JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT **21**, 17–31 (1991)

Market Failure in Incentive-Based Regulation: The Case of Emissions Trading*

SCOTT ATKINSON

Department of Economics, Brooks Hall, University of Georgia, Athens, Georgia 30602

AND

TOM TIETENBERG

Department of Economics, Miller Library 234, Colby College, Waterville, Maine 04901

Received October 12, 1990; revised December 31, 1990

Among the existing evaluations of the US EPA's emissions trading program, a consensus has emerged. While the program has resulted in significant cost savings, it has not even approximately achieved a cost-effective allocation of the control responsibility. The cost savings have been smaller and the trades fewer than might have been expected at the outset of the program. In this article we explore one hypothesis which purports to explain the divergence between the cost-minimizing and the observed pattern of trades for nonuniformly mixed pollutants. The "trading process hypothesis" attributes some significant proportion of this divergence to the nature of the process by which emission reduction credits are traded under the bubble policy. An examination of actual bubble trades reveals that the actual trading process is sequential and bilateral and, hence, differs considerably from the implicit process modeled in the existing empirical studies. Simulations of this more realistic trading process suggest that the resulting equilibria deviate considerably from cost-effective allocations of the control responsibility. © 1991 Academic Press, Inc.

INTRODUCTION

The theoretical cost effectiveness of an emissions trading approach to pollution control was formally developed in the early 1970s by Baumol and Oates [5] and Montgomery [14]. Under this approach polluters would be free to trade emission rights among discharge points as long as the prevailing air quality standards were not violated. Beginning with Atkinson and Lewis [1] a series of empirical studies have established that the potential savings to be achieved by implementing emissions trading were substantial.[1]

On December 11, 1979 the U.S. Environmental Protection Agency promulgated rules launching the bubble policy, whereby existing sources could trade emissions permits among themselves.[2] Expected to be the centerpiece of the subsequently named "Emissions Trading Program," the bubble policy followed by some three years the creation of its successful predecessor, the offset policy, whereby new sources were required to acquire offsetting emission permits from existing emitters as a precondition for entering an area with air quality worse than allowed by the

---

*An earlier version of this paper was presented at the AERE Workshop on Natural Resource Mechanisms, Madison, Wisconsin, June 7, 1990.

[1] These are summarized in [17].

[2] These rules can be found in 44 *Federal Register* 71780.

0095-0696/91 $3.00

ambient standards.[3] In contrast to the offset policy, which was designed to reduce the conflict between air quality improvement and economic growth, the bubble policy attempted to increase the cost-effectiveness of the regulatory policy for controlling existing sources of air pollution; it is this component of the Emissions Trading Program which most closely resembles the concept addressed in the empirical and theoretical work.

Despite an obvious willingness to expand the domain of application of the emissions trading concept,[4] in one key respect the bubble policy has failed to fulfill expectations. Among the existing evaluations of the emissions trading program,[5] a consensus has emerged that while the bubble policy has resulted in significant cost savings, it has not even approximately achieved a cost-effective allocation of the control responsibility. The cost savings have been smaller and the number of trades fewer than might have been expected at the outset of the program.

In this article we explore one hypothesis which purports to explain the divergence between the cost-minimizing and the observed pattern of trades for nonuniformly mixed pollutants.[6] The "trading process hypothesis" attributes some significant proportion of this divergence to the nature of the process by which emission reduction credits are traded under the bubble policy. An examination of external bubble trades reveals that the actual trading process is sequential and bilateral and, hence, differs considerably from the implicit process modeled in the existing empirical studies. Simulations of this more realistic trading process suggest that the resulting equilibria deviate considerably from cost-effective allocations of the control responsibility.

To explore this issue we develop and employ an original algorithm which mimics the essential elements of the trading process as it actually occurs under the bubble policy. This algorithm represents a considerable departure from the previous empirical work on emissions trading where mathematical programming techniques, typically linear or quadratic cost-minimization algorithms, are used to calculate the cost-effective equilibrium.

Our empirical results suggest that treating the mathematical programming equilibrium as if it were the emissions trading equilibrium is a serious misrepresentation of the emissions trading process. Furthermore our results provide support for the trading process hypothesis by demonstrating that a significant proportion of the gap between the expected cost savings and achieved cost savings can be explained by the dynamics of the trading process. Specifically we show that a sequence of bilateral trades which conforms to current EPA regulations produces substantially less cost savings for an important class of pollutants than the cost-effective allocation. A sensitivity analysis which recalculates the results for a variety of assumptions about target air quality levels and the nature of the trading process suggests that the results are robust, at least for our data set.

---

[3] The offset policy was initiated on December 21, 1976. The rules can be found in 41 *Federal Register* 55254.

