

FACULTY OF ECONOMICS AND
APPLIED ECONOMIC SCIENCES
CENTER FOR ECONOMIC STUDIES
ENERGY, TRANSPORT & ENVIRONMENT



KATHOLIEKE
UNIVERSITEIT
LEUVEN

WORKING PAPER SERIES
n° 2004-15

**The Relative Efficiency of Market-based
Environmental Policy Instruments with
Imperfect Compliance**

S. Rousseau (K.U.Leuven – CES)
S. Proost (K.U.Leuven – CES; CORE - UCLouvain)

November 2004



secretariat:

Isabelle Benoit
KULeuven-CES

Naamsestraat 69, B-3000 Leuven (Belgium)

tel: +32 (0) 16 32.66.33

fax: +32 (0) 16 32.69.10

e-mail: Isabelle.Benoit@econ.kuleuven.ac.be

<http://www.kuleuven.be/ete>

The Relative Efficiency of Market-based Environmental Policy Instruments with Imperfect Compliance¹

Sandra Rousseau and Stef Proost

*Centre for Economic Studies
K.U.Leuven
Naamsestraat 69
B-3000 Leuven
Belgium*

*E-mail: Sandra.Rousseau@econ.kuleuven.ac.be
Stef.Proost@econ.kuleuven.ac.be*

Abstract

This paper examines to what extent incomplete compliance of environmental regulation mitigates the distortions caused by pre-existing labour taxes. We study the relative cost efficiency of three market-based instruments: emission taxes, tradable permits and output taxes. In a first-best setting and given that monitoring and enforcement is costless, we find that the same utility levels can be reached with and without incomplete compliance. However, allowing for violations makes the policy instruments less effective. The nominal tax rate needs to be higher or the number of permits issued smaller, in order to obtain the required emission reduction. Including monitoring and enforcement aspects, and more specifically fines, into the model in a second-best setting, provides us with a new means of collecting tax revenues and of lessening existing tax distortions. We show that the relative position of grandfathered tradable permits vis-à-vis emission taxes improves considerably when allowing for incomplete compliance in a second-best setting.

Keywords: Environmental policy, instrument choice, monitoring and enforcement

¹ We would like to acknowledge the financial support of the DWTC research program - Indicators for sustainable development - contract nr. HL/DD/015 (*Law and economics of the Choice of Environmental Policy Instruments*). Moreover we would also like to thank Carole M. Billiet, Firouz Gahvari, Inge Mayeres Laurent Franckx, Anthony Heyes, Erik Schokkaert, Dirk Heremans, Thomas Ulen and Bruno De Borger for their useful comments and suggestions as well as two anonymous referees for comments on a previous version.

1. INTRODUCTION

When designing environmental regulation the government has to decide, among other things, which policy instrument is best suited for the job. This choice is determined, in part, by the cost efficiency of the instrument, the political ease with which it can be implemented, its distributional impact and the technological feasibility. In this paper we focus on the efficiency costs associated with environmental policy instruments and, more specifically, how these costs are influenced by incomplete compliance if pre-existing distortionary taxes are present in the economy.

Over the past decennia, it has become increasingly clear that firms do not always comply with existing regulation. Magat and Viscusi (1990), as an example, report an average level of non-compliance of 25 percent in the US pulp and paper industry between 1982 and 1985. Limited governmental monitoring activities are partly to blame for this phenomenon, as noted by Russell et al. (1986) and Harrington (1988). Moreover, fines for detected violations are often quite low. For example, the median administrative fine imposed by the US EPA in 1995 was \$4000, while the average fine was \$10181 and the maximum fine was \$125000 (Lear, 1998). In the US, less than 200 firms were fined in 1995.

This non-compliance by firms obviously makes environmental policies less effective. Therefore, the regulator should take monitoring and enforcement issues into consideration when selecting policy instruments. Sandmo (2002) has studied the impact of imperfect compliance on the ranking of policy instruments in a partial equilibrium context. We will focus on the general equilibrium context. Recent overviews of the monitoring and enforcement literature can be found in Cohen (2000), Polinsky and Shavell (2000) and Heyes (2001). The seminal paper of Becker in 1968 on the economics of crime was the start of rapid growth in both theoretical developments and empirical studies of monitoring and enforcement.

In recent years considerable attention has been paid to the interaction between environmental policies and pre-existing tax distortions caused by, for example, labour taxes. It was shown that environmental policy instruments can increase distortions caused by existing taxes. In fact, reducing pollution, no matter how, has almost always a hidden cost when tax distortions exist (Goulder, 1995 and Bovenberg and de Mooij, 1994).

Goulder et al. (1997) have used analytical and numerical general equilibrium models to explore the choice between pollution taxes and quotas in the presence of distortionary labour taxes. Schmutzler and Goulder (1997) studied the choice between emission taxes and output taxes under costly monitoring. They show how the costs of monitoring emissions or output, influence the efficiency of both instruments. They find that *“pure emissions taxes are usually not optimal with monitoring costs and pure output taxes are optimal under sufficiently high monitoring costs, sufficiently limited options for emission reduction by means other than output reduction, and sufficiently high substitutability of the output”*. However, Schmutzler and Goulder do not allow for incomplete compliance by the firms.

This paper extends the work of Goulder et al. (1999). In a second-best setting with pre-existing labour taxes, they use a general equilibrium model to examine the costs of achieving pollution reductions under a range of environmental policy instruments. They compare the overall efficiency impact of emission taxes, emission quota, fuel taxes, performance standards and mandated technologies. The ability to generate revenues appears to have a significant influence on the cost efficiency of environmental policy instruments. They find that *“prior taxes can eliminate the cost advantage of market-based instruments (emissions taxes and permits) over technology mandates or performance standards, particularly if the former policies fail to generate revenues and (fail to) use the revenues to finance cuts in the prior distortionary taxes”*. This failure to generate revenues is the main drawback of emission quota compared to an emission tax. Output taxes have a distinct disadvantage vis-à-vis emission taxes because they do not encourage firms to abate their emissions. Moreover, cost differences between instruments turn out to be highly sensitive to the extent of pollution abatement: for most policies, abatement costs converge to the same value only when the required emission reduction approaches 100 percent.

In a first-best setting, without distortionary taxation, and for a given actual reduction in emissions, the costs of the different instruments are the same with and without perfect compliance since we assume that monitoring and enforcement is costless. However, if the possibility of noncompliance by firms is incorporated, we find that a stricter policy is needed in order to reach a particular environmental goal. The policy instrument will be less effective and making the regulation more stringent compensates this.

Moreover, in a second-best setting, the ability to generate revenues and alleviate existing distortions appears to have a significant influence on the cost efficiency of environmental policy instruments. Including monitoring and enforcement aspects, and more specifically fines, into the model, provides us with a new means of collecting revenues for the government. We show that the relative position of grandfathered tradable permits vis-à-vis emission taxes improves considerably when allowing for incomplete compliance in a second-best setting.

Compared to Goulder et al. (1999) we limit the set of policy instruments. We only focus on market-based instruments: emission taxes, grandfathered tradable permits² and output taxes.

In section 2 the model is specified, including monitoring and enforcement issues. In section 3 we derive and compare the gross efficiency costs of marginal policy changes in an analytical model. In section 4 we describe the numerical model, which is a general equilibrium model for the American economy and addresses the air pollutant NO_x. This model is used to calculate the relative cost efficiency of different policy instruments for several levels of emission reductions in section 5. In section 6 we conclude.

2. THE MODEL

As indicated in the introduction, the model originates from the work of Goulder et al. (1999). We particularly focus on the monitoring and enforcement aspects associated with the firms' compliance decisions. Subsequently, we discuss the behaviour of households, firms and government.

² Grandfathered tradable permits are distributed freely among firms; rather than auctioning or selling the tradable permits to firms. The grandfathered, or free, permits are often distributed according to historical emission levels. After the initial distribution, permits can be traded amongst firms and a market price for permits will occur.

2.1 Households

In this static, general equilibrium model a representative household derives utility from the consumption of a polluting good D , a clean good N and from leisure. Leisure is equal to the household's time endowment T less labour supply L . Consumer utility is affected negatively by the environmental damage caused by the emissions E resulting from the production of the polluting good. The household's utility function is:

$$U = u(D, N, T - L) - \phi(E) \quad (1)$$

where $u(\cdot)$ is a quasi-concave utility function for non-environmental goods. The function $\phi(\cdot)$ represents the disutility from emissions and is weakly convex. The separability restriction in equation (1) implies that demand for D and N and supply of labour do not vary with changes in E .³ Moreover, we only consider efficiency aspects and ignore distributional concerns.

The household's budget constraint is:

$$p_D D + N = (1 - t_L)L + G + \Pi \quad (2)$$

where p_D is the demand price of D . This price is equal to unity in the absence of regulation. The price of the non-polluting good p_N is constant, equal to unity and not affected by the environmental policy. The non-polluting good N is the numeraire in the theoretical model. The firms' profits Π are redistributed to the households. The households also receive a lump-sum transfer G from the government.

³ Relaxing this assumption would complicate the tax-interaction effect discussed below. If, compared with consumption, leisure is a relatively strong (weak) substitute for environmental quality, then this effect is weakened (strengthened). There is little empirical evidence on the relative ease of substitution between leisure, overall consumption, and environmental quality. Under these circumstances it seems reasonable to assume separability, which implies that changes in environmental quality do not affect the relative attractiveness of consumption and leisure. (Goulder et al., 1999)

For a model without this separability assumption see Mayeres and Proost (1997). In their model the demand for production goods and the supply of labour depend on the changes in environmental quality and feedback effects occur. Mayeres and Proost look at a congestion type of externality and derive optimal tax and public investment rules.

