EMBARQ
The WRI Center for Sustainable Transport

WORLD RESOURCES INSTITUTE

Market Design

cts MÉXICO

Center for Sustainable Transport

Market Design Considerations

M.Ph. Jorge A. Macías Mora
M.A.. Aimee Aguilar Jaber
Lic. Ulises Hernández Jiménez

Introduction

The introduction of an adequate designed market instrument may result in a cost-effective allocation by solving the problem of information and incentives that are posed by a command and control approach. As Tietenberg (2006) puts it, "When emission permits are transferable, those plants that can control most cheaply find it in their interest to control a higher percentage of their emissions because they can sell the excess permits. Buyers can be found whenever it is cheaper to buy permits for use at a particular plant than to install more control equipment. Whenever an allocation of control responsibility is not cost-effective, further opportunities of trade exist". Following the main objective of our project, we aim to design a market that can allow automobile industry firms to trade fuel efficiency averages of their sales in Mexico; this in order to comply with the fuel efficiency standard that is expected to be implemented in this country. This document provides a first approach to the considerations that should be taken into account for the design of such market.

The first part of this document describes the main considerations for the market design. This part will expand on four conditions that according to Robert W. Hahn and Robert N. Stavins (Hahn, R. 1995) are necessary in order to construct an efficient permit market. Further in the second chapter, the document will explore the two general forms that markets take and which set the bases for the different market settings and variations: The common Cap & Trade (C&T), and the Performance Standard Rate (PSR).

A third part provides an analysis of Banking and Borrowing, two instruments that can be used to introduce inter-temporal cost efficiency mechanisms into a market. The third part focus on instruments that allow for inter-temporal cost efficiency and that enhance market competitiveness.

I. Tradable Permit Systems: Building an Efficient Market

Tradable permit systems are relatively new instruments in the scope of environmental policy making. Along with pollution charges, market friction reductions, and government subsidy reduction, tradable permit systems fall into the environmental regulation category of market-based instruments. “Market-based instruments are regulations that encourage behavior through market signals rather than through explicit directives regarding pollution control levels or methods” (Stavins: 2002).

The main motivation of using a market-based (or incentive-based) mechanism for pollution control is that, contrary to command-and- control regulations, in which all firms abide the same share of the emission burden, incentive based policies (such as tradable permits) are meant to internalize the relative costs of abatement that firms face when searching compliance (even-out their marginal cost). One of the most important advantages that these kinds of mechanisms offer is the possibility of emission reduction at the lowest possible cost.

The logic behind tradable permit systems is that when emission sources are regulated and required to lower their emissions they have the possibility to sell extra-emission reductions or buy emission reduction shortfalls from each other. These systems can be basically of two types: credit programs, or cap-and-trade systems. In the first type, firms create their own emission credits when they reduce emissions to a lower than required level. In the second type, an overall level of permitted emissions is set by the regulator and permits that account for a certain amount of emissions are allocated among the sources that make up the system.

As Stavins and Hahn, R. (1995) indicates, “The full measure of potential cost savings offered by tradable permits in green house gases can be realized only if efficient markets develop”. Thus, a basic concern on the design of a market is to foster the required conditions in order for it to be efficient. Stavins and Hahn recognize four conditions that are necessary for a successful permit market. We have decided to explore such conditions as the starting point to the identification of specific considerations to take into account for the design of the market in question. These four conditions are the following:

1) Compliance.- A sufficient degree of compliance must be achieved

2) Transaction Costs.- Transaction costs must be low enough not to prevent efficient permit exchanges from taking place

3) Perdurability.- The policy must be seen as one that will remain in place for a significant period of time

4) Market Competitiveness.- The market for permits must be competitive

The aim of this chapter is to explore the above conditions in order to 1) understand their importance in the endeavor of designing an efficient permit tradable market system 2) identify the main factors that create the context under which such conditions can be ensured. We have deemed prudent to revise, when necessary, the relation between having this conditions in the framework of the regulation itself (such as in the case of a fuel efficiency standard), and having them in the permit tradable market system that accompanies the regulation.

a. Compliance

As T.H. Tietenberg states, "establishing emissions trading systems that allocate the control responsibility cost-effectively is of little value if sources regularly fail to comply with the terms of their permits" (Tietenberg,T. 2006). As it will be seen in this section, adequate monitoring and enforcement mechanisms are necessary conditions for the achievement of compliance in any environmental regulation and particularly in an emission trading system.

Monitoring

An effective system of monitoring is one of the most important predecessors of an adequate enforcement and a critical input for compliance. Accurate monitoring is the only way to effectively detect the violation of an actor to a regulation, which as it will be seen, is the first necessary step in the process of effective enforcement.

One way in which monitoring could be carried out is by the direct measurement of emissions. The most important problem that could be brought by this procedure is that it is expensive for the regulator. In the case of carbon emissions, as there is a close relation between carbon content in fuel and CO2 emissions, it is possible to monitor fuel consumption rather than CO2 emissions directly, which is usually less expensive. In the case where the regulator decides to directly monitor fuel flows, it must set an efficient and reliable system that provides fuel consumption information, along with an organized system for sources to be monitored. This is still expensive and time consuming.

A widely used monitoring method that can avoid the complications of direct measurement of emissions is self-reporting of emission levels (or fuel consumption). This reporting is done by the actors and cross-checked by the regulator through other sources of information. This cross-checking can be done through governmental data-bases, or by private sources. "Self-reporting has the virtue that it is relatively inexpensive, and can provide remarkably accurate information… (however, this only happens)… when it is backed by the appropriate sanctions for misrepresentation" (Tietenberg,T. 2006). Also, for self-reporting monitoring to be adequate, it should include clear requirements and revision criteria for reports and information periodically submitted by firms, as well as the certification of the methodology by which emission or fuel consumption levels will be measured, and certification of specialized actors whose measurements will be recognized by the regulator. Finally, a simple and accurate system of permit certification by the regulator is crucial to the efficiency of the tradable permit market.

Enforcement

Effective enforcement has to do with the fact that authorities involved in the environmental regulation are able to detect violations and follow through the process by which a sanction will be assigned and paid by the non-compliant actor, or by which a new compliance period will be negotiated. According to Tietenberg, enforcement at the national level (which is the scope of this analysis), can be divided into the next four steps:

1) Detecting the violation

This step can be achieved by using distinct methods such as self-certification, on-site inspections, or direct monitoring of polluting flows. It is important to mention that regular updates on the information usually make it easier to detect violation and more likely to achieve compliance. Another thing that is important to keep in mind is that "the degree of scrutiny…(necessary for effectively detecting the violation)…can be tailored to the likelihood of noncompliance" (Tietenberg,T. 2006).

2) Notifying the source

This step plays a crucial role in the path towards compliance because, no matter which is the response of the non-compliant to the notification, it serves as the start for non-compliance to be dealt with. In the case where the actor decides to return to compliance, "notifying the source of violation initiates the process of regaining compliance" (Tietenberg,T. 2006) . In this case, it sets a start to any possible negotiations between the non-compliant and the regulator. In the case where the non-compliant actor, decides not to comply, it serves as the beginning of any legal procedures that the regulator will have to enroll in.

3) Negotiating a compliance schedule

During this step, the regulator must have the ability to engage in negotiations with the non-compliant actor that establish clear conditions and deadlines for the actor to return to compliance.

4) Applying sanctions for noncompliance when appropriate

The effective application of noncompliance sanctions represents the crucial mechanism for the compliance of any policy. This is because when the cost of non-compliance is cero (no sanctions are effectively applied) every other action that an actor would take towards compliance becomes relatively costly. In fact every other step in the process described in this section could have been effectively achieved, but if the course followed by the regulator in the case of noncompliance does not lead to the appliance of a sanction when any other agreement has been reached, then compliance to the policy will not be achieved. In this last step, the institutional and legal frameworks, as well as the administrative mechanisms surrounding the tradable permit market have to facilitate the execution of non-compliance sanctions.

The four steps described in this section are necessary for the effective enforcement of any environmental regulation. As it has been said, the theory behind a tradable permit market states that this is an instrument by which regulated actors can comply with an environmental regulation in a cost-effective way. Because a high level of weakness in the four steps of enforcement results in non-compliance, then a mechanism that makes compliance cost-effective simply cannot exist. In other words, in the case of environmental regulations where violations cannot be accurately detected and/or where compliance orders and noncompliance sanctions are not made effective, emission permits (or credits) have no market value and therefore will not be traded.

As Tietenberg well says, “it could be very misleading to assume perfect enforcement when comparing regulatory approaches…(since)…not only will some violations inevitably go undetected, but not all detected violations result in compliance” (Tietenberg,T. 2006). Because of this, when considering the enforcement mechanisms, one should evaluate the degree to which they assure compliance, even after some delay, rather than assessing the regulation in the measure of perfect enforcement. Under the assumption of imperfect enforcement, the degree of compliance assurance becomes a matter directly linked with the cost-minimizing behavior of the actors involved. In other words, the success of enforcement mechanisms depends in a great extent to their capacity to make compliance less costly, relatively to noncompliance. The regulator can make noncompliance more costly by 1) increasing the probability that violations are detected and effectively sanctioned 2) increasing the level of the sanctions. The important point for increasing enforcement effectiveness is that actors involved will have to perceive that the value of the sanction level, times the probability of being effectively sanctioned, is greater than the costs of complying.

The use of tradable permit systems as instruments of environmental regulation have a two way relation with the level of enforcement achieved. On the one hand, as it has already been mentioned, without a level of enforcement that makes non-compliance have a cost, a tradable permit market cannot even exist. In this same sense, even when noncompliance costs are necessarily high to allow the creation of a tradable permit system, weak enforcement mechanisms may give rise to low price permits and therefore compromise the potential environmental benefits of the regulation and the market as a whole; "lower permit prices means fewer emissions controlled by other firms"(Tietenberg, T. 2006). On the other hand, when enforcement mechanisms have already proved to be sufficiently efficient as to allow the existence of a permit market and to maintain the value of permits sufficiently high for the regulation to achieve environmental benefits, a tradable permit system can 1) make the cost of compliance lower for those actors who would otherwise choose to lobby in order to make the regulation more lax 2) incentive compliance by actors that face relatively high abatement/lobby costs and that would otherwise choose the noncompliance alternative. In the first case, a tradable permit system can increase the relative cost of lobbying and prevent the regulator from having to choose between a low degree of enforcement and a low degree of environmental benefits. In the second case, it could directly increase the cost of noncompliance, reduce the incidence of violations and increase the degree in which the enforcement of the regulator is perceived as certain by the actors.

b. *Transaction Costs*

One crucial condition for a market to be efficient is that its mechanisms are designed in such a way that the costs of participating in permit trading (transaction costs) do not exceed the gains generated from buying and selling permits. Talking about transaction costs includes those costs of "researching the market, finding buyers or sellers, negotiating and enforcing contracts for permit transfers, completing all the regulatory paperwork, and making and collecting payments" (Tietenberg,T. 2006). Of course, the extent to which the presence of transaction costs will be translated into lower amounts of trading in a given market depends on the magnitude of these costs in comparison to the benefits of transactions. To some degree, this magnitude can be predicted by identifying factors that tend to generate systems with higher administrative burden. For example, Hargrave recognized the following four factors:

1) *Large number of actors that participate in a market*. The larger the number of actors the more information will be needed in order for them to find trading partners and decide to carry out trading transactions.

