Optimal Design of Permit Markets with an *ex ante* Pollution Target

Sergey Rabotyagov
Center for Agricultural and Rural Development and
Department of Economics
Iowa State University
Email: rabotyag@iastate.edu

Hongli Feng
Center for Agricultural and Rural Development and
Department of Economics
Iowa State University
Email: hfeng@iastate.edu

Catherine L. Kling
Center for Agricultural and Rural Development and
Department of Economics
Iowa State University
Email: ckling@iastate.edu

Draft—Please do not cite!

Selected Paper for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, CA, July 23-26, 2006.

Draft—Please do not cite!

Optimal Design of Permit Markets with an *ex ante* Pollution Target

Abstract

The optimal pollution permit trading system is examined when the regulator, faced with incomplete information on firms' abatement costs and delivery coefficients, seeks to minimize expected total abatement costs to meet an *ex ante* pollution target. Intuitively, we find that the optimal trading ratio and permit cap are set such that there will be more pollution when abatement costs are high and less pollution when abatement costs are low. Surprisingly however , even when the delivery coefficients are known with certainty, the optimal trading ratio will not necessarily equal the delivery coefficient, nor will it be optimal for the total permit quantity to equal the given pollution target. Instead, the trading ratio will tend to be larger when there is negative correlation between firms' abatement costs and/or there is positive correlation between abatement costs and the delivery coefficient. The result that the optimal permit trading system under these circumstances depends on the regulator's information on firms' abatement costs contrasts sharply with a previously well-established attribute of permit trading systems: no information on firms' abatement costs is needed in order to design an optimal policy and achieve the least cost of reaching a pollution target (*ex post*). Our results demonstrate that whether an *ex ante* or *ex post* pollution target is used has fundamental implications for the design of the permit trading system. Finally, while not descriptive of all pollution problems, the class of pollutants for which this model applies is large.

Key words: delivery coefficient, *ex ante* pollution target, *ex post* pollution target, permit trading, total permit cap, trading ratio.

1. Introduction

A highly celebrated property of emissions trading markets is that decentralized decisions made by firms will achieve a preset emissions target at the least possible cost and no information on the firm's abatement costs is required to achieve this outcome (Baumol and Oates; Montgomery). Montgomery demonstrates that this property extends to the class of non-uniformly mixed pollutants, pollutants whose damages differ based on their location. He shows that if the regulatory authority allows firms to trade emissions according to the ratio of delivery coefficients (the effect that a source's emissions have on resulting pollution concentrations) and sets the pollution cap equal to the desired pollution standard, the least cost property is retained. No information on firms' costs is needed to achieve the regulators objective of cost minimization.[1]

The basic model underlying these findings assume that the regulator is interested in minimizing the cost of meeting an *ex post* environmental standard. While *ex ante* uncertainty regarding firm's abatement costs is commonly used to motivate the attractiveness of the permit system, the objective function is typically specified in *ex post* terms – a given environmental target invariant with respect to realizations of any sources of uncertainty. As has long been recognized, characterization of the objective function in this way requires that the pollution control level is independent of the actual realization of costs --- no tradeoff between abatement costs and benefits (pollution levels) is permitted[2].

[1] The total permit quantity can be set at the socially efficient level, a legally mandated requirement, or any other level deemed appropriate by the regulator.
[2] An important exception to the use of *ex post* objective functions occurs in the literature related to non point source effluents (Griffin and Bromley; Segerson; Malik, Letson, and Cruthfield, 1993; Horan et al., 2001; Horan 2001; Horan and Shortle, 2005). Due to the uncertainty assumed in delivery coefficients, these authors have specified the regulators objective function in *ex ante* terms, though they assume that abatement costs are known with certainty.

In this paper, we study the optimal design of a permit trading system when the regulator is uncertain about the firms' abatement costs and specifies her objective function based on minimizing expected costs subject to meeting an expected emissions level. We called this an *ex ante* target. Our model is applicable to cases where the regulator possesses some information about the firms' abatement costs, but is uncertain about their magnitude either due to existence of genuine aleatory uncertainty in abatement costs, or for reasons of asymmetric information, where regulator's uncertainty is epistemic in nature.

Whether the regulator has (or should have) the freedom to design a permit market that allows the aforementioned flexibility is an empirical question that will have a case by case answer. However, there are many examples where averages over time or space define standards. Examples include carbon monoxide (with both an 8-hour and 1-hour average standard), nitrogen dioxide (an annual arithmetic mean), ozone (1- and 8-hour averages), lead (quarterly average during the phase-out), and sulfur dioxide (annual means, a 24- hour average and a 2-hour average) (http://www.epa.gov/airs/criteria.html). Examples from water pollution abound as well; values for arsenic, cadmium, cyanide, and selenium emissions in storm water under the National Pollutant Discharge Elimination System (a key regulatory program that regulates point sources of water effluents) trigger need for action only when the annual average exceeds the benchmark (http://www.epa.gov/npdes/pubs/msgp2006_factsheet-proposed.pdf).

Several striking findings emerge from our model. First, the optimal emissions cap does not necessarily equal the regulators emission target. Further, the optimal trading ratio depends on the moments of the uncertain costs as well as the delivery coefficients.

Surprisingly, even when the delivery coefficients are assumed to be known with certainty, it is not optimal to set trading ratios equal to the simple ratio of delivery coefficients --- the basic Montgomery solution. Instead, the regulator can lower expected costs by including some information on the uncertain abatement costs in the formation of the trading ratio[3].

These somewhat surprising findings come directly from the fact that our regulator's objective function is specified in *ex ante* terms: she minimizes expected costs subject to an expected pollution level. This allows the regulator flexibility that is not present when emission levels must be met with certainty.[4] In essence, this allows the regulator to consider, at least to some degree, the actual cost realizations of firms: if costs are unexpectedly high (low), the resulting pollution levels will be higher (lower) than they would be without this flexibility. Intuitively, once the regulator is interested in both costs and benefits, it becomes optimal for the regulator to design the system so that if costs are unexpectedly high (a big positive stochastic shock), higher than expected pollution levels are permitted. In considering this tradeoff, the regulator recognizes that the ultimate abatement levels chosen by firms will depend upon their cost realization and therefore the ultimate emissions level become stochastic from the regulator's perspective. By choosing the parameters of the trading program to be a function of the moments of the distribution of costs, the regulator can lower total expected abatement costs.

[3] That the optimal trading ratio depends on both the regulators information about costs and the delivery coefficients is consistent in spirit with the findings from Horan and Shortle, Horan, and Malik, Letson, and Crutchfield although we do not assume perfect information on costs.

