ISSN 1835-9728

Environmental Economics Research Hub Research Reports

A Theoretical Model of Optimal Compliance Decisions under Different Penalty Designs in Emissions Trading Markets

Phillia Restiani and Regina Betz

Research Report No. 86

December 2010

About the authors

Phillia Restiani is a PhD student at the School of Economics and Centre for Energy and Environmental Markets at UNSW

Regina Betz is Joint Director (Economics), Centre for Energy and Environmental Markets (CEEM), School of Economics, Australian School of Business, University of NSW

E: r.betz@unsw.edu.au

Environmental Economics Research Hub Research Reports are published by The Crawford School of Economics and Government, Australian National University, Canberra 0200 Australia.

These Reports present work in progress being undertaken by project teams within the Environmental Economics Research Hub (EERH). The EERH is funded by the Department of Environment and Water Heritage and the Arts under the Commonwealth Environment Research Facility.

Crawford School of Economics and Government
THE AUSTRALIAN NATIONAL UNIVERSITY
http://www.crawford.anu.edu.au

TABLE OF CONTENTS

Abstract

This paper employs a theoretical model to examine compliance incentives and market efficiency under three penalty types: the fixed penalty rate, which uses a constant marginal financial penalty; the make-good provision (quantity penalty), where each missing permit in the current period is to be offset with a ratio (restoration rate) in the following period; and a mixed penalty, which combines the two penalty types. Using a simple two-period model of firm's profit maximisation, we analyse compliance decisions and the efficient penalty level under each penalty type. Firms' compliance strategies are modelled as an irreversible investment in abatement measures and permit buying in the market. Our findings indicate that the penalty type does not affect compliance decisions provided that the efficient penalty level is applied. Market efficiency is retained regardless of penalty types. Nevertheless, the mixed penalty design provides the strongest compliance incentives. Hence this finding supports the practice in which this penalty design is widely used in the existing and the proposed trading schemes. Furthermore, we discuss the policy implications of the findings with regard to permit price discovery process and the Australian proposal of tying the penalty level to the permit price.

Keywords: emissions trading, penalty design, compliance

Acknowledgements
We thank, without implicating, AJ Bostian, Robert Marks and Martin Jones for their comments and suggestions.
This research was supported financially through the Environmental Economics Research Hub, which is funded by the Australian Commonwealth Environmental Research Facilities (CERF) program.

1. Introduction

In the past few decades, emissions trading schemes have played an important role as a market-based instrument used to tackle the problem of controlling air pollution. The critical issue of climate change has put the issue of the design for emissions trading markets into the spotlight, as efforts to put a price on carbon have seen a growing number of schemes implemented or developed all over the world. Emissions trading can achieve the targeted emissions reduction efficiently if the scheme is designed properly, and penalty design is a crucial feature needed to maintain the integrity of the environmental goal. A poorly designed scheme might not achieve its efficiency and, even worse, might distort the existing market without meeting its emissions reduction target.

Despite the large body of literature regarding enforcement in the context of emissions trading schemes, very little research discusses the effect of different penalty types as an element of enforcement. This essay looks at how penalty designs, in terms of penalty types and penalty levels, might affect compliance incentives and market efficiency. Three penalty types are considered: the fixed penalty rate; the make-good provision; and the mixed penalty, a combination of both penalty types. We use a simple analytical model at the firm level to assess different compliance incentives related to each penalty type and further analyse the implications of having a different penalty level from the efficient level. Market efficiency in the model is evaluated in terms of the efficient production level, the abatement level, and permit holding under the conditions of firms' compliance and profit maximisation.

The rest of the paper is organised as follows: Section two provides an overview of the studies on enforcement in emissions trading schemes and clarifies the motivation of this essay. Section three describes each penalty design and its application in the existing trading schemes. Section four explains the basic assumptions used in each model, and sections five through seven discuss the model for each penalty type. Section eight discusses the implications of the results, and section nine concludes the findings.

2. Enforcement in Emissions Trading Schemes

The use of tradable pollution permits as a market-based instrument has gained more popularity in recent decades due to its advantages over the command-and-control approach in achieving environmental goals at the least possible cost. It was Coase (1960) who first proposed the idea of transferable property rights as a response to Pigouvian taxes to address externalities such as pollution. These transferable property rights are argued to offer more flexibility by allowing the market to distribute the rights to its highest value users. The concept of tradable permits was first applied in the context of water pollution (Dales, 1968) and air pollution (Crocker, 1966). A general theoretical framework by Montgomery (1972) proves that a tradable permit system can achieve efficiency for a given environmental target or emissions cap.

Although the actual implementation of an emissions trading scheme began in the mid-1970s with the US Environmental Protection Agency Emissions Trading (EPA ET) for stationary sources, a large-scale system was not created until the US Sulfur Dioxide (SO_2) permit trading or Acid Rain Program was initiated in 1995 (Ellerman et al., 2003). The programme, which tackles SO_2 as a local pollutant, is more successful than a standard approach not only in achieving its emissions target (addressing the effectiveness criterion), but also in cutting abatement costs (addressing the efficiency criterion) (Ellerman et al., 2000).

Recently, there have been more emissions trading schemes implemented to address global pollutants, such as greenhouse gases, wherein the concentration of pollutants in a particular area (hot spots) is not a problem. For example, the European Union introduced an emissions trading scheme (EU ETS) in 2005 covering more than 30 countries today, and the Regional Greenhouse Gas Initiatives (RGGI) scheme, which began in 2009, is the first large-scale mandatory cap-and-trade system for greenhouse gases in the US, covering ten states. Australia developed its first trading scheme in 2003 with the implementation of the New South Wales Greenhouse Gas Reduction Scheme (GGAS). Whereas the EU ETS and RGGI are cap-and-trade system, the GGAS requires participants, who are electricity retailers and other individual participants, to meet a benchmark level of emissions reductions by undertaking project-based emissions reduction activities. The scheme is basically a baseline-and-credit system in which a credit is awarded to a facility that reduces emissions beyond the pre-specified emissions baseline or benchmark. These credits must first be certified and can then be used for compliance or traded with another facility (New South Wales Greenhouse Gas Reduction Scheme, 2008). Although the scheme claims to have made significant reductions, from 8.65 ton CO_2/capita to 7.27 ton CO_2/capita (Independent Pricing and Regulatory Tribunal, 2009), it has been severely criticised for a number of design problems, such as the fungibility of its emissions reductions activities, imputed emissions, its methods of calculating the baseline, and its complicated baseline rules, which are believed to result in a price that is much lower than the true scarcity price of carbon (MacGill et al., 2006).

In spite of the potential that emissions trading markets offer, in practice, some issues can have adverse effects on the efficiency of the market. Stavins (1995) points out some examples of these issues, such as market power in the permit market, market power in the product market, non-profit-maximising behaviour, pre-existing regulatory environments, and the degree of monitoring and enforcement. It is argued that the presence of transaction costs will impact efficiency through higher marginal abatement costs for permit buyers (Stavins, 1995) and a reduced number of trading participants (Gangadharan, 2000). Likewise, under the presence of market power, dominant firms might manipulate permit markets to their own advantage, making total pollution control costs more expensive than the efficient level (van Egteren and Weber, 1996, Hahn, 1984).

It is important to recognise that the environmental effectiveness and economic efficiency of a tradable permit system depend, among other things, on the enforcement mechanism used to encourage the compliance of market participants. An enforcement mechanism can include a number of elements: penalty design in terms of level and type, reporting procedures, the verification of reports, monitoring, and sanctioning. Furthermore, each of these elements entails some cost. Three different penalty designs can be distinguished: 1) a fixed financial penalty rate per missing permit, which can be thought of as a ‘price penalty’; 2) a make-good provision requiring firms to surrender missing permits at a given restoration rate or make-good factor, which can be thought of as a ‘quantity penalty’; or 3) a mixed approach combining the price and quantity penalties (henceforth referred to as a mixed penalty). In general, most existing trading schemes have shown compliance rates that are very high compared to those achieved under the regulatory emissions standard approaches. The chosen enforcement mechanism may not only have a direct impact on firms’ compliance decisions (whether a firm chooses to be compliant or non-compliant) but may also indirectly impact permit prices, which might in turn influence the ability of the programme to achieve potential cost savings and related economic benefits (Murphy and Stranlund, 2006). Other factors apart from the penalty design itself that might influence compliance decisions under emissions trading programs are the risk attitudes of market participants, the probability of an audit, flexibility in banking (saving permits for future use) or borrowing (using future permits in the current compliance period), initial allocation rules, trading rules such as auction rules, and any form of market failure, including market power, transaction costs and uncertainties.

The work on compliance decision and enforcement builds from Becker’s (1968) on the economics of crime and punishment. The first theoretical work on enforcement in the area of environmental policy was conducted by Downing and Watson (1974) and focused on standards and effluent fees. Further work on pollution permits was conducted by Malik (1990), who examined market efficiency in the presence of non-compliance and found that compliance decisions will affect the demand for permits and can shift the equilibrium permit price upward or downward, resulting in lower market efficiency.

Following those early works, numerous studies on enforcement models in emissions trading markets have been conducted. Under the presence of market power, the initial allocation of permits to the dominant firm can be used as an enforcement tool in which the regulator can control policy parameters specifically for the price-setting firm rather than adjusting them for all firms (van Egteren and Weber, 1996). However, when marginal enforcement cost is increasing in the initial allocation of permits to the dominant firm, then the initial permits should be distributed such that the dominant firm becomes a net buyer (Chavez and Stranlund, 2003). Keeler (1991) studies compliance decisions under marginal penalty functions with different shapes. Baldursson and von der Fehr (2004) and Stranlund (2008) take into account the influence of risk aversion on compliance. The observed phenomenon of high compliance rates in spite of less frequent inspections or non-severe penalties for discovered violations has been explained using dynamic enforcement models by Greenberg (1984), Harrington (1988), Landsberger and Meilijson (1982), and Stranlund et al. (2005). Theoretical analyses of compliance rules in the context of the Kyoto Protocol and its effects on permit price are assessed by Nentjes and Klaasen (2004) and Godal and Klaasen (2006). Furthermore, Stranlund et al. (2005) study the effect of high penalties on reporting violations (submitting false data) rather than permit violations (failing to hold sufficient allowances) under permit banking provisions.

These existing studies have emphasised the effects of monitoring, different audit probabilities and penalty rates, targeted enforcement, self-reporting, and cheating as important factors in the enforcement of emissions trading schemes. However, we believe that even in cases where we have perfect monitoring and sanctioning mechanisms as well as costless sanctioning, the behaviour of market participants might still be influenced by different penalty types.

This paper aims to use a simple analytical model to assess how different penalty types, specifically the fixed penalty rate, make-good provision, and mixed penalty design, can affect compliance decisions and the efficiency of emissions trading markets. We seek to contribute to the existing literature by focusing on the following aspects:

1) The effects of penalty design on compliance incentives

   To our knowledge, only a few of the existing studies focus on penalty design. Nentjes and Klaasen (2004) look at the compliance incentives associated with the Kyoto Protocol, which include both a make-good provision and a fixed penalty rate. However, they do not undertake a theoretical analysis and do not focus on emissions trading. Moreover, they ground their analysis in the cost of reputation protection for buyers and sellers. Likewise, Godal and Klaasen (2006) use a game theoretical approach under the scenario of market power and US participation, and consider how these may affect committed parties on the road to final compliance under the Kyoto Protocol. Some studies discuss the use of an intertemporal trading ratio to discourage the borrowing of permits that has a similar function to the ratio in the make-good provision (Kling and Rubin, 1997, Stranlund et al., 2005). However, Kling and Rubin (1997) do not focus their model on enforcement, while Stranlund et al. (2005) emphasize the use of tying the penalty to reporting violations. In contrast to those studies, we do not consider reporting violations, but rather focus on the equilibrium in a perfectly competitive permit market under different penalty designs.

2) We abstract from enforcement and monitoring costs or audit probability

   As we want to isolate the effects of the chosen penalty type on compliance decisions, we assume that the violating firms will always be discovered and penalised. Meanwhile, numerous studies, such as those of Stranlund and Dhanda (1999), Sandmo (2002), and Arguedas (2008), consider audit probability as an important variable.

3) Our emphasis is on permit violation with regard to two main compliance strategies

   The compliance strategies are simplified to irreversible investment in emissions reduction measures and permit trading. We model an investment decision as an irreversible decision to highlight that, once the decision is made, it cannot completely be undone, because it has created a positive or zero sunk cost; e.g. the installed equipment cannot be removed simply, and its scrap value is insignificant or zero. Furthermore, we follow Kolstad's (1996) definition of irreversibility, in which today's choices restrict tomorrow's choices. As such, we consider a two-period model in which the investment decision must be made in the first period.

