i, i \in \{L, H\}$, then she uses Bayes' theorem to assess

$$\begin{aligned} &\pi_{\alpha+\delta|L} = \frac{(\alpha+\delta)\phi}{(\alpha+\delta)\phi + \alpha(1-\phi)} = \frac{(\alpha+\delta)\phi}{\alpha+\delta\phi}; \\ &\pi_{\alpha+\delta|H} = \frac{(1-\alpha-\delta)\phi}{(1-\alpha-\delta)\phi + (1-\alpha)(1-\phi)} = \frac{(1-\alpha-\delta)\phi}{1-\alpha-\delta\phi}; \end{aligned} \tag{5}$$

where $\pi_{\alpha+\delta|i}$ is the probability that the firm is a $\alpha+\delta$ given signal $i \in \{L, H\}$. The expected damage levels given signals are

$$\begin{aligned} &E_L[\theta] = \pi_{\alpha+\delta|L} E_{\alpha+\delta}[\theta] + (1-\pi_{\alpha+\delta|L}) E_{\alpha}[\theta] \\ &\quad = (\alpha+\delta)\theta_L + (1-\alpha-\delta)\theta_H + (\theta_H - \theta_L)\delta\frac{\alpha(1-\phi)}{\alpha+\delta\phi} \geq (\alpha+\delta)\theta_L + (1-\alpha-\delta)\theta_H; \\ &E_H[\theta] = \pi_{\alpha+\delta|H} E_{\alpha+\delta}[\theta] + (1-\pi_{\alpha+\delta|H}) E_{\alpha}[\theta] \\ &\quad = \alpha\theta_L + (1-\alpha)\theta_H + (\theta_L - \theta_H)\delta\frac{(1-\alpha-\delta)\phi}{1-\alpha-\delta\phi} \leq \alpha\theta_L + (1-\alpha)\theta_H; \\ &E_H[\theta] - E_L[\theta] = \frac{(\theta_H - \theta_L)(1-\phi)\phi\delta^2}{(1-\alpha-\delta\phi)(\alpha+\delta\phi)} \geq 0. \end{aligned} \tag{6}$$

The following is shown in the appendix.

Proposition 1. *Optimal taxes are* $t_L = E_L[\theta]$ *under signal L and* $t_H = E_H[\theta]$ *under signal H, as provided in (6).*

Henceforth we will label $t_L = E_L[\theta]$ and $t_H = E_H[\theta]$ as signal-conditioned expected damage taxes, or SCED taxes. Note that, from (4) and (6), $\theta_L \le E_{\alpha+\delta}[\theta] \le t_L \le t_H \le E_{\alpha}[\theta] \le \theta_H$. The SCED taxes are squeezed between the taxes that would apply were firm type known to the regulator. This observation will be central when we study abatement incentives.

To this point we have assumed that the lotteries represented by firms of types α and $\alpha+\delta$ are fixed. These lotteries represent the informativeness of the signals given. We will show next that the level of dispersion in signals determines the level of dispersion in optimal taxes.[5] This is relevant because, as we will show later, dispersion in taxes encourages abatement.

Let a different conditioner be used so that the signal structure changes, $X \to X'$. The different conditioner may perhaps be a different scientific test or attribute of an asset. The nature of the change is that $\alpha \to \alpha - \eta$ where $\eta \ge 0$ so that the fraction of L signals for a non-abating firm declines. Since only the signal changes, we must preserve unbiasedness in the unconditional probability of low damage. Thus, $(\alpha+\delta)\phi + \alpha(1-\phi) = \alpha + \delta\phi$, and we must have $\alpha+\delta \to \alpha+\delta+\hat{\eta}$, $\hat{\eta} = (1-\phi)\eta/\phi$, so that $\alpha(1-\phi)+(\alpha+\delta)\phi \equiv (\alpha-\eta)(1-\phi)+(\alpha+\delta+(1-\phi)\eta/\phi)\phi$ as required. Under SCED taxation and the new signal structure, $t_L = E_L[\theta]$ in (6) becomes

$$t_L(\eta) = \pi_{\alpha+\delta+\hat{\eta}|L} E_{\alpha+\delta+\hat{\eta}}[\theta] + (1-\pi_{\alpha+\delta+\hat{\eta}|L}) E_{\alpha-\eta}[\theta], \tag{7}$$

where

[5] This is related to observations by Blackwell [1] and Epstein [5] that the value of information and utility of information in decision making are related to the dispersion of the signals the information conveys.

$$\pi_{\alpha+\delta+\tilde{\eta}|L} = \pi_{\alpha+\delta|L} + \frac{(1-\phi)}{\alpha+\delta\phi}\eta;$$
$$E_{\alpha-\eta}[\theta] = E_{\alpha}[\theta] + (\theta_H - \theta_L)\eta; \qquad E_{\alpha+\delta+\tilde{\eta}}[\theta] = E_{\alpha+\delta}[\theta] - \frac{(1-\phi)(\theta_H - \theta_L)}{\phi}\eta; \tag{8}$$

so that

$$\begin{aligned} t_L(\eta) &= \left(\pi_{\alpha+\delta|L} + \frac{(1-\phi)}{\alpha+\delta\phi}\eta\right)\left(E_{\alpha+\delta}[\theta] - \frac{(1-\phi)(\theta_H - \theta_L)}{\phi}\eta\right) \\ &\quad + \left(1 - \pi_{\alpha+\delta|L} - \frac{(1-\phi)}{\alpha+\delta\phi}\eta\right)\left(E_{\alpha}[\theta] + (\theta_H - \theta_L)\eta\right) \\ &= t_L(\eta=0) - 2\frac{\delta(1-\phi)(\theta_H - \theta_L)}{\alpha+\delta\phi}\eta - \frac{(1-\phi)(\theta_H - \theta_L)}{(\alpha+\delta\phi)\phi}\eta^2 \le t_L(\eta=0). \end{aligned} \tag{9}$$