The households choose D , N and L in order to maximise their utility subject to the budget constraint, taking environmental damages as given. From the resulting first-order conditions the uncompensated demand and labour supply functions are obtained:

$$D(p_D, 1-t_L, G+\Pi), N(p_D, 1-t_L, G+\Pi) \quad \text{and} \quad L(p_D, 1-t_L, G+\Pi) \quad (3)$$

Substituting these equations into (1) gives the indirect utility function:

$$V = v(p_D, 1-t_L, G+\Pi) - \phi(E) \quad (4)$$

2.2 Firms

The competitive firms maximise profits and are risk neutral. Household labour, which is the only factor of production, is used to produce goods D and N . The marginal product of labour is assumed to be constant in each industry. Output is normalised such that the marginal products and the wage rate equal unity. Firms decide on the production of D and N , on abatement expenditures and on the amount of taxes paid (or permits bought). Due to non-compliance by firms, the reported emissions on which taxes are paid (or emissions covered by permits) do not equal the total actual emissions in the economy. Economy-wide emissions E equal

$$E = (e_o - a)D = eD$$

with e total emissions per unit of D , which equal the difference between the baseline emissions e_o per unit of output and the reduction a in per-unit emissions due to abatement. Firms can reduce their emissions per unit of output by using abatement equipment that is produced directly from labour. The total cost C of abatement is:

$$C = c(a).D \quad (5)$$

where $c(a)$ is a convex function representing the per-unit cost of abatement activity.

Thus, total emissions fall as a result of reduced production of the polluting good D (the output-substitution effect) and increased abatement activity (the abatement effect). In the numerical model we will incorporate a third way of emission reduction, namely via input-

substitution. This means that firms will have the option to change their input mix and use less of the most polluting input.

We now concentrate on the implications of noncompliant behaviour by firms. We assume that firms can be noncompliant in a continuous way. The regulator anticipates the firms' inclination to cheat and will therefore pursue a monitoring and enforcement policy in order to deter them. We consider three different instruments: an emission tax, tradable permits and an output tax.

2.2.1 Emission tax

Considering an emission tax t_E , firms have to choose the fraction θ_E of their emissions they wish to report to the government. It is obvious that reported emissions never exceed actual emissions if firms behave rationally ($0 \leq \theta_E \leq 1$).

Every firm is inspected with a fixed probability \bar{p}_E regardless of its compliance status. The environmental inspection agency audits a certain percentage of all firms at random. A noncompliant firm, however, faces an additional probability of inspection proportional to the size of its violation. This variable inspection probability equals $\eta_E [1 - \theta_E]$. It reflects the regulatory practice of following up complaints. These complaints can be issued by neighbours, interest groups or civil servants. We assume that these complaints are positively correlated to the seriousness of the infraction. The total probability of detection

$$\begin{aligned}
 p_{det,E} \text{ for a firm is:} &= \bar{p}_E + \eta_E \left[\frac{E - \theta_E E}{E} \right] \\
 &= \bar{p}_E + \eta_E [1 - \theta_E]
 \end{aligned} \tag{6}$$

We assume that performing inspections⁴ and levying fines are costless. Moreover, measurement and judicial errors are absent in our model. Consequently, all violators are caught if they are inspected.

⁴ We normalise the inspection costs to zero because we do not have reliable estimates of their levels. Later we will perform a sensitivity analysis and perform the empirical exercise with strictly positive inspection costs.

Once the violator is caught, the firm has to pay a penalty rate π on the evaded taxes. This rate is assumed to exceed one since firms will have to pay at least the tax rate on unreported emissions. The total payment is called the fine $F_E (= f_E D)$ and is equal to

$$F_E = [E - \theta_E E] t_E \pi_E \quad \text{with } \pi_E \geq 1$$

2.2.2 Tradable permits

In a system of grandfathered tradable permits⁵, the government determines an overall acceptable level of emissions and issues permits⁶ correspondingly. These permits are distributed freely among firms proportional to their baseline emissions. The government, therefore, does not receive any revenues from issuing the permits.

Firms can obtain more permits by trading. However, since all firms are homogeneous, no trade will occur. This policy can be interpreted as a virtual tax t_E^v on emissions, where the ‘revenues’ are rebated in a lump-sum fashion⁷. These revenues represent the rents associated with grandfathered permits. The limited supply of emission permits imply a reduced output, which gives rise to economic rents.

However, there can be reasons for non-compliance even if monitoring is costless. If the inspection agency is ineffective due to measurement errors (see Polinsky and Shavell, 2000) or uncertainty about the legal rules (Craswell and Calfee, 1986), firms can choose to be non-compliant. Including these possibilities into our model would lead us to differentiate between the probability of inspection, the probability of detection and the probability of punishment. We choose not to make this distinction in order to keep the computations tractable.

⁵ Instead of distributing the permits freely, or grandfathering, the government could also decide to auction the permits. Such an auction would generate revenues for the government. In our model, without transaction costs, auctioned tradable permits are equivalent to an emission tax. We, therefore, decide to model only grandfathered tradable permits.

⁶ We assumed that one permit allows a firm to produce one unit of emissions.

⁷ We model the permits as an emission tax with lump-sum redistribution of the tax revenues (but not the fine revenues). The only difference between permits and the emission tax is that government revenues from emission taxes are absent for permits.

Initially firms will claim that the emissions produced E equal the number $\theta_p E$ of permits they own (*with* $0 \leq \theta_p \leq 1$). However, this is not necessarily true. With a certain probability of detection, depending on a fixed inspection frequency and the seriousness of the violation, firms are caught lying. This is:

$$p_{\text{det},p} = \bar{p}_p + \eta_p \left[\frac{E - \theta_p E}{E} \right] = \bar{p}_p + \eta_p [1 - \theta_p]$$

Violators then have to buy extra permits and pay a fine, proportional to the size of the violation.

$$F_p = [E - \theta_p E] t_E^v \pi_p \quad \text{with} \quad \pi_p \geq 1$$

We assume that the monitoring and enforcement parameters are equal across the different policy instruments. Consequently, we will drop the indices and use \bar{p}, η and π from now on.

2.2.3 Output tax

The government can also decide to tax the polluting good D directly at a tax rate t_D . Under this regulatory scheme, firms will never abate their emissions. The output tax is, after all, levied on the amount of output D and this is independent of the abatement expenditures by firms. Firms will initially pay taxes on reported output $\theta_D D$ with $0 \leq \theta_D \leq 1$. With a certain probability of detection, depending on the level of the violation, firms are caught.

$$p_{\text{det},D} = \bar{p} + \eta \left(\frac{D - \theta_D D}{D} \right) = \bar{p} + \eta (1 - \theta_D)$$

Detected violators have to pay the overdue taxes increased with penalty payments or

$$F_D = (D - \theta_D D) t_D \pi \quad \text{with} \quad \pi \geq 1.$$

2.3 Government

The government levies a proportional tax of t_L on labour earnings, regulates emissions and provides a fixed nominal lump-sum transfer G to the households. The government budget is assumed to be balanced. Adjusting the labour tax⁸ t_L offsets any revenue consequences from environmental policies. In the numerical model we will assume that rent or capital income is taxed at the same rate as labour income.

The expression for the government budget is:

$$t_E \theta_E E + t_D \theta_D D + t_L L + p_{\text{det},m} F_m = G \quad (7)$$

with $m \in \{E, p, D\}$ depending on the policy instrument used by the government. We assume that the government uses only one instrument at a time.

We assume that there are no costs attached to performing inspections and levying fines. We return to this assumption later. Remark also that the government, under a grandfathered permit system, does not receive any income from the environmental policy except for the fine payments.

3. GROSS EFFICIENCY COST OF THE DIFFERENT ENVIRONMENTAL INSTRUMENTS

In this analysis we focus on the gross efficiency cost of different environmental policies. This cost is the monetary equivalent of the loss in utility. It is a gross concept in that it does not include the change in environmental damages. Next we analyse the gross efficiency cost of three environmental policies with imperfect compliance and obtain the results of Goulder et al. (1999) as a special case.

⁸ In a one-consumer setting, the optimal tax structure is to use only the lump-sum tax G and to have $t_L = 0$ or to use a profit tax in the presence of pure profit. In this case the Goulder et al. problem becomes trivial because we can return to a first-best if Pigouvian taxes can be used (Mayeres and Proost (1997)). However, this simple framework will allow us to better isolate the effects of the environmental policy.

3.1 The gross efficiency cost of emission taxes

Firstly, we discuss the impact of emission taxes on the firms. We concentrate on the changes in actual emissions and in the fraction of emissions reported. Secondly, we derive the gross efficiency costs of emission taxes, tradable permits and output taxes.

3.1.1 The impact of the emission tax on the firms

Let us consider a revenue-neutral tax t_E on emissions. The firm's profit per unit of D is:

$$p_D - \{ 1 + c(a) + t_E \theta_E e + p_{\text{det},E} f_E \} \quad (8)$$

with

$$p_{\text{det},E} f_E = [\bar{p} + \eta[1 - \theta_E]] [e - \theta_E e] t_E \pi$$

$\theta_E e$ = reported emissions per unit of output

π = penalty payments on overdue taxes

\bar{p} = fixed inspection frequency

η = variable inspection parameter

and, since we work in competitive markets, profits are zero in equilibrium.

We consider the case in which both the emission tax and the penalty are positive and finite. The other scenarios are discussed in appendix A. The firms will never report zero emissions because then the firm would always have to pay the complete tax plus the penalty. It could always do better by reporting truthfully because then it would not have to pay the fine. However, there exists an internal solution that is even better, which we will derive now.

The firm's problem can be defined as choosing actual emissions e and fraction of emissions reported θ_E to minimise unit costs of environmental compliance. This can be

broken into two stages: firstly, the optimal choice of θ_E for any given e and secondly, the optimal choice of actual emissions e for given optimal θ_E .⁹

First, for given e , the firm chooses θ_E to minimise the proportion of tax it actually pays:

$$\rho(\pi, \bar{p}, \eta) \equiv \min_{\theta_E} \left\{ \theta_E + \pi [1 - \theta_E] [\bar{p} + \eta [1 - \theta_E]] \right\} \quad (9)$$

This gives
$$\theta_E^*(\pi, \bar{p}, \eta) = \left(1 - \frac{1 - \pi \bar{p}}{2\pi\eta} \right) \quad \text{for } f_E, p_{\det, E} \neq 0 \quad (10)$$

The proportion of tax reported is increasing in the penalty $\left(\frac{d\theta_E^*}{d\pi} = \frac{1}{2\pi^2\eta} > 0 \right)$, the fixed

inspection frequency $\left(\frac{d\theta_E^*}{d\bar{p}} = \frac{1}{2\eta} > 0 \right)$ and the variable inspection parameter

$\left(\frac{d\theta_E^*}{d\eta} = \frac{1 - \bar{p}\pi}{2\pi\eta^2} > 0 \text{ if } \bar{p}\pi \leq 1 \right)$. Rationality dictates that $\theta_E^* \leq 1$; or firms never report more

than their actual emissions. This implies that the condition $\bar{p}\pi \leq 1$ must hold¹⁰.