Our theoretical model is built mainly on the work of Malik (1990) and Baldursson and von der Fehr (2004). Malik examines compliance decisions shaped by a marginal penalty rate as a function of violation and permit price, whereas Baldursson and von der Fehr consider how the initial allocation impacts the level of investment in pollution reduction under the assumption of risk aversion. Baldursson and von der Fehr also incorporate the effects of aggregate-level and firm-level risks on the choice of investment level. However, they do not allow for non-compliance in their model. Although they take into account the idea of irreversible investment, they also add the option of undertaking incremental abatement measures that force firms to be compliant. This reduces the irreversibility effect of investment. We combine the models to examine the effects of penalty design on compliance decisions and investment level under the assumption of risk neutrality.

3. Penalty Design

At present, different types of penalty designs have been adopted in emissions trading schemes (see Table 1). In general, three basic types of penalties can be distinguished: 1) a fixed financial penalty rate per missing permit (price penalty); 2) a make-good provision requiring firms to surrender missing permits at a given restoration rate or make-good factor (quantity penalty); and 3) a combination of the two penalty types, which we call a mixed penalty.

A fixed penalty rate, henceforth referred to as an FPR, can provide an incentive and a clear signal of the maximum cost of compliance for firms. The FPR acts as an indication of the maximum compliance costs for firms in choosing their compliance strategy: whether to invest in an abatement technology or to trade in the permit markets. Under some circumstances, a particular level of a fixed penalty rate may also act as a safety valve. The effective safety valve is triggered when the permit price rises above the chosen penalty level. In such a case, the emissions target is not achieved and firms pay the penalty, which is similar to a tax. In this light, we can view the safety valve as a hybrid instrument, a mix between a tradable permit and an emissions tax. This idea is presented by Jacoby and Ellerman (2004) and is similar to a concept proposed by Roberts and Spence (1976).

There are two ways of implementing a safety valve. Firstly, firms can buy additional permits from the government or the market at a fixed price to meet their obligations and remain compliant. The limitation of this approach is that it does not guarantee that the targeted level of emissions reduction will be achieved, because firms can buy as many permits as they want at this trigger price. Secondly, the companies can be temporarily (e.g. for a month) exempted from their obligation to surrender permits. This leeway will similarly compromise any progress made towards reaching the emissions target and undermine the cap-and-trade system. The two approaches have different implications with regards to the compliance status, because the first approach does not automatically ensure that firms will be in compliance once the safety valve is triggered whereas the second approach does. A critical issue with the design of safety valves is the price level, which can be used to maintain the emissions target. When the trigger price is set at a relatively low level, the trigger price becomes an effective price cap (price limit) on the cost of polluting or a binding price ceiling for the permit price. Hence, it indicates the maximum compliance costs. If it is set at a relatively high level, the trigger price acts as a fixed penalty rate that deters firms from polluting. The issue of penalty level is crucial to the FPR design in practice because the regulator will not necessarily have the perfect knowledge with regard to damage costs, firms' marginal abatement costs, or even the current emissions levels, which are all important in setting the theoretical equilibrium permit price and the penalty level based on that permit price.

The second type of penalty design is the make-good provision, henceforth referred to as the MGP, in which firms must compensate for their missing permits in one period at a particular make-good factor or restoration rate in the following period. In the Kyoto Protocol, there is also an additional rule that suspends the non-compliant firms in any particular year, keeping them from selling their permits in the following year (Betz et al., 2006). This MGP ensures that the environmental goal is achieved, because it reduces the allowable aggregate emissions in the following year should it be exceeded in the current year. Assuming that the number of permits allocated by the regulator remains the same every year, non-compliance under this penalty design will create either a future increase in the demand for permits to make up for non-compliance during the current year or a decrease in the number of permits sold in the market because the non-compliant firms are not allowed to sell any permits. Thus, non-compliance in the current year will exert an upward pressure on future permit prices. Furthermore, this type of penalty introduces additional uncertainties, as the compliance costs are uncertain because they are linked to the future permit price, which is unknown. This is not the case for financial penalties, in which the magnitude of the marginal penalty rate is fixed and is publicly announced at the outset.

Most emissions trading schemes employ a mix of the FPR and MGP penalty design. For instance, this is the case with the EU Emissions Trading Scheme, which is currently the largest trading programme (in terms of coverage) with over 12,000 installations. This mix of penalty mechanisms acts as a double penalty for market participants: non-compliance triggers both a fixed penalty and the requirement to surrender the required permits at a future date. On the one hand, this guarantees that the environmental target will be attained, but on the other hand, it may increase the cost of compliance. Some examples of penalty design in existing trading schemes are listed in Table 1.

Table 1 Penalty Designs in Existing Emissions Trading Schemes

Penalty type	Schemes	Pollutants	Sector coverage	Penalty level	Compliance Rate and Permit price
FPR	NSW GGAS	6 GHGs**	Electricity generators 41 participants	A$12.50 incl. taxes	95% compliance rate, 0.01% carried forward shortfalls in 2008 Average $5.85 spot price[a]
	Chile	PM	680 sources emitting $>1000m^3/h$	Penalty fee	Low, then high[b]
	LA RECLAIM	NOx,SOx	292 facilities for NOx and 32 facilities for SOx (2009)	$500/violation/day, determined by court	95% for NOx, 97% for SOx (2009) $809-4780 for NOx, $653-1488 for SOx (2009)[c]
MGP	US NOx Budget Program	NOx	2568 units of power plants and large combustion sources in eastern US	Automatic quota reduction at 1:3	Nearly 100% (2008) $825(Jan) - $592 (Dec)[d]
Mixed penalty	US Acid Rain	SO_2	3456 electricity-generating units	Penalty $2000/ton + MGP 1:1	100% (2008) $509(Jan) - $179 (Dec)[e]
	EU ETS	CO_2	Over 10,000 installations	€100 (2008) + MGP 1:1	98% compliance rate, 3% failed to submit verified emissions[f]
	Australian CPRS*	6 GHGs**	Stationary energy, transport, fugitive emissions, industrial processes, waste and forestry sectors at the start	Predetermined value or max. 110% of benchmark average auction price increased by 5% in real terms annually and MGP 1:1[g]	-
	RGGI (10 participating states) 2009	CO_2	Fossil fuel electricity generators above a size threshold of 25 MW	MGP 1:3 + penalty set by each state 3 year control period	June 2010 reserve price $1.86[h]
	WCI (7 Western US States and 4 Canadian Provinces) *	6 GHGs**	electricity generation, commercial and industrial combustion, and industrial process emissions; gas and diesel for transportation; residential fuel uses	MGP 1:3 + penalty set by each state[i]	-
	UK Carbon Reduction Commitment *	CO_2	Large non-energy-intensive businesses and public sector entities that are not covered by the EU ETS	Safety valve, linked to EU ETS, first set at £40[j]	Allowance price set at £12 in introductory phase
	New Zealand ETS	6 GHGs**	Forestry first, then all sectors by 2013	Penalty NZ$30 + MGP 1:1 and can be raised to NZ$60 + MGP 1:2[k]	-

Note: * schemes are not implemented yet; ** CO2 - Carbon dioxide, CH4 – Methane, N2O - Nitrous oxide, PFCs – Perfluorocarbons, HFCs – Hydrofluorocarbons, SF6 - Sulphur hexafluoride

Source: [a] Independent Pricing and Regulatory Tribunal (2009), [b] Montero et al. (2002), [c] Haimov (2010), [d] US EPA (2009b), [e] US EPA (2009a), [f] EU Directive 2003/87/EC (2010b), Community Independent Transaction Log (2010a), [g] The Parliament of the Commonwealth of Australia (2009), [h] Regional Greenhouse Gas Initiative(2009), [i] Western Climate Initiative (2010), [j] UK Department of Energy and Climate Change (2010), [k] New Zealand Government (2007)

These different penalty designs create different compliance incentives and have different effects on market efficiency. Compliance rates are generally very high, and have reached 100% in the US Acid Rain Program. The data show that higher penalty levels, through either an FPR, an MGP, or a mix of both, will encourage higher compliance rates. Slightly lower compliance rates than under the US Acid Rain Program, as experienced in the Los Angeles Regional Clean Air Incentives Market (LA RECLAIM) and Chile's TSP-Emissions Trading Program, are due more to the monitoring and enforcement problems in those schemes. LA RECLAIM uses a complicated procedure and an ad-hoc court approach in deciding the final compliance status of a violating firm (EPA Clean Air Markets Division, 2006). This approach clearly reduces the influence of the very high marginal penalty rate on compliance and increases administration costs, although the compliance rate is still fairly high. In Chile, permit allocations are made using a proxy-based benchmark approach that has performed poorly with very limited historical emissions information. This problem was exacerbated by poor institutional capacity, which made enforcement more difficult for the programme (Montero et al., 2002). Typically, as markets develop over time and more firms reveal their emissions history, the authority's enforcement capacity is also enhanced.

As shown in Table 1, most of the existing emissions trading programmes use a combination of the FPR and the MGP. Most of the schemes will also publish the non-compliance statuses of the firms in question, providing an additional incentive for compliance due to the risk for firms of losing their reputation. Apart from high compliance rates in those existing schemes, very limited information is available on the actual efficiency of those markets compared to their potential efficiency. In the EU ETS, given the generous allocation and mixed penalty design, the compliance rate is very high. The New South Wales Greenhouse Gas Abatement Scheme also shows a very high compliance rate, with only 0.01% of permit shortfalls being carried forward. It is important to note that, as the New South Wales Greenhouse Gas Abatement Scheme is a baseline-and-credit system rather than a cap-and-trade system (the latter being the focus of our study), it may offer different compliance incentives as participating firms act as suppliers of credit and the credit supply is not fixed.

Drawing from the available empirical data, it is difficult to determine how market efficiency is impacted by penalty design. Because information on firms' actual emissions levels and marginal abatement costs are usually not known in practice, it is difficult to determine whether existing permit markets have achieved their full potential efficiency gains. Our theoretical analysis of an emissions trading scheme uses a stylised model that allows us to focus on a few choice variables related to compliance decisions under the assumed market setting. The simplicity of the model also provides insights that are generally applicable, provided the assumptions are met. The comparative statics of the model also provide us with straightforward implication of the effect of a particular variable on the variable of interest. Therefore the analysis of compliance decisions and market efficiency under different penalty design merits the use of a theoretical model.

4. Model Assumptions

Consider an emissions trading scheme that consists of *n* firms that are price takers in both the permit market and a downstream market, which is independent of the permit market. For a given quantity of outputs, the production activity of firm *i* generates emissions e_i and revenues in which the price of the commodity, τ, is exogenous.

Firms are required to have a permit for each unit of pollution that they produce, and these permits can be obtained through endowments (as in the case of free initial allocation or grandfathering) or purchase (as in the case of auctioning). There is a central authority that conducts spot checks of reported data to prevent cheating and enforcement, ensuring that firms that produce more emissions than is allowed by the number of permits that they hold are penalised.

The central authority initially allocates free permits to firms or auctions those permits. Without loss of generality, we take the free initial allocation of permits (grandfathering) as a basic model, although the model also applies when permits are auctioned by holding the amount of free permits at zero. We define an initial stock of permits at time *t* in the market, $\overline{S}_t$, which is fixed over time and is the sum of gratis permits given to each firm, $\overline{S_t} = \sum s_{it}$. The aggregate emissions level in the business-as-usual (BAU) scenario, E_t, is the sum of firms' BAU emissions levels $E_t = \sum e_{it}$. Since the authority seeks to reduce the aggregate emissions level, the emissions cap or the permit supply is kept lower than the BAU emissions level, $\overline{S}_t < E_t$. The total number of permits in the market at the end of a compliance period *t* is denoted by $\overline{L_t}$ and should be the same as the initial amount of stock because we are considering a closed permit market that does not allow for linking with other permit systems.

$$\overline{L_t} = \sum l_{it} = \overline{S_t} = \sum s_{it} < E_t \; ; \; i \in \{1,2,\ldots,n\} \qquad \textbf{(1)}$$

where l_{it} is the number of permit holdings of firm *i* at the end of compliance period *t*.

The compliance strategies consist of:

1) Investing in an abatement technology and/or
2) Trading permits.

To inspect the effect of irreversible investment on enforcement in a permit market, we use a stylised two-period model. We model investment as an irreversible decision that will commit firms to undertaking abatement measures in the first period. Rather than including a lump sum investment cost in the model, the investment decision is indicated by a positive abatement level a_i that will operationalise the same marginal abatement costs over time $c_{it}(a_{it}) = c_i(a_i)$. As investment decisions are irreversible and cannot be undone (and the associated costs cannot be recovered), we require that $a_i \geq 0$.

