ISSN 1835-9728

Environmental Economics Research Hub

Research Reports

The Effects of Penalty Design on Market Performance: Experimental Evidence from an Emissions Trading Scheme with Auctioned Permits

Phillia Restiani and Regina Betz

Research Report No.87

December 2010

About the authors

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

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

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.

The views and interpretations expressed in these Reports are those of the author(s) and should not be attributed to any organisation associated with the EERH.

Because these reports present the results of work in progress, they should not be reproduced in part or in whole without the authorisation of the EERH Director, Professor Jeff Bennett (jeff.bennett@anu.edu.au)

Crawford School of Economics and Government

THE AUSTRALIAN NATIONAL UNIVERSITY

http://www.crawford.anu.edu.au

TABLE OF CONTENTS

Abstract

This paper investigates the behavioural implications of penalty designs on market performance using an experimental method. Three penalty types and two penalty levels are enforced in a laboratory permit market with auctioning, including the Australian Carbon Pollution Reduction Scheme proposed design of tying the penalty rate to the auction price. Compliance strategies are limited to undertaking irreversible abatement investment decisions or buying permits. We aim to assess how penalty design under the presence of subjects' risk preferences might affect compliance incentives, permit price discovery, and efficiency. In contrast to theory, we find that penalty levels serve as a focal point that indicates compliance costs and affects compliance strategies. The make-good provision penalty provides stronger compliance incentives than the other penalty types. However, the theory holds with regard to permit price discovery, as we find no evidence of the effect of penalty design on auction price. Interestingly, risk preference does not directly affect compliance decision, but it does influence price discovery, which evidently is a significant factor in compliance decisions as well as efficiency. Most importantly, a trade-off between investment incentives and efficiency is observed.

Keywords: emissions trading, penalty design, experiment, auction, irreversible investment, abatement, compliance

Acknowledgements

We thank, without implicating, AJ Bostian, Ben Greiner and Martin Jones for their comments and suggestions, as well as participants at the 2010 Asia Pacific Economic Science Association (ESA) Conference in Melbourne, 2010 International Energy Workshop in Stockholm, Institute of Information Systems and Management Seminar at the Karlsruhe Institute of Technology in Karlsruhe, and the 2010 International ESA Conference in Copenhagen.

This research was supported financially through the Environmental Economics Research Hub, which is funded by the Australian Commonwealth Environmental Research Facilities (CERF) programme.

1. Introduction

In order to achieve its environmental effectiveness, an emissions trading scheme requires a penalty that encourages compliance in the permit market. In spite of the wide recognition of its importance, the task of evaluating the efficacy of penalty design using empirical data is almost insurmountable due to differences in design features of the trading schemes as well as related market structures. Therefore, the use of a laboratory experiment offers an advantage in controlling for these design features and market parameters, which enables the isolation of the variables of interest.

This essay employs an experimental method to investigate the behavioural implications of penalty design on subjects' compliance decisions and is based on a theoretical model contained in Restiani and Betz (2010). In particular, we aim to assess how a specific penalty design in terms of penalty levels and penalty types in an auctioned permit market might induce different compliance incentives as well as market performance under the presence of subjects' risk preferences. We consider the fixed penalty rate, the make-good provision, and the mixed penalty design as penalty types as well as low and high penalty levels for the first two penalty types. As a test-bed of a proposed policy design in the Australian Carbon Pollution Reduction Scheme, the third penalty type includes a novel trait of tying the penalty rate to the auction price. To isolate the effects of penalty design, we abstract from exogenous uncertainties, such as shocks in the emissions levels and changes in product prices, in the experiment. Furthermore, a two-period model is employed to highlight the effect of irreversible investment decisions as a compliance strategy other than permit buying. The evaluation of market performance is carried out by assessing price discovery in the permit market, the efficient investment level, compliance rates, and static efficiency.

The sections in this paper are organised as follows. The second section conducts a literature review of penalty design in emissions trading schemes in practice, theory, and related experiments. This section also explains the motivation and the contribution of the study. Section three describes the experimental design, which is followed by the presentation of the experiment's hypotheses in section four. The following section, section five, displays results as shown by statistical summaries and the convergence path of some variables. Section six tests whether the results are statistically different from the equilibrium and further tests the hypotheses. Following results of the hypothesis testing, regression models are performed to control for other potentially influential variables in section seven. Section eight discusses the findings, and the last section concludes.

A penalty design that ensures the compliance of market participants is one of the key market design elements that enables an emissions trading scheme to deliver environmental effectiveness and economic efficiency. When a firm does not have the number of permits required for the levels of greenhouse gases it has reported, it will need to pay the penalty and/or 'make good' on its permit shortfall. From an economic perspective, it is interesting to see how firms choose to comply with regulations in a particular setting to maximise their profits.

Generally, three types of penalties are widely used in existing emissions trading schemes. The first is the fixed penalty rate (FPR) system, which sets a constant fine for each missing permit. For example, the New South Wales Greenhouse Gas Abatement Scheme and the Los Angeles Regional Clean Air Incentives Market (LA Reclaim) for NOx and SOx pollutants use this type. The second penalty type is the make-good provision (MGP), which requires firms to make up for their permit shortfalls according to a particular ratio. Under this system, firms do not have a direct financial penalty to pay. Examples of this system include the US Ozone Transport Commission NOx Budget Trading Program, which imposes a 3:1 ratio. The last penalty type combines the two types (mixed penalty). This is the most widely used penalty design, which serves as a double penalty to ensure that the relevant environmental goals are attained. This approach has been used in the European Union Emissions Trading Scheme (EU ETS) and some US emissions trading schemes. These practices are intended to prevent the continuous carrying-over of permit shortfalls, which in the end might undermine a scheme's reduction targets in the long term.

An important question with regard to penalty design is its effectiveness. A penalty is designed such that firms will choose to comply because the cost of being compliant is lower than the cost of not complying. Thus, a firm will prefer to buy enough permits to cover its emissions and avoid a penalty. Unfortunately, the equilibrium permit price, as a benchmark for penalty levels, is normally unknown to both the regulator and the firm. As the distance between the permit price and the penalty level decreases, the marginal benefit of being non-compliant increases. Thus, the question emerges of what the penalty level should be in the beginning, when the regulator does not have much information about the equilibrium permit price.

When compliance rates are used to measure the effectiveness of a penalty design, the existing schemes prove to have very high compliance rates. As permit prices can be quite volatile, the distance between penalty levels and permit prices varies accordingly. The Australian Carbon Pollution Reduction Scheme (CPRS) proposal links the penalty level to auction prices in an attempt to ensure that the penalty level would remain slightly above the expected permit price.

2. Literature Review

There have been many theoretical studies of enforcement in the context of pollution control. Among others, Malik (1990) employs a stylised enforcement model that focuses on audit probability and the magnitude of the penalty. By allowing for non-compliance rather than seeking an optimal enforcement scheme, the model shows that non-compliance will alter the equilibrium permit price and market efficiency. Only a few studies discuss penalty types with regard to emissions trading schemes. Nentjes and Klaasen (2004) discuss the compliance incentives under the Kyoto Protocol but they look at the emissions trading scheme as an implicit compliance incentive rather than focusing on the penalty design within the trading scheme itself. It is argued that in cases in which a seller's reputation costs are lower than that of buyers and in which seller liability applies[1], the provision to trade will induce overselling on the seller's part, resulting in a lower compliance rate. Nevertheless, this conclusion is only valid when no further penalty is enforced in response to seller non-compliance. Furthermore, this condition is not fulfilled in the existing trading schemes, because penalty costs are the same for all firms and reputational costs constitute an additional penalty. Additional penalty costs also exist due to different penalties for reporting violations in each Member State of the EU ETS. A study by CPB (2003) discusses the restoration rate or the make-good ratio as a means to induce early action rather than delaying in investment. A general equilibrium model is used to analyse the appropriate restoration rate under some particular scenarios and the degree of delay in six blocks of countries. The study suggests that the interpretation of the results is highly dependent on the particular setting of the model.

Trading schemes can have very different design elements, cover different industries, and operate in different market structures that are not directly comparable to one another. Thus, it is very difficult to assess the effectiveness of a particular penalty design empirically using field data. Experimental economics offers an approach in which subjects' decision-making can be observed while the parameters and environment in which a market operates are controlled within the laboratory. To our knowledge, there have been a limited number of experimental studies focusing on enforcement in the context of emissions trading schemes. Cason and Gangadharan (2006) use a dynamic enforcement model to assess the interactions among banking, uncertainty regarding emissions, and compliance. In the experiment, subjects were required to self-report their emissions level. The penalty design involves a higher audit probability and a fine for subjects found to falsely report their emissions levels. Their results show that a banking provision induces higher non-compliance. Murphy and Stranlund's study (2007) investigates the effects of targeted enforcement by differing marginal penalties, in terms of audit probabilities and penalty levels, for different characteristics of firms. They confirm the results of Stranlund and Dhanda's (1999) theoretical model that targeted enforcement does not increase the effectiveness of the enforcement scheme, although firms that are expected to be net buyers show a higher level of non-compliance than those that are expected to be net sellers.

[1] The seller's liability rule states that a sanction will be imposed on permit sellers that have oversold their permits without making sufficient emissions reductions.

There are a couple of experiments in tradable green certificates that also study the effects of penalty levels. Schaeffer and Sonnemans (2000) examine compliance incentives for mandatory and voluntary market participants under the provision of permit banking and/or borrowing, while Ivanova (2007) looks at the price effect of penalty levels in an oligopolistic market. Both studies conclude that penalty levels affect compliance rate and price levels. Nevertheless, incomplete design of treatment variables, non-balanced sessions for each treatment, and the absence of control on voluntary and mandatory demand of permits in Schaeffer and Sonnemans (2000) might all induce confounding effects on their findings. On the other hand, the oligopolistic setting in Ivanova (2007) does not suit our interest, as we investigate a competitive permit market. Furthermore, the specific roles of buyers and sellers are not assigned to the subjects in our experiment, in contrast to Schaffer and Sonnemans' and Ivanonva's respective experiments.

The existing literature seems to focus more on different audit probabilities and marginal penalties, targeted enforcement, and cheating as the main elements of enforcement in an emissions trading scheme. Employing a rather different perspective on the enforcement model, our study focuses on the types and levels of penalties given perfect monitoring. Hence, we do not consider the effect of audit probability. It is plausible that different penalty designs might, per se, have different effects on firms' behaviour, especially under the presence of uncertainties regarding permit price and risk aversion. In particular, our study aims to contribute to the literature by investigating the following aspects:

1) The effect of penalty levels

   Since emissions trading schemes are newly created markets, penalty levels provide information about the maximum compliance costs, which will facilitate price discovery in the markets. From this perspective, the penalty level can also be seen as an indication of a price cap, although it does not have binding force when the cap level is triggered. Some have argued that the presence of non-binding price controls, which are price ceilings (or floors) set above (or below) the equilibrium price, can be useful in lowering the costs of uncertainty and thus increasing the efficiency of emissions trading markets (Jacoby and Ellerman, 2004, Burtraw et al., 2010, Fell and Morgenstern, 2009, Szolgayova et al., 2008). An experiment by Isaac and Plott (1981) studying non-binding price controls indicates that price controls do not serve as a signalling price or a focal point. However, the study does not find conclusive evidence that price ceilings (or price floors) will bias prices below (or above) the competitive equilibrium. In contrast, Smith and Williams's (1981) experiment with a double auction market reveals that non-binding price controls affect price convergence. As mentioned previously, experiments in tradable green certificates find that a higher penalty level, combined with banking provision (Schaeffer and Sonnemans, 2000) or under the presence of market power (Ivanova, 2007), raises permit price.

Thus, the level of the penalty may influence the price discovery process in the market because it may steer the direction of prices.

2) The effect of penalty types

The penalty type itself may have different effects on a firm's behaviour. To our knowledge, no experiments have been done to test the effect of penalty type. Experiments focusing on emissions trading schemes that have used enforcement models have only employed fixed penalty rates. Given that the make-good provision (MGP) can be seen as a quantity penalty that allows non-compliant firms to 'borrow' from future permits, the cost of compliance under the MGP is reliant on future permit prices. Thus greater uncertainty about future permit prices will put more pressure on the cost of borrowing as well as uncertainty regarding marginal penalty rates. In contrast, a fixed penalty rate implies a fixed per unit cost of violation (fixed marginal penalty rate). Hence, the different nature of the two penalty types might affect firms differently in choosing their compliance strategies.

3) We abstract from any uncertainties other than those which arises from subjects' decisions.

Our focus is on the effects of penalty design on the performance of an emissions trading market. To isolate the effects of the treatment variable, we do not introduce any form of uncertainties in the experiment. The only uncertainties that might arise are those that stem from subjects' decisions during the experiment. Hence, it allows us to isolate the effects of penalty types and levels on compliance rates and compliance strategies.

We simplify the spectrum of a firm's compliance strategies to two options: irreversible investment decisions and permit holdings. Investment decisions are modelled as irreversible in order to parallel with real-world conditions in which once the investment is made, e.g. installing a more efficient turbine, the decision cannot be reversed and the scrap value of the installed abatement equipment is insignificant.

4) The mixed penalty design, which ties the penalty rate to the auction price.

The proposed Australian model uses a mixed penalty design in which the penalty level is set very close to the auction price, and the make-good factor is one. This design may encourage strategic bidding with regard to the auction price intended to drive down firms' compliance costs. Given its policy relevance, we also consider this penalty design in our experiment.

To sum up, this experiment aims to investigate the effect of penalty design on compliance incentives and market performance. In order to evaluate the effects of penalty design on market performance, we analyse permit prices and standard deviation of prices, the incentives of each penalty design on investment levels and compliance rates, and lastly how the penalty design affects efficiency. The details of the experimental design and the hypotheses upon which the evaluation is based are elaborated in the following sections.