[4] The emissions trading concept has recently been applied to controlling ozone depletion and nonpoint sources of water pollution, to reducing the lead content in gasoline, and has been prominently featured in the acid rain provisions in the Clean Air Act Amendments passed in the Fall of 1990.

[5] See [6, 8, 11, 17, 18].

[6] This hypothesis was first suggested in [9] and explored theoretically in [13]. This latter article suggests the need for an empirical algorithm of the type developed in this paper.

THE ISSUE

The Effectiveness of the Bubble Policy

While the offset and netting policies seem to have worked pretty much as expected, the bubble policy has not lived up to expectations. In light of the large number of existing sources in nonattainment areas that could presumably take advantage of its flexibility, the failure of the bubble program to stimulate more trades and to yield more cost savings is contrary to what was expected on the basis of the cost-effectiveness theorem and the empirical studies examining the potential for cost savings. Estimated potential savings of 50% were common and one study [10] even indicated that some 95% of the control cost could be saved by instituting a cost-effective policy. It was a false hope.

While many explanations for this divergence have been offered,[7] the empirical significance of each of these explanations remains unclear. One empirically unexamined explanation, the one providing the focus for this paper, is the hypothesis that the sequential, bilateral trading process followed by the sources of nonuniformly mixed pollutants under the bubble policy cannot achieve a cost-effective equilibrium.

Though the trading process hypothesis may appear to contradict one of the oldest and most important theorems of environmental economics, the cost-effectiveness theorem, in fact it does not. This theorem demonstrates the *existence* of a market equilibrium for a well-defined system of ambient permits which yields the cost-effective allocation of the control responsibility for achieving the ambient standards at specified air quality monitors, the least-cost solution. It says nothing about the details of the trading process which would yield that equilibrium. In fact since individual sources affect each air quality monitor differently with nonuniformly mixed pollutants, simultaneous trading of the permits by all firms would be a necessary condition for the market to always reach the cost-effective equilibrium.

The Nature of the Trading Process

An examination of external trades under the bubble policy (those involving transactions where the buyer and sellers are different corporate entities) reveals that they are bilateral, not multilateral, and sequential, not simultaneous. Furthermore, with very few exceptions, these trades either hold constant or reduce emissions, rather than allow increases.[8] These distinctions are important. Bilateral, sequential trades involving nonuniformly mixed pollutants between non-contiguous sources necessarily imply that air quality will change near both the purchasing and selling sources. Relative to the standards imposed in the State Implementation Plan (SIP), pollutant concentrations would fall in the neighborhood of the source selling the ERCs (due to the surplus emissions reductions necessary to create the credits) and pollutant concentrations would rise in the neighborhood of the

---

[7] See [6, 7, 11, 17].

[8] The exceptions occurred when EPA allowed sources to use an "allowable emissions" rather than an "actual emissions" baseline. As a result sources with emissions lower than allowed could trade the excess, causing actual emissions in the airshed to rise. This is described in detail in [17, 196–197; 11, 52–60].

acquiring source (since the ERC is a substitute for the higher level of control required by the SIP). Whereas simultaneous, multilateral trades can instantaneously capitalize on all offsetting increases and decreases among surrounding sources subject to the ambient standards, bilateral sequential trades cannot. Requiring that air quality be better than the standard implies that every bilateral, sequential trade must *individually* assure that the air quality constraints are not violated, a much more restrictive condition. Furthermore, the requirement that no trade can increase emissions means foregoing the opportunity to worsen air quality at receptors where air quality is significantly better than required by law, an important aspect of the cost-effective allocation of control responsibility.

While this a priori reasoning is sufficient to demonstrate that bilateral, sequential emissions trades are not likely to result in a cost-effective allocation of the control responsibility, it is not sufficient to demonstrate the empirical significance of the divergence. It is that issue to which the analysis in this paper now turns.

THE DATA AND METHOD

The data employed in this study comprise the 27 largest point sources of particulate emissions in the St. Louis Air Quality Control Region (AQCR), accounting for approximately 80% of total particulate emissions.[9] We consider their control costs and the impact of their emissions on each of nine receptors.

Transfer Coefficients and Control Cost Data

Source-receptor transfer coefficients, employed in the diffusion model to map source emissions into air quality for each of the trades, are derived using a Gaussian diffusion model developed by Martin and Tikvart [12]. The meteorological input data required for the model are the pollution dispersion characteristics which include location, stack height, average mixing height, stack exit conditions, stability wind rose (speed, direction, and stability class), and pollutant decay rates.[10] The output consists of an $(m \times n)$ matrix which gives the contribution of each of n sources to the predicted annual arithmetic average pollutant ground-level concentrations at each of m receptors. Transfer coefficients,

$$\{a_{ij}\} \qquad i = 1, \dots, m, j = 1, \dots, n,$$

with units of mg/m^3/ton/day, are obtained by dividing the concentration at the ith receptor due to the jth source by the number of tons emitted by the jth source.