Similarly, under ex post taxation and using $\pi_{\alpha+\delta+\tilde{\eta}|H} = \pi_{\alpha+\delta|H} - (1-\phi)\eta/(1-\alpha-\delta\phi)$, then t_H becomes

$$\begin{aligned} t_H(\eta) &= \pi_{\alpha+\delta+\tilde{\eta}|H} E_{\alpha+\delta+\tilde{\eta}}[\theta] + (1 - \pi_{\alpha+\delta+\tilde{\eta}|H}) E_{\alpha-\eta}[\theta] \\ &= t_H(\eta=0) + 2\frac{\delta(1-\phi)(\theta_H - \theta_L)}{1-\alpha-\delta\phi}\eta + \frac{(1-\phi)(\theta_H - \theta_L)}{(1-\alpha-\delta\phi)\phi}\eta^2 \ge t_H(\eta=0), \end{aligned} \tag{10}$$

so that

$$\begin{aligned} t_H(\eta) - t_L(\eta) &= t_H(\eta=0) - t_L(\eta=0) + 2\frac{\delta(1-\phi)(\theta_H - \theta_L)}{(1-\alpha-\delta\phi)(\alpha+\delta\phi)}\eta \\ &\quad + \frac{(1-\phi)(\theta_H - \theta_L)}{\phi(1-\alpha-\delta\phi)(\alpha+\delta\phi)}\eta^2 \ge t_H(\eta=0) - t_L(\eta=0). \end{aligned} \tag{11}$$

From equations (9)-(11), the following results can be easily obtained.

Proposition 2. *Under SCED taxation, all of* t_H, $-t_L$, *and* $t_H - t_L$ *are increasing and convex in* η.

A more informative signaling system, as defined by an increase in η, will render the SCED taxes more dispersed. We will defer comments on Proposition 2 until we study abatement problems that SCED taxes may generate. Our next goal is to characterize the losses that SCED taxes are optimal in guarding against.

4. Losses from the WFP

Under the SCED taxes given in (6), we will develop a loss function capturing statistical type I and II errors, namely, the losses when a tax induces firms that I) should produce (under the lottery in damages the firm provides) to shut down, and II) should shut down to produce. Consider I) first. This can occur when the tax is too high, i.e., when the firm is type $\alpha+\delta$ and either the signal is *i*) H or *ii*) L.[6] For *i*), the expected benefit from production would be $B-c-E_{\alpha+\delta}[\theta]\geq 0$ since the firm should produce. But $B-c-t_H<0$, so that the firm does not produce. Therefore, the error occurs on firm set $c\in(B-t_H, B-E_{\alpha+\delta}[\theta]]$. The expected loss from non-production is $B-c-E_{\alpha+\delta}[\theta]\geq 0$ over this set of firms. Finally, the probability of the error is $1-\alpha-\delta$ on fraction ϕ of firms. For *ii*), the expected benefit from production is as before. But $B-c-t_L<0$ so that the error occurs on firm set $c\in(B-t_L, B-E_{\alpha+\delta}[\theta]]$. The probability of the error is $\alpha+\delta$ on fraction ϕ of firms. Summing the two cases and integrating over c, aggregate loss from this sort of error is

$$\begin{aligned}\mathcal{L}_I &= (1-\alpha-\delta)\phi\int_{c\in(B-t_H, B-E_{\alpha+\delta}[\theta]]}\left(B-c-E_{\alpha+\delta}[\theta]\right)dF_c(c)\\ &\quad +(\alpha+\delta)\phi\int_{c\in(B-t_L, B-E_{\alpha+\delta}[\theta]]}\left(B-c-E_{\alpha+\delta}[\theta]\right)dF_c(c).\end{aligned}\tag{12}$$

Now consider production by firms that should shut down but the tax is too low, i.e., when the firm is an α and the signal is either *i*) L or *ii*) H. For case *i*), the expected benefit from production would be $B-c-E_\alpha[\theta]<0$ since the firm should not produce. But $B-c-t_L\geq 0$, so that the firm produces. Therefore, $c\in(B-E_\alpha[\theta], B-t_L]$. The expected loss from production is $-(B-c-E_\alpha[\theta])\geq 0$ over this firm set. The probability of misdiagnosis is α on fraction $1-\phi$ of firms. For case *ii*), the expected benefit from production is as before. Now, though, $B-c-t_H$

[6] When the signal is H and the type is α then satisfaction of the inequalities $B-c-E_\alpha[\theta]\geq 0$ and $B-c-t_H<0$ would lead to type I error. But (6) relates that $t_H>B-c\geq E_\alpha[\theta]$ cannot occur.

≥ 0 so that the firm set in question is $c \in (B - E_{\alpha}[\theta], B - t_H]$ with probability of misdiagnosis $1-\alpha$ on fraction $1-\phi$ of firms. Summing the two cases and integrating over c, the aggregate loss from this sort of error is

$$\begin{aligned} \mathcal{L}_{II} &= \alpha(1-\phi)\int_{c\in(B-E_{\alpha}[\theta],B-t_L]}\big(c+E_{\alpha}[\theta]-B\big)\,dF_c(c) \\ &\quad +(1-\alpha)(1-\phi)\int_{c\in(B-E_{\alpha}[\theta],B-t_H]}\big(c+E_{\alpha}[\theta]-B\big)\,dF_c(c). \end{aligned} \tag{13}$$

Summing, total expected loss from misdirected firms is[7]

$$\mathcal{L} = \mathcal{L}_I + \mathcal{L}_{II}. \tag{14}$$

The loss function is described in Fig. 3, with $B-c$ on the horizontal axis and conditional loss on the vertical axis. Intervals Ia and Ib are where type I errors occur; see (12). In the loss function, two weightings are applied. One is a cost weighting as represented through the integration in (12) and the downward-sloping line segment in Fig. 3. The other is a weighting by the overall probability that wrong signals of this sort occur, and this second weighting differs across the intervals Ia and Ib. Interval IV represents firms that do not produce and should not produce under both signals, so there are no losses for such firms. Intervals IIa and IIb are where type II errors occur, and the structure in (13) is similar to that for type I error. On the right is interval III, representing firms such that production should and does occur (given limited information as represented by the lotteries) under either signal so there could be no loss.[8] The possibility of type I and II errors does not disappear when the signal is completely informative because it informs only on firm type and the types continue to emit heterogeneous signals.