The resulting proportion of the tax actually paid is:

$$\rho^*(\pi, \bar{p}, \eta) = 1 - \frac{(1 - \bar{p}\pi)^2}{4\pi\eta} \quad (11)$$

We find that the proportion of tax actually paid is increasing in the penalty π

$\left(\frac{d\rho^*}{d\pi} = \frac{1 - \bar{p}^2\pi^2}{4\pi^2\eta} \geq 0 \right)$, the fixed inspection frequency \bar{p} $\left(\frac{d\rho^*}{d\bar{p}} = \frac{1 - \bar{p}\pi}{2\eta} \geq 0 \right)$ and the

⁹ We implicitly assume that the firm first decides on the number of emission and only afterwards on the amount of reported emissions. We solve this problem by backward induction.

¹⁰ If this condition does not hold and $\bar{p}\pi > 1$, there is overdeterrence and the expected fine of dishonest reporting is larger than the cost of paying all due taxes. All firms, therefore, report their taxes honestly. This same result, truthful reporting, could also have been obtained, and at lesser cost, by setting $\bar{p}\pi = 1$.

variable inspection parameter $\eta \left(\frac{d\rho^*}{d\eta} = \frac{(1-\bar{p}\pi)^2}{4\pi\eta^2} \geq 0 \right)$. These signs hold since rationality implies $\bar{p}\pi \leq 1$.

We can now define the effective tax rate as a fraction of the nominal tax rate, $\tau_E = \rho^*(\pi, \bar{p}, \eta)t_E$.

Next, in a second step, the firm chooses the actual emissions e to minimise its environmental costs:

$$\gamma(\tau_E) \equiv \min_e [c(e_o - e) + \tau_E e] \quad (12)$$

This gives us the following expression:

$$c'(e_o - e^*) = \tau_E = \rho^* t_E \quad (13)$$

For each firm, abatement activity occurs until the marginal abatement cost per unit of D equals the effective emission tax rate. If we want to find out how the optimal firm emissions depend on the effective tax rate τ_E , we have $\frac{de^*}{d\tau_E} = -\frac{1}{\frac{dc}{de} \frac{de}{d\tau_E}}$ and

therefore, $\gamma'(\tau_E) = e^*$.

The profit per unit of output is equal to $p_D - \{1 + c(e_o - e^*) + \tau_E e^*\}$ and perfect competition then ensures that $p_D = 1 + \gamma(\tau_E)$.

3.1.2 Derivation of the gross efficiency costs

We now consider an incremental, revenue-neutral increase in the nominal emission tax rate t_E , starting from a strictly positive tax rate. The effect of this policy reform on the product price is:

$$\frac{dp_D}{dt_E} = \rho^*(\pi, \bar{p}, \eta)e \quad (14)$$

Revenues from the emission tax will be employed to finance cuts in the distortionary labour tax t_L .

Using the requirement of balancing the governmental budget and deriving the efficiency cost of an environmental increase in the emission tax (see appendix B for more details), we obtain:

$$\begin{aligned} -\frac{1}{\lambda} \frac{dV}{dt_E} = & \rho^*(\pi, \bar{p}, \eta)t_E \left[-\frac{dE}{dt_E} \right] - M \rho^*(\pi, \bar{p}, \eta) \left[E + t_E \frac{dE}{dt_E} \right] \\ & + [1 + M] t_L \left[-\frac{\partial L}{\partial p_D} \right] \frac{dp_D}{dt_E} \end{aligned} \quad (15)$$

where λ is the marginal utility of income and M is defined as:

$$M \equiv \frac{-t_L \frac{\partial L}{\partial t_L}}{L + t_L \frac{\partial L}{\partial t_L}} \quad (16)$$

This is the (partial equilibrium) efficiency cost from raising an additional dollar of labour tax revenue. The numerator is the efficiency loss from an incremental increase in t_L . This equals the wedge between the gross wage (equal to the value of the marginal product of labour) and the net wage (equal to the marginal social cost of labour in terms of foregone leisure), multiplied by the reduction in labour supply. The denominator is the marginal labour tax revenue (from differentiating $t_L L$).

We find, using $E = [e_0 - a]D$, that:

$$\frac{dE}{dt_E} = -\frac{da}{dt_E} D + [e_0 - a] \frac{dD}{dt_E} \quad (17)$$

Substituting (17) in (15), using (14), gives us:

$$\begin{aligned}
-\frac{1}{\lambda} \frac{dv}{dt_E} = & \underbrace{\rho^*(\pi, \bar{p}, \eta) t_E \frac{da}{dt_E} D}_{\text{abatement}} + \underbrace{\rho^*(\pi, \bar{p}, \eta) t_E e \left[-\frac{dD}{dt_E} \right]}_{\text{output-substitution}} \underbrace{-M \rho^*(\pi, \bar{p}, \eta) \left[E + t_E \frac{dE}{dt_E} \right]}_{\text{revenue-recycling}} \quad (18) \\
& + \underbrace{[1+M] \rho^*(\pi, \bar{p}, \eta) t_L e \left[-\frac{\partial L}{\partial p_D} \right]}_{\text{tax interaction}}
\end{aligned}$$

As in Goulder et al. (1999) we find that the gross cost of an increase in the emission tax, recycled via a decrease in labour taxes, can be decomposed into four effects. The reduction in emissions is achieved via a combination of two effects: the reduction of emissions per unit of output (abatement effect) and the substitution away from the consumption of D (output-substitution effect). These two effects are called the primary costs. In a first-best setting, the relative cost-effectiveness of different policies can be explained fully in terms of differences in primary costs.

In a second-best setting, with distortionary taxes, two additional cost terms come into play. The first term is the efficiency gain from the (marginal) revenue-recycling effect. This is the product of the marginal excess burden of taxation and the marginal revenues (if any) from the policy. It represents efficiency gains associated with using these revenues to finance cuts in distortionary taxes.

The second extra term is the efficiency loss from the tax interaction effect. This effect has two components. First, the new policy can increase the price of D , implying an increase in the cost of consumption and thus a reduction in the real wage. This reduces labour supply and produces a marginal efficiency loss that equals the tax wedge between the gross and the net wage multiplied by the reduction in labour supply. In addition, the reduction in labour supply contributes to a reduction in tax revenues, which has an efficiency cost of M times the lost tax revenues, equal to the change in labour supply times the labour tax rate.

Incorporating incomplete compliance in the model means that the effective tax rate τ_E equals the nominal tax rate t_E times the factor $\rho^*(\pi, \bar{p}, \eta) \leq 1$. The effective tax rate is, therefore, lower than the nominal one. Some firms can get away with only paying taxes on their reported emissions and not on their actual emissions. Since the effective tax rate is lower, all the different efficiency effects will also be lower. The four effects are deflated

with the same factor $\rho^*(\pi, \bar{p}, \eta)$. Therefore, we can say that the gross efficiency cost of a marginal increase in the nominal emission tax is lower in the model with monitoring and enforcement. However, this does not mean that utility is higher for the model with partial compliance. Remember that we did not take the environmental effects of the policy into account.¹¹ For a given level of the emission tax, it is obvious that the environmental quality is worse in the partial compliance case. In order to reach the same environmental emission reduction as with full compliance, one needs to set a comparatively higher tax rate.

When π approaches infinity¹², there is full compliance and the factor $\rho^*(\pi, \bar{p}, \eta)$ equals one. The results of Goulder et al. are a special case of our results. Note that, for a given actual reduction in emissions, the costs of the emission tax in this paper and in Goulder et al. (1999) are equivalent, except for the resource costs involved in inspection and monitoring.

3.2 The gross efficiency cost of tradable permits

Remember that a grandfathered tradable permit system can be interpreted as a virtual tax t_E^v on emissions, where the ‘revenues’ are rebated in a lump-sum fashion. We analyse the gross cost of a decrease in the number of permits that is distributed via an increase in the virtual tax t_E^v .

The profit for the firm per unit of D is:

¹¹ In order to include environmental effects, we would need to value them. Determining the consumers’ willingness-to-pay for environmental improvement is a difficult and complicated exercise; see, for example, Mitchell and Carson (1989). The valuation of the environmental quality will influence the benefits associated with an environmental policy. Moreover, it is often difficult to assess the impact of increases in emissions on overall environmental quality since this relationship is influenced by, for example, the timing of the emissions and the characteristics of the receiving environmental medium.