Firms are assumed to be price takers not only in output markets but also for permit markets and they differ in their marginal abatement costs, $c_i(a_i)$, which are continuous, increasing, and convex. $c_i(0) = 0$, $c_a > 0$ and $c_{aa} > 0$. A firm's output level is expressed as a function of capital, k_{it}, which is a production input that will have the main influence on emissions levels.

$$q_{it} = f(k_{it}),\ \partial q_{it}/\partial k_{it} > 0,\ \partial^2 q_{ti}/\partial k_{it}^2 < 0 \qquad \textbf{(2)}$$

Likewise the firm's initial emissions levels are a function of their output level and technology parameter θ_{it}.

$$e_{it} = h\left(q_{ti}(k_{ti}), \theta_{it}\right) \qquad \textbf{(3)}$$

We assume that an increasing production level will increase the firm's emissions levels at a decreasing rate and that the same applies for the effect of better technology in decreasing emissions levels.

$$\partial e_{it}/\partial q_{it} > 0,\quad \partial^2 e_{it}/\partial q_{it}^2 < 0,\quad \partial e_{it}/\partial \theta_{it} < 0,\quad \partial^2 e_{it}/\partial \theta_{it}^2 < 0 \qquad \textbf{(4)}$$

The regulator chooses the type of penalty in the form of a fixed penalty rate, denoted by f, or a restoration rate, ρ.[1]

The decision-making process in each period is as follows:

---

[1] Many emissions trading schemes allow for banking and/or borrowing to provide more flexibility for firms dealing with price fluctuations due to external shocks. However, to keep the model simple and tractable, we do not allow for banking or borrowing in this paper or for trading in futures markets.

1) Information stage

    During this stage, firms receive both public or global information and private information pertaining to the permit market. When the information is public, a regulator announces the penalty type, its level and the initial allocation mechanism. When grandfathering occurs, the number of free initial permits allocated to each firm is fixed during each period and denoted by $\overline{s_i} \geq 0$. When an FPR is used, the authority announces the penalty rate f, whereas the restoration rate, ρ, is revealed if a make-good provision is used. In the mixed penalty design, both f and ρ are declared. The penalty design variables are exogenous to firms. Firms also acknowledge that the probability of their being caught in a violation and the probability of their being sanctioned are equal to one (under perfect monitoring and penalty enforcement). Furthermore, penalty enforcement is assumed to be costless because this is not our focus. This information is also available to all firms. Thus, if their emissions level e_i exceeds the number of permits held l_i, the penalty will be automatically enforced.

    Apart from public information, firms also receive private information regarding marginal abatement cost $c_i(a_i)$, product price τ and capital rent r. Production level q_i is used to generate the firm's total revenue; at the same time, production activity creates pollution and determines the firm's initial emissions level, e_i, before an abatement measure is adopted.

2) Firms decide whether to invest in an abatement measure.

3) Firms trade permits in the market.

    When grandfathering occurs, firms learn about the permit price of the secondary market only after trading permits has taken place.. When permits are auctioned off, the auction price provides an early signal of the expected permit price. In this model, permit price, p, is an exogenous variable because firms are price takers.

4) Compliance checks are carried out by the regulator, and penalties are imposed on non-compliant firms.

These stages are exactly the same for the two-period model that we have, with the exception that investment decision can only be made in the first period. The theoretical framework also builds on the assumption that the law of one price applies so that the auction price (in the primary market) is equal to the permit price in the spot markets (secondary markets). In the following sections, we will analyse the efficient market conditions that correspond to the three different forms of penalty design. It is important to note that the term “optimal” always refers to efficiency, which requires the attainment of potential cost savings in the market. We will use the terms “optimal” and “efficient” interchangeably in our discussion.

5. Fixed Penalty Rate

With a fixed penalty rate the two-period model can be simplified into a static model, because there are no differences in market structure across the two periods and non-compliance will have the same effect by the end of each period, unlike with the MGP. Without loss of generality, we remove the time subscript *t* from our variables.

Let e_i be firm *i*’s initial emissions, $\bar{s}_i$ be firm *i*’s initial permits, a_i be firm *i*’s abatement level, l_i be firm *i’s* number of permit holding, and $c_i(a_i)$ be firm *i*’s abatement investment costs. The firm’s violation level is denoted by

$$v_i = e_i(q_i(k_i)) - a_i - l_i \geq 0 \qquad \textbf{(5)}$$

The firm’s total costs are expressed as

$$C_{(a_i, l_i \mid \bar{f}, \bar{s}\ \ , \bar{p})} = c_i(a_i) + pd_i + fv_i \qquad \textbf{(6)}$$

where $d_i = l_i - \bar{s}_i$. d_i specifies the number of permits traded by firm *i* in the market.

When $d_i > 0$, the firm is a net buyer, which means that by the end of a compliance period that firm buys more permits than it sells. Accordingly, if $d_i < 0$, the firm in question is a net seller. The firm’s profit function prior to an investment decision and trading in permits is expressed as

$$B_i = \tau\, q_i(k_i) - r\, k_i \qquad \textbf{(7)}$$

where τ and r symbolise the price of the good and the capital rent. These parameters are exogenous to the firm and hence are not the firm’s choice variables.

As we are interested in compliance decisions, the firm is allowed to choose a non-negative violation $v_i \geq 0$. Firm *i*’s profit maximisation function is

$$\underset{a_i, l_i, K_i}{Max} \quad \Pi_i = \tau\, q_i(k_i) - r\, k_i - c_i(a_i) - p[l_i - \bar{s}_i] - f\, v_i \qquad \textbf{(8)}$$

$$\text{subject to} \quad v_i \geq 0, e_i \geq 0,\ a_i \geq 0, l_i \geq 0, k_i \geq 0$$

The profit maximisation problem yields a Lagrangian equation:

$$L = \tau\, q_i(k_i) - r\, k_i - c_i(a_i) - p[l_i - \bar{s}_i] - f\, v_i + \lambda\, v_i \qquad \textbf{(9)}$$

We derive the Kuhn-Tucker conditions as follows:

$$\frac{\partial L}{\partial k_i} = \tau \frac{\partial q}{\partial k} - r - [f - \lambda]\left(\frac{\partial e}{\partial q}\frac{\partial q}{\partial k}\right) \le 0 \qquad \textbf{(10)}$$

$$\frac{\partial L}{\partial a_i} = -c_a + f - \lambda \le 0 \qquad \textbf{(11)}$$

$$\frac{\partial L}{\partial l_i} = -p + f - \lambda \le 0 \qquad \textbf{(12)}$$

$$\frac{\partial L}{\partial \lambda} = v_i = e_i(q_i(k_i)) - a_i - l_i \ge 0 \qquad \textbf{(13)}$$

$$k_i \ge 0,\ k_i \frac{\partial L}{\partial k_i} = 0; \qquad a_i \ge 0,\ a_i \frac{\partial L}{\partial a_i} = 0;$$

$$l_i \ge 0,\ l_i \frac{\partial L}{\partial l_i} = 0; \qquad \lambda \ge 0,\ \lambda \frac{\partial L}{\partial \lambda} = 0 \qquad \textbf{(14)}$$

Assuming an internal solution, we need to hold (11) and (12) equal to zero to obtain

$$\lambda = f - c_{ai} = f - p \qquad \textbf{(15)}$$

A rational firm will choose its optimal investment in abatement so that marginal abatement costs are equalised across firms and will be the same as the equilibrium permit price *p**. This is a common finding and highlights the advantage of a permit market.

$$c_{ai} = p^* \qquad \textbf{(16)}$$

An optimal compliance decision for the firm is obtained when (12) holds as equality when (13) equals zero. The Kuhn-Tucker conditions in (14) require that $\lambda \ge 0$ when $\frac{\partial L}{\partial \lambda} = 0$, which implies that compliance is efficient as long as the fixed penalty rate is equal to or higher than the equilibrium price.

$$f \ge p \qquad \textbf{(17)}$$

This is the usual condition for achieving perfect compliance. Using this logic, tying the marginal penalty rate to the equilibrium price will ensure that the penalty rate is above but very close to the permit price and hence ensure perfect compliance in the market and market efficiency as a result. However, this may in fact introduce additional uncertainty into the firm's decision-making process, because the marginal penalty rate will always change following the equilibrium permit price. In the efficient equilibrium, the firm's abatement level is equal to the difference between its emissions levels and number of permits that it holds. When the penalty level is set lower than the permit price, firms will chose to be non-compliant, and the aggregate emissions level will be higher than the number of permits in the market. A high emissions level will also drive down the permit price, as permit demand declines, until it reaches the low penalty level. Hence, the emissions target will not be achieved. However, when the penalty level is set higher than the permit price, firms will be profit-maximizing by being compliant.

At the equilibrium, there exists an efficient aggregate investment level for all firms such that

$$c_a\left(a_i\left(p^*\right)\right) = c_a^*\left(a_i^*\right) = p^* \tag{18}$$

$$v_i = 0 \Rightarrow e_i = a_i\left(p^*\right) + l_i\left(p^*\right)$$

$$\therefore a_i\left(p^*\right) = e_i - l_i\left(p^*\right) \tag{19}$$

When incorrect price signals are transmitted through investment, or firms have incorrect knowledge about the equilibrium permit price, firms may over- or under-invest. Consequently, this will affect the demand for permits, which in turn will be reflected in the permit price. Hence, it will change the compliance incentives associated with a given penalty level in later periods.

Proposition 1 *Firms will find it optimal to comply as long as the marginal penalty rate is set greater than or equal to the equilibrium permit price. Q.E.D.*

Corollary 1 *As the penalty rate increases, the amount of abatement will also increase or the output level will decrease accordingly until the efficient level is achieved. At the equilibrium, the firm's marginal net benefit after taking into account the cost of compliance is equal to the firm's marginal cost of production.*

Proof: When λ from (15) is substituted in to (10) and we hold (10) to zero, we obtain

$$\tau \frac{\partial q_i}{\partial k_i} = r + p \frac{de_i}{dq_i} \frac{\partial q_i}{\partial k_i} \tag{20}$$

Under a perfectly competitive permit market, we achieve the usual profit-maximising condition in which the firm's marginal benefit from undertaking a production activity, as expressed by the marginal revenue on the left-hand side, is equal to the sum of its marginal costs of production (capital rent) and its marginal compliance costs in the permit market on the right-hand side. As seen in the equation, this marginal cost is increasing in production level as the emissions level increases accordingly.

We now use a simulation of the comparative statics to graph these effects, using the following functions and parameters: $c = a^2,\ e = q^{0.9},\ \ q = k^{0.5},\ \tau = 0.05,\ r = 0.15,\ p = 0.02$. We can see that, as the penalty rate f increases, the violation level v decreases and then drops to zero when $f \geq p$, the permit price (0.02), holding other variables constant (Figure 1a). This model confirms that firms will find it optimal to comply when the cost of being compliant is lower than the benefit of being non-compliant.

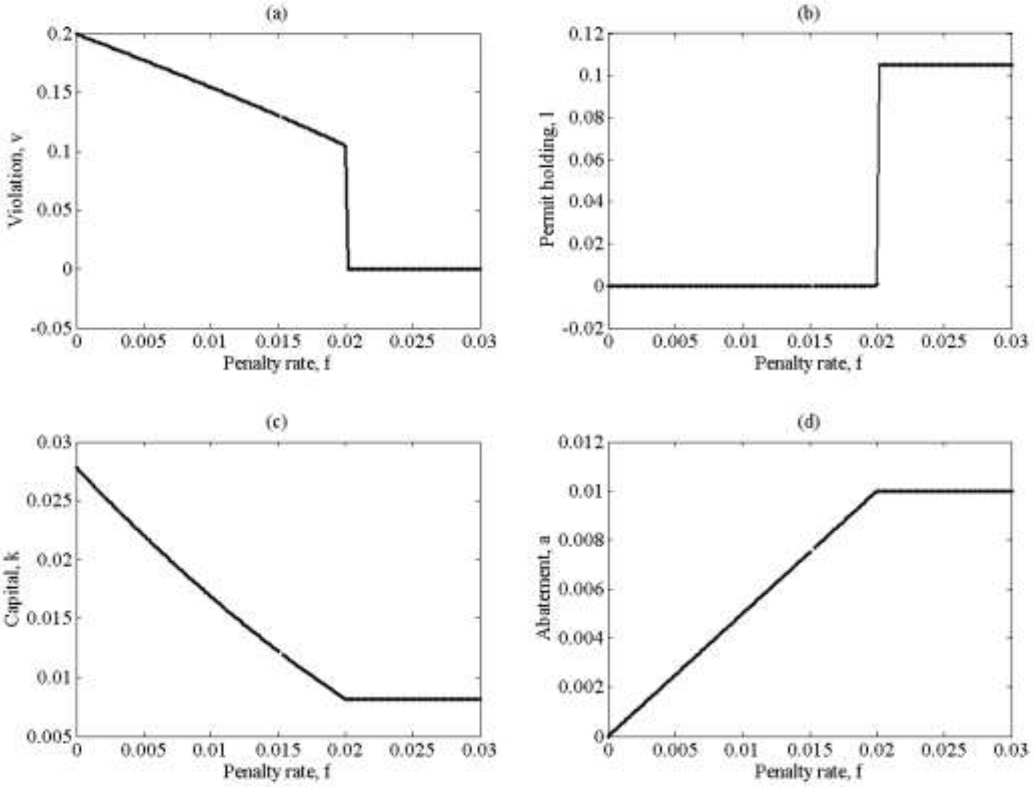


Figure 1 The Effects of Penalty Rates on Violation Levels, Compliance Strategies, and Production Levels under the FPR

Since the model simplifies compliance strategies to making investment decisions in abatement measures and/or holding permits, it can explain very well the effects of the level of the penalty rate on both of these options (Figure 1b and 1d). When the penalty rate is less than the equilibrium permit price, firms will not need to hold a single permit because it will be cheaper for them to either violate when their abatement cost is more expensive or to comply by investing in abatement measure when the abatement cost is cheaper than the permit price.