5. Abatement, absent the WFP

From this point on we extend the H-L damage model by assuming a firm can alter the extent

---

[7] Optimizing $\mathcal{L}$ over choices of t_L and t_H provides an alternative proof of Proposition 1; see the appendix.

[8] Obviously a firm cannot be subject to both error types. While the intervals overlap, they are for different firm types and that additional dimension is omitted in the figure.

of damage done through an abatement activity. That is, the firm becomes type α or type $\alpha+\delta$ by making an abatement choice. The timeline is now that 0) nature draws the production cost c and the abatement cost, call it e, for a firm, 1) then the firm makes a non-observable investment decision on abatement that determines firm type, 2) then the regulator receives a signal $i, i \in \{L, H\}$, about a firm's production environment and levies a SCED tax on the unit, and 3) then the producer makes the production decision. Fig. 4 provides the timeline. In this section, however, we assume that firms always operate (i.e., $B-c$ is sufficiently high) since our intent is to understand how SCED taxation in the presence of noisy signals on damage can affect incentives even absent the WFP. We will later show that the problem we identify here and the WFP have similar effects on incentives.

Signals emitted by firms are affected by the abatement action. Under no action, the regulator observing (ex post) the producer's operation will receive an L signal with probability $\alpha \in (0,1)$ and an H signal otherwise. Under the action, denoted as letter a, the regulator observing the producer's operation will receive an L signal with probability $\alpha+\delta$, and an H signal otherwise. The cost of the action, e, is heterogeneous across firms, having strictly increasing distribution $J(e):[\underline{e},\overline{e}] \to [0,1]$. This distribution is known to the regulator, but no more is known to the regulator apart from the observed signal. So a firm is now characterized by the couple (c,e), and the firm's lottery characterization of α or $\alpha+\delta$ is an endogenous choice. For convenience we will assume that e and c are independent.

The regulator's assessed SCED tax upon receiving a signal $i \in \{L, H\}$ is t_i. The sum of expected tax and action cost to the firm upon not taking the action is set at $\alpha t_L + (1-\alpha)t_H$ since t_L occurs with probability α and t_H occurs with probability $1-\alpha$. The expected tax upon taking the action is $(\alpha+\delta)t_L + (1-\alpha-\delta)t_H + e$, so that the action will be taken if $(t_H - t_L)\delta \geq e$. The indifferent producer is that with e value

$$\hat{e} = (t_H - t_L)\delta, \tag{15}$$

while the set that chooses to incur the abatement sunk cost has measure $J(\hat{e})$. For later comparison, the efficient solution is when the threshold abatement cost satisfies $(\alpha+\delta)\theta_L +$ $(1-\alpha-\delta)\theta_H + e = \alpha\theta_L + (1-\alpha)\theta_H$, with solution

$$\tilde{e} = (\theta_H - \theta_L)\delta. \tag{16}$$

To identify the problem with SCED taxation, assume that taxes are as in (6), except that endogenous $J(\hat{e})$ substitutes for exogenous ϕ. Thus, taxes are

$$\begin{aligned}
t_L &= (\alpha+\delta)\theta_L + (1-\alpha-\delta)\theta_H + (\theta_H - \theta_L)\delta\frac{\alpha[1-J(\hat{e})]}{\alpha+\delta J(\hat{e})} \geq \theta_L; \\
t_H &= \alpha\theta_L + (1-\alpha)\theta_H + (\theta_L - \theta_H)\delta\frac{(1-\alpha-\delta)J(\hat{e})}{1-\alpha-\delta J(\hat{e})} \leq \theta_H; \\
t_H - t_L &= \frac{(\theta_H - \theta_L)[1-J(\hat{e})]J(\hat{e})\delta^2}{[1-\alpha-\delta J(\hat{e})][\alpha+\delta J(\hat{e})]} \geq 0;
\end{aligned} \tag{17}$$

so that $\hat{e} = (t_H - t_L)\delta \leq (\theta_H - \theta_L)\delta = \tilde{e}$.

Notice that t_L is decreasing in e because the probability the action is taken given L is increasing in the fraction taking the action. Similarly, t_H is decreasing in e. This means that the taxes are (weakly at any rate) too high. Noting that $\hat{e} \leq \tilde{e}$, means damage is given as

$$\begin{aligned}
&J(\hat{e})E_{\alpha+\delta}[\theta] + \left[1-J(\hat{e})\right]E_{\alpha}[\theta] = \alpha\theta_L + (1-\alpha)\theta_H - \delta J(\hat{e})(\theta_H - \theta_L) \\
&\geq \alpha\theta_L + (1-\alpha)\theta_H - \delta J(\tilde{e})(\theta_H - \theta_L),
\end{aligned} \tag{18}$$

and is socially excessive. In summary,

Proposition 3. *Let the tax imposed on a firm be the SCED. Assume that e and c are independently distributed and the support of c is such that no firm will shut down upon emitting signal H. Then, relative to perfect signals on firm type, noisy signals lead to underinvestment in abatement and an increase in mean damage.*

The central point of this paper is the contrast between Proposition 1 (where abatement opportunities were not available) and Proposition 3. When in the presence of abatement opportunities, setting taxes equal to SCED will distort incentives. From (15) and (17) we have

$$\hat{e} = \frac{J(\hat{e})[1-J(\hat{e})](\theta_H - \theta_L)\delta^3}{[1-\alpha-\delta J(\hat{e})][\alpha+\delta J(\hat{e})]}. \tag{19}$$

Since $t_H - t_L$ is not monotone in e we cannot be sure from (19) whether there is a unique fixed-point solution. Before going further, let us consider what the condition relates. Nature moves first in describing the distribution of firm characteristics. The regulator's rule of setting taxes equal to SCED is known to firms. Firms make production and abatement decisions that are consistent with that rule and in light of the noise externalities that firms impose on each other. The game is Bayesian where the regulator interprets signals to arrive at fixed-point problem (19), solves it for the set abating, and then sets tax values using (17). Thus, (19) and (17) represent a perfect Bayesian equilibrium in a game of imperfect information as explained in Fudenberg and Tirole [6]. Three cases illustrate much of the content in (17) and (19).