Based upon Standard Industrial Classification code and source type, all area source and mobile sources were included as part of the background. Thus this study focuses on allocating the pollution control responsibility among the following major point sources: combustion facilities (primarily industrial and steam-electric power-plant boilers), industrial process sources, and solid-waste disposal sources (incineration and open-burning).

---

[9] These data have been used in earlier articles to examine other aspects of the emissions trading program. See [2, 3].

[10] For a more complete discussion see [19].

Before control cost data were developed, the applicability of control measures to each source was considered. In order to determine the compatibility of control devices with each source, consideration must be given to the temperature and volume of the effluent gas stream, type and efficiency of existing pollution controls, fuel usage requirements, and maximum process rate. A number of control measures were examined: wet scrubbers, mechanical collectors, electrostatic precipitators, mist eliminators, fabric filters, afterburners, and fuel substitution.

The costs of each device (in 1969 dollars) are obtained from the Control Technique Documents prepared by the U.S. Department of Health, Education, and Welfare [20].[11] The total annual cost of control includes annualized capital and installation cost (based upon a rate of interest and rated life of the device), as well as the annual average operating and maintenance costs. The resulting emission rate and control cost data can be found in Atkinson and Lewis [1].

Using the total control costs and tons removed by the devices which define the lower bound of total costs, we fit a continuous cost function for each source. Specifically the function

$$C_j = \beta_j X_j^2 + e_j \qquad j = 1,\dots,27 \tag{1}$$

was estimated using ordinary least squares (OLS), where C_j is the long run total control cost borne by the jth firm for X_j tons of controlled emissions, β_j is its corresponding population parameter, and e_j is an i.i.d. error term with zero mean and constant variance.

The Least Cost Strategy

We obtain the least-cost solution using the Mathematical Programming Optimization System (MPOS) algorithm to solve the quadratic programming problem

$$\begin{aligned} \min_{X_j} \mathrm{TC} &= \sum_j \hat{\beta}_j X_j^2 \\ \text{s.t. } A\mathbf{x} &\geq \mathbf{b}, \\ \text{and } 0 \leq \mathbf{x} &\leq \mathbf{x}^*, \end{aligned}$$

where

TC = total regional marginal cost;

A = a (9×27) matrix of transfer coefficients $[a_{ij}]$, $i = 1,\dots,9$, $j = 1,\dots,27$;

$\mathbf{x}$ = a (27×1) vector $(X_1,\dots,X_{27})'$;

$\mathbf{x}^*$ = a (27×1) vector $(X_1^*,\dots,X_{27}^*)'$ of maximum available emission control for each source; and

$\mathbf{b}$ = a (9×1) vector $(b_1,\dots,b_9)'$ of required improvement in air quality degradation (so the standard is not violated) at each air quality receptor.

---

[11] Since the purpose of this study is to explore a hypothesis that can be compared in order of magnitude terms to existing studies, we ran the algorithm on a familiar, admittedly dated, data set. To the extent that technological choices and costs are now different in real terms, we would be unable to forecast what would happen in St. Louis today.

The *potential* cost savings are calculated by subtracting this control cost from the corresponding control cost for the pre-trade equilibrium, the SIP allocation. The *actual* cost savings are calculated by subtracting the control cost for each scenario from the corresponding SIP control cost. The percent of potential savings is then calculated by dividing the actual savings by the potential savings and multiplying the result by 100.

The SIP Strategy

In accordance with the Clean Air Act of 1970, as amended in 1977, each state has submitted to the federal government an SIP which describes its basic air pollution control strategy for achieving the ambient air quality standards. For the purposes of this study, a set of emission regulations suggested in the SIP guidelines and representative of those employed by many states has been selected to form what we refer to as the SIP control strategy. The particulate standards include a heat input standard for fuel combustion sources (0.30 lb particulate matter/million BTUs), a process weight standard for industrial process sources (46.72 lb/hour of particulates/million lb/hour process weight), and a refuse-charged emission standard for solid waste disposal sources (0.20 lb particulates/100 lb of refuse processed). For the simulations involving different ambient air quality standards, the benchmark SIP control responsibilities were proportionally scaled. In each simulation the SIP allocation satisfies the ambient standards with at least one (usually only one) receptor reporting air quality equal to the standard.