Considering that the model focuses on compliance decision and penalty design, the long-run incentive of making an early investment is not captured in the model. Rather, investment decision is merely a compliance strategy. Thus, any compliance strategies in this model, either abatement investment or permit holding, are expressed mainly as costs (permit selling is carried out more as a strategy to minimize compliance cost). In this sense, increasing the penalty rate renders higher compliance costs, as firms need to hold permits in accordance with their output level. Consequently, firms reduce their output levels with increasing penalty rates up to the point where the penalty rate is equal to the equilibrium permit price, then firms achieve their efficient output level as expressed by the amount of capital use (Figure 1c).

The permit market under FPR design will achieve its efficiency when firms choose their best compliance strategies at the equilibrium permit price. When a firm decides to make an investment and also buy permits, it should choose an efficient mix of investment in abatement level and number of permit holding. However, if the firm should choose either one of the available compliance strategies based on the information on the equilibrium permit price, then its best compliance strategy should either be investing in an abatement measure or buying permits. In a permit market with a free allocation of permits (grandfathering), the choice of the firm's best compliance strategy will divide firms into two groups: net buyers and net sellers.

Proposition 2 *In a permit market with a fixed penalty rate in which the stock of permits is less than the aggregate emissions under business as usual, and assuming a free allocation of permits, the firm's best compliance strategy is to be a net seller when its marginal abatement cost is lower than the equilibrium permit price, and to be a net buyer when its marginal abatement cost is higher than the equilibrium permit price.*

Proof: Let E and Q denote the aggregate amount of emissions and the total production level, respectively, under the business-as-usual scenario.

Total emissions per period:

$$E = \sum e_i(q_i) \qquad \textbf{(21)}$$

Total output per period:

$$Q = \sum q_i \qquad \textbf{(22)}$$

When the total permit supply, which is equal to the total number of permit holding at the end of a compliance period, is lower than the aggregate emissions level under the business as usual scenario, some degree of abatement is required in the market.

$$S = L < E \qquad \textbf{(23)}$$

Then some abatement level is required

$$A = E - L \qquad \textbf{(24)}$$

where

$$A = \sum a_i \qquad \textbf{(25)}$$

This holds both globally and for individual firms, as shown in equation (19).

Given that $v_i = 0$ and $\overline{f} < p^*$, let $i = \{1,..., j, j+1,..., n\}$ and

$c_{a1} < c_{a2} < ... < c_{aj} < c_{aj+1} < ... < c_{an}$:

a) Firms $i \in \{1,..., j\}$ are net sellers of permits, with $c_{ai} \leq p^*$.

   Then, making an investment decision in abatement measures and selling the freely allocated permits is the best compliance strategy for firm *i*, $\forall i \in \{1,..., j\}$.

   For net sellers, denoted by *ns*, total emissions are:

$$E_i^{ns} = \sum_{i=1}^{j} e_i = \sum_{i=1}^{j} a_i^{ns}\left(p^*\right) \qquad \textbf{(26)}$$

b) Firm $i \in \{j+1,..., n\}$ are net buyers of permits, with $c_{ai} > p^*$.

   Then, buying permits in the secondary market is the best compliance strategy for firm *i*, $\forall i \in \{j+1,..., n\}$ and firm *i* are net buyers in the permit market.

   For net buyers (*nb*), total emissions are as follows:

$$E_i^{nb} = \sum_{i=k}^{n} e_i(p^*) = \sum_{i=k}^{n} l_i^{nb}\left(p^*\right) \qquad \textbf{(27)}$$

The efficient aggregate emissions level in the market is

$$E(p^*) = E^{ns} + E^{nb} = A + L = \sum_{i=1}^{j} a_i^{ns} + \sum_{i=j+1}^{n} l_i^{nb} \qquad \textbf{(28)}$$

where

$$a_i = 0 \ \forall \ i > j \quad \text{and} \quad l_i = 0 \ \forall \ i < j+1 \qquad Q.E.D.$$

Corollary 2 *Market efficiency is achieved when all firms choose their profit-maximising compliance strategy at the equilibrium permit price.*

Proof: Let the optimum aggregate profit in the market at the equilibrium permit price be

$$\Pi_{market}^* = \sum \Pi_i^*\left(a_i, l_i, k_i\right) = \Pi_1^* + \Pi_2^* + ... + \Pi_n^* \qquad \textbf{(29)}$$

Suppose we have one firm *i* =1 which does not choose its best compliance strategy such that $\Pi_1 < \Pi_1^*$; hence, $\Pi_{market} = \Pi_1^* + \Pi_2^* + ... + \Pi_n^* < \Pi_{market}^*$ *Q.E.D.*

6. Make-Good Provision

Under the make-good provision penalty design (MGP), a restoration rate ρ determines the ratio at which a firm should have to compensate for its missing permits. For instance, if a firm has 3 missed permits and $\rho = 2$, then in the next period a firm should hold 6 more permits (additional to what it needs to surrender for that period). Hence this penalty design allows a borrowing provision to the trading scheme. When $\rho = 1$, the make-good provision allows for perfect borrowing from one period to another. However $\rho > 1$ implies that there is an additional cost of borrowing. It can be said that the borrowing cost increases with a higher restoration rate. In practice, the presence of a discount rate can encourage firms to shift emissions today to the future in order to push the costs further in the future (Kling and Rubin, 1997). Nevertheless, the absolute cost of this borrowing provision in fact also depends on the permit price in the following period. As a result, to analyse the effects of a MGP we need to develop a dynamic model.

Under a MGP, firms will not be penalized with a penalty fee when they have missed permits at a particular period *t.* For the sake of simplicity, let *t* = 1, 2, where period one is the first year and period two is the last year of a phase of an emissions trading scheme. Compliance is ensured only through a restoration rate, which influences the initial allocation in period two. Thus the model is set such that the violation in the second period should be equal to zero. In practice, the regulator normally establishes a massive fine and/or serious legal consequences for violation at the end of a trading stage to ensure the firm’s compliance in the market. In some countries, criminal charges or even incarceration are imposed in order to deter non-compliance in the market. Hence, firms need to keep their total violations equal to zero. However, we do not take into account this legal implication as an additional compliance cost, rather we guarantee a condition of perfect compliance by setting the second period violation at zero.

It is assumed that the total number of permits will be equal to the total number of initial permits, which is constant in both periods.

$$S_1 = \sum s_{i1} = S_2 = \sum s_{i2} \qquad \textbf{(30)}$$

Furthermore, the total number of permits is lower than the total initial emissions of all firms to create an incentive to invest in abatement technology.

$$\sum_{t=1}^{2} \sum_{i=1}^{j} e_{it} = \sum E_t > \sum S_t \qquad \textbf{(31)}$$

The decision-making stages in this model are similar to those in the FPR model. However, the following differences exist with regard to the investment decision:

a) Firms can only make investment decisions in the first period as a compliance strategy. This condition reflects the irreversible nature of investment. If firms choose to invest, the reduction in emissions will take place immediately, and the same abatement costs will also be incurred in the second period.

$$\sum c_{it}(a_{it}) = 2c_i(a_i) \tag{32}$$

If firms do not invest in the first period, they can only achieve compliance through permit trading.

b) After the investment decision stage, firms choose their compliance in period one by choosing their permit holdings in period one, denoted by l_{i1}.

c) In period two, firms choose their compliance strategy by determining their permit holdings in period two, denoted by l_{i2}.

Let firm *i*'s violation level in the first period be

$$\boldsymbol{v}_{i1} = \boldsymbol{e}_{i1}(\boldsymbol{q}_{i1}(\boldsymbol{k}_{i1})) - \boldsymbol{a}_{i1} - \boldsymbol{l}_{i1} \geq 0 \tag{33}$$

The violation level in the second period is denoted by

$$\boldsymbol{v}_{i2} = \boldsymbol{e}_{i2}(\boldsymbol{q}_{i2}(\boldsymbol{k}_{i2})) - \boldsymbol{a}_{i2} - \boldsymbol{l}_{i2} = 0 \tag{34}$$

Thus, firm *i*'s total violation level for both periods is

$$\sum_{t=1}^{2} v_{it} = \sum (e_{it} - a_{it}) - l_{i1} - l_{i2} \tag{35}$$

When the firm violates in the first period, its initial permit allocation in the second period is reduced proportionately by a factor of ρ.

$$v_{i1} > 0 \rightarrow s_{i2} = (s_{i1} - \rho\, v_{i1}) < s_{i1} \tag{36}$$

Firm *i*'s maximisation problem is

$$\underset{k_{i1},k_{i2},a_i,l_{i1},l_{i2}}{Max} \quad \Pi_i = \sum \tau\, q_{it}(k_{it}) - \sum r\, k_{it} - \sum c_{it}(a_{it}) \\ - p_1[l_{i1} - \bar{s}_{i1}] - p_2[l_{i2} - \bar{s}_{i1} + \rho\, v_{i1}] \tag{37}$$

subject to $\boldsymbol{v}_{i1} \geq 0, \;\; \boldsymbol{v}_{i2} = 0$

The Lagrangian equation for the profit maximisation problem is given by

$$L = \sum \tau \, q_{it}\left(k_{it}\right) - \sum r \, k_{it} - \sum c_i\left(a_{it}\right) - p_1\left[l_{i1} - \bar{s}_{i1}\right] \\ - p_2\left[l_{i2} - \bar{s}_{i1} + \rho \, v_{i1}\right] + \lambda_1 \, v_{i1} + \lambda_2 \, v_{i2} \tag{38}$$

The Kuhn-Tucker conditions are

$$\frac{\partial L}{\partial k_1} = \tau \frac{\partial q_{i1}}{\partial k_{i1}} - r - \left[p_2 \rho - \lambda_1\right]\left(\frac{de_{i1}}{dq_{i1}} \frac{\partial q_{i1}}{\partial k_{i1}}\right) \le 0 \,,$$

$$\frac{\partial L}{\partial k_2} = \tau \frac{\partial q_{i2}}{\partial k_{i2}} - r + \lambda_2\left(\frac{de_{i2}}{dq_{i2}} \frac{\partial q_{i2}}{\partial k_{i2}}\right) \le 0 \tag{39}$$

$$\frac{\partial L}{\partial a_i} = -2c_{ai} + p_2 \rho - \lambda_1 - \lambda_2 \le 0 \tag{40}$$

$$\frac{\partial L}{\partial l_{i1}} = -p_1 + p_2 \rho - \lambda_1 \le 0 \tag{41}$$

$$\frac{\partial L}{\partial l_{i2}} = -p_2 - \lambda_2 \le 0 \tag{42}$$

$$\frac{\partial L}{\partial \lambda_1} = e_{i1}\left(q_{i1}\left(k_{i1}\right)\right) - a_i - l_{i1} \ge 0 \,,$$

$$\frac{\partial L}{\partial \lambda_2} = e_{i2}\left(q_{i2}\left(k_{i2}\right)\right) - a_i - l_{i2} = 0 \tag{43}$$

$$k_{i1} \ge 0 \,,\ k_{i1} \frac{\partial L}{\partial k_{i1}} = 0 \,; \qquad k_{i2} \ge 0 \,,\ k_{i2} \frac{\partial L}{\partial k_{i2}} = 0 \,;$$

$$a_i \ge 0 \,,\ a_i \frac{\partial L}{\partial a_i} = 0 \,; \quad l_{i1} \ge 0 \,,\ l_{i1} \frac{\partial L}{\partial l_{i1}} = 0 \,; \quad l_{i2} \ge 0 \,,\ l_{i2} \frac{\partial L}{\partial l_{i2}} = 0 \,;$$

$$\lambda_1 \ge 0 \,,\ \lambda_1 \frac{\partial L}{\partial \lambda_1} = 0 \,; \quad \lambda_2 \ge 0 \,,\ \lambda_2 \frac{\partial L}{\partial \lambda_2} = 0 \tag{44}$$

Proposition 3 *The firm chooses its efficient level of investment in abatement measures by equalising its marginal abatement cost to the permit prices in both periods. This follows the result under a fixed penalty rate design.*

Proof: Assuming an interior solution, from (41) and (42), we obtain $\lambda_1 = p_2 \rho - p_1$ and $\lambda_2 = -p_2$. These two equations are substituted back into equation (40) to obtain

$$-2c_{ai} + p_1 + p_2 = 0 \tag{45}$$

Likewise, equations (41) and (42) are substituted into equation (39) to obtain

$$\tau \frac{\partial q_{i1}}{\partial k_{i1}} = r + p_1 \frac{de_{i1}}{dq_{i1}} \frac{\partial q_{i1}}{\partial k_{i1}}$$

$$\tau \frac{\partial q_{i2}}{\partial k_{i2}} = r + p_2 \frac{de_{i2}}{dq_{i2}} \frac{\partial q_{i2}}{\partial k_{i2}} \qquad \textbf{(46)} \; \textit{Q.E.D.}$$

This shows that the profit-maximising firm will increase its production level until the marginal revenue of production is equal to the sum of the marginal production cost and the marginal compliance cost under a permit market. In this sense, the degree of emissions increase that corresponds to a production increase is the key to the equation. Although we consider decreasing emissions levels as a result of abatement, we have not discussed the effect of the firm's technology θ_i on its emissions level: firms with cleaner technology have an advantage at a given level of capital.