3. Experimental Design

We consider five treatments with the type and level of penalty as our treatment variables. For each penalty type – the fixed penalty rate (FPR) and the make-good provision (MGP) –, we consider a low and a high penalty level. Additionally, we study a mixed penalty that combines the FPR and MGP mechanisms. For this mixed penalty, a low MGP ratio is used, and the penalty rate is linked to the auction price, as in the Australian model.

Table 1 Treatments for the Penalty Design Experiment

Penalty Design	Penalty levels	
	Low	High Level
Fixed penalty rate (FPR)	1.2 x equilibrium permit price Treatment I (AFL)	3 x equilibrium permit price Treatment II (AFH)
Make-good provision (MGP)	1:1 ratio Treatment III (AML)	1:3 ratio Treatment IV (AMH)
Mixed penalty	MGP low level + FPR (1.2 x auction price) Treatment V (AFM)	

The experiment procedure consists of three sessions, each of which involves two groups of eight subjects. The groups remain the same for the whole session. Thus, we have six observation groups for each treatment. The subjects are randomly allocated to a group so that they do not know whether the people next to them belong to the same group. In each session, the subjects participate in Holt and Laury's (2002) lottery-choice game before taking part in the emissions trading game[2]. However, the payoff from the Holt and Laury's experiment is only determined after the emissions trading game is concluded to avoid any endowment effects.

The subjects are undergraduate and postgraduate students at the University of New South Wales who were recruited through the ORSEE online recruitment system (Greiner, 2002). Each subject could

[2] The Holt and Laury's (2002) experiment asked subjects to choose 10 paired lottery choices A and B for which the probability of a higher payoff from both choices was increased. A consistent risk preference attitude will require a change from the safer lottery A to lottery B somewhere within the 10 pairs. The combinations of safe and risky choices constitute an index from 1 to 9 in which a higher number represents greater risk aversion.

only participate in one session as this experiment featured a between-subject design. Thus, a total of 240 students participated in the experiment. Each session lasted for 2.5-3 hours, and subjects earned an average of $24.48 for the emissions trading game and $34.20 for the whole session.

The emissions trading game was programmed using the University of Zurich's Z-Tree program (Fischbacher, 1999). Eight subjects in one group played six repeated rounds, and each round was comprised of 2 sub-periods. Although we use terms related to the emissions trading scheme context in this essay, the subjects received instructions worded using neutral terminology. Emissions or other environmental terminology was not used at all in the instructions (Appendix).

The subjects were told that they were firms that needed a license for each unit of good X that they produced. Firms' production levels before taking abatement measures represented their emissions levels. Permits expired at the end of each sub-period. If a firm did not hold enough permits, it would incur a penalty. The regulator would determine the emissions cap, which was set at 50% of firms' initial emissions levels, and this cap remained the same for both sub-periods. Firms had two compliance strategies from which they might choose: 1) making investment decision in abatement technology, and 2) holding enough permits to cover their emissions levels. Firms were allowed to undertake both measures although each firm's optimal compliance strategy was only either one of them.

The key features of the emissions trading game are as follows:

1. Stages of the game

   There are four stages in the emissions trading game, as shown in Figure 1.

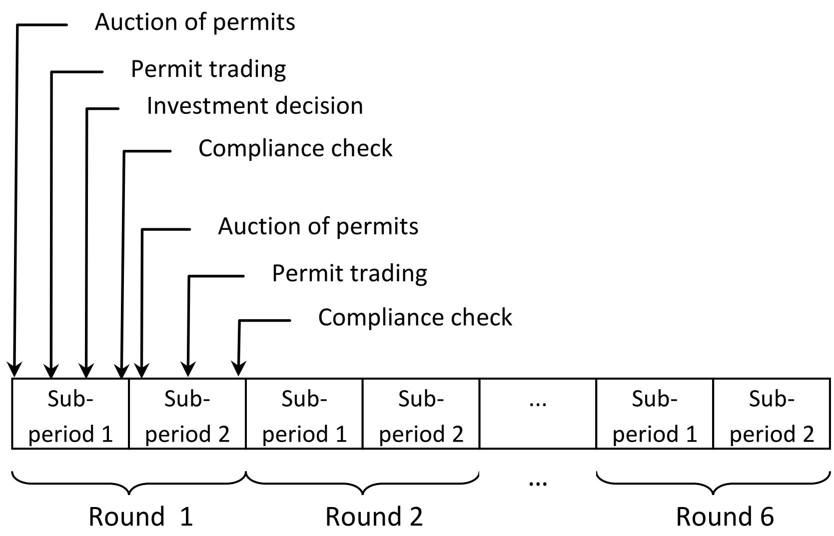


Figure 1 Stages in the Emissions Trading Game

a) Stage 1: Auction of permits

The initial allocation of permits to firms is conducted through an auction using an ascending clock auction[3]. The auction supply is fixed at 80 permits in each sub-period. At this stage, each firm needs to submit a non-negative bidding quantity during each bidding round. As the bidding rounds continue, firms are given the information about the gap between the aggregate demand and supply. This information may lead firms to collaborate to drive down the auction price (Klemperer, 2002). Nevertheless, this is highly unlikely considering the number of subjects and the random matching of those subjects in each session.

The bidding price starts at Experimental Dollar (EX$) 18 and increases in EX$5 increments. We start with a bidding price lower than the lowest firm's marginal abatement cost to allow all firms to submit a positive bidding quantity. We choose these price parameters so that we will have enough bidding rounds to allow for a better price discovery process but still keep the auction stage from becoming too long and the experiment from taking too much time.

When the aggregate demand of permits equals the aggregate supply, then the last bidding price becomes the auction price and firms are allocated permits as many as their last bidding quantity. If the aggregate demand for permits is less than the supply in a bidding round, then the auction price is the price in the penultimate bidding round, and firms are allocated their last bidding quantity plus any remaining excess supply. The excess supply is allocated according to the order of the fastest bidders. In this fashion, we follow the Virginia NOx auction model, which gives bidders incentives to submit their bids promptly (Holt et al., 2007, Porter et al., 2009).

The clock auction format is used due to its simplicity and transparency. This dynamic auction type, which includes multiple bidding rounds, allows subjects to think carefully in submitting their bids as bidding price increases (Compte and Jehiel, 2007). An ascending auction is also likely to allocate the goods to bidders with the highest value because they can always rebid and top lower-value bidders who may have bid aggressively in earlier bidding rounds (Klemperer, 2002).

b) Stage 2: Permit trading

[3] In an ascending clock auction, the price is increased with each tick of the clock, and the bidding price increases as long as the aggregate bidding quantity (aggregate demand) is higher than the total supply.

After receiving their permit allocation at the auction, firms trade permits using a posted-offer continuous double auction mechanism for one minute[4]. This mechanism allows firms to either buy or sell. Firms are free to accept any submitted (buy or sell) public offers at any time during the trading stage, although improved bidding rules are used to encourage faster convergence of offer prices[5]. Trade can only take place for each unit of license at a time. This double auction mechanism is a widely used trading institution in economic experiments and has proven to be highly efficient (Ledyard and Szakaly-Moore, 1994). Following each trade, firms receive updated public information regarding standing offers and trading prices as well as private information about their money and permit holdings. At the end of the trading stage, the average trading price from that sub-period is revealed to firms as a point estimate of the current price.

c) Stage 3: Investment decision (only for sub-period one)

The investment decision can be seen as a way to activate firm's abatement technology; hence, the investment cost is represented as the marginal abatement cost rather than as a lump sum capital cost. The decision to invest in abatement technology will ensure that the firm is compliant for both sub-periods in a round. To reflect the irreversibility of investments, we only allow firms to make investment decisions in the first sub-period; they cannot be changed in the second sub-period. During this stage, firms also find out whether they have an excess or shortfall of permits before they make investment decisions. Partial investments are not allowed. This is meant to encourage firms to learn about their best compliance strategies. If firms are short permits and decide to invest, they will be over-compliant, because investment automatically guarantees compliance and firms can no longer sell their permits at that stage in a sub-period.

d) Stage 4: Compliance check

The compliance check is the last stage in each sub-period where subjects learn about their earnings for that sub-period, their compliance status, and the penalties imposed on them if they are non-compliant.

2. Players' characteristics:

---

[4] We allow a relatively short trading period for the spot markets as the main permit allocation process should have taken place at the auction. Consequently, permit trading serves as a secondary trading institution used to 'clean up' the results of the auction if it does not yield the expected allocative efficiency.

[5] Improved bidding rules require that the buy offer be higher than the current highest-standing buy offer, whereas the sell offer should be less than the lowest-standing sell offer.

All firms produce a homogenous product and have the same production level of 20 units in each sub-period throughout the experiment. Firms are only differentiated by their constant marginal abatement cost (MAC), which is one of these values: $c_i \in EX\$\{20, 25, 30, 35, 40, 45, 50, 55\}$. The MAC is randomly allocated to each firm in each round and the set of MACs remains the same in all rounds. Based on the magnitude of a firm's marginal abatement costs, there are two types of firms: low-cost firms with MACs of EX$ 20-35 and high-cost firms with MACs of EX$ 40-55.

At the beginning of each sub-period, firms receive the same total revenue of EX$2800 from their production activity as the price of the good is exogenous and the same for everyone. Thus, the marginal revenue is constant at EX$140, but firms have different marginal benefits of each unit of good.

3. Information structure

   At the beginning of sub-period one in a round, subjects receive common information about their initial emissions level, the emissions cap, and the penalty design. This common information is known to all firms and remains the same during all rounds. At this stage, the subjects also receive private information about their marginal abatement costs, available money, and required number of licenses. In sub-period two, subjects are also reminded of their investment decisions and their compliance status in sub-period one. The information structure basically enables participants to estimate optimal decisions whether to invest in abatement technology or buy permits in the market. Subjects can even calculate where the equilibrium price should be under the assumption of risk neutrality. Nevertheless, different risk attitudes and different expectations regarding prices may create uncertainty in permit price.

4. Banking and borrowing are not allowed.

   Because the focus of our experiment is on the penalty design, we simplify our two-period model to abstract from the effect of banking. By allowing neither banking nor borrowing, we attempt to keep the market structure the same for both sub-periods; banking might create upward pressure on the expected permit price in the first sub-period and will add more noise to the results. Hence the expected permit price should remain the same across the two sub-periods in one round.

5. Penalty

   The enforcement of the penalty design in the emissions trading game is conducted as follows:

   a) In the FPR treatment, the penalty is imposed at the end of each sub-period during the compliance check. If a firm is non-compliant, the penalty costs are deducted from that firm's earnings. The penalty rate is EX$ 45 for the low level FPR and EX$ 114 for the high level FPR.

b) In the MGP treatments, the penalty is enforced differently for the two sub-periods.

   i) Non-compliance in sub-period one has no financial penalty but the violating firms need to surrender the quantity of the missing permits by a ratio. For example, in the high level MGP treatment using a ratio of 3:1, if a firm is two permits short, then it must hold six additional permits in sub-period two.

   ii) As firms cannot further compensate ("make good") for non-compliance at the end of sub-period two, we attempt to deter non-compliance by imposing an enormous financial penalty which is equivalent to the firm's total revenue (EX$ 2800).

6. Payoff

Firms can maximise their payoff by minimising their compliance costs or by maximising their profits from selling permits during the trading stage. The payoff function is the same for all firms.

Payoff =

\+ total revenue

\+ cash balance in sub-period one of the same round

– number of licenses bought in auction x auction price

– investment costs

– trading price of licenses bought during trading stage

\+ trading price of licenses sold during trading stage

– penalty costs

This payoff is accumulated for all rounds, and subjects' earnings are shown at the end of each round. Nevertheless, the amount of money that subjects receive in the beginning of each round (sub-period one) is always equal to the total revenue. This helps us to avoid the issue of wealth effects for the subjects as the round goes on.

4. Hypotheses

In the competitive equilibrium, the permit price should lie between EX$35-40 under perfect compliance. Considering the design of the auction bidding price, the auction price should reach its equilibrium at a price of EX$38. At this equilibrium, the best compliance strategy for low-cost firms is an investment decision, while the high-cost firms should comply by buying permits. The equilibrium permit price is achieved when each firm chooses its best compliance strategy.

The findings from our theoretical models (Restiani and Betz, 2010) indicate that firms will find it optimal to comply as long as the penalty rate is set higher than the equilibrium permit price (*Proposition 1*) or as long as the make-good ratio (restoration rate) is higher than or equal to one when permit prices remain the same in both sub-periods (*Corollary 3*). The mixed penalty design, which uses both types of penalties, supports those findings, and the presence of the double penalty ensures that firm compliance is still achieved. For this penalty regime, the comparative static analysis indicates that, although the level of one penalty type is varied, compliance is still maintained due to the presence of the other penalty type (*Proposition 7*).

Based on those findings and the parameters of the experimental design, we derive the following hypotheses:

Hypothesis 1: The auction price should remain the same in all treatments because the supply and demand structure remains the same.

Based on the Law of One Price, it is expected that the auction price will be the same as the trading price.

Hypothesis 2: In the FPR treatments, investment levels and compliance rates should be the same at 100% regardless of the penalty levels because the penalty rate is set higher than the theoretical equilibrium permit price.

Hypothesis 3: The make-good ratio should not affect investment levels or compliance rates in the MGP treatments as long as it is set equal to or higher than one under the assumption that prices remain the same in the two sub-periods.

Hypothesis 4: At the high penalty level, the compliance rates and investment levels in the FPR treatments should be the same as those in the MGP treatments. At the low penalty level, similar results are expected, although the penalty level is slightly higher in the low level FPR treatment (with a factor of 1.2), whereas the make-good ratio is 1.