The Trading Algorithms

The SIP allocation of the control responsibility is taken as the trading baseline in our algorithm just as it is in the emissions trading program. Sources are allowed to create ERCs by controlling more than required by the SIP. These surplus ERCs can then be traded to other sources.

The cost-effectiveness of the trades depends on the information available to the traders at the time of the trade, the rules governing the trades, and the sequence in which the trades take place. To capture the effects of various assumptions about information, rules, and sequencing on the ability of these trades to achieve significant cost reductions, we have constructed several trading scenarios for each of two target ambient air quality levels. The two air quality levels correspond roughly to the primary and secondary ambient standards for particulates in the geographic area which serves as the setting for our simulation model. In this simulation meeting the secondary standard requires more control than meeting the primary standard.

The simulations begin with what we refer to as the simultaneous, full information scenario. In order to isolate the decline in cost savings that can be attributed to the EPA rule prohibiting trades that result in increased emissions[12] this scenario allows simultaneous trading. Due to the simultaneous trading assumption, the possible cost savings for the simultaneous, full information scenario are derived

[12] In fact this scenario somewhat understates the decline in cost savings since the new EPA rules require trades which improve air quality in nonattainment areas, not merely hold it constant. This implies that total emissions must actually decline as a result of the trade.

from the same quadratic programming model used to derive the least cost solution with the added constraint that total emissions must be no greater than those permitted by the SIP allocation. Comparing the regional control costs for this allocation with that of the least cost allocation produces an estimate of the percentage of maximum cost savings achievable under current EPA rules.

The second scenario, called the sequential, full-information scenario, models the effects of bilateral trading in circumstances most likely to maximize cost savings under sequential trading. In the algorithm which simulates this scenario the first step involves creating a matrix of the cost savings from each possible pairwise *feasible* trade, assuming that only this trade takes place. (Feasible trades are defined as those which assure that the ambient standard is met after the trade as well as before.) The pairwise trade producing the largest cost savings is selected and assumed to be consummated. The emission and air quality vectors are updated to reflect the post-trade emission patterns, and the cost savings is recorded. The matrix is then recalculated for the remaining sources (eliminating the row and column corresponding to the two trading partners who have already consummated a trade). The recalculation of this matrix is necessary because the post-trade air qualities would have changed due to the rearrangement of emissions implied by the trade, changing in turn the set of feasible future trades. These iterations continue until the last feasible trade has been consummated.

The amount of the feasible trade for each scenario is calculated in several steps. For each potential buyer and seller, an attempt is made to determine the trade that would equalize marginal control costs. Since the fitted total cost function for the jth source from (1) is

$$\hat{C}_j = \hat{\beta}_j X_j^2 \qquad j = 1,\dots,27,$$

where $\hat{\beta}_j$ is the OLS estimator of β_j, we can obtain the fitted marginal cost function for the jth source as

$$\widehat{MC}_j = 2\hat{\beta}_j X_j \qquad j = 1,\dots,27.$$

Consider two sources who wish to trade emissions. Without loss of generality let source 2 be the low marginal cost source, so that

$$2\hat{\beta}_1 X_1 > 2\hat{\beta}_2 X_2. \tag{2}$$

Given the direction of the inequality, source 2 would sell permits to source 1. To find the traded amount, T, we assume trading takes place until the marginal costs are equal for each firm. Since this would make (2) hold as an equality,

$$2\hat{\beta}_1(X_1 - T) = 2\hat{\beta}_2(X_2 + T).$$

Solving for T yields

$$T = \left(\hat{\beta}_1 X_1 - \hat{\beta}_2 X_2\right)/\left(\hat{\beta}_1 + \hat{\beta}_2\right),$$

where the numerator is positive from (2).

An actual trade of size T may not be feasible, however, if the firm selling permits does not have enough control capacity with T added as part of its required control level or if additional emissions of T by the source buying permits would violate air quality. We address the first problem by requiring that a seller's previous control responsibility plus any additional responsibility as a result of a proposed trade must be no larger than its control capacity. We address the second problem by running the resulting emission levels through the diffusion model to assure that the trade is feasible (i.e., satisfies the ambient air quality standards). If the potential trade is feasible, the potential cost savings are recorded and the algorithm considers the next pair of potential traders. If the trade is not feasible, the size of the trade is reduced by 5%. The feasibility of the scaled-back trade is then checked by running the proposed post-trade emissions through the diffusion model to see if the ambient standards would be violated at any receptor. If the trade is still not feasible, the iterative process is repeated until a feasible trade is found or the proposed trade is ruled inadmissable after 20 unsuccessful iterations.