2) *Low availability of needed data*. In the extent in which actors can obtain detailed information about other actors participating in the market they will not have to gather this information through means other than the system built around the permit market are reduced. The availability of data can reduce transaction costs even when there are a large number of actors in the market.

3) *High complexity of reporting requirements*. Reporting requirements should not be too time and/or labor consuming

4) *Confusing accounting of emission quotas*. A system where firms hold emission permits for a small amount of green house gases and in a well defined market, results in simpler procedures for the firm to account its emission quotas.

All of the factors above make it necessary for firms to engage in a series of activities that can compensate for the information and simplicity that the tradable permit system cannot provide, and which are extremely important for them to engage in trade. Such activities constitute costs, in both resources and time, which will have to be subtracted from the gains that actors could have from trade. This will delay transactions, inhibit transactions as a result, and may even prevent the tradable permit system to function at all.

Again, one cannot assume that a tradable permit system with no transaction costs is possible to be built. However, adequate trading mechanisms that minimize transaction costs as much as possible will minimize the obstacles that can deviate any market from the outcome that its system of incentives would allow it to reach in the absence of such costs. In the case where such incentives would allow the market to generate a cost-effective distribution of abatement costs among the regulated actors, designing market mechanisms that do not balance-off such incentives (even when they can diminish them to some extent) becomes a crucial matter in the search for market efficiency.

Although not mentioned in Han and Stavin`s four conditions, we deem necessary to include in this subsection not only transaction costs but administrative costs as well. The reason to do this is to make sure that the market mechanism makes the fuel efficiency standard not only cost-effective for regulated actors who trade permits but for the regulator and the society as well.

The minimization of administrative costs follows a parallel logic to the minimization of transaction costs: If a regulator sets up a permit market system as the instrument that will foster flexibility and cost-effectiveness to an environmental regulation, but the gains of making compliance more cost-effective are balanced-off by the costs incurred by the regulator to set and maintain the system, then the whole purpose of the market instrument is wiped out. In the case of a market for permits, administrative costs include setting up the registry system to keep track of permits, the design and enforcement for the rules for trading, the design and enforcement of a compliance system to match permit holdings to actual deviations from the standard, revision and collection of reports with the fuel efficiency of vehicles and the sale`s record, and the procedures for managing noncompliance. The four factors that were identified by Hargrave as leading to high transaction costs are also a helpful guideline to recognize conditions under which administrative costs will have a tendency to be high, and therefore put in danger the potential social gains of a tradable permit instrument.

c. *Perdurability*

Perdurability of an environmental regulation and the instruments chosen to articulate it depends on how much certainty the regulated actors have upon the future existence and well functioning of both the regulation and the instruments. In this sense, the actions of the regulator must be perceived not only as effective but also as long lasting.

In the case of a fuel efficiency standard with a tradable permit system, perdurability means firstly that the standard has to be seen as a policy that will be revised in terms of its progressing level but whose existence along time should not be put into question by the regulated agents. If this conditions mentioned are not met, agents will under estimate the value of compliance (even to a 0 value) and postpone all actions required to comply (buying permits, investing in technology or engaging in a sale composition change).

As a second necessary step is that of fostering perdurability in the permit market. This is directly linked with the institutional and legal framework mechanisms that facilitate the execution of permit trading contracts. Contracts of trade in the market should be perceived as valid to the actors involved in the buying and trading of permits. Also, such framework has to offer compensating mechanisms in the case where contracts are not executed. If this does not occur, permits will not be seen as assets that are worth acquiring, since they hold no real future value. In the case where banking and/or borrowing are included as mechanisms of the market, certainty of the existence of the standard and the market along time gains even more importance.

d. *Market Competitiveness*

As Stavins and Hahn describe, “(t)he degree of competition in the permit market will also affect the extent to which potential cost savings are likely to be realized” (Hahn, R. 1995). Therefore, fostering competitiveness in the permit market is essential for this instrument to effectively provide flexibility and cost-effectiveness to the fuel efficiency standard.

An important clarification worth to make before following this analysis is to define what will be understood by a high degree of competition. Based on Tietenberg´s work, a market with a high degree of competition will be a market in which market power concentration is avoided (Tietenberg, T. 2006). Market power concentration refers to, in this context, to the ability by one or more actors to influence the price of permits, and/or the ability of one or more actors to leverage in the permit market in order to gain power in the out-put market (Tietenberg, T. 2006).

Based on the above definitions, an important question can be identified when analyzing the likelihood of market power in a certain permit market system: how direct is the relation between competitors in the permit and the output market? In the case where few actors are competitors among each other in both markets "the permit market would be a relatively (or completely) inefficient vehicle for a predatory source to use to inflict harm on… (its out-put market)… competitors" (Tietenberg, T. 2006). However, in a market where various or all actors compete with each other in both markets, mechanisms that avoid leverage in the permit market that can directly change the power relation among actors in the out-put market is central to the positive effects of the permit market instrument.

It is important to say that the determination of the value of the sanction for under-compliance that will be inflicted by the regulator is a very relevant factor also in the fostering of competitiveness in the permit market. As it has been already mentioned in this chapter, sanctions must be high enough to promote compliance and generate permits that hold value. However, they should be low enough to prevent permit holders to set a price that is too high for the market to exist. In this way, sanctions can help in the avoidance of market power concentration because they can prevent firms that hold permits initially to have the power to influence permit prices and do leverage in this market to gain power in the output market.

II. Absolute and Relative Structures

Text book market structures always result in an efficient allocation; in practice the diverse market structures have shown some particular feats that make them suitable for specific industries. The objective here is to study the specific characteristics of the diverse structures, in order to set a regulation on emissions and to analyze the impact of the diverse trade mechanisms on the overall performance and cost-effectiveness of a market instrument in a fuel economy regulation. The general structures that we analyzed have their main divergence in the way they set their emissions cap. The common Cap & Trade (C&T) system sets its emission limit at an absolute target that puts a ceiling overall emissions while, the Performance Standard Rate (PSR) restricts emission per unit produced. This simple difference, in the way the targets are set, represents important disparities in the institutional arrangements needed, implications for the industry, environmental results, market uncertainty and overall functioning of the market. We compare the benefits and drawbacks of these two basic markets in order to portrait the conditions and circumstances under which this basic systems work better. We base our conclusions on a modified version of the partial equilibrium model presented on Gielen, 2002.

Cap & Trade (Absolute Target)

The most commonly used and studied market structure is the Cap & Trade system with an absolute target. Under this system the government restricts emissions relative to the status quo or a projected baseline level and allows agents to trade by issuing emission permits equal to the absolute emission target. Thus, participating polluters face an absolute cap on their emissions, and then decide, either, to reduce emissions, buy additional permits at the permit market, pay a fine or tax for under-compliance to the regulator, or a combination of the three, according to the relative cost of each option. The former is shown in Figure-1 which shows a graphic portrait of the Cap & Trade scheme aiming for an absolute target.

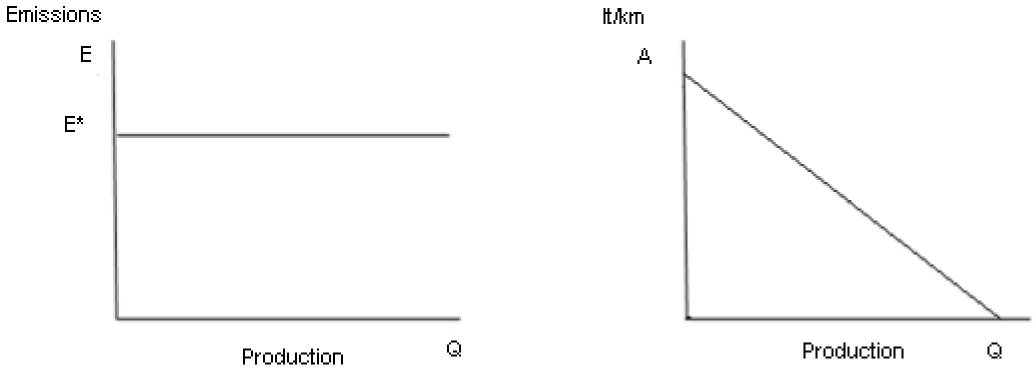


Figure 1: Traditional Cap & Trade Scheme (Absolute Target)

In the traditional Cap & Trade Scheme a level of emissions or absolute target (E) is set as the maximum amount to be emitted (E*). This maximum absolute target (E*) represents the overall emissions by the industry and is the result of the product of a total quantity forecasted. To ease the comparison, the abatement variable (a) is going to be defined as the efficiency level but in terms of lts/km of fuel (i.e. fuel economy), as a consequence, a higher level of *a* means a lower level of abatement. The forecasted value of the absolute target E^e is given by the product of an emission factor (EF) which is assumed to be constant, the expected abatement per unit (a^e) and the expected quantity produced (q^e), thus making $E^e = EFq^e a^e$. The EF is the factor by which the abatement level (expressed in lt/km can be converted into gms of CO_2 (emissions).

The variables (a^e, q^e) are forecasted and its real value will be revealed at the end of the time frame hence they may differ from the optimal level (a*,q*) which is the cap level that corresponds to the correct prediction of abatement and production costs. Variation from social optimum by the forecasted parameters may be possible and thus we should assign error terms leaving $\Delta a^{*} = \mu a^{*}$ and $\Delta q^{*} = \omega q^{*}$.