Example 1. At one extreme suppose that $\alpha = 0$ so that those not taking the action cannot attain the L signal. Then $\pi_{a|L} = 1$, i.e., those attaining the L signal do damage $\theta_H + \delta(\theta_L - \theta_H)$ and are assessed a tax at that amount. On the other hand, $t_H = \theta_H + (\theta_L - \theta_H)(1-\delta)J(e)\delta/[1-\delta J(e)]$ $< \theta_H$ because some action takers receive an H signal. In equilibrium, $\hat{e} =$ $[1-J(\hat{e})](\theta_H - \theta_L)\delta^2/[1-\delta J(\hat{e})]$ where the right-hand side is decreasing in $\hat{e}$. Thus, in this case there is a unique solution to the equilibrium fraction of action takers. The value of $J(\hat{e})$ increases with an increase in either δ or θ_H and decreases with an increase in θ_L. ■

Example 2. At the other extreme, suppose that $\alpha + \delta = 1$ so that those taking the action always attain the L signal. Then $t_H = \alpha\theta_L + (1-\alpha)\theta_H$ because those emitting the H signal certainly do not take the action. For L signal emitters, $t_L = \theta_L + (\theta_H - \theta_L)\delta(1-\delta)[1-J(e)]/[1-\delta+\delta J(e)] >$

$(\alpha+\delta)\theta_L+(1-\alpha-\delta)\theta_H$. This is because some not taking the action do emit a low signal. If $\alpha+\delta=1$ then (19) reduces to $\hat{e}=J(\hat{e})(\theta_H-\theta_L)\delta^2/[1-\delta+\delta J(\hat{e})]$. The right-hand side is increasing in $\hat{e}$ and, furthermore, the derivative may have value exceeding 1. Indeed, with $J(e=0)=0$ then the value $\hat{e}=0$ is a solution. In that case any strictly positive solution will be unique whenever $J(e)/[1-\delta+\delta J(e)]$ is concave, i.e., whenever $J_e(e=0)(\theta_H-\theta_L)(1-\delta)\delta^2>$ $[1-\delta+\delta J(e=0)]^2$ and $(1-\delta)J_{ee}(e)+\delta\{J(e)J_{ee}(e)-2[J_e(e)]^2\}<0$. Notice that $J_{ee}(e)<0$ is sufficient to ensure a unique positive solution when $\alpha+\delta=1$, but this is a strong requirement. In the terminology of Caplin and Nalebuff [4], the distribution function $J(e)$ is -1-convex if $[J(e)]^{-1}$ is convex. For sufficiently smooth distributions, $[J(e)]^{-1}$ is convex whenever $JJ_{ee}-2(J_e)^2<0$. Thus, a δ-weighted average between concavity on $J(e)$ and -1-convexity on $J(e)$ ensures a unique positive solution when $\alpha+\delta=1$. ■

Example 3. Finally, suppose for simplicity that $J(e)$ is the uniform distribution on $[0,1]$ while $\alpha=0.5$. Then (19) resolves to $\delta^2\hat{e}^2-(\theta_H-\theta_L)\delta^3\hat{e}+(\theta_H-\theta_L)\delta^3-0.25=0$ with solutions $0.5(\theta_H-\theta_L)\delta\pm 0.5\delta^{-1}\sqrt{(\theta_H-\theta_L)^2\delta^4-4(\theta_H-\theta_L)\delta^3+1}$. If and only if $0.25>(\theta_H-\theta_L)\delta^3$, then there is a unique (strictly) positive root.[9] Suppose that $(\theta_H-\theta_L)\delta=1$ so that, by (16), the efficient solution is for all to act. Then $\delta^2\hat{e}^2-\delta^2\hat{e}+\delta^2-0.25=0$ with solutions $0.5(1\pm\delta^{-1}\sqrt{1-3\delta^2})$. Since $\delta\in(0,0.5]$ when $\alpha=0.5$, there is no solution on $\hat{e}\in(0,1]$.[10] Thus, for $\delta\in(0,0.5)$ the only solution on $e\in[0,1]$ is $\hat{e}=0$. That is, the Nash equilibrium is for no one to abate when the optimal solution is for all to abate. Is it a stable solution? From (19) and

---

[9] $\hat{e}=0$ is one of three roots, but we have already cancelled through by it.

[10] Since $0.5+0.5\delta^{-1}\sqrt{1-3\delta^2}<1$ if and only if $0.5<\delta$, which is not possible, we need not consider the positive root. On the other hand, $0.5-0.5\delta^{-1}\sqrt{1-3\delta^2}>0$ if and only if $\delta>0.5$.

$(\theta_H - \theta_L)\delta = 1$, private benefit less private cost has value $e(1-e)(0.25-\delta^2 e^2)^{-1}\delta^2 - e$. A deviation to $\Delta e > 0$ at the limit, $\mathrm{Lim}_{\Delta e \downarrow 0} \Delta e$, gives a change in private benefit of

$$\delta^2 \frac{0.25 - 0.5e + \delta^2 e^2}{(0.25 - \delta^2 e^2)^2}\bigg|_{e=0} - 1 = 4\delta^2 - 1 \le 0. \tag{20}$$

If $\delta \in (0, 0.5)$ then a small $\Delta e > 0$ increase in e away from 0 decreases private gains so that the solution is stable.

