CEIS Centre for Economic and International Studies
TOR VERGATA

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 8, Issue 2, No. 163 – February 2010

Risk Management and Regulation Compliance with Tradable Permits under Dynamic Uncertainty

Pasquale L. Scandizzo and Odin K. Knudsen

Risk Management and Regulation Compliance with Tradable Permits under Dynamic Uncertainty

by

Pasquale L. Scandizzo and Odin K. Knudsen

Abstract

In this paper, we explore the effects of dynamic uncertainty on the risk management of regulated industries and emission market. We consider as major sources of uncertainty the stochastic growth of demand for the industry output (e.g. electric energy) and the ensuing lack of information on the pollution levels of individual firms, their behavior and the behavior of the regulator. These sources of uncertainty are common in pollution permit trading as not only does the market respond to the volatility of fundamentals but also to the vagaries of the institutional structure, created by public policy and enforced through regulation. The paper shows that in the presence of strategic behavior on the part of the agents involved, even though both the level and the volatility of output increases over time, trading of permits is a highly effective instrument of risk management, since it allows the firms to pool the risks arising from the volatile environment, thereby simplifying enforcement, reducing emissions and improving resource allocation. Moreover, uncertainty plays a subtle influencing role, since on one hand it broadens the regulator's deterrent power over potential polluters, while on the other it reduces the expected value of the sanction for the individual firm.

Key words: risk; permits; regulation; enforcement; dynamic uncertainty; option; pricing; equilibrium.

JEL classification: K34, H40, Q52

Introduction

Policies to achieve environmental quality have particular importance as the challenge of mitigating climate change and reducing emissions has taken on currency. Two instruments have received particular support from economists: marketable permits and emission taxes or charges (Pigou, 1920 and 1932; Crocker, 1966; Dales, 1968; Montgomery,1972;Kneese and Schultze, 1975). In theory, pollution taxes or tradable permits will minimize the costs of achieving a targeted level of pollution (Baumol and Oates, 1988). They will also provide incentives for adoption and diffusion of new and cheaper technologies (Milliman and Prince, 1989).

Real options analysis has been used to determine the value of flexibility or exit and the timing of capital investments of a regulated firm under uncertainty (Teisberg, 1994), finding that investments of utilities will be delayed when there is asymmetry between profits and losses due to regulation. Using real options valuation, it has been also found that a major reason US electrical utilities delay the decision to invest is to gather more information on regulatory restructuring (Ishii and Yan, 2004). With few exceptions (Knudsen and Scandizzo, 2005), these real options analyses have assumed that the regulators will enforce perfectly the regulations and do not optimize their own behavior. The real options approach to investment decisions has not been extended to a utility having the provision to buy and sell tradable pollution permits under regulatory uncertainty on enforcement. Non-compliance is a substantial issue as much of regulation relies on self-reporting or incomplete monitoring by the regulatory (Shapiro, 1984). As non-compliance on individual taxation will attest, the presence of fines for violation of emission standards is not necessarily sufficient motivation for compliance, especially if monitoring is imperfect.

In this paper, we use a real options approach to examine strategic behavior under two sources of dynamic uncertainty: market or demand uncertainty and regulatory uncertainty. These uncertainties have particular relevance to the design of emissions trading permits under cap and trade systems as is currently in operation for carbon in the EU Emissions Trading System (ETS) and potentially for a United States system under policy discussion by states and the federal government. They are also important to take into account the costs and benefits of regulation, as advocated, for example, by Crals and Vereek (2005).

The model developed in the paper may have more general implications for law enforcement under uncertainty. It is based on the idea, common in the literature on the economics of crime and punishment (see, for example Becker, 1968, Graetz et al, 1986, Benoit et al. 1995) that the agent's behavior is rationally based on the comparison of expected benefits and

costs. Within a framework of uncertainty evolving over time, however, the model hypothesizes that the agent responds to the credible threat of the authorities to impose norms and sanction non compliance, by treating regulatory prescriptions as a form of institutional risk that can be managed by using appropriate risk management strategies. From the point of view of a single firm, regulatory uncertainty may depend on various problems. For example, the firm may not to be able to control compliance with the level of accuracy requested by the regulations (e.g. in the case of pollution, its emissions may be irregular). The firm may not have full confidence in the measurements necessary to monitor compliance, because they may be mistaken, or because the monitoring agents may themselves behave strategically. Also, the firm may find sources of uncertainty in the behavior of the regulator as well as on its possible changes over time, on a variety of issues: the setting of the norm, the interpretation of the rules and regulations, the effectiveness of enforcement, the legal and social consequences (as in "name and shame") of sanctions and the possibility to appeal and reverse any specific decision on regulation and sanctioning.

From the point of view of the regulator, enforcement is also likely to result in strategic behavior. Monitoring and enforcement costs, in fact, are lumpy and different for the different agents subject to the regulation. As such, the regulator will be inclined to undertake these costs discontinuously, and only if she believes that the evidence of non compliance is sufficiently large that a threshold for action appears to be crossed. In turn, this threshold may depend on the regulator's budget, her policy priorities and, ultimately, her perception of the expected social payoff. In this respect, the behavior of the regulator appears to belong to the general class of control models with discontinuous adjustments or , as put in a recent contribution (Stokey, 2009) of the "economics of inaction". More specifically, we can interpret the deterrent power of the enforcer as a call option (the faculty, but not the obligation to sanction under evidence of non compliance), which constitutes a contingent liability for the potential violator. This option will only be exercised by the enforcer if a threshold of evidence of non compliance is crossed.

Pollution permits and trading are becoming increasingly important as a market friendly instrument to control pollution at lower costs. Although such schemes have had their birth with sulfur dioxide trading in the United States, they really did not hit international prominence until the Kyoto Protocol came into force. By building into Protocol carbon emission trading and with the emergence of the European Trading System (ETS), pollution trading became a multi-billion dollar market. Despite their growth, the

economics underlying these pollution markets are still only imperfectly understood. Although it is assumed that these markets promote least cost means of meeting targets on carbon emissions and the economics of pricing of permits and penalties has received wide attention in the literature, several areas need further exploration, specially for what concerns the role of uncertainty.

Fundamental and regulatory uncertainty are important for the problem of non compliance, which is the main question underlying the theory of law enforcement under uncertainty . Non compliance in the case of emission regulations has been investigated for example by Malik (1990), who analyzes the implications of cheating on permit prices, by Stranlund and Dhanda (1999), who look at the problem of monitoring, and by Hatcher(2005), who relates the problem of cheating to the mode of sanctioning violations. Other studies have treated the regulation of pollution when emissions are stochastic (Plourde and Yeung 1989; Xepapadeas 1992). Beavis and Walker (1983) in an earlier study on the problem of enforcement find that both the mean and variance of emissions will rise if permit prices drop or if inspection frequency rises. Mrozeki and Keelers' (2004) model of firms pooling risks from stochastic emissions through tradable permits, Montero's (2005) model of pollution controls under imperfect information and Wirl and Noll (2007) investigation of firm's decisions on permits and abatement are studies closer in approach and spirit to our analysis . The novelty of our results, however, consists in the demonstration that permit markets tend to be more effective under dynamic uncertainty determined by a host of concurring factors, including stochastic growth of output, larger and more volatile emissions over time and firms and regulators behaving strategically to cope with the ensuing risks from a dynamically changing random environment .

With respect to the existing literature, our analysis achieves three novel results. First, it shows that under dynamic uncertainty the power of the regulator to determine compliance is based on a credible threat to sanction non compliers. This threat, in turn, is a real option and, as such, depends positively on the degree of uncertainty, that is on the volatility of both the size and the timing of the sanction that the regulator is perceived as having the faculty (but not the obligation) to apply. Regulation and firms' strategic behavior thus becomes a form of risk management whereby economic agents share the risk of emission, sanctioning and noncompliance. Second, we demonstrate that a market for permits is compatible with a full market equilibrium in a regulated industry, and that in such a market, permit trading can achieve significant allocation economies through risk pooling. Third, our

analysis shows that heterogeneity of firms and uncertainty may favor the success of a permit emission system, which , by allowing trade to occur across firms in different states of nature, may achieve social emission targets with comparably lower fines and allocation costs than alternative systems .

Dynamic uncertainty in the permit system

While emission permit systems have been variously used in several countries before and after the Kyoto protocol, the European Emission Trading System (ETS) appears the most interesting from the point of view of well organized provisions and regulatory practices, designed to achieve maximum regulatory effectiveness. It is the very nature of ETS as a system of tradable assets in contingent markets, however, to be prone to uncertainty. Because ETS regulatory practices are based on a self assessment approach combined with stiff penalties, furthermore, uncertainties arise from the inevitable confrontation between the regulator and the potential non complier. In this respect, at least at first sight, ETS presents itself as a punitive system, rather than one based on incentives to comply through gradually aligning the objectives of the firms to the goals of the program. This punitive stance is such that, even once the ex ante uncertainties on the level of the allowances and on the size of the penalties are removed, at least three main sources of uncertainty remain. First, if pollution is stochastic, firms face intrinsic uncertainty and may not be able to perfectly comply, since they cannot predict the amount of pollution that they will generate without error. The uncertainty is compounded by the fact that most sectors affected, and chiefly the energy industry, are subject to stochastic growth, with multiple sources of random changes, such as demand, technology, shifting patterns of resource uses, public sentiment, laws and regulations. Thus, not only uncertainty on pollution levels is high but it is also dynamic, with emissions becoming larger on average and more volatile as time goes on. Second, monitoring and verification can only be carried out with limited precision, depending on the volatility of emissions, and the capabilities and costs of the monitor/verifier. Third, sanctions can be imposed only at some cost, and the effectiveness of the enforcement will itself be uncertain, depending on the strength of the evidence on non compliance, the costs of collecting the fines, and the possible legal challenges[1].

[1] For example, the Netherlands Emission Authority (NEA) distinguishes five fsources of uncertainty on emission monitoring and reporting: (i) primary data flow from the

Dynamic uncertainty in the ETS , therefore, derives from the fact that emissions are associated with growth, and both growth and volatility respond to various random factors and tend to change over time. These characteristics are exemplified by the electricity sector, by far the biggest sector covered by the ETS and the most sensitive to external influences on demand, such as temperature, precipitation (hydro generation), and wind (power generation). Electricity demand, even though strongly affected by population and income growth, varies by the time of the day and the day of the year and is influenced by climatic factors such as temperature and precipitations.