This sequential, full-information scenario is complemented by two partial information scenarios. Whereas the full information scenario assumes that the traders have complete information on the feasibility of all possible trades and the possible cost-savings to be derived from each, the partial information scenarios assume that the traders have more limited information. In particular the partial information scenarios differ in the manner by which the sequence of trades is determined. In the first of these scenarios the algorithm sequences trades by selecting the remaining firm with the lowest marginal cost as the next seller in each iteration, while in the second scenario the next trader is selected randomly.

The first partial information scenario begins by identifying the source with the lowest marginal control cost at its SIP allocation of control responsibility. This source is assumed to be the first seller of emission reduction credits. Once the "best" set of feasible trades is found for each of the remaining 26 buyers for this iteration by the process described above, the total costs savings attributable to each potential bilateral trade is calculated by subtracting the total cost of the post-trade allocation for that pair from the total cost of the SIP allocation. The trading partner which maximizes the cost savings is selected and that trade is assumed to be consummated. The emission vectors are then updated to reflect this trade and these two sources are removed from further trading consideration.

The algorithm proceeds by identifying the lowest remaining marginal cost source, assuming the first trade has been consummated. The "best" set of feasible trades with all remaining purchasers is calculated in the same manner as just described. The feasible trade which maximizes the cost savings for this iteration is selected and assumed to be consummated. The iterations continue until the set of remaining feasible trades is empty.

The second partial information scenario differs from the first in that it involves a random selection of the first seller of permits rather than selecting the source with the lowest marginal cost. This scenario is designed to capture the fact that considerations other than emissions trading frequently motivate further control (e.g., plant modification). This scenario also has the virtue that it allows us to test the sensitivity of the partial information scenarios to the assumption that the firm with the lowest marginal cost initiates the process.

In each run of this scenario the first seller is randomly selected using a discrete uniform distribution with range [1, 27].[13] The best trading partner is chosen by

---

[13] This is produced by transforming output from the IMSL subroutine GGUBFS, using initial seed 97783221.

calculating the potential cost savings from each possible feasible pairwise trade with that trader. Once this partner is identified, the trade is consummated, the cost savings calculated and the emissions and air quality vectors updated. A new random seller is then selected from the remaining list and the process is repeated. The iterations continue until no more trades are possible.

Because this scenario is potentially sensitive to the luck of the draw in selecting the first trader, the algorithm was run 500 times for each ambient air quality standard. For this scenario, therefore, we report the mean cost-savings achieved by these 500 runs and the standard deviation.

To discover the sensitivity of the results to the assumption that all firms are able to pick out the best *feasible* trade, the partial information scenarios are rerun assuming that the "best" trades are selected without considering ex ante whether the proposed trade would satisfy the ambient air quality constraints. Instead we select the most cost-reducing trade and then employ the above iterative scaling procedure if necessary to assure that the ambient standards are met. A failure to recognize infeasible trades prior to selecting a trading partner represents yet another potential source of foregone opportunities to produce cost savings. To distinguish them we refer to the scenario which assumes traders can determine the feasibility of any trade in advance as "partial information–feasible ex ante" and the less forward-looking scenario as "partial information–feasible ex post."

To complete the analysis, the control cost associated with the resulting post-trade allocation is calculated for each of the scenarios and compared with the cost associated with the corresponding SIP allocation and the least cost allocation. From this information we compute the percentage of the potential cost savings that this trading process is able to capture. (A cost-effective trading process would, by definition, capture 100%.)

RESULTS

The Constraint-Emissions Rule

Our first results, presented in Table I isolate the magnitude of the effect of EPA's constant-emissions rule, the requirement that no trade can result in an emissions increase. The column labeled "Primary" represents the air quality level corresponding to the primary ambient air quality standard while that labeled "Secondary" approximately represents the air quality level associated with the

TABLE I

The Effect of the Constant–Emissions Rule: Control Costs ($/Day) and % of Potential Cost Savings Achieved

	Secondary		Primary	
Scenario	$/day	% Savings	$/day	% Savings
Simultaneous, full information	15,531	66%	835	91%
Sequential, full information	18,852	50%	1,043	88%
Least cost benchmark	8,255	100%	188	100%
SIP benchmark	29,612	0%	7,284	0%

more stringent secondary standard. The entries for each scenario represent the control costs associated with that scenario and the percentages of the potential cost savings achieved. The least cost benchmark, which could only be achieved if all trades were multilateral and simultaneous, represents the best possible allocation.

These results strongly suggest that the EPA constant-emissions rule does bear some responsibility for the failure of the sequential trades to achieve a cost-effective allocation of control. When the air quality standards are stringent (as in our "secondary" case), the divergence from cost-effectiveness attributable to this factor can be significant. With this rule in place some 34% of the potential cost savings are lost immediately. The effect is less dramatic when the ambient air quality standards are less stringent (as in our "primary" case).

The sequential, full information scenario establishes the degree to which a bilateral, sequential trading process could, under the most optimistic of assumptions about the amount of information held by the traders, secure the remaining potential cost savings. The estimates suggest that a substantial proportion of the remaining potential cost savings could be achieved by a bilateral, sequential trading process, providing the traders are fully informed about the trading opportunities and those traders able to achieve the greatest cost savings are able to consummate the earliest trades. Even with the constant-emissions rule in effect trading could conceivably capture from 50% to 88% of the potential savings. This conclusion is not very sensitive to the stringency of the ambient standards.

Limited Information Trading

The sequential, full information scenario is rather optimistic, probably too optimistic, about both the amount of information available to traders and the optimal use of that information to sequence bilateral trades in an effective manner. To understand the sensitivity of these results to assumptions about the amount of information held by traders, the partial information–feasible ex ante and ex post scenarios were created.

The general conclusion which emerges from these results is that limited information and trader sequencing can have a large effect on the amount of savings captured by the trading process. As seen from the results for the partial information–feasible ex post scenarios (the scenarios which assume traders cannot anticipate the feasibility of the trade in advance), the ability to anticipate trades which will trigger violations of the ambient air quality standards is a particularly valuable type of information. In its absence, less than 20% of the potential savings is captured by three of the four scenarios.

Which of the scenarios are the most realistic? In our opinion the full information, sequential trading scenario should be viewed as an upper bound on the cost savings to be reasonably anticipated from the bubble policy while the partial information–feasible ex post scenarios should be viewed as a lower bound. The full information scenario is rather optimistic because it assumes that the trades producing the greatest cost savings always precede the others. Since installing many of the control devices modeled in this paper involves a considerable investment by each source, it seems reasonable to believe that the timing of those investments would be a function more of such firm-specific considerations as the strength of demand for the firm's product than of the size of the potential savings in the emissions trading market. To the extent this point is valid, the early trades would not necessarily be the trades producing the greatest cost savings. Once consummated early suboptimal trades tend to reduce future opportunities considerably.

Yet the partial information–feasible ex post scenarios are probably excessively pessimistic. They ignore the feasibility of the best potential trade in selecting the trading partners. It is not difficult to imagine that environmental managers have sufficient knowledge of air quality conditions in their area to avoid seeking trades which ultimately would be infeasible.

Perhaps the most surprising result is that for the primary standard the random trader scenario achieves a greater proportion of the potential savings than does the low MC trader scenario. This is not as counterintuitive as it may first appear. Selecting the firm with the lowest marginal cost to be the first seller in the partial information scenarios turns out to be a particularly ineffectual strategy because it fails to take account of the second derivative for this firm's cost function. If marginal cost rises rapidly with the amount controlled, the lowest marginal cost firm may generate little savings by trading with another firm. While this trading partner could potentially generate much greater savings by subsequently acquiring more ERCs from other sellers, any such trade is precluded since firms in the current period are eliminated from future trading. Because the full information scenario implicitly considers the second derivative by calculating the largest potential cost savings for each pair of potential traders prior to the first trade, it is able to achieve substantially greater savings.

It is also important to remember that the results in Table II for the random trader scenario are based upon the mean savings of 500 runs. Since the standard deviation of these runs is approximately one-half of the savings, many of the random runs do result in higher control costs.

Why do these trading strategies achieve such a low proportion of the potential cost savings? The empirical results suggest a number of reasons. First, some

TABLE II
Limited Information Scenarios: Control Costs ($/Day) and % of Potential Cost Savings Achieved

	Partial information–feasible ex ante		Partial information–feasible ex post	
	Secondary	Primary	Secondary	Primary
Random				
Mean ($/day)	24,354	3,854	26,004	4,012
Stan dev ($/day)	1,663	1,700	1,816	1,859
% Savings	25%	48%	17%	43%
Low MC				
Control cost ($/day)	21,249	6,314	25,670	6,754
% Savings	39%	13%	18%	7%
Least cost benchmark				
Control cost ($/day)	8,255	188	8,255	188
% Savings	100%	100%	100%	100%
SIP benchmark				
Control cost ($/day)	29,612	7,284	29,612	7,284
% Savings	0%	0%	0%	0%

cost-effective trades which could be consummated in a multilateral, simultaneous trading environment cannot be consummated in a bilateral, sequential trading environment because they would violate ambient air quality standards. Forcing individual trades to satisfy the air quality constraint, as is done by the current policy, is a much more restrictive condition than requiring only the final equilibrium to satisfy those constraints.