[4] In this way, our model and findings are in the spirit of Roberts and Spence (1976) and Montero (2005) who each recognize that rigidity of a quantity mechanism may be socially costly. Roberts and Spence (1976) propose a penalty for exceeding the emissions cap, while Montero (2005) models incomplete enforcement to provide a softening of the quantity constraint.

One of the most interesting findings from our work is that when the regulators problem is framed as one of minimizing expected costs subject to achieving an expected emission level, it is no longer necessarily optimal to set the emissions cap equal to the desired emissions level. One is an *ex ante* concept (the desired emissions level) while the other is an *ex post* construct (the emissions cap). This can be viewed as a two-stage decision where in the first time period the regulator settles on a desired pollution target and then, based on the reaction functions of firms, sets the number of permits and trading ratio to implement the market.

In the next section of the paper, we present the basic model of firms' behavior under a tradable emissions program and the regulator's problem. In section 3, we examine the optimal permit market design under three different assumptions. First, we consider the case when the delivery coefficient is known. This provides results that contrast with the *ex post* standards studied in Baumol and Oates and Montgomery, highlighting the importance of using *ex ante* targets and objective functions. Second, we consider the important case when the delivery coefficient is uncertain. While this latter feature is typically viewed as a characteristic of non point sources, there are likely many point sources where the true impact of emissions from the source are known with less than perfect certainty such as air sheds where dispersion of particulates may depend on stochastic weather conditions. Final remarks and conclusions complete the paper in section 4.

2. The Model

Suppose there are two firms acting as sources of emissions and the environmental impacts of the two firms' emissions are not identical. Specifically, we assume that the impact of the first firm on the resulting pollution level is direct, and one unit of Firm 1's emissions increases the resulting pollution level by one unit. The impact of Firm 2 is described by the delivery coefficient d , that is, one unit of Firm 2's emissions increases the resulting pollution level by d units. Specifically, the total resulting pollution level is $e_1 + de_2$, where e_i for $i = 1, 2$ represents Firm *i*'s emissions. We model both the situation in which the delivery coefficient is fixed and known by the regulator, as well as a more realistic case where the delivery coefficient is random. In the latter case the regulator, however, knows the distribution of the delivery coefficient: the mean of the delivery coefficient is $E(d) = \mu$ and its variance is $Var(d) = \sigma_d^2$. The model lends itself to multiple interpretations including (1) two firms located spatially apart whose emissions contribute differentially to loadings at the receptor (Baumol and Oates), (2) two firms whose emissions contribute differentially to loadings for reasons other than spatial location, such as production process or concentration of emissions released, or (3) two firms, one of which is a point source and the other is a nonpoint source with an uncertain delivery coefficient.

The abatement cost function for Firm *i* is $C_i(e_i^0 - e_i; \theta_i)$, where, for $i = 1, 2,$ e_i^0 represents the initial (unregulated) emissions levels for firm *i* and $e_i^0 - e_i$ represents the abatement of Firm *i* after the implementation of a permit trading program. The abatement cost function is assumed to be increasing and convex in abatement, that is, $C_i^{'} > 0$ and $C_i^{"} \geq 0$. The parameter (θ_i) in the cost function captures the information uncertainty regarding the costs of pollution abatement on the regulator's side. We assume that the regulator has some, albeit incomplete, information on abatement costs. While throughout our model the formal depiction of regulator's uncertainty is unchanged, we can endow our formal modeling of uncertainty with two different interpretations: asymmetric information or stochastic information not revealed at the design stage of the permit market, but revealed at the time permit trading decisions are made. Formally, when making decisions, firms know θ_1 and θ_2 while the regulator knows only their distribution: the means (zero), variances (σ_1^2 and σ_2^2), and covariance, ($\mathrm{cov}(\theta_1, \theta_2)$). Furthermore, the regulator is assumed to know the covariances, if any, between the delivery coefficient and the cost parameters: $\mathrm{cov}(d, \theta_1)$ and $\mathrm{cov}(d, \theta_2)$. Such correlations may arise, for example, when weather affects the efficacy and cost of abatement as well as its spatial movement.

2.1. The regulator's problem

Our paper focuses on a permit trading program design for the case where the environmental goal is to reach an environmental target at the lowest costs when the target is set as an *ex ante* emissions level, rather than an *ex post* standard. [5] While we believe the conditions of uncertainty we model are representative of a broad variety of environmental pollutants, water quality provides a strong motivating example. Imagine there are two sources of effluent that enter a river: source 1 is a large "point" source that is located at the river's edge and source 2 is a "nonpoint" source that is located some distance from the river. Given the proximity of source 1 to the river, its delivery

[5] We focus on the design of permit trading programs in the context of cost-effectiveness for the same reason as typically provided in the literature. Pollution targets are often set by political processes as in the case of sulfur permit trading program or water quality trading programs (Horan and Shortle 2005) and in practice the social damage of pollutants is often unknown making cost minimization the most relevant policy approach.

coefficient is known with certainty to be unity whereas the nonpoint nature of source 2 means that the delivery coefficient is uncertain. Assume that rainfall variability drives the uncertainty in the value of d. Finally assume that the abatement costs of both sources are unknown to the regulator.

2.1.1. Ex ante and ex post pollution targets versus total pollution permits

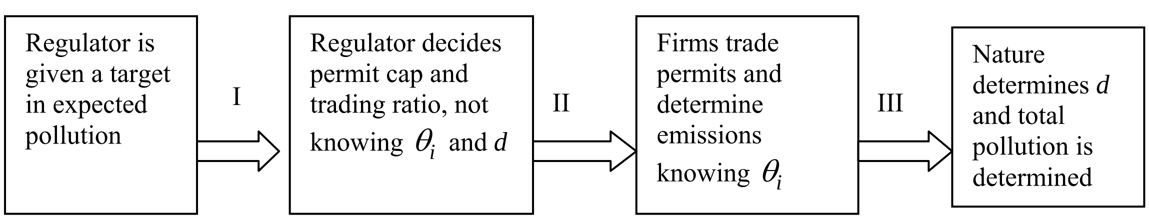
Since cost minimizing firms equate marginal abatement costs with permit prices to choose their emission levels, once uncertainty is introduced into the cost functions, there is uncertainty in emission levels and it is necessary to clearly differentiate between *ex ante* and *ex post* measures of emissions as well as other constraints that relate to the design of an emissions trading system. Four different measures of interest are:

$$(1)\quad \begin{array}{ll} \text{(Ex ante pollution constraint)} & E[e_1]+E[de_2]\le \bar{P}_{ante}, \\ \text{(Ex post pollution constraint)} & e_1+de_2\le \bar{P}_{post}, \\ \text{(Permit market constraint)} & e_1+te_2\le \bar{P}_{permit}, \\ \text{(Actual realization of pollution)} & e_1+de_2 = P_{actual}. \end{array}$$