If instead $\alpha = 0$, then (19) resolves to $\delta \hat{e}^2 - [1 + (\theta_H - \theta_L)\delta^2]\hat{e} + (\theta_H - \theta_L)\delta^2 = 0$. If, as before, $(\theta_H - \theta_L)\delta = 1$ so that the efficient solution is for all to act then (19) becomes $\delta \hat{e}^2 - (1+\delta)\hat{e} + \delta = 0$ with solutions $[(1+\delta)/(2\delta)]\left(1 \pm \sqrt{(1-\delta)(1+3\delta)/(1+\delta)^2}\right)$. The positive root is outside $\hat{e} \in [0,1]$.[11] The negative root is inside the unit interval. In that case, private benefits are $\delta(1-e)(1-\delta e)^{-1} - e$ and the derivate at that root is

$$\delta \frac{\delta - 1}{(1-\delta e)^2}\bigg|_{e=[(1+\delta)/(2\delta)]\left(1-\sqrt{(1-\delta)(1+3\delta)/(1+\delta)^2}\right)} - 1 < 0, \tag{21}$$

so that the equilibrium is stable under a small deviation. ■

Problems related to that in (15)-(16) have arisen elsewhere. Studying a competitive labor market, McGuire and Ruhm [10] showed that the inability to observe drug abuse directly, rather than only through accidents, elicited an insufficient wage incentive to cure drug dependency. Hennessy [7] and Bogetoft and Olesen [2] showed that noisy grading of produce in the presence of competitive trading after production elicited insufficient incentive to invest in quality enhancement. In our case, the problem arises with SCED taxation and not with competitive

---

The roots are on either side of the unit interval.

[11] The larger root is at most one if and only if $\sqrt{(1-\delta)(1+3\delta)} \le \delta - 1$, which is not possible. The smaller root is non-negative if and only if $\delta \ge 0$. Finally, the smaller root is at most one if and only if $0 \le \delta$ so the smaller root is on the unit interval.

trading. As such, the problem is one the regulator has much control over. Indeed, the problem originates from a regulatory attempt to increase social welfare.

6. Abatement, with the WFP

In this case $B-c$ might possibly be so small as to deter production. It is under this possibility, but without abatement opportunities, that SCEDs give optimal taxes. Firm profit under signal $i \in \{L,H\}$ and cost c is $\max[0, B-t_i-c]$ so that the expected revenue enhancement effect of investing in abatement is $(\alpha+\delta)\max[0,B-t_L-c]+(1-\alpha-\delta)\max[0,B-t_H-c]-\alpha\max[0,B-t_L-c]-(1-\alpha)\max[0,B-t_H-c]=\delta(\max[0,B-t_L-c]-\max[0,B-t_H-c])$. When this exceeds the cost of investing in abatement, then investment occurs, i.e., (15) becomes

$$\hat{e}=\left(\max[0,B-t_L-c]-\max[0,B-t_H-c]\right)\delta, \tag{22}$$

where independence between c and e ensures that the t_i are as in Proposition 1.

Fig. 5 graphs the right-hand side of (22) as the value of $B-c$ changes. It is the continuous, piece-wise linear function labeled EPBA, for expected private benefit from abatement. It is flat to t_L, with slope δ on $(t_L, t_H]$, and flat on (t_H, ∞). Graphed also is the expected social benefit from abatement, $\left(\max[0,B-\theta_L-c]-\max[0,B-\theta_H-c]\right)\delta$ and labeled ESBA. It is flat to θ_L, with slope δ on $(\theta_L, \theta_H]$, and flat on (θ_H, ∞). Observe that

$$\theta_H-\theta_L \geq t_H-t_L \geq \max[0,B-t_L-c]-\max[0,B-t_H-c]. \tag{23}$$

This inequality, together with the logic underpinning Proposition 3, leads immediately to

Proposition 4. *Assume that e and c are independently distributed. Assume too that SCED taxes are imposed. Relative to the case without the WFP under noisy signals, under the WFP and noisy signals there is even more underinvestment in abatement and even larger mean damage.*

The contrast between Proposition 1 and Proposition 3 identified a problem with SCED taxes when abatement decisions are to be made. But it assumed away the problem motivating the SCED taxes, leaving the potential loophole that the abatement problem becomes less relevant in the presence of the WFP. Proposition 4 shows that the abatement problem becomes worse in this case. Optimal use of signals to mitigate WFP losses involves guarding against a tax that is too severe when the signal is H and too permissive when the signal is L. But then the incentive to invest in achieving a lower level of expected damage is dulled, and so fewer firms make the investment.

7. Naïve taxation and other policy issues

We have seen that levying Pigouvian taxes at the SCED levels reduces the incentive to abate and, furthermore, the extent of reduction is stronger when the WFP is relevant. If the WFP never occurs, then optimal taxes are easy to identify. Just compare (15) with (16) to see that one should set $t_H = \theta_H$ and $t_L = \theta_L$.[12] We call this naïve taxation in that the regulator may be viewed as naïvely assuming that an L signal means that the monitored firm always causes damage to extent θ_L. In this case, a little knowledge is a dangerous thing in that a comparatively sophisticated regulator who uses Bayes' rule creates a problem that a regulator who taxes θ_i upon seeing $i \in \{L, H\}$ would avoid. However, in stretching out the penalties like that then WFP losses likely increase. From (14), loss due to the wrong-firms problem becomes

$$\begin{aligned}\mathcal{L} = &(1-\alpha-\delta)\phi\int_{c\in(B-\theta_H,B-E_{\alpha+\delta}[\theta]]}\big(B-c-E_{\alpha+\delta}[\theta]\big)dF_c(c)\\ &+(\alpha+\delta)\phi\int_{c\in(B-\theta_L,B-E_{\alpha+\delta}[\theta]]}\big(B-c-E_{\alpha+\delta}[\theta]\big)dF_c(c)\\ &+\alpha(1-\phi)\int_{c\in(B-E_\alpha[\theta],B-\theta_L]}\big(c+E_\alpha[\theta]-B\big)dF_c(c)\\ &+(1-\alpha)(1-\phi)\int_{c\in(B-E_\alpha[\theta],B-\theta_H]}\big(c+E_\alpha[\theta]-B\big)dF_c(c).\end{aligned} \tag{24}$$