Using a simple model to illustrate the scheme, we start with the familiar cost minimization problem of a representative firm, subject to a C&T program. The firm minimizes production and abatement cost C(q,a) with *q* as the quantity produced and *a* as the level of abatement. Furthermore, the firm does not have market power and therefore the permit price Pp is given to it. Total emission *E* is determined by the quantity produced and the level of abatement, and the firm is restricted to emission level *E**. Thus we have the following problem for the individual firm:

$$1)\quad \mathbf{Min\ C(q,a) = c(q,a) + Pp\ (E(q,a) - E^{*}) + Pp(E^{*} - E^{e})}$$

$$\text{with } \mathbf{E^{e} = (EF)a^{e}q^{e}} \text{ and } \mathbf{E^{*} = (EF)a^{*}q^{*}}$$

With c_q, c_{qq}, c_a, E_a E_q, $E_{aa,}$ > 0

The diverse elements that constitute firm´s total costs (Equation 1) are thus defined accordingly as the sum between production costs, social costs and the associated cost due to forecasted error. The forecast error $Pp(E^* - E^e) = PpEF(\Delta a^* q^* = \Delta q^* a^* + \Delta a^* q^* + \Delta q^* \Delta a^*)$

Considering the variation/error from the social optimal:

$$2)\quad \mathbf{\Delta a^{*} = \mu a^{*}} \text{ and } \mathbf{\Delta q^{*} = \omega q^{*}}$$

$$3)\quad Min\ C(q,a) = c(q,a) + Pp\ (E(q,a) - \mathbf{(EF)a^{*}q^{*}}) + PpEF(\mathbf{\mu q^{*}}a^{*} + \mathbf{\omega}a^{*}\mathbf{q^{*}} + \mathbf{\omega\mu q^{*}}c^{*})$$

With c_q, c_{qq}, c_a, E_a E_q, $E_{aa,}$ > 0

FOC

$$4)\quad c_a = -Pp[(E]_a) - PpEF\ (\mu + \omega + \mu\omega)q^{*}$$

$$5)\quad c_q = -Pp[(E]_q) - PpEF(\mu + \omega + \mu\omega)a^{*}$$

The First Order Conditions (FOC) gives us interesting results. The first equation tells us that the marginal costs of increasing abatement will be equal to the market value of the emissions reduced by this reduction in efficiency. The error term $(\mu + \omega + \mu\omega)$, which is the result of the variation from forecasted and optimal unitary abatement and production, can take values along $-\infty < \mu, \omega < +\infty$. Consequently, under perfect competition and in the case forecasted and social optimal values are the same, μ and ω become zero then marginal abatement costs are equal for all polluters and total abatement costs are minimized.

In the case $(\mu + \omega + \mu\omega) > 0$, ($E^e$) is overestimated and the marginal cost (c_a) takes a smaller value than in the case where optimal abatement costs are minimized and we will have smaller than optimal abatement level.

In the case $(\mu + \omega + \mu\omega) < 0$, ($E^e$) is underestimated and the marginal cost (c_a) takes a greater value than in the case where optimal abatement costs are minimized. Abatement is greater than the optimal abatement level.

The second FOC states that marginal costs of production equal the value of production forgone due to pollution, this is, the marginal addition to emissions of raising output times the permit price or shadow price of emissions minus the contribution given by the resulting adjusted unit efficiency given by $(\mu + \omega + \mu\omega)a^*$. Conversely, in the case μ and ω become zero then marginal production costs are equal for all polluters and costs are minimized.

In the case $(\mu + \omega + \mu\omega) > 0$, ($E^e$) is overestimated and the marginal cost of production (c_q) takes a smaller value than in the case where optimal costs are minimized and we will have a greater than optimal production level.

In the case $(\mu + \omega + \mu\omega) < 0$, ($E^e$) is underestimated and the marginal cost of production (c_q) takes a higher value than in the case where optimal costs are minimized and we will have a smaller than optimal production level.

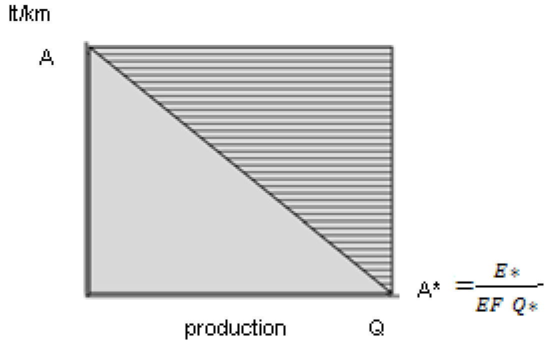
Graphic Analysis of an Absolute Target

It is important to consider that deviations from the optimal cap level can be caused because of a deviation of a^e from a* or deviations from q^e from q* (or both). The different cases and magnitudes of these deviations will cause different economic and environmental results for the market in question and in a broader sense for the economy.

As it has been shown in Figure 1, an absolute cap, that maintains permitted emissions at a fixed level, leaves firms to face the regulation by deciding among different combinations of abatement and production. The deviations in a^e and q^e from their optimal values will change the slope of the equation that describes the relation of abatement and production that exists for the compliance of the industry with a certain cap level imposed.

Parting from the definition $E^e = EF\ A^*(1 + \mu)Q^*(1 + w)$ we know that for any given level of absolute target E^e chosen by the regulator, the equation that will describe the relation between the level of abatement and production in the industry is the following: A*=($E^{\uparrow}e/(EF\ Q*(1 + \mu)(1 + \omega))$) .In the case where the regulator has chosen a level of absolute target $E^e = E*$ (has correctly predicted A* and Q*) the relation is solely A*=$\frac{E*}{EF\ Q*}$. However, when A* and Q* have not been correctly predicted, then the $E^e \neq E*$ and therefore the industry will choose combinations of abatement and production that either too high or too low to make emission reduction cost effective.

Figure 2 below, offers a graphic illustration of strict and lax policies under an absolute target scheme, compared to an optimal policy, where the regulator was able to estimate correctly the real abatement costs and the level of production where emission reduction results cost-effective.



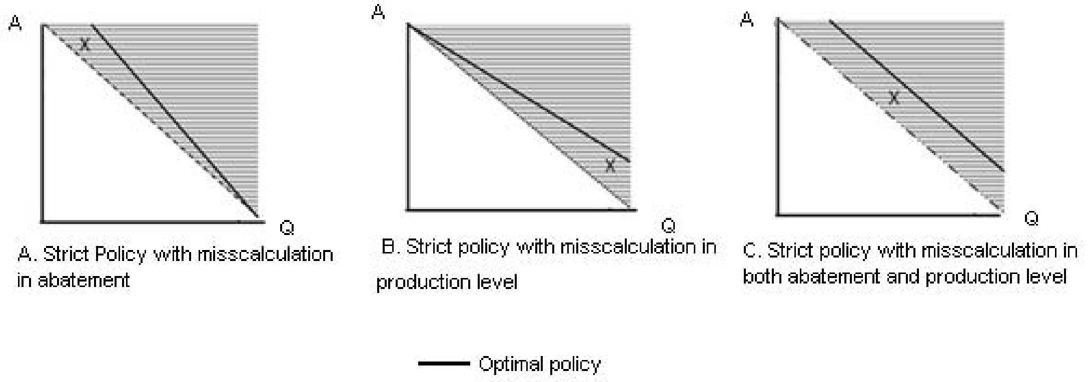
Firgure 2: Strict and Lax Policies under an Absolute Target Scheme

Table 1 summarizes the miscalculations of Q and A levels made by the regulator and the type of policy associated to them.

	Strict Policy	**Optimal Policy**	**Lax Policy**
Color	░	■	▤
Conditions	$a^e < a^*$ and /or $q^e < q^*$ ($\mu<0$ and/or $w<0$)	$a^e = a^*$ and $q^e = q^*$ ($\mu=0$ and $w=0$)	$a^e > a^*$and/or $q^e > q^*$ ($\mu>0$ and/or $w>0$)

Table 1: Summary of Strict, Optimal, and Lax policies under an Absolute Target Scheme

Strict Policies



A. Strict Policy with misscalculation in abatement

B. Strict policy with misscalculation in production level

C. Strict policy with misscalculation in both abatement and production level

Figure 3: Strict Policies under an Absolute Target Scheme

- If $a^e < a^*$, the regulator calculated that firms were more efficient than they are, and therefore allowed a number of emissions that is too low for the technological possibilities of the industry. Since $\mu<0$, the slope of the set of (Q, A) coordinates that will reach compliance will be smaller than in the optimal policy. The average emissions per unit of abatement at every level of production will be lower than optimal. This can be seen in Figure 4 where each unit of production is more expensive in terms of the price of emissions than with an optimal policy. The extra limitation in lt/ km produced because of the error in the policy is shown in the area E.

- If $q^e < q^*$, the regulator calculated an optimal production, given the costs faced by producers, that was smaller than it really is. He therefore assigned a maximum number of emissions coherent with this quantity of production. Since $w<0$, the slope of the set of (Q,A) coordinates that will reach compliance will be greater than in the optimal policy. The average emissions per unit of abatement at every level of production will be lower than optimal. This can be seen in Figure 4 where each unit of production is more expensive in terms of the price of emissions than with an optimal policy. The extra limitation in lt/ km produced because of the error in the policy is shown in the area E

Consequences of a Strict Policy

A strict regulation is characterized by a supply shortage in the emission permit market, as illustrated in Figure 4.

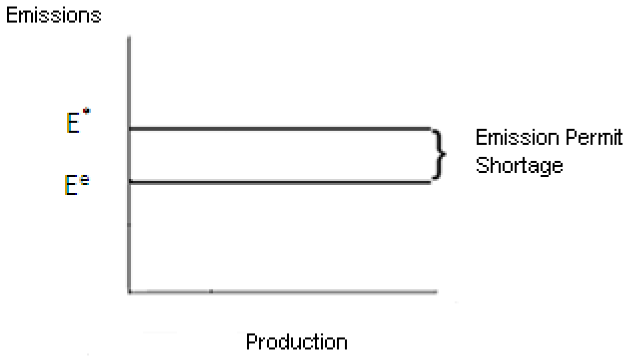


Figure 4. Strict Absolute Target

The shortage of emission permits reduces the industry average level of inefficiency that firms can have in order to produce. As shown in Figure 5, production of the industry falls from Q* to Q^e. The price of a strict policy is therefore the quantity that would have been produced if the regulation had been optimal and that is not produced because the regulation was too strict.