Within this context of stochastic changes, monitoring and verification (M&V) is an important source of dynamic uncertainty. Although their dynamic character is not explicitly acknowledged, M&V uncertainties are recognized as important in several ETS documents. As the UK Environmental Agency reports (2004)[2], ".. the EU Monitoring Guidelines (EU 2004) clearly require operators to include in their M&R plan description of "*a list of tiers to be applied for activity data, emission factors, oxidation and conversion factors for each of the activities and fuel types/materials*" (Section 4.2, Annex I of the M&R Guidelines). In the case of activity data involving metering, tiers tend to be benchmarked in terms of permissible uncertainty (confirmed as "*the 95% confidence interval around the measured value*" in Section 4.3, Annex I of the M&R Guidelines). Tiers related to the other factors tend to be benchmarked in terms of given default values or derivation according to good practice outlined in Section 10 of Annex I of the M&R Guidelines, including application of CEN standards and employing ISO 17025 accredited laboratory services. Table 2 in Annex I of the M&R Guidelines does provide an "*Informative table with the uncertainty ranges typically found for different metering devices*" supposedly to "*assist the selection of appropriate tiers for activity data*". Statement is made that "*The table may inform competent authorities and operators about the possibilities and limitations for applying appropriate*

---

source to the NEA, (ii) installation (sources), (iii) measurement (raw data), (iv) calculation (data processing), (v) reporting emissions.

[2] These remarks are based on the comments of the Monitoring, Reporting and Verification (MRV) Group involving UK Government, Competent Authority, Verifier, UKAS and industrial representation. The Group was established to consider relevant requirements related to ETS implementation and specifically the Commission's Decision of 29th January 2004 (the Monitoring and Reporting Guidelines).

tiers for the determination of activity data". The problem is that the extremes of the ranges specified often correspond to different tiers, so whereas the lower end of the range may correspond to the specification required of a highest tier, the higher end does not. In such cases, this results in it being unclear whether a given device will comply with a desired tier uncertainty benchmark unless a specific uncertainty analysis is carried out."

The dynamic character of M&V uncertainties firstly derives from the fact that, quite apart from technical difficulties, the process of verification creates its own uncertainties. For example , the German Emission Trading Authority cites (Landgrebe, 2009), as the main source of unsatisfactory verification reports: (i) inconsistencies between real installation, monitoring plan & monitoring report, (ii) obvious misstatements of operators not identified, (iii) additional requirements of CA ignored and, (iv) verification statements missing or not comprehensive. While accreditation can be considered a first step in reducing the uncertainties arising from these inaccuracies, according to the same source, by no means it can be considered a resolution to the problem.

A second element of M &V stochastic dynamics is due to the fact that uncertainties linked to imperfect verification and the need to "verify the verifiers" open the possibility that both the firms and the authorities behave strategically. The firms may try to capture weak verifiers, by promising gains through consulting or procurement contracts. The authorities will face a decision on whether to challenge the inadequate verification reports, proceed to costly second round controls, take the reports at their face or simply consider unreliable reports as lack of evidence. This implies that sanctions may not be imposed not because of an ascertained high degree of compliance, but because the evidence of non compliance is not sufficiently robust. The ensuing erratic tendency to inaction may be reinforced by the fact that some EU-25 have incorporated ETS permits into the existing installation environmental law code , so that operators or companies, in addition to pay the fines for non compliance, could be indicted, with additional uncertainties on the type of offence involved. In Germany, for example, the following violations are listed as administrative offenses under section 19 of GHG Emissions Trading Act, and are subject to penalty fines up to 50.000 Euros per offence: (a) operations without correct permit, (b) non notification of changes to installation or operator, (c) non cooperation with competent authority inspection. Also, a fraudulent claim could be considered under Section 263 of the Federal Penal Code and bring to a prison sentence up to 5 years for the violator (Thomas , 2005).

The EU Commission proposal amending the ETS, delivered less than five years after adoption of the ETS Directive, is also a case in point of the dynamic uncertainties arising from the continuous confrontation between the regulator and the companies over time, as both regulation and enforcement evolve to reflect the changing circumstances and behavioral patterns. The amendment of the ETS agreed on in December 2008, in fact, involves both regulatory uncertainty and "business uncertainty". Free allocation is associated with larger regulatory uncertainty, in the sense that firms do not know a priori their allowances. Business uncertainty is prevalent in the event of auctioning, since in this case the allowances themselves will become a strategic variable to bid for in a competitive confrontation with the other firms .

Finally, as highlighted in the Bank of England Financial Markets Law Committee's assessment of October 2009, uncertainty looms over the fundamental legal principles that underpin the ETS system. This uncertainty is dynamic in the sense that only time, enforcement and jurisprudence can solve most of the ambiguities involved. Specifically, the assessment highlighted the following issue – "that nothing in the ETS provides any indication of the legal nature of emission allowances. Emission allowances have aspects of both administrative grants or licenses and of private property…" As emissions allowances have the properties of both licenses and private property, different interpretations of the legal status of allowances might occur in various member states of the EU . The potential consequences of this ambiguity are uncertainty over "which law properly governs the creation, transfer and cancellation of [an] allowance, and whether (and how) security rights can be created over that allowance". The assessment also refers to uncertainty over the tax and accounting treatment of allowances, their disposal following a registered holder's insolvency, whether allowances should be subject to regulation as an investment and whether allowances can be stolen.

The Model

The model is based on the idea that trading permits under dynamic uncertainty allows firms to behave strategically, by accounting for output and prices uncertain evolution over time and anticipating the regulator's behavior in implementing the regulation.

We begin from an industrial sector base, such as the power sector, where demand is assumed to be exogenous and stochastic and output to

adjust to demand in every period. Because of demand and output stochastic growth, emissions are also growing stochastically and dynamic uncertainty ensues. Specifically the output (and demand) of the industrial sector Q is assumed to be a random variable following a stochastic process of the Brownian motion variety. This type of process is often used in physics and, more recently, has been used in economics, to represent in a parameter-parsimonious way, the end product of a multiplicity of random changes evolving over time. It has the property of being the limit of the sum of many discrete-time stochastic processes with stationary independent increments. We use the so called geometric Brownian motion specification, more popular in economics because it determines values that cannot be negative (such as prices), and can be represented by a single equation:

(1) $dQ = \alpha Q dt + \sigma Q dz$

dz being a log-normally distributed random variable with mean zero and variance equal dt. The parameters α and σ^2 represent respectively the drift or trend in demand and the variance. Within the sector, firms (depicted by the subscript *i*) are technologically heterogeneous with emissions y_i assumed for simplicity to be proportional to their output Q_i:

(2) $y_i = Q_i u_i$

The firm has a share w_i of sector output (i.e. $Q_i = w_i Q$) and therefore is responsible for $w_i\, u_i\, Q$ of the sector's emissions where u_i is the emissions per unit of output of the *i*th firm. The more technologically clean the firm, the lower is the emissions factor u_i. Because industry output is stochastic, both the firm's output and its pollution level will be stochastic as well.

The government is assumed to have implemented a cap and trade system where each firm is allocated through the so called "granfathering") an emissions allowance Y_i that it can supplement through market purchases or sales to other firms in the industry at a market determined price p. To deter firms from exceeding their emission allowances, the regulator is authorized to impose a fine at an ad valorem rate of γ of the amount of emissions that exceed the allowance plus the net amount of permits derived from the trading of emissions. This fine is imposed on the emission level detected as exceeding an allowance Y_i (granted to the firm at no cost) for a

given amount of time τ, that is, the present value of fine equals: $\gamma(\int_0^{\tau} e^{-\delta t} w_i Q u_i dt - q_i - Y_i) = \gamma(\frac{w_i Q u_i}{\delta}(1-e^{-\delta\tau}) - q_i - Y_i) = \gamma(\frac{w_i Q u_i}{\delta}\theta - q_i - Y_i)$.

where $\theta = 1 - e^{-\delta\tau}$.

Allowances Y_i are fixed in such a way that their sum is the targeted value of emissions and is therefore a fraction of total emissions expected by the regulator. We assume that the regulator fixes the pollution target based on full information on the total emission level and rational expectations about the total output of the industry i.e.: $Y = \sum_{i=1}^{N} Y_i = \alpha \frac{\theta Q}{\delta} \sum_{i=1}^{N} w_i u_i$, with $0 < \alpha < 1$. This does imply, however, that in distributing the allowances the regulator has similar detailed information on the individual firms, i.e. in general, $Y_i \neq \alpha_i \frac{\theta Q}{\delta} w_i u_i$. As a consequence, firms will be uncertain ex ante on the level of allowances they will be given and ex post on the possibility of challenging the allocation by appealing to authorities or by non complying. Because output is stochastic, they will also be uncertain on whether they will be able to respect the individual pollution level that they have been given. Detecting non compliance and imposing the fine to the *ith* firm requires the regulator to commit a fixed cost $V_i \geq 0$. This detection/enforcement cost is firm specific and depends on factors such as its location, its size, its technological characteristics, the perceived appropriateness of the allowances received and the firm's power to keep non transparency in a political economy sense. It can be seen as deriving from monitoring problems, which range from measurement costs to the need to verify the verifiers, as explained for the ETS case.

Because of this detection cost, the regulator, in attempting to limit emissions, does not collect the fine at any possible evidence of non compliance, but relies on the deterrent value of the threat to impose this sanction. Furthermore, because the output of emissions is stochastic, a temporary violation of the allowance limit may be of random nature, so that repeated observations may be necessary before the regulator may safely conclude that the violation has effectively occurred. Thus, because collecting information is costly, the socially efficient strategy for the regulator would be to engage in detection operations, such as repeated observations, report verification and other analyses for the ith firm, only for the cases where the costs incurred would appear to be equal or smaller than the expected benefits. In these cases, the regulator would be rationally committing the amount of resources corresponding to the detection costs in the expectation

of receiving the uncertain benefits from applying the sanction. In other words, committing resources to find out whether the ith firm were effectively complying with regulations would be equivalent to an investment with given costs and uncertain benefits. Under dynamic uncertainty, in particular, the regulator may be seen as holding a real option constituted by the faculty, but not the obligation, to invest in detection and enforcement in the case of suspected non compliance. This option will be exercised only if the allowance is expected to be exceeded by a significantly large margin, because only in that case the costs of detecting the opportunity to impose the penalty will be matched by sufficient benefits in terms of tax collection and pollution abatement. [3]

Note that this does not mean that the regulator is invariably imposing the penalty only if this is large enough to outweigh the costs, but that she tends to commit resources to find out whether the penalty should be applied so that, on average, this can be expected to happen. This behavior is what should be expected under the assumption of rationality, since the balance between detection costs and penalty revenues is a first best strategy for the regulator, if she is exhibits no preferences for firms on the basis of any characteristic that may be related to detection costs. The same behavior also ensures that the regulator pursues the best strategy to maximize social welfare, provided that the penalty correctly reflects the social cost of non compliance.

The firm, on its side, maximizes its expected net worth, as measured by its (expected) net present value, but must also decide whether to purchase emission permits to cover the possibility of an increase in emissions as demand, for example for electricity, randomly increases due to fluctuations in weather. It must also manage the threat of the regulator imposing potentially stiff penalties on its excessive emissions. Over the longer term, the firm can invest in cleaner technologies, but in the short term it faces both the vagaries of the market for permits and the possibility of fines from the regulator.

Firms are assumed to be heterogeneous because of their size, their unit emission rates, their productions costs and their abatement costs. Each firm maximizes its expected net present value Π_i:

[3] Note that in many regulatory or legal situations, enforcement is not absolute. A simple reference to vehicle speeding where the police may decide to let some unknown violation of speed limits take place or in over law enforcement, e.g. taxes or recreational drugs, where certain violations may not be worth the effort of imposing the sanction.