Hypothesis 5: The mixed penalty treatment should yield the same compliance rates as the FPR and MGP treatments.

5. Results

5.1. Subject Risk Preferences

The Holt & Laury (2002) lottery choice experiment shows that more than 75% of the subjects are risk neutral, slightly risk averse, or risk averse, as expected. We observe that some subjects made inconsistent risk preference choices by changing from one lottery to another more than once. However, only about 20% (or fewer) of the subjects in each treatment showed these inconsistencies (Figure 2). The data on subjects' risk preferences are later used in the estimation models because their risk attitude may be an important determinant of compliance strategy and auction prices.

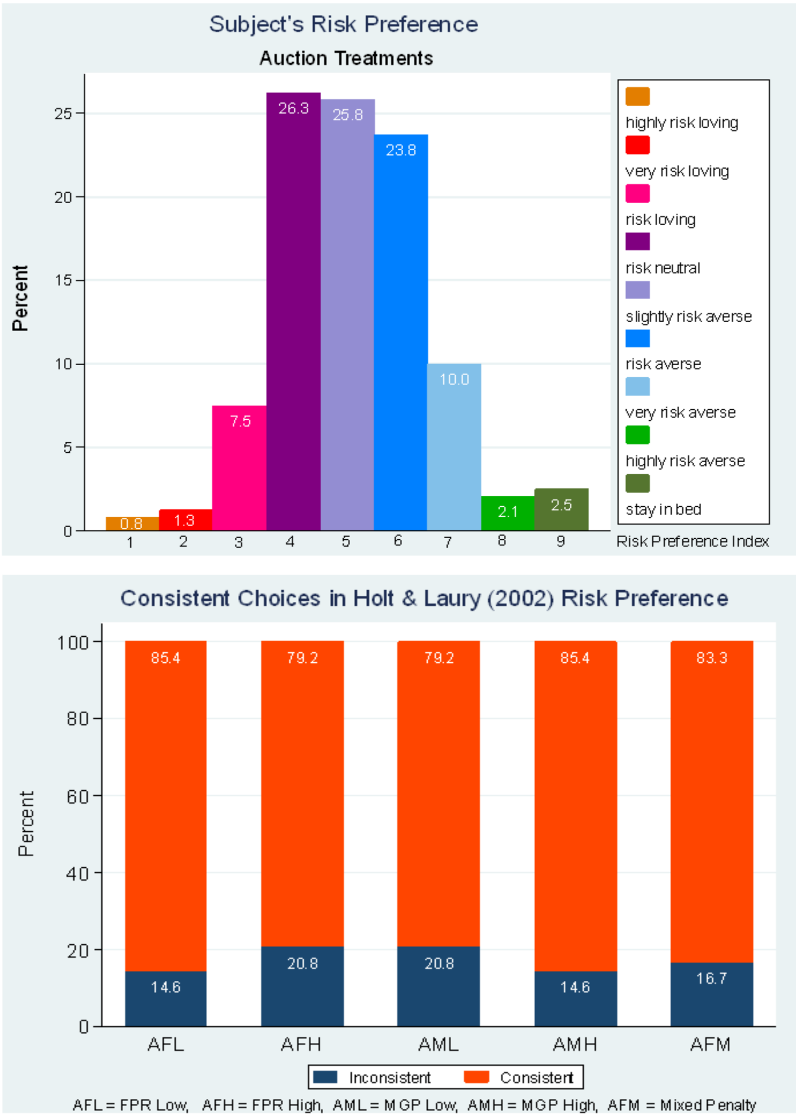


Figure 2 Results from the Lottery Choice Experiment

5.2. Convergence Path and Statistics Summary

This section discusses the statistics summary for the main variables of interest pertaining to prices, compliance strategies, and efficiency. Graphically presenting a particular variable over a period facilitates a closer inspection of the data at the group level and illustrates the existence of a discernible convergence path.

5.2.1. Price Variables

The results show that a high penalty level for each penalty type encourages a higher auction price (Figure 3). It is also clear that the auction has served as the primary market for distributing permits instead of the spot market (trading stage), with trade volume at less than 15% of the total number of permits in the market. Furthermore, the mean trading price is lower than the auction price in all treatments. Correspondingly, treatments with a high penalty level also have higher mean trading prices. The mixed penalty rate, which can be viewed as a double penalty regime, results in the highest mean trading prices, which are close to the auction price.

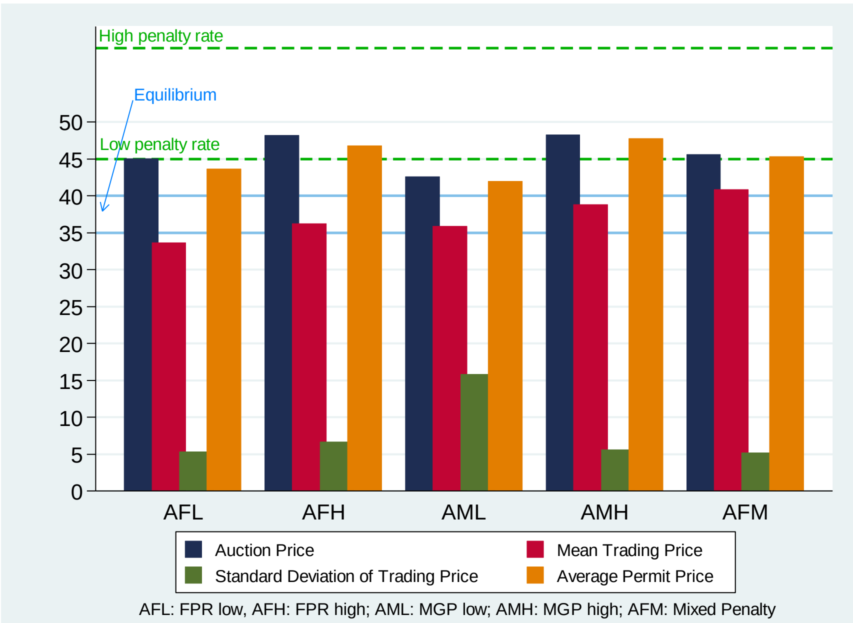


Figure 3 Price Variables by Treatment

When the standard deviation of prices is considered as a measure of strong price signals, the statistics in Table 2 reveal that there is a pattern across treatments that is similar to that of auction prices. High penalty levels in the FPR penalty treatment are related to higher standard deviations for all three measures of prices. However, a more ambiguous link is found in the MGP treatments: the trading prices in the high MGP treatment have a lower standard deviation, but both auction price and average permit price have higher standard deviations. In this sense, the FPR provides a stronger price signal than does the MGP penalty, while the mixed penalty performs fairly well in this regard.

The difference between auction prices and mean trading prices may indicate that the spot market is used to dispose of any unwanted excess permits that a subject obtains. The resale value of the permit may be lower because subjects are keen to minimise the loss that they accrued by acquiring such excess permits at the auction. On the contrary, this can also highlight how the permit demand has been falsely increased at the auction, possibly due to strategic bidding behaviour rather than to a real need to acquire permits for compliance purpose. The data show that in 120 out of 360 observations, the auction price is equal to or higher than the mean trading price, which confirms that some buyers at the auction realise gains by trading in the spot markets. However, most of the time those buyers make losses as mean trading prices tend to be lower than the average auction price. A lower price at a resale market is not unusual as the opportunity of having a resale market induces a common-value character at the auction. Hence, subjects' bidding quantities are not only motivated by their private valuation of the permit but may be biased by the speculative wish to realise some gains at the resale market or the secondary market (Haile, 2003, Garratt and Tröger, 2006). As a result, inefficiency occurs at the auction and this inefficiency is increasing in the uncertainty regarding the common-value or the resale value (Goeree and Offerman, 2003).

Table 2 Statistics Summary for Prices Variables (EX$)

Treatment	Mean auction price	S.d.[a] auction price	Mean trading prices	Mean s.d. of trading prices in each sub-period	Ave.[b] permit price	S.d. of ave. permit price	Trade volume (permits)
FPR Low (AFL)	45.01	11.59	33.63	5.36	43.66	10.72	11.75
FPR High (AFH)	48.21	11.79	36.25	6.70	46.80	11.19	10.25
MGP Low (AML)	42.58	21.73	35.91	15.82	41.94	20.82	6.35
MGP High (AMH)	48.28	27.67	38.85	5.63	47.77	27.00	8.22
Mixed Penalty (AFM)	45.57	14.09	40.85	5.23	45.30	13.97	6.81
Optimum	38	0	35-40	0	35-40	0	0

Notes: [a] s.d. = standard deviation, [b] ave. = average

The average permit price reflects the volume-weighted average based on the auction and trading price. As the volume traded in the secondary market is low, it is not very different from the auction price. The average permit price in all treatments is constantly higher than the range of prices in the efficient equilibrium. The average prices lie around the low level FPR and remain well below the high level FPR.

Given the complexity of our experimental design, it is expected that a learning curve might affect subjects' decision making during the experiment. This learning effect is not directly observable from mean values but will be more discernable through a convergence path for permit prices over time. Based on the scatter plot and the fitted line of the average permit price using *lowess smoothing*[6] in Figure 4, we observe a general convergence path for permit price in all treatments, although this convergence pattern is stronger in the FPR and mixed penalty treatments than in the MGP treatments. Furthermore, an end-game effect[7] is evident in both of the MGP treatments as the price plummeted during the last round of the experiment.

As shown in Figure 4(a), the permit price in the FPR low level treatment (AFL) begins above the low level penalty rate, (EX$45) in the first round and then rises slightly in the second round before starting to fall until round five, when it flattens out around the equilibrium price. A similar trend is also apparent in the FPR high level treatment (AFH). Nevertheless, the starting point of the permit price in the first round is slightly higher than in the AFL treatment. Moreover, in the two last rounds, the price stabilises above the equilibrium range and just slightly below the low penalty level of EX$45. The range of the scatter plot is also higher for AFH than for AFL.

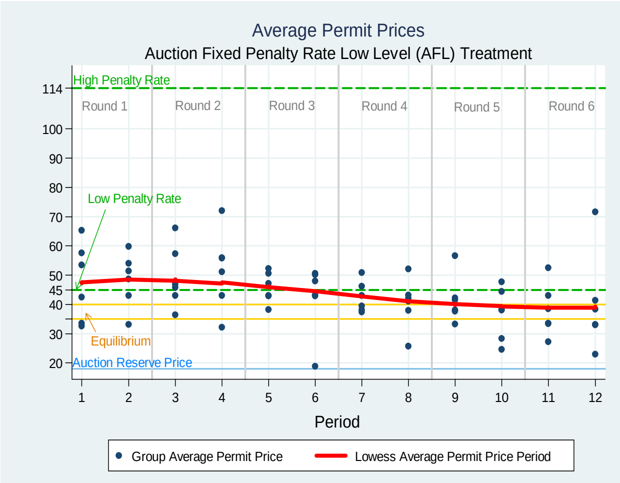
In contrast, the convergence path for permit prices is markedly different for the low and high level MGP treatments (Figure 4 (c) and (d)). The price in the low level MGP treatment (AML) begins at a very high level at EX$70 in the first sub-period of the first round. Then it starts to fall and continues to do so until the third round. Despite some fluctuation, the price then remains near the equilibrium until round five. Afterwards, there is an end-game effect in the last round. On the other hand, a more pronounced pattern of convergence of permit price over time is observed in the high level MGP treatment (AMH). The starting price in the first round is close to those of the FPR treatments at EX$50. Subsequently, the price remains adjacent to the low-level penalty rate at EX$45 for four rounds before the end-game effect takes place in the last round. The range of the scatter plot in the MGP treatment is much larger than that in the FPR treatments, and more outliers are observed with a maximum permit price of EX$193 in AMH. This is due to the penalty design in sub-period two for the MGP treatments in which subjects will lose a total revenue of EX$2800, which is equal to a marginal penalty of EX$140 for 20 units of permits. If a subject is less than 20 units short on permits, the value of the marginal penalty will increase accordingly. Therefore, the permit price in sub-period two of the MGP treatments can rise to a very high level.

---

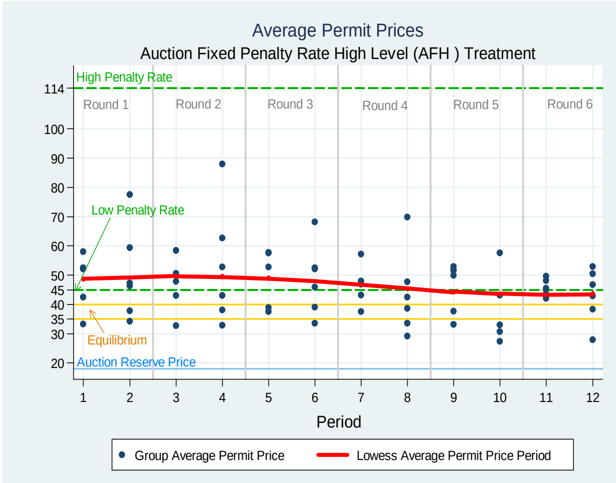
[6] Lowess is a locally weighted regression of y (dependent variable) on x (independent variable). It is normally used for scatterplot smoothing and is desirable because it tends to follow the data.

[7] An end-game effect is a systematic change in behaviour that occurs as an experiment with repeated rounds reaches its conclusion.