If the pollution target is specified in an *ex ante* manner, the first equation in (1) describes the constraint and indicates that the expected pollution has to be less than or equal to a prefixed target ($\bar{P}_{ante}$). Under this constraint, the *ex post* realization of pollution levels can be greater or less than the target. In contrast, if the constraint is specified as *ex post*, the realized *ex post* pollution levels must be less than a pollution target ($\bar{P}_{post}$) in each realization. A third relevant constraint relates to the permit market. Here, t is the trading ratio for the emissions of the two firms --- 1 unit of Firm 2's emissions is equivalent to t units of Firm 1's emissions --- and $\bar{P}_{permit}$ is the total permit allocation, denominated in terms of Firm 1's emissions. Thus, this constraint requires that total emissions (weighted by the trading ratio) be less than or equal to the total pollution permits. Note that all three constraints are important in the regulator's decisions while the firms are only concerned with the permit market constraint. Finally, the last equation specifies the actual realization of pollution given firms' emissions decisions and the realization of the delivery coefficient.

With perfect information, there is no distinction between *ex ante* and *ex post* and we know from Montgomery that efficiency dictates that we set $t = d$, resulting in $\overline{P}_{ante} = \overline{P}_{\text{post}} = \overline{P}_{permit}$. That is, the three constraints collapse into one. However, when there is incomplete information, either $\overline{P}_{post}$ or $\overline{P}_{ante}$ may be used as a target in pollution reduction policies resulting, as we will shoe, in very different efficient designs for a permit program. If it is legally stipulated or the damage function dictates that pollution not exceed the prefixed standard, then $\overline{P}_{post}$ is the relevant constraint for cost minimization. This is the commonly analyzed case when total pollution is limited to a prefixed cap, regardless of firms' abatement costs. There are many examples when the regulator is not constrained by such an inflexible target however. For example,

Given one of these forms of *ex ante* targets, the regulator potentially has the flexibility to issue permits and set the trading ration, t, to achieve the expected pollution target at least cost. The following chart illustrates the decision process and the occurrence of events:



This sequential timing process makes clear that the actual pollution, P_{actual}, varies with the realization of firms' abatement costs and/or the delivery coefficient whereas and $\overline{P}_{ante}$ (or $\overline{P}_{\mathrm{post}}$) must be set before the realization of these uncertainties and will not change when the uncertainties are resolved.

2.1.2. Minimizing expected abatement costs under an ex ante pollution constraint.

For the situation analyzed most in the permit trading literature where the delivery coefficient is known and an *ex post* pollution target is used, the regulator must set the trading ratio equal to *d* and $\overline{P}_{permit} = \overline{P}_{\mathrm{post}}$. Otherwise, there is no guarantee that the target will be met. This is because from (1), we know that

(2) $$\overline{P}_{permit} - \overline{P}_{post} = (t-d)e_2 .$$

If $t = d$, then $\overline{P}_{post} = \overline{P}_{permit}$, regardless of the value of e_2. However, if the regulator is to set $t \neq d$, then she needs to adjust $\overline{P}_{permit}$ as well so that the *ex post* pollution target will be met. However, any adjustment will depend on the magnitude of e_2, which is assumed unknown to the regulator when designing the permit market (due to uncertain abatement costs).

Interestingly, it is not even feasible to use an *ex post* pollution constraint if the delivery coefficient is uncertain. For any given $(t,\ \overline{\mathrm{P}}_{permit})$, the value of $\overline{P}_{post}$ will vary with d and e_2. It is obvious that (2) will not hold for all possible value of d and e_2 in a permit trading program with trading ratio and total permits, $(t,\ \overline{\mathrm{P}}_{permit})$. In this case, an *ex ante* constraint is the only meaningful policy option.

On the other hand, when an *ex ante* pollution target is used, the realization of total pollution can be higher or lower than the target. Thus, even though the regulator cannot directly control the realization of total pollution, she may be able to set the parameters of the permit system (t and $\overline{P}_{permit}$) in conjunction with her (incomplete) knowledge of the firms' abatement costs to generate higher than average emission levels when firms' abatement costs turn out to be high and vice versa. Formally, we set up the regulator's problem as follows. First, the regulator obtains the firms' permit demand (reaction functions) as functions of the trading ratio (*t*) and the total permit endowment ($\overline{P}_{permit}$), that is, $e_i^*(t, \overline{P}_{permit}; \theta_1, \theta_2)$ for $i = 1, 2.$ The reaction functions, with random parameters θ_1 and θ_2, are then incorporated into the regulator's decision of the optimal trading ratio and total number of permits. Specifically, the regulator's problem is:

$$(3) \qquad \min_{t, \overline{P}_{permit}} \quad E[TC] \equiv E\left[C_1(e_1^*(t, \overline{P}_{permit}; \theta_1, \theta_2) + C_2(e_2^*(t, \overline{P}_{permit}; \theta_1, \theta_2)\right],$$

subject to *Ex ante pollution constraint in (1).*

We explore the design of such a system in this paper by first examining firms' decisions on the permit market and then deriving the optimal permit trading ratio and total permits.

2.2. Firms' emissions decisions in a permit trading market

Should an emissions trading program be introduced, the firms will face the permit market constraint in (1). Suppose the initial permit endowments allocated to Firm *i* (and denominated in Firm *i*'s emissions) are $\overline{e}_i$ for $i = 1, 2$; and $\overline{e}_1 + t\overline{e}_2 = \overline{P}_{permit}$. Through trading, both firms can hold the permits denominated in terms of another firm's emissions, and the trading ratio is used to convert between the two types of permits. The trading program requires that each firm's actual emissions do not exceed its holding of permits. Let y_i, denominated in terms of Firm *i*'s emissions, denote the equilibrium quantities of permits traded. Assuming that each firm takes permit prices as given, then Firm 1's problem would be as follows

$$\begin{aligned} &\min_{e_1, y_1, y_2} C_1(e_1^0 - e_1) - p_1 y_1 + p_2 y_2 \\ &\textit{subject to} \;\; e_1 - y_1 + t y_2 \le \bar{e}_1 \end{aligned} \tag{4}$$

The solution to this problem satisfies: $MC_1 \equiv C_1'(e_1^0 - e_1^*) = p_1$, the standard result that marginal abatement costs equal the price of permits. Similarly, for Firm 2, we have $MC_2 \equiv C_2'(e_2^0 - e_2^*) = p_2$. In both firms' problems, we can derive $p_1 / p_2 = 1/t$, indicating that the ratio of permit prices must be equal to the trading ratio. Otherwise, costless arbitrage opportunities would be available to firms. Then, we have

$$\frac{MC_2}{MC_1} = t \,. \tag{5}$$

From (5) and the permit market constraint in (1) , we can solve for the firms' optimal emissions as a function of t and P_{permit}. Recall that when emissions decisions are made in the permit trading market, firms have complete information, i.e., θ_1 and θ_2 are known with certainty. Thus regardless of the realization of θ_1 and θ_2, firms will equate the ratio of marginal costs with the trading ratio. When an *ex post* target must be met, we know from Montgomery that it is efficient to set t=d, resulting in

$$\frac{MC_2}{MC_1} = d \tag{6}$$

and $\overline{P}_{post} = \overline{P}_{permit}$ under every realization. Any gains in setting t at a level other than d in an *ex ante* targeting program would need to be weighed against the efficiency costs of not attained the equality in (6). This is an issue we will return to in the next section.