(3) $\Pi_i = \frac{Pw_iQ}{\delta} - \frac{C_i(w_iQ)}{\delta} - \frac{w_iQK_i(u_i)}{\delta} - pq_i - F(Q_i)$**,**

where P is price of output, $C_i(w_iQ)$ is a concave cost function ($\frac{dC_i}{d(w_iQ)} \geq 0, \frac{d^2C_i}{d(w_i^2Q^2)} > 0$), C_i indicating irreversible investment and operating costs , $K_i(u_i)$ is the abatement cost per unit of output ($\frac{\partial K_i}{\partial u_i} < 0, \frac{\partial^2 K_i}{\partial u_i^2} > 0$) and $F(Q_i)$ is the liability threat coming from the regulator. The discount factor δ is the difference between the risk free rate of interest *r* and the drift α of the stochastic process or $\delta = r - \alpha$. The term pq_i represents the value of permits acquired by the firm in pollution trading.

The last term of (3), $F(Q_i)$, requires a longer explanation. Because of uncertainty on both pollution levels and the regulator's behavior, the value of the term depends on the circumstances. If the regulator is expected to impose immediately a fine on the firm, it will have the value of the expected penalty. On the other hand, the value of the term is zero if the regulator is expected to decide that the violation is only suspected, is minor and/or will not persist and therefore is not worth the detection costs, either directly to the regulator or indirectly to society. If there is a violation that could engage the regulator at some future time, then the option has a value in between these two extremes.

To the firm, the possible imposing of the fine is a risk, which takes the form of a contingent liability, whose value depends on the option held and/or exercised by the regulator. The value of this option depends on uncertainty, the period of the option and the underlying asset value, in this case the possible fine, along with other parameters such as the interest rate. At the time the firm makes its decision, its contingent liability equals its expectation of the regulator's behavior under alternative states of the nature. Because of dynamic uncertainty, such an expectation is measured by the value of the option that the regulator holds to impose the fine on the firm if the latter does not comply with the regulations.

The regulator thus projects a credible threat based on the value of the call option that she is expected to exercise. As explained earlier, under dynamic (i.e. time dependent) uncertainty, the power to impose a fine on non compliant firms can be interpreted as a call option for the regulator, who will

be expected to exercise it, given a measure of non compliance. More specifically, the regulator will be expected to exercise her power to fine a non compliant firm, when the emissions reported/expected for the latter are sufficiently high that the expected amount of the fine exceeds costs of detection and enforcement.[4]

We assume that firms not only differ in technology and emission levels, but also, idiosyncratically, in the amount of transaction or economic costs that they generate when the regulator tries to detect and sanction non compliance. At any point in time, the option is taking on a value that depends on the expected revenue from the fine, detection costs and on the time when the firm's liability option held by the regulator is expected to be exercised.

We can express these concepts analytically. Following Dixit and Pindyck (1994) the value of the call option, i.e. the credible threat that the regulator transmits to firms that may be non compliant $F(Q_i)$ at time T can be expressed as:

$$(4) \quad F(Q_i) = \sup E_Q\left[\left(\int_{\tau+T}^{\infty} e^{-\rho(s-\tau)}\gamma(w_i Q_s u_i \theta - Y_i - q_i)ds - V_i\right)\right]$$

For *i= 1 to N.*

In (4) γ is the *ad valorem* rate of the fine, $\theta = 1 - e^{-\delta\tau}$ is the discount factor for the period of time for which the fine is levied, while Y_i is the pollution cap imposed by the regulator on the ith firm and V_i the cost of implementing the sanction against the ith firm. Notice that the fine is raised on the difference between the present value of the firm's emission and the total value of the individual emission cap and of the quantity of permits (expressed in emission tons) owned by the firm. Once non compliance is detected, we assume that an irreversible investment in enforcement costs has to be made by the agency in charge. The ith firm thus expects the regulator to invest in detection and enforcement on the basis of its expected emissions over a given period of time[5].

Given these premises, we can now state the following proposition:

[4] This is referred to as the value matching condition.

[5] Because output is assumed to follow a geometric Brownian motion and the emissions are proportional to output, expected emissions are the same as point emissions.

Proposition 1, For each firm, the value of the contingent liability given by the threat of the regulator's action is equal to the expected present value of the total amount of the penalty to be paid in the case of non compliance.

A full proof of this proposition is presented in the Appendix . Here, we notice that two alternative cases can be distinguished:

1. the firm's expected emissions have exceeded or are just at the trigger point of the regulator $Q_i^* u_i = w_i Q^{*(i)} u_i$.

2. the firm's expected emissions have not exceeded the trigger point.

In case 1., the value of the contingent liability is simply the value of the fine at the time it is imposed:

(5a) $F(Q_i) = \gamma(w_i \frac{Q^*}{\delta} u_i \theta - Y_i - q_i) - V_i$ if $Q \geq Q^{*(i)}$

In case 2. the value of the contingent liability is the present expected value of the fine discounted from the time at which it is expected to be imposed by the regulator:

(5b) $F(Q_i) = Ee^{-rt_i}[\gamma(\frac{Q_i^*}{\delta} u_i \theta - Y_i - q_i) - V_i]$ if $Q < Q^{*(i)}$

where:

(5c) $Q_i^* u_i \theta = Q^{*(i)} w_i u_i \theta = \frac{\beta_1 \delta}{\beta_1 - 1}(Y_i + q_i + \frac{V_i}{\gamma})$

In (5a)-(5c), $\delta = \alpha - \rho$, and $Ee^{-rt_i} = (\frac{w_i Q u_i}{Q_i^* u_i})^{\beta_1}$ is the expected discount factor corresponding to the (stochastic) time of enforcement t_i for the ith firm. In fact, expression (5b) can be considered comprehensive of both cases, and is reduced to expression (5a) when the discount factor at time of enforcement is unity. Thus, the higher the allowances granted to the ith firm, the higher the costs of monitoring its emissions, the higher the number of permits acquired, the higher will be the threshold level of emissions and the farther away the time at which the regulator may be expected to impose the fine.

Consider the value of the firm under the threat of sanction in the case of expected action by the regulator some time in the future (the second case above). The expected present value of the firm is affected by the regulation as a contingent liability (the expression in (5b)) . As demonstrated in Appendix, within a wide range of values of the ratio between the emission level and the threshold of action of the regulator, the value of the option

increases as the volatility of the underlying process (i.e. output and emissions) increases. Thus, within this range, the higher the uncertainty, the higher the impact of the threat of being sanctioned by the regulator for a given level of the fine or a given degree of non compliance. An increase in uncertainty, by increasing the liability option and the number of potential non compliers, will also increase the aggregate liability option of the industry. However, in the area of low uncertainty (the vicinity of the point of zero vola), we can find ourselves also in the opposite regime, since , if volatility is sufficiently low and allowances and enforcement costs are sufficiently high, increased uncertainty will decrease the value of the option, thereby reducing the number of potential non compliers.

We now explore the behavior of the firm in the face of this contingent liability. We assume that each firm maximizes net present value by setting abatement costs, thereby selecting emission levels, by purchasing a certain number of permits, and by selecting the firm output level as a share of the industrial sector output. Market prices for output and for the permits are determined by the conditions that supply and demand must be equal in both markets or: $\sum_i w_i = 1$ and $\sum_i q_i = 0$.

Proposition 2. At the optimum each firm equates marginal abatement costs to the expected present value of the fine, corrected for the degree of expected non compliance in the absence of abatement.

Assume initially that permits are non tradable (i.e. $q_i = 0$ in (3)). In order to maximize its expected present value, each firm will use its available abatement equipment up to the point where the marginal cost of reducing the emission equals the marginal reduction in the threat by the regulator. Using expression (3) and (5b), we obtain:

(6) $$\left|K_i^{'}\right| = \gamma(\frac{Q}{Q^{*(i)}})^{\beta_1}\theta(Q^{*(i)}/Q) = (Ee^{-rt_i}\gamma)\theta(Q^{*(i)}/Q)$$

Because of the existence of a threshold of intervention from the regulator (see expression (5c)), we can distinguish two types of compliers. The first type, that we may call “effective compliers”, are the firms whose emissions are below the allowances received, i.e. those that satisfy the

constraint: $\frac{w_i Qu_i}{\delta}\theta < Y_i + q_i$. The second type, that we may simply call " compliers" are those satisfying the weaker requirement: $Y_i + q_i \le \frac{w_i Qu_i}{\delta}\theta < \frac{\beta_1}{\beta_1 - 1}(Y_i + q_i + \frac{V_i}{\gamma})$. In the absence of abatement, there may be both expected compliers ("just" compliers and over-compliers) and non compliers, since the allowances may equal, exceed or fall short of the firm's expected emissions. If overall over-compliance is not too large (see Appendix for the demonstration), the higher uncertainty, the higher the threshold level of enforcement and, as a consequence, the higher the number of non compliers. For a firm that would expect to over-comply ($Q^{*(i)} > Q$) in the absence of abatement, the expression on the RHS of (6) is formed of two parts: (i) the expected value of the fine (first term in parenthesis) and, (ii) the ratio between the threshold of action of the regulator and the firm expected output. For firms that are expected to "just comply" and for non compliers, on the other hand, expression (6) collapses to the simpler equation: $\left|K_i^{'}\right| = \gamma$. Thus, the marginal cost of abatement will be lower than the expected value of the fine for firms that would over-comply without abatement, the reason for this being that these firms not only do not need to abate to comply, but they can also reduce their abatement to reflect the over-allocation of permits. As for the uncertainty, the marginal value of the liability option depends positively on β_1 through the exponent and negatively through the threshold level $Q^{*(i)}$, which is at the denominator and is higher the higher the multiplier $\beta_1/(\beta_1 - 1)$. The first effect, however, prevails on the second one if the ratio between the expected and the threshold level of output is not too low as compared to the degree of uncertainty (for a formal proof, see Appendix). High over-compliance, on the other hand, combined with low volatility, may determine a situation where an increase in uncertainty will generate a higher marginal threat, thus causing a higher degree of abatement.

Comment. If each firm is seen in isolation, as it is the case when no permit market is considered, the expected present value of the fine is different for each firm and depends on its allowances and abating costs as well as on the costs that the regulator would incur to monitor and enforce the regulation in its specific case. With diseconomies of scale in abatement, larger fines will tend to reduce the emission level of the firm, as abatement marginal cost raises to meet the marginal value of the threat. Under high or moderate

uncertainty, a decrease in volatility of industry demand has a similar effect. Notice here the relationship with the literature on crime and punishment, where an increase in the probability of enforcement increases the expected penalty on the potential offender, thereby increasing her willingness to comply. In the case considered, a decrease in volatility of industry demand, by raising the marginal value of the option to sanction by the regulator, increases the probability that the option itself is exercised any time in the near future and the fine is actually applied. At the same time, a less volatile environment reduces the number of firms potentially liable, thereby reducing the average risk, and increasing the capacity of the regulator to elicit compliance from any specific firm. This result can also be interpreted as suggesting that the level of the sanction and the probability of enforcement are not the only two variables of importance in determining the degree of law abiding of a group of decentralized agents. Equally important may be the volatility of the sanction application. We must notice again that the opposite relation between volatility and compliance may occur in a low uncertainty environment. In this case, an increase in volatility, by boosting the expected level of the fine for over-compliant firms, may increase their willingness to abate even further and increase their level of compliance.