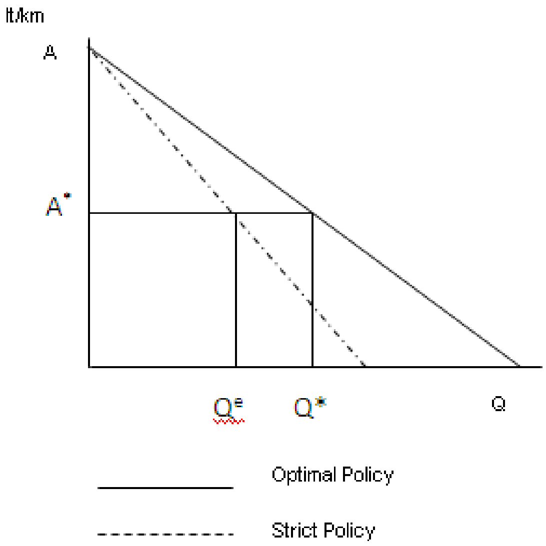
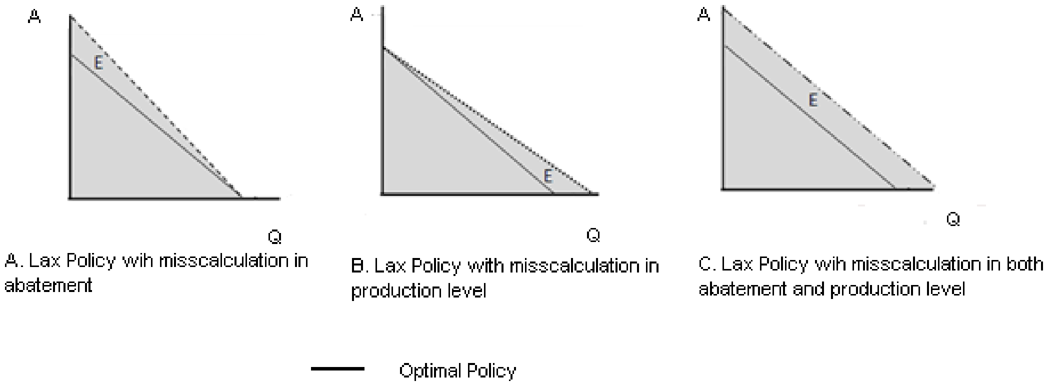


Figure 5. Output deviation in an Absolute Target Strict Policy

Lax Policies



A. Lax Policy wih misscalculation in abatement

B. Lax Policy with misscalculation in production level

C. Lax Policy wih misscalculation in both abatement and production level

Figure 6. Lax Policies under an Absolute Target Scheme

- If $a^e > a^*$, the regulator calculated that firms were less efficient than they actually are, and therefore allowed a number of emissions that is too high for the technological possibilities of the industry. Since $\mu > 0$, the slope of the set of (Q,A) that will reach compliance will be greater than in the optimal policy. The average emissions per unit of abatement at every level of production will be higher than optimal. This can be seen in Figure 6 where each unit of production is less expensive in terms of the price of emissions than with an optimal policy. The extra lt/ km produced because of the error in the policy are those shown in the area E.

- If $q^{e >} q^*$, the regulator calculated an optimal production, given the costs faced by producers, that was higher than it really is. He therefore assigned a maximum number of emissions coherent with this quantity of production. Since $w > 0$, the slope of the set of (Q,A) coordinates that will reach compliance will be smaller than in the optimal policy in a way in which the policy results lax. The average emissions per unit of abatement at every level of production will be higher than optimal. This can be seen in Figure 6 where each unit of production is less expensive in terms of the price of emissions than with an optimal policy. The extra lt/ km produced because of the error in the policy are those shown in the area E.

Consequences of a Lax Policy

As it is exemplified in Figure 7, a lax policy is characterized by an oversupply of emission permits in the market.

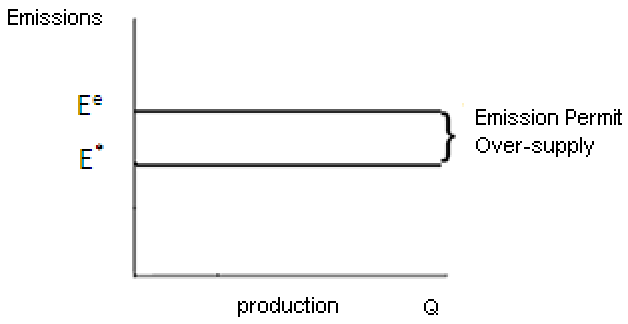


Figure 7. Lax Absolute Target

The oversupply of emission permits increases the industry average level of inefficiency that firms can have in order to produce. As shown in Figure 7, production of the industry raises from Q* to Q^e, while the level of inefficiency of the firms is the same, therefore the extra production units produce extra emissions. The price of a lax policy is therefore the quantity of emissions that would have not been produced if the regulation had been optimal and that are produced because the regulation was too lax.

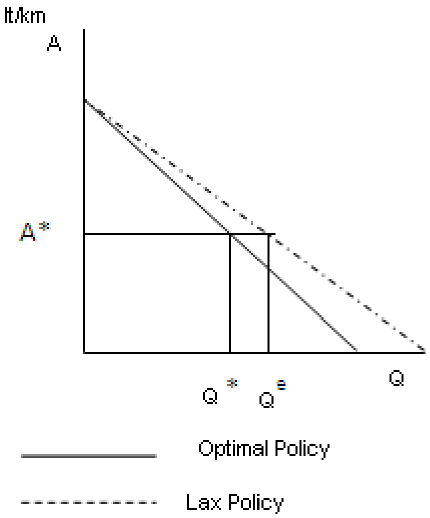


Figure 7. Output deviation in an Absolute Lax Policy

Performance Standard Rate (Relative Target)

The basic idea of a PSR trading system is that the government defines emission limits in terms of a maximum allowable level of emissions per unit of production or a minimum fuel consumption level km/lt. If a CO_2 emitting source is able to generate fewer emissions per unit of activity than allowed by the PSR, the source is allowed to sell the difference between the allowed volume of emissions and the actual level of emissions. Conversely, if a source generates more emissions per unit of activity than allowed, it can buy the difference on the market for CO2 allowances.

The system of rate-based or PSR trading provides incentives for technological innovation and diffusion of CO2-efficient production techniques. There is an incentive because any reduction in CO2-intensity (CO2 emissions per unit of output) can directly be sold on the market for emissions allowances.

There are quite a lot benefits from this approach. Trading with relative targets is more easily combined with existing regulation and policies therefore it has greater political acceptability. Polluters only have to pay for emissions above their relative target and not for their remaining emissions, which is a concept more easily communicated than an absolute level.

Relative C&T mechanisms allow entry and expansion at no extra costs as long as emissions per unit of output or input are below the relative target. Grandfathering permits in a classical C&T scheme with absolute targets does not allow for free entry. Instead, firms will have to buy additional permits if they want to enter a market or increase their emissions. Figure 7 portraits how this scheme works.

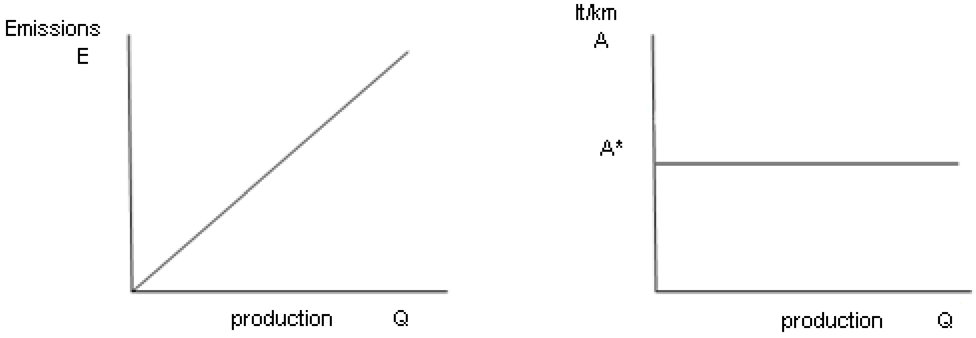


Figure 8: Performance Standard Rate (Relative Target)

Given that under the Performance Standard Rate a fixed efficiency standard (A*) is required for each unit, the efficiency requirements do not increase as the quantity produced (Q) increases as shown in the second graph in Figure 2. On the other hand, total emissions (E) keep a direct relationship with the quantity produced. As quantity produced increases total emissions increase as well.

In a scheme with trading around relative targets like PSR, emissions as such are no longer restricted, but emissions per unit of output. Firms are allowed to produce as long as they fulfill the required unit abatement level. To capture this type of policy target we introduce the relative target (per unit of output) A*. The firm now minimizes the following cost function:

$$1)\quad Min\, C(q,a) = c(q,a) + Pp\left(\frac{E(q,a)}{qEF} - a^*\right)q + Pp(a^* - a^e)q$$

With c_q, c_{qq}, c_a, E_q, E_{aa}, $E_a > 0$

Considering the variation/error from the social optimal in the expected abatement level:

$$2)\quad \Delta a^e = \mu a^*$$

$$3)\quad Min\, C(q,a) = c(q,c) + Pp\left(\frac{E(q,a)}{qEF} - a^*\right)q - Pp(\mu a^*)q \quad *$$

A firm will buy permits if its emissions per unit of output (total emissions E divided by the total quantity produced q) are higher than the allowed relative target a^e.

FOC

$$4)\quad c_a = -Pp\left({}^{E_a}/_{EF} - (1+\mu[\square)q]^* \right)$$

$$5)\quad c_q = -Pp\left({}^{E_q}/_{EF} - (1+\mu)a^*\right)$$

From equation 10 we can see that Gielen´s result still holds for the analysis in marginal costs of production. There is a production subsidy given by the term (1+μ)a*, thus, in the absence of errors in quantity , consumers pay less in PSR than in the absolute target. Production levels in PSR are therefore bigger. Nonetheless, equation 9 gives us a very interesting result. Abatement in PSR (whenever μ equal cero) is attained at a lower marginal cost than in the absolute target case.

Graphic Analysis of a Relative Target

In the case of a relative target, the deviations from an optimal target are only caused by deviations of the expected achievable abatement from the real achievable abatement of the industry. In this case, firms will choose different combinations of production and emissions per unit of production that will allow that, after having traded average efficiency certificates, they can comply with the average efficiency required.