Before deriving an expression for the optimal design of the permit system under *ex ante* targets, we must assume functional forms. For tractability, we assume one firm faces a linear abatement cost function while the other faces an increasing convex abatement cost function, as specified below,

(7) $$C_1(e_1^0 - e_1, \theta_1) = (a + \theta_1)(e_1^0 - e_1),$$

(8) $$C_2(e_2^0 - e_2, \theta_2) = (b + \theta_2)(e_2^0 - e_2) + c(e_2^0 - e_2)^2.$$

With the above cost functions, we can derive firms' optimal emissions from equation (5) and the permit market constraint in (1),

(9) $$e_1^*(t, \overline{P}_{permit}; \theta_1, \theta_2) = \frac{2c(\overline{P}_{permit} - e_2^0 t) - (b + \theta_2)t + t^2(a + \theta_1)}{2c}$$

(10) $$e_2^*(t, \overline{P}_{permit}; \theta_1, \theta_2) = \frac{2ce_2^0 + (b + \theta_2) - t(a + \theta_1)}{2c}.$$

Clearly, the amount of emissions generated by the firms depend on regulatory decisions on t and $\overline{P}_{permit}$ as well as the values of the parameters θ_1 and θ_2. Note in this special case that e_2^* does not vary with $\overline{P}_{permit}$, implying that when $\overline{P}_{permit}$ is altered, e_1^* will absorb all the changes in $\overline{P}_{permit}$, i.e.,

(11) $$\frac{\partial e_1^*}{\partial \overline{P}_{permit}} = 1 \quad \text{and} \quad \frac{\partial e_2^*}{\partial \overline{P}_{permit}} = 0.$$

This feature come directly from the linearity assumption, and while not likely typical of real world situations, it makes the analysis tractable with no obvious loss of generality.

From (9) and the permit market constraint: $e_1^* + te_2^* = \bar{P}_{permit}$, it is straightforward to show the following

$$(12) \qquad \frac{\partial e_1^*}{\partial t} \le 0 \text{ iff } -1 \le \varepsilon_{e_2,t} \equiv \frac{t}{e_2^*}\frac{\partial e_2^*}{\partial t}; \qquad \text{and} \qquad \frac{\partial e_2^*}{\partial t} < 0.$$

That is, as t increases e_2^* decreases. For a given $\bar{P}_{permit}$, the sign of $\partial e_1^*/\partial t$ is the opposite of $\partial\left(te_2^*\right)/\partial t$. If the elasticity of e_2^* with respect to t is greater than or equal to -1, then $\partial e_1^*/\partial t \le 0$ for any given $\bar{P}_{permit}$. If the elasticity is less than -1, then $\partial e_1^*/\partial t > 0$. These reactions to changes in t and $\bar{P}_{permit}$ will be taken into account by the regulator in the design of an optimal permit market.

3. Optimal permit trading ratio and total permits

With analytical solutions ((9) and (10)), for the firms choice of emission levels, it is straightforward to solve *ex ante* optimization problem (3) [6]. When the regulator chooses a trading ratio and a total permit quantity that minimizes *ex ante* expected abatement costs the optimal trading ratio is a function of the regulator's prior information on the covariance structure of abatement cost uncertainties and the delivery coefficient, or specifically, [7]

$$(13) \qquad t^* = \mu + \frac{1}{a^2 - \sigma_1^2}\left(\mu\sigma_1^2 + a\,\mathrm{cov}(d,\theta_1) - \mathrm{cov}(\theta_1,\theta_2)\right).$$

[6] The problem is a standard optimization problem with one constraint and so the details on the derivation of the solutions are not presented. To simplify our discussions, interior solutions are assumed throughout the paper, unless otherwise noticed.

[7] In order for the optimal trading ratio to be positive, we must have $a\,\mathrm{cov}(d,\theta_1) + \mu a^2 > \mathrm{cov}(\theta_1,\theta_2)$, given that $a^2 - \sigma_1^2$>0. This is not an unreasonable assumption, as will be clear in our later analysis.

To derive the optimal permit cap, we first note that as long as the programs is intended to reduce emissions, all constraints in (1) will be binding. Further, since the permit market constraint must hold for every level of firms' emissions, it also must hold for expected emissions levels: $E[e_1^*] + tE[e_2^*] = \overline{P}^*_{permit}$. Taking the difference of this equation and the *ex ante* pollution constraint in (1), we obtain $\overline{P}^*_{permit} - \overline{P}_{ante} = E[e_2^*]t^* - E[de_2^*]$. Note that $E[de_2^*] = E[d]E[e_2^*] + Cov(d, e_2^*)$, $E[d] = \mu$, and $Cov(d, e_2^*) = \frac{Cov(d, \theta_2) - t^* Cov(d, \theta_1)}{2c}$. Then, we obtain:

(14) $$\overline{P}^*_{permit} - \overline{P}_{ante} = E[e_2^*](t^* - \mu) + \frac{Cov(d, \theta_2) - t^* Cov(d, \theta_1)}{2c}$$

Equations (13) and (14) imply that with certain values of θ_1, θ_2 and d (say, $d = \mu$), the optimal trading ratio would be set equal to the delivery coefficient, i.e., $t^* = d(= \mu)$, and the total permit quantity allocated to firms would equal the pollution target, i.e., $\overline{P}^*_{permit} = \overline{P}_{ante}$. However, in general, the second moment terms will matter. The implication is that, given that the regulator does not know firms' abatement costs and/or the delivery coefficient *ex ante*, the total permit quantity issued to firms can be higher or lower than the *ex ante* pollution target depending on the abatement cost parameters and on the correlation structure of cost and pollution impact uncertainties. This, of course, differs starkly with the case of an *ex post* permit trading design, where the total quantity of pollution permits issued is the same as the target pollution level that the regulator sets out to achieve.

3.1 Case of a known delivery coefficient

To gain further understanding of these findings, we next investigate the role of uncertainty in the delivery coefficient in the optimal solution by studying the optimal trading ratio and permit quantity when d is known with certainty.