Proposition 3. The demand of permits will be homogeneous of degree zero in the permit price and the fine level and will be positive or negative, depending on the whether the emission effect exceeds or falls short of the allowances and the enforcement costs.

Consider now the introduction of a permit market and assume that each firm is allowed to sell or buy permits, by taking short or long positions in their contingent value as substitutes for abatement. We first find the value of the number of permits as long or short positions that maximize the value of the firm, by taking the first derivative of (3) under condition (5b) with respect to q_i and equating it to zero:

$$(7)\quad \frac{\partial \Pi_i}{\partial q_i} = -p + \gamma[\frac{\beta_1 \delta}{(\beta_1 - 1)}(Y_i + q_i + \frac{V_i}{\gamma})]^{-\beta_{1i}} (w_i u_i \frac{Q}{\delta})^{\beta_1} = 0$$

Equation (7) is a condition for a maximum since, (as shown in Appendix), $\frac{\partial^2 \Pi_i}{\partial q_i^2} < 0$ for all non zero values of q_i. It says that at the maximum, the firm equates the marginal benefit of abating its pollution level (the term in the RHS of equation (6)) and thereby reducing its contingent liability, with the marginal cost of permits (i.e. their market price). Solving for the amount of permits q_i yields:

$$(8)\ q_i = (\frac{\gamma}{p})^{\frac{1}{\beta_1}} (\frac{\beta_1 - 1}{\beta_1}) w_i u_i \theta \frac{Q}{\delta} - Y_i - \frac{V_i}{\gamma},$$

Comment. The possibility of trading permits allows the firms react to the threat of the fine by engaging in risk management. The heterogeneity of firms, in particular, opens the possibility of risk pooling (Mrozeki et al., 2004) . The quantity of permits demanded by the ith firm will be a function of the uncertainty (represented by the β parameter). In the case of maximum uncertainty $(\beta_1 = 1, \quad \sigma = \infty)$, the demand for permits will be zero, while it will equal to the pollution level of the ith firm net of the allowances and the enforcement costs in the case of perfect certainty $(\beta_1 = \infty, \sigma = 0)$. The demand for permits, furthermore, will be inelastic (perfectly elastic only in the case of maximum uncertainty) and will decrease towards zero as uncertainty decreases. Thus, with higher uncertainty we should expect each firm to benefit more from trade as its consumer (producer) surplus, given by the excess of the willingness to pay over the actual price paid (or the excess of the price received over the minimum acceptable), will be higher, coeteris paribus, because of the corresponding higher elasticity. In a less uncertain environment, vice versa, demand (supply) of permits will tend to be both larger and less reactive to price or fine changes. Note that firms may be both under and over-compliant, in the sense that they may decide to own less or more of the quantity of permits that would be strictly needed to abide by the regulation. The introduction of permits thus creates a phenomenon that is typically absent in the mechanism of crime and punishment: the transfer of credits from “virtuous” units to “vicious” ones, within a new market-like framework where “virtue” is motivated, at least to some extent, by “vice” willingness to pay.

Proposition 4: The equilibrium price for permits is the expected present value of the fine for a representative firm.

In equilibrium, short and long positions should cancel each other (demand equals supply), i.e. $\sum_{i=1}^{N} q_i = 0$ for N firms. As shown in the Appendix, this implies the following equilibrium price:

$$(9) \qquad p^* = \gamma[\frac{U}{\frac{\beta_1}{\beta_1 - 1}(Y + \frac{V}{\gamma})}]^{\beta_1} = (\frac{U/N}{U^*/N})^{\beta_1}\gamma$$

Comment. First, note that the equilibrium price is a decreasing function of uncertainty if the degree of over-compliance of the representative firm is not too high as compared to the degree of uncertainty (see Appendix for the proof), i.e. if not too many allowances have been distributed. In this case, because the demand for permits has elasticity equal to $\frac{1}{\beta_1}$ w.r.t. price, higher uncertainty will increase consumer surplus (the gain for net buyers of permits) both because of a higher elasticity and a lower price. The opposite will occur for the producer surplus of over-compliant firms that sell permits. In (9), the numerator on the RHS of the expression in parenthesis $U/N = \frac{1}{N}\frac{Q\theta}{\delta}\sum_{i=1}^{N} w_i u_i$ denotes total expected emissions of a representative firm, i.e. the expected emissions that each firm would have if firms were all identical, while the denominator of the same expression, $\frac{\beta_1}{\beta_1 - 1}(Y + \frac{V}{\gamma})/N$ is the correspondent value of the threshold of intervention for the emission level (the correspondent quantity for the production level is $Q^* = \frac{U^*}{\sum_{i=1}^{N} w_i u_i}$), i.e. the value of emissions that would prompt the regulator to intervene for the representative firm in the absence of permits. Note that both U and U^* are measured in emission units (e.g. tons of CO2). The ratio $\frac{U}{U^*} = g$, therefore , is a pure number and, as the ratio between the expected emission level and the threshold of intervention of the regulator for a representative firm, it can be interpreted as an indicator of the degree of compliance of the

industry, as compared with the threshold of action of the regulator. The measure ranges from zero (no emissions from the representative firm or very high threshold of intervention for maximum (over)-compliance) to one (the representative firm just at the verge of sanctioning or minimum compliance).

In particular, we can use this indicator to link the degree of expected compliance of the representative firm to the target level of emissions. Since the regulator aims at reducing expected emissions of a given amount, the target level Y is fixed as a fraction of total expected emissions[6] . Assuming rational expectations : $Y = \alpha \frac{\theta Q}{\delta} \sum_{i=1}^{N} w_i u_i$, where $< \alpha \leq 1$. Applying the definition of g , substituting and simplifying, we obtain:

$$(10) \quad g = \frac{U}{U^*} = \frac{\beta_1 - 1}{\beta_1} \frac{1}{\alpha + \frac{V}{\gamma U}}$$

In order for the representative firm to be considered compliant, g must be less than 1, which implies $\frac{\beta_1 - 1}{\beta_1} - \frac{V}{\gamma U} < \alpha \leq 1$. Since the term $\frac{\beta_1 - 1}{\beta_1}$ is inversely related to the volatility of the underlying process, this implies that the target fraction of total emissions will be bounded from below by the difference between an inverse measure of the uncertainty level (measured by the term in β_1), and the cost - revenue ratio (enforcement costs as a ratio to the total value of the fines collectible). The greater the uncertainty (the lower the term in β_1), and the higher, coeteris paribus, the cost-revenue ratio, the higher will be the reduction of emissions achievable. Notice that the uncertainty term will range from zero (total uncertainty or $\beta_1 = 1$) to 1 (perfect certainty or $\beta_1 = \propto$), while the cost revenue ratio ranges from a minimum value of zero (no enforcement costs) to any arbitrary large number. A more ambitious target (a lower α) can thus be achieved at higher levels of uncertainty, regardless of the cost–revenue ratio, since this term may only reduce the limit required for compliance. This result depends on the fact that a lower target implies smaller allowances, and this has the consequence of making the regulator more willing to engage in enforcement, the higher is uncertainty (since smaller allowances will lower her threshold

[6] In the case of ETS, the target reduction established in the Kyoto protocol was 8% of the historical level (the 90's) over the period 2007-2012, but each member country subsequently negotiated a different target level and for some countries (Greece, Portugal and Spain) the target was set above the historical level. Even in these cases, however, it can be argued that the agreement aimed at achieving a reduction of expected emissions.

of action the higher is uncertainty). A similar effect is also obtained by lower enforcement costs.

Note also that the target emission level can be more ambitious, coeteris paribus, the lower is the level of the fine γ. Intuitively, this is due to the fact that the price is the product of a discount factor (which is larger the larger is uncertainty) multiplied by the level of the sanction. A lower level of the target emissions α will be sustainable if a larger amount of permits will be bought and sold, and this in turn will require a lower equilibrium price (i.e. a lower expected value of the fine for the representative firm) to prevail. Given an arbitrary target level of emissions, on the other hand (i.e. any $0 < \alpha \leq 1$), compliance can be achieved by setting the fine γ at an appropriate level. An upper bound for such a level can be derived from expression (10), by imposing $g < 1$ and solving the ensuing dis-equation for the level of fine, to obtain: $\gamma < \frac{V}{U} / [\frac{\beta_1 - 1}{\beta_1} - \alpha]$. Compliance thus requires the fine to be bounded from above by a term inversely proportional to the difference between the (inverse) measure of uncertainty and the target ratio. Thus, the higher uncertainty, and the higher the target, the higher will the level of the fine required to achieve compliance. In order to have a positive level of the fine, however, the target ratio α cannot be larger than the measure of uncertainty, which is comprised between zero and 1, and will have to be lower and lower the higher the level of uncertainty. On the other hand, higher uncertainty and/or a more ambitious target will imply a less stringent upper bound on the fine.

It can also be shown (see Appendix) that:

(11) $g^{\beta_1} = Ee^{-rt}$

Where Ee^{-rt} is the expected value of the discount factor for the stochastic time of the regulator's intervention for the representative firm in absence of permits. This value will decline with uncertainty unless allowances are too large as compared with expected emissions and with the level of uncertainty. The time t, in fact, represents waiting time in the sense of Dixit (1993a, pp.54-57), and denotes the first time the fine would be applied starting from the initial level of emissions U.[7] Since holding

[7] Note that for a Markovian process this level would also be the expected level of emissions at the time the discount factor is calculated.

permits would postpone the day of reckoning, the time in (13) is the earliest time at which the average firm will be sanctioned.

The market for permits creates an interdependence among the firms, which allows to exploit their heterogeneity to achieve emission reduction through inter-firm re-allocation of resources. While heterogeneity of emission levels, allowances and enforcement costs motivates trading, however, what ultimately drives the level of permit price for each given distribution of market shares w_i, is the aggregate emission performance of the industry. In this context, the notion of the representative firm emerges directly from the analysis and not as an assumption as in the famous definition by Alfred Marshall (1925). Expression (9), in fact, predicts that the equilibrium price of permits will equal the expected present value of the fine *ad valorem* rate (expressed as a value per ton of emissions) for the representative firm. As expression (10) shows, the discount factor is a function of the ratio between the average emission level and the level at which the regulator would impose a fine on a representative firm. If this ratio were unity, the regulator would be exactly at the threshold of sanctioning for the representative firm, and the equilibrium price of permits would be equal to the fine *ad valorem* rate γ. It should be the objective of the regulator, however, to achieve a level of g is less than 1, so that operational compliance would be obtained and there would be no need to sanction.

Table 1 shows how the equilibrium price would evolve under alternative values of β and $g\frac{\beta_1}{\beta_1 - 1} = \frac{U}{Y + \frac{V}{\gamma}}$, if the value of the fine were fixed at 100 Euro per ton.[8] Note that the pollution ratio g is a hybrid term between the ratio of the targeted emissions cap $Y = \sum Y_i$ relaxed by the aggregate enforcement costs (converted to emission tons by the ad valorem fine rate) to the total emissions U of the sector.