¹² In reality we do not observe infinitely high fines since the wealth of individuals and firms is limited. For an analysis of fines when wealth varies among individuals see Polinsky and Shavell, 1991.

$$p_D - \left\{ 1 + c(a) + t_E^v \theta_p e + \left[\bar{p} + \eta [1 - \theta_p] \right] \left[e - \theta_p e \right] t_E^v \pi \right\} \quad (19)$$

The optimal number of permits for a firm, as a fraction θ_p of actual emissions, is:

$$\theta_p^*(\pi, \bar{p}, \eta) = \left(1 - \frac{1 - \pi \bar{p}}{2\pi\eta} \right) \quad \text{for } f_p, p_{\det, p} \neq 0 \quad (20)$$

Consequently, the optimal emissions are determined by:

$$\begin{aligned} c'(e_o - e^*) &= \tau_E^v = \rho^v t_E^v \\ \text{with } \rho^v &= 1 - \frac{(1 - \bar{p}\pi)^2}{4\pi\eta} = \rho^* \\ \tau_E^v &= \text{effective virtual tax rate} \end{aligned} \quad (21)$$

The effect of an incremental increase in the virtual tax can be derived in a similar manner as for the emission tax. This gives:

$$\begin{aligned} -\frac{1}{\lambda} \frac{dv}{dt_E^v} &= \underbrace{\rho^* t_E^v \left[\frac{da}{dt_E^v} \right]}_{\text{abatement}} D + \underbrace{\rho^* t_E^v e \left[-\frac{dD}{dt_E^v} \right]}_{\text{output-substitution}} - \underbrace{[\rho^* - \theta_p] M \left[E + t_E^v \frac{dE}{dt_E^v} \right]}_{\text{revenue-recycling}} \\ &\quad + \underbrace{[1 + M] \rho^* e t_L \left[-\frac{\partial L}{\partial p_D} \right]}_{\text{tax interaction}} \end{aligned} \quad (22)$$

Compared to the emission tax case, we have the same abatement, output-substitution and tax-interaction effect. The revenue-recycling effect, however, will be smaller. The existence of a revenue-recycling effect contrasts with Goulder et al.'s (1999) results. They found no such effect for tradable permits because of the full compliance assumption. Under incomplete compliance, the collected fines generate revenues for the government. That is why we observe a revenue-recycling effect. This effect reduces the disadvantage of grandfathered tradable permits vis-à-vis emission taxes in a second-best world. Another interpretation of this result is the observation that the fine has some characteristics of a (non-linear) tax, albeit one that is imposed only with a certain probability (Sandmo, 2002).

3.3 The gross efficiency cost of an output tax

We now consider the output tax t_D on the polluting output D . Firms have no incentive to invest in abatement technologies under this regulatory policy. Only the quantity produced of good D has an effect on the taxes paid.

The profit for the firm per unit of D is:

$$p_D - \left\{ 1 + c(a) + t_D \theta_D e + [\bar{p} + \eta[1 - \theta_D]] [e - \theta_D e] t_D \pi \right\} \quad (23)$$

The optimal output to report for a firm, as a fraction θ_D of actual production, is:

$$\theta_D^*(\pi, \bar{p}, \eta) = \left(1 - \frac{1 - \pi \bar{p}}{2\pi\eta} \right) \quad \text{for } f_D, p_{\det, D} \neq 0 \quad (24)$$

Consequently, the optimal emissions are determined by:

$$\begin{aligned} c'(e_o - e^*) &= \tau_D = \rho^D t_D \\ \text{with } \rho^D &= 1 - \frac{(1 - \bar{p}\pi)^2}{4\pi\eta} = \rho^* \\ &= \text{proportion of output tax actually paid} \\ \tau_D &= \text{effective output tax} \end{aligned} \quad (25)$$

The gross efficiency cost of a revenue-neutral increase in the output tax t_D is equal to:

$$-\frac{1}{\lambda} \frac{dv}{dt_D} = \underbrace{\rho^* \left[-\frac{dD}{dt_D} \right] t_D}_{\text{output-substitution}} \underbrace{- M \rho^* \left[D + t_D \frac{dD}{dt_D} \right]}_{\text{revenue-recycling}} + \underbrace{[1 + M] \rho^* t_L \left[-\frac{\partial L}{\partial p_D} \right]}_{\text{tax interaction}} \quad (26)$$

The abatement effect is absent. The output tax does not induce firms to abate their emissions. However, emissions of the hazardous pollutant decrease by the reduction of the output of good D . In a setting of incomplete compliance, this output-substitution effect is smaller than under complete compliance. The same holds for the revenue-recycling and tax interaction effects. Both are multiplied by the same factor $\rho^*(\pi, \bar{p}, \eta)$.

4. NUMERICAL MODEL

Next we illustrate the results of the theoretical analysis with an empirical exercise. The main objective of this exercise is not to model the American economy as realistically as possible. We want to highlight the impact of incomplete compliance on the relative cost efficiency of market-based instruments. Therefore, we chose to use the same model and data as Goulder et al. (1999), who treated the perfect compliance case. This facilitates the comparison and interpretation of the results.

Subsequently, we describe the model. Some additional assumptions with respect to the theoretical model are made.

4.1 Description of the model

We use a general equilibrium model for the American economy, calibrated to the 1990 situation. The environmental pollution problem addressed is the air pollutant NO_x .

Deviating from the theoretical model, we incorporate intermediate inputs in the production model. This yields a new channel for emission reduction: the input-substitution effect. Emissions can be reduced not only by abatement and output substitution but also by input substitution. This means that the firms can alter the mix of intermediate inputs and use less of the polluting input.

We distinguish two different intermediate goods: a polluting (D) and a clean (N) intermediate good. In our application the polluting good can be thought of as being energy. There are two final consumption goods: CD represents output from industries that use D more intensively and CN represents output from industries that use N more intensively. The production relationships between different commodities are:

- Final goods CD and CN are produced using goods D and N
- Intermediate goods D and N are produced using labour L and goods D and N .

The structure of the numerical model is directly based on the previously discussed theoretical model. Labour is the numeraire.

We assume that the representative household has a nested constant-elasticity-of-substitution (CES¹³) utility function (the definitions of the variables and parameters can be found in appendix C):

$$Utility = U(leis, CD, CN, E) = \left(\alpha_{leis} leis^{\frac{\sigma_u-1}{\sigma_u}} + \alpha_C C^{\frac{\sigma_u-1}{\sigma_u}} \right)^{\frac{\sigma_u}{\sigma_u-1}} - \phi(E) \quad (27)$$

where *leis* depicts leisure, the σ 's represent different elasticities of substitution, the α 's are distribution parameters and *C* is composite consumption.

The household maximises utility with respect to the budget and time constraint:

$$p_{CD}CD + p_{CN}CN = p_L [1 - t_L]L + \Pi_{tot} [1 - t_R] + p_C G \quad (28)$$

$$T = \text{total time endowment} = L + leis \quad (29)$$

with $\Pi_{tot} [1 - t_R]$ the total after-tax rent or capital income. The tax rate on rent income t_R equals the labour tax t_L .

We now consider the production side of the economy. A CES-form is used for the production functions in all industries $j \in \{D, N, CD, CN\}$:

$$X_j = \delta \left[\sum_i \alpha_{ij} X_{ij}^{\frac{\sigma_j-1}{\sigma_j}} \right]^{\frac{\sigma_j}{\sigma_j-1}} \quad (30)$$

We assume a constant-returns-to-scale production technology. Labour and rent income are taxed at the same rate ($t_L = t_R$)¹⁴. We work with linear marginal abatement costs.

Firms choose profit-maximising production X_j and abatement A_j subject to the constraints imposed by environmental regulation and taking input prices pr_i and output prices p_j as

¹³ Further reading on CES functions can be found in Keller (1976).

¹⁴ In the numerical exercise both tax rates are reduced if we have extra revenue from the environmental policy. It would also be possible to adapt only the labour tax rate. This would not change the results for the instruments with zero profit for the firms. For the other instruments we expect some small changes.

given. Profits equal the value of output minus expenditures on inputs, labour and abatement, less any tax and fine payments. This gives the following expression:

$$\Pi_j = [p_j - t_j \theta_{X_j}] X_j - \sum_i p r_i X_{ij} - t_E \theta_{E_j} E_j - A_j - p_{\text{det}j} F_j \quad (31)$$

The resulting emissions per industry are:

$$E_j = \beta_D X_{Dj} \left[1 - \alpha_e \left[\frac{A_j}{\beta_D X_{Dj}} \right]^\chi \right] \quad (32)$$

with β_D representing the pollution content of good D and parameters α_e and χ describing the emission abatement technology¹⁵.

To obtain a general equilibrium, supply must equal demand for all produced goods, government revenue must equal government transfer payment, and pollution emissions must equal a specified target. Since production and abatement functions are linearly homogeneous, the supply of each good is perfectly elastic at given factor prices and tax rates. Under these conditions we can reduce the set of equilibrium conditions to three equations:

$$\text{aggregate labour demand equals aggregate supply}^{16}: L = \sum_j X_{Lj} + \sum_j A_j,$$

government revenue equals expenditures:

$$REV \equiv p_L t_L L + t_R \sum_j \Pi_j + \sum_j t_j \theta_{X_j} X_j + \sum_j t_E \theta_{E_j} E_j + \sum_j p_{\text{det}j} F_j = p_C G$$

$$\text{and pollution levels equal the target level: } \sum_j E_j = [1 - \text{reduc}] \overline{E_{\text{tot}}}^{17}.$$

¹⁵ In the numerical exercise we choose $\chi = 0.5$, which implies linear marginal abatement costs.

¹⁶ We assume that one unit of labour is needed to produce one unit of abatement.

¹⁷ In the GAMS model we only use the government budget and emission target conditions. By Walras' law, if these two conditions are satisfied, the third condition must also hold. We have used this third condition, labour market clearance, as a check on the optimality of the obtained solution.

4.2 Data

In Table 1 we summarise the benchmark data set of Goulder et al. (1999), which represents the United States' economy in 1990. Production data were obtained from the Commerce Department Bureau of Economic Analysis.

The polluting intermediate good D covers fossil fuels (oil, coal and natural gas), while the clean good N includes all other intermediates. The final good CD is a composite of the consumer goods whose production involves intensive use of fossil fuels (consumer utilities, motor vehicles and gasoline), while the good CN embraces all other final goods.

	D	N	CD	CN	Leisure time	Total output value
D	91 441	111 843	156 881	6 264		366 429
N	88 073	4 741 098	464 160	2 670 486		7 963 817
Labour	186 915	3 110 876			1 832 106	5 129 897
Total output value	366 429	7 963 617	621 041	2 676 750		
Emissions (millions of Kilograms)	23 000					

Table 1: Input-output flows (in millions of 1990 dollars per year except when otherwise noted)
Source: Goulder et al. (1999)

The parameter values used in the model can be found in appendix C. The distribution parameters α for production and utility functions were calibrated in GAMS based on the assumed elasticities of substitution and the restriction that the benchmark data must be replicated in the absence of a new environmental policy.