3.1.1 *Difference between pollution target, total permits, and actual pollution.*

When the delivery coefficient is a known constant, we know from (1) that the gap between the total permits allocated and the pollution target is:

(15) $$\bar{P}_{permit} - \bar{P}_{ante} = E[e_2](t - d) .$$

Thus, if t≠d, then the total permit quantity will also deviate from the *ex ante* target so that the *ex ante* pollution constraint will be met. Similarly, we can derive

(16) $$P_{actual} - \bar{P}_{permit} = -e_2(t - d)$$

That is, if t>*d*, then the actual pollution will be less than the permit allocated. This occurs because 1 unit of Firm 2's emissions contributes d units to total actual pollution, but 1 unit of Firm 2's emissions requires t units of permits in the market constraint. Adding up the previous two equations, we have

(17) $$P_{actual} - \bar{P}_{ante} = (t - d)\left(E[e_2] - e_2\right).$$

Plugging in the reaction function of Firm 2 from equation (10) yields

(18) $$P_{actual} - \bar{P}_{ante} = (t - d)\frac{t\ \theta_1 - \theta_2}{2c}.$$

This relationship implies that, for any given θ_2, the higher θ_1 is, the higher the actual pollution will be if $t > d$. It will be optimal for the functional forms considered here for the regulator to set $t > d$ given an *ex ante* pollution constraint. Before providing the analytic solution we explain the intuition as follows. The regulator knows that, for any realization of θ_1 and θ_2, the emissions that would result in the least abatement cost for any given level of pollution would satisfy (6), that is,

(19) $$MC_2 = (b + \theta_2) + 2c(e_2^0 - e_2) = d(a + \theta_1) = d * MC_1 .$$

The relationship in (19) implies that the marginal cost of keeping total pollution under a given level would be determined by θ_1 regardless of the value of θ_2 and the allocation of emissions among the two firms. Thus, a higher θ_1 would imply a higher marginal abatement cost to control pollution under a given level. For a regulator who is required to meet a pollution level in expectation (i.e., on average), therefore, it makes sense to design policies that require lower abatement (higher pollution level) when marginal abatement cost (θ_1) turns out to be high and vise versa. Having a trading ratio that is greater than the delivery coefficient would achieve this goal, since equation (16) implies that $\frac{\partial P_{actual}}{\partial \theta_1} =$ $(t - d)\frac{t}{2c} > 0$ if $t - d > 0$.

In contrast, we know from the relationship in (19), for any given θ_1, the variation in θ_2 will not change the marginal costs of controlling the total pollution under a given level. This happens even though equation (16) implies that total pollution level change with θ_2 and $\frac{\partial P_{actual}}{\partial \theta_1} = -(t - d)\frac{1}{2c}$. In other words, for any given θ_1, there will be more pollution when θ_2 is lower and there will be less pollution when θ_2 is higher. However, since the marginal cost is the same for the extra pollution as for the reduction in pollution, setting $t > d$ will not affect the total cost of reaching an *ex ante* pollution target. Overall, given the effects of θ_1 and θ_2 on abatement costs, the optimal trading ratio should be greater than the delivery coefficient, assuming θ_1 and θ_2 are uncorrelated.

This reasoning is confirmed by the optimal solution for the trading ratio, presented below for the case with known delivery coefficient,

(20) $$t^* = d + \frac{d\sigma_1^2}{a^2 - \sigma_1^2} \text{ if } d \text{ is known and } \operatorname{cov}(\theta_1, \theta_2) = 0.$$

If θ_1 is known with certainty, i.e., $\sigma_1^2 = 0$, then the optimal trading ratio should equal the delivery coefficient. However, if $\sigma_1^2 > 0$, then $t^* > d$ under the assumption that $a^2 - \sigma_1^2 > 0$. Further, t^* increases with σ_1^2. That is, the more Firm 1's costs vary, the more there is to gain (in terms of reduction in the total expected abatement costs) by setting a higher trading ratio.

3.1.2 The total pollution effect and the deadweight loss effect

Equation (20) implies that the optimal trading ratio should be set greater than the deterministic delivery coefficient. Since this result stands in sharp contrast with traditional permit trading theory, we examine the effects of the trading ratio in more detail. Note that setting $t \neq d$ will result in two types of changes in the permit trading program. First, as the regulator adjusts t, she effectively changes the actual total pollution that is allowed in a given trading program. This change occurs because one unit of permit is no longer necessarily the same as one unit of pollution since t is used in the permit market constraint, not d. We refer to this effect as the *total pollution effect.* Second, when $t \neq d$, the permit market equilibrium satisfies $MC_2 = t * MC_1$ $(= t * a)$, but not $MC_2 = d * MC_1$. We refer to this effect as the *deadweight loss effect* since this change increases in the total abatement costs through suboptimal emission distributions among the firms.

Figure 1 and Figure 2 illustrate the intuition and magnitude of these effects. For simplicity, the delivery coefficient in the Figures is set to one, which is known by the regulator. In both figures, the total length of the horizontal axis represents the total permits available and the solid downward sloping line is the marginal abatement cost curve of Firm 2 as emissions are increased (i.e., abatement is decreased) for the case where $\theta_2 = 0$. In Figure 1, the marginal abatement cost curve for Firm 1 (for $\theta_1 = 0$, i.e., $MC_1 = a$) is represented by the horizontal line that intersects with Firm 2's marginal cost curve at B^0. When $t = d = 1$, we set $\bar{P}_{permit}^{t=1} = \bar{P}_{ante}$ by equation (15). At B^0, $MC_2 =$ $d * MC_1 = MC_1$. Thus, B^0 represents the permit market equilibrium, indicating the split of the emissions by the two firms with Firm 1's emissions read from the right (O_1) and Firm 2's emissions read from the left (O_2). As the condition in (6) is satisfied at B^0, the *ex post* abatement cost is minimized to reach a total pollution level of $P_{permit}^{t=1}$. By "*ex post*", we mean that minimization is with respect to known values of θ_1 and θ_2 (in the case of certainty $\theta_1 = \theta_2 = 0$).

When the trading ratio is set greater than the known delivery coefficient several changes occur in Figure 1. First, the optimal total permit endowment increases to $\bar{P}_{permit}^{t>1}$, which is reflected by the shifting out of the right boundary of Figure 1 from O_1 to O_1'. Second, the new permit market equilibrium is represented by point B', indicating a reduction in e_2. Third, we can no longer obtain e_1 from right (O_1') to the equilibrium point (B'), since the permit market constraint now requires that the total permits be greater than or equal to the weighted sum of emissions (with the weight on e_2 equal to t), not the simple sum of emissions from the two firms. To reflect the weighting, we use the dotted downward sloping curve to represent $t * e_2$ for every e_2 on MC_2. Then, Firm 1's emissions can be obtained by reading from right (O_1') to B''.