Table 1 shows that it is not only the stringency of the target emissions that determines price but the uncertainty of emissions (driven by industrial output) and the regulators behavior costs. For plausible values of the pollution ratios g and the uncertainty parameter, in fact, the price of permits will tend to be a fraction of the value of the fine and this fraction will be

[8] This value is the fine under the second phase of the European Trading Scheme. However, under the ETS the fine is supplemented by the need to cover the gap in emissions to allowances by a market purchase. Also, purchases are allowed from the mechanisms of Clean Development and Joint Implementation. We have abstracted from these complications for simplicity of exposition.

lower the higher the amount of allowances distributed beforehand. Note that for a standard deviation of 0.1, which corresponds to a value of 2 of the parameter β_1 under the assumption that the underlying stochastic process has no drift, market price of permits is only 21% of the fine even if the degree pollution ratio is as high as 95%, while it falls to 6.75 Euros (a figure close to the current price of one ton of carbon under ETS) for a pollution ratio equal to 50%. This means that with emissions uncertainty and a regulator with the option to sanction, the rate of the fine does influence the price of permits but does not drive the price of permits to anywhere near its level except when high uncertainty on the stochastic emissions is combined with average emissions very close to the average threshold of regulator's intervention. Note also that the relationship between equilibrium price and uncertainty, even though mainly negative in the relevant range of values, is not monotonic . The price, in fact, tends to reach a maximum, as uncertainty increases, for a volatility value that is higher the lower the pollution ratio (the higher the degree of expected compliance).

Table 1

Equilibrium Permit Prices for Different levels of Volatility and Pollution Ratios (Fine level=100 Euros)

Expected emissions/equivalent allowances ($g \frac{\beta_1}{\beta_1 - 1} = \frac{U}{Y + \frac{V}{\gamma}}$)

St.Dev. (of log Q)	0,95	0,90	0,85	0,80	0,75	0,70	0,50	0,30	0,10
2,8109	0,90	0,85	0,80	0,75	0,71	0,66	0,47	0,28	0,09
1,2267	3,88	3,66	3,45	3,24	3,02	2,81	1,98	1,16	0,36
0,6681	9,06	8,51	7,97	7,43	6,90	6,38	4,33	2,41	0,68
0,4899	12,54	11,72	10,92	10,12	9,33	8,56	5,62	2,97	0,75
0,3931	15,08	14,02	12,98	11,96	10,96	9,99	6,34	3,18	0,72
0,3299	17,02	15,73	14,48	13,26	12,08	10,93	6,71	3,20	0,65
0,2845	18,54	17,05	15,60	14,20	12,85	11,55	6,85	3,11	0,57
0,2499	19,76	18,07	16,44	14,88	13,38	11,94	6,85	2,95	0,48
0,2225	20,75	18,88	17,08	15,36	13,72	12,16	6,75	2,76	0,40
0,2002	21,57	19,52	17,56	15,70	13,93	12,26	6,58	2,56	0,34
0,1815	22,26	20,03	17,92	15,92	14,04	12,27	6,37	2,35	0,28
0,0844	25,35	21,61	18,26	15,27	12,62	10,30	3,82	0,85	0,03
0,0425	25,78	20,82	16,61	13,08	10,13	7,72	2,04	0,27	0,00
0,0000	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Note: The calculations are based on the assumptions of a 5% interest and 1% growth rate .

Proposition 5. In equilibrium, demand for (supply of) permits for the i-th firm equals the difference between its "fair " endowment of permits (adjusted for enforcement costs) and its actual adjusted endowment.

We define as a "fair " endowment a share of total permits equal to the firm emission share over the industry total. Since the firm specific enforcement costs cause a proportional increase in the amount of emissions allowed, we can define a similar "fair " share for the enforcement costs divided by the ad valorem fine. The sum of the two variables can be denoted as "permit equivalents" or permits adjusted for enforcement costs. Substituting p^* in the expression for q_i in (8), we find the expression for the demand (supply) of permits for each level of production for the i-th firm:

$$(12) \quad q_i^* = (\omega_i Y - Y_i) + (\frac{\omega_i V - V_i}{\gamma})$$

where $\omega_i = \frac{w_i u_i}{\sum_i w_i u_i}$ is the *i-th* firm's share of the emission levels of industry.

Comment. When the permit market is in equilibrium, the level of the fine acts on the demand (supply) of permits only through the enforcement costs. The larger the fine, the less important will be enforcement costs and the unambiguously smaller will be the demand (supply) of permits for each firm. Remarkably, in equilibrium firm demand (supply) of permits will not depend on stochastic demand Q of the industry, nor on uncertainty, but only on the extent to which its emission cap (i.e. its deductible) and the enforcement costs that it generates are smaller (greater) than its share of the corresponding totals. Note that the ratio between enforcement costs and the fine represents an allowance equivalent, that is the tons of emissions corresponding, at the level of the sanction γ, to the allowances that would have the same effect of the enforcement costs. An over-allocation of allowances (that is a greater share of allowances with respect to one's pollution share) yields a gain for the firm in question equal to the value (at the equilibrium price of permits) of the excess rights received. An under-allocation implies a corresponding loss. The difference in enforcement costs across firms will act in the same way. Firms presenting higher enforcement

costs than their fare share of the total will experience a gain from the corresponding higher level of equivalent allowances, while the opposite will occur for firms easier to control than average. The demand or supply of permits, therefore, will depend on the distribution performed through grandfathering and on the distribution of enforcement costs, and this will determine a distribution of gains and losses across firms.

From (12) , by multiplying both sides for the firms' shares w_i and summing over all firms, it also follows that in equilibrium, the representative firm will be a net seller or a net buyer of permits according to whether the average share of total emissions is greater or smaller than the average share of total allowances i.e.: $\sum_i w_i q_i^* = Y[\sum_i w_i(\omega_i - \frac{Y_i}{Y})]$. If allowances perfectly reflect emission shares, therefore, the representative firm would buy no permits and any increase in uncertainty would have no average impact on the industry. However, in most common situations where uncertainty is high, not too many allowances have been distributed and/or enforcement costs are not too large, increased uncertainty would cause price elasticity to be higher and permit prices to be lower, thereby redistributing the gain from trading in favor of net buyers. If lower allowance shares were assigned to higher emission firms, on the other hand, the representative firm would be a net buyer, so that average industry surplus from trading would increase with an increase in uncertainty. The opposite would occur if allowances were mostly distributed to higher polluters.

Proposition 6. In equilibrium, permits will be distributed in such a way, that all firms will be equally vulnerable to the regulator's action, i.e. the critical level of output at which the regulator is expected to apply the fine is the same for all firms.

Substituting the equilibrium level of permits for each firm in (14) back into (5b), we find that, once equilibrium is reached in the permit market, the output level at which the ith firm will expect the regulator to act is:

$$(13)\ w_i \frac{Q^{*(i)}}{\delta} u_i \theta = w_i \frac{Q^*}{\delta} u_i \theta = \frac{\beta_1 \omega_i}{\beta_1 - 1}(Y + \frac{V}{\gamma}), \quad \text{i.e.}$$

$$\rightarrow \frac{\theta Q^{*(i)}}{\delta} = (\frac{1}{\sum w_j u_j}) \frac{\beta_1}{\beta_1 - 1}(Y + \frac{V}{\gamma}) \text{ for all } i.$$

Comment. At the level of equilibrium permits each firm will produce excess emissions equal to its "fair share" of total emissions: $\omega_i(\frac{Q}{\delta}u\theta - Y - \frac{V}{\gamma})$. Thus, the level of compliance of each firm will be a proportion of the level of compliance of the representative firm. While the proportion will be different for each firm, whether the firm would be compliant or not will not depend on the firm's characteristics. The equilibrium price of permits is the expected present value of the fine for the representative firm, and all firms buy permits up to the point where the reduction of the threat of sanction equals the permit price, each firm expecting the regulator to impose the fine as if it were a representative firm. The risk pool created by the market for permits has thus the consequence of simplifying expectations upon the regulator's decision to act, since it removes the uncertainty due to the differences in allowances and monitoring costs across firms. All these differences, in fact, are traded away through the permit market and eventually all firms look at the threat of regulator's sanction as if they were integrated into one large enterprise. Intuitively, since at the margin the price of permits equals the willingness to pay of each firm to reduce its risk of being sanctioned, this risk is eventually the same for all firms. Thus, in a permit system all firms will tend to be compliant, in the sense that they will all tend to be below the level at which the regulator will intervene ($Q\sum w_j u_j < \frac{\beta_1}{\beta_1 - 1}(Y + \frac{V}{\gamma})$). This might explain the fact that the record of the ETS system looks so good in terms of the degree of compliance (more than 95% of all firms compliant in most of the countries involved).

The Firm's Output level and Market Equilibrium

We now go back to the optimizing problem for the firm and find the value of market share w_i. We assume that we are dealing with a regulated or semi-regulated industry such as, for example, a power utility. We hypothesize that the industry bases its production plans on the expected output price . In the short run the number of firms is given and each firm targets its share of expected stochastic market demand, taking into account the administered price that it will be allowed to charge, investment costs plus the costs involved by the threat of pollution fines and/or of emission

permits. If the bids for production are too low (the sum of the shares offered by the firms falls short of unity) as compared to demand, the government increases the administered price, while it lowers it if the output exceeds expected demand. Market equilibrium is thus reached when the market price for output is such that expected demand is equal to the productive capacity allocated and each firm produces its optimal share of total output.

Proposition 7. In equilibrium, the market share of the i-th firm is such that each firm has the same total marginal cost.

A full proof of this proposition is contained in Appendix. Here we briefly note that when each firm holds its optimum market share, output price must equal total marginal cost :

$$(14) \quad P = C_i^{'} + K_i + \gamma u_i \theta [\frac{U}{U^*}]^{\beta_1 - 1_1} = TMC_i \quad i = 1,2...N \,,$$

where TMC_i denotes total marginal production cost for the ith firm in equilibrium. This cost is formed by direct marginal cost $C_i^{'}$, abatement cost per unit of output K_i and the marginal increase in the value of the liability option in response to an increase in output.

Recalling that $\sum w_i = 1$ and that $U = \frac{1}{N} \sum w_i u_i \theta \frac{Q}{\delta}$:

$$(15) \quad P = \sum w_i C_i^{'} + \sum w_i K_i + p \frac{U^*}{Q} = TMC \,,$$

where TMC is total marginal cost in equilibrium for the representative firm given by direct marginal costs plus abatement costs per unit of output plus the average cost of emissions per unit of output . Substituting (15) back into (14), we conclude that $TMC_j = TMC$.