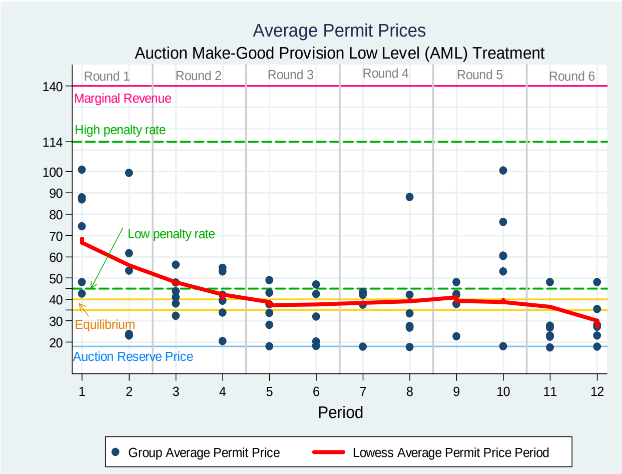
The Mixed Penalty treatment (AFM) shows a convergence path that is more similar to that of the FPR treatments than to that of the MGP treatments. In the first round, the price starts slightly higher at around EX$55. After that, the price continuously falls until it stabilises right around the efficient equilibrium in the last round. Nevertheless, the range of permit prices in this treatment is obviously larger than those in FPR treatments, possibly due to the effect of having MGP element in the penalty design. Given that the penalty rate is linked to the auction price, we expect to see strategic bidding behaviour drive down the auction price and hence the compliance cost. Although relatively low prices are observed in period ten (sub-period two of round five), these prices rise again in the next period. Hence, there is no clear evidence of that sort of collusion to drive down auction prices.



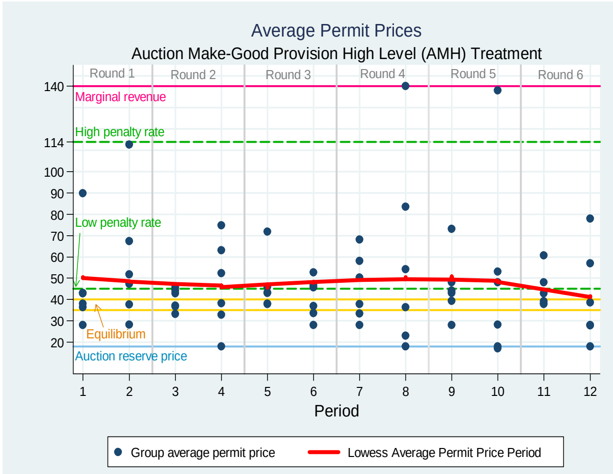
(a)



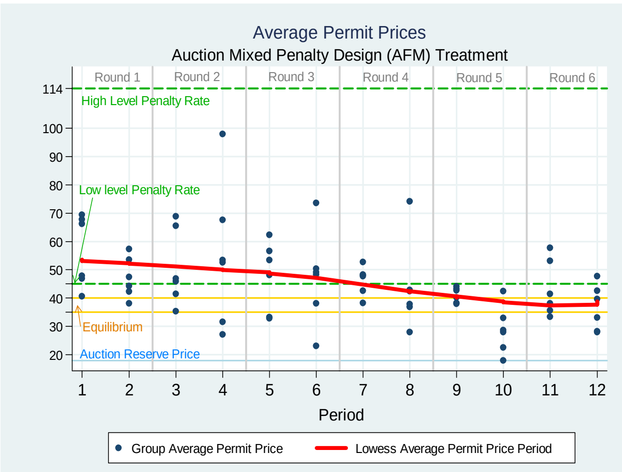
(b)



(c)



(d)



(e)

Note:

Each round consists of two sub-periods.

Period with odd numbers represent sub-period one of a round, while the even numbers represent sub-period two of a round.

Figure 4 Convergence Path of Average Permit Price over Time

5.2.2. Compliance Strategy and Compliance Rate

We examine the main effects of penalty design on environmental effectiveness using two variables: investment level and compliance rate. The variables are expressed in terms of the number of firms and compared to the optimal level. At the equilibrium, there should be four investing firms – and thus the other four permit buying firms - and eight compliant firms, yielding the optimal scale of unity (Figure 5).

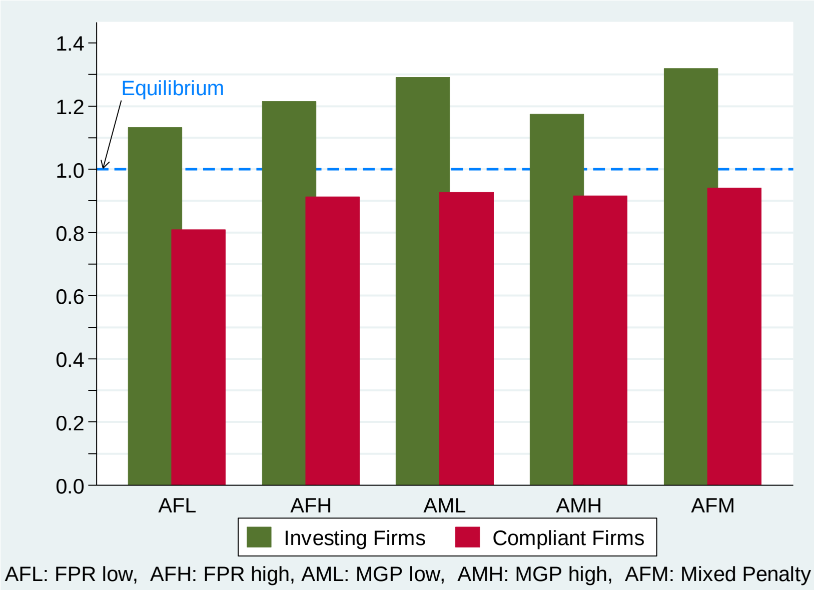


Figure 5 Investment and Compliance Levels by Treatment

As with the standard deviations of trading prices, the investment level is also higher with high level FPR treatments, but this is not the case with MGP treatments, which behave in the opposite manner. Over-investment is observed in all treatments as the mean investment level is higher than the optimal level. It seems plausible that a high penalty level will lead to a higher investment level because the mixed penalty design, which involves double penalties, exhibits the highest investment level. Nevertheless this inference will only be valid after a regression model is performed to control for influencing factors.

Interestingly, the observed over-investment is not translated into full compliance (where 100% of firms are compliant), although the trend in the investment level is also shown in the number of compliant firms. A high penalty level encourages a higher number of compliant firms in the FPR treatments, and the mixed penalty displays the highest compliance level.

Logically, if over-investment is prevalent, then there will be less of a demand for permits, which will consequently temper permit prices and render cheaper compliance costs for permit-buying firms.

Nevertheless, this is not observed in the experiment, as prices remain high regardless of over-investment. This result highlights the inefficiency at the auction market as well as in the secondary market, which fail to distribute permits to subjects who require them. The fact that full compliance is not realised in the market, despite over-investment, also means that some investing firms still hold excess permits at the end of a sub-period and put buyers in a short position. Considering that the trade volume on the secondary market is pretty low, which imply that subjects had enough time to trade but only realised a few trades, this inefficiency is more likely to be attributed to the auction market. Nevertheless, it is plausible that permit buying firms at the auction were attempting to achieve gains from arbitrage trading, but mostly were unsuccessful. The competitive nature of clock auctions (Compte and Jehiel, 2007) can push prices to a fairly high level, which makes it more difficult for firms to make a profit in the spot market.

Comparing the processes that firms use to choose the best compliance strategy in the FPR and MGP treatments, we conjecture that the "quantity-penalty" nature of the MGP treatments creates much more price uncertainty for firms and more deterrent effects, resulting in more volatility in the market. Over time, this volatility makes it more difficult for firms to arrive at the best compliance strategies because price signals are more scattered. As shown in the standard deviation statistics (Table 3), although the low MGP treatment and the mixed penalty provide very strong compliance incentives, they also make it more difficult for firms to make investment decisions, as revealed by the high standard deviations of the investment level figures.

Table 3 Statistics Summary for Investment Levels and Compliance Rates

Treatment	**Mean investment level**[a]	**S.d.**[b] **investment level**	**Compliance rate**[a]	**S.d.**[b] **compliance rate**
FPR Low (AFL)	1.130	0.2481	0.810	0.1512
FPR High (AFH)	1.215	0.2232	0.913	0.1020
MGP Low (AML)	1.292	0.3168	0.927	0.0956
MGP High (AMH)	1.174	0.2125	0.917	0.1168
Mixed Penalty (AFM)	1.319	0.2560	0.941	0.0938
Optimum	1.000	0	1.000	0

Notes: [a] compared to the optimal level, [b] s.d. = standard deviation

In the theoretical equilibrium, firms with low marginal abatement costs (low MAC firms) should choose investment decisions as their sole compliance strategy, whereas those with high marginal abatement costs (high MAC firms) should not invest and should just buy permits, using this as their best compliance strategy. However, when the auction price is higher than the equilibrium, firm decisions about the best compliance strategies will not be as easy as expected, especially for the high MAC firms.

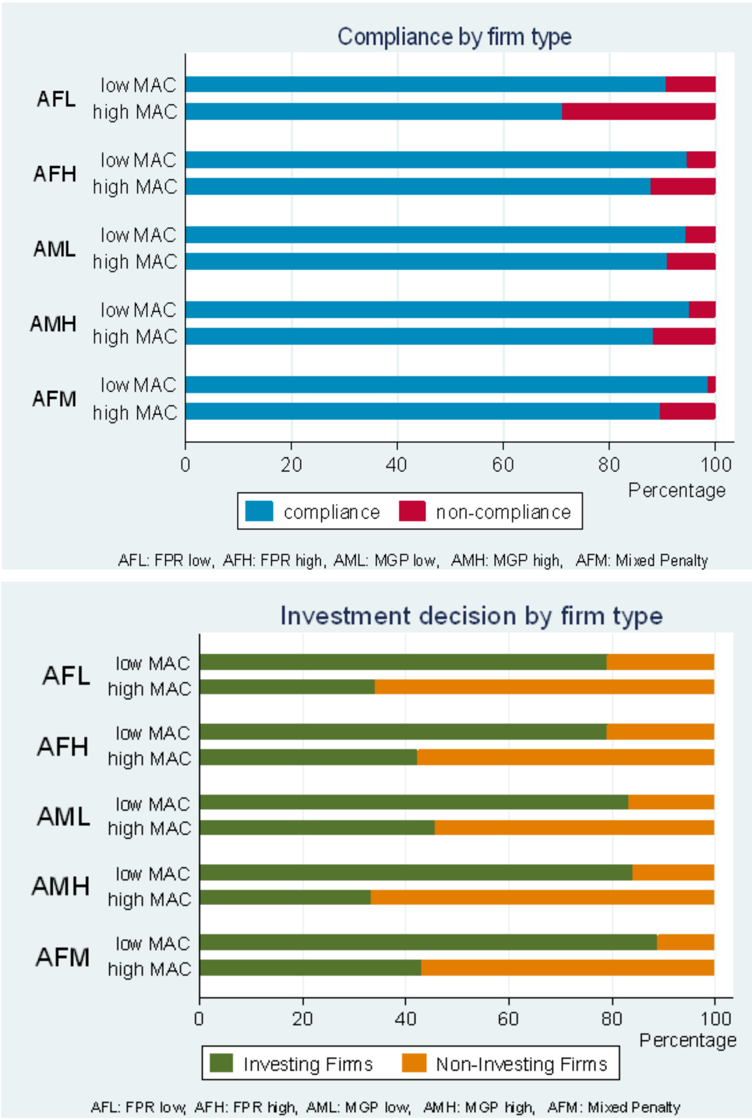
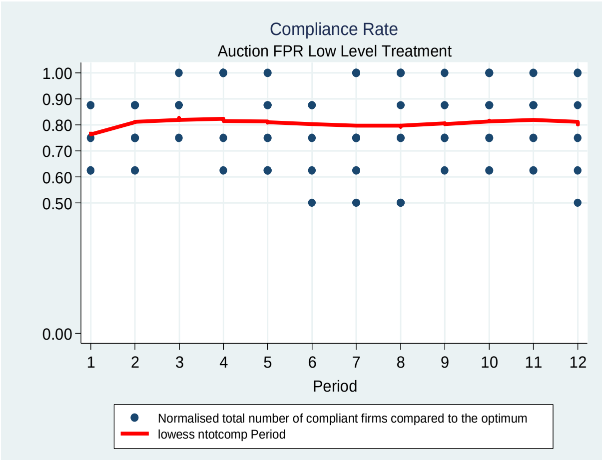


Figure 6 Compliance Rates and Investment Decisions by Firm Type

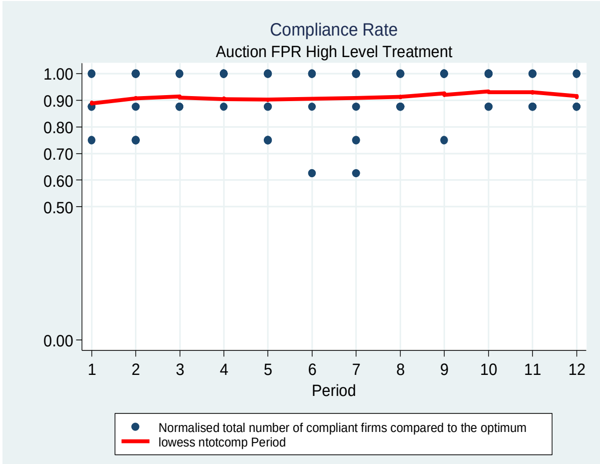
The data confirm that across treatments, firms with low MAC have higher compliance rates than do high MAC firms (Figure 6). Hence, buying permits is not always perceived as the best compliance strategy for high MAC firms, because prices fluctuate at a higher level than the theoretical equilibrium permit price. Although the low MAC firms do not always choose investment as their best compliance strategy, the compliance rate for these firms is still higher than that of the high MAC firms. The non-parametric Kruskal Wallis test statistics verify that the differences in compliance rates and investment decisions by firm type are highly significant (p-value = 0.0001). For both firm types, the mixed penalty (AFM) provides the strongest compliance incentives. In terms of penalty type, MGP treatments also provide better compliance incentives than do FPR treatments, and yet they also induce stronger investment incentives for high MAC firms, which explain the incidence of over-investment in general.