Although we try to derive general relationships, we must commit to certain parameter values in running the model. The central case values for pollution-related parameters are identical to Goulder et al. (1999). Pollution takes place every time a unit of D is used in the production process.

We also had to determine the monitoring and enforcement strategy. Accurate and realistic data are very difficult to obtain. From Belgian data we know that, yearly, ten percent of all firms are inspected ($\bar{p} = 0.1$), and we assume that the variable inspection parameter η equals 0.5. Furthermore, court experience tells us that violators have to pay twice the amount of evaded taxes ($\pi = 2$) in Belgium. This will be our reference point. Moreover, in this model monitoring and enforcement are costless. In section 6 we discuss the implication of these assumptions and perform a sensitivity analysis.

5. COMPARING THE GROSS COST OF DIFFERENT POLICY INSTRUMENTS

We now determine the impact of different policies that all produce the same net emission reduction. Using initial prices we define the welfare gain as that sum of money which the households would have accepted in the initial position as equivalent to the impact of the reform, and we call this the equivalent gain, or equivalent loss, (King – 1983). It is defined by¹⁸:

$$EG = \text{equivalent gain} = EF(p^R, U^N) - EF(p^R, U^R) \quad (33)$$

with EF being the expenditure function. In the reference situation we assume that there are no emission reductions and no environmental policy. The elements of the reference price vector p^R are all assumed to be equal to unity. Therefore, we can rewrite the equivalent gain, using $EF(p^R, U^N) = \frac{U^N}{p_u}$ ¹⁹, as follows²⁰:

$$EG = U^N - U^R. \quad (34)$$

In each new scenario we impose a certain level of emission reduction, e.g. 10 percent. For every policy instrument we then determine the optimal size so that the required emission reduction is achieved. With each scenario a new utility level is associated. It is obvious that after the introduction of the environmental regulation utility will be lower. This holds

¹⁸ with $EF(\cdot)$ = expenditure function, R = the reference value and N = the new value

¹⁹ with p_u the price of utility or $p_u U = p_L [1 - t_L] L + \Pi_{tot} [1 - t_R] + p_C G$.

²⁰ This holds for the specific functional forms we use in our model, where the marginal utility of income equals p_u , but this is not a general result.

because we ignore any utility effect of improved environmental quality (i.e. leaving out the benefit of the policy, which gives us the gross cost, rather than the net welfare change).

First we compute the efficiency losses in a first-best setting and only later we look at the second best setting. This allows us to distinguish between the impact of including distortionary taxes and the impact of incomplete compliance. To facilitate comparisons across instruments we take the emission tax as a reference point. We will compare the equivalent loss of all instruments with that of the emission tax in each scenario.

5.1 First-best setting: gross costs with and without perfect compliance

We first consider the equivalent gain in a first-best setting²¹ ($t_L=0$) with perfect compliance. This exercise mimics the one performed by Goulder et al. (1999). Only primary costs will apply. The losses (or costs) under the different policy instruments are shown in figure 1. The differences across policies are expressed as the ratio of total losses of the policy in question to total costs under the emission tax. Consequently, the curve for the emission tax is constant at unity.

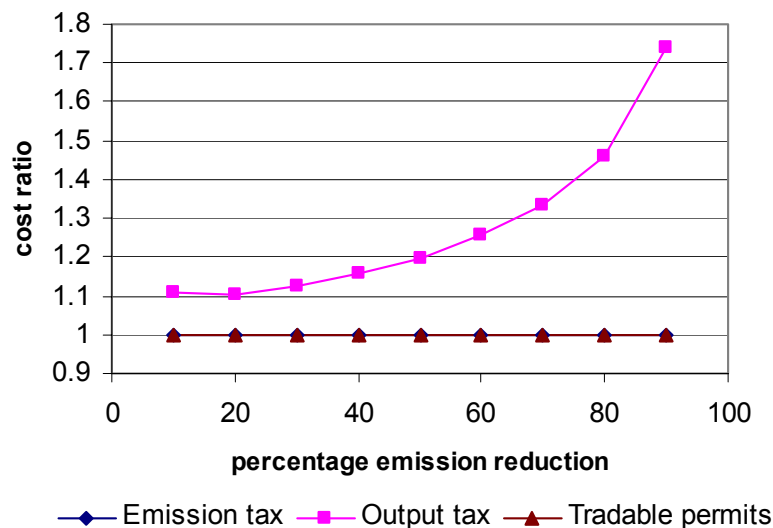


Figure 1: The cost ratio of first-best policy alternatives with perfect compliance to the cost of a first-best emission tax with perfect compliance

²¹ Any government revenues are returned in a lump-sum fashion to the households.

The curve for the tradable permits is equal to unity and therefore coincides with the curve for the emission tax. This means that, in a first-best setting, grandfathered tradable permits are just as efficient as emission taxes. This holds because the tax interaction and revenue-recycling effects do not prevail in the absence of distortionary taxes and thus the source of the cost differences, the revenue-recycling effect, is absent.

Next we consider the output tax. Its first-best cost exceeds that of the emission tax because the abatement effect is absent. Firms will not reduce emissions by installing abatement equipment because it does not help them to comply with the policy and it is costly.

We now compare these results with the results obtained in a model with imperfect compliance and investigate how monitoring and enforcement aspects influence the relative cost efficiency of the instruments. Since monitoring and enforcement are costless, the industries are perfectly competitive and the reductions in emissions are equal, we obtain exactly the same utility levels in the simulation with and without perfect compliance. Including monitoring and enforcement costs would alter the results. The costs of the environmental regulation would then be higher under incomplete compliance than under complete compliance.

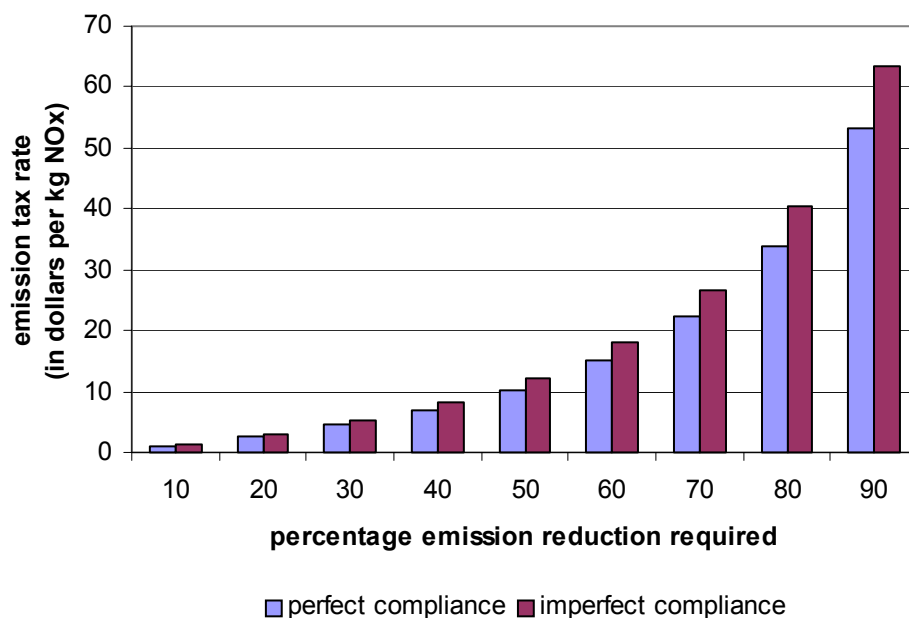


Figure 2: Comparing emission tax rates under perfect and imperfect compliance

Even though the utility levels are equal, the tax rates necessary to obtain the required emission reduction differ for the two scenarios (see figure 2). The tax rate under perfect

compliance is 84 percent of the rate under imperfect compliance. Given the values chosen for the monitoring and enforcement parameters, this 84 percent is the proportion of tax actually paid or $\rho^*(\pi, \bar{p}, \eta)$. From the theoretical analysis we recall (equation (12)):

$$\tau_E = \rho^* t_E \text{ with } \rho^*(\pi, \bar{p}, \eta) = 1 - \frac{(1 - \bar{p}\pi)^2}{4\pi\eta}.$$

The effective tax rate under imperfect compliance is a fraction ρ^* of the nominal tax rate. As expected, the firms' noncompliance forces the government to set a higher tax rate in order to obtain the desired environmental result.

Analogously, we find that, in order to obtain a particular environmental quality goal, the regulator has to issue fewer permits under incomplete compliance than under full compliance. Moreover, under incomplete compliance, the nominal output tax will also be 1.19 (1/0.84) times the nominal tax under full compliance.

Looking at expression (11), we compute the proportion of emissions reported:

$$\theta_E^*(\pi, \bar{p}, \eta) = \left(1 - \frac{1 - \pi\bar{p}}{2\pi\eta}\right) = 0.60$$

Given the monitoring and enforcement policy, firms will report 60 percent of their emissions. Analogously, under an output tax, firms will report 60 percent of their output (see expression (24)). Under a tradable permit scheme, firms will only cover 60 percent of their emissions with permits (see equation (20)).

Allowing firms to partially comply with environmental regulation, leads to stricter policies. In order to obtain a certain environmental goal, the government has to take into account that some firms will decide to violate the regulation.

5.2. Second-best setting

In a second-best setting we take the effect of pre-existing distortionary labour taxes ($t_L = 0.4$) into account and use the policy revenues for decreasing this distortionary tax. The resulting cost ratios under full compliance are represented in figure 3. We compare the

results with the first best emission tax. Again we have results for three policy instruments: the emission tax, the output tax and grandfathered tradable permits.