Parting from the definition $E^e = EF\ A^*(1+\mu)Q^*$ we know that for any given level of relative target A^e chosen by the regulator, the equation that will describe the relation between the level of emissions and production in the industry is the following: $\frac{E^e}{EF} = A*(1+\mu)Q*$ .In the case where the regulator has chosen a level of relative target $A^e = A*$ (has correctly predicted abatement costs) the relation is solely $\frac{E^e}{EF} == A*Q*$ However, when $A^e \neq A*$ the industry will choose to emit a level of pollution that is not cost effective.

It is important to understand that under a relative target scheme, a fuel efficiency level is required for every vehicle. The pollution emitted when achieving a fuel efficiency level equal or below the requirement is not paid for. Therefore, a PSR functions as an inefficiency subsidy for every unit produced, equal to the emissions that each vehicle can make without the firm being charged for.

Figure 9 below, offers a graphic illustration of strict and lax policies under a relative target scheme, compared to an optimal policy, where the regulator was able to estimate correctly the real abatement per unit where emission reduction results cost-effective.

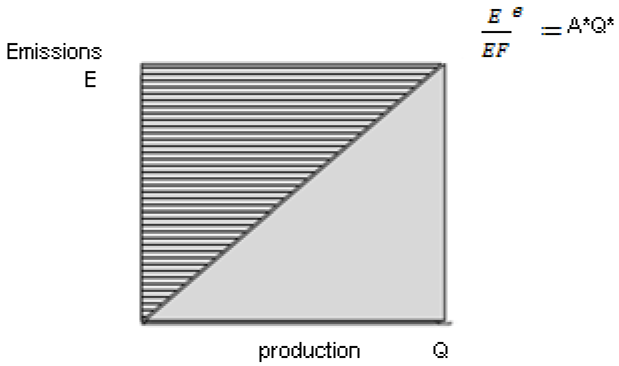


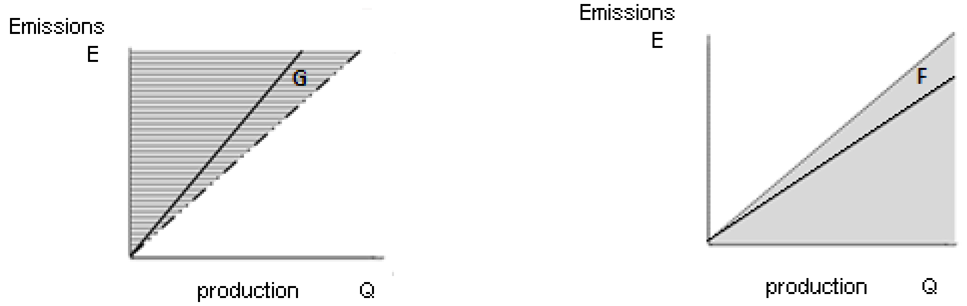
Figure 9. Strict and Lax Policies under a Relative Target Scheme

.

	Strict Policy	Optimal Policy	Lax Policy
	░	■	▤

Color			
Conditions	$a^e < a^*$	$a^e = a^*$	$a^e > a^*$

The way in which each type of policy shifts the slope of the relation between emissions and production can be seen in the Figure 10.



a. Strict Policy under a Relative Target Scheme

b.Lax Policy under a Relative Target Scheme

Figure 10. Comparing a Strict and a Lax Policy under a Relative Target Scheme

- If $a^e < a^*$, the regulator calculated that firms were more efficient than they are, and therefore allowed a number of emissions, per vehicle, that is too low for the technological possibilities of the industry. Since $\mu<0$, the slope of the set of (Q,E) coordinates that will reach compliance will be smaller than in the optimal policy in a way that vehicles that with an optimal policy would be efficient are now inefficient. In this case, for any given Q*, the level of E^e chosen will be smaller than E*. This can be seen in Figure 11 where vehicles that are in the area G are now in the stripped area. In this case firms end up paying units of emission that for a cost-efficient emission reduction to happen would have not been paid for; the inefficiency subsidy per unit is less than optimal

- If $a^e > a^*$, the regulator calculated that firms were more inefficient than they actually are, and therefore allowed a number of emissions, per vehicle, that is too high for the technological possibilities of the industry. Since $\mu>0$, the slope of the set of (Q,E) coordinates that will reach compliance will be greater than in the optimal policy. In this case, for any given Q*, the level of

E^e chosen will be higher than E*.Vehicles that with an optimal policy would be inefficient are now efficient. This can be seen in Figure 11 where vehicles that are in the area F are now in the solid area. In this case, firms do not pay for units of emission that for a achieving an optimal emission reduction they would have had to pay; the inefficiency subsidy per unit is higher than optimal.

Consequences of a Strict Policy

As it has been said, under a Performance Standard Rate, a strict policy means that, under the level of relative target chosen, some of the vehicles of a firm`s fleet that, with an optimal policy, would be efficient (would not pay for their emissions) now pay an extra cost for their production. In the same way, inefficient vehicles will pay even more of their emissions than under an optimal relative target. Shortly, less emission per vehicle would be subsidized. This fact is shown graphically in Figure 11 where it can be seen that at any given level of production of the industry, a number of lt/km of fuel consumption that under an optimal policy (A*) would not be charged, under the chosen per unit level (A^e) would have to be paid for. The number of additional lt/km of fuel charged because of the strict policy is equal to ($A^* - A^e$)x Q*.

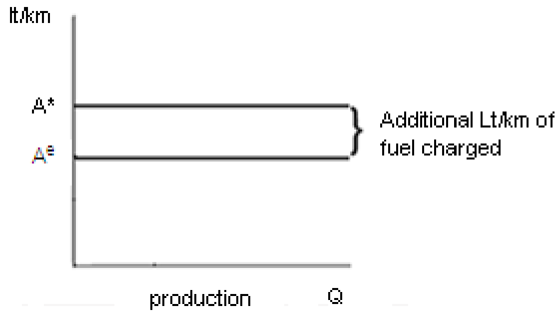


Figure 11: Strict Relative Target

The problem that can be caused by the establishment of a deviation under a relative target scheme is that firms will face higher than optimal costs for a certain level of production. Because abatement costs were under-calculated, one part of the pollution emissions will not be reduced by technology improvements. Firms will have to cover an extra charge that they can pay by paying the regulator`s fine. However it is also probable that some of the cost can be confronted by changing the composition of the fleet. In this case a relative target that is too strict can lead to firms to stop producing many vehicles that they could have offered to the consumers under and optimal policy. Figure 12 below illustrates this deviation by showing that for any level of production the industry will have to emit less, per vehicle, than they would do under an optimal policy.

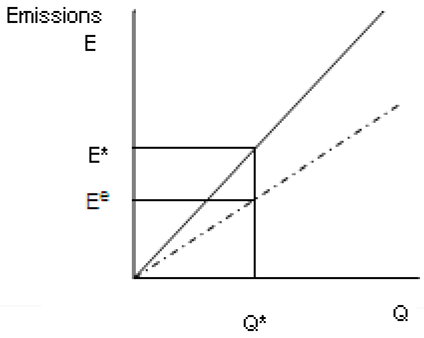
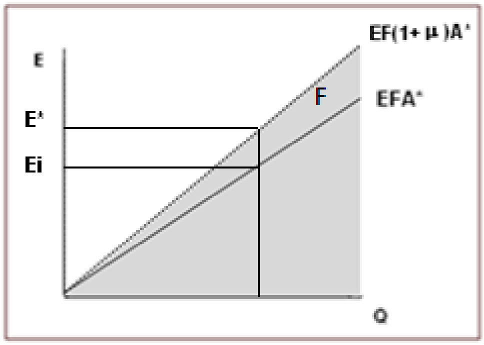


Figure 12: Deviations in a Relative Target Scheme

Consequences of a Lax Policy



Comparing Basic Structures (C & T vs PSR)

Production & Emissions

In the discussion about the performance of a Cap & Trade with an absolute target and the Performance Standard Rate the error terms are central for the market structure implications. As we have seen before, in the case that the social optimum values for abatement and production level are the same as the forecasted, μ and ω take a cero value and the FOC take the following forms:

C & T (Absolute Target)

$$1)\quad c_a = -Pp\llbracket(E\rrbracket_a)$$

$$2)\quad c_q = -Pp\llbracket(E\rrbracket_q)$$

Performance Standard Rate

$$r_a = -P_p\left({}^{E_a}/_{EF} - q^*\right) \quad 3)$$

$$r_q = -P_p\left({}^{E_q}/_{EF} - \alpha^*\right) \quad 4)$$

From what can be seen in equations 11-14, abatement and production level under PSR will be higher than under a C&T scheme given that the terms in excess in equations 13 and 14 work as subsidies.

In the presence of deviation from the social optimum, μ and ω take a non-cero value, PSR will be higher (smaller) than under a C&T scheme if (E_a/EF- (1+μ)q*) is bigger (smaller) than (E_a-EF($\mu + \omega + \mu\omega)q)$. In production, the quantity under PSR will be higher (smaller) than under a C&T scheme iff (E_a/EF- (1+μ)q*) is smaller (bigger) than (E_a-EF($\mu + \omega + \mu\omega)q)$.

A further effect of the error terms is that, in their presence, marginal costs become different for each firm increasing the burden for some and being more relaxed for others.

Entry Barriers

Absolute targets might provide barrier against entrants, especially when the established firms receive their permits for free. Entry barriers might be raised if capital markets do not work perfectly and therefore entrants have to pay higher capital costs than the established firms. With relative targets, entrants have to meet the same relative cap as the incumbent firms therefore they are not faced with any entry costs in excess.

This is an important result for market efficiency since competition is not hindered by the PSR market structure. PSR will also give incentives for new competitors with lower emissions per unit to enter.

Uncertainty

With emission trading under an absolute cap, the emission target is given, regardless of the a priori uncertain level of economic growth. Instead, the costs related to achieving the emissions target are uncertain. A higher economic growth than expected and the related production level will result in higher costs which will have to be made in order to achieve the target.

With a relative cap, the emission result is uncertain but the production costs to achieve the fixed target are known for each firm. Higher than expected economic growth increases the level of emissions above those expected but do not increase the cost of compliance. Consequently, firm costs will not rise as much as in the case of the absolute cap and for firms there is more certainty on the costs they will bear per produced unit.