Second, both current practice and our simulations require that emissions must either be held constant or be reduced by each trade. Reaching the cost-effective allocation from the SIP allocation necessarily involves some trades that allow emissions to increase; these occur when the acquiring source is near air quality receptors that are already recording air quality significantly better than that required by the standards.[14] Current practice rules these cost-effective trades out.

Third, all of the partial information strategies require traders to secure all the ERCs they can from their trading partner. Greater cost savings could result if some of these were reserved for subsequent trades. Early traders do the best they can under the circumstances, but the inability to make additional future trades leads to a suboptimal solution.

POLICY IMPLICATIONS

According to our results the type of bilateral, sequential trading that occurs under the bubble policy results in substantially smaller cost savings than would be expected from the more sophisticated, but uncommon, cost-effective trading process that analysts have historically presumed to exist. Our results also suggest that a substantial proportion of this "expectations gap" may be explained by the bilateral, sequential nature of the trading process and the amount of information about market opportunities the traders possess.

Though our results show smaller cost savings than the traditional literature, in a sense they may be excessively optimistic because we have not included transactions costs in our trading algorithm. Emission reduction credit trades now frequently involve a broker. Brokers' fees are typically calculated as a percentage of the value of the transaction.[15] Add to this the costs of accomplishing the air quality modeling that is frequently required for trades involving noncontiguous sources of nonuniformly mixed pollutants and the transactions costs can become very large. These costs serve to further reduce the cost savings from trades and therefore the incentive to trade, particularly when the potential savings are small.

These results suggest one reason why a disproportionately high proportion of the existing trades have been internal (between entities sharing a common corporate parent). Internal trades avoid many of the problems discussed in this paper. Usually the trading units involved in internal trades are contiguous, thereby avoiding the problem of violating an air quality constraint near the acquiring source. Constant-emissions trades would be the norm for internal trades so the constant-emissions rule would play no role. Information on future credit creation and use is much easier to acquire within a firm. Unfortunately internal trades

---

[14] For evidence on this point see [4, 10, 15].

[15] This information supplied by Thomas Brooks, a broker at AER*X, at a presentation at the AERE Workshop on Natural Resource Market Mechanisms, Madison, WI, June 7, 1990.

represent a relatively small proportion of the total potential trades envisioned by the empirical studies on trading possibilities.

Though we have no specific evidence on the point, it seems reasonable to presume that the ability of emissions trading to achieve potential cost savings is greater for uniformly mixed pollutants than for nonuniformly mixed pollutants, such as the case modeled in this paper. Emissions trading of uniformly mixed pollutants does not face the additional constraints imposed by the need to locally monitor ambient air quality standards, since all emissions impact air quality equally. Furthermore since cost-effectiveness for these pollutants involves the same level of aggregate emissions as the SIP allocation, forcing all trades to hold emissions constant does not exact a cost-effectiveness penalty as it does for nonuniformly mixed pollutants. This presumption is reinforced by the evidence from studies of actual trading which suggests that much of the activity under the bubble policy involves uniformly mixed pollutants.[16]

The importance of this point is underscored by the nature of the gases which are possible candidates for emissions trading in the future. Many proposed applications of the emissions trading concept on the global scale involve pollutants that can be treated as uniformly mixed. For the gases responsible for both ozone depletion and global warming the amount of the gases emitted is an important policy consideration, but the location of those emissions is not. Global emissions trading markets are a distinct possibility.

It would be a mistake, however, to conclude that bilateral, sequential trades involving uniformly mixed pollutants would automatically be cost-effective. Although such trades avoid the cost of meeting the air quality constraints by holding emissions constant, they do not avoid the problem associated with early traders consummating suboptimal trades. In the absence of good information on future market opportunities, early traders may opt for trades which in the glare of hindsight serve to undermine greater subsequent savings. Sequencing can have a drastic effect on the total potential cost savings that can be achieved. This problem is particularly acute in the emissions trading market because of the lumpiness of the pollution control investments and the thinness of the markets.

The current regulatory environment exacerbates the inherent difficulties in this market. Most states have not established emission banks in which created ERCs could be stored for subsequent use or sale. Some states even confiscate any created credits that have not been used within a 2 or 3 year period. By denying sources the opportunity to save credits for subsequent sale, these policies encourage early sale by creating a "use it or lose it" tradeoff. Our results suggest that by altering the normal pattern of sequencing, the failure to establish emission banks could have large adverse effects.