Proposition 4 *Firms will find it optimal to comply in period 1 as long as the restoration rate is set at* $\rho \geq \frac{p_1}{p_2}$ *. This proposition holds when we consider a zero discount rate, which otherwise would have made different implications.*

Corollary 3 *When* $p_1 = p_2$*, increasing the restoration rate in the make-good provision will lower the firm's total violation level in period 1 as the cost of borrowing increases up until the restoration rate equals one, beyond which the firm will find it optimal to have a zero total violation level. The penalising effect of the borrowing cost will be higher when* $p_1 < p_2$*. Accordingly, when* $p_1 > p_2$*, the optimal restoration rate should be even higher than when* $p_1 = p_2$*.*

Proof: Based on equations (41) and (43), we derive

$$\frac{\partial L}{\partial l_{i1}} = p_1 - p_2\rho + \lambda_1 \geq 0$$

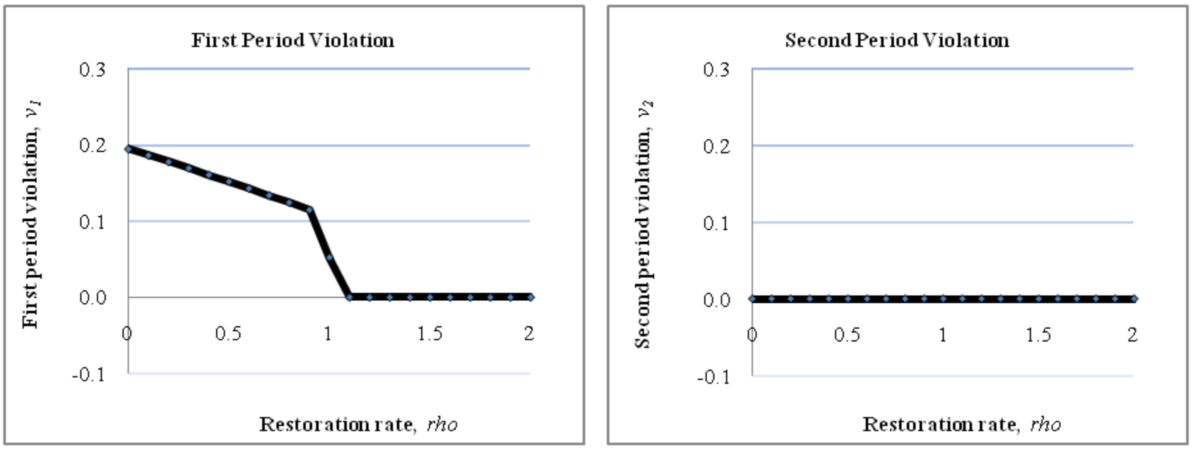
$$\frac{\partial L}{\partial l_{i1}} = \frac{\partial L}{\partial \lambda_{i1}} \geq 0 \Rightarrow e_{i1}\left(q_{i1}\left(k_{i1}\right)\right) - a_i - l_{i1} = v_{i1} = p_1 - p_2\rho + \lambda_1 \qquad \textbf{(47)}$$

Taking the first derivation of equation (47) with respect to ρ, we obtain the marginal effect of the restoration rate on first-period violations. Since $p_2 \geq 0$ and $v_{i2} = 0$, increasing the restoration rate will decrease the firm's total violation rate in the first period.

$$\frac{\partial v_{i1}}{\partial \rho} = -p_2 < 0 \tag{48}$$

Optimal compliance in the first period is achieved by setting equation (47) equal to zero. Since Kuhn Tucker's first order condition requires $\lambda_1 \geq 0$, we find that $\rho \geq \frac{p_1}{p_2}$. When we have non-zero discount rates, even higher levels of restoration rates are required, as the discount rate will reduce the value of the second period permit price. *Q.E.D.*

For illustrative purposes, we conduct a simulation that involves keeping the permit prices the same in both periods, $p_1 = p_2$. We use a comparative static analysis with the same parameters and functions as in the FPR design: $c = a^2,\ e = q^{0.9},\ q = k^{0.5},\ \tau = 0.05,\ r = 0.15,$ $p_1 = p_2 = 0.02$. The results (Figure 2, below) indicate that, when the restoration rate is zero – which means that firms are not penalised for their missed permits in period one –, the violation level reaches the maximum level; this is correlated with the maximum production level. As the restoration rate increases, the violation level decreases and then drops to zero after the restoration rate equals one.



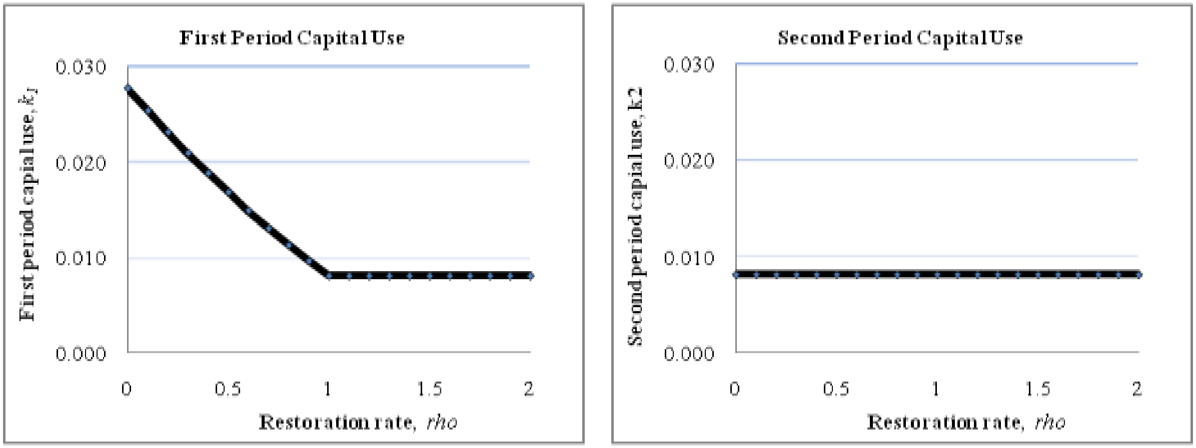
(a) The effect of restoration rates on the first period violation levels

(b) The effect of restoration rates on the second period violation levels

Figure 2 The Effect of Restoration Rates on Violation Levels under the MGP

This result confirms that, when restoration rate equals one –implying perfect borrowing across the two periods under zero interest rate –, then firms are indifferent between violating in the first or second period. On the other hand, when the restoration rate is greater than one, the firm finds it more expensive to violate in period one because there is a higher cost of borrowing, and thus the firm keeps its first period violation rate to zero. As for the second period, the restoration rate has no effects in this model and the violation levels always equal to zero as the setting of the model.

As shown in Figure 3, the restoration rate will affect the amount of capital use in the first period, but not in the second period. Since the model requires perfect compliance in the second period, the capital use in the second period is the same as the optimal capital use for the static model. In the first period, firms still have both options of compliance strategies and permit buying decision is influenced by the restoration rate. When the restoration rate is less than one, firm incurs cheaper compliance costs by shifting its permit buying to the second period. Hence, increases in the restoration rate are matched by decreases in the first period capital use because the restoration rate represents increasing compliance costs. This pattern continues until the restoration rate equals one, after which firms find it optimal to have a zero first period violation by holding an optimal amount of capital under an emissions trading scheme.

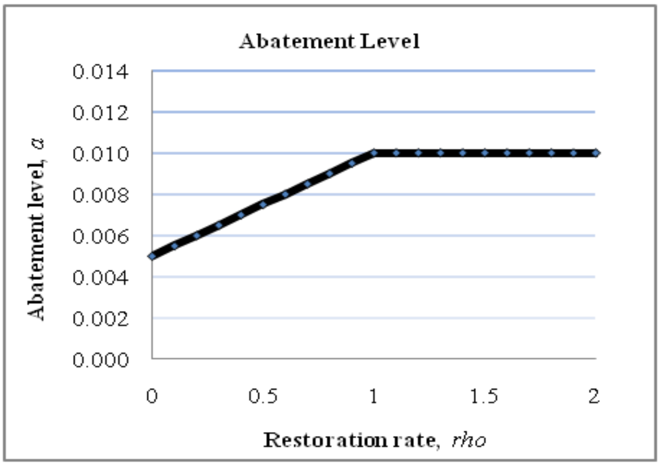


(a) The effect of restoration rates on the first period capital use (production levels)

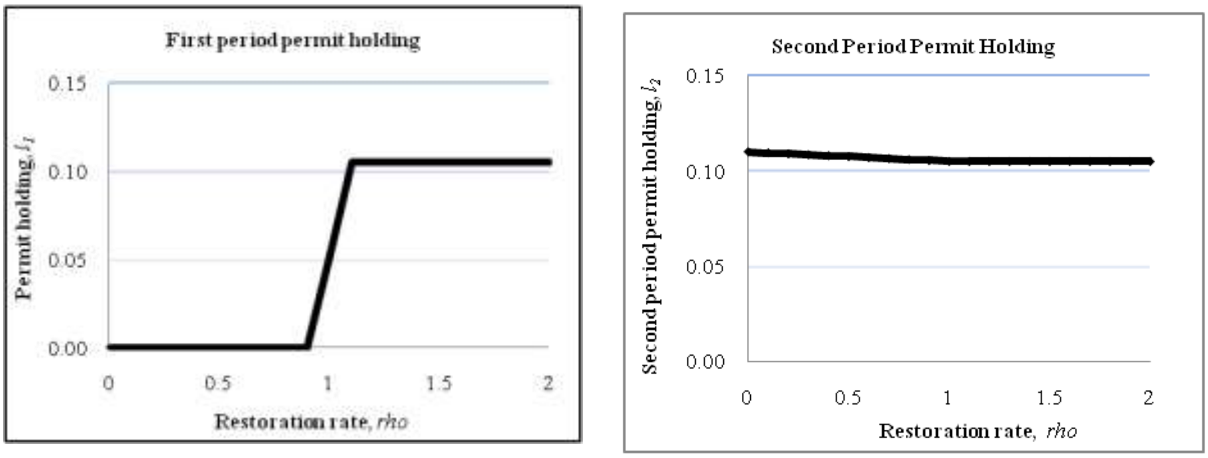
(b) The effect of restoration rates on the second period capital use (production levels)

Figure 3 The Effect of Restoration Rates on Production Levels under the MGP

With regard to the choice of compliance strategies, the effect of restoration rate is as expected (Figure 4). Holding everything else constant, profit-maximising firms will increase their abatement levels with a higher restoration rate, and firms find the efficient abatement level when the restoration rate reaches one. Accordingly, in the first period, firms will not hold any permits when the restoration rate is still below one, because it will be cheaper for them to invest in abatement or violate and shift the purchase of permits to the second period. When the restoration rate exceeds one, then firms should hold the efficient number of permits. This result also explains why the violation level in the first period is positive when the restoration rate is less than one. Likewise, in the second period, firms hold the efficient number of permits regardless of the restoration rate after the restoration rate reaches one.