Comment. Market equilibrium for production completes the effect of equilibrium in the market for permits. The choice of output shares by the firms and output price determination, in fact, ensures that all firms face the same trade-off between emissions and output size at the margin. If it were

not so, in fact, for a given output price, a firm with lower (higher) total marginal costs could profitably bid for a higher (lower) market share so that the situation would not be of equilibrium. With no heterogeneity of any sort, the equilibrium share of the output would be equal for each firm. Heterogeneity thus determines a higher or lower share depending, respectively, on the lower or higher unit cost of production, on the lower or higher cost of abatement and/or the lower or higher emission level compared to the average. Because a key element of total marginal cost is the unit cost of permits in equilibrium, and this can be positive or negative depending on whether the firm is a net buyer or a net seller of permits, uncertainty plays a major role in this respect. On one hand, in fact, for a given nominal level of the fine, higher uncertainty will imply a lower price for permits, thus reducing marginal costs and output prices for all net buyers and increasing it for all net sellers. On the other hand, as we have seen in the discussion of expression (10), higher uncertainty may also require a lower fine (or a higher level of allowances) for compliance to be achievable.

Conclusions

In this paper, we have explored the effects of dynamic uncertainty in a regulatory environment on the risk management of firms and regulators. The primary uncertainty considered arises because the output of the industry, and related emissions, grow stochastically in response to stochastic growth of demand over time. Output stochastic behavior depends on the intrinsic characteristics of the industry (exemplified by the power sector) and by the fact that a host of random variations affect both demand and supply. Moreover, the combination of this uncertainty and imperfect monitoring generates further uncertainty in the behavior of the agents involved. In turn, this implies that perfect enforcement of regulations is not optimal and that a degree of non compliance and moderate sanctioning may be more efficient if it is indirectly controlled by a market allocation mechanism, such as a system of tradable permits.

With respect to the usual approach of the economics of crime and punishment, representing the threat of a sanction as a real option has the effect of introducing an additional dimension in the analysis: the volatility of

an underlying stochastic process. Because an option that is not too far from exercise (too much "out of the money") tends to have a greater value the larger the volatility on the underlying asset, the more uncertain is the sanction, the greater is its overall deterring strength. In the case of regulated emissions, this implies a paradox. On one hand, in fact, over the range of interest of the policy maker, an increase in uncertainty on the output and related emission levels tends to increase the total value of the option to sanction held by the regulator. On the other hand, increased uncertainty would reduce the value of the individual expected penalty, so that the number of perspective non compliers would increase. This would expand the power of the regulator, as measured on her potential to sanction a large number of firms in the industry. At the same time, however, it would force her to a costly policy of individual monitoring and enforcement for a number of non compliant firms that would be larger the larger is uncertainty.

A market for permits, on the other hand, by allowing the firms to manage the risks of non compliance by pooling and sharing them according with the firms' characteristics, would allow the regulator to set and calibrate her enforcement policies only with reference to the representative firm. This would allow the achievement of a target on emission reduction, by threatening to impose the sanction for non compliers at a common expected time and level of aggregate production (though in proportion to each firm's size). Under these conditions, general compliance may be achieved and the demand (supply) of permits of each firm in equilibrium will not depend on uncertainty, but will be larger the smaller is the fine threatened as a sanction for non compliance.

With respect to the economic costs that have to be borne to internalize the pollution externality, the degree of uncertainty appears a crucial variable with a subtle influencing role on the degree of compliance and the size of the fine. In general, higher uncertainty favors the reduction of emissions by making firms more reactive to the threat of being sanctioned (higher price elasticities of both demand and supply of permits) and more willing to pool risks and share the costs of meeting the target by engaging in permit trading. Thus, higher dynamic uncertainty may make market adjustment more effective and compliance more likely. Greater uncertainty, however, tends to reduce the market price of permits and the corresponding expenditure (revenue) from trading of each firm, since in equilibrium demand or supply of permits will only depend on pollution levels, endowments, enforcement costs and the size of the fine. Because demand and supply of permits will be more elastic, the higher the uncertainty under which the firms will operate, the higher will be the surplus gained by net buyers and the lower the surplus

gained by net sellers. By reducing the expected present value of the fine, increasing uncertainty reduces the incentive to abate, increases the incentive to share emission risks and redistributes gains across firms in favor of potential non compliers. Higher uncertainty and higher heterogeneity across firms are thus both assets for the regulator to achieve emission control through a market for permits, by providing commitment incentives with lower costs for monitoring or other enforcement activities.

In conclusion, the traditional approach to optimal law enforcement relies on the assumption that compliance is an increasing function of the expected value of the sanction. In that context, if uncertainty of enforcement decreases, the only possibility to maintain unaltered the deterrent power of the sanction is to increase its value. Dynamic uncertainty, however, increases the risk over time to be sanctioned for non compliance for a larger and larger number of firms .As a consequence, the introduction of a market for pollution permits can be seen as an effective policy to exploit agents' increasing willingness to engage in risk management to cover themselves from volatile sanctions under dynamic uncertainty.

References

— Baumol, W. J. and Oates, W. E. 1988. The Theory of Environmental Policy, Second edition. Cambridge University Press

— Beavis, B. and Walker M., 1983. 'Random Wastes, Imperfect Monitoring, and Environmental Quality Standards', Journal of Public Economics 21, 377–387.

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

— Benoit, J. and Osborne, M.J., 1995. Crime, punishment and social expenditure. *Journal of Institutional and Theoretical Economics* 151 (2), 326–347.

— Crawls E. and Vereeck L. 2005 ."Taxes, Tradable Rights and Transaction Costs", The European Journal of Law and Economics, 20(2), 199-223.

— Crocker, T. D. 1966. "The Structuring of Atmospheric Pollution Control Systems," in Wolozin H. ed. The Economics of Air Pollution. Norton Press.

— Dales, John H. 1968. Pollution, Property and Prices. University of Toronto Press.

— Dixit, A. K. and Pindyck, R. S. 1994. Investment under Uncertainty. Princeton Princeton University Press.

— Field, B. C. 1997. Environmental economics: An introduction. 2nd ed. Hightstown, N.J.: McGraw-Hill.

— Graetz, M.J., Reinganum, J.F. and Wilde, L.L., 1986. The tax compliance game: toward an interactive theory of law enforcement. *Journal of Law, Economics and Organization.* 22, 1–32.

— Ishii, Jun and Yan, J. 2004. "Investment under Regulatory Uncertainty: US Electricity Generation Investment Since 1996." Working Paper, Center for the Study of Energy Markets.

— Kahn, J. R. 1998. The economic approach to environmental and natural resources. 2nd ed. Fort Worth, Tex.: Dryden.

— Kees V. K. 2007. "Towards a compliance and enforcement strategy in the Netherlands" , Netherlands Emission Authority

— Knudsen, O.K., and Scandizzo P.L. 2005." Bringing Social Standards into Project Evaluation Under Dynamic Uncertainty", Risk Analysis 25 (2) , 457–466.

— Landgrebe, J. 2009 . Verification in the EU ETS: German Perspective, German Emissions Trading Authority (DEHSt)

at the Federal Environment Agency (UBA)ICAP China Conference Beijing, 12-13.

— Malik A.S. ,1990 "Markets for pollution when firms are non compliant", Journal of environmental and Economic Management, 18, 97-106.

— Marshall, A., 1925, Principles of Economics. 8th ed. London: Macmillan

— Milliman, S. R. and Prince, R. 1989. "Firm Incentives to Promote Technological Change in Pollution Control," Journal of Environmental Economics and Management. 17, 247-65.

— Montero, J. P. 2005. " Pollution markets with imperfectly observed emissions", RAND Journal of Economics Vol. 36, No. 3, Autumn 2005 pp. 645-660

— Montgomery, W. D. 1972. “Markets in Licenses and Efficient Pollution Control Programs,” Journal of Economic Theory, 395-418.

— Mrozeki, J.R. and Keeler A.G, 2004. “Enforcing Tradable Permits Regulation When Emissions Are Stochastic”, Environmental & Resource Economics 29: 459–481

— Pigou, A. C. 1920. The Economics of Welfare. Macmillan and Company.

— Plourde, C. and D. Yeung (1989), ‘A Model of Industrial Pollution in a Stochastic Environment’,Journal of Environmental Economics and Management 16, 97–105.

— Pollock, D.S.G. (2008) , Environmental Economics, www.le.ac.uk/users/dsgp1/COURSES/ENVIRONS/coase.pdf

— Stranlund J.K., Dhanda K.K. (1999) Endogenous Monitoring and Enforcement of a transferable Emission Permit System. Journal of environmental and Economic Management, 18, 97-106.

— Shapiro, C. 1984. Review of “Controlling Industrial Pollution: The Economics and Politics of Clear Air” by Robert W. Crandall, Journal of Economic Literature, Vol. 22, 2, 625-27.

— Stavins, R. N. 1998. “What can We Learn from the Grand Policy Experiment? Lessons from S02 Allowance Trading.” The Journal of Economic Perspectives, 12, 69-88.

— Stokey, N.2009. The Economics of Inaction. Princeton University Press

— Tietenberg,T. 2000. Environmental and natural resource economics, 4th ed. New York: Harper-Collins.

— Teisberg, E. O. 1994, “An Option Valuation Analysis of Investment Choices by a Regulated Firm.” Management Science. 40, 535-48.

— Thomas, T. 2005 BP’s perspectives on EU ETS, www.perf.org./ppt/Thomas.ppt

— Weber, D. W. 2002. “Pollution permits: a discussion of fundamentals”. The Journal of Economic Education

— Weishaar, Stephan , 2007, “CO_2 emission allowance allocation mechanisms, allocative efficiency and the environment: a static and

dynamic perspective", The European Journal of Law and Economics, 24 (1), 29-70.

— Wirl, F., J. Noll, 2007. "Abatement and Permits when Pollution is Uncertain and Violations are Fined", Environmental and Resource Economics, September, 1502-1573.

— Woerdeman, E. , Couwenberg, O. and Nentjes A., 2009, "Energy prices and emissions trading: windfall profits from grandfathering?" , The European Journal of Law and Economics

— Xepapadeas, A. 1992, 'Optimal Taxes for Pollution Regulation: Dynamic, Spatial and Stochastic Characteristics', Natural Resource Modeling 6(2), 139–170.

APPENDIX

Detailed Proofs of the Paper's Propositions

Proof of Proposition 1. For each firm, the value of the contingent liability given by the threat of the regulator's action is equal to the expected present value of the total penalty to be paid in the case of non compliance.

The value matching condition (Dixit and Pyndick, 1994, p. 141) for the ith firm requires that when the regulator decides to exercise its option to fine (i.e. at the production level $Q^{*(i)}$,different for each firm), the following equation is satisfied :

(A1) $$A_i Q^{*\beta_1} = [\gamma(\frac{w_i u_i \theta Q^*}{\delta} - Y_i - q_i) - V_i]$$

where for simplicity we have dropped the (i) superscript to denote the threshold level of the regulator's intervention for the output in the case of the i-th firm.