The two effects of setting $t > d$ on the total abatement cost of meeting the *ex ante* pollution target are illustrated by the shaded areas in Figure 1. As to the deadweight loss effect, note that the marginal abatement cost curve is still the horizontal line a, not the horizontal line ta. However, firms make their decisions based on the latter, which leads to too few emissions (i.e., too much abatement) by Firm 2, resulting in deadweight loss as reflected by the shaded triangle. As to the total pollution effect, it is easy to see that, the same pollution on as the permit market with $t = d = 1$ and total permits $\overline{P}_{permit}^{t=1} = \overline{P}_{ante}$. This is because the increase in total permits would just equal $e_2(t-d)$, canceling the extra weight on e_2 due to the use of $t > d$.[8] So, the pollution effect is zero. Not surprisingly, when there is no uncertainty, a welfare loss will occur if $t \neq d$.

Under uncertainty however, the total pollution effect will not necessarily be zero and an optimally designed permit market will try to achieve a balance of this effect with the deadweight loss effect. To show how the regulator can reduce total *ex ante* expected abatement costs by setting $t > d$, we use the illustration in Figure 2, which is the same as

---

[8] Graphically, the detailed reason is as follows. Given that B^0 is the intersection of the marginal cost curves for the case $\theta_1 = \theta_2 = 0$, from (9) we know that the emissions determined by B^0 are also the expectation of emissions. Then, from (15) the distance $B'B''$ is equal to the distance that the total permit boundary is moved to the right from O_1 to O_1'. Thus, Firm 1's emissions under $t > d = 1$, which is the distance going left from O_1' to B'', is equal to the distance going left from O_1 to B'. Therefore, the total emissions from the two firms is still the distance O_2O_1.

Figure 1 except that it illustrates a case where θ_1 can take on two values ($+\hat{\theta}_1$>0 and $-\hat{\theta}_1$) with equal probability. As is clear from the figure, there will be deadweight loss (the two shaded triangles) regardless of whether Firm 1's marginal abatement cost are high ($\theta_1 = +\hat{\theta}_1$) or low ($\theta_1 = -\hat{\theta}_1$). However, the distortion is higher when the realization of Firm 1's marginal abatement cost is higher since MC_2 is set equal to MC_1 multiplied by t, that is, the distortion is multiplicative.

Setting $t > d$ will result in a cost saving from less abatement when marginal cost is high. In Figure 2, the total pollution difference between setting $t > d$ and $t = d$ is the width of the shaded rectangle and the cost saving is represented by the larger shaded rectangle**.**[9] Similarly, setting $t > d$ will result in an extra cost from more abatement when marginal cost is low. The smaller shaded rectangle in figure 2 represents this cost. When the difference between the cost saving and the extra cost is positive, and when the difference is greater than the deadweight loss (the sum of the two shaded triangles), the regulator reduces total abatement cost with $t > d$. As illustrated in Figure 2, the area of the larger rectangle is larger than the sum of the areas of the smaller rectangle and the two shaded triangles, resulting in a welfare gain from setting $t > d$.

3.1.3 Effects of the covariance structure on the optimal permit trading program.

[9] The detailed explanation is similar to that in footnote 3 except that here the distance between $A'A''$ is smaller than the distance O_1O_1', while in the situation in footnote 3, distance $B'B''$ is equal to the distance O_1O_1'. Similarly, for the case when marginal abatement cost is know, the distance between $D'D''$ is greater than the distance O_1O_1', which indicate less total pollution than when $t = d$.

In the previous subsection, we have discussed the case where the covariance of θ_1 and θ_2 is zero. A nonzero $\mathrm{cov}(\theta_1,\theta_2)$ will affect the optimal design of a permit market, specifically when d is uncertain

(21) $$t^* = d + \frac{1}{a^2-\sigma_1^2}\left(d\sigma_1^2 - \mathrm{cov}(\theta_1,\theta_2)\right).$$

If $\mathrm{cov}(\theta_1,\theta_2) \leq 0$, then clearly $t^* \geq d$. If $\mathrm{cov}(\theta_1,\theta_2) > 0$ and large enough, it is possible to have $t^* < d$. The intuition is as follows. If θ_1 and θ_2 are positively correlated, then they tend to take similar values. When both θ_1 and θ_2 are large, from equation (19) we know that e_2 will be large as well. If e_2 is so large that it is greater than $E[e_2]$, then (17) implies that $P_{actual} > \overline{P}_{ante}$ if and only if $t < d$. Thus, to minimize total abatement costs which call for higher pollution when marginal abatement cost is high (i.e., θ_1 is large), the optimal trading ratio should satisfy $t^* < d$. Similarly, when both θ_1 and θ_2 are small, e_2 can be smaller than $E[e_2]$. In such a case, setting $t^* < d$ will save abatement costs.

The covariance and the trading ratio move in the opposite directions since (21) implies that $\frac{\partial t^*}{\partial\,\mathrm{cov}(\theta_1,\theta_2)} = \frac{-1}{a^2-\sigma_1^2} < 0$. We know from (18) that the total pollution effect will be smaller when the cost shocks move in the same direction than when they move in the opposite direction. Intuitively, larger positive correlation reduces the trading ratio because a smaller total pollution effect can just overcome the effect of a smaller deadweight loss effect which requires a smaller deviation from the know delivery coefficient.

The magnitude of the optimal trading ratio is also affected by other factors. First, as the variance of the marginal abatement cost of Firm 1 and $\frac{\partial t^*}{\partial \sigma_1^2} = \frac{t^*}{a^2 - \sigma_1^2} > 0$. This is because as θ_1 becomes more variable, P_{actual} also becomes more variable and the total pollution effect will be larger. That is, the difference is larger between the abatement costs saved when θ_1 is high and the extra abatement costs incurred when θ_1 low. In terms of Figure 2, as θ_1 becomes more variable, the width of both rectangles increase and so the total pollution effect increases. In addition, the difference between their heights also increases. A larger total pollution effect can outweigh a higher deadweight loss and so the optimal trading ratio can be set higher.