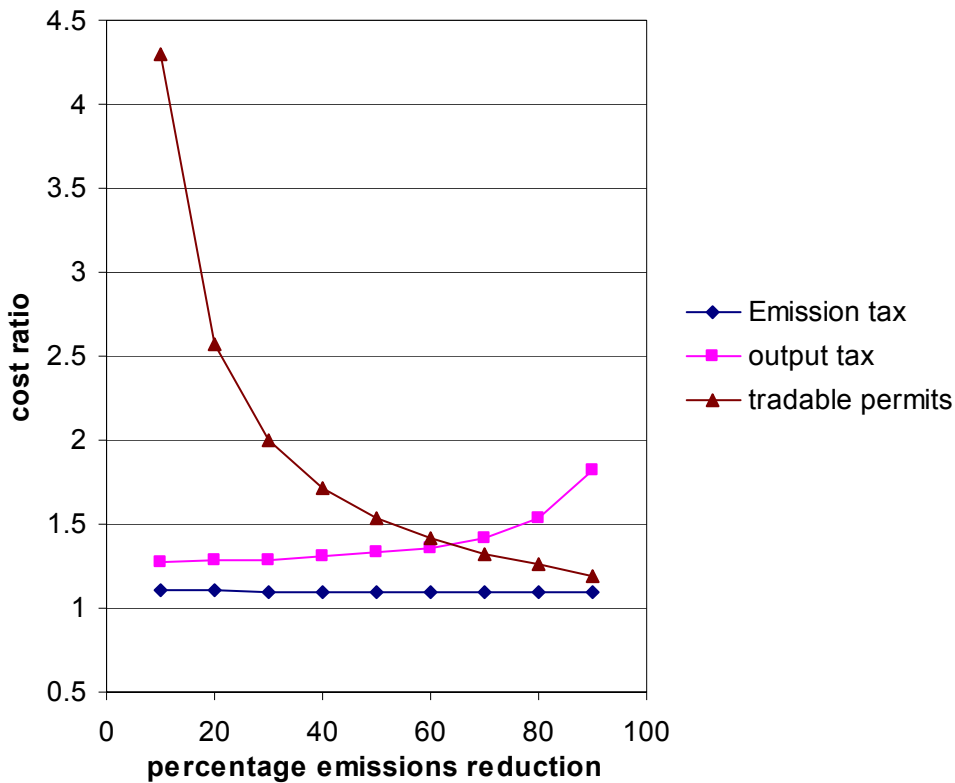


Figure 3: The cost ratio of second-best policy alternatives with perfect compliance to the cost of the first-best emission tax with perfect compliance.

We find that, under second best, the efficiency costs of an emission tax are ten percent higher than first best. The tax-interaction and revenue-recycling effects contribute to a 10 percent increase in the costs of an emission tax. The tradable permit system performs considerably worse due to the lack of a revenue-recycling effect. This policy does not generate revenues for the government that can be used to alleviate existing distortions. Finally the relative position of the output tax does not change vis-à-vis the emission tax. These results duplicate the ones of Goulder et al. (1999).

We now investigate how these results change when we allow firms to violate the environmental policy. The cost curves for imperfect compliance in a second-best world are shown in figure 4. They are all expressed relative to the cost of emission taxes in a first-best world. In general, we find that the compliance rate equals 60 percent for the three instruments. This is identical to the results in a first-best setting since the same

monitoring and enforcement parameters are chosen. Moreover, we find that allowing firms to violate environmental regulation increases welfare slightly (given the assumption of costless monitoring and enforcement). The firms have more flexibility in their decision-making. Some firms prefer the risk of getting caught and paying fines to paying taxes.

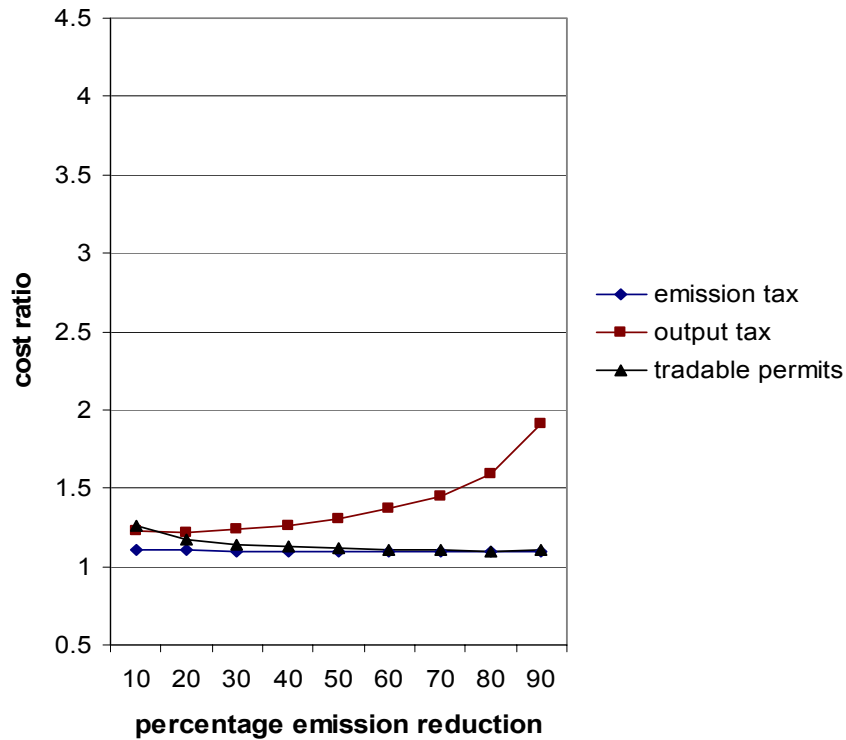


Figure 4: The cost ratio of second-best policy alternatives with imperfect compliance to the cost of first-best emission tax with perfect compliance.

We also notice two important effects on the relative cost efficiency of the environmental policy instruments. First, when we examine the cost of using emission taxes, the cost increase due to the labour market distortion is approximately the same (10%). Secondly, the relative cost difference of grandfathered tradable permits compared to emission taxes has decreased spectacularly. It even disappears completely when the required emission reduction approaches 100 percent. The reason for this result is the revenue-recycling effect that now exists for tradable permits too: revenues are generated through fines and these can be used to alleviate the distorting labour tax. Clearly, fines have some characteristics in common with taxes (Sandmo, 2002). Imperfect compliance does not affect the relative disadvantage of output taxes compared to emission taxes.

We now look at the results for grandfathered tradable permits in more detail. Firstly, an important consequence of allowing imperfect compliance is the effect on the distortionary

labour tax. The labour tax in equilibrium is lower: 0.396 with incomplete compliance instead of 0.401 with full compliance. This lower tax rate provides the households with more income and they increase their consumption. The production sectors accordingly react and expand as well. Labour demand and supply increase and leisure time decreases. Secondly, the quota rents (or profits) are significantly lower under imperfect compliance since violating firms have to pay fines, which include paying for additional permits. Total profits in the economy with incomplete compliance are reduced to 4.9 percent of the profits under perfect compliance. Total fine payments in the economy equal 6.9 billion dollar. Finally, we notice that abatement expenditures under incomplete compliance are slightly higher than under full compliance. Additional abatement allows the firms to compensate for the higher production while still achieving the desired emission goal.

6. SENSITIVITY ANALYSIS

In the following two sections we shortly discuss how the monitoring and enforcement parameters and the assumption of costless monitoring influence our results.

6.1 Monitoring and enforcement parameters

As mentioned in section 4.2, data on the monitoring and enforcement parameters are very difficult to obtain. We assumed that, yearly, ten percent of all firms are inspected ($\bar{p} = 0.1$) and that the variable inspection parameter η equalled 0.5. With respect to the penalty we assumed that violators have to pay twice the amount of evaded taxes ($\pi = 2$). We now look briefly at the implications of changing these parameters.

The three parameters influence the reported fraction θ_E^* of emissions, the reported fraction θ_D^* of output and the reported fraction θ_p^* of permits positively (see expressions (11), (20) and (24)). The more stringent the monitoring and enforcement policy, the more compliant firms will be.

Moreover, increasing the parameters \bar{p} , η and π also raises the fraction ρ of taxes actually paid (see expressions (12), (21) and (25)). From the theoretical analysis we recall

(section 3.1.1) that the effective tax rate τ_E under imperfect compliance is a fraction ρ of the nominal tax rate t_E . Therefore, for a given nominal tax rate, we will find a higher effective tax rate if the monitoring and enforcement policy becomes more rigid. The gross efficiency costs of augmenting the nominal level of the different instruments will increase if \bar{p} , η and π increase.

It is important to see that the relative cost efficiency of the different instruments in the numerical exercise will not change if we modify the monitoring and enforcement policy. The same utility levels will be achieved since monitoring and enforcement is costless. However, the environmental policy will be less strict if the monitoring and enforcement policy is more stringent. Nominal taxes will be lower or the number of permits issued will be larger. The environmental policy instruments and the monitoring and enforcement instruments are, to some extent, substitutes.

6.2 Monitoring costs

We now include positive monitoring costs for the government into the model. Therefore, we allow for a fixed budget assigned to the environmental protection agency. This reflects regulatory practice. However, we do not model how this budget is spent. The addition of these monitoring and enforcement costs reduces the utility level that can be achieved since government expenditures increase. Nonetheless, the relative positions of the three environmental policy instruments will not change.

The relative cost efficiency of these instruments would change if one included variable and fixed monitoring and costs for firms as well as government. After all, not only the inspection agency has costs but also the firm that is inspected. For example, it has to accompany the auditors, prepare documents or analyse control samples. Equivalently, the government is not the only one to have enforcement costs. For example, firms have to pay for legal assistance and for gathering evidence. Moreover, third parties can be involved and households can be injured by the environmental offence. In chapter 3 we compare different policy instruments in a partial equilibrium setting while taking into account positive monitoring and enforcement costs.

Allowing for positive monitoring and enforcement costs for firms would influence the amount of reported emissions θ^* . Firms will report more emissions because they want to decrease the expected monitoring and enforcement costs. This decision is only influenced by the amount of variable inspections (through parameter η) since they cannot influence the number of fixed inspections (parameter \bar{p}). Moreover, firms would also abate more and emit less since the resulting proportion of tax actually paid ρ^* increases if these costs are included.