The smooth pasting condition (Dixit and Pyndick, 1994, p.141) requires that the equality is satisfied also by the derivatives of both sides of (A1) w.r.t. to Q:

(A2) $$\beta_1 A_i Q^{*\beta_1 - 1} = \gamma \frac{w_i u_i \theta}{\delta}$$

Re-writing (A2) as $A_i Q^{*\beta_1} = \gamma \frac{w_i u_i \theta Q^*}{\delta \beta_1}$ and substituting into (A1) yields:

(A3) $$\gamma \frac{w_i u_i \theta Q^{*(i)}}{\delta \beta_1} = [\gamma (\frac{w_i u_i \theta Q^{*(i)}}{\delta} - Y_i - q_i) - V_i]$$

Solving for $Q^{*(i)}$:

(A4) $$Q^{*(i)} = \frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}$$

This is the threshold level of industry output at which the regulator will intervene to fine the i-th firm..

Substituting (A4) into (A2), and solving for the constant A_i yields:

(A5) $$A_i = \gamma \frac{w_i u_i \theta Q^{*(i)1-\beta_1}}{\delta \beta_1} = \gamma \frac{w_i u_i \theta}{\delta \beta_1} [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{1-\beta_1}$$

Thus, the expression for the option to fine the ith firm can be written, by substituting into it the expression found for the constant A_i:

(A6) $$A_i Q^{\beta_1} = (\frac{Q}{Q^{*(i)}})^{\beta_1} \gamma \frac{w_i u_i \theta Q^{*(i)}}{\delta \beta_1}$$

The LHS of expression (A6) is the present value of the option, while the second part of the RHS is the value of the option at the moment of its exercise. The first part of the RHS, therefore, is the expected discount factor (with stochastic time) that maps the future value of the option when it is expected to be exercised to the present value scale, i.e.:

(A7) $$A_i Q^{\beta_1} = E e^{-rt_i} (\gamma \frac{w_i u_i \theta Q^{*(i)_1}}{\delta \beta_1}) = E e^{-rt_i} [\frac{\gamma (Y_i + q_i) + V_i}{\beta_1 - 1}]$$

where $Ee^{-rt_i} = (\frac{Q}{Q^{*(i)}})^{\beta}$ and t_i indicates the time at which the option is expected to be exercised for the ith firm.

Differentiating Ee^{-rt_i} with respect to β_1, we obtain:

(A8) $$\frac{\partial E e^{-rt_i}}{\partial \beta_1} = \gamma (\frac{Q}{Q^{*((i)}})^{\beta_1} [\ln(\frac{Q}{Q^{*(i)}}) + \frac{1}{(\beta_1 - 1)}]$$

while differentiating the term in parenthesis in (A6) yields:

(A9) $$\frac{\partial [\frac{1}{\beta_1 - 1} (Y_i + q_i + \frac{V_i}{\gamma})]}{\partial \beta_1} = -\frac{1}{(\beta_1 - 1)^2} (Y_i + q_i + \frac{V_i}{\gamma})$$

Combining (A8) and (A9), we obtain:

$$\text{(A10)}\quad \frac{\partial(A_i Q^{\beta_1})}{\partial \beta_1} = \frac{1}{\beta_1 - 1}(Y_i + q_i + \frac{V_i}{\gamma})\gamma(\frac{Q}{Q^{*(i)}})^{\beta_1}[\ln(\frac{Q}{Q^{*(i)}}) - \frac{1}{(\beta_1 - 1)}]$$

This derivative is lower than zero for $\left|\ln(\frac{U}{U^*})\right| < \frac{1}{(\beta_1 - 1)}$, indicating an increase in the value of the liability with an increase in uncertainty, provided that the option is not too much "out of the money", i.e. the degree of compliance (the extent of the pollution level being below the target) is not too high.

Proof of Proposition 2. At the optimum each firm equates marginal abatement costs to the expected present value of the fine, adjusted for the degree of compliance.

Consider expression (6) in the text:

$$\text{(A11)}\quad \Pi_i = \frac{Pw_i Q}{\delta} - \frac{C_i(w_i Q)}{\delta} - \frac{w_i Q K(u_i)}{\delta} - pq_i - \gamma[\frac{\theta w_i Q u_i}{\frac{\beta_1 \delta}{\beta_1 - 1}(Y_i + q_i + \frac{V_i}{\gamma})}]^{\beta_1}(\frac{Y_i + q_i + \frac{V_i}{\gamma}}{\beta_1 - 1})$$

We can re-write it as follows:

$$\text{(A12)}\quad \Pi_i = \frac{Pw_i Q}{\delta} - \frac{C_i(Qw_i)}{\delta} - \frac{w_i Q K_i(u_i)}{\delta} - pq_i - \gamma\frac{\theta w_i u_i^{\beta_1}}{\delta \beta_1}[\frac{\beta_1 \delta}{\beta_1 - 1}\frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i \theta}]^{1-\beta_1} Q^{\beta_1}$$

Differentiating with respect to u_i and equating to zero yields:

(A13) $$\frac{\partial \Pi_i}{\partial u_i} = -\frac{w_i Q K_i^{'}}{\delta} - \gamma \frac{\theta w_i}{\delta}[\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{1-\beta_1} Q^{\beta_1} = 0 \rightarrow$$

by virtue of (A7)

(A14)

$$\left|K_i^{'}\right| = \gamma\theta(\frac{Q}{Q^{*(i)}})^{\beta_1 - 1} = Ee^{-rt_i}\gamma(Q^{*(i)}/Q) = Ee^{-rt_i}\gamma(\frac{\beta_1}{\beta_1 - 1}(Y_i + q_i + \frac{V_i}{\gamma})/w_i \frac{Q}{\delta}\theta u_i)$$

Using a quadratic approximation to the abatement function:

(A15) $$u_i^* = \gamma \frac{1}{k_{i2}} Ee^{-rt_{ii}} \frac{Q^{*(i)}}{Q} + \frac{k_{i1}}{k_{i2}}$$

Consider the derivative of the RHS of (A14) w.r.t. β_1:

(A16) $$\frac{d[Ee^{-rt_i}\gamma(\frac{Q^{*(i)}}{Q})]}{d\beta_1} = (\frac{Q}{Q^{*(i)}})^{\beta_1}\gamma[\ln(\frac{Q}{Q^{*(i)}}) + \frac{1}{\beta_1}]$$

For $\frac{Q}{Q^{*(i)}} < 1$, $\left|\ln(\frac{Q}{Q^{*(i)}})\right| < \frac{1}{\beta_1}$, the derivative in (A16) will be positive, indicating that the marginal value of the threat decreases with uncertainty. Since the value of the total liability instead increases with uncertainty (see (A10)), this result implies that larger uncertainty is matched by a larger total amount of emissions, provided that the pollution rate is sufficiently close to the target. This will be due both to the increased amount of emissions of the representative firm and the larger amount of firms that will be expected to be non compliers.

Proof of Proposition 3. The demand of permits will be homogeneous of degree zero in the permit price and the fine level and will be positive or negative, depending on the whether the emission effect exceeds or falls short of the allowances and the enforcement costs.

We can re-write expression (A11) as follows:

(A17) $$\Pi_i = \frac{P w_i Q}{\delta} - \frac{C_i(Q w_i)}{\delta} - \frac{w_i Q K(u_i)}{\delta} - p q_i - \gamma \frac{\theta w_i u_i}{\delta \beta_1}[\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{1-\beta_1} Q^{\beta_1}$$

Differentiation with respect to q_i yields:

(A18) $$\frac{\partial \Pi_i}{\partial q_i} = -p - \gamma \frac{\theta w_i u_i}{\delta \beta_1} \frac{1}{w_i u_i \theta} (1-\beta_1) \frac{\beta_1 \delta}{\beta_1 - 1} [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{-\beta_1} Q^{\beta_1}$$

Simplifying and equating to zero, we obtain the first order conditions to maximize Π_i with respect to q_i:

(A19) $$\frac{\partial \Pi_i}{\partial q_i} = -p + \gamma [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{-\beta_1} Q^{\beta_1} = 0$$

Equating to zero and solving for q_i :

(A20) $$q_i = (\frac{\gamma}{p})^{\frac{1}{\beta_1}} (\frac{\beta_1 - 1}{\beta_1}) w_i u_i \theta \frac{Q}{\delta} - Y_i - \frac{V_i}{\gamma}$$

This value of $q_i = \arg\max \Pi_i$ if and only if :

(A21) $$\frac{\partial^2 \Pi_i}{\partial q_i^{\,2}} = -\beta_1 \gamma [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{-(\beta_1 + 1)} Q^{\beta_1} < 0$$

Substituting into (A21) the value of q_i obtained in (A20), we obtain:

(A22) $$\frac{\partial^2 \Pi_i}{\partial q_i^{\,2}} = -\beta_1 \gamma [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + (\frac{\gamma}{p})^{\frac{1}{\beta_1}} (\frac{\beta_1 - 1}{\beta_1}) w_i u_i \theta \frac{Q}{\delta} - Y_i - \frac{V_i}{\gamma} + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{-(\beta_1 + 1)} Q^{\beta_1}$$

Simplifying:

(A23) $$\frac{\partial^2 \Pi_i}{\partial q_i^{\,2}} = -\beta_1 \gamma [\frac{\beta_1 \delta}{\beta_1 - 1}((\frac{\gamma}{p})^{\frac{1}{\beta_1}}(\frac{\beta_1 - 1}{\beta_1})\frac{Q}{\delta})]^{-(\beta_1+1)} Q^{\beta_1}$$

For positive values of its components the value of the second derivative in (A23) is always less than zero, so that the condition in (A21) denotes a maximum.

Proof of Proposition 4: The equilibrium price for permits is the expected present value of the fine for a representative firm.

In equilibrium $\sum_{i=1}^{N} q_i = 0$, so that, using (A20), and summing over all firms, we obtain:

(A24**)** $$\sum q_i = (\frac{\gamma}{p})^{\frac{1}{\beta_1}}(\frac{\beta_1 - 1}{\beta_1})\sum w_i u_i \theta \frac{Q}{\delta} - Y - \frac{V}{\gamma} = 0$$

where $Y = \sum Y_i$ and $V = \sum \frac{V_i}{\gamma}$. Denote also with $U = \frac{1}{N}\sum w_i u_i \theta \frac{Q}{\delta}$ the expected level of emissions of an average firm over the time span on which the fine is applied and with $U^* = \frac{1}{N}\frac{\beta_1}{\beta_1 - 1}(Y + \frac{V}{\gamma}) = \frac{\sum w_i u_i}{N}\theta Q^*$ the average value of the threshold of intervention, i.e. the value of emissions that would prompt the regulator to intervene for an average firm without any permits. Solving (A24) for the equilibrium price p, yields:

(A25) $$p = \gamma(\frac{U}{U^*})^{\beta_1}$$

Recalling (A7):

(A26) $$p = Ee^{-rt}\gamma$$

where Ee^{-rt} is the expected value of the discount rate for the (stochastic) time at which the regulator can be expected to levy a fine on an average firm with no permits. This value decreases with uncertainty, if the level of emissions is not too low as compared to the threshold of intervention of the regulator. This can be seen by differentiating (A19) w.r.t. β_1:

$$\text{(A27)} \quad \frac{dp}{d\beta_1} = \gamma\left(\frac{U}{U^*}\right)^{\beta_1}\left[\ln\left(\frac{U}{U^*}\right) + \frac{1}{(\beta_1 - 1)}\right]$$

Since $U < U^*$, it must be $\left|\ln\left(\frac{U}{U^*}\right)\right| < \frac{1}{(\beta_1 - 1)}$ for the term in the square brackets to be positive, provided that the degree of compliance is "moderate". The derivative in (A27) is thus positive, indicating that the permit price declines with uncertainty.