Table X. Summary of the differences between an absolute and a relative target

Type of Target	Absolute	Relative
Production and Abatement	***Restricts production*** in the case of a ***non-expected economic expansion*** of the industry and ***restricts abatement*** in the case of an ***un-expected economic contraction.*** Because the more firms produce the higher abatement per unit is needed to comply, ***abatement through sales composition*** can ***compensate less*** for technological improvement abatement	Un-expected ***economic expansions or contractions*** of the industry ***do not interfere*** with its performance. Because abatement per unit needed to comply is constant, ***abatement through sales composition*** can ***compensate more*** for technological improvement abatement
Entry Barriers	***Might provide barrier against entrants***, especially when the established firms receive their permits for free and /or capital markets do not work perfectly	***Entrants*** have to meet the same relative cap as the incumbent firms therefore they are ***not faced with any excessive costs***.
Uncertainty	Magnitude of ***emission reduction is certain***, but the ***cost of compliance*** is ***uncertain*** (depends on the level of production)	Magnitude of ***emission reduction is uncertain*** (depends on the level of production), but the ***cost of compliance is certain***

Attribute Based Performance Standard Rate

Performance Standard Rate (PSR) may be modified from its basic structure by generating various categories and asking different relative emission targets for each one of them. This type of scheme has been adopted in the structure of some fuel economy standards for new vehicles. Two variables that have been used as the attribute according to which different fuel efficiency requirements are set are motor weight[1] and foot-print[2].

As it can be seen in the figure below, emissions grow with the total quantity produced as in the basic structure case. The only difference is that now we have different abatement levels set accordingly to W (the attribute variable chosen), as can be seen in the right hand side part of Figure 15 below.

---

[1] Definition:

[2] Definition:

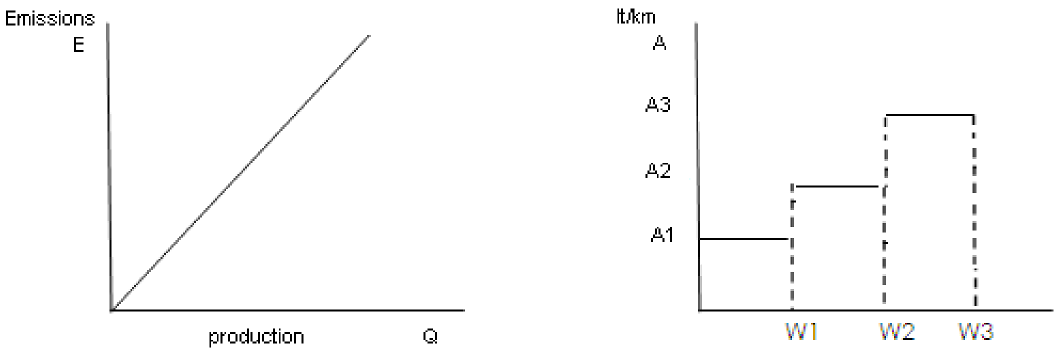


Figure 15:Attribute Based Performace Standard Rate

In both a motor-weight or foot-print based Performance Standard Rate, it is assumed that at higher levels of the attribute, it is harder for a vehicle to be more efficient. Therefore, as it is shown by Figure 15 firms with a vehicle fleet that has a lower motor-weight or foot-print average will fall into a category for which a lower average of fuel consumption would be allowed (lt/km). Contrastingly, firms with a vehicle fleet that has a relatively high motor-weight or foot-print average will fall into a category which allows a higher average of fuel consumption.

An attribute based Performance Standard Rate can have different number of categories. The possible number of categories starts from two and ends in one for each level of motor-weight/foot-print inside the range that goes from the lighter/smaller vehicle in the market to the heaviest/biggest vehicle. In this last case, the relation between the average fuel consumption and the attribute is continuous instead of by tiers of a range of attribute levels per each fuel consumption requirement. Figure 16 illustrates this example.

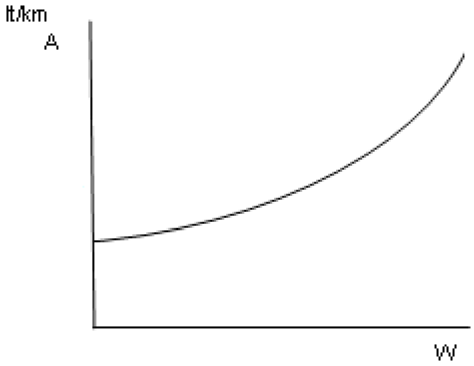


Figure 16: Continuous Attribute Based Performance Standard Rate

We will discuss a model that aims to portray the characteristics of a PSR by Tiers Structure and compare it with the basic systems. The firm will have the problem as to maximize its benefits as shown in equation 15:

$$1)\quad Max\ \pi = \sum P_i\, q_i - \sum c(q_i, a_i) - \sum [\![Pp \left({}^{E(q_i, a_i)}\!/_{q_i EF} - a_i{}^{*} \right) q_i]\!] - \sum Pp(a_i{}^{*} - a_i{}^{e}) q_i$$

With c_q, c_{qq}, c_a, E_q, $E_{aa,} > 0$, $E_a < 0$

where, $c(q_i,a_i)$ refers to the production cost, $PP(E(q_i,a_i)/q_iEF - a_i{}^*)q_i$ is the social cost when valued at the social optimal level $a_i{}^*$, and the last term refers to the costs associated to the difference between the forecasted actual emission standard $a_i{}^{e}$ and the optimal level $a_i{}^*$.

Considering the variation/error from the social optimal in the expected abatement level:

$$2)\quad \Delta a_i^e = \mu a^{*}$$

$$3)\quad Max\ \pi = \sum P_i\, q_i - \sum c(q_i, a_i) - \sum [\![Pp \left({}^{E(q_i, a_i)}\!/_{q_i EF} - a_i{}^{*} \right) q_i]\!] - \sum Pp(\mu_i a_i{}^{*}) q_i$$

A firm will buy permits if its emissions per unit of output (total emissions E divided by the total quantity produced q) are higher than the allowed relative target a^e.

FOC

$$4)\quad c_{ai} = -Pp \left({}^{E_{ai}}\!/_{EF} - (1 + \mu_i) q_i{}^{*} \right)$$

5) $$P_i = -c_{qi} - Pp\left(E_{q_i}/_{EF} - (1+\mu_i)a_i^{*}\right)$$

The error terms are tier related and therefore they tend to add up among tiers. To see what would happen inside the firm among two tiers in the abatement level, we allow i= 1, 2.

6) $$c_{a1} = -Pp\left(E_{a1}/_{EF} - (1+\mu_1)q_1^{*}\right)$$

7) $$c_{a2} = -Pp\left(E_{a2}/_{EF} - (1+\mu_2)q_2^{*}\right)$$

Assuming that marginal costs of abatement in equilibrium should be equal among tiers and that there is an overestimation of the optimal permit level of emission in tier 2 ($\mu_1 = 0$, and $\mu_2 > 0$), we find that the marginal cost of abatement will be less in the second tier than in the first. The firm will be better up moving their abatement efforts and production towards the products on the second tier where it is easier for them to achieve the relative target until marginal costs of abatement equate among tiers.

8) $$P_1 = -c_{q1} - Pp\left(E_{q1}/_{EF} - (1+\mu_1)a_1^{*}\right)$$

9) $$P_2 = -c_{q2} - Pp\left(E_{q2}/_{EF} - (1+\mu_2)a_2^{*}\right)$$

($\mu_1 = 0$, and $\mu_2 > 0$),

Given the overestimation of the optimal permit level of emissions in tier 2, the relative price diminishes compared to tier 1. Thus even though the absolute price for tier 2 may be higher, the existence of an error in the optimal level of allowed emissions may favor the demand of a tier. To illustrate this point, in terms of fuel economy standards, the presence of errors in the optimal level not only do not provide incentives to switch to smaller vehicles but even provide incentives to sell a greater number of bigger vehicles.

III. Temporal Dimension (Banking and Borrowing)

Following Tietenberg (2006), Greenhouse Gases are specifically suitable for emission trading design since they are considered as uniformly mixed accumulative pollutants. That is, pollutants that accumulates in the environment because their injection rate exceeds the assimilative capacity. Location and timing does not matter since it is the cumulative level that does. The policy objective is to ensure that the cumulative level of emissions E would never be higher than some ceiling $\bar{\mathbf{E}}$.

The unimportance of emissions timing has several immediate design implications. Permit dating is not important. With undated permits the design automatically allows unlimited banking and borrowing. For a uniformly mixed accumulative pollutant a common form of the pollution target-emission rate relationship can be described as:

1)

$$\bar{E} \geq b_0 + \sum_{j=1}^{J}\sum_{t=1}^{t}(e_{jk} - r_{jk})$$

This formulation assumes no depreciation of the pollution stock over time, that is no assimilative capacity. The cost-effective allocation in this context is the one that has the lowest associated present value of control costs among all those allocations that satisfy the pollution constraint. The decision problem can be symbolically stated as:

2)

$$min \sum_{t=1}^{T}\sum_{j=1}^{J}\frac{C_j(r_{jt})}{(\llbracket 1+\rho)\rrbracket^{t-1}}$$

Subject to:

3)

$$\bar{E} \geq b_0 + \sum_{j=1}^{J}\sum_{t=1}^{t}(\bar{e}_{jk} - r_{jk})$$

And

$r_{jt} \geq 0$ t=1,…,T

Where ρ is the discount rate used to translate future costs into their present value equivalents and T is the length of the planning horizon. Obtaining the Kuhn Tucker conditions in order to derive the cost-effective allocation we obtain that they shall satify:

4)

$$\frac{\partial C_j(r_{jt})}{\partial r_{jt}} \frac{1}{([1+\rho)]^{t-1}} - \lambda \geq 0$$

j= 1, ... , J t=1,…,T

5)

$$r_{jt}\left[\frac{\partial C_j(r_{jt})}{\partial r_{jt}} \frac{1}{([1+\rho)]^{t-1}} - \lambda\right] = 0$$

j= 1, ... , J t=1,…,T

6)

$$\bar{E} - b_0 - \sum_{j=1}^{J} \sum_{t-1}^{t} (\bar{e}_{jk} - r_{jk}) \geq 0$$

If T is long enough, as t increases eventually a year is reached in which the ambient constraint (equation 6) becomes binding and allowable emissions cease from then on. The marginal control costs at that point are those associated with complete control; they are no longer necessarily equalized.