In addition to affecting the optimal design of a permit market, the covariance structure of the cost and the delivery coefficient can also affect the expected total abatement costs. As a benchmark, we use the permit trading program with a trading ratio equal to the known delivery coefficient and a total permit cap equal to the *ex ante* pollution target. We denote the regulator's *ex ante* expected total abatement costs as $E[TC\left(t^*, \bar{P}^*_{permit}\right)]$ and $E[TC\left(d, \bar{P}_{ante}\right)]$ for the optimal permit trading program and the benchmark case. Taking the difference between the total expected abatement costs, we obtain

(22) $$\Delta TC^{d,t^*} = E[TC(d, \bar{P}_{ante})] - E[TC(t^*, \bar{P}^*_{permit})] = \frac{\left(\mathrm{cov}(\theta_1, \theta_2) - d\sigma_1^2\right)^2}{4c(a^2 - \sigma_1^2)} \geq 0 .$$

The difference between the permit caps are given by (15), that is,

(23) $$\Delta P^{d,t^*} = \bar{P}_{ante} - \bar{P}^*_{permit} = -(t^* - d)E[e_2^*]$$

First, note that $\frac{\partial \Delta TC^{d,t^*}}{\partial \operatorname{cov}(\theta_1,\theta_2)} = \frac{\operatorname{cov}(\theta_1,\theta_2) - d\sigma_1^2}{2c(a^2-\sigma_1^2)} = \frac{-(t^*-d)}{2c}$, which has the same sign as $\Delta P^{d,t^*}$. In the case $t^* > d$, the potential of the optimal program to save costs relative to the benchmark case is higher if the covariance is smaller. Ceteris paribus, this implies a negative correlation would mean a higher cost saving potential than a positive correlation. This is consistent with our earlier discussion that a negative correlation is associated with a larger total pollution effect.

Next, we observe that $\frac{\partial \Delta TC^{d,t^*}}{\partial \sigma_1^2} = \frac{(d\sigma_1^2 - \operatorname{cov}(\theta_1,\theta_2))d\sigma_1}{c(a^2-\sigma_1^2)} + \frac{\sigma_1\left(d\sigma_1^2 - \operatorname{cov}(\theta_1,\theta_2)\right)^2}{2c(a^2-\sigma_1^2)^2}$ $= \frac{(t^{*2}-d^2)}{2c}$, which is positive if $t^* > d$ (i.e., $\Delta P^{d,t^*} < 0$) and negative otherwise. Thus, the expected cost gains from our approach are positively related to the variance of Firm 1's abatement cost if the regulator finds it optimal to allocate more permits than the target expected pollution level, and the higher the variability in abatement costs of Firm 1, the higher expected abatement cost savings can be realized from implementing our solution relative to the benchmark.

3.2 Case of an uncertain delivery coefficient

The delivery coefficient is likely to be known for some pollutants (e.g. carbon dioxide), but there are many pollutants where delivery coefficients will be uncertain. While uncertain delivery coefficients clearly characterize nonpoint source popllution, many point sources can also have uncertain delivery coefficients; for example, wind and weather uncertainty can affect air pollution deposition rates. Many water pollutants exemplify this notion well. The fate and transport of water pollutants is subject to both stochastic elements related to weather as well as scientific uncertainty concerning the physical diffusion process.

In the case of an uncertain delivery coefficient, there is an additional source of flexibility due to the fact that our approach is able to utilize distributional characteristics of the delivery coefficient and the abatement cost parameters to avoid potentially socially costly situations. Consider a case where the delivery coefficient turns out to be high, and simultaneously, Firm 1 receives a positive cost shock ($\theta_1 > 0$), but θ_2, on the other hand, turns out to be low. To achieve an *ex post* target, higher abatement costs would result relative to an *ex ante* objective. With an *ex ante* target, the trading ratio is anchored around the *expected* delivery coefficient and is appropriately modified to account for the joint distribution of abatement cost parameters and the delivery coefficient, and so the regulator's objectives can be achieved at the lowest possible expected abatement cost.[10]

Whether the optimal trading ratio for this case, given by (13), is greater or smaller than the known delivery coefficient depends on the values of the covariance terms. This result provides an interesting contrast to actual practice where emissions from uncertain sources have been discounted due to the stochastic nature of the emissions. That is, if the regulator is asked to reduce total expected pollution to a certain level and a permit trading system is used to achieve this goal. Then, from standard economic theory and prior studies (e.g. Horan 2001), the regulator should use a trading ratio that gives less weight to

[10] In the nonpoint source literature, as one of the defining features of nonpoint source pollution is its inherent unobservability (Segerson, 1988), the focus has been on the trading in expected, as opposed to actual, emissions from a nonpoint source (e.g., Horan et al, 2001). In this case, basically, another layer of uncertainty would be added to the design of the permit market: both firms and the regulator only know the distribution of emissions given any action taken by the firms. We can show that, like the uncertainty on firms' abatement costs and the delivery coefficient, this uncertainty will also be reflected in the optimal trading ratio and the optimal total number of permits.

emissions from uncertain sources (in our context, t should be set less than d). However, our results suggest that this is suboptimal if $\mu\sigma_1^2 + a\,\mathrm{cov}(d,\theta_1) - \mathrm{cov}(\theta_1,\theta_2) > 0$.

Compared to equation (21) for the simple case of a known delivery coefficient, in equation (13) the delivery coefficient is replaced by its expectation, μ , and an additional term, $\mathrm{cov}(d,\theta_1)$, comes into play. We find that the trading ratio moves in the same direction as $\mathrm{cov}(d,\theta_1)$: $\frac{\partial t^*}{\partial\,\mathrm{cov}(d,\theta_1)} = \frac{a}{a^2 - \sigma_1^2} > 0$. The impact of this covariance takes place through the total pollution effect. Suppose $\mathrm{cov}(d,\theta_1) > 0$, that is, if the delivery coefficient is expected to be high, the marginal cost of abatement by Firm 1 is also expected to be high. This means for given emissions, there will be more total pollution. In other words, for a given total pollution target, more abatement has to be undertaken. To ameliorate the pressure for more abatement, the trading ratio is increased and so more emissions will be allowed when both the abatement cost and delivery coefficient is high. By the same logic, when both abatement cost and delivery coefficient are low, a higher trading ratio will restrict the amount of emissions that are allowed. However, the cost savings from extra pollution is higher than the increased cost from more abatement and so total abatement costs are reduced.

The optimal permit allocation gap with an uncertain delivery coefficient is given by equation (14). Compared to a known delivery coefficient as given in equation (15), there are two additional covariance terms. The terms indicate that if e_2 and d are positively correlated, then the optimal total permit cap should be even higher and vice versus. Thus, with an uncertain delivery coefficient, there is an additional reason that the optimal total permit cap might differ from the *ex ante* pollution target.

4. Conclusions

In this paper, we examine the optimal trading ratio and the optimal total permit allocation for implementation of an emissions trading program under the goal of cost-effectiveness, where the environmental objective is to be achieved in expectation. We formulate the regulator's problem in terms of minimizing expected costs of achieving an given expected level of pollution reduction. This is done in the spirit of the well-known "efficiency without optimality" approach. We solve for the firms' permit demands (reaction functions) as functions of the trading ratio and the total initial permit endowment. Firms' reaction functions are then substituted into the expected total abatement cost and the expected emissions constraint which comprise the regulator's problem. Finally, we obtain the optimal trading ratio and the number of total emission permits.