When looking at the gross efficiency costs we see that those will increase when monitoring and enforcement costs are positive. Since firms are more compliant, the gross efficiency costs also increase and the effective tax rate faced by the firms is higher. Moreover, we would also distinguish an extra tax interaction effect because the government needs funds to cover monitoring and enforcement costs. These monitoring and enforcement costs are influenced by the number of inspections made even though we assume that the fixed inspection frequency \bar{p} , the variable inspection parameter η and the penalty parameter π are exogenous. Allowing the government to optimise its monitoring and enforcement strategy would be an interesting extension. However, the extensive modelling of monitoring and enforcement costs and policy falls beyond the scope of this paper.

7. CONCLUSION

In this paper we examine to what extent incomplete compliance with environmental regulation mitigates the distortions caused by pre-existing labour taxes. In a second-best setting we find that policies, which do not generate revenues for the government in a setting of perfect compliance, perform much better under incomplete compliance. We include the possibility that firms are violating environmental rules and are fined with a certain probability. These fines provide the government with income that is used to alleviate existing distortions in the labour market. In our simulation, grandfathered tradable permits greatly improve their position vis-à-vis emission taxes and output taxes due to these fine payments. Clearly fines have some characteristics in common with a (non-linear) tax as was mentioned in Sandmo (2002).

In a first-best setting and given that monitoring and enforcement is costless, we find that the same utility levels can be reached with and without incomplete compliance. However, the nominal tax rate needs to be higher or the number of permits issued smaller, in order to obtain the required emission reduction.

This is a first attempt to integrate monitoring and enforcement considerations into the choice of policy instruments in a second-best setting. Obviously it would be interesting to include quantity instruments into the framework. This would require us to think about how to make non-compliance comparable across instruments and would involve using some other assumptions for the expected fine function. This would pose new challenges for the numerical model.

Also broader monitoring and enforcement policy options need to be considered. Fines can be replaced by firm closure, imprisonment or other non-monetary sanctions.

Furthermore, it would be useful to take into account the possibility of measurement errors. In reality the measurement equipment of the inspection agency is not perfectly accurate. False positives and false negatives are encountered. Some violators remain undetected even if they are inspected, while some innocent firms will be sentenced.

Another possible extension is to distinguish between the probability of detection and the probability of punishment. In practice we often see that minor violations are left unpunished. This is because convicting a firm is not costless. So it is possible that judges decide to drop the case because it is not worth the time and money to follow up. Therefore, the firms will make their decisions based on the probability of punishment and not the probability of detection.

Further extensions can consist of changing the assumptions of the economic model. We could work with heterogeneous firms per sector. Or we could incorporate heterogeneous consumers and take distributional aspects into account. Finally, we could introduce imperfect compliance for the labour tax too. This is not unrealistic as the shadow economy counts for 10 to 25% of GDP in Western economies (Sandmo, 1981).

Appendix A: Firm behaviour

For $p_{det} = 0$ or $f = 0$, we get a corner solution and the reported emissions will equal zero. Since violators are not punished for lying about their emissions, they maximise their profits by reporting no emissions and therefore paying no taxes.

We investigate the influence of the emission tax t_E and the penalty π on the optimal behaviour of firms. More specifically we look at changes in actual emissions e^* ($= e_o - a^*$) and in the fraction of emissions reported θ_E^* . These results are summarised in table 1 for different scenarios. We distinguish four scenarios depending on the value of the emission tax rate t_E and the penalty π . Each scenario is now discussed in turn.

t_E	π	θ_E^*	e^*
> 0	> 1	$\theta_E^* = \left(1 - \frac{1 - \pi \bar{p}}{2\pi\eta}\right)$	$c'(e_o - e^*) = t_E \left(1 - \frac{(1 - \pi \bar{p})^2}{4\pi\eta}\right) = \tau_E$
$= 0$	≥ 1	$\theta_E^* = 0$	$e^* = e_o$
> 0	$= 1$	$\theta_E^* = \left(1 - \frac{1 - \bar{p}}{2\eta}\right)$	$c'(e_o - e^*) = \left(1 - \frac{(1 - \bar{p})^2}{4\eta}\right) t_E$
> 0	$\rightarrow \infty$	$\theta_E^* \rightarrow 1$	$c'(e_o - e^*) \rightarrow t_E$

Table 2: Reaction of firms to changes in the emission tax and the penalty

Scenario 1: $t_E > 0$ and $\pi > 1$

This scenario is discussed more fully in the body of the text.

For any given e , the firm chooses θ_E to minimise the proportion of tax it actually pays:

$$\theta_E^*(\pi, \bar{p}, \eta) = \left(1 - \frac{1 - \pi \bar{p}}{2\pi\eta}\right) \quad \text{for } f_E, p_{\det, E} \neq 0$$

Next, in a second step, the firm chooses the actual emissions e to minimise its environmental costs. This gives us the following expression:

$$c'(e_o - e^*) = \tau_E = \rho^* t_E$$

Scenario 2: $t_E = 0$ and $\pi \geq 1$

Firms will not invest in abatement nor will they report any emissions. Therefore, we focus on a strictly positive emission tax from now on.

Scenario 3: $t_E > 0$ and $\pi = 1$

When the penalty equals its lower bound or $\pi = 1$ and $\bar{p} \neq 0$ and $\eta \neq 0$, violating firms still have to pay their overdue taxes. The reported emissions will not be zero because the reported emissions influence the probability of detection. The total amount paid, taxes on reported emissions plus overdue taxes, is minimised by reporting a fraction $\left(1 - \frac{1 - \bar{p}}{2\eta}\right)$ of actual emissions (see table 1). This result is obtained by minimising $t_E \theta_E e + p_{\det} f$ with respect to θ_E . Next, the firm chooses the actual emissions e to minimise its environmental costs. The marginal abatement cost equals the actual emission tax rate $\tau_E = \rho t_E$ with

$$\rho^* = 1 - \frac{(1 - \bar{p})^2}{4\eta} \quad \text{since the penalty } \pi \text{ equals } 1.$$

Scenario 4: $t_E > 0$ and $\pi \rightarrow \infty$

When the penalty π goes to infinity, the firms will be reporting more and more truthfully. The marginal abatement cost will, in the limit, equal the emission tax rate. A summary of these findings can be found in table 1.

Appendix B: Deriving expression (16)

First we derive an expression for the change in labour tax necessary to maintain government budget balance following the increase in emission tax.

Totally differentiating the government budget gives (holding G constant):

$$\frac{dt_L}{dt_E} = -\frac{\rho^*(\pi, \bar{p}, \eta) \left[E + t_E \frac{dE}{dt_E} \right] + t_L \frac{\partial L}{\partial p_D} \frac{dp_D}{dt_E}}{L + t_L \frac{\partial L}{\partial t_L}} \text{ with } \frac{dE}{dt_E} = \frac{\partial E}{\partial t_E} + \frac{\partial E}{\partial t_L} \frac{dt_L}{dt_E} \quad (35)$$

Using the expression of the (partial equilibrium) efficiency cost, we obtain:

$$\frac{dt_L}{dt_E} = -\frac{(1+M)}{L} \left\{ \rho^*(\pi, \bar{p}, \eta) \left[E + t_E \frac{dE}{dt_E} \right] + t_L \frac{\partial L}{\partial p_D} \frac{dp_D}{dt_E} \right\} \quad (36)$$

We are interested in the gross effect of the tax increase t_E on welfare. Differentiating utility $V = v(p_D, 1-t_L, G+\pi) - \phi(E)$ with respect to t_E and ignoring the terms in ϕ gives

$$\frac{dv}{dt_E} = \frac{\partial v}{\partial p_D} \frac{dp_D}{dt_E} + \frac{\partial v}{\partial t_L} \frac{dt_L}{dt_E}$$

Using Roy's identity, $\frac{dp_D}{dt_E} = \rho^* [e_o - a]$ and with λ equal to the marginal utility of income,

gives
$$\frac{dv}{dt_E} = -\lambda \rho^* D[e_o - a] - \lambda L \frac{dt_L}{dt_E}$$

$$-\frac{1}{\lambda} \frac{dv}{dt_E} = \rho^* D[e_o - a] + L \frac{dt_L}{dt_E}$$

and using $E = (e_o - a)D$ gives
$$-\frac{1}{\lambda} \frac{dv}{dt_E} = \rho^* E + L \frac{dt_L}{dt_E}$$

This is the efficiency cost (ignoring environmental benefits) for an incremental increase in t_E , expressed in monetary terms.

Appendix C: The numerical model

The numerical model was written in GAMS.

This is the description of the model when an emission tax is levied.

1. Sets

$i = \{D, N, L\}$ - inputs

$j = \{D, N, CD, CN\}$ - outputs

$k = \{\text{leis}, C, CD, CN\}$

2. Parameters

α_e	effectiveness of abatement technology	/0.155/
α_{ij}	distribution parameter for input i in production of good j (via calibration)	
α_k	distribution parameter for the utility function (via calibration)	
β_i	pollution content of good i used	/D 0.062768, N 0, L 0/
χ	curvature parameter for abatement	/0.5/
δ, μ	scaling parameters	
$\overline{E_{tot}}$	maximum amount of emissions possible (in millions of kg)	/23000.028/
$reduc$	emission reduction	
σ_j	elasticity of substitution in production of good j	/D 0.8, N 0.8, CD 0.9, CN 0.9/
σ_c	elasticity of substitution between consumption goods	/0.85/
σ_u	elasticity of substitution between consumption and leisure	/0.96/
T	total time endowment	/2129897.1/

3. Variables

A_j	abatement expenditure in industry j
AD_i	aggregate demand for good i
b_{ij}	use of input i per unit of output of good j
b_{CD}, b_{CN}	relative share of consumption of CD and CN to total consumption
C	aggregate demand for composite consumption good
$CDdem$	aggregate demand for energy-intensive goods
$CNdem$	aggregate demand for non-energy intensive final goods
E_j	actual pollution emitted from production of good j
$Etot$	total actual emissions
η	variable inspection parameter
F_j	fine per sector j
G	lump-sum transfer
L	aggregate labour supply
$leis$	leisure or non-market time
\bar{p}	fixed probability of detection
p_{det_j}	probability of detection per sector j
p_C	price of composite good
p_j	price of output j
pr_L	price of labour
pr_i	price of input i
π	fine paid on overdue taxes
Π_j	profit per industry j
Π_{tot}	total profits or total pollution quota rents
REV	government revenue
t_E	emission tax
t_j	tax on output j
t_L	labour taxation rate
t_R	rent taxation rate