(a) The effect of restoration rates on abatement levels



(b) The effect of restoration rates on the first period permit holdings

(c) The effect of restoration rates on the second period permit holdings

Figure 4 The Effect of Restoration Rates on Compliance Strategies under the MGP

Proposition 5 *When the restoration rate is set at the optimal level, this restoration rate affects the firm's compliance strategies because it forces the firm to be compliant by either making an efficient investment in abatement measures or holding the efficient number of permits.*

Proof From equation (40) and (43) we derive the effect of the restoration rate on the amount of abatement.

$$\frac{\partial L}{\partial a_i} = 2c_{a_i} - p_2\rho + \lambda_1 + \lambda_2 \geq 0 \ , \ \frac{\partial L}{\partial \lambda_1} = e_{i1}\left(q_{i1}\left(k_{i1}\right)\right) - a_i - l_{i1} \geq 0 \ , \ \frac{\partial L}{\partial a_i} = \frac{\partial L}{\partial \lambda_1} \geq 0$$

By re-arranging the terms, we obtain a new function, $g(a)$, that is expressed in terms of abatement level.

$$g(a) = 2{c_{a_i}}^{-1}(a_i) + a_i = e_{i1} - l_{i1} + p_2\rho - \lambda_1 - \lambda_2 \tag{49}$$

Taking the first derivative of $g(a)$ with regard to the restoration rate yields

$$\frac{\partial g(a_i)}{\partial \rho} = p_2 \geq 0 \tag{50}$$

As $p_2 \geq 0$, the increasing restoration rate will increase the firm's abatement level until it reaches the efficient abatement level under the optimal restoration rate. When $\rho < \rho^*$, it will be optimal for the firm to be non-compliant, and hence, it has a positive violation in the first period. With regard to permit holding, a non-compliant firm does not need to buy permits, as $v_{i1} > 0 \Rightarrow l_{i1} = 0$. However, when $\rho \geq \rho^*$, compliance becomes an optimal strategy for the firm, and hence it will hold the efficient number of permits, $v_{i1} = 0 \Rightarrow l_{i1} > 0$. *Q.E.D.*

7. Mixed Penalty Design

Under the mixed penalty design (MIX), firms will be penalized with a restoration rate when they violate in the first period and will also be fined with a fixed rate, f, for their total violation level in both periods. The assumptions in this model follow from both the Fixed Penalty Rate (FPR) and Make-Good Provision (MGP) model. The important difference is that we allow for non-compliance in the second period, which is not the case in the MGP model. As before, it is assumed that we have a zero discount rate.

Firm *i*'s maximisation problem:

$$\begin{aligned} \underset{k_{i1},k_{i2},a_i,l_{i1},l_{i2}}{Max} \quad \Pi_i &= \sum \tau\, q_{it}(k_{it}) - \sum r\, k_{it} - \sum c_i(a_{it}) \\ &\quad - p_1[l_{i1} - \bar{s}_{i1}] - p_2[l_{i2} - \bar{s}_{i1} + \rho\, v_{i1}] - f[v_{i1} + v_{i2}] \end{aligned} \tag{51}$$

subject to $\boldsymbol{v_{i1}} \geq 0,\ \ \boldsymbol{v_{i2}} \geq 0$

The Lagrangian equation from the profit maximisation problem:

$$\begin{aligned} L = &\sum \tau\, q_{it}(k_{it}) - \sum r\, k_{it} - \sum c_i(a_{it}) - p_1[l_{i1} - \bar{s}_{i1}] \\ &- p_2[l_{i2} - \bar{s}_{i1} + \rho\, v_{i1}] - f[v_{i1} + v_{i2}] + \lambda_1\, v_{i1} + \lambda_2\, v_{i2} \end{aligned} \tag{52}$$

The first-order conditions are obtained from the Kuhn-Tucker conditions:

$$\frac{\partial L}{\partial k_1} = \tau \frac{\partial q_{i1}}{\partial k_{i1}} - r - [p_2 \rho + f - \lambda_1]\left(\frac{de_{i1}}{dq_{i1}}\frac{\partial q_{i1}}{\partial k_{i1}}\right) \le 0$$

$$\frac{\partial L}{\partial k_2} = \tau \frac{\partial q_{i2}}{\partial k_{i2}} - r - [f - \lambda_2]\left(\frac{de_{i2}}{dq_{i2}}\frac{\partial q_{i2}}{\partial k_{i2}}\right) \le 0 \qquad \textbf{(53)}$$

$$\frac{\partial L}{\partial a_i} = -2c_a + p_2 \rho + 2f - \lambda_1 - \lambda_2 \le 0 \qquad \textbf{(54)}$$

$$\frac{\partial L}{\partial l_{i1}} = -p_1 + p_2 \rho + f - \lambda_1 \le 0 \qquad \textbf{(55)}$$

$$\frac{\partial L}{\partial l_{i2}} = -p_2 + f - \lambda_2 \le 0 \qquad \textbf{(56)}$$

$$\frac{\partial L}{\partial \lambda_1} = e_{i1}(q_{i1}(k_{i1})) - a_i - l_{i1} \ge 0$$

$$\frac{\partial L}{\partial \lambda_2} = e_{i2}(q_{i2}(k_{i2})) - a_i - l_{i2} \ge 0 \qquad \textbf{(57)}$$

$$k_{i1} \ge 0, \quad k_{i1}\frac{\partial L}{\partial k_{i1}} = 0; \qquad k_{i2} \ge 0, \qquad k_{i2}\frac{\partial L}{\partial k_{i2}} = 0;$$

$$a_i \ge 0, \; a_i \frac{\partial L}{\partial a_i} = 0; \qquad l_{i1} \ge 0, \; l_{i1}\frac{\partial L}{\partial l_{i1}} = 0; \qquad l_{i2} \ge 0, \; l_{i2}\frac{\partial L}{\partial l_{i2}} = 0;$$

$$\lambda_1 \ge 0, \; \lambda_1 \frac{\partial L}{\partial \lambda_1} = 0; \qquad \lambda_2 \ge 0, \; \lambda_2 \frac{\partial L}{\partial \lambda_2} = 0 \qquad \textbf{(58)}$$

Proposition 6 *Under the mixed penalty design, the same results as in FPR and MGP models are derived in which the firm maximizes its profit by equalizing its marginal benefit after compliance cost to its marginal cost of production.*

Corollary 4 *The efficient level of investment in abatement measures is attained by setting the firm's marginal abatement cost* equal to the permit price in both periods.

Proof: Assuming an interior solution, we derive $\lambda_1 = p_2 \rho + f - p_1$ and $\lambda_2 = f - p_2$ from (55) and (56). We substitute these equations into equation (54) to obtain the efficient choice of abatement level:

$$-2c_a + p_1 + p_2 = 0 \qquad \textbf{(59)}$$

In the same way, we substitute λ_1 and λ_2 into the Lagrange derivative of capital to obtain

$$\left(\tau - p_1 \frac{de_{i1}}{dq_{i1}}\right)\frac{\partial q_{i1}}{\partial k_{i1}} - r = 0$$

$$\left(\tau - p_2 \frac{de_{i2}}{dq_{i2}}\right)\frac{\partial q_{i2}}{\partial k_{i2}} - r = 0 \qquad \textbf{(60)} \quad \text{Q.E.D.}$$

As with the MGP, firms will be profit-maximising when the marginal revenue in each period is equal to the sum of capital rent and marginal compliance cost.

To determine the efficient penalty level in the MIX, we need to look at each penalty element separately and then assess what happens when we vary their levels. Based on the first conditions (FOC), we see two differences between the MIX and the MGP. Firstly, the fixed penalty rate now appears in both equations of partial derivative to permit holding in each period (55) and (56). Secondly, the sign of the partial derivative to the Lagrange multiplier with regard to the second period violation has changed because the MIX allows for non-zero violation. These changes lead to different implication in setting the efficient level of each penalty.

Proposition 7 *Under the mixed penalty design, the firm's compliance in the first period can be achieved by setting either the fixed penalty rate or the restoration rate at an efficient level. However, compliance in the second period is only attained by setting* $f \geq p_2$.

Corollary 5 *In the presence of a double penalty in the mixed penalty design, stronger compliance incentives are observed than in other models and market efficiency is retained as in the other models.*

Proof: Focusing on compliance decisions in the first period, we set equations (55) and (57) equal to zero to attain

$$e_{i1}(q_{i1}(k_{i1})) - a_{i1} - l_{i1} = v_{i1} = p_1 - p_2\rho - f + \lambda_1 = 0 \qquad \textbf{(61)}$$

Since $\lambda_1 \geq 0$, we can determine the efficient level of the fixed penalty rate.

$$f \geq p_1 - p_2\rho \qquad \textbf{(62)}$$

Likewise, we solve for the efficient level of the restoration rate under the mixed penalty design.

$$\rho \geq \frac{p_1 - f}{p_2} \qquad \textbf{(63)}$$

Both equations reveal that even if either $f = 0$ or $\rho = 0$, the other penalty type will still have a positive value and hold the firm's compliance at the efficient level. In the second period, however, firm's compliance relies only on the FPR.

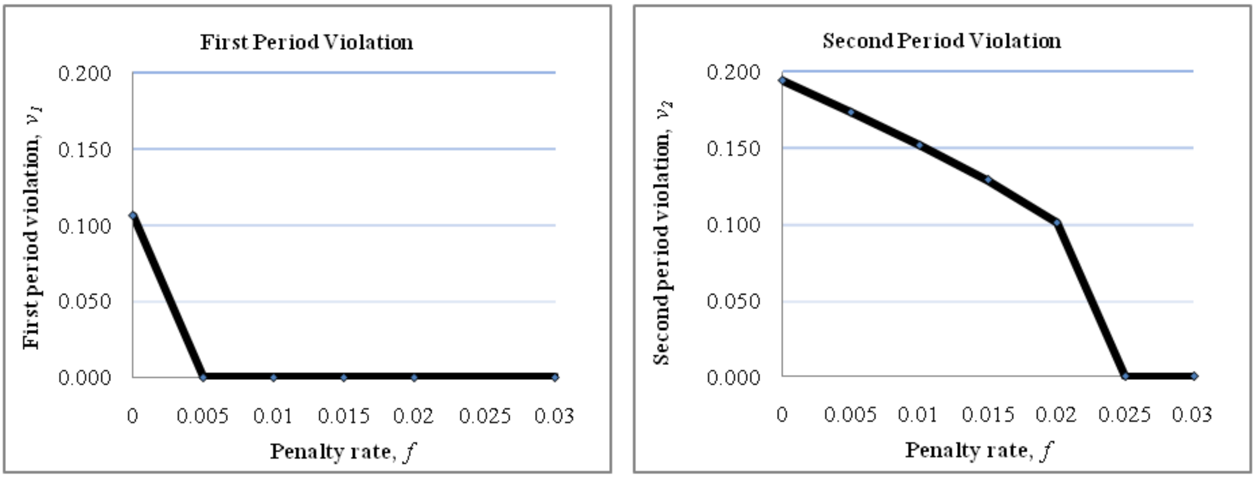
From (56) and (57), we obtain

$$e_{i2}\left(q_{i2}\left(k_{i2}\right)\right)-a_{i2}-l_{i2}=v_{i2}=-p_2+f-\lambda_2=0 \qquad \textbf{(64)}$$

Since $\lambda_2 \geq 0$, we derive the efficient fixed penalty rate for the second sub period

$$f \geq p_2 \qquad \textbf{(65)} \quad \text{Q.E.D.}$$

As in the previous penalty designs, we run a simulation; the same parameters and functions are used: $c=a^2$, $e=q^{0.9}$, $q=k^{0.5}$, $\tau=0.05$, $r=0.15$, $p_1=p_2=0.02$. Additionally, we set $\rho=1$ to determine the effect of an increasing penalty rate under the MIX. In general, the same effect as that in the FPR is obtained; a higher penalty rate decreases violation level until the efficient penalty rate is achieved. The difference is the efficient level of penalty. As seen from Figure 5, the presence of a restoration rate changes the efficient penalty level in the first period to a lower rate than it would be under the FPR alone because the restoration rate increases the cost of being non-compliance. Obviously, the efficient level of penalty will change accordingly depending on the restoration rate as determined by equation (62). On the contrary, a much higher efficient level of penalty rate is required in the second period under a restoration rate than under an FPR. This illustrates the implication of the comparative static result in equation (65).