[12] Bogetoft and Olesen [2] make a related point when motivating the use of competition through contracts rather than competition through competitive trading after production in markets for

It may, however, be that stretching the taxes delivers no additional abatement. This would occur were $J(\tilde{e}) - J(\hat{e}) = 0$, i.e., were no firms deterred from abating in any case. Suppose that equilibrium e is completely inelastic to the shift from SCED taxation to naïve taxation so that we may write $J(e)$ as ϕ, exogenous. Then the change in losses due to the WFP is

$$\begin{aligned}\Delta\mathcal{L} = &(1-\alpha-\delta)\phi\int_{c\in(B-\theta_H,B-t_H]}\left(B-c-E_{\alpha+\delta}[\theta]\right)dF_c(c)\\ &+(\alpha+\delta)\phi\int_{c\in(B-t_L,B-\theta_L]}\left(B-c-E_{\alpha+\delta}[\theta]\right)dF_c(c)\\ &+\alpha(1-\phi)\int_{c\in(B-t_L,B-\theta_L]}\left(c+E_\alpha[\theta]-B\right)dF_c(c)\\ &+(1-\alpha)(1-\phi)\int_{c\in(B-\theta_H,B-t_H]}\left(c+E_\alpha[\theta]-B\right)dF_c(c).\end{aligned} \tag{25}$$

Since $B-c \geq E_{\alpha+\delta}[\theta]$ on $c\in(B-\theta_H, B-t_H]$ and $B-c \leq E_\alpha[\theta]$ on $c\in(B-t_L, B-\theta_L]$, the first and third right-hand terms are non-negative. The other two right-hand terms are of ambiguous sign and would be non-negative if $\int_{c\in(\theta_L,E_{\alpha+\delta}[\theta]]}dF_c(c) = \int_{c\in(E_\alpha[\theta],\theta_H]}dF_c(c) = 0$. Therefore, the unsuccessful attempt to elicit more abatement can be at the real cost of further distortions in the firm sets producing and not producing.

Circumstances are readily identified where SCED taxation and a subsidy on abatement would be an improvement over either naïve taxation or SCED taxation by themselves. Viewing Fig. 5, we can break the analysis into five cases.

Case i), $c \geq B-\theta_L$: This is not an interesting case because production should never occur and would never occur under naïve taxation or SCED taxation.

Case ii), $c\in[B-t_L, B-\theta_L)$: In this case, the gap between social and private benefit from abatement is

$$\begin{aligned}&\left(\max[0,B-\theta_L-c]-\max[0,B-\theta_H-c]-\max[0,B-t_L-c]+\max[0,B-t_H-c]\right)\delta\\ &=(B-\theta_L-c)\delta.\end{aligned} \tag{26}$$

---

graded produce.

Were it possible, and in addition to SCED taxation, providing a small direct subsidy on abatement to this amount would be optimal. Since abatement is not observable, the subsidy might be in the form of public involvement in technology awareness and/or skill development. Note that the subsidy in (26) depends on c so that, even within the Case ii) interval, the value of c would need to be known in order to implement the subsidy.

Case iii), $c \in [B - t_H, B - t_L)$: Equation (26) assumes the value $(E_L[\theta] - \theta_L)\delta$. Were it feasible, and in addition to SCED taxation, this cost-independent amount would be the optimal subsidy on abatement.

Case iv), $c \in [B - \theta_H, B - t_H)$: The subsidy would be $(B - c - \theta_L + E_L[\theta] - E_H[\theta])\delta$ in this case, and quite difficult to implement.

Case v), $c < B - \theta_H$: This case has already been considered in detail in Section 5. The abatement subsidy would be $(\theta_H - \theta_L + E_L[\theta] - E_H[\theta])\delta$, but naïve Pigouvian taxation without an abatement subsidy would also work.

When costs span two or more of these cases, no single subsidy level will support first best. Finally, a further approach would be to tackle the problem directly by subsidizing research and development of a more informative test. Proposition 2, when viewed together with (15)-(16), shows that the level of abatement that occurs under imposition of SCED taxes is increasing and *convex* in η, implying increasing marginal benefits from an innovation that increases the value of test informativeness.

8. Conclusion

The intent of this paper has been to tease out consequences of Pigouvian taxation in the presence of noisy signals. For policy guidance, we are of the view that a bundle of policies may prove most practical in remedying losses from distorted production and abatement levels. This may involve signal-conditioned expected damage taxes, subsidies where possible on abatement activities, and subsidies to develop more effective testing technologies.

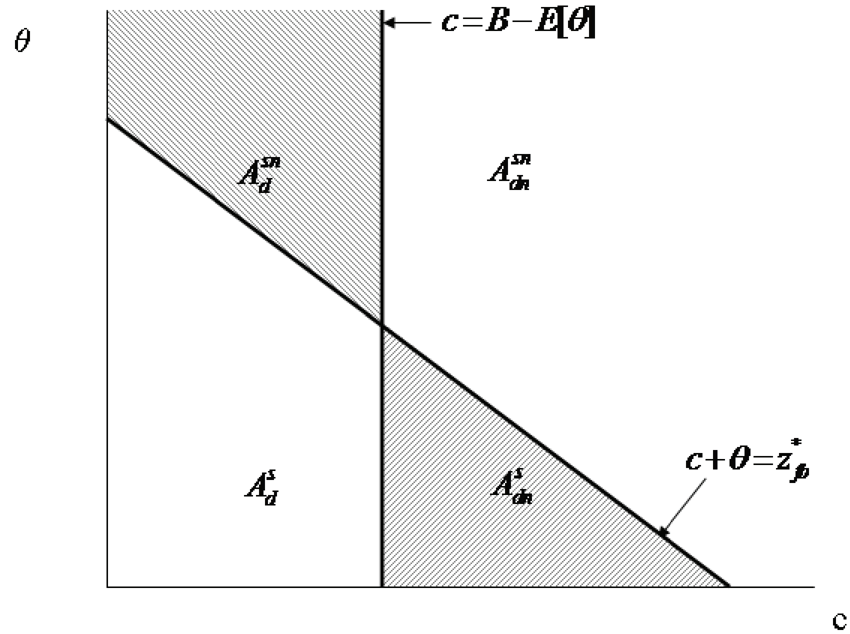


Figure 1. Sets of open and shut firms under average damage tax.