7)

$$\lambda\left[E - b_0 - \sum_{j=1}^{J} \sum_{t=1}^{t} (\bar{e}_{jk} - r_{jk})\right] = 0$$

$r_{jt} \geq 0$; $\lambda \geq 0$ j= 1, ... , J
t=1,…,T

The cost-effective control level shall satisfy:

8)

$$\frac{\partial C_j(r_{jt})}{\partial r_{jt}} = \frac{(1+\rho)\left(\partial C_j(r_{jt-1})\right)}{\partial r_{jt-1}}$$

j= 1, ... , J t=1,…,T

This condition implies that in a cost effective allocation, marginal pollution control costs would rise over time at rate ρ and the amount emitted would decline over time. This condition is known as Hotellings rule. In each time period, the marginal costs of control would be equalized across all sources. Across time, the present value of marginal costs shall be equalized.

In theory, banking and borrowing allows the Hotelling´s rule condition for efficiency to bind. Banking and borrowing favors the present value of marginal costs to be equalized. This is a necessary condition to achieve the least marginal cost across time. In order to make this statement a bit clearer we shall imagine that banking and borrowing has the same efficiency effect as allowing a specific firm to trade with itself in a different period in time.

Let's say firm 1 has an excess of permits in time 1 and knows that in time 2 standards are going to be more stringent. Firm 1 may diminish their fulfilling costs by banking some permits if they find this option less costly than trading them in time 1. The same is true while borrowing, if a firm finds that in Time 2 costs to fulfill the standard will be less than in time 1 then they may find less costly to borrow permits from Time 2.

In practice the presence of Hotelling's rule in a market depends on whether the standard is seen as forever lasting (perdurability) and the institutional arrangements that allow banking and borrowing are seen as enforceable. In case that the perdurability of the standard is not institutionalized and no limits to borrowing and banking are set, firms will tend to borrow from further periods of time until the standard quits functioning. The environmental goals would not be met unless the standard is not understood as long lasting and enforceable. Limits to borrowing must thus be set and enforceability and perdurability promoted.

On the other hand limits on borrowing and free banking, on an uncertain world on further periods of time´s costs will make firms to find optimal to bank permits in excess. This institutional setting may foster early implementation, increase environmental performance of the standard but will at the same time increase firms´ costs.

It must also be noted that since monitoring of the standard fulfillment is normally performed ex-post some banking and borrowing is required in order to improve firms´ certainty on the standard realization.

IV. Lessons and Experiences: What has previous implementation told us in the frame of the four necessary conditions to build an efficient tradable permits market?

First of all, previous experience in tradable permit systems has shown that this instrument holds real potential for successfully bringing cost-effectiveness to the compliance of an environmental regulation. However, the implementation of various tradable permit systems (both credit and cap-and-trade programs) have shown that any attempt to design a market should include a detailed analysis of the circumstances in which such market will be constructed. This is because conditions regarding the characteristics of firms to be regulated, the particular relation among them in different markets, the legal frame that surround the regulation, and the institutional capacities of the regulator, all have implications on the likeliness of building an efficient tradable permit/credit market. More specifically, on the likeliness of achieving compliance, low transaction costs, perdurability, and competitiveness.

The objective of this chapter is to revise the experience of regulators that have implemented tradable permit systems in different contexts in order to 1) recognize circumstances and characteristics that are directly related with the success or failure of fostering each of the necessary conditions 2) Identify possible structures and procedures that have been used in order to face major obstacles for the achievement of these conditions.

a. Compliance: Monitoring and Enforcement

In the first chapter of this document we have explained the importance of compliance in making a tradable permit system possible, as well as some of the theoretical characteristics that adequate monitoring and enforcement (the two main tools for compliance achievement) have. In this first subsection we revise specific lessons that tradable permit systems such as RECLAIM, the SO2 Trading Program and the Alaskan IFQ have taught, in the practice of monitoring and enforcement.

Monitoring

Empirical experience for tradable permit markets has emphasized the fact that building effective monitoring mechanisms is a complex task. “Effective monitoring systems are composed of data, data management, and verification components” (National Research Council 2002). The first lesson to be learned in this matter is the importance of using available information technology in order to meet the goal of designing programs that reach adequate monitoring. Both the SO2 Trading Program (and Acid Rain Program) and the Regional Clean Air Incentives Market (RECLAIM) have brought light in the building of such an information technology system.

In the SO2 Trading Program, information gathering and diffusion is done through the web and even special software (a windows-based program called Allowance Trading System, or ATS) has been developed for these means. In order to gather information, firms are requested to set up CEMS, which stand for continuous emission monitoring systems. “All CEMS must be in continuous operation and must be able to sample, analyze, and record data at least every 15 minutes and then reduce the data to 1-hour averages” (Stranlund et al. 2002). The information is processed by the Data Acquisition Handling System after each firm`s CEMS sends it to the EPA. The agency has also created, within ATS, systems that allow electronic transactions.

To sum up, the SO2 Trading Program has proved that an advanced and user friendly information technology system constitutes the structure that will allow accurate and timely information gathering and diffusion. It has also confirmed that a monitoring system with those characteristics will facilitate transactions to be supervised and increase the likeliness for such transactions to happen. Kruger describes the importance of the technology information by saying that “(t)he implementation backbone of the Acid Rain Program is the data systems that track allowances holdings. EPA processes about 90 percent of allowance transactions within 24 hours of receipt, using just two Acid Rain staff” (Kruger et al, 2002).

The case of RECLAIM is also interesting in the subject of technology information systems. Just as the SO2 Trading Program, RECLAIM has placed emphasis in the development of accurate and advance technology devices through which monitoring mechanisms can be put into action. However, a remarkable characteristic of such mechanisms is that monitoring requirements vary depending in the type of firm. For example, “NOx sources are classified into four categories depending upon emissions levels: major sources, large sources, process units, and equipments. SOx sources are classified into three categories: major sources, process units, and equipment” (Stranlund et al. 2002). All major sources are required to set up CEMS as their monitoring systems, while other sources are allowed to place monitoring systems that are less accurate, but cheaper, than CEMS. Also, only major sources need to use a Remote Terminal Unit (RTU) to transfer information to the Air Quality Management District. Other emission measuring devices as fuel flow meters are also required depending on the type of source; only large NOx sources and both NOx and SOx process units and equipment sources are obliged to use fuel flow meters.

The different categories of firms by level of emissions has allowed the RECLAIM program to make large emitters relatively more accountable for the costs and accuracy of the monitoring mechanisms needed for the system to work. Therefore, the RECLAIM program is a good example of a way to build a monitoring system that combines technological sophistication with the aim of avoiding the costs of the program to be too high for the participants as a whole.

A second lesson that can be learned from the SO2 Trading Program has to do with the adequate design of the formats that will be used to gather information; in other words, with the correct identification of the information that should be collected and disseminated. In the SO2 Program emissions data has been selected, and “EPA publishes summary reports in the web that include SO2, NOx, CO2, and other information” (Kruger et al, 2002). As from 1999, quarterly certified emission information began to be posted in the EPA website. As Kruger points out, emission information has proved to be helpful for participants in the program to take better informed decisions.

The SO2 Trading Program has also successfully recognized the allowance data that is necessary for the well functioning of the market. Information is posted in the web site and it includes the number of transactions, the type of transaction (e.g. auction, private), name of the buyer and his/her account information, name of the seller and his/her account information, confirmation information for the transaction, and the serial numbers of the allowances that were traded.

Enforcement

A central message from preceding practice in the implementation of tradable market systems is that "Penalties should be commensurate with the danger posed by noncompliance" (National Research Council, 2002). Chapter one of this document has explained how the theory behind tradable permit systems establishes that when penalties are too low, the marginal cost of non-compliance is smaller than the marginal cost of compliance and this leads to firms not to comply. Nevertheless, practice has brought attention as well to the link between penalties that are too high and non-compliance.

The fact is that, when it comes to fostering effective enforcement, the regulator must consider that the level of abatement required for complying, the magnitude of penalties and the procedures of payment of such penalties by the actors must match its real ability of enforcement. In other words, if a standard is so strict (or existing allowances are so few), that too many actors will have to pay fines, the regulator could find it beyond its capacities to ensure that most non-compliers actually pay and reduce the extent in which it can prove actors that compliance is necessary. Also,"(u)nrealistically high penalties… (or unclear payment procedures)… are likely to consume excessive enforcement resources as those served with penalties seek redress through the appeals process" (National Research Council, 2002). In the same subject of reducing the resources that are wasted in legal procedures (both by the regulator as by the sources),"(c)riminal penalties should be reserved for falsification of official reports and the most serious violations" (National Research Council, 2002).

In the aim of building an enforcement system that can be effectively put into practice, the Alaskan IFQ[3] program (a program that aims at the limitation of fishing endangered species), has distinguished between two levels of non-compliance in order to establish its penalties. When a source has exceeded by less than 10 percent its IFQ level, the sanction is simply that the excess fishing should be subtracted from the IFQ level of the next period. Instead, "(o)verages greater than 10 percent are considered a violation and are handled by enforcement personnel" (National Research Council, 2002). This procedure could be seen as a type of quota borrowing mechanism with limited periods and limited amount.

In the case of the SO2 Trading Program, the regulating agent has been stricter with penalties than in the previous example because 1) there is not a non-compliance level under which sources will not pay a monetary sanction 2) every non-compliant source faces both a monetary sanction and also has to off-set additional emissions during future years. However, strict penalties in this case have not been an issue in accomplishing an efficient level of enforcement because "the unit penalties in the SO2 program are unique in the fact that they are applied automatically" (Stranlund et al. 2002); this reduces the costs and efforts required for the regulator to enforce penalties even if the number of non-compliers is high[4].

---

[3] Individual Fising Quota

[4] The penalty was set at $2000 per ton of excess emissions in 1990, as is indexed to inflation.