When we consider the learning effect of compliance decisions, we note that the convergence path for compliance rates over time does not show a very strong effect compared to that of the average permit price (Figure 7). The *lowess* regression curves seem to have fairly stable patterns, although the compliance rates vary across treatments. We have six observations for each treatment in each period, and fewer scatter points imply that there are some repeated values for compliance rates.

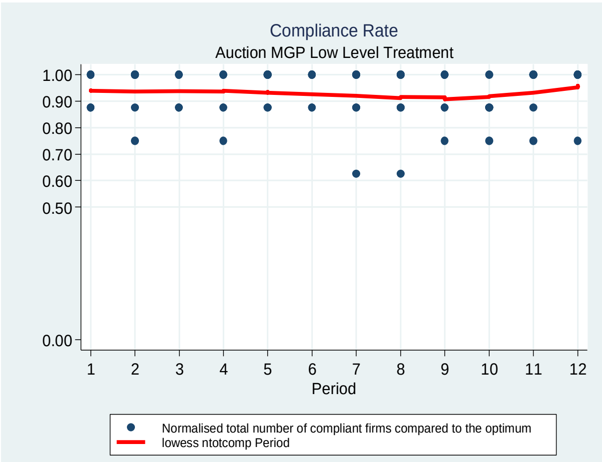
The low FPR treatment exhibits a fairly low compliance rate of 75% in the first round before stabilising around the 80% level in later rounds. Nevertheless, the scatter points are more dispersed in the low FPR, implying a higher standard deviation of compliance rates. Higher levels of compliance rates are clearly noticeable in both MGP treatments, in which they stay above the 90% level throughout all rounds, although more variance in the lower rates is also apparent in the high MGP treatment. A slightly different convergence path for compliance rates occurs with the mixed penalty treatment, where compliance rates begin very high at around 95% before decreasing to around 90% in the first half of the session and then stabilising back at the 95% level. The lowest standard deviation of compliance rates is achieved under the mixed penalty regime.



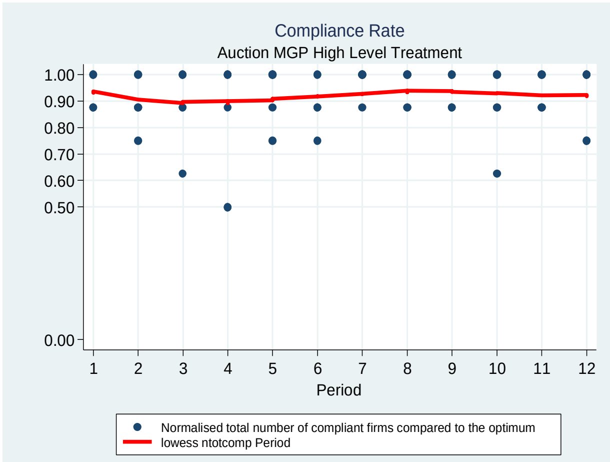
(a)



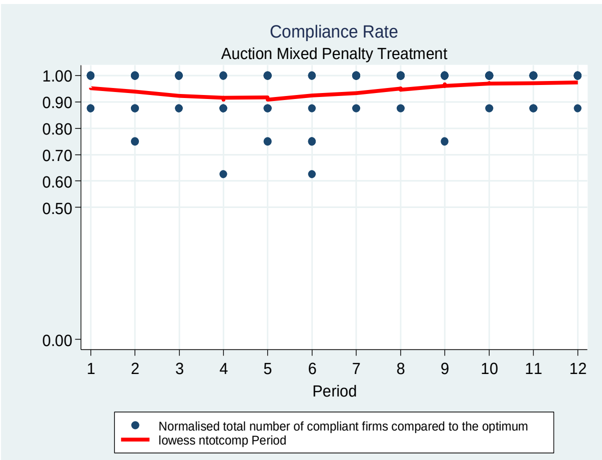
(b)



(c)



(d)



(e)

Figure 7 Convergence Path of Compliance Rates over Time

5.2.3. Efficiency

At the end of each sub-period, efficiency is the variable that sums up the measure of market performance under non-compliance. As permit prices rise, firms need to incur higher costs to buy permits, and hence, the efficiency of the market is compromised. In this case, static efficiency is measured in terms of the actual group earnings compared to the theoretical optimum. This efficiency measure is chosen rather than the usual cost savings measure to obtain normalised values for efficiency since higher prices will yield some negative values for efficiency in terms of cost savings. The data show that a low penalty level results in greater efficiency and that the FPR treatment performs better than the MGP and the Mixed Penalty treatments in this regard (Figure 8).

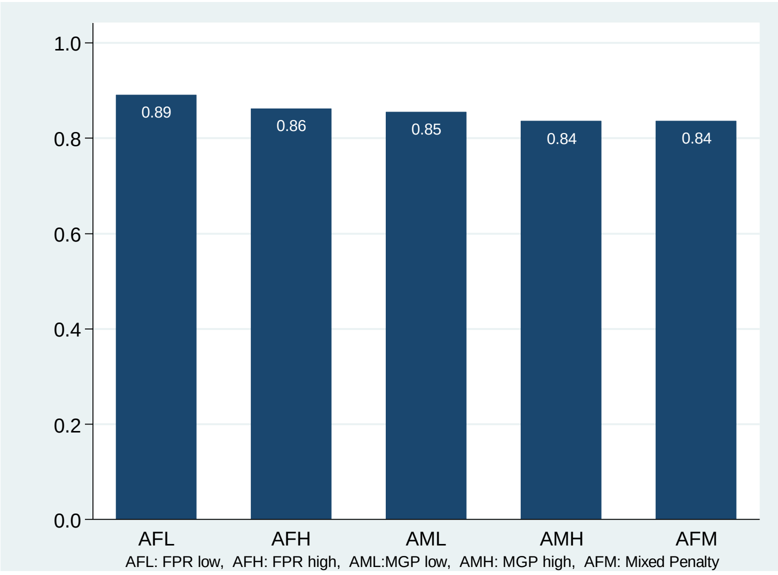


Figure 8 Efficiency by Treatment

6. Test of Treatment Effects

6.1. Test of Significant Difference to the Theoretical Equilibrium

The statistical summaries of prices and compliance strategies illustrate how the mean values generally deviate from the theoretical optimal values. In spite of the observable differences, further tests are required to confirm that these differences are statistically significant. Therefore this section looks at how the hypotheses are tested using the non-parametric Wilcoxon signed rank test. This test is performed to evaluate auction prices, investment levels, and compliance rates. It should be noted that the Wilcoxon signed rank test evaluates whether the medians (rather than the means) of those variables are significantly different from the hypothesised values.

Table 4 Test of Auction Price Equal to Theoretical Equilibrium

Treatment	p-value from Wilcoxon Sign Rank test for H_0: Auction price= 38		
	All periods	Period 1-6 (Round 1-3)	Period 7-12 (Round 4-6)
FPR Low (AFL)	0.0000***	0.0000***	0.1144
FPR High (AFH)	0.0000***	0.0000***	0.0003***
MGP Low (AML)	0.4452	0.1770	0.7582
MGP High (AMH)	0.0031***	0.0065**	0.1049
Mixed Penalty (AFM)	0.0000***	0.0000***	0.2722

Notes: * significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 72 for the whole session and 36 for half sessions.

The Wilcoxon test statistics reveal that the medians of the auction prices are significantly different from the optimal level at EX$38 for all treatments except the low level MGP treatment. The chosen optimal value is not only a mid-value within the optimal range (EX$ 35-40) but also the theoretical optimal auction price given by our experimental design.

In order to assess whether learning effect may affect the results, the data set is split into two half-sessions (round 1-3 and round 4-6). Similar results as for the whole data set are achieved for the first half of the session. In the second half of the session, the significant difference to the optimal equilibrium is only maintained for the high level FPR treatment. This result is in line with the convergence path for average permit price that illustrates how the prices start to enter the equilibrium range in round three or round four and remain there until an end-game effect takes place in the last round in the case of MGP treatments. Thus, there is a general indication of learning effect over rounds as subjects learn to arrive at the equilibrium auction price in the second half of the session.

The same test is also performed to assess whether the investment and compliance rates are significantly different from the optimal value of unity. Highly significant test statistics are obtained for all treatments.

Table 5 Test of Investment Level Equal to Theoretical Equilibrium

Treatment	p-value from Wilcoxon Sign Rank test for H_0: normalised investment level = 1		
	All periods	**Period 1-6 (Round 1-3)**	**Period 7-12 (Round 4-6)**
FPR Low (AFL)	0.0001***	0.0041**	0.0092**
FPR High (AFH)	0.0000***	0.0000***	0.0000***
MGP Low (AML)	0.0000***	0.0000***	0.0000***
MGP High (AMH)	0.0000***	0.0003**	0.0000***
Mixed Penalty (AFM)	0.0000***	0.0000***	0.0000***

Notes: * significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 72 for the whole session and 36 for half sessions.

Table 6 Test of Compliance Rate Equal to Theoretical Equilibrium

Treatment	p-value from Wilcoxon Sign Rank test for *Ho*: normalised compliance rate = 1		
	All periods	**Period 1-6 (Round 1-3)**	**Period 7-12 (Round 4-6)**
FPR Low (AFL)	0.0000***	0.0000***	0.0000***
FPR High (AFH)	0.0000***	0.0000***	0.0000***
MGP Low (AML)	0.0000***	0.0001***	0.0000***
MGP High (AMH)	0.0000***	0.0000***	0.0002***
Mixed Penalty (AFM)	0.0000***	0.0000***	0.0082**

Notes: * significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 72 for the whole session and 36 for half sessions.

One major difference between those two variables is the direction of the deviation from the optimal equilibrium. While the medians investment levels lie above the theoretical equilibrium, the median compliance rates mostly stay below the optimal perfect compliance rate.

6.2. Hypothesis Testing of Treatment Effects

6.2.1. The Effect of Penalty Design on Auction Price

Result 1: There are no differences in auction prices across treatments (consistent with Hypothesis 1) [8]*. Furthermore, auction prices remain above the optimal equilibrium level in earlier rounds but then converged to the equilibrium range in later periods.*

Support: The Kruskal-Wallis non-parametric test is used to test whether all five treatments have the same underlying distribution in terms of auction price. Each observation group is assumed to be independent, and no further assumptions are made with regard to the distribution of the data. Group-level data is collected for each round, and hence, the test is run over 360 observations. Since the test yields a p-value of 0.1537, we cannot reject the null hypothesis that auction prices come from the same underlying population distribution.

The Kruskal-Wallis test is the analog of the ANOVA test that is based on a normal distribution. These tests cannot determine whether only one or some of the samples exhibit a distribution that is significantly different from those of the rest of the samples. To answer that question, another test should be performed using a pairwise comparison. We employ the Wilcoxon rank sum test (Wilcoxon-Mann-Whitney test) and the Kolmogorov-Smirnov test to inspect the presence of treatment effects. Whereas the Wilcoxon-Mann-Whitney test evaluates whether the medians of two samples represent two populations with different median values, the Kolmogorov-Smirnov test looks at the differences between the underlying population distributions of the two samples (Sheskin, 2004).

In most cases, consistent results are obtained from the two tests (Table 7). Whereas the Kolmogorov-Smirnov test only confirms significant difference between median auction price in the low FPR and the low MGP treatment, more significant test results are obtained from the Wilcoxon-Mann-Whitney test. Significant test statistics at the 5% level verify lower medians of auction price in the MGP than the FPR treatment for both penalty levels, especially in earlier rounds. The median auction price is also lower in the low MGP compared to the Mixed Penalty treatment. However, this lower median auction price in the MGP treatment is achieved through very volatile markets, as the standard deviation of auction price is the highest in the MGP treatment.

[8] *Hypothesis 1* states that the auction price should remain the same in all treatments because the supply and demand structure remains the same.

Table 7 Test Statistics for Treatment Effects for Auction Price

Pairwise comparison	p-value All Periods		p-value by half-session			
			Period 1-6 (Round 1-3)		Period 7-12 (Round 4-6)	
	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]
Penalty level in FPR (AFL & AFH)	0.1213	0.419	0.4841	0.965	0.0646	0.257
Penalty level in MGP (AML & AMH)	0.1706	0.213	0.5824	0.413	0.2045	0.615
Penalty type, low-level penalty (AFL & AML)	0.0413*	0.014**	0.0634	0.257	0.2468	0.083
Penalty type, high-level penalty (AFH & AMH)	0.0530*	0.062	0.0339*	0.083	0.4889	0.257
Mixed penalty and low FPR (AFM & AFL)	0.9085	0.992	0.8157	0.825	0.7795	1.000
Mixed penalty and low MGP (AFM & AML)	0.0483*	0.062	0.0376*	0.150	0.3816	0.257

Notes: [a] KS = Kolmogorov-Smirnov test
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 144 for the whole session and 72 for half sessions.

Separate tests on the half session data unveils similar inferences as the whole session data: we cannot reject the null hypothesis of no differences in auction prices, except for the comparison of penalty type in the high level penalty and the comparison of the mixed penalty treatment to the low MGP treatment for the first half of the session.

6.2.2. The Effect of Penalty Level in the Fixed Penalty Rate Treatment

Result 2: There are differences between compliance rates but not between investment levels for the low and high level penalty in the FPR treatment. The compliance rate is statistically higher in the high level penalty treatments (which is inconsistent with Hypothesis 2)[9].