Proof of Proposition 5. In equilibrium, demand for (supply of) permits for the i-th firm equals the difference between its "fair " endowment of permits (adjusted for enforcement costs) and its actual adjusted endowment.

Substituting equation (A20) into equation (A14) yields:

$$\text{(A21)} \quad q_i = \left(\frac{\gamma}{\gamma\left(\frac{U}{U^*}\right)^{\beta_1}}\right)^{\frac{1}{\beta_1}}\left(\frac{\beta_1 - 1}{\beta_1}\right) w_i u_i \theta \frac{Q}{\delta} - Y_i - \frac{V_i}{\gamma}$$

Simplifying and applying the definitions of U and $U*$, we obtain:

$$\text{(A22)} \quad q_i = \omega_i\left(Y + \frac{V}{\gamma}\right) - \left(Y_i + \frac{V_i}{\gamma}\right)$$

where $\omega_i = \frac{w_i u_i}{\sum_{j=1}^{N} w_j u_j}$ is the emission share of the i-th firm.

Proof of Proposition 6. In equilibrium, permits will be distributed in such a way, that all firms will be equally vulnerable to the regulator's action, i.e. the critical level of output at which the regulator is expected to apply the fine is the same for all firms.

Substituting the expression for q_i found in (A22) into (A4), we find:

$$(A23) \qquad Q^{*(i)} = \frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + \omega_i (Y + \frac{V}{\gamma}) - (Y_i + \frac{V_i}{\gamma}) + \frac{V_i}{\gamma})}{w_i u_i \theta}$$

Simplifying, we obtain:

$$(A24) \quad w_i u_i \theta Q^{*(i)} = \frac{\beta_1 \delta}{\beta_1 - 1} \omega_i (Y + \frac{V}{\gamma}) = \frac{\beta_1 \delta}{\beta_1 - 1} \frac{w_i u_i}{\sum w_i u_i} (Y + \frac{V}{\gamma})$$

which also implies:

$$(A25) \quad Q^{*(i)} = \frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y + \frac{V}{\gamma})}{\theta \sum w_i u_i} = Q(\frac{U^*}{U}) = Q / g \qquad \text{for all } i's .$$

i.e. in equilibrium all threshold points are the same proportion of expected output level and are therefore equalized across firms. As a consequence:

(A26) $\left|K_i^{'}\right| = Ee^{-rt}(\gamma \theta w_i Q^{*(i)}) = p \theta w_i Q^*$, which summed over all $i's$ yields:

$$(A27) \quad \sum \left|K_i^{'}\right| = p \theta Q^* \quad <\Rightarrow \quad \frac{\left|K_i^{'}\right|}{\sum \left|K_i^{'}\right|} = w_i$$

From (A26) , considering a quadratic approximation of the abatement function, it directly follows:

(A27) $u_i^* = \gamma \frac{\theta}{k_{i2}} E e^{-rt} \frac{Q^*}{Q} + \frac{k_{i1}}{k_{i2}}$

Proof of Proposition 7. In equilibrium, the market share of the i-th firm is such that each firm has the same total marginal cost.

Consider again the expression for the present value of profit in (A9) and re-write it by isolating w_i in the fifth term on the RHS as follows:

(A28) $\Pi_i = \frac{P w_i Q}{\delta} - \frac{C_i(Q w_i)}{\delta} - \frac{w_i Q K(u_i)}{\delta} - p q_i - \gamma \frac{\theta w_i^{\beta_1} u_i}{\delta \beta_1} [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{u_i \theta}]^{1-\beta_1} Q^{\beta_1}$

Maximization requires that the derivative of Π_i with respect to w_i be zero:

(A29) $\frac{\partial \Pi_i}{\partial w_i} = \frac{PQ}{\delta} - \frac{C_i' Q}{\delta} - \frac{Q K(u_i)}{\delta} - \gamma \frac{\theta w_i^{\beta_1 - 1} u_i}{\delta} [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{u_i \theta}]^{1-\beta_1} Q^{\beta_1} = 0$

where $C_i' = \frac{\partial C_i}{\partial (Q w_i)}$ is the marginal cost of production for the i-th firm.

The second order condition for a maximum is respected if:

(A30) $\frac{\partial^2 \Pi_i}{\partial w_i^2} = -\frac{C_i'' Q^2}{\delta} - \gamma \frac{(\beta_1 - 1)\theta w^{\beta_1 - 1} u_i}{\delta} [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{u_i \theta}]^{1-\beta_1} Q^{\beta_1} < 0$

Since $\beta_1 \geq 1$, a sufficient condition for (A23) to be true is that $C_i'' = \frac{\partial^2 C_i}{\partial (wiQ)^2} \geq 0$, i.e. that there are no economies of scale internal to the firm.

Finally, to prove the proposition, expression (A29) can be re-written as follows:

$$(A31) \quad P = C_i^{'} + K_i + \gamma u_i \theta [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{(Y_i + q_i + \frac{V_i}{\gamma})}{w_i u_i \theta}]^{1-\beta_1} Q^{\beta_1 - 1}$$

And, substituting the equilibrium value of q_i given by (A22) and simplifying:

$$(A32) \quad P = C_i^{'} + K_i + \gamma u_i \theta [\frac{\beta_1 \delta}{\beta_1 - 1} \frac{((Y + \frac{V}{\gamma}))}{\sum w_i u_i \theta}]^{1-\beta_1} Q^{\beta_1 - 1} = C_i^{'} + K_i + \gamma u_i \theta [\frac{U}{U^*}]^{\beta_1 - 1_1}$$

Recalling that $[\frac{U}{U^*}]^{\beta_1} = Ee^{-rt}$, (A32) can be written:

$$(A33) \quad P = C_i^{'} + K_i + Ee^{-rt} \gamma u_i \theta [\frac{U^*}{U}] = C_i^{'} + K_i + p \frac{u_i}{\sum w_j u_j} \frac{U^*}{Q}$$

Approximating the cost function with a quadratic expansion , using (A32), we also obtain:

$$(A34) \quad P = c_{1i} + c_{2i} w_i Q + K_i + p\omega_i \frac{U^*}{Q}$$

where $c_{1i} = \frac{\partial C}{\partial (w_i Q)}$ for and $c_{2i} = \frac{\partial^2 C}{\partial (w_i Q)^2}$ at an appropriate point. Solving for w_i:

$$(A35) \quad w_i = \frac{1}{c_{2i} Q} [P - c_{1i} - K_i - p\omega_i \frac{U^*}{Q}]$$

Multiplying both sides of equation (A33) by w_i and summing over i:

(A36) $$P\sum w_i = \sum w_i C_i^{'} + \sum w_i K_i + Ee^{-rt}\gamma \sum w_i u_i \theta \frac{U^*}{U}$$

Recalling that $\sum w_i = 1$ and that $U = \frac{1}{N}\sum\ w_i u_i \theta \frac{Q}{\delta}$:

(A37) $$P = \sum w_i C_i^{'} + \sum w_i K_i + Ee^{-rt}\gamma \frac{U^*}{Q} = \sum w_i C_i^{'} + \sum w_i K_i + \frac{(pU^*)}{Q}$$

In the case of the quadratic approximation to the cost function, using (A33), summing over all i's , and applying the condition $\sum w_i = 1$, for all $w_i \geq 0$, yields:

(A38) $$\sum\ w_i = 1 = \frac{1}{Q}[P\sum \frac{1}{c_{2i}} - \sum \frac{c_{1i} + K_i}{c_{2i}} - p\frac{U^*}{U}\sum \frac{\omega_i}{c_{2i}}]$$

Solving for P:

(A38) $$P = \frac{1}{\sum \frac{1}{c_{2i}}}[Q + \sum \frac{c_{1i} + K_i}{c_{2i}} + p\frac{U^*}{U}\sum \frac{\omega_i}{c_{2i}}]$$

Note that if $c_{1i} = c_1$, $c_{2i} = c_2$ and $K_i = K$, $P = [\frac{c_2 Q}{N} + c_1 + K] + p\frac{U^*}{Q}$, that is, under no cost heterogeneity, market price equals average marginal cost plus the value of the allowances and equivalent monitoring costs evaluated at the permit equilibrium prices.

Expression (A39) can be written as:

(A40) $$P = [\frac{Q}{\sum c_{2i}^{-1}} + \sum \frac{(c_{1i} + K_i)c_{2i}^{-1}}{\sum c_{2i}^{-1}} + p(\sum \frac{\omega_i c_{2i}^{-1}}{\sum c_{2i}^{-1}})\frac{U^*}{U} = \frac{Q}{\sum c_{2i}^{-1}} + \sum \psi_i (c_{1i} + K_i) + p(\sum \psi_i \omega_i)\frac{U^*}{U}$$

where $\psi_i = \frac{c_{2i}^{-1}}{\sum c_{2j}^{-1}}$.

Substituting (A37) into (A33), we obtain:

(A41) $$(C_i^{'} - \sum w_j C_j^{'}) + (K_i - \sum w_j K_j) + p\frac{U^*}{U}(u_i - \sum w_j u_j) = 0$$

For the quadratic approximation to the cost function, (A40) becomes:

(A41)

$$w_i = \frac{c_{2i}^{-1}}{\sum c_{2j}^{-1}} + \frac{1}{Qc_{2i}} \sum (\frac{c_{2j}^{-1}(c_{1j}+K_j) - c_{1i}\sum c_{2j}^{-1}}{\sum c_{2j}^{-1}} -) + (\sum \frac{c_{2j}^{-1}\omega_j - u_i \sum c_{2j}^{-1}}{\sum c_{2j}^{-1}}) p\frac{U^*}{U}]$$

(A42) can be written, using the definition in (A41) $\psi_i = \frac{c_{2i}^{-1}}{\sum c_{2i}^{-1}}$:

(A43)
$$\begin{aligned} w_i &= \psi_i + [(\sum \psi_j (c_{1j}+K_j) - (c_{1i}+K_i)] + p\frac{U^*}{U}[\sum \psi_j u_j - u_i] \\ &= \psi_i + (\bar{c_1} - c_{1i}) + (\bar{K} - K_i) + p(\bar{u} - u_i)\frac{U^*}{U} \end{aligned}$$

where $\bar{c_1} = \sum \psi_j c_j$, $\bar{K} = \sum \psi_j K_j$ and $\bar{u} = \sum \psi_j u_j$.

Note that if $c_{1i} = c_1$ and $c_{2i} = c_2$ for all i, (A43) becomes:

(A44) $w_i = \frac{1}{N} - \frac{1}{c_2}[(K_i - \bar{K}) + p(u_i - \bar{u})\frac{U^*}{Q}]$

and if $u_i = u$ for all i, (A43) becomes simply $w_i = \frac{1}{N}$.