The conclusion that the EPA constant-emissions rule should be eliminated does not automatically follow from our analysis. While our results do suggest that this rule diminishes the cost-effectiveness of the bubble policy, it is not without merit. Recent results suggest, for example, that the additional emission reductions attributable to this policy may yield sufficient benefits to make them worthwhile.[17] In addition for pollutants such as sulfur oxides, most of the extra emissions allowed

---

[16] See, for example, [17; p. 56].

[17] See [16].

under cost-effectiveness (but prohibited under current policy) are released from tall stacks and are transported long distances, causing acidification of distant lakes.

Our results also suggest a methodological conclusion in addition to the policy conclusions. The traditional approach of using mathematical programming models to characterize actual emissions trading equilibria may produce very misleading results. Actual trading processes have several characteristics which lead to equilibria which differ considerably from the programming equilibria. The existing studies have led to unrealistic expectations.

Virtually all regulatory reform goes through stages. An initial stage characterized by optimism dictated by high expectations is followed by a pessimistic stage arising from the realization that the high expectations were inflated. The third stage, the one we believe we are now in, is characterized by the realization that, though not perfect, the reform is still an improvement, indeed a substantial improvement, on what it replaced. Only by learning from the nature of its initial shortcomings can the reform ultimately be targeted and tailored to those circumstances where it works best.

REFERENCES

1. Scott E. Atkinson and Donald H. Lewis, A cost-effectiveness analysis of alternative air quality control strategies, *J. Environ. Econom. Management* **1**, 237–250 (1974).
2. Scott E. Atkinson and T. H. Tietenberg, The empirical properties of two classes of designs for transferable discharge permit markets, *J. Environ. Econom. Management* **9**, 101–121 (1982).
3. Scott E. Atkinson and T. H. Tietenberg, Approaches for reaching ambient standards in nonattainment areas: Financial burden and efficiency considerations, *Land Econom.* **60**, 148–159 (1984).
4. Scott E. Atkinson and T. H. Tietenberg, Economic implications of emissions trading rules for local and regional pollutants, *Canad. J. Econom.* **20**, 370–386 (1987).
5. William J. Baumol and Wallace E. Oates, The use of standards and pricing for protection of the environment, *Swedish J. Econom.* **7**, 42–54 (1971).
6. Daniel J. Dudek and John Palmisano, Emissions trading: Why is this thoroughbred hobbled? *Columbia J. Environ. Law* **13**, 217–256 (1988).
7. Robert W. Hahn and Gordon L. Hester, Where did all the markets go? An analysis of EPA's emission trading program, *Yale J. Regul.* **6**, 109–153 (1989).
8. Robert W. Hahn, Economic prescriptions for environmental problems: How the patient followed the doctor's orders, *J. Econom. Perspect.* **3**, 95–114 (1989).
9. Alan J. Krupnick, Wallace E. Oates, and Eric Van de Verg, On marketable air pollution permits: The case for a system of pollution offsets, *J. Environ. Econom. Management* **10**, 233–247 (1983).
10. Alan J. Krupnick, Costs of alternative policies for the control of nitrogen dioxide in Baltimore, *J. Environ. Econom. Management* **13**, 189–197 (1986).
11. Richard A. Liroff, "Reforming Air Pollution Regulation: The Toil and Trouble of EPA's Bubble," Conservation Foundation, Washington, DC (1986).
12. D. O. Martin and J. A. Tikvart, A general atmospheric diffusion model for estimating the effects on air quality of one or more sources, *J. Air. Pollution Control Assoc.* **18**, 68–148 (1968).
13. Albert M. McGartland and Wallace E. Oates, Marketable permits for the prevention of environmental deterioration, *J. Environ. Econom. Management* **12**, 207–228 (1985).
14. W. David Montgomery, Markets in licenses and efficient pollution control programs, *J. Econom. Theory* **5**, 395–418 (1972).
15. Wallace E. Oates and Albert M. McGartland, Marketable pollution permits and acid rain externalities: A comment and some further evidence, *Canad. J. Econom.* **18**, 668–675 (1985).
16. Wallace E. Oates, Paul R. Portney, and Albert M. McGartland, The *Net* benefits of incentive-based regulation: A case study of environmental standard setting, *Amer. Econom. Rev.* **79**, 1233–1242 (1989).

17. T. H. Tietenberg, "Emissions Trading: An Exercise in Reforming Pollution Policy," Resources for the Future, Inc., Washington, DC (1985).
18. T. H. Tietenberg, Economic instruments for environmental protection, *Oxford Rev. Econom. Pol.* **6**, 17–33 (1990).
19. TRW Systems Group, "Air Quality Implementation Planning Program; Vol. 1, Operator's Manual," Washington, DC (1970).
20. United States Department of Health, Education and Welfare, "Control Techniques for Particulate Air Pollutants" (1969).