Support: A test of treatment effects is required only for two groups of independent samples. The Wilcoxon rank sum test and the Kolmogorov-Smirnov test are conducted to assess whether the statistics from the two samples are statistically different. The test statistics (Table 8) show that there is no significant difference between the investment levels at the low and the high level penalty and consistent results are obtained from both non-parametric tests. On the contrary, the test statistics for the compliance rates are highly significant for the whole session and the half-sessions.

[9] *Hypothesis 2* states that in the FPR treatments, investment levels and compliance rates should be the same at 100% regardless of the penalty levels because the penalty rate is set higher than the theoretical equilibrium permit price.

Table 8 Test Statistics for Treatment Effects of Penalty Levels in the FPR Treatment

Variable	p-value All Periods		p-value by half-session			
			Period 1-6 (Round 1-3)		Period 7-12 (Round 4-6)	
	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]
Investment level	0.0658	0.419	0.3471	0.615	0.0899	0.615
Compliance rate	0.0000***	0.001***	0.0032**	0.043*	0.0017**	0.005**

Notes: [a] KS = Kolmogorov-Smirnov test
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 144 for the whole session and 72 for half sessions.

6.2.3. The Effect of Penalty Level in the Make-Good Provision Treatment

Result 3: Penalty level does not significantly affect either investment levels or compliance rates in the MGP treatment (consistent with Hypothesis 3)[10].

Support: The Kolmogorov-Smirnov test shows that we cannot reject the null hypothesis that the two samples come from populations with the same underlying distribution. Nevertheless, the Wilcoxon test results show that the median investment level is just significantly different at 5% level. The results from the half session's data confirms that the difference is only significant for earlier rounds. Hence, the test statistics from the Kolmogorov-Smirnov show more conservative estimates than those from the Wilcoxon test, and we conclude that, in general, no significant differences are found in both investment levels and compliance rates.

Table 9 Test Statistics for the Effect of Penalty Levels in the MGP treatment

Variable	p-value All Periods		p-value by half-session			
			Period 1-6 (Round 1-3)		Period 7-12 (Round 4-6)	
	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]
Investment level	0.0513*	0.419	0.0400*	0.083	0.3731	0.615
Compliance rate	0.7987	1.000	0.2051	0.965	0.3706	1.000

Notes: [a] KS = Kolmogorov-Smirnov test
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 144 for the whole session and 72 for half sessions.

[10] *Hypothesis 3* states that the make-good ratio should not affect investment levels or compliance rates in the MGP treatments as long as it is set equal to or higher than one under the assumption that prices remain the same in the two sub-periods.

6.2.4. The Effect of Penalty Type

Result 4: At the high penalty level, different penalty types do not provide different compliance incentives because there are no significant differences in terms of investment levels and compliance rates between the FPR and the MGP (which is consistent with Hypothesis 4) [11]*. On the other hand, different compliance rates are observed at the low penalty level in which the MGP treatments have higher compliance rates than the FPR treatments (which is inconsistent with Hypothesis 4). However, the same distinction does not exist between investment levels.*

Support: As shown by Table 10, we obtain consistent test statistics for the high level penalty treatments. We do not find enough evidence to reject the null hypothesis that the two samples are derived from the same population distribution. Meanwhile, the low level penalty treatment constantly demonstrates that compliance rates are significantly higher in the MGP penalty than in the FPR penalty for the whole session and the half-session data. Similar results are obtained for the median of investment levels, particularly in the first half-session as the subjects still learn to decide on their optimal compliance strategy.

Table 10 Test Statistics for Treatment Effects by Penalty Type

Variable	p-value All Periods		p-value by half session			
			Period 1-6 (Round 1-3)		Period 7-12 (Round 4-6)	
	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]	Wilcoxon rank-sum	KS[a]
Low-level penalty						
Investment level	0.0070**	0.213	0.0255*	0.022*	0.0739	0.615
Compliance rate	0.0000***	0.000***	0.0000***	0.002**	0.0041**	0.043*
High-level penalty						
Investment level	0.3116	1.000	0.4039	0.615	0.5217	0.615
Compliance rate	0.5571	1.000	0.9708	1.000	0.4004	0.965

Notes: [a] KS = Kolmogorov-Smirnov test
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 144 for the whole session and 72 for half sessions.

[11] *Hypothesis 4* states that at the high penalty level, the compliance rates and investment levels in the FPR treatments should be the same as those in the MGP treatments. At the low penalty level, similar results are expected, although the penalty level is slightly higher in the low level FPR treatment (with a factor of 1.2), whereas the make-good ratio is 1.

6.2.5. The Effect of Double Penalty in the Mixed Penalty Design

Result 5: The mixed penalty design provides the same investment and compliance incentives as the low MGP treatment. However, the same conclusion cannot be drawn when the comparison is made to the low FPR treatment since significant differences are found with regard to investment levels and compliance rates (inconsistent with Hypothesis 5) [12].

Support: As the mixed penalty comprises of two penalty types with some modification, the evaluation of the effect of this penalty design is carried out using these elements. Hence, we assess the treatment effects by making a comparison with the low level MGP (AML) treatment and with the low level FPR (AFL) treatments. Unlike the AFL treatments, the mixed penalty design uses variable penalty rates over time because the penalty rate is linked to the auction price. As shown in Table 11, the test statistics are only significant when we draw a comparison with the AFL treatment. Once again, this result highlights the different compliance incentives provided by low level FPR and low level MGP.

Table 11 Test Statistics for Treatment Effects of the Mixed Penalty Design

Variable	p-value All Periods		p-value by half session			
			Period 1-6 (Round 1-3)		Period 7-12 (Round 4-6)	
	Wilcoxon rank-sum	Kolmogorov-Smirnov	Wilcoxon rank-sum	Kolmogorov-Smirnov	Wilcoxon rank-sum	Kolmogorov-Smirnov
Mixed Penalty and low FPR						
Investment level	0.0001***	0.014**	0.0106**	0.083	0.0020**	0.083
Compliance rate	0.0000***	0.000***	0.0013***	0.022*	0.0000***	0.000***
Mixed Penalty and low MGP						
Investment level	0.3040	0.947	0.9439	0.615	0.1499	0.965
Compliance rate	0.2715	0.847	0.2956	0.965	0.0087**	0.083

Notes: [a] KS = Kolmogorov-Smirnov test
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level
Number of observation is 144 for the whole session and 72 for half sessions.

7. Estimation models

The test of the treatment effects in the previous chapter cannot fully capture the relationship between our treatment variables and a particular variable of interest (the dependent variable) because the test procedure only takes into account the variation in one particular variable without holding other

[12] *Hypothesis 5* states that the mixed penalty treatment should yield the same compliance rates as the FPR and MGP treatments.

variables constant. Moreover, the subjects' characteristics might have some influence. In order to control for all of those factors and to isolate the treatment effects, we ran regression models to further examine the effects of our treatment variables on auction price, investment decisions, firm compliance through permit-buying, and efficiency.

Auction price is chosen as the first dependent variable, because it is the first price signal that subjects receive before deciding on their compliance strategy through investment decisions and permit-buying in the spot market. Subsequently, investment decisions and firm compliance status of the non-investing firms (permit-buying firms) are also important dependent variables that represent the two available compliance strategies. Hence, regressions are performed on those dependent variables so that the treatment effects of penalty type and penalty level can be carefully verified. Finally, efficiency is a crucial measure of market performance, as it is a major criterion of the economic success of an emissions trading scheme.

7.1. Auction Price

Considering that the auction is the first stage in each sub-period in the emissions trading game, we can only include treatment variables and subjects' characteristics in each observation group as dependent variables. The estimation is performed using group-level data collected in each sub-period with a total of 360 observations. The regression model for auction price is estimated using a robust panel data random-effects model.

Model 1 represents the basic model that contains the treatment variables for penalty design as the main regressors, i.e. a dummy for the FPR treatment, a dummy for the high-level FPR treatment, a dummy for the MGP treatment, and a dummy for the high-level MGP treatment. In the mixed penalty design, the dummy for high FPR is set at zero (the low level), although the penalty rate is actually varied. This measure is taken as the penalty rate is directly linked to the auction price; hence, it is not independent of the auction price. Including the penalty rate as a regressor would produce bias estimates toward higher significance. In view of the complexity of our experiment, round and sub-period two are also included as regressors and used to examine subjects' learning curves over time.

In Model 2, the effects of risk-related variables are taken into account. The group risk preference index represents the aggregated value of each subject's Holt & Laury's (2002) risk preference index for each observation group. We also employ the same aggregation approach for the variable inconsistent risk preference choice.

Model 3 incorporates additional control variables related to subjects' income variables that include age, household income, number of household members, and individual income. Since most of the students are not financially independent, we control for the effects of these income variables on subjects' risk preferences. These demographic variables are measured in terms of the mean values of interval or ordinal variables.

An additional set of demographic variables pertaining to gender and education are added to Model 4. These variables are the number of females in a group, study degree (e.g. undergraduate, Master's, or PhD program), a dummy for majors related to economics, number of years in school, and full-time enrolment status. At the individual level, study program is a categorical variable that indicates whether a subject is undertaking an undergraduate, master's, or doctoral program. For group-level data, a mean value is used for each group. The same approach is used for the other demographic explanatory variables.

The estimate summary (Table 12) shows that the signs of the explanatory variables are intuitive and consistent across all models. The coefficients of the MGP treatment are always much smaller than those of the FPR treatment, although neither are significantly different than zero. On the contrary, the coefficients of the high level MGP are about 2 to 3 times larger than those of the high level FPR. The coefficients for penalty design are not statistically significant except in Model 3 for the high MGP treatment. Thus, we can conclude that the quantity-penalty nature of the high level MGP evidently raises the demand for permits vis a vis other penalty designs. Significant constant terms are also evident across models. The summary statistics also display the values of rho and theta to represent the influence of group-specific effects as an inherent term in a random-effects model.

The variable round has a negative sign, indicating that the learning curve has a negative effect on auction prices. This implies that at the beginning of the experiment, permits were in very high demand, possibly due to cautious behaviour (the subjects wanted to ensure that they would achieve compliance) or due to the required time for subjects to learn about the equilibrium permit price. The coefficient of round is also statistically and economically significant. Unsurprisingly, the auction price is slightly lower in sub-period 2 because over-investment in sub-period 1 effectively reduces permit demand. However, the coefficient is not statistically significant. In terms of the goodness of fit of the model, the overall correlation value is relatively small, and the model performs better in explaining the variation between observation groups than the variation within the same group.

Based on the risk-related variables in Model 2, we see that subjects that make inconsistent choices during the Holt & Laury experiment may engage in irrational bidding behaviour that increases the demand for permits at the auction. The magnitude of the coefficient is also economically significant because the presence of only one irrational subject will increase the auction price by more than EX$2. The group risk preference index also shows the expected relationship: greater risk aversion is associated with more cautious bidding behaviour, which then results in a lower auction price – albeit not with statistical significance. According to the R^2 value, the goodness of fit is slightly greater than in Model 1, and the between correlation is triple enhanced.

Table 12 Estimates Summary for Auction Price Model

Regressors for auction price	Model 1 (basic)	Model 2 (Model 1 + Risk)	Model 3 (Model 2 + Income)	Model 4 (Model 3 + gender and study)
Dummy for FPR	2.9861	4.3953	5.0095	3.4850
	(2.6218)	(3.0792)	(3.2187)	(3.5141)
Dummy for high level FPR	3.1944	2.3382	2.8419	3.3898
	(3.7962)	(3.0537)	(3.6890)	(3.5971)
Dummy for MGP	0.5556	1.1085	0.6905	3.4868
	(3.5854)	(3.4998)	(3.5133)	(3.1911)
Dummy for high level MGP	5.6944	6.7694	8.1732*	5.2603
	(4.7587)	(4.2066)	(3.9863)	(3.9259)
Round	-2.4024***	-2.4024***	-2.4024***	-2.4024***
	(0.6777)	(0.6796)	(0.6835)	(0.6885)
Dummy for sub-period two	-0.3611	-0.3611	-0.3611	-0.3611
	(1.7946)	(1.7997)	(1.8101)	(1.8232)
Group risk preference index		-0.3280	-0.5365*	-0.6538**
		(0.2126)	(0.2198)	(0.2501)
Number of subjects with inconsistent risk choices		2.5873*	2.7198**	2.1594*
		(1.1067)	(0.9997)	(0.9004)
Constant	50.6167***	59.2014***	72.1731***	115.0062*
	(4.5917)	(10.9804)	(15.9237)	(48.2215)
Observation	360	360	360	360
Within correlation	0.0580	0.0580	0.0580	0.0580
Between correlation	0.0897	0.2752	0.3570	0.4581
Overall correlation	0.0627	0.0904	0.1026	0.1177
Chi^2	15.4591	35.8100	47.9306	104.2794
Rho (% variance due to group-specific effect)	0.0926	0.0716	0.0816	0.1021
Theta[13]	0.3294	0.2792	0.3044	0.3496

Notes: The numbers in parentheses represent the standard errors of the estimates.
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level

Although the estimations in Model 3 seem to explain better compared to the other models, the constant term of the model also increases by 20%. In addition to the previously significant variables,

[13] Theta =[0,1]. When theta equals zero, the regression can be performed directly using OLS estimators. If theta is one, then fixed-effect estimators are more appropriate.

the group risk preference index and the dummy for high level MGP also become significant. The coefficients of the latter variables have intensified as much as the constant term. Although none of the additional income variables have significant test statistics, the control of those risk-related variables noticeably improves the model. The signs of the income variables are also as expected. Higher group age and individual income are negatively correlated with auction price, indicating more sensible bidding behaviour that drives down auction price. In contrast, higher household income and more household members contribute to higher auction prices. Nevertheless, these relationships are statistically not different from zero. With regard to the goodness of fit as measured using the R^2 value, this model yields a slightly higher overall correlation than Model 2.