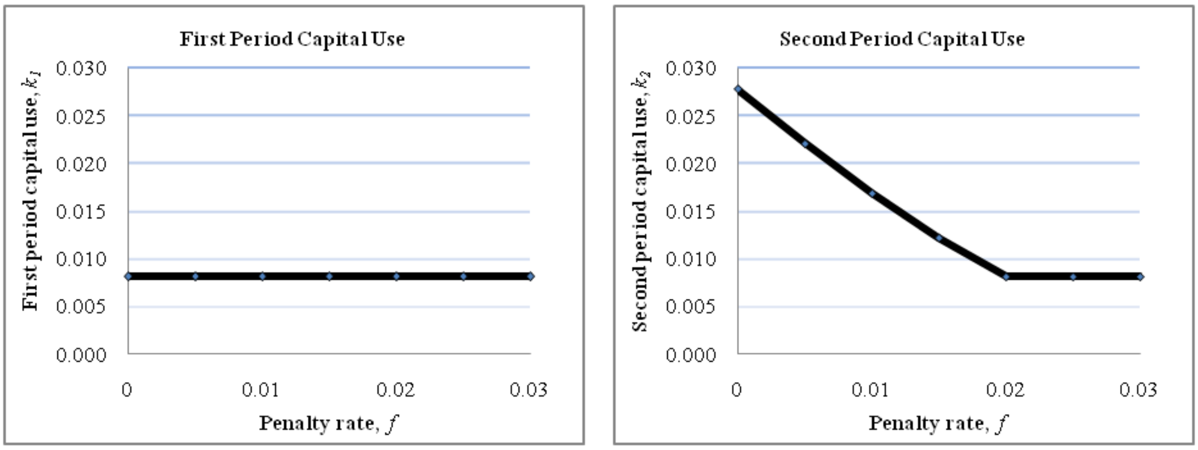


(a) The effect of penalty rates on the first period violation level under the MIX

(b) The effect of penalty rates on the second period violation level under the MIX

Figure 5 The Effect of Penalty Rates on the Firm's Violation Levels under the MIX

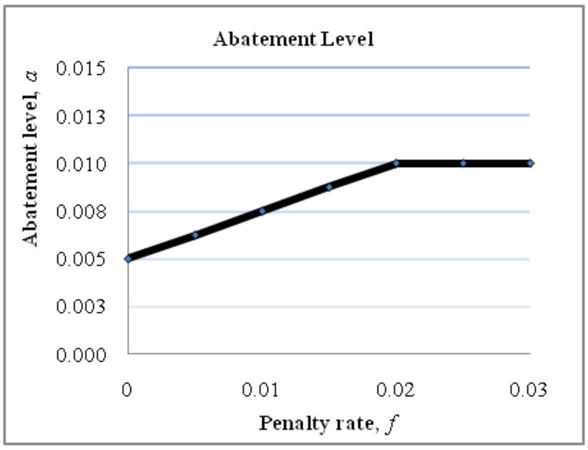
The simulation results of the comparative statics on capital use show that the double penalty in the first period forces the firm to choose its efficient production level regardless of the penalty rate. When the effect of restoration rate is removed in the second period, the penalty rate has a similar effect as that of the FPR.



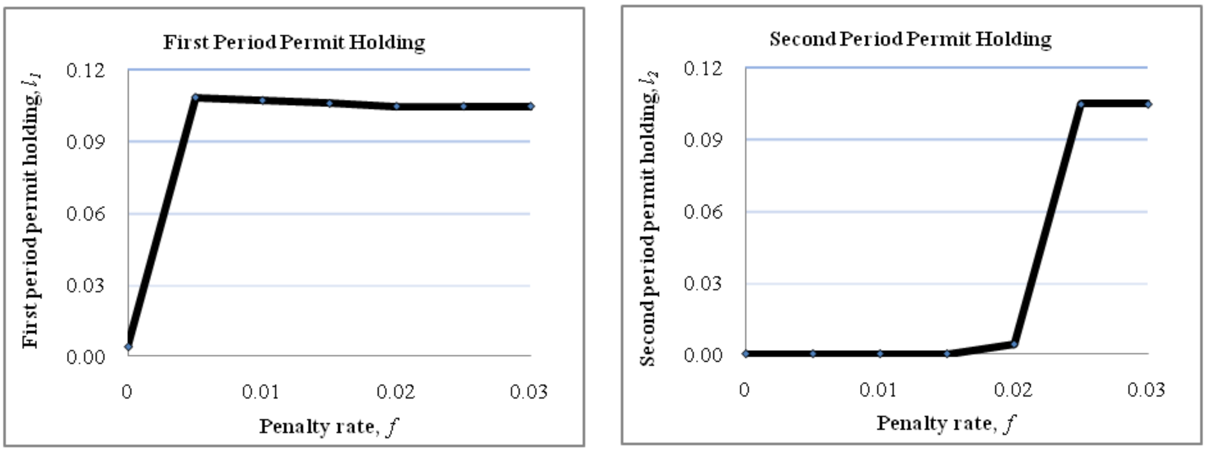
(a) The effect of penalty rates on the first period capital use (production levels) under the MIX

(b) The effect of penalty rates on the second period capital use (production levels) under the MIX

Figure 6 The Effects of Penalty Rates on Production Levels under the MIX



(a) The effect of penalty rates on abatement levels under the MIX



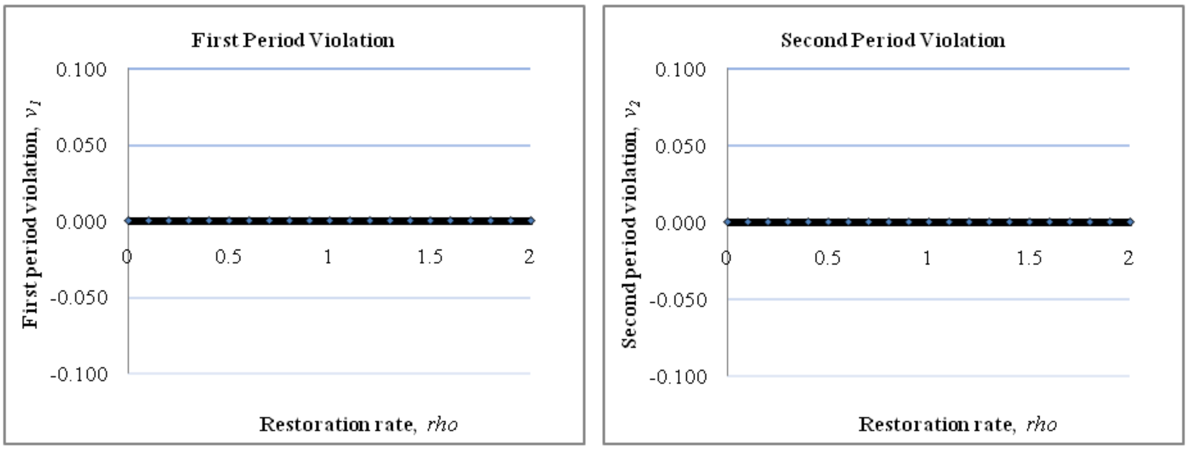
(b) The effect of penalty rates on the first period permit holding under the MIX

(c) The effect of penalty rates on the second period permit holding under the MIX

Figure 7 The Effect of Penalty Rates on Compliance Strategies under the MIX

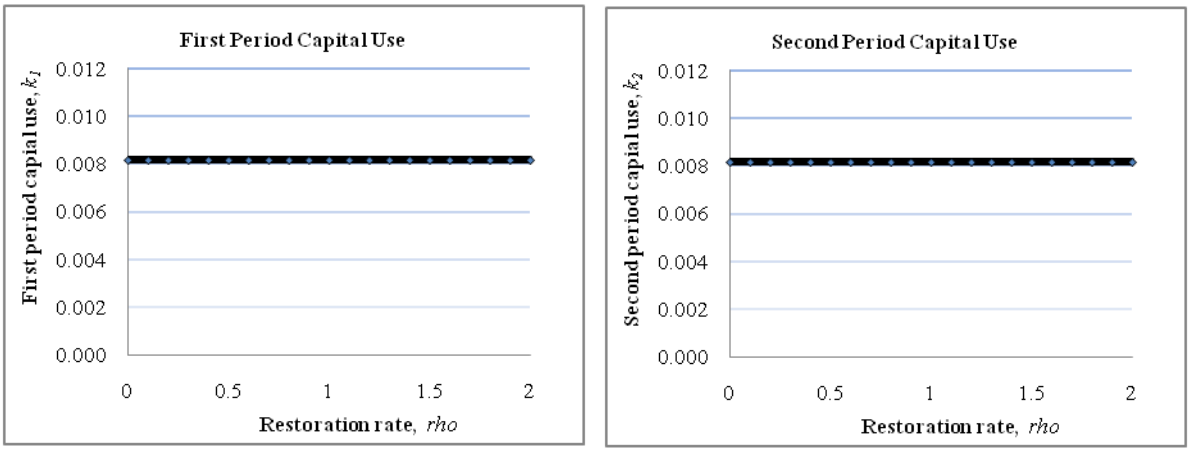
Compliance strategies under the mixed penalty design are affected by the penalty rate in fairly the same way as with the violation level and production level (capital use). Contrary to the effect on the second period capital use, increases in the penalty rate raise abatement levels until the efficient level is reached. The first period permit holding is increasing in the penalty rate until it reaches the efficient level. As in the case of violation level, a higher level of efficient penalty rate is required before permit holding in the second period achieves its equilibrium.

The restoration rate does not have a significant effect when an efficient level of penalty rate is enforced. A simulation of the comparative statics analysis is performed using the same parameters and functions as before, but with the penalty rate set at $f = 0.04$, which is twice the permit price.



(a) The effect of restoration rates on the first period violations under the MIX

(b) The effect of restoration rates on the first period violations under the MIX

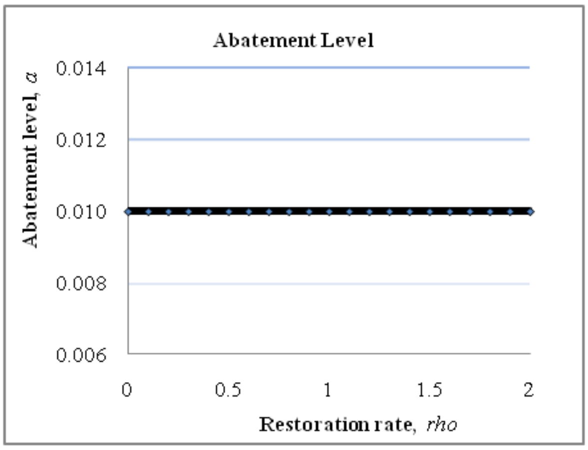


(c) The effect of restoration rates on the first period permit holding under the MIX

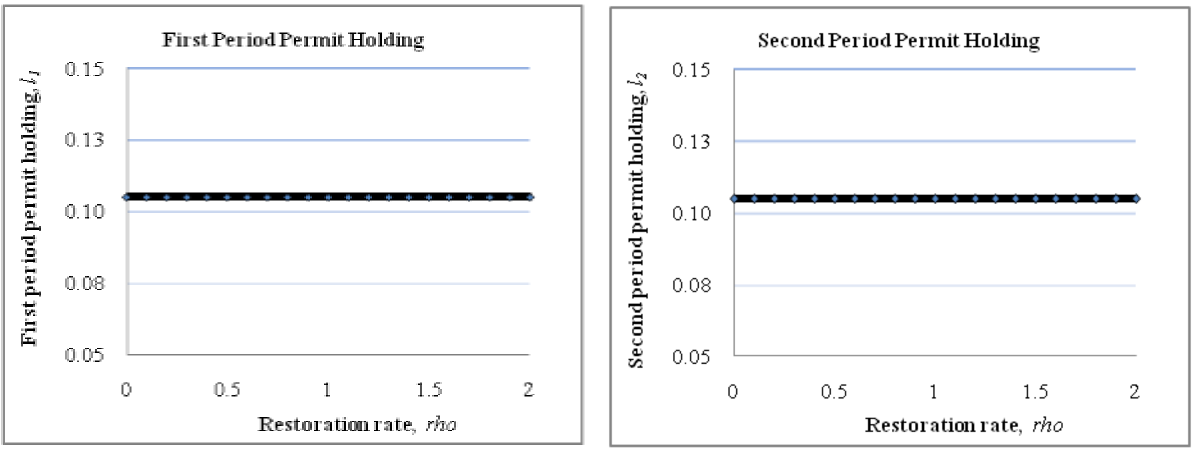
(d) The effect of restoration rates on the second period permit holding under the MIX

Figure 8 The Effect of Restoration Rates on Violation and Production Levels under the MIX

When we look at the violation level and production level, the simulation results indicate that the restoration rate does not play a role in determining the efficient level of those variables given that the penalty rate is established at the efficient level. Likewise, Figure 9 shows that all compliance strategy variables reach their efficient levels immediately at the beginning. Hence, the MIX model seems to guarantee that firms arrive at the efficient level of abatement and permit holding immediately, which is different compared to the other models. Nevertheless, it is worth noting that the strong compliance incentive under the MIX does not compromise market efficiency as proven by equation (60).



(a) The effect of restoration rates on abatement level under the MIX



(b) The effect of restoration rates on the first period permit holding under the MIX

(c) The effect of restoration rates on the second period permit holding under the MIX

Figure 9 The Effect of Restoration Rates on Compliance Strategies under the MIX

8. Discussion

As mentioned earlier, a theoretical model has the advantage of the simplicity of stylised facts. Nevertheless, some key issues related to the complexity of an emissions trading scheme need to be considered in order to gain an understanding of the implications that our results might have in practice.

One of the important functions of an emissions trading scheme is its role in the process of price discovery of the regulated pollutant. Typically, there is very little information about the permit price at the beginning of a trading scheme. At that point, the fixed penalty rate (*f*) is practically the first price signal received by firms, apart from their own marginal abatement cost, of the maximum compliance cost. We can thus see the penalty rate as a focal point that serves as the first external reference on which firms can base their compliance decisions. These decisions will be adjusted as firms receive more price signals from the permit markets. On this ground, the initial allocation rule might actually influence the process of price discovery. This shows that the assumption of perfect information is crucial for Montgomery's finding (1972) that the mechanism used to distribute initial permits to each firm should not affect the firm's behaviour in making an optimal decision. In his model he assumed that each firm should be able to calculate the equilibrium permit price. In the case of grandfathering, price signals are generated by the secondary market. When permits are initially auctioned, firms will gain more signals at the earlier stage of a trading scheme about the expected permit price. Thus it is expected that there will be a faster convergence path to the efficient equilibrium when permits are auctioned off.