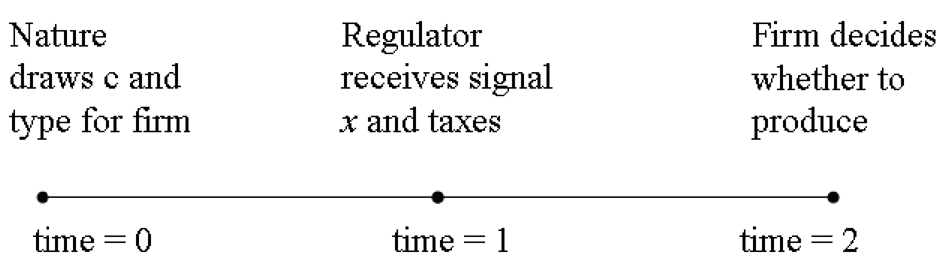


Figure 2. Timeline under Pigouvian taxes.

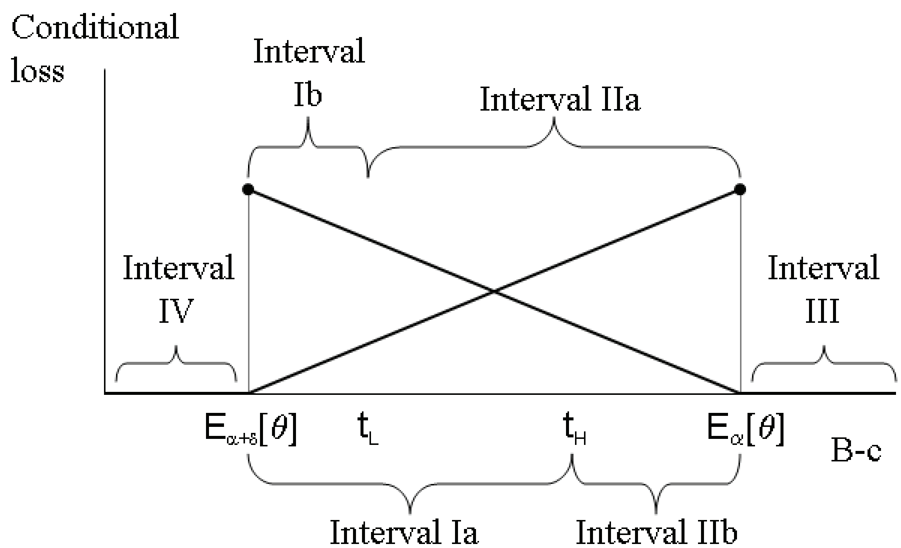


Figure 3. Losses from WFP across firm costs.

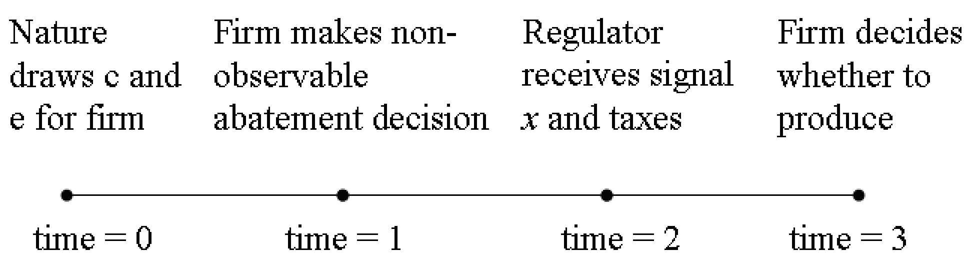
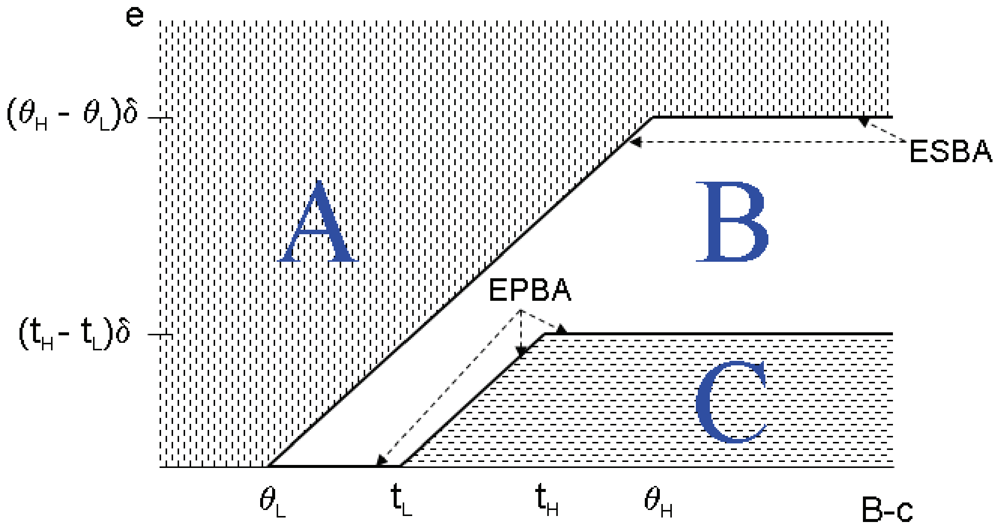


Figure 4. Timeline with opportunity to abate.



Area A=firms that should not and do not abate under noise and WFP
Area B=firms that should abate but do not abate under noise and WFP
Area C=firms that abate under noise and WFP

Figure 5. Set of abaters under noise and WFP.