The inclusion of more demographic variables in Model 4 has slightly enhanced the overall correlation of the model at the expense of the escalated constant term, which is at about 60% of that in Model 3. As in the previous model, neither of the demographic variables is statistically significant. The significant independent variables remain the same except for the dummy for the high MGP treatment.

Overall, the regression models show that the penalty design does not significantly affect the price discovery process in the market with respect to auction price. The only exception is one model in which the high MGP treatment contributes significantly to a much higher auction price. These results are essentially in line with the test statistics for treatment effect for *Hypothesis 1*, in which the Wilcoxon rank sum test verifies the effect of the high MGP treatment, especially in the first half of the session. More importantly, the risk-related variables have the largest marginal effect on auction price. Higher risk aversion moderates speculative bidding behaviour and drives down auction price closer to the equilibrium. On the contrary, subjects with irrational risk choices can inflate auction price. The variable round confirms the presence of learning effects because it remains statistically significant across models.

7.2. Investment Decision

After participating at the auction, subjects can trade in the secondary market before deciding to make an investment in abatement technology in sub-period one. At the investment decision stage, each subject is given information about his or her final permit holdings for that sub-period and whether they hold an excess or shortfall of permits for compliance. During this stage, if firms have a shortfall of permits, then the only way to achieve compliance and avoid a penalty is by making an investment decision.

In view of the decision process, the penalty design treatment variables, price variable, firm type, and firm long permit position are used to regress individual investment decisions. Unlike in the estimation models for auction price, the penalty rate is used rather than a dummy variable for the high FPR because the penalty rate is now independent of the investment decision, which is not the case for the auction price estimation. The individual-level data from sub-period one are used in the estimation. We employ panel data probit and logit estimation models because the investment decision is a binary choice. The regression summary is shown in Table 13.

The first four models are estimated using a probit model, and the results are consistent across all four. Model 1 is run using cluster-robust standard error OLS estimators. Models 2 to 4 are based on a robust random-effects probit model with bootstrapped estimates. A logit model similar to Model 4 is run in Model 5. Due to the nature of binary choice models, the interpretation of the estimation results is not straightforward in terms of magnitude. Nevertheless, the sign of the estimates indicates the effect of the regressors on the dependant variable.

The estimates reveal that different penalty types provide different incentives with regard to investment decisions. The FPR treatment has a negative but trivial effect on investment. In contrast, the MGP treatment has a significantly positive effect on investment. These findings are very reasonable, because firms might have a lower marginal financial penalty in the FPR treatment than in the highly punitive MGP treatment. The effect of the penalty rate is almost negligible, but positive as expected. On the contrary, a high MGP level has a negative effect on investment decisions, even though it is not statistically significant. However, the regression models can better explain the effect of the MGP treatment because they control for other factors that might influence investment decisions.

According to the experimental design, high MAC firms should not invest and choose to comply by buying permits. The regression models confirm the theory as this variable has a negative sign and is highly significant. A similar effect is also produced by firm permit position. When firms learn that they have an excess of permits, they do not invest. In this sense, the firms show rational investment behaviour.

The auction price has a crucial effect on firm investment behaviour. Conversely, trading prices have essentially no effect on investment decisions. This proves that the main price signal for compliance strategy is determined at the auction as the primary market rather than in the secondary market.

We find that learning does not have an effect on investment decisions as indicated by the estimate for round. The additional risk-related variables in Model 4 and Model 5 are also not significant. Other demographic variables such as the subject’s gender and age are also tested as regressors, but the results show that they are insignificant.

Table 13 Estimates Summary for Investment Decisions

Regressor for investment decisions	Model 1 Probit OLS cluster	Model 2 Probit RE bootstrap	Model 3 Probit RE bootstrap	Model 4 Probit RE bootstrap	Model 5 Logit RE bootstrap
Dummy for FPR	-0.045	-0.0746	-0.0713	-0.0515	-0.0534
	(0.2573)	(0.2573)	(0.2579)	(0.2767)	(0.5008)
Penalty rate	0.0023	0.0031	0.0032	0.003	0.0064
	(0.0026)	(0.0029)	(0.0029)	(0.0031)	(0.0056)
Dummy for MGP	0.5013*	0.5857**	0.5871**	0.5832**	1.0922**
	(0.197)	(0.2037)	(0.2033)	-0.1949	(0.3596)
Dummy for high-level	-0.3369	-0.3787	-0.3755	-0.3455	-0.5245
MGP	(0.1775)	(0.2152)	(0.2137)	(0.1765)	(0.34)
High MAC firm	-0.8266***	-0.9084***	-0.9067***	-0.8914***	-1.6401***
	(0.097)	(0.1296)	(0.1316)	(0.1347)	(0.2509)
Auction price	0.0121***	0.0142***	0.0132***	0.0138***	0.0247***
	(0.0034)	(0.0032)	(0.0033)	(0.0036)	(0.0063)
Mean trading price	0.0000	-0.0002	-0.0002	0.0000	0.0000
	(0.0014)	(0.0019)	(0.0019)	(0.0019)	(0.0036)
Permit surplus	-0.1191***	-0.1393***	-0.1394***	-0.1406***	-0.2623***
	(0.008)	(0.0113)	(0.0114)	(0.0102)	(0.0194)
Round			-0.0179		
			(0.0396)		
Group risk preference				0.0065	
index				(0.0467)	
Subjects with				0.3338	
inconsistent risk choices				(0.1798)	
_cons	-1.0329***	-1.2810***	-1.1820***	-1.3977***	-2.5122***
	(0.3073)	(0.2813)	(0.3538)	(0.3478)	(0.5691)
No. obs.	1440	1440	1440	1440	1440
No. subjects	240	240	240	240	240
Log likelihood	-448.63	-431.01	-430.859	-429.065	-422.93
R^2	0.5433				
Wald chi^2	303.1957	227.3476	221.2588	285.8775	229.7005
% Correctly predicted	88.75				

Notes: The numbers in parentheses represent the standard errors of the estimates.
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level

The statistics of the models show that the addition of more explanatory variables does not substantially increase the goodness of fit of the model. Overall, the models have considerably high predictive power (88.75%), as shown by the basic model (Model 1).

To sum up, the investment decision is influenced by price signals at the auction, firm type (high or low MAC firm), and permit holding position toward compliance. Furthermore, MGP as a penalty type provides significantly stronger investment incentives compared to other penalty designs.

7.3. Compliance Decisions through Permit-Buying

Since an investment decision automatically ensures firm compliance, this section discusses estimation models for compliance only through permit-buying. Therefore, only observations associated with those subjects who do not make investment decisions are used in the regression. Considering that compliance status is a binary variable, the regressions are performed using probit and logit estimators for random effect panel data.

With regard to penalty design, the summary of estimates in Table 14 shows results similar to those of the investment decision model. Subjects tend to be more non-compliant in the FPR treatment than in the MGP treatment. Nevertheless, unlike in the investment model, the penalty rate provides a highly significant compliance incentive for permit buyers. This finding is in line with *Result 2*, in which the compliance rate is higher when the penalty level is higher in the FPR treatment.

The MGP treatment generates the highest marginal effect on compliance, and this effect is also highly significant. A higher make-good ratio also increases the likelihood of subject’s compliance although this effect is statistically not different to zero. The test statistics of the models validate *Result 3* in which the penalty level in the MGP treatment has no effect on compliance rates.

There is evidence of learning over time as the coefficient of round is statistically significant across models. The opposite effect is observed with the variable sub-period two. In line with the experimental design, the estimates reveal that subjects find it more difficult to be compliant only through permit-buying, particularly in sub-period two. This effect is undoubtedly true in MGP treatments in which even slight non-compliance by the end of sub-period one can put very high pressure on permit demand in sub-period two. Knowing that permit buyers will attempt to avoid penalties in the second sub-period, permit sellers have the advantage of selling the permit at a much higher price than the theoretical equilibrium. Nevertheless, the effect is statistically not different from zero.

Table 14 Estimates Summary for Compliance Decisions

Regressors for compliance decisions	Model 1 Probit OLS cluster robust	Model 2 Probit RE bootstrap	Model 3 Probit RE bootstrap	Model 4 Probit RE bootstrap	Model 5 Logit RE bootstrap
Dummy for FPR	-0.0872	-0.1416	-0.1397	-0.142	-0.2593
	(0.1653)	(0.1911)	(0.2206)	(0.2189)	(0.3500)
Penalty rate	0.0087***	0.0089**	0.0088***	0.0089***	0.0152***
	(0.0021)	(0.0028)	(0.0024)	(0.0025)	(0.0046)
Dummy for MGP	0.9548***	0.9796***	0.9776***	1.0025***	1.6834***
	(0.2019)	(0.2354)	(0.2383)	(0.2298)	(0.4696)
Dummy for high level	0.0779	0.1307	0.1306	0.1235	0.1954
MGP	(0.1801)	(0.1870)	(0.1796)	(0.2176)	(0.3814)
Round	0.051	0.0749*	0.0750*	0.0727*	0.1263*
	(0.0291)	(0.0334)	(0.0331)	(0.034)	(0.0514)
Auction Price	-0.0088***	-0.0103***	-0.0102***	-0.0086**	-0.0175***
	(0.0025)	(0.0028)	(0.0026)	(0.0029)	(0.0043)
Dummy for sub-			-0.0094		-0.0225
period two			(0.0762)		(0.1396)
Mean trading price				-0.0031	
				(0.0018)	
_cons	0.0802	0.1508	0.1559	0.1912	0.2811
	(0.2639)	(0.3028)	(0.2984)	(0.3093)	(0.5910)
N	1114	1114	1114		1114
Log likelihood	-592.4348	-572.8482	-572.8431	-570.8979	-572.347
R^2	0.0632	0.0461^	0.0461^	0.0493^	0.0456^
Chi2	41.7655	45.5528	62.4192	62.1237	60.0678
% correctly predicted	74.78				

Notes: The numbers in parentheses represent the standard errors of the estimates.
^ indicates estimated R^2 = (log likelihood – constant-only log likelihood) / constant-only log likelihood
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level

It is not surprising that auction prices generate a negative effect on compliance incentive, because higher permit prices increase compliance costs, causing the marginal benefit of being non-compliant to increase accordingly. This effect is highly significant, although it is much smaller in magnitude than the influence of the MGP treatment. The trading price also has a negative effect on compliance, but the magnitude of the effect is only half that of the auction price and is not statistically significant. It is important to point out that the coefficients on auction price and penalty rates counterbalance each other. Across models, we can see that their magnitudes are not very different. In all models except for Model 4, the marginal effect of the penalty rate is smaller than that of the auction price. In Model 4, the inclusion of the mean trading price moderates the marginal effect of the auction price.

The estimates are fairly consistent across models in terms of the sign of the coefficients. Statistically, the random effects binary choice models yield more consistent estimates than the OLS estimator. The OLS model shows that the model has fairly good predictive power with about 75% of the data correctly predicted. However, the value of R^2 is relatively small and slightly reduced in the random effect models.

Overall, the compliance decisions of net buyers are influenced by penalty designs (the level of penalty rate and the MGP penalty) and auction price. The adverse effect of a high auction price on the compliance decision surpasses the positive incentives given by penalty rates. The estimates explain the incidence of lower compliance rates in the FPR compared to the MGP and the mixed penalty treatment. Over time, subjects learn to make better compliance decisions.

7.4. Efficiency

Regression models for efficiency are performed using Tobit estimators because the range of the possible values is truncated. As previously explained, efficiency is measured in terms of the actual group earnings compared to the theoretical optimum. Interestingly, we find some observations in which the efficiency level is higher than one, due to the auction price being below the optimal equilibrium. Those observations are mostly associated with the MGP treatment. There are some reasons why this might have happened. First, a low auction price might occur due to over investment, which would naturally lower permit demand. Furthermore, some subjects who have decided to make investment simply do not actively participate at the auction by submitting zero bidding quantity even at a low bidding price because they feel that they have chosen their best compliance strategy and hence are not interested in the outcome of the auction market. Low prices also emerge in sub-period one of the MGP treatment, in which the financial penalty for non-compliance is zero. Thus, zero compliance cost does not provide an incentive for the subjects to actively participate in the auction. Therefore, we only use left censoring at zero in the estimation models. Group-level data is used to estimate the models because we would like to assess market-level efficiency instead of individual-level efficiency.

The estimation results are fairly similar across models (Table 15). The first model is performed with cluster-robust standard error estimators and with each observation group as the cluster identity. The penalty design treatment variable, price variables, and time variables are used as explanatory variables. Surprisingly, the dummy for the FPR treatment and the penalty rate have almost negligible effects on efficiency. As in the previous estimation models, the MGP treatment and auction price are highly significant in both economic and statistical terms.