Another important issue to address in practice is the assumption of perfect information on the regulator's part that is used in the model. By and large, a regulating authority does not necessarily have all the required information on the firm's characteristics or emissions inventories, let alone its marginal abatement costs. Thus the authority makes its choice of marginal penalty rate under imperfect information (e.g. uncertainty about future emissions, perceptions about the risk of illiquidity in permit market). In such a situation, high penalty levels might lead to overinvestment in reduction measures because the cost of potentially being non-compliant for firms will be high compared to the cost of reducing emissions under the presence of uncertainties regarding permit prices. This effect may increase when the number of permits is fixed, which means that the supply of permits will be inelastic in the short run. On the contrary, the penalty level may be set lower than the true equilibrium price, acting as a price cap that provides lower investment incentives.

Under the absence of perfect information about the equilibrium permit price, there is no certainty that the penalty rate will always be above the equilibrium permit price, which is a very crucial issue especially if the penalty rate also functions as a price cap. Hence, the question about the level of the penalty rate becomes relevant. This concern has prompted the idea of tying the penalty rate to an auction price to bring the penalty rate closer to the permit price and at the same time guarantee that the penalty rate will always be higher than the permit price given that the auction price is a good proxy of the permit price. The Australian federal government has put forward this concept in its proposal for the Australian trading scheme (the Carbon Pollution Reduction Scheme), in which the penalty rate is suggested to be capped at 110% of the benchmark average auction price. Nevertheless, it is important to note that, in practice, strategic bidding behaviour might drive down auction prices as firms understand that their bids will determine the maximum cost of compliance. Furthermore, this penalty design might create additional cost uncertainty as the level of penalty varies in auction prices that implies an uncertainty regarding the maximum cost of compliance, which in turn will influence compliance decisions. In an extreme case, market players might collude to drive down the auction price to zero, creating a zero compliance cost. However, it is unlikely that the regulator would allow this to happen, as the auction process should be designed to prevent such collusion from occurring and most auctions set a reserve price for that purpose. Most existing schemes, as shown by the rules applied by some Member States in the European Emissions Trading Scheme, also have additional penalties making violations a criminal offense and thus encourage further compliance by market players (Schleich et al., 2009). Firms are also exposed to additional reputational costs when they are non-compliant. In spite of this, it is worth noting the potential drawbacks of tying the penalty rate to the equilibrium permit price.

The use of a mixed penalty system should not affect the efficient compliance strategies of firms, whether compliance is achieved through investment in abatement or through permit trading. Although this penalty design is perceived as a double penalty, under perfect knowledge about the equilibrium permit price and as long as the level of both the penalty rate and the restoration rate are set at the efficient level, in theory the efficient market condition is retained. However, this double penalty might have more deterrent effect for risk averse market players and may encourage over-investment in abatement.

When a penalty rate and a restoration rate co-exist as a penalty design, firms arrive at the optimal compliance strategies earlier than they do under the other penalty designs. Based on the comparative statics analysis, the fixed penalty rate seems to have a more prominent effect on compliance strategy than the restoration rate. Yet, the theoretical result rejects some concern that the MIX model will yield a lower efficiency level given the double penalty.

As shown in Table 1, the mixed penalty design seems to be favoured in practice, such as in the US Acid Rain Program, the EU ETS, the RGGI and the emerging schemes. This seems to be in line with the theoretical findings as stronger compliance incentives encourage faster convergence toward the optimal compliance strategies.

9. Conclusion

Penalty design is an important element of permit markets that ensures that the market is capable of achieving both environmental effectiveness and economic efficiency. Our model shows that different penalty designs in the form of the fixed penalty rate (FPR) and the make-good provision (MGP) do not yield different results in terms of firms optimal compliance strategy as long as the penalty level is set at the efficient level such that the penalty rate is greater than the permit price and the restoration rate is greater than ratio of the permit price in the first period to the permit price in the second period. For the mixed penalty design, perfect compliance in the first period can be achieved by setting either the penalty rate or the restoration rate at the efficient level. Nevertheless, the penalty rate element evidently plays a more important role to ensure compliance in the second period.

Efficiency in a permit market should be maintained regardless of the penalty design, provided that each firm chooses its best compliance strategy and maximizes its profit by equalizing its marginal benefit to its marginal cost of production including compliance cost. Hence, if there is a firm in the market that does not choose its best compliance strategy, then market efficiency will be compromised.

Lastly, it is important to note that the final effects of different penalty designs on market efficiency are also influenced by the firms' risk attitudes, which are reflected in both investment levels and the number of permit holding. On this ground, either the distance of the penalty level from the equilibrium permit price or the penalty type itself, which is not an issue in theory, might be a crucial design element for the regulator to consider. We therefore suggest that further research in this area is important and that experimental testing seems to be a fruitful approach to further investigate penalty design implications.

References

ARGUEDAS, C. 2008. To Comply or Not To Comply? Pollution Standard Setting Under Costly Monitoring and Sanctioning. *Environmental & Resource Economics,* 41**,** 155-168.

BALDURSSON, F. M. & VON DER FEHR, N.-H. M. 2004. Price volatility and risk exposure: on market-based environmental policy instruments. *Journal of Environmental Economics and Management,* 48**,** 682-704.

BECKER, G. S. 1968. Crime and Punishment: An Economic Approach. *The Journal of Political Economy,* 76**,** 169-217.

BETZ, R., ROGGE, K. & SCHLEICH, J. 2006. EU Emissions Trading: An Early Analysis of National Allocation Plans for 2008-2012. *Climate Policy,* 6**,** 361-394.

CHAVEZ, C. A. & STRANLUND, J. K. 2003. Enforcing Transferable Permit Systems in the Presence of Market Power. *Environmental and Resource Economics,* 25**,** 65-78.

COASE, R. H. 1960. The Problem of Social Cost. *Journal of Law and Economics,* III**,** 1-44.

COMMUNITY INDEPENDENT TRANSACTION LOG. 2010a. Emissions trading: EU ETS emissions fall more than 11% in 2009. Available: http://europa.eu/rapid/pressReleasesAction.do?reference=IP/10/576 [Accessed 7 Aug 2010].

COMMUNITY INDEPENDENT TRANSACTION LOG 2010b. Verified Emissions for 2008-2009 and allocations 2008-2009.

CROCKER, T. D. (ed.) 1966. *The Structuring of Atmospheric Pollution Control Systems,* New York: WW. Norton & Co.

DALES, J. H. 1968. *Pollution, Property and Prices,* Toronto, University of Toronto Press.

DOWNING, P. B. & WATSON, W. D. 1974. The economics of enforcing air pollution controls. *Journal of Environmental Economics and Management,* 1**,** 219-236.

ELLERMAN, A. D., JOSKOW, P. L. & HARRISON, D. J. 2003. Emissions Trading in the US: Experience, Lessons, and Consideration for Greenhouse Gases.

ELLERMAN, A. D., JOSKOW, P. L., MONTERO, J. P., SCHMALENSEE, R. & BAILEY, E. M. 2000. *Markets for clean air: The U.S. acid rain program.,* New York, Cambridge University Press.

EPA CLEAN AIR MARKETS DIVISION 2006. An Overview of the Regional Clean Air Incentives Market (RECLAIM). US EPA.

GANGADHARAN, L. 2000. Transaction Costs in Pollution Markets: An Empirical Study. *Land Economics,* 76**,** 601-614.

GODAL, O. & KLAASSEN, G. 2006. Carbon trading across sources and periods constrained by the Marrakesh Accords. *Journal of Environmental Economics and Management,* 51**,** 308-322.

GREENBERG, J. 1984. Avoiding tax avoidance: A (repeated) game-theoretic approach. *Journal of Economic Theory,* 32**,** 1-13.

HAHN, R. W. 1984. MARKET POWER AND TRANSFERABLE PROPERTY RIGHTS. *Quarterly Journal of Economics,* 99**,** 753-765.

HAIMOV, M. M. S., NGUYEN, D., CHEN, F., HYNES, C., LEE, T., REMILLARD, K.-K. L., THOMAS, S. & TSAI, S. 2010. Annual RECLAIM Audit Report for 2008 Compliance Year. South Coast Air Quality Management District.

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

INDEPENDENT PRICING AND REGULATORY TRIBUNAL 2009. Compliance and Operation of the NSW Greenhouse Gas Reduction Scheme during 2008.

JACOBY, H. D. & ELLERMAN, A. D. 2004. The safety valve and climate policy. *Energy Policy,* 32**,** 481-491.

KEELER, A. G. 1991. Noncompliant firms in transferable discharge permit markets: Some extensions. *Journal of Environmental Economics and Management,* 21**,** 180-189.

KLING, C. & RUBIN, J. 1997. Bankable permits for the control of environmental pollution. *Journal of Public Economics,* 64**,** 101-115.

KOLSTAD, C. D. 1996. Fundamental irreversibilities in stock externalities. *Journal of Public Economics,* 60**,** 221-233.

LANDSBERGER, M. & MEILIJSON, I. 1982. Incentive generating state dependent penalty system : The case of income tax evasion. *Journal of Public Economics,* 19**,** 333-352.

MACGILL, I., OUTHRED, H. & NOLLES, K. 2006. Some design lessons from market-based greenhouse gas regulation in the restructured Australian electricity industry. *Energy Policy,* 34**,** 11-25.

MALIK, A. S. 1990. Markets for pollution control when firms are noncompliant. *Journal of Environmental Economics and Management,* 18**,** 97-106.

MONTERO, J.-P., SANCHEZ, J. M. & KATZ, R. 2002. A Market-Based Environmental Policy Experiment in Chile. *Journal of Law and Economics,* XLV**,** 267-287.

MONTGOMERY, W. D. 1972. Markets in licenses and efficient pollution control programs. *Journal of Economic Theory,* 5**,** 395-418.

MURPHY, J. J. & STRANLUND, J. K. 2006. Direct and market effects of enforcing emissions trading programs: An experimental analysis. *Journal of Economic Behavior & Organization,* 61**,** 217-233.

NEW SOUTH WALES GREENHOUSE GAS REDUCTION SCHEME 2008. Introduction to the New South Wales Greenhouse Gas Reduction Scheme (GGAS). *In:* GOVERNMENT, N. S. W. (ed.).

NEW ZEALAND GOVERNMENT 2007. The Framework for a New Zealand Emissions Trading Scheme. Ministry for the Environment and The Treasury.

REGIONAL GREENHOUSE GAS INITIATIVE 2009. RGGI Fact Sheet. Regional Greenhouse Gas Initiative.

ROBERTS, M. J. & SPENCE, M. 1976. Effluent charges and licenses under uncertainty. *Journal of Public Economics,* 5**,** 193-208.

SANDMO, A. 2002. Efficient environmental policy with imperfect compliance. *Environmental and Resource Economics,* 23**,** 85-103.

SCHLEICH, J., ROGGE, K. & BETZ, R. 2009. Incentives for energy efficiency in the EU Emissions Trading Scheme. *Energy Efficiency,* 2**,** 37–67.

STAVINS, R. N. 1995. Transaction Costs and Tradeable Permits. *Journal of Environmental Economics and Management,* 29**,** 133-148.

STRANLUND, J. K. 2008. Risk aversion and compliance in markets for pollution control. *Journal of Environmental Management,* 88**,** 203-210.

STRANLUND, J. K., COSTELLO, C. & CHÁVEZ, C. A. 2005. Enforcing Emissions Trading when Emissions Permits are Bankable. *Journal of Regulatory Economics,* 28**,** 181-204.

STRANLUND, J. K. & DHANDA, K. K. 1999. Endogenous Monitoring and Enforcement of a Transferable Emissions Permit System. *Journal of Environmental Economics and Management,* 38**,** 267-282.

THE PARLIAMENT OF THE COMMONWEALTH OF AUSTRALIA 2009. Carbon Pollution Reduction Scheme Bill.

UK DEPARTMENT OF ENERGY AND CLIMATE CHANGE 2010. The CRC Energy Efficiency Scheme User Guide.

US EPA. 2009a. Clean Air Markets: Emission, Compliance, and Market Data. Available: http://www.epa.gov/airmarkets/progress/ARP_1.html [Accessed 7 Aug 2010].

US EPA 2009b. NOx Budget Trading Program - Basic Information. *In:* AGENCY, U. E. P. (ed.).

VAN EGTEREN, H. & WEBER, M. 1996. Marketable Permits, Market Power, and Cheating. *Journal of Environmental Economics and Management,* 30**,** 161-173.

WESTERN CLIMATE INITIATIVE 2010. Design for the WCI Regional Program.