$\theta_{E_j} E_j$	reported emissions from production of good j
$\theta_{X_j} X_j$	reported output of good j
U	total consumer utility
X_j	aggregate supply of good j
X_{ij}	use of good i in production of good j

4. Equations

4.1 Production – firm behaviour

Output:
$$X_j = \delta \left(\sum_i \alpha_{ij} X_{ij}^{\frac{\sigma_j-1}{\sigma_j}} \right)^{\frac{\sigma_j}{\sigma_j-1}}$$

Profit:
$$\Pi_j = (p_j - t_j \theta_{X_j}) X_j - \sum_i p r_i X_{ij} - t_E \theta_{E_j} E_j - A_j - p_{\det.j} \cdot F_j$$

Total profits:
$$\Pi_{tot} = \sum_j \Pi_j$$

Expected fine:
$$p_{\det.j} \cdot F_j = \left(\bar{p} + \eta \frac{(E_j - \theta_j E_j)}{E_j} \right) \cdot (E_j - \theta_j E_j) t_E \pi$$

Emissions:
$$E_j = \beta_D X_{Dj} \left[1 - \alpha_e \left(\frac{A_j}{\beta_D X_{Dj}} \right)^\chi \right]$$

Total emissions:
$$E_{tot} = \sum_j E_j$$

Reported emissions:
$$\theta_{E_j} E_j = \left(1 - \frac{1 - \pi \bar{p}}{2\pi\eta} \right) E_j$$

First-order condition (A_j):
$$A_j = \beta_D X_{Dj} (\rho t_E \alpha_e \chi)^{\frac{1}{1-\chi}}$$

First-order condition (X_{ij}):

$$b_{ij} = \frac{X_{ij}}{X_j} = \frac{\alpha_{ij}^{\sigma_j}}{\delta} \left\{ \frac{pr_i + \beta_i t_E \rho \left[1 - \alpha_e \left(\frac{A_j}{\beta_D X_{Dj}} \right)^\chi (1 - \chi) \right]}{p_j - t_j \theta_{Xj}} \right\}^{-\sigma_j}$$

$$\text{Price: } p_j = t_j \theta_{Xj} + \sum_i b_{ij} pr_i + \frac{\rho t_E E_j + A_j}{X_j}$$

4.2 Household behaviour

$$\text{Utility: } U = U(leis, CD, CN, E) = \left(\alpha_{leis} leis^{\frac{\sigma_u - 1}{\sigma_u}} + \alpha_C C^{\frac{\sigma_u - 1}{\sigma_u}} \right)^{\frac{\sigma_u}{\sigma_u - 1}}$$

$$\text{Composite consumption (1): } C = \mu \left(\alpha_{CD} CD^{\frac{\sigma_c - 1}{\sigma_c}} + \alpha_{CN} CN^{\frac{\sigma_c - 1}{\sigma_c}} \right)^{\frac{\sigma_c}{\sigma_c - 1}}$$

$$\text{Budget constraint: } p_{CD} CD + p_{CN} CN = pr_L (1 - t_L) L + \Pi_{tot} (1 - t_R) + p_C G$$

$$\text{Time constraint: } T = leis + L$$

$$\text{Distribution: } b_{CD} = \frac{CD}{C} = \frac{1}{\mu} \left[\alpha_{CD} + \alpha_{CN} \left(\frac{\alpha_{CD} p_{CN}}{\alpha_{CN} p_{CD}} \right)^{1 - \sigma_c} \right]^{\frac{\sigma_c}{\sigma_c - 1}}$$

$$b_{CN} = \frac{CN}{C} = \frac{1}{\mu} \left[\alpha_{CN} + \alpha_{CD} \left(\frac{\alpha_{CN} p_{CD}}{\alpha_{CD} p_{CN}} \right)^{1 - \sigma_c} \right]^{\frac{\sigma_c}{\sigma_c - 1}}$$

$$\text{Composite price: } p_C = p_{CD} b_{CD} + p_{CN} b_{CN}$$

$$\text{Leisure: } leis = \frac{pr_L (1 - t_L) T + p_C G + (1 - t_R) \Pi_{tot}}{pr_L (1 - t_L) + p_C \left[\frac{\alpha_{leis} p_C}{\alpha_c pr_L (1 - t_L)} \right]^{-\sigma_u}}$$

$$\text{Labour: } L = T - leis$$

Numeraire: $pr_L = 1$

4.3 Government behaviour

Revenue: $REV = t_L L + t_E \sum_j \theta_{Ej} E_j + \sum_j t_j \theta_{Xj} X_j + t_R \Pi_{tot} + \sum_j p_{detj} \cdot F_j$

Tax: $t_R = t_L$

4.4 Aggregate demand and supply

Inputs: $AD_i = \sum_j X_{ij}$

Outputs: $X_{CD} = CD$

$X_{CN} = CN$

$X_i = AD_i \quad \text{for } i = \{D, N\}$

4.5 Equilibrium conditions

Labour market: $L = AD_L + \sum_j A_j$

Emissions: $Etot = (1 - \text{reduc}) \overline{Etot}$

Government: $REV = p_c G$

Bibliography

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

Bovenberg, A.L. and **de Mooij**, R.A. (1994). Environmental levies and distortionary taxation.

American Economic Review, vol. 84, p. 1085-1089.

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

Goulder, L. H. (1995). Environmental taxation and the 'double dividend': a reader's guide. *International Tax and Public Finance*, vol. 2, p. 157-183.

Goulder, L., **Parry**, I., and **Burtraw**, D. (1997). Revenue-raising vs. other approaches to environmental protection: the critical significance of pre-existing tax distortions. *RAND Journal of Economics*, vol. 28, nr. 4, p. 708-731.

Goulder, L. H., **Parry**, I.W.H., **Williams III**, R.C., and **Burtraw**, D., (1999). The cost-effectiveness of alternative instruments for environmental protection in a second-best setting, *Journal of Public Economics*, vol. 72, p. 329-360.

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

Heyes, A. (2001). *The law and economics of the environment*. Edward Elgar. p.424

Keller, W.J. (1976). A nested CES-type utility function and its demand and price-index functions. *European Economic Review*, vol. 7, p. 175-186.

King, M.A. (1983). Welfare analysis of tax reforms using household data. *Journal of Public Economics*, vol. 21, p. 183-214.

Lear, K.K. (1998). *An empirical estimation of EPA administrative penalties*. Working paper, Kelley School of Business, Indiana University.

Magat, W. and **Viscusi**, W.K. (1990). Effectiveness of the EPA's regulatory enforcement: the case of industrial effluent standards. *Journal of Law and Economics*, vol. 33, p. 331-360.

Mayeres, I. and **Proost**, S., (1997). Optimal tax and public investment rules for congestion type of externalities, *Scandinavian Journal of Economics*, vol. 99, nr. 2, p. 261-279.

Polinsky, A.M. and **Shavell**, S. (2000). The economic theory of public law enforcement. *Journal of Economic Literature*, vol. XXXVIII (March), p. 45-76

Pechan, E.H. and associates, (1996). Emission Reduction and Cost Analysis Model for NO_x (ERCAM-NO_x). http://www.pechan.com/projects/Reports_Publications/

Russell, C.S., **Harrington**, W. and **Vaughan**, W.J. (1986). *Enforcing pollution control laws*. Resources for the Future.

Sandmo, A. (1981). Income tax evasion, labour supply, and the equity-efficiency tradeoff. *Journal of Public Economics*, vol. 16, p. 265-288.

Sandmo, A. (2002). Efficient environmental policy with imperfect compliance. *Environmental and Resource Economics*, vol. 23(1), p. 85-103.

Schmutzler, A. and **Goulder**, L.H. (1997). The choice between emission taxes and output taxes under imperfect monitoring. *Journal of Environmental Economics and Management*, vol. 32, p. 51-64.



The Center for Economic Studies (CES) is the research division of the Department of Economics of the Katholieke Universiteit Leuven. The CES research department employs some 100 people. The division Energy, Transport & Environment (ETE) currently consists of about 15 full time researchers. The general aim of ETE is to apply state of the art economic theory to current policy issues at the Flemish, Belgian and European level. An important asset of ETE is its extensive portfolio of numerical partial and general equilibrium models for the assessment of transport, energy and environmental policies.

ETE WORKING PAPER SERIES 2004

N°2004-15	Rousseau S., Proost S. (2004), The Relative Efficiency of Market-based Environmental Policy Instruments with Imperfect Compliance
N°2004-14	Eyckmans J., Finus M. (2004), An Almost Ideal Sharing Scheme for Coalition Games with Externalities
N°2004-13	De Borger B., Mayeres I. (2004), Taxation of car ownership, car use and public transport: insights derived from a discrete choice numerical optimisation model
N°2004-12	De Borger B., Proost S. (2004), Vertical and horizontal tax competition in the transport sector
N°2004-11	de Palma A., Dunkerley F., Proost S. (2004), Imperfect Competition and Congestion in a City with asymmetric subcenters
N°2004-10	Calthrop E., Proost S. (2004), Regulating on-street parking
N°2004-09	de Palma A., Proost S. (2004) Imperfect competition and congestion in the City
N°2004-08	Pepermans G., Willems B. (2004), Ramsey Pricing in a Congested Network with Market Power in Generation: A Numerical Illustration for Belgium
N°2004-07	Delhaye E. (2004), Traffic safety: speed limits, strict liability and a km tax
N°2004-06	Eyckmans J., Finus M. (2004), An Empirical Assessment of Measures to Enhance the Success of Global Climate Treaties
n°2004-05	Eyckmans J., Meynaerts E., Ochelen S. (2004), The Environmental Costing Model: a tool for more efficient environmental policymaking in Flanders
n°2004-04	Saveyn B., Proost S. (2004), Environmental Tax Reform with Vertical Tax Externalities in a Federal State
n°2004-03	Rousseau S. (2004), Timing of environmental inspections: Survival of the compliant