Table 15 Estimates Summary for Efficiency

Regressor for efficiency	Model 1	Model 2	Model 3
	Tobit	Panel data Tobit	Panel data Tobit
Dummy for FPR	-0.0003	-0.0024	0.0094
	(0.0231)	(0.0297)	(0.0117)
Penalty rate	0.0000	0.0001	-0.0004**
	(0.0002)	(0.0002)	(0.0001)
Dummy for MGP	-0.0437**	-0.0395*	-0.0786***
	(0.0156)	(0.0184)	(0.0127)
Dummy for high-level MGP	0.0153	0.0154	-0.0058
	(0.0199)	(0.0232)	(0.0111)
Auction Price	-0.0059***	-0.0059***	-0.0055***
	(0.0004)	(0.0004)	(0.0002)
Mean trading price	-0.0003	-0.0003	-0.0001
	(0.0002)	(0.0003)	(0.0001)
Round	0.0062**	0.0061*	0.0003
	(0.0024)	(0.0025)	(0.0021)
Dummy for sub-period two	-0.0697***	-0.0690***	-0.0678***
	(0.0113)	(0.0103)	(0.0071)
Compliance rate			0.5168***
			(0.0373)
Investment level			-0.2020***
			(0.0166)
_cons	1.1733***	1.1709***	0.9885***
	(-0.0324)	(0.0313)	(0.0324)
N	360	360	360
Log likelihood	383.5838	385.8185	470.3238
Chi2	180.0935	492.9965	1445.322

Notes: The numbers in parentheses represent the standard errors of the estimates.
^ indicates estimated R^2 = (log likelihood – constant-only log likelihood) / constant-only log likelihood
* significant at 5% level, ** significant at 1% level, *** significant at 0.1% level

The existence of a learning effect is also confirmed because the coefficient of round is highly significant. This indicates that, over time, subjects learn to make better decisions in the game that contribute to higher efficiency in the market. The coefficient on sub-period two has a negative sign, indicating that efficiency in this sub-period tends to be lower than in sub-period one. The high financial penalty in the second sub-period for MGP treatment might be the underlying reason for this effect. Otherwise, it seems that investing firms attempt to make a profit selling in sub-period two by buying more permits at the auction. Nevertheless, this attempt does not seem to be successful; the mean trading prices are lower than the auction price 70% of the time. Hence, efficiency is reduced for both buyers who cannot obtain the required permits and sellers who cannot achieve their preferred trading price.

Model 2 is very similar to Model 1, except for the fact that the regression is run with a panel data Tobit estimator. The estimation results are not very different, but the goodness of fit of the model is increased as shown by the value of the log likelihood.

The inclusion of the investment level and compliance rates variables in Model 3 changes the model estimates considerably. The sign of the FPR treatment is now positive, although it is not statistically significant. The penalty rate, on the other hand, becomes highly significant. This is reasonable as higher penalty rates are related to higher costs when non-compliance occurs. Similarly, the effect of the MGP treatment is almost doubled in this model, while the effect of high MGP levels is not significant. It can be inferred that the MGP penalty type and the level of the penalty rate are the main determinants of efficiency.

The important finding that we obtain from this model is that the investment levels and compliance rates have the opposite effect on efficiency. Higher compliance levels in the market contribute to greater efficiency. On the contrary, higher investment levels will reduce efficiency, due to higher investment costs than necessary. As seen in Figure 6, it is also clear that some high MAC firms decide to make investment decisions that entail higher total investment costs than are optimal in the market. Furthermore, over-investment prevails in the market and exacerbates inefficiency.

8. Discussion

In line with the theory, our results show that different penalty designs do not necessarily indicate different permit prices. We find that auction prices do not significantly vary across treatments, as shown by *Result 1* and the regression models for auction price. The important determinants of auction price are the subject's risk attitude and rational thinking ability. This finding will have an important implication on the design of trading schemes in practice in which there is a tendency that auctions will increasingly be used as the initial allocation mechanism. Despite the advantages of having a system with auctioned permits, it is important to ensure that the extent of speculative or strategic bidding behaviour can be moderated with an appropriate auction design. Some studies have shown that reducing uncertainties regarding the expected permit price might be the key to address the problem.

Result 2 of this experiment has shown that, contrary to the theory that predicts that there will be no difference in compliance rates as long as the penalty level is set above the equilibrium permit price, significantly higher compliance levels emerge with the high level FPR. This study reveals that subjects learn about different maximum compliance costs related to high penalty rates, which in turn affects their compliance strategy. However, high penalty rates do not provide a different incentive with regard to investment decisions.

In contrast, *Result 3* confirms the theory that high penalty levels (make-good ratios) in the make-good provision penalty type do not produce different investment levels and compliance rates. It is believed that the feature of our experimental design ,which forces compliance by imposing huge financial penalty at the end of the second sub-period, affect the result. This provision leaves little room for subjects to be more speculative unlike in the FPR penalty type.

Result 4 indicates that when a comparison is drawn across penalty types with similar penalty levels, different compliance rates are only confirmed for low level penalty treatment. In this case, the MGP treatment provides a stronger compliance incentive than the FPR treatment even though the make-good ratio is slightly lower (a restoration rate of unity) than the penalty rate factor (1.2 of equilibrium price).

The statistics of the estimation models for compliance decision verify *Result 2* and *Result 4* as both the penalty rates and the MGP treatment are the significant penalty design variables.

With regard to the Mixed Penalty design, *Result 5* reveals that this double penalty encourages higher investment levels and compliance rates compared to the baseline low level FPR treatment. Nevertheless, the regression models prove that the MGP treatment is the only significant penalty design variable that affects investment decisions. The models also verify that subjects behave rationally in making their investment decisions.

A trade-off between efficiency and compliance is revealed as the regression models show that the MGP treatment has both statistically and economically significant effects on efficiency. Hence, the penalty design that encourages higher compliance levels might encourage lower efficiency levels when over-investment occurs in the market. As auction price plays a significant role in investment decisions, the presence of risk aversion might in practice indirectly contribute to this over-investment, leading to inefficiency in the market.

It is important to point out that our experimental design does not take into account banking of permits, which is allowed in almost all existing trading schemes. The presence of banking might smooth out the uncertainties regarding the permit price and thus facilitates better price discovery and a convergence path of the permit price. Nevertheless, another experiment also shows that this provision might have an adverse effect on the compliance rate and emissions level due to the perceived benefit of underreporting under the condition of imperfect enforcement, in which the audit probability is less than one (Cason and Gangadharan, 2006).

Another feature that might change compliance incentives under different penalty designs is the presence of a discount rate and the use of longer sub-periods to allow its effect to take place. The presence of a discount rate will reduce future costs, and when the costs and emissions cap in the trading scheme are stationary over time, firms might delay investment and shift their emissions towards the present, as pointed out by Kling and Rubin (1997). Therefore, it will be interesting to see the choice of firms' compliance strategies when the restoration rate is set equal to the discount rate to counterbalance the effect of decreasing costs due to the discount rate. Furthermore, longer compliance periods will also enhance the irreversible nature of investment in our model, in which risk preference might play a more important role in this case.

Finally, the use of a different auction design might reduce the inefficiency due to overbidding. Sealed-bid auctions have been recommended as a format that might facilitate bidding behaviour closer to bidders' private valuations. The downside of this auction format is that inexperienced bidders do not have an opportunity to revise their bids and there is a potential of having a winner's curse problem. An alternative format that might enhance efficiency is the use of an Anglo-Dutch format as suggested by Klemperer (2002). In this format, an auction is first run with an ascending clock auction until two bidders, or a few subjects in this case, are left. The auction is then continued using a sealed-bid format.

9. Conclusion

Although the importance of penalty design as an enforcement tool has been widely recognised in practice, the assessment of its efficacy on compliance rates in particular and market performance in general is difficult to make due to the unfeasibility of comparing existing trading schemes with different market structures and design features. A laboratory experiment offers the advantage of providing insight by isolating the effect of the penalty design in question as the market design features are held under control.

Our findings reveal that under the presence of subjects' risk preferences and some degree of permit price uncertainty, penalty levels provide an indication of total costs of compliance. This, in turn, affect firms' choice of compliance strategies as well as the compliance rate, although it does not influence the price discovery process. Surprisingly, risk preference does not have a direct role in influencing subjects' compliance decision. Nevertheless, it affects the permit price, which evidently is a significant determinant of compliance decisions and efficiency.

The make-good provision evidently induces higher investment and compliance levels than the fixed penalty rate does. It is important to point out that there is a trade-off between higher investment levels and efficiency because the penalty design that encourages higher investment levels for compliance purposes also corresponds to an adverse effect on efficiency. The inefficiency attributed to penalty design has not received sufficient attention, as the focus has been placed on compliance. These tradeoffs should be considered before a policy is implemented; otherwise, the efficiency of a trading scheme will be compromised.

In practice, the mixed penalty design is widely used in order to encourage higher compliance rates. The findings from our laboratory experiment support that view, as the presence of a FPR element and a MGP element induce higher compliance rates. Nevertheless, only the MGP element in the mixed penalty provides higher investment incentives at the expense of lower efficiency levels.

The main findings from this experiment are summarised as follows.

Table 16 A Summary of the Main Findings of the Penalty Design Experiment

No.	Main Findings	Supports
1.	Penalty design does not affect permit price	Result 1 and regression models for auction price
2.	Higher compliance rates are found with higher penalty levels in the FPR penalty, but not in the MGP penalty	Result 2 and Result 3
3	MGP penalty type induces higher investment incentives than the FPR	Result 4 and regression models for investment decision
4	The mixed penalty design provides higher investment and compliance incentives relative to the low FPR penalty.	Result 5
5	There is a trade-off between investment incentives and efficiency levels as the MGP penalty corresponds to a lower efficiency level and a higher investment level.	Regression models for efficiency

The mixed penalty design in this study also shed some light on the effects of tying the penalty rate to the auction price as proposed by the Australian Carbon Pollution Reduction Scheme. We do not find the occurrence of bid shedding at the auction, and both investment levels and compliance rates remain relatively higher compared to other penalty designs. It is only the MGP element of the double penalty, rather than the two penalty types (the FPR and the MGP element), that contributes to lower efficiency levels. With the presence of a reserve price at the auction, the proposed Australian model seems to serve its purpose of providing strong compliance incentives.

Overall, we believe that the experiment has provided valuable insights into how a specific penalty design can have an impact not only on compliance rates but also on market efficiency. Furthermore, a laboratory experiment can serve as a test-bed for policy makers to test how well the proposed design features of a trading scheme might work.

References

BURTRAW, D., PALMER, K. & KAHN, D. 2010. A symmetric safety valve. *Energy Policy,* 38**,** 4921-4932.

CASON, T. N. & GANGADHARAN, L. 2006. Emissions variability in tradable permit markets with imperfect enforcement and banking. *Journal of Economic Behavior & Organization,* 61**,** 199-216.

COMPTE, O. & JEHIEL, P. 2007. Auctions and information acquisition: sealed bid or dynamic formats? *RAND Journal of Economics (RAND Journal of Economics),* 38**,** 355-372.

CPB. February II 2003. *RE: Restoration Rates to Enforce Early Action.* Type to VROM.

FELL, H. G. & MORGENSTERN, R. 2009. Alternative Approaches to Cost Containment in a Cap-and-Trade System. *SSRN eLibrary*.

FISCHBACHER, U. 1999. z-Tree - Zurich toolbox for Readymade Economic Experiments. Institute for Experimental Research in Economics, University of Zurich.

GARRATT, R. & TRÖGER, T. 2006. Speculation in Standard Auctions with Resale. *Econometrica,* 74**,** 753-769.

GOEREE, J. K. & OFFERMAN, T. 2003. Competitive Bidding in Auctions with Private and Common Values. *The Economic Journal,* 113**,** 598-613.

GREINER, B. 2002. The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics.

HAILE, P. A. 2003. Auctions with private uncertainty and resale opportunities. *Journal of Economic Theory,* 108**,** 72-110.

HOLT, C., SHOBE, W., BURTRAW, D., PALMER, K. & GOEREE, J. 2007. Auction Design for Selling CO2 Emission Allowances under the Regional Greenhouse Gas Initiative Final Report.

HOLT, C. A. & LAURY, S. K. 2002. Risk Aversion and Incentive Effects. *American Economic Review,* 92**,** 1644-1655.

IVANOVA, G. 2007. Experimental test of the effect of imposing a penalty on buyers in a specific market structure. *51st Annual Conference of the Australian Agricultural and Resource Economics Society* Queenstown, New Zealand.

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

KLEMPERER, P. 2002. What Really Matters in Auction Design. *The Journal of Economic Perspectives,* 16**,** 169-189.

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

LEDYARD, J. O. & SZAKALY-MOORE, K. 1994. Designing organizations for trading pollution rights. *Journal of Economic Behavior & Organization,* 25**,** 167-196.

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

MURPHY, J. J. & STRANLUND, J. K. 2007. A laboratory investigation of compliance behavior under tradable emissions rights: Implications for targeted enforcement. *Journal of Environmental Economics and Management,* 53**,** 196-212.

NENTJES, A. & KLAASSEN, G. 2004. On the quality of compliance mechanisms in the Kyoto Protocol. *Energy Policy,* 32**,** 531-544.

PORTER, D., RASSENTI, S., SHOBE, W., SMITH, V. & WINN, A. 2009. The design, testing and implementation of Virginia's NOx allowance auction. *Journal of Economic Behavior & Organization,* 69**,** 190-200.

RESTIANI, P. & BETZ, R. 2010. A Theoretical Model of Optimal Compliance Decisions under Different Penalty Designs in Emissions Trading Markets. The Commonwealth Environmental Research Facilities (CERF) EERH *Research Report No. 87*.

SCHAEFFER, G. J. & SONNEMANS, J. 2000. The Influence of Banking and Borrowing under Different Penalty Regimes in Tradable Green Certificate Markets - Results from an Experimental Economics Laboratory Experiment *Energy & Environment,* 11**,** 407-422.

SHESKIN, D. J. 2004. *Handbook of Parametric and Nonparametric Statistical Procedures,* Boca Raton, Florida, Chapman & Hall/CRC.

SMITH, V. L. & WILLIAMS, A. W. 1981. On Nonbinding Price Controls in a Competitive Market. *The American Economic Review,* 71**,** 467-474.

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.

SZOLGAYOVA, J., FUSS, S. & OBERSTEINER, M. 2008. Assessing the effects of CO2 price caps on electricity investments--A real options analysis. *Energy Policy,* 36**,** 3974-3981.