In the case of RECLAIM, a particular characteristic regarding enforcement is that penalties vary in accordance to the changes in permit prices. This is done with the aim of preventing the incidence of non-compliance to rise at a high-speed when there is a sudden rise in the permit price. In business as usual circumstances "(n)on-compliant facilities may face penalties of up to $500 for every 1,000 pound exceedance for every day the exceedance persists…when credit prices are high, if the annual average price of credits per ton reaches $8000, then the $500 penalty can be applied to every 500 pounds of excess emissions, effectively doubling the available penalties" (Stranlund et al. 2002). As Stranlund explains, the procedure of case-by-case sanction appliance implemented in RECLAIM creates uncertainty for the actors as to which are the non-compliance consequences they face. This uncertainty makes it difficult to calculate the extent in which sanctions will dissuade firms from choosing not to comply, but it also makes them base their decisions in their expected sanction (which changes when circumstances in the market change) rather than in a fixed amount that is never adjusted. Still, the fact that penalties are not automatically charged has increased the likeliness of deficient enforcement in this program.

b. *Transaction Costs*

Previous experience has shown that transaction costs do have an important responsibility in diminishing the cost-effectiveness potential of a tradable permit market. Both these facts are conclusions in the analyses of the Regional Clean Air Incentives Market (RECLAIM) and the Lead Phasedown Program (the first one carried out by Lata Gangadharan and the latter by Suzi Kerr and David Maré). Previous experience has also helped classify and clarify where and how transaction costs rise during the implementation and daily practice of a tradable permit system. Alghough, as Kerr and Meré assert, "the relative weight of each…(type of transaction cost)… will vary with the quantity traded, number of trades, and length of trading relationship" (Kerr and Maré, 1998). In a general sense, it is important to have in mind that the first years of any program will present the highest transaction costs, because in later years previous information and trading relations will help to diminish them. Also, all types of transaction costs will be translated into a diminishing probability of trading in the market and a loss in the cost-effectiveness of the program. For example, Gangadharan calculated the in RECLAIM transaction costs meant 32% less probability that trading occurred. In the Lead Phasedown Program, even when the market proved to be simple and active, Kerr and Maré estimated that the cost-effectiveness lost because of transaction costs was about 10%-20%.

The aim of this subsection is to 1) revise the types of transaction costs that have been most problematic in existing tradable permit systems and the extent of their consequences 2) the conditions under which these transaction costs have been of a greater magnitude 3) review some ways in which the design of mechanisms through which these markets work have been more, or less, successful in minimizing the existence of transaction costs.

Market information and search costs

The first type of transaction costs that have been found to hinder trade in a permit market are those related with actors being able to access information regarding the market (e.g prices, market rules, conditions to base their optimization plans), and information about their potential trading partners (e.g. reputation, willingness to trade, financial situation, etc). The study in the RECLAIM case shows that having a decentralized information mechanism where each source is expected to search for trading partners and market information by itself tends to have greater transaction costs than those like the Acid Rain Program in which sources are "helped in this respect by centralized auctions arranged by the regulatory authority" (Gangadhara, 2000).

The RECLAIM experience also demonstrated that "search costs and information costs are high, as the firms do not participate in similar input or output market" (Gangadhara, 2000). Contrary to this case, in the Lead Phasedown Program, in which "the level of trading between firms far surpassed those observed in earlier environmental markets,… the firms were homogenous (they were all refineries), and this meant that they found it easier to search for trading partners" (Gangadhara, 2000).

Nevertheless, even when in general the Lead Phasedown Program had less market information and search costs, it shows important differences between those costs faced by the different type of refineries that participated in it. In this respect, "(l)arge refineries, large companies, and companies with many refineries, achieve high levels on all measures of efficiency. Transaction costs seem to most affect refineries in small companies or in companies with few refineries, with trading efficiency generally under 60 percent and dropping frequently below 50 percent… (Therefore)…, when most potential traders in a planned market are small, non-integrated, and unsophisticated, the market is less likely to be efficient."(Kerr and Maré, 1998).

Finally, the study of the Lead Phasedown Program also brings light to transaction costs associated with firms optimizing their production plans. Kerr and Maré calculated the difficulties that actors had in this respect by analyzing the trading reports from the the General Accounting Office records. They found that "around 20% contained errors that did not lead to commercial advantage and thus were probably not deliberate" (Kerr and Maré, 1998).

Regarding mechanisms that have proven to minimize market information and search transaction costs Tietenberg points out price transparency, which means making prices public and which was done in the Acid Rain Program. He also states that, by "providing mechanisms for sharing information on available technologies…(the regulator)…can reduce duplication of effort for larger firms and provide a larger menu of control options for smaller firms" (Tietenberg,T. 2006). In this same subject, Gangadhara proposes the development of brokerage services (although brokerage costs would have to be low enough not to diminish trading incentives); the introduction of an electronic bulletin board system where actors could publish terms of trade in order for potential trading partners to have access to them; the periodically distribution of market statistics summaries from the regulator to the participating actors; and the use of a well-designed centralized trading system when possible.

Cost of private information about the validity of the rights being traded

As Kerr and Meré explain, the case of the Lead Phasedown Program showed that if the regulator is not able to prevent invalid permits to be traded in the market, firms have to face costs from researching the validity of permits and this too disincentives their willingness to trade. The problem of invalid permits was caused, in this case, because the regulating institution (the Environmental Protection Agency) verified traded permits instead of having a verification process that would certify such permits before trading took place. Both deliberated and accidental trading of invalid permits occurred as a consequence of such a procedure. It is therefore necessary for the regulator to establish an efficient and clear permit certification procedure in which firms can rely. Also, as Gangadhara points out, in the aim of minimizing costs from private information search about the validity of the rights being traded, it is important for the regulator to have an active role in the monthly (or periodical) analysis of transaction data to verify that the validity of the permits being traded in the market.

Costs of negotiation

The costs of negotiation refer to those associated with the process that firms have to follow after finding a trading partner in order to complete a trading transaction. The Lead Phasedown assessment shows that larger companies prefer to negotiate with other large companies. Because of this reason, relatively small companies often find it hard to access the market because of the relatively high negation costs that other firms face when trading with them. This same study concludes, in this respect, that the problems with negotiation costs are less likely the more liquid the market is. A regulator can minimize negotiation costs by simplifying trading requirements. By making such requirements more accessible for smaller firms the regulator can homogenize the terms of negotiation between big and small firms, and in this way grant easier access to the market for these last ones.

Costs of confidentiality release

The costs of confidentiality release are the costs generated when, in order to participate in the permit market, firms have to make public information regarding future incomes, technological capacity, etc., that otherwise would be kept confidential. Kerr and Meré found that in the case of the Lead Phasedown Program, in order to avoid these costs, some companies engaged mostly in trading only between their own refineries. "A manager in one company said that the price information was extremely sensitive and consequently 95% of their trading was done internally"(Kerr and Maré, 1998).

It has been said that when firms compete in similar input and output markets it is likely for search and market information costs to be reduced. However, having firms competing in the same permit and output market could be problematic when it comes to costs of confidentiality release. This is because in the extent to which the information released for the well functioning of the permit market can influence the behavior of other firms in the output market, the more costly it results for firms to give away this information. In the aim of reducing confidentially release costs it is important for the regulator to differentiate between the information that should be made public for the firms that participate in the permit market and the information that is needed by the regulator but that is not necessary to make public.

c. Perdurability

As seen before, perdurability of an environmental regulation and its associated tradable permit market has to do with the certainty upon the policy and the market. Based on previous experience in the implementation of tradable permit systems, two important lessons can be pointed out in this subject. This subsection has the objective to explain these two lessons.

The first lesson that has been demonstrated by preceding designs has to do with setting a governance structure that provides certainty upon the institutional legitimacy of both the policy and the market. "For those resource regimes in the United States, it is common for the goals to be set by the government (either at the national or state level) and considerable "top-down" management to be in evidence" (National Research Council, 2002). For example, although the RECLAIM regulation was set by a regional authority, EPA supervision has meant that the system is subject to national environmental control. In the case of trading programs in Fisheries, local creation and managing is always supported by national supervision from the Secretary of Commerce.

A second point concerning perdurability is associated with choosing an adequate legal framework and definition for the permits being created by environmental regulations that use the tradable permit market instrument. Designers of tradable permit markets had to make sure that permit holders were granted the rights associated with the holding of such permits. They knew that "confiscation of the rights granted under these permits could undermine the entire process…(of credit trading)" (National Research Council, 2002). The legal inconsistency that was created was that, in a start, permits were aimed at, granting permit holders with private use of resources. As stated in many countries' applicable legislation (including Mexico), the "public trust doctrine" establishes that certain resources are public and therefore, their privatization is not possible; under such circumstances, permit market designers could not tie the rights to dispose of such resources (abusus), for being reserved to public agents, with the rights to trade or commercialize (usus and fructus), as granted under the permit, for being reserved to permit holders. In order to solve this legal flaw, designers have opted to modify the initial definition of permits and made sure that these do not stand for full property rights over public resources. For example, according to the title of the U.S. Clean Air Act dealing with the sulfur allowance program: "An allowance under this title is a limited authorization to emit sulfur dioxide…" (National Research Council, 2002) implicitly entering thereby into some kind of public-private usufruct agreement.

Competitiveness

As it has been seen, fostering the three first conditions for an efficient tradable permit market to exist is a hard task and even best practices leave some margin where non-compliance, transaction costs and a certain lack of perdurability can exist. Also, a complicated issue is that the input and/or out-put markets in which participants of the permit market compete may already have some competition problems which can transmit and even be magnified by the existence of the permit market. Under this context, market power is likely to become a problematic issue when offering a tradable permit market as an instrument that can support a policy. "The concentration of permits in the hands of a few either can reduce the efficiency of the tradable permits system or can be used as leverage to gain economic power in other markets" (National Research Council, 2002).

The avoidance of such concentration is therefore a very important issue if a tradable permit system is being proposed as an instrument to better implement an environmental regulation. In the effort to prevent market power from emerging in a tradable permit system, designers and regulators have used different kinds of approaches. Some actions confront the problem by targeting accumulation directly, while others center in providing alternative options for players competing with the market power actor. This section intends to provide some examples of both these types of actions.

Accordingly to the approach of targeting accumulation directly, New Zealand fisheries have set limits to the percentage of quotas (from the available quotas in the market) that each fishery can accumulate. The limit has ranged between 20% and 35%. Fisheries in Iceland have used the same approach but have set different limits depending the specie that is being fished (10% for cod and 20% for other fish) (National Research Council, 2002). Another measure that targets accumulation of permits/quotas directly is the zero-revenue auction, which has been utilized in the SO2 Trading Program. A zero-revenue auction means that actors are required to put certain percentage of allowances for sale once a year. In the other hand, in the case of measures that focus in providing alternative options for players competing with the market power actor, the SO2 Trading Program has also set precedence by having a reserve of allowances. This reserve can be sold in the case in which one or various actors are not willing to sell permits in the market. (National Research Council, 2002).