We find that, the optimal trading ratio is anchored around the expectation of a ratio of delivery coefficients, and is modified by variance and covariance terms stemming from joint distributions of the random components. Importantly, our model produces a trading ratio that is not a function of the optimal solution, but instead is a function of regulator's prior information on abatement costs, environmental impacts, and actual emissions.

Furthermore, we find the optimal number of total permits need not coincide with the exogenously given pollution target. Regardless of the uncertainty in the delivery coefficients, permit trading will ensure that that the sum of emissions adjusted by the *trading ratio* is equal to the total permits issued. However, for the regulator, the relevant constraint is that the expected pollution (the sum of emissions adjusted by the *delivery* *coefficient*) is equal to a preset standard. Given that the regulator does not know firms' reaction functions and the delivery coefficients *ex ante*, the total permit number issued to firms can be higher or lower than the target depending on the abatement cost parameters and on the correlation structure of cost and pollution impact uncertainties. Our approach ensures that the total permit allocation is set in such a way that when the abatement costs are high, the stringency of the quantity instrument is reduced. It is this added flexibility that allows for an improved performance of a pollution trading program in terms of *ex ante* expected abatement costs.

Our model lends itself to two different but equally important interpretations. One is that the regulator and the polluting firms are facing genuine cost (and, consequently, abatement) uncertainty and therefore the regulator can form actual *ex ante* expectations regarding the pollution levels. When the framing of environmental policy in terms of average pollution levels is appropriate, this model improves on the simple pollution trading design. The second interpretation is one of asymmetric information, where the regulator is uncertain of firms' pollution abatement costs. In this case, the model provides a regulator a way to introduce some flexibility into a framework of pollution trading. This may prove useful to a regulator who is charged by the political process to institute a pollution trading scheme, yet is worried about abatement costs being excessive.

References:

Baumol, William and Wallace Oates. ***The Theory of Environmental Policy***, Cambridge University Press, second edition 1988.

Horan, Richard D. 2001. "Differences in Social and Public Risk Perceptions and Conflicting Impacts on Point/Nonpoint Trading Ratios." ***American Journal of Agricultural Economics*** 83(4):934-41.

Horan, R.D. and J.S. Shortle. 2005. "When Two Wrongs Make a Right: Second-Best Point-Nonpoint Trading Ratios." ***American Journal of Agricultural Economics***, 87(2): 340-352.

Malik, Arun; David Letson; and Stephen Crutchfield, 1993. "Point/Nonpoint Source Trading of Pollution Abatement: Choosing the Right Trading Ratio." ***American Journal of Agricultural Economics*** 75(4): 959-967.

Montgomery, W. D., 1972. "Markets in Licenses and Efficient Pollution Control Programs." ***Journal of Economic Theory*** 5, 395-418.

Segerson, Kathleen. 1988. "Uncertainty and Incentives for Nonpoint Pollution Control." ***Journal of Environmental Economics and Management***, 15: 87-98 (1988).

Yates, A.J., Cronshaw, M.B., 2001. "Pollution Permit Markets with Intertemporal Trading and Asymmetric Information." ***Journal of Environmental Economics and Management*** 42, 104-118.

Figure 1. The effects of setting $t^* > d = 1$ under the *ex ante* pollution constraint $e_1 + de_2 = P_{ante}$ and the permit market constraint $e_1 + te_2 = P_{permit}$

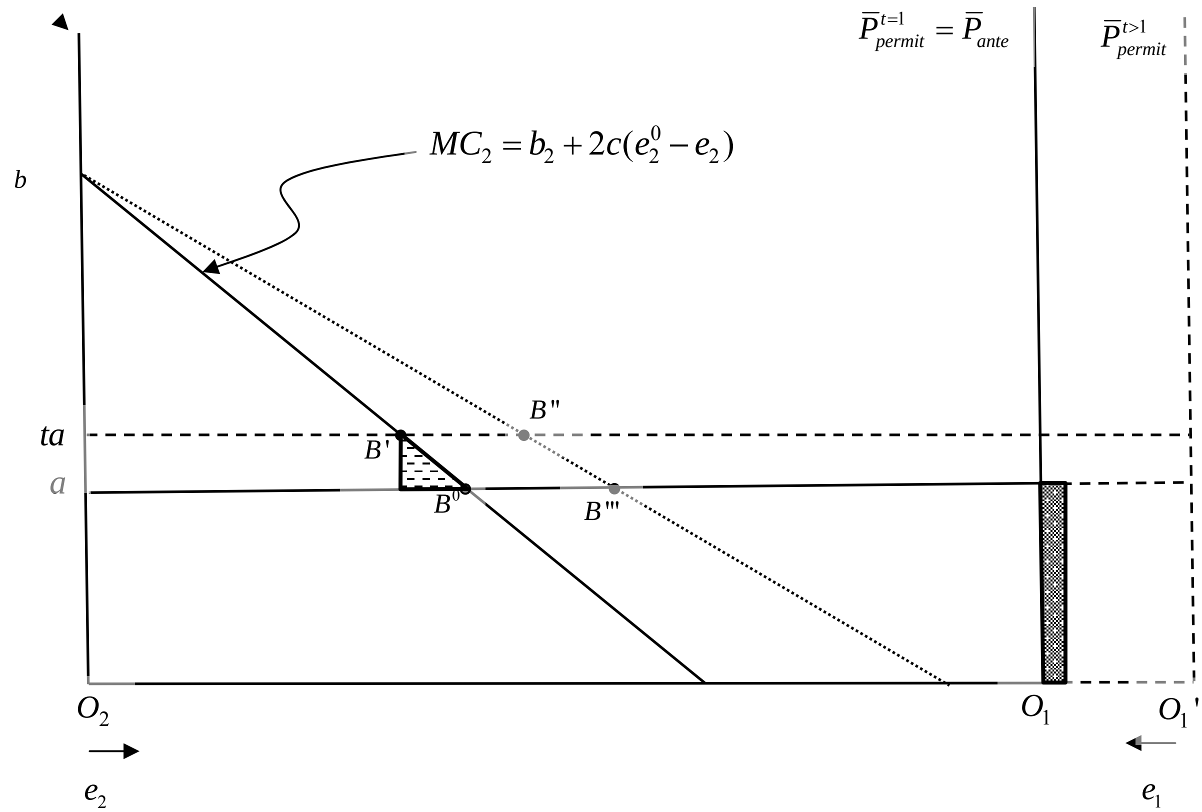


Figure 2. A comparison of the welfare effects when θ_1 is high versus when θ_1 is low for a given value of θ_2 ($\theta_2 = 0$).