Appendix

Demonstration that tax $t = E[\theta]$ is optimal under no conditioning information: If tax t is imposed such that $B - c < t$, then the firm shuts down and no economic surplus is generated. If tax t is imposed such that $B - c \geq t$, then production occurs. But social surplus could be positive or negative. The regulator's task is to choose t to maximize social surplus, or

$$\max_t \quad \int_{\underline{c}}^{B-t} \int_{\theta} (B - c - \theta) dF_c(c) dF_\theta(\theta). \tag{A1}$$

Rearranging provides

$$\max_t \quad \int_{\underline{c}}^{B-t} \int_{\theta} (B - c - \theta) dF_c(c) dF_\theta(\theta) = \max_t \quad \int_{\underline{c}}^{B-t} (B - c) dF_c(c) - E[\theta] \int_{\underline{c}}^{B-t} dF_c(c). \tag{A2}$$

The differential is

$$\left(E[\theta] - t \right) f_c(c = B - t) = 0, \tag{A3}$$

with solution $t = E[\theta]$ since $f_c(c = B - t) \neq 0$ on $c \in [\underline{c}, \overline{c}]$. To show that this is indeed the maximum, note that the second-order condition is $-f_c(c = B - t) +$ $\left(E[\theta] - t \right) df_c(c = B - t)/dt = -f_c(c = B - t) < 0$ when evaluated at $t = E[\theta]$. ■

Proof of Proposition 1. The proof is as in the demonstration above, except that the measure $F_\theta(\theta)$ is discrete and becomes a signal-conditioned measure. (A3), then, becomes $E_L[\theta] - t = 0$ under signal L and $E_H[\theta] - t = 0$ under signal H. ■

Alternative Proof of Proposition 1. From (12), (13), and (14),

$$\begin{aligned}
\mathcal{L} = {} & (1-\alpha-\delta)\phi\int_{c\in(B-t_H,B-E_{\alpha+\delta}[\theta]]}\big(B-c-E_{\alpha+\delta}[\theta]\big)dF_c(c) \\
& +(\alpha+\delta)\phi\int_{c\in(B-t_L,B-E_{\alpha+\delta}[\theta]]}\big(B-c-E_{\alpha+\delta}[\theta]\big)dF_c(c) \\
& +\alpha(1-\phi)\int_{c\in(B-E_{\alpha}[\theta],B-t_L]}\big(c+E_{\alpha}[\theta]-B\big)dF_c(c) \\
& +(1-\alpha)(1-\phi)\int_{c\in(B-E_{\alpha}[\theta],B-t_H]}\big(c+E_{\alpha}[\theta]-B\big)dF_c(c).
\end{aligned} \tag{A4}$$

Differentiate and set equal to 0:

$$\begin{aligned}
\frac{d\mathcal{L}}{dt_L} &= \big[(\alpha+\delta)\phi(t_L-E_{\alpha+\delta}[\theta])-\alpha(1-\phi)(E_{\alpha}[\theta]-t_L)\big]f_c(c=B-t_L)=0; \\
\frac{d\mathcal{L}}{dt_H} &= \big[(1-\alpha-\delta)\phi(t_H-E_{\alpha+\delta}[\theta])-(1-\alpha)(1-\phi)(E_{\alpha}[\theta]-t_H)\big]f_c(c=B-t_H)=0.
\end{aligned} \tag{A5}$$

Given positive support for c on $[\underline{c},\overline{c}]$, (A5) resolves to the tax levels as given in (6). Convexity at the solution to (A5) is immediate upon applying (A5) to second derivatives of (A5).

References

[1] D. Blackwell, Equivalent comparisons of experiments, Annals Math. Stat. 24 (2, June) (1953) 265–273.

[2] P. Bogetoft, H.B. Olesen, Incentives, information systems, and competition, Amer. J. Agr. Econ. 85 (1, February) (2003) 234–247.

[3] R. Cabe, J.A. Herriges, The regulation of non-point-source pollution under imperfect and asymmetric information, J. Environ. Econ. Manage. 22 (2, March) (1992) 134–146.

[4] A. Caplin, B. Nalebuff, Aggregation and imperfect competition: On the existence of equilibrium, Econometrica 59 (1, January) (1991) 25–59.

[5] L.G. Epstein, Decision making and the temporal resolution of uncertainty, Int. Econ. Rev. 21 (2, June) (1980) 269–283.

[6] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge MA, 1991.

[7] D.A. Hennessy, Information asymmetry as a reason for food industry vertical integration, Amer. J. Agr. Econ. 78 (4, November) (1996) 1034–1043.

[8] J.G. Klatt, A.P. Mallarino, J.A. Downing, J.A. Kopaska, D.J. Wittry, Soil phosphorus, management practices, and their relationship to phosphorus delivery in the Iowa Clear Lake agricultural watershed, J. Environ. Qual. 32 (6, November/December) (2003) 2140–2149.

[9] A.P. Mallarino, B.M. Stewart, J.L. Baker, J.D. Downing, J.E. Sawyer, Phosphorus indexing for cropland: Overview and basic concepts of the Iowa phosphorus index, J. Soil Water Conserv. 57 (6, November/December) (2002) 440–447.

[10] T.G. McGuire, C.G. Ruhm, Workplace drug abuse policy, J. Health Econ. 12 (1, April) (1993) 19–38.

[11] K. Millock, D. Sunding, D. Zilberman, Regulating pollution with endogenous monitoring, J. Environ. Econ. Manage. 44 (2, September) (2002) 221–241.

[12] K. Segerson, Uncertainty and incentives for nonpoint pollution control, J. Environ. Econ. Manage. 15 (1, March) (1988) 87–98.

[13] J.S. Shortle, R.D. Horan, The economics of nonpoint pollution control, J. Econ. Surveys 15 (3, July) (2001) 255–285.

[14] A.P. Xepapadeas, Observability and choice of instrument mix in the control of externalities, J. Pub. Econ. 56 (3, March) (1995) 485–498.