DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

THE FEASIBILITY OF MARKETABLE EMISSIONS PERMITS IN THE UNITED STATES

Roger G. Noll

Prepared for the Regulation Conference,
International Institute of Management
Berlin, Federal Republic of Germany, July 1981

SOCIAL SCIENCE WORKING PAPER 397

July 1981

ABSTRACT

Economists have long advocated the use of economic incentives, rather than detailed regulations, as a means for combatting environmental pollution. In the late 1970s, environmental regulators in the United States began experimenting with one such method — emissions permits that, within important limits, can be traded among sources of pollution. This paper explores the feasibility of an extreme version of the marketable permits approach, one in which all source-specific regulations are replaced with tradable emissions permits. First, the general argument for a marketable permits system is presented, including a discussion of the legal procedures that are required by each of the major alternative methods for effecting improvements in environmental quality. Then, the implementation problems of a permits market are explored. Because this is partly an empirical problem, this analysis is presented in the context of an example: particulate sulfates in the Los Angeles atmosphere. Finally, some specific design possibilities are presented, and compared to the early experiments with tradable permits.

THE FEASIBILITY OF MARKETABLE EMISSIONS PERMITS IN THE UNITED STATES*

Roger G. Noll

Environmental Quality Laboratory
California Institute of Technology

In the decade of the 1970s, the United States Government vigorously pursued a policy to reduce substantially the pollution of the nation's air and water. Beginning with the Clean Air Act in 1963, a series of laws were passed that expanded the role of the federal government in setting and enforcing environmental policy goals, that set substantially more ambitious goals in terms of the purity of air and water, and that increased dramatically the resources available for writing and enforcing pollution regulations. Although all the returns are not in and considerable disagreement about the exact figures remains, a reasonable estimate of the cost of environmental regulation is in the range of thirty to fifty billion dollars annually. In some industries, such as chemicals, electric power generation and mining, the cost of compliance with environmental regulation accounts for as much as 10 percent of investment expenditures, and 5 percent of

* The research reported in this paper was supported by a grant from the California Air Resources Board. The author gratefully acknowledges the helpful comments and criticisms on an earlier draft by Glen R. Cass and Robert W. Hahn.

annual sales.[1]

Whether these expenditures are worthwhile depends, of course, on the benefits from environmental regulation. As with costs, there is some disagreement about the effect of environmental policies on pollution, but the consensus is that, while progress has been made, the nation is still a long way from attaining the stated goals of environmental policy.[2] Indeed, there is enough agreement on this point that all of the major groups in the environmental policy arena -- business, environmental protection organizations, politicians, and regulatory authorities -- now generally favor substantial reform of the regulatory process. And the most widely supported direction of reform is to introduce a greater measure of decentralized decisionmaking, guided by economic incentives, as a substitute for regulatory rules promulgated by government.

This paper reviews one area of environmental policy — air pollution controls -- in which the transition to a more incentive-based regulatory mechanism has been underway since 1977, when amendments to the Clean Air Act enabled the Environmental Protection Agency (EPA) to adopt a more flexible approach to regulation than the traditional method of setting standards for each specific source of pollution. EPA's new approach is commonly called marketable (or

---

[1] For some data on these points, see Cost of Government Regulation Study for the Business Roundtable (Arthur Anderson and Co., March 1979).

[2] Annual reports of the U.S. Government's Council on Environmental Quality show steady but slow decline in the concentration of most air and water pollutants.

tradable) emissions permits, whereby pollution sources can buy, sell or trade emissions of various pollutants as long as the overall objectives of environmental policy are not sacrificed. To understand fully how marketable permits are being implemented requires some knowledge of the details of the standard-setting process that is being replaced; hence the first section of the paper describes the most important features of the old approach. This section is intended to summarize a vast literature on American regulatory institutions. The second section of the paper discusses the advantages and problems associated with implementing more decentralized, market-oriented approches to environmental regulation. Most of the important questions about implementation depend upon empirical, rather than theoretical, issues, and so the discussion in this section is in the context of a particular case study: the regulation of sulfur oxides emissions in the Los Angeles air shed. The third section of the paper describes the methods that EPA has adopted for making the transition to a marketable permits system, and evaluates the progress to date. This section also contains some conclusions about how best to implement this change.

I. AIR POLLUTION REGULATION IN THE UNITED STATES

The environmental legislation that was passed in the United States between the mid-1960s and late 1970s was but a part of a larger political change. For a number of reasons, the issue of preventive health and safety measures became of high importance to political leaders. At the same time that environmental laws were being written, so, too were tough laws regarding consumer products and occupational health and safety. The environmental laws that were passed in this context were defended on the grounds that pollution represented a significant threat to public health.[3]

In this milieu, the principal objective of air pollution controls became to reduce emissions to the point that they produced no deleterious health effects. Federal regulators were charged with the responsibility of establishing ambient air quality standards -- e.g., maximum limits on the concentration of various pollutants in the atmosphere — that were low enough to guarantee the absence of hazards to health. This is, of course, a highly controversial policy objective, for it implies an infinite value to human health: *any* expense is justified to reduce air pollution as long as *any* adverse health effect, however minor, is observed. Moreover, if there is no threshold of pollution below which there are no health effects, the implication of the policy is zero emissions -- a goal that probably is technically impossible to achieve in a modern, industrialized society. Because of the difficulty of determining the maximum concentration that produces no health effects, very few ambient air quality standards have yet been established. Nevertheless, the goal of zero adverse health effects remains the objective of air pollution regulation and the basis for ambient air quality standards.

---

[3] For a detailed study of this issue, see Lester B. Lave and Eugene P. Seskin, *Air Pollution and Human Health* (Baltimore: Johns Hopkins University Press for Resources for the Future, 1977).

Setting Source-Specific Standards

For regions that do not satisfy ambient air quality standards, the next step is to develop a plan for achieving compliance. For regions that do meet the standards, the next step is to adopt a policy for preventing significant deterioration of air quality.

In this phase the full complexity of the American federal system is apparent. The role of the federal government is essentially to define the characteristics of an acceptable plan, allowing the states to work out the details within the federal guidelines. The resulting state strategy for reducing pollution is called a State Implementation Plan (SIP), and it must be approved by EPA or else the latter can take over air quality regulation in the state.

For areas that are not in compliance with ambient air quality standards, the basic approch of the SIP is to specify abatement standards for each source of pollution, where a source is normally defined as a point at which emissions occur. Thus, a complicated production facility, such as an oil refinery or a steel mill, may have several points at which emissions are released, and hence be regulated as several different independent sources. Federal ground rules require that states identify the significant sources, develop regulations for controlling their emissions, and adopt a timetable for bringing these sources into compliance.

A major problem in developing source-specific regulations is that the relationship between the amount and pattern of emissions and measured ambient air quality is often quite complex and poorly understood. In only a few parts of the nation has the relationship between emissions and air quality been estimated relatively well. In most areas, the basic data that are necessary to estimate such a relationship — long-term time series of emissions and air quality measurements -- either are unavailable, are unreliable, or have only been collected for a few years. As a result, the relationship between source-specific standards and the extent to which gains are made in achieving ambient air quality standards is highly uncertain. In response to this difficulty, EPA allows a state to adopt a SIP without direct reference to whether the resulting emissions will satisfy ambient air quality standards. If a region does not satisfy the ambient air quality standards, its plan is acceptable if it requires that sources adopt the best available control technology by a specified target date, usually in the mid to late 1980s. The presumption is that this will attain air quality objectives. If not, state regulators must later adopt more draconian measures, such as shutting down some polluting facilities or curtailing automobile travel; however, whether these measures will ever actually be imposed is a matter of uncertainty and disagreement.

Regardless of whether a region is in compliance with air quality standards, a plan must be developed by state and local authorities for preventing further substantial deterioration of air quality due to the entry or expansion of polluting facilities. This, too, in practice amounts to setting in place a process for writing source-specific standards. Although the definition of what constitutes substantial degradation varies among regions, in many cases the standards for new sources amount to a requirement that there

be no net increase in emissions — that is, a net zero discharge requirement for new sources.

The SIP procedure gives state and local regulatory authorities the job of identifying the appropriate standards for each source, old and new, in their jurisdiction. Because the best abatement technology normally depends on specific technical features of a source, in most instances general standards for a broad class of sources are not attempted. With a few exceptions, state and local governments are free to identify the best abatement technology for each source, given its special characteristics. The few exceptions are standards for extremely important sources that the federal government has decided to set itself, such as emissions limits for automobiles and for newly constructed coal-burning electric generation facilities -- although even in these cases the states can set their own standards as long as they set even more demanding regulations than the federal standards.

In principle, two kinds of source-specific standards can be set. Input standards specify the technical method to be adopted by the firm for reducing emissions. Examples are requiring facilities that burn coal or oil to install stack-gas scrubbers or specifying the maximum permissible amount of sulfur and other impurities that are allowed in fuel. Performance standards specify the maximum permissible emissions from a source, but allow the regulted entity to choose whatever control technology it prefers so long as the emissions ceiling is not exceeded.

In practice, source-specific standards as they have been developed in the United States are a hybrid of these two approaches.

Regardless of how a standard is formally expressed, the method for developing it is to try to identify the best control technology and to establish the emissions that would result if that technology were used. Even if the standard is expressed in performance terms, a firm that prefers to use another method to achieve its allowed emissions level normally must obtain advance approval. The regulatory authorities will require proof that the alternative method is as effective as the one upon which the standard was based.

There are two reasons for a heavy reliance by regulators on specifying the abatement technology to be adopted by polluters. One is that the states must convince EPA that the SIP represents a good-faith effort to achieve air quality objectives. Because relatively few areas know the relationship between emissions and air quality and, therefore, can make reasonably accurate predictions of the air quality results of their SIP, the most effective strategy for demonstrating that a SIP is reasonable is to show that all of the sources will be required to use the best technology for reducing emissions. A second reason is that source-specific standards must be developed in a formal regulatory process according to a set of procedural standards.[4] An enormous body of law in the United States -- Constitutional, statutory and common — protects the private sector against arbitrary and unreasonable confiscation of wealth by the government. Decisions of a regulatory authority must be based upon the evidence submitted in a

[4] For a more complete discussion of the nature of the regulatory process and how process affects policy, see Roger G. Noll, "Breaking Out of the Regulatory Dilemma: Alternatives to the Sterile Choice," Indiana Law Journal 51, no. 3 (1976):686-99.

formal, public process in which anyone who is significantly affected by the decision has an opportunity to participate. The decisions must be reasonably and rationally based on the evidence, and must be clearly derived from the statutory responsibilities of the regulator. The regulatory authority bears the burden of proving that the regulation is not capricious and arbitrary (e.g., is rationally based upon substantial evidence).

The decisions of the regulator can be appealed in the courts. To withstand these legal challenges, regulators must be able to prove the reasonableness of each standard. Normally this means that the regulator must show that the standard is feasible, and that it will make progress towards reducing emissions. If the regulator fails to show that emissions will be substantially reduced (e.g., a performance standard) by a specific, verifiable technique (e.g., an input standard), the regulated firm and environmentalist groups have grounds on which to appeal the decision in the courts. In fact, appeals of standards are commonplace, in part because the evidentiary basis is often shaky, but also in part because firms can delay the imposition of a costly standard by exercising their full rights of appeal through the judicial system.

Problems of the Present Approach

The standard-setting regulatory process has proven to have a number of important shortcomings.[5]

[5] For a more complete discussion of the problems of the present regulatory system, see Allen V. Kneese and Charles L. Schultze, Pollution, Prices and Public Policy (Washington, D.C.: The Brookings Institution, 1975).

First, the process is expensive and time-consuming, requiring numerous technical studies by all of the adversaries in a regulatory process, rebuttals of these studies, and formal hearings. The costs and delays inherent in the process have made progress in developing SIPs much slower than was anticipated. Moreover, the attention on technical, source-specific standards has deflected attention -- and resources -- away from developing good data on emissions and air quality, and hence has inhibited air quality modeling. Thus, over a decade after having begun to regulate air quality seriously, most local and state regulators are still largely in the dark about the likely effects of the regulations that are being proposed and adopted.

Second, the process is economically inefficient in that vast differences emerge among firms in costs per unit of abated emissions. Some production processes are easier to understand than others, and the costs and efficacy of abatement techniques are known with differing degrees of certainty. As a result, the maximal feasible extent to which regulators can be successful legally in forcing abatement varies from firm to firm. These differences are especially great between old and new sources of emissions.[6]

Third, the policy erects entry barriers against new and expanding firms. These barriers are of two forms: the generally higher standards applied to new sources, and the requirement to go through a time-consuming process in which production plans are made a

[6] See Bruce A. Ackerman and William T. Hassler, Clean Coal/Dirty Air (New Haven: Yale University Press, 1981).

public record that is available to competitors. To the extent that entry is retarded, competition and technological change are inhibited.

Fourth, the process creates perverse incentives with respect to technological innovation in abatement methods.[7] Because of the repermitting process, the adoption of an innovation is delayed until regulatory approvals are obtained for each source that can use it. Moreover, the expense of the process detracts from the attraction of cost-reducing abatement technologies. And, while manufacturers of abatement equipment have an incentive to invent ways to increase the abatement of emissions, polluting firms have no incentive to reduce their emissions below their currently active standard. Consequently, the latter can be expected to resist all improvements in abatement technology that involve any additional expense. In addition, the incentives for improving abatement technology focus primarily on separable technical fixes — e.g., specific pieces of equipment or other changes in input -- as opposed to innovations in manufacturing processes because the formal regulatory process is better equipped to deal with the former.

As a practical matter, elimination of the quasi-judicial regulatory process is not possible in the United States, deriving as it does from Constitutional principles having widespread political support. But even if it were possible to simplify the process and

[7] For a well-documented sample, see Richard E. Ayers, "Enforcement of Air Pollution Controls on Stationary Sources under the Clean Air Act Amendments of 1970," Ecology Law Quarterly 4, no. 3(1975):441-78).

make it more like European environmental regulatory processes,[8] some of the problems of standards remain. The absence of appropriate economic incentives to improve the methods of emissions control and the dependence on officials outside of polluting entities to make technical abatement decisions create situations in which inefficient results are likely. The delays and formal burden of proof in the American system contribute to these problems, but are not the only important cause of them.

II. ADVANTAGES AND PROBLEMS OF TRADABLE EMISSIONS PERMITS

Tradable emissions permits are one example of a general category of environmental regulatory methods that rely on market incentives, rather than source-specific regulations, to achieve policy objectives. The other leading examples are emissions taxes and abatement bounties.

Tradable emissions permits are somewhat related to the regulatory system currently in place. Source-specific standards either state directly or imply a legal ceiling on emissions rates for each source. Thus, they can be interpretted as permits to release given amounts of emissions. A natural way to conceive of a tradable emissions permits system is as a modification of the present implicit

[8] For descriptions of European regulatory systems with contrasts to the American process, see Blair T. Bower, "Mixed Implementation Incentive Systems for Water Quality Management in France, the Ruhr and the United States," Conference on International Comparisons of the Implementation Process in Environmental Regulation, March 1981; and Steven Kelman, Regulating America, Regulating Sweden (Cambridge, Mass.: MIT Press, 1981).

emissions permits in which firms in a region can rearrange the pattern of emissions by exchanging their permits. Firms would report to regulators changes in their permit holdings in order to facilitate enforcement; however, they would not need advance approval for transactions.

Emissions taxes are a charge per unit of emissions that is levied on all sources in a region. An abatement bounty is like an emissions tax, except that firms are paid a subsidy in proportion to the emissions reductions that they achieve.

The underlying principle of all of these systems is the same. Polluters are given an incentive to reduce emissions, but the choice of the abatement technique is left to their discretion. And, because all polluters face the same set of incentives, rational decentralized responses will lead to the attainment of any given regional emissions target at minimum total costs.

A second advantage of incentive-based approaches is that they avoid a major part of the process costs of controlling pollution. With taxes, bounties or tradable permits, government need not be in the business of specifying the technology of abatement or setting the emissions limits for each specific source. The principal task of regulators under incentive-based systems is to define overall objectives in terms of environmental quality and total emissions, and to enforce compliance with the system that generates the incentives, including the measurement of performance by each source. Both of these responsibilities are, of course, features of the system of source-specific standards.

A third advantage of incentive-based approaches is that they do a better job of promoting advances in abatement technology. In all three systems, emissions are costly to a firm because they result in either taxes, the necessity to hold monetizable emissions permits, or a foregone opportunity to collect a bounty for reducing them. Consequently, polluters have a continuing incentive to search for less costly, more effective abatement methods. Thus, a firm that invents a better abatement technique finds a willing market, rather than a source of opposition to the adoption of the new method.

Advantages of Tradable Permits

The special advantages -- and problems -- of tradable emissions permits have been widely discussed in general, theoretical terms in the economics literature.[9] Only a brief review will be presented here.

One advantage of tradable permits in comparison with taxes and bounties is that they do not necessarily require involvement in the fiscal processes of government. The source-specific standards approach gives away its implicit emissions permits, and so, too, can a system in which the permits are tradable. This avoids the political issues associated with either taxing or subsidizing industry. Indeed, literally any wealth effect of taxes or bounties can be reproduced by

[9] See J. H. Dales, Pollution, Property and Prices (Toronto: University of Toronto Press, 1968); W. David Montgomery, "Markets in Licenses and Efficient Pollution Control Programs," Journal of Economic Theory 5 (1972):395-418; and Thomas H. Teitenberg, "Transferable Discharge Permits and the Control of Stationary Source Air Pollution: A Survey and Synthesis," Land Economics 56(1980):391-416.

the choice of allocating permits: auctioning them is equivalent to an emissions tax, while granting permits equal to original emissions and then purchasing them is equivalent to an emissions-reduction bounty. Thus, the tradable permits approach allows a separation between the equity issues of who should pay and the efficiency issue of how to minimize the costs of achieving the policy objective.

A second characteristic — not necessarily an advantage or disadvantage -- of tradable permits has to do with the ways in which the uncertainties of the regulatory process are distributed.[10] Like the system of source-specific standards, tradable emissions permits specify the total emissions of a given pollutant in a given geographic area. Unlike source-specific standards, the geographic distribution of emissions — and hence of pollution — is not specified. Consequently, in the absence of a good model for estimating the distribution of emissions that a market would produce and another good model for estimating the effects of this pattern of emissions on pollution, tradable emissions permits are somewhat more uncertain than source-specific standards in terms of the resulting quality of the environment. Similarly, in the absence of the same kind of modeling capability, a given tax on emissions or bounty on emissions reductions will produce uncertainty in the distribution of emissions — and on

[10] For a more complete treatment of this subject, see A. Michael Spence and Martin L. Weitzman, "Regulatory Strategies for Pollution Control," in Anne F. Friedlander, editor, Approaches to Controlling Air Pollution (Cambridge, Mass.: MIT Press, 1978); and Marc J. Roberts and A. Michael Spence, "Effluent Charges and Licenses under Uncertainty," Journal of Public Economics 5 (1976):193-208.

total emissions as well, since neither system imposes limits on total emissions. Hence, taxes and bounties are more uncertain than tradable permits in terms of their effects on environmental quality.

A similar story can be told on the cost side. In the American system of setting hybrid standards having both performance and input dimensions, compliance costs are estimated as part of the regulatory process. These estimates vary, sometimes dramatically, so they must be regarded as uncertain even after the standard is adopted. A tradable permits system in which the implicit permits in current standards are made marketable has greater cost uncertainty; however, because the permits are voluntarily exchanged, the greater uncertainty arises because polluters may find an amount of cost-reducing trades of permits that is unknown in advance. Hence the added uncertainty arises solely because of the possibility of lower costs.

In the case of taxes and bounties, a distinction must be made between resource costs and the expenditures on taxes or bounties. Because the amount of abatement resulting from any given tax or bounty is uncertain, so, too, is the resource cost of abatement. However, the total cost is bounded by the product of the tax (or bounty) and the initial amount of emissions. If firms do opt for zero discharge, it is because the cost of complete abatement is less than the maximum possible tax or bounty payment. A similar argument shows that the expenditures on taxes or bounties are also uncertain, but bounded in the same fashion.

From the standpoint of a firm, the sum of emissions taxes and abatement costs is likely to be less uncertain than the sum of

abatement costs and the costs of obtaining permits in a system of marketable permits. The announcement of an emissions tax conveys more information about potential tax liabilities and abatement costs than is conveyed by the announcement of the tradability of emissions permits. The reason is that in the latter case the price of the permits must still be determined by the market and will always have some degree of random variability.

A third feature of tradable emissions permits is that they have advantages with respect to enforcement requirements. Taxes and bounties -- at least in their most efficient forms -- depend on total emissions. Hence the quantity of emissions must be measured in a manner that produces a legally enforceable measurement of total quantity. In some cases, direct and continuous measurement can be avoided, such as by making occasional spot checks of processes that produce a constant rate of emissions or by performing a mass balance analysis on inputs and outputs to the production process. But in most instances the measurement requirements for estimating total emissions are quite demanding.

Tradable permits do not require an estimate of total emissions, but a determination of whether the firm is releasing more or less emissions than are permitted. Spot checks, when backed by appropriately calculated noncompliance fines that make compliance an optimal strategy, are more likely to be feasible with tradable permits than the other incentive-based methods. Even if continuous monitoring is adopted, the technical requirements are easier, for all that need be measured is whether emissions are above or below a given limit.

Mechanisms that are very simple, even simpler than the methods now in use -- e.g., optical scanning of stack emissions, or chemical dosimeters like litmus paper or the radiation detectors worn by workers in areas that can be contaminated by radiation -- can be used to detect whether a firm is in compliance without measuring the actual amount or emissions. In this way, tradable permits are comparable to any performance-based, source-specific emissions standard. And, like performance standards, tradable permits usually present more difficult enforcement problems than input standards because the latter require only an observation that a required technical fix is in place.

A fourth characteristic of tradable permits is that they can relatively easily accommodate economic growth. As economic expansion in a region proceeds, the optimal amount of emissions — and the optimal pattern of abatement among sources -- is likely to change. A system of source-specific standards cannot readily accommodate this reality, because to do so requires rewriting the standards for every source in the region. Consequently, present procedures and policies impose far heavier burdens on new and expanding firms than on old sources, with dileterious effects on economic change as described above. Emissions taxes, too, must be adjusted as the economic structure of a region changes, for if not environmental quality will deteriorate in direct proportion to economic growth -- a result that is not likely to be correct. If firms are allowed to enter before the tax is adjusted to their presence, legal problems can arise, not to mention political problems, if the new level of the tax makes the entrant economically unviable. In any case, the ex ante emissions tax

gives false signals to potential entrants. Moreover, unless the emissions consequences of the entrant are examined fairly closely, as in the present source-specific standards, regulators do not know how the tax should be adjusted. Abatement bounties present similar problems, but in addition they actually encourage the entry of polluting industries into heavily polluted areas.[11] If a firm can expect its abatement costs to be subsidized, it will have diminished incentive to take into account the social costs of environmental degradation in selecting a location for a facility. Indeed, if the abatement bounty is high enough to produce voluntary abatement, investments to control emissions can enhance profitability by the maximal amount only in heavily polluted areas (where presumably the bounty is highest). As with taxes, the solution to the problem is ex ante review of entrants to decide through a regulatory process whether entry is desirable, thereby undermining the advantages of an incentive-based regulatory system.

Tradable emissions permits allow entry to occur without advance approval by the acquisition of permits in the market. Regulators may then, ex post, adjust the number of permits, or may even have a long-term plan for gradually changing the number of permits in order to accommodate growth.

In the context of existing regulatory policy in the United States, with ambient air quality standards based upon the goal of

---

[11] For a comparison of the effects of taxes and bounties on entry, see R. Talbot Page, "Failure of Bribes and Standards for Pollution Abatement," Natural Resources Journal 13 (1976):677-704.

eliminating adverse health effects, environmental quality will not be allowed to deteriorate with economic growth, even though this might be the more efficient option. The advantages of a system of tradable permits in this case are even greater, for it is the only method that can accommodate entry without ex ante review of the emissions of the entrant. In order to prevent an increase in total emissions, taxes and bounties would have to be adjusted for all sources by an amount necessary to leave emissions unchanged. Thus, only tradable permits escape source-by-source regulatory review of each new or expanding polluting entity.

The preceding discussion suggests that the case for experimenting with tradable emissions permits is strong, justifying close examination of the implementation problems associated with such a system. Indeed, there are some aspects of a tradable permits approach that raise important design questions.

Design Problems for Tradable Permits

The main purpose of a tradable permits system is to convey to polluters — new and old — appropriate price signals about the social cost of emissions so that each can select a combination of capital investments, operating practices and emissions releases that minimize the sum of abatement costs and permits costs. The economic efficiency of the system depends on firms being able to buy and sell permits relatively easily, with incidental transactions costs, at competitive prices. The principal implementation problems associated with a tradable permits system are related to the question of whether these conditions for an efficient market can be satisfied.

One problem is the possibility of "thin" markets — that is, markets in which transactions are rare, and in which few firms are willing to buy or sell. In such a situation, the transactions costs of trading permits can prevent the market from being much of an improvement over source-specific standards. If a firm that seeks to buy permits must invest substantial time and resources in finding a potential trading partner, and then engage in bilateral negotiations to determine a price, the ability of the permits market to find a cost-minimizing total cost of achieving ambient air quality standards is undermined. Moreover, infrequent trades arranged through negotiations are less likely to convey clear price signals to potential entrants, firms contemplating expansion, or sources considering further abatement and the sale of some emissions permits.

A second problem is related to the structure of the permits market. In some air sheds, one or two firms can account for a very large share of emissions. Moreover, there is some tendency for regulators to require somewhat greater abatement efforts from the largest firms. In this situation, if a tradable permits system is initiated by making tradable the emissions permits that are implicit in current standards, it is conceivable that only one or two firms will be seeking to buy permits, with all other firms seeking to be sellers. If so, the market may not settle on the competitive equilibrium price, but a monopsonistic price instead. More generally, the degree to which a market diverges from the competitive ideal depends on the initial allocation of permits, and in any situation it is technically possible to pick an initial allocation that produces a monopoly or a monopsony.[12] Thus, a design problem for a tradable permits market is to avoid an initial allocation that has this property.

A third problem has to do with the definition of markets and permits. As discussed briefly above, the relationship between emissions and pollution is often very complex. Pollution at any given receptor point is the consequence of emissions from several locations, and orten depends on their interactions as well. Similarly, every source of pollution has a unique pattern of polluting effects, which, because or interactions, may also depend on emissions from other sources. In general, to achieve theoretical efficiency (ignoring transactions costs and possible market imperfections) requires a separate market for each point where pollution damage occurs, and a separate transformation function for each source of pollution that maps its holdings in pollution permits at any source to its emissions allowances. Of course, this degree of complexity is impractical to implement. Hence, an important design problem is to make simplifications in the definition of permits and regions in which permits are valid that do not sacrifice too much in the way of the potential efficiencies ot a market mechanism. At one extreme, a large geographic region can be treated as one market, with the implication that the region will be treated as one large mixing bowl in which emissions from all sources are uniformly spread across the region. As

---

12 For a complete analysis of this problem, see Robert W. Hahn, "An Assessment of the Viability of Marketable Permits," Ph.D. Dissertation, California Institute of Technology, May 1981.

a description of reality, no pollution problem -- not even emissions into standing bodies of water -- has this fully mixed property; however, as a practical matter it may be a workable assumption. A somewhat more complicated strategy is to define a few receptor points at which pollution is measured and require firms to purchase emissions permits for pollution at each receptor point where their emissions cause pollution.

The best way to organize the market -- the definition of a permit and the sources that must hold it -- depends only in part on the physical aspects of the pollution problem. It also depends on the economic incentives operating upon sources. If abatement cost functions for all sources lead to more or less the same degree of abatement (e.g., they are all reducing emissions by roughly the same proportion), a permits market that is defined crudely, even wildly incorrectly, as a mixing bowl may still be workable. In the worst case -- in which each receptor point is polluted by only one source -- the cost-minimizing distribution of emissions may still produce approximtely the same amount of abatement at all sources.

In most regions, pollution problems exhibit both kinds of characteristics: localized, single-source pollution, and effects from the combined emissions of many sources. A plume from a smokestack may be the primary cause of pollution on receptors a few miles downwind, but as distance from the stack increases its emissions will mingle with the releases from other facilities. To take an extreme example, the problem of acid rain in Canada, New York and New England is probably the cumulative etfect of emissions from literally thousands of sources, some more than a thousand miles away. Whether a tradable permits market is workable, then, depends on the relative importance of the local versus long-distance effects, and on the likely pattern of abatement that will emerge from the market.

A fourth issue in the design of a tradable permits system is its flexibility with respect to changes in ambient air quality or total emissions targets. Because the relationship between emissions and air quality and the effect of air quality on health are not well understood, there is a good chance that new knowledge will cause regulators to want to change emissions levels. A decision to create more permits is relatively straightforward to deal with; regulators can give away or sell some net increment to the total emissions rights in an area. But a decision to reduce the number of permits raises potential difficulties. The heart of the issue is still another dimension of the definition of an emissions permit. Is its lifetime perpetual, or of fixed duration? Can it be redefined by fiat, or as an outcome of a regulatory process, or must changes in the number of permits be accomplished by purchase by the state? Obviously, the ease with which the number of permits can be changed depends on the answers to these questions. Moreover, a constraining factor on building into the permits system a mechanism for changing the number of permits is the effect of the mechanism on the willingness of firms to hold permits. If polluting entities are made to believe that the value of an emissions permit is subject to significant change at the whim of the state, abatement strategies — in terms of both the amount of abatement and its distribution between long-term capital investments

and changes in operating methods — are likely to be affected.

Finally, some account needs to be taken of so-called air pollution episodes: periods when meteorological conditions are exceptionally unfavorable, and so air pollution builds up over a number of days. To limit emissions to a level consistent with good air quality on these worst days is irrational; it is far less costly to curtail economic activity for a few days a year than to build in abatement capacity that would keep air quality high regardless of the weather. The current practice is to announce the degree of unfavorability of conditions a day in advance, and to invoke special regulations when conditions look especially bad. To do something much more complicated than this is of dubious value, because the frequency and magnitude of our pollution episodes is not very high, and will be lower still as limits on emissions are lowered.

The tradable permits system could easily adopt the present approach to episodes, with the emissions permits applying only in the vast majority of days when there is no special condition. Alternatively, separate emissions permits markets could be implemented, one for normal conditions, and one or more for episodes, with regulators announcing each day which permits apply tomorrow. Because this problem is relatively easy compared to the other four, it will be ignored for the remainder of this paper.

Variants in System Design

The design features available to attack the first four problems are as follows.

1. Permit Life. Regulators could elect to make the durability of emissions permits uncertain by stating that they were valid until a formal regulatory procedure declared them to be invalid or changed the amount of emissions allowed by a single permit. Such a system would create incentives among firms to adopt production methods with some flexibility in emissions, and to hold more permits than were actually used. Alternatively, regulators could define the time period in which a permit is valid. At one extreme, permits could be perpetual, requiring regulators to buy them back to reduce total emissions. Or, regulators could assign a fixed life. If regulators decided to alter the number of permits, they could do so by allowing firms to trade in old permits for new at a specified exchange rate. Finally, regulators could have several different kinds of permits: some perpetual, some of a fixed, long-term duration, and some with a short life (e.g., one year). Some periodic variability in the number of permits could be accomplished through the process of reissuing the permits with the short life; somewhat greater variability could be introduced as the intermediate-duration permits expired.

2. Market Definition. An emissions permit pertains to a particular geographic area. The size of the region and the variety of permits a source must hold for a given emissions allowance is a design feature of the system. Regulators could define emissions permits as freely tradable among all sources in a wide geographic area. Alternatively, a region could be subdivided into smaller areas, with trades between areas either barred or permitted according to some transformation of the value of a permit across area boundaries. Or,

markets could be defined according to the location of receptors. In each area of the region, a coefficient would be estimated that related the effect of a unit of emissions on ambient air quality at a receptor point. Sources could then be required to hold permits to pollute at a receptor point equal to their quantity of emissions multiplied by the corresponding coefficient.

3. Market Initialization. Regulators must select a method for initially distributing the permits. One possibility is to give them away according to some rule. Examples of allocation rules are: in proportion to precontrol emissions, in proportion to emissions allowed under existing standards, or equal to the expected equilibrium distribution of emissions if abatement costs were minimized. Alternatively, permits could be given to entities other than sources of pollution: the poor, schools, etc., presumably any of which would then elect to sell them. Or the government could allocate the permits by auctioning them. The latter two options suggest that sources of pollution would have to pay for permits; however, this is not necessarily the case for a state auction. Ownership of permits could be conferred on sources according to one of the rules for giving permits away, but sources could then be required to use an auction process to allocate the permits among themselves, with the revenues from the auction divided among the sources in proportion to their ownership shares.

4. Market Operation. Once an initial allocation has been made, provisions must also be adopted for later transactions. Government could leave the problem of organizing a continuing market

to the private sector. Alternatively, given the record-keeping requirements of the government for purposes of enforcement, the government could act as a marketplace by providing information about potential buyers and sellers to anyone requesting it. Or the government could be more than a passive marketing agent by actually requiring regular opportunities for reallocation of permits. This could be accomplished by forcing periodic reauctioning (with proceeds redistributed among the sources) of some fraction of the permits. A reauctioning process fits naturally with a system in which permits have fixed durations, for then the replacement of old permits by new ones can be accomplished through an auction of the same sort as used to accomplish the initial allocation.

Solving the Design Problem: A Case Study

The importance of the first four potential problems of a tradable permits system depends on the empirical features of the emissions problem that the system is designed to solve. The following discussion uses a particular example — the control of sulfate particulates in Los Angeles -- to illustrate how these issues can be assessed. This analysis is based upon relatively complete information about abatement costs, emissions inventories, and the relationship between emissions and air quality throughout the region. Los Angeles probably has the most sophisticated regulatory system for air pollution in the world, in part because local agencies have been collecting emissions and air quality information for two decades and in part because these data have been extensively used by research scholars to study the Los Angeles air pollution problem. This

information, of course, is especially helpful for illustrating the way that issues of designing a permits market might be resolved and for designing a particular set of market institutions for this pollutant in this region. It is not necessary, however, to have all of this information in order to move towards a tradable permits system. Following the discussion of the Los Angeles sulfate problem, attention will be turned to methods of approaching the same design problems when the available information is less reliable.

The problem of sulfate particulates in Los Angeles is somewhat unusual in that the state, not the federal government, is solely responsible for its regulation. Sulfate particulates are suspected of being a health hazard and having other damaging effects, but the principal justification for controlling them in Los Angeles is that they account for a very large part -- probably more than half -- of the reduced visibility due to air pollution in Los Angeles. There is no federal ambient air quality standard for sulfate particulates; however, the state has adopted a standard of 25 micrograms per cubic meter, averaged over a 24-hour period.

Although sulfates are released directly into the atmosphere by some sources, by far the most important cause of sulfates is the release and subsequent atmospheric oxidation of SO_2, nearly all of which is associated with petroleum products that contain sulfur as an impurity. There is a federal ambient air quality standard for SO_2; however, Los Angeles is not in violation of it. Hence, the state standard for sulfates is the binding constraint on SO_2 releases.

To control sulfate particulates in Los Angeles requires controlling emissions from about forty different categories of sources. The most important sources are electric utilities that burn oil to generate electricity, petroleum refiners, coke calciners, glass manufacturers, a steel mill, industries that are heavy fuel burners, and mobil sources burning gasoline. A tradable emissions permit system must be designed to account for emissions from these major sources.

The tools with which to undertake an analysis of the design of a permits market in Los Angeles are a detailed model of the relationship between emissions and air quality, and estimates of the abatement cost functions for all major sources in the region. The abatement cost functions provide estimates of the costs to each source of various degrees of abatement of its sulfur oxides emissions. A firm seeking to minimize the sum of its expenditures on permits and its abatement costs would elect to abate up to the point at which the marginal cost of abatement equaled the market price of a permit; therefore, the abatement cost functions provide a means for predicting the quantity of permits that each source would seek to hold at any given permit price. When all of the abatement cost functions are combined, the relationship between abatement and permit prices for the entire region can be estimated. Thus, given a limit on total emissions for the entire region -- e.g., the number of permits to be issued — the abatement cost data yield a prediction about the price of a permit, the distribution of remaining emissions in the air shed, and the expenditures on abatement (in total and by source).

The abatement cost information was gathered in the following manner.[13] First, public regulatory records and publications were searched to find cost estimates for various abatement methods for each source. Then, preliminary abatement cost functions were estimated and circulated among industry representatives and regulators for comments. The responses were then used to revise the cost estimates. For most sources, a few discrete abatement options were discovered, each with differing costs and levels of abatement. Thus, for most sources the abatement cost function is a step function.

The model relating emissions to air quality is based upon an analysis of detailed measurements of emissions and air quality.[14] Air quality estimates are made for each of the seventeen monitoring stations in the region, based upon meteorological conditions and the pattern of emissions among the sources in the area. The structure of the model is such that the geographical location of the sources and measuring stations is specified, so that the effects of changing geographical patterns of emissions can be estimated. Thus, the patterns of abatement and emissions predicted by the cost model under varying assumptions about the design of the permits market can be fed

---

13 For a more complete discussion and description of the cost data, see Robert W. Hahn, "Data Base and Programming Methodology for Marketable Permits Study," Open File Report 80-8, Environmental Quality Laboratory, California Institute of Technology, 1981.

14 For a detailed description of the model, see Glen R. Cass, "Methods for Sulfate Air Quality Management with Applications to Los Angeles," Doctoral Dissertation, California Institute of Technology, December 1977; and Glen R. Cass, "Sulfate Air Quality Control Strategy Design," Atmospheric Environment 15, no. 7, 1227.

into the air quality model to predict the results in terms of the concentration of sulfate particulates at each of the measuring stations.

One important result of the cost and air quality studies is that mobile sources — autos and trucks -- do not need to be dealt with directly in a permits system. Accurate air quality forecasts can be developed if mobile sources are redefined as fixed traffic sources along major arterial streets, using normal traffic densities and average auto emissions to calculate the emissions from these pseudo-fixed sources. Moreover, by far the least expensive method for reducing sulfur emissions from vehicles is to reduce the sulfur content of fuel. Consequently, it is feasible -- with little loss of efficiency — to allocate responsibility for mobile source emissions to distributors of refined products. Indeed, because Los Angeles refines more fuel than is consumed locally, responsibility for mobile sources can be pushed even further back in the production process to refiners. This is an important advantage, for each automobile emits a tiny amount of sulfur. The transactions cost of forcing vehicle owners to purchase emissions permits in very small denominations are probably roughly equal to the price of the permit, for the latter is unlikely to be more than a few dollars a year. Moreover, to be efficient, the auto permits would have to be related to use of fuel, which creates a very difficult enforcement task. Thus, allocating responsibility for mobile sources to distributors or refiners greatly improves the performance of the permits markets. This result is true for sulfur emissions in any region; however, it is not necessarily

true for controls on hydrocarbons and NO_x, the main components of photochemical smog. The reason is that in the latter case some cost-effective methods for reducing emissions are in control of the vehicle owner. Thus, to place responsibility on others -- for example, auto manufacturers -- would entail some loss of efficiency. This would have to be balanced against the greater transactions costs for including in the market literally millions of holders of small amounts of permits.

The next issue to be attacked is the possibility of an imperfectly competitive market structure. The first step in attacking this issue is to simulate the competitive allocation of permits. This is achieved by finding the minimum-cost allocation of abatement responsibilities among the sources that achieves a target level of total emissions. A permits market begins with some initial allocation rule. For each source, the difference between the initial allocation and the competitive allocation is the amount of permits it will buy or sell. By examining these differences, the structure of both the supply and demand sides of the market can be observed.

Although numerous market simulations have been made under varying assumptions about ambient air quality standards and the availability of substitutes for petroleum fuels, three examples will be presented here.[15] One assumes that the state's ambient air quality

---

15 For more details about the choice of emissions limits and the various simulations, see Robert W. Hahn and Roger G. Noll, "Designing a Market for Tradable Emissions Permits," California Institute of Technology, Social Science Working Paper No. 398, July 1981.

standards will be satisfied all of the time, the second assumes that the standard will be violated approximately two weeks per year, and the third assumes that the emissions allowed under regulations now in place become freely tradable. All cases assume that the availability of natural gas as a fuel for industrial boilers and electric utility generation will be as it was in the early 1970s, which means neither freely available (as it would be under total deregulation of energy) nor completely curtailed (as it was about to be in the late 1970s before natural gas price regulation was eased). Under this assumption, the controls on sulfur oxides emissions that were established in 1977 would produce emissions of about 300 tons of SO_2 equivalent per day in Los Angeles; to meet the standard all of the time requires that emissions be cut in half, but to meet it all but two weeks per year, on average, requires a further reduction of only about 50 tons per day. Thus, the three cases represent a major change, a minor reduction, and no change in currently enacted (but not yet fully in place) source-specific standards.

The single largest source of emissions is an electric utility. In 1973, prior to controls, this source accounted for approximately 28 percent of emissions in Los Angeles. Table 1 shows the share of permits that this firm would be expected to hold under two simulated market structures for the cases described above.

TABLE 1

FRACTION OF TOTAL EMISSIONS ACCOUNTED FOR BY LARGEST PERMIT HOLDER IN LOS ANGELES

	A. Competition (percent)	B. Monopsony (percent)
1. Make existing permits tradable with historical gas supplies	48	33
2. Violate standard two weeks/year with historical gas supplies	43	40
3. Satisfy standard all of the time with historical gas supplies	32	32

Source: Hahn, An Assessment of the Viability of Marketable Permits.

The shares reported in Table 1 should not be taken too literally. Two important sources of error could have an important effect on these estimates. First, treating each source as having a few discrete choices of abatement methods and hence facing an abatement cost function that is a step function, creates only a handful of potential emissions equilibria that are feasible for a particular firm. In reality, the abatement cost functions are likely to be smoother than the functions used in the model. Second, the shares of permit holdings for the largest firm are likely to be underestimates. Among the major source categories in Los Angeles, abatement costs are best known — and least likely to be overestimated -- for electric utilities. This means that even greater efficiency gains may be possible by substituting abatement elsewhere for the emissions reductions at utilities that are calculated from the existing cost data.

With these caveats in mind, the results in Table 1 illustrate the possibility of serious market imperfections, depending upon the selection of an emissions target and an initial allocation of the permits. Column A shows the cost-minimizing allocation of permits under the three emissions targets described above. This allocation is the competitive equilibrium. If the initial allocation process is an auction so that all firms are buyers, the share of the largest source is the share shown in Column A. Other initial allocations can raise this figure substantially. For example, suppose the allocation is designed to retain present emissions levels and is a proportion of precontrol emissions. In this case, the largest source, assuming the market were competitive, would seek to increase its share of holdings by 20 percent of the total number of permits (the difference between 48 percent on Line 1 Column A and the 28 percent share of baseline emissions). This would make this source an almost complete monopsonist, e.g., the only source of demand for permits at the competitive equilibrium price (almost all other firms would be sellers). The potential inefficiency of a monopsonist is that it will systematically understate its demand in order to force the price of permits down. This is achieved by engaging in excessive abatement, the extra costs of which are made up in the effects of pushing down permit prices.

Column B shows the results from the most extreme degree of monopsony that is possible for each of the three cases. Here it is

assumed that the largest source has an initial allocation of no permits, and that all other firms are given permits in a manner that causes them to seek to be sellers at any price equal to or above the monopsony equilibrium. The discreteness in the options available to the utility strongly influence these numbers: the actual emissions produced in Lines 1 and 2 of Column B are identical, and in Line 3 the monopsony and competitive equilibria are the same. Such extreme results should not be expected to emerge in the real world. Nevertheless, the pattern of the results — a greater divergence between competitive and monopsony shares for higher total limits on emissions — is likely to be robust for this particular case. The reason is that in the range of the competitive equilibrium for emissions limits around the most stringent standard, the supply of permits from other firms to the largest source is very elastic, even with discrete options in the abatement cost analysis. This undermines the opportunity of the monopsonist to take advantage of its high market share: overabatement will not force much of a drop in permit prices, and hence the gains from the latter will not generate much of an offset against the higher abatement costs that are necessary to allow the firm to reduce its demand for permits.

The tentative conclusion from this analysis is that for the particular case at hand, monopsony appears to be a serious design concern only if regulators do not conform to the existing ambient air quality standards. The actual allocation rule is certain to be less likely to cause monopsony than the extreme case analyzed here, yet even under this extreme assumption imperfections in the permit market appear relatively unimportant if the emissions limit is low. On the other hand, market imperfections could be important if existing permits were simply made tradable unless the initial allocation were designed to guard against it.

Multiple Permits Markets

Another important implementation issue is the degree of geographical resolution in the definition of an emissions permit. In the case of sulfates in Los Angeles, a permit could be defined as a license to emit a given amount of sulfur oxides anywhere in the air shed. Or, permits could have a varying value depending on the geographic location of the source holding them, with the relationship based upon estimates of the damage created by emissions from different sources. Finally, the permits could be defined in terms of the resulting pollution at each of the seventeen sites at which air quality has been measured, with an air quality model being used to calculate the number of permits a given source must hold at each receptor site for each unit of emissions that it releases. The last alternative is the most complicated to implement, for it requires that sources participate in seventeen markets and that the state be continuously available to run air quality simulations whenever any firm seeks to change its emissions. While such a system is difficult -- perhaps impossible -- to operate in the real world, its results can nevertheless be simulated. The differences between the first system (all emissions are treated equally) and the last system (each receptor is associated with a separate market) thus provides a measure of the potential gross gains from a fine-tuned method of defining the permits.

Table 2 presents some of the results of these simulations.[16]

TABLE 2

COMPARISON OF UNIVERSAL AND RECEPTOR-SPECIFIC PERMITS
(costs in $ millions)

Baseline Emissions Target in Tons/Day SO_2 Equiv.	Annualized Costs of Competitive Equilibrium Abatement for: A. Universal Permits	Receptor-Specific Permits that Produce: B. Same Air Quality For Each Receptor	C. Uniform Air Quality Equal to Worst Receptor
150	682	682	682
250	565	557	545
300	515	513	505

Source: Hahn, "An Assessment of the Viability of Marketable Permits."

The case analyzed here is one in which natural gas availablity is low. This case is likely to produce the greatest differences in abatement costs among various methods for organizing the permits market. Given the historically available gas supply to Los Angeles, abatement costs tend to be about 60 percent of the costs if natural gas supplies are low. Column A shows the annualized expenditures on

[16] For a more complete analysis, see Hahn and Noll, "Designing a Market for Tradable Emissions Permits."

abatement costs in the Los Angeles area under the competitive equilibrium distribution of permits if there is no geographical resolution in the permit system. Column B shows the costs if firms are required to buy pollution permits for each of the seventeen measuring stations in the air shed, subject to the condition that the air quality results at each station will be the same as the outcome from the system reported in Column A. Thus, the difference between A and B is the gain, if any, arising solely from geographical relocation of permits in a system that takes account of the specific polluting effects of emissions from each location in the region.

Column C further relaxes the system, allowing pollution at all measuring stations to be constrained only by the air quality achieved at the most polluted station under the allocation corresponding to Column A. Thus, emissions can be reallocated and total emissions increased as long as pollution does not increase at the location that is most polluted under the Column A allocation. Again, the results are affected by the discreteness in abatement options assumed in the model; however, the general result from the analysis is that there is little to be gained from fine-tuning the definition of permits. The reasons are twofold: the simple market allocates emissions relatively evenly over the region, and leaves relatively little differences among measuring stations in terms of the air quality results. Hence, there is little opportunity for improving the efficiency of the allocation through adopting a more complicated market system.

While these results may not be generalizable to other pollutants in other areas, they nevertheless suggest an important

lesson. In order for a tradable permits system to be able to capture significant gains from a fine-tuned, complicated definition of permits and their markets, it is <u>not</u> sufficient that different sources have a significantly different pattern of effects. It must also be the case that the equilibrium allocation of permits in a crude, simplified market be such that emissions would be concentrated in a particular location.

If the comparison is to be made between a simple tradable permits system and the existing regulatory arrangement, the issue is whether a tradable permits system increases or reduces the geographic concentration of emissions compared to the present pattern. Because the present source-specific standards system tends to force some activities to overabate (for example, electric utilities), while leaving other sources virtually unregulated, it should not be surprising to find that a tradable permits system evens out the pattern of emissions in comparison to the present system, and therefore makes more even the geographical distribution of measured pollution within a region. Moreover, if most sources face relatively low abatement costs for a substantial fraction of their emissions, a market system will lead to substantial abatement from all sources. In such a case, it should not be surprising to find situations in which there is relatively little additional gain from moving from a simple to a complex system.

<u>Thin Markets</u>

The final major potential source of a failure in the permits market is that transactions will be too infrequent to convey meaningful price signals to polluting firms, to make relatively easy the acquisition of permits for entry and expansion of polluting facilities, and to allow a firm to avoid the expense of organizing the market and engaging in extensive bilateral negotiation every time it desires to make a trade. This is an especially difficult design problem to get a firm grip on in advance of operating the market, for the indicators of the extent of market transactions are so crude. One measure is the number of firms accounting for existing and expected emissions. In Los Angeles ten companies account for approximately 85 percent of the sulfur oxides emissions under current standards, assuming mobile sources are assigned to the oil refiners operating in the air shed.[17] Most major industrial polluters emit relatively small amounts of sulfur, so that the market for small quantities of permits is likely to be reasonably well-functioning; however, a major expansion or entry of an oil refinery, an offshore oil terminal, or an electric utility generation facility would be especially difficult to accommodate because so few sources have sufficient numbers of permits to be potentially significant sellers to the new source.

A second problem in anticipating the extent of a problem of market thinness is that there is likely to be a systematic tendency to underestimate the possibilities for transactions. A substantial source of demand and supply in the market for permits will be factors that are not measurable in advance. Examples are innovations in abatement technology, entry, exit, contraction and expansion of

[17] See Cass, "Methods for Sulfate Air Quality Management with Applications to Los Angeles."

polluting entities, and opportunities for more efficient abatement methods that may be known to existing sources but that have not yet appeared in the public domain (e.g., process changes).

In Los Angeles, the problem is even more difficult because the local air pollution control authority has explicitly adopted the policy of attempting to write standards in inverse order of their costs per unit abatement. Thus, with few exceptions, the standards in place are the least expensive abatement methods available, and the pending standards are the least expensive remaining possibilities. Consequently, most of the demand for trades, and the gains from a permits market, are unlikely to be measured using existing cost information. Therefore the extent to which the thinness of the market is a potential problem is likely to be overstated. For example, given historical natural gas supplies and no abatement methods other than those whose costs are known publicly, the annual cost savings from a competitive reallocation of the emissions permits that are currently in place is estimated to save only about $20 million per year in abatement costs, which is about 7 percent of the total.[18] This indicates a strong possibility of a thin market, indeed; however, it is sure to be an underestimate of the potential savings, and hence the desire to trade.

Whether the market is thin, initially and in the future, again depends on the design of the system. A few examples illustrate this point. (1) If existing emissions (or some proportion of them) are

[18] Calculations in Hahn, "An Assessment of the Viability of Marketable Permits."

simply made tradable, a thin market is a more likely prospect than if an auction process is used for the initial allocation. (2) Fine-tuned, multiple-market systems are more likely to face a problem of thinness than single markets defined over a broad geographic area. (3) If permits are perpetual with no periodic reallocation process, a decision to make a major purchase or sale would then require that the firm wishing to make a market undertake the time and expense of organizing and negotiating a trade. At the other extreme, if permits have a fixed life and are reallocated by auction, a convenient time and place is established for facilitating major redistributions of permits should changes in underlying economic and technological conditions warrant it.

Selecting a Design: Conclusions

The preceding discussion should make clear that selecting a design for a system of marketable emissions permits is not a purely technical, scientific matter. Working out the details requires considerable judgment, including an assessment of which risks a political entity would prefer to run. The following observations present an analysis of two polar cases: a design for sulfur oxides emissions in Los Angeles, an area in which information is comparatively rich, and a design for the general problem of emissions in control in an information-poor environment.

In Los Angeles, market imperfections apparently are a far more important design issue than is the selection of appropriate definitions for the geographical extent of a permits market. Fine-tuning the system in terms of multiple markets for sulfur emissions

promises little gains, yet there will be formidable problems of transactions costs and market structure. Consequently, a system in which permits are simply stated in terms of allowable quantities of SO_2 equivalent emissions anywhere in the region appears to be the most desirable.

The Los Angeles air shed has a relatively large number of sources producing a small quantity of emissions, but only a dozen or so firms account for about 85 percent of the total. Consequently, imperfectly competitive and thin markets are a potential problem for large transactions. The implication is that the method selected for initializing and maintaining the system should encourage an active, competitive market.

The most attractive method for the initial allocation is an auction mechanism. This provides a thick market (all permits are transacted in the initial distribution) and, because all polluters are placed on the same side of the market, it minimizes the likelihood of monopolistic imperfections. The mechanics of the proposed mechanism are as follows. Each source would be asked to write down the number of permits it would seek to purchase at each of several prices. The firm would be free to choose as many price gradations as it wanted. It could write down one price-quantity pair (e.g., X tons per day at any price up to $Y per ton). It could provide a step function of several jumps, such as X tons per day for prices between $Y and $Z ($Y larger), and X+W tons for prices below $Z. Or, it could bid a continuous demand function for permits. Permits would then be allocated to the highest bidders at the quantities requested, descending down the price bids until the permits were completely allocated. Among the bids not receiving any permit allocations, the one with the highest bid price determines the price of all permits (e.g., the allocation mechanism is a second-price auction). This process is the most likely theoretically to produce a competitive allocation of the permits.

A separable equity issue accompanying the auction is the allocation of the net costs of the permits. Whereas the permit price determined above could actually be paid to the state, an alternative is to pay the revenues according to a previously arranged provisional permit allocation, as described above. As a political matter, the chances of implementing a tradable permits system are probably greater if the revenues do not accrue to the governmental treasury, and if the provisional allocation of the rights to receive the revenues from the permit auction does not reward firms who have been most resistent to current environmental regulations. For example, if ownership rights for purposes of allocating the auction revenues were based upon emissions under current standards, firms that had succeeded in fighting regulation or that had managed to induce regulators to impose relatively undemanding standards on them would receive relatively large quantities of rights in emissions in comparison with firms that had been more cooperative and engaged in more costly abatement techniques. From an equity standpoint, provisionally allocating the permits prior to the auction on the basis of either precontrol emissions or the expected competitive allocation are both superior. The latter approach — estimating the competitive allocation and

granting ownership in permits on the basis of it — is more difficult, for it requires establishing in a formal regulatory process what the cost-minimizing allocation would be. This is comparable to setting source-specific emissions standards, and could be upset by the procedural requirements of the regulatory process. If any single firm objected to the allocation, the entire system could be delayed by continuing litigation until its appeals were resolved.

The primary difficulty with a system based on precontrol emissions is that for one important source — oil refineries -- the methods that they would adopt today with no regulation would produce far less emissions than the best methods of 1970. A method based on precontrol emissions would, in any event, end up giving most of the permits to oil companies, who would then probably sell them in large quantities and high profit to public and municipal utilities. This does not seem to be very acceptable politically.

Obviously, the equity aspects of initialization are thorny. Perhaps the best alternative is to base allocations on an amended list of existing emissions, with the few remaining uncontrolled sources being put through an emissions standard process before the initial allocation of tradable permits is made. Whatever the choice, the allocation would proceed as follows. The ratio of target emissions to the baseline emissions would be calculated. Each firm would then receive a provisional allocation of permits equal to this ratio times its baseline emissions. New or expanded sources would receive allocations based upon actual emissions when operations began.

After the provisional allocation is made, each firm would make bids on permits, and thereby receive a final allocation of emission permits according to its bids at the price of the highest excluded bid. The firm would pay for these permits at the established price, and receive revenues at the same price for the permits which it held provisionally. The net payment for a particular source would be the product of the auction price and the difference between its final allocation based upon the bidding procedure and its provisional allocation based upon precontrol emissions. For all firms taken together, the net payment would be zero.

To provide a continuing opportunity for entry and expansion, permits could be separated into vintages according to useful life. The permits would be declared the binding control on emissions beginning at a specified date after the auction -- perhaps a year or two after it takes place to give sources ample time to engage in capital investments to accommodate their permit holdings. All permits could then be valid for another fixed period after the system is in place, such as for an additional year. Then, a predesignated portion of the permits could expire each year -- for example, 10 percent of the permits to expire each year. Prior to the expiration date, the regulators would determine how many permits would be issued to replace the expired ones, based upon considerations of cost and air quality. The new permits could be allocated by the same auction procedure as was used for the initial allocation. Provisional allocations for purposes of distributing auction revenues would be based upon holdings of the expiring permits, but the final allocation would be based upon

a second-price auction.

Between the formal auctions, government regulators could maintain a public file of the current holdings of permits of various vintages, and could serve as a clearinghouse for information about firms that wish to buy or sell their holdings. More risk-averse firms (or firms wanting a long-term emissions commitment) could seek to sell permits with short remaining lives and buy permits with longer lives. Entry and expansion of polluting sources could still take place through the clearinghouse; however, the presence of an annual auction would probably end up being the primary mechanism for a new major source to acquire the necessary permits.

The preceding mechanism appears to cope best with the specific design problems for the case at hand. A key element of the Los Angeles problem, however, is the finding that the geographic allocation of emissions does not make much difference in terms of satisfying air quality objectives. In the absence of information about the relationship between emissions and air quality, or in a world in which localized effects are understood and known to be potentially important, how might the design of the system differ?

The mechanism described above has several features that are well-suited to the case of poor information. First, the periodic auction is the best way to protect against market imperfections owing to either market concentration or market thinness. Second, the concept of having some of the permits expire each year is especially appropriate when the relationship between emissions and air quality is poorly understood, because the process by which expiring permits are converted into new ones allows the regulator continually to adjust the number or permits (and hence air quality). Of course, the more uncertain are regulators about the emissions to air quality relationship, the greater is the degree of variability in the number of permits that they would desire. One method for achieving greater potential for variability is to have two types of permits: long-term (perhaps ten years, as above, or even perpetual) and short-term (as short as one year). Whereas under the system described above only 10 percent of the permits expire each year, with a two-permit system a much higher proportion could be assigned to the short-term category and therefore varied in quantity from year to year. A second use of short-term prmits would be to facilitate an economically efficient approach to achieving the ultimate air quality objective. Regulators could announce a strategy to reissue short-term permits at some ratio of new to expiring permits that is less than one, thereby gradually winding down the total emissions in the area. The process could be based upon an emissions target that is established before the system begins to operate (assuming the relationship between emissions and air quality is known well enough to make this feasible), or the winding down process could be open-ended, with regulators announcing a fixed percentage reduction in short-term permits until ambient air quality standards are achieved (or changed).

The preceding arrangements still leave unsolved one potential failing of a market in permits: the possibility of localized effects from a single or a few sources that elect to buy a large number of permits rather than to abate. Because this result in the context of a

permits market is a consequence of a cost-minimizing process of reducing emissions, the appropriate response to the problem may be to allow some degree of localized violations of air quality standards. Nevertheless, in the context of existing air pollution regulation, rather than regulation seeking to make some sort of optimal trade-off between benefits and costs, the ambient air quality standard is an inflexible policy objective.

If enough firms contribute to a localized effect, one possible solution is a system of multiple permits markets. This would require a formal regulatory determination of the coefficients relating emissions to air quality for the relevant sources and localized pollution hot spots. As long as the number of these localized problems is relatively small, allowing firms to participate in only a few markets, this approach may prove workable. But it would require a substantial evidentiary burden on the regulators in defining the permits, the markets, and the mathematical relationship of sources to each. In the absence of good information about the relationship between emissions and air quality, regulators may not be able to sustain such findings legally, or may be able to do so only after a long legal battle.

A second, probably more fruitful approach is to set up a permits market, but to overlay minimum standards on the sources that are suspected of having important localized pollution effects. Only very large sources or sources emitting at or near ground level can be expected to cause a violation of ambient air quality standards all by themselves; hence regulators could deal on a case-by-case basis with these sources. The idea would be to let the permits market allocate emissions, but to set an upper bound on the number of permits that could be held by some specific sources or a lower bound on stack heights. These standards would be set according to the same procedural and evidentiary requirements that apply to the present source-specific regulatory standards; however, the process of implementing them would not need to delay the implementation of a general permits markt. The long legal process to set an upper bound on a particular source could be underway while the market operated, and could be directed only at the major sources that in equilibrium were observed to hold a greater number of permits than the proposed upper bound.

Any of the preceding systems -- numerous short-term permits, some standards overlaying the market for permits -- creates uncertainty among emissions sources in picking an optimal abatement strategy, or indeed in being willing to experiment with a change in the regulatory system for controlling emissions. Uncertainty about the future state of regulatory stringency affects the selection of abatement methods by a regulated entity -- both the total amount abated, and the choice among capital-intensive technologies versus changes in operating procedures. In general, uncertainty should make firms somewhat more reluctant to abate at all, and somewhat more likely to adopt more flexible abatement methods that allow relatively inexpensive adjustments in emissions as policy changes occur.

One good argument for a stable regulatory policy is to create a more certain decisionmaking environment for regulated firms. But if

long-term goals cannot be very accurately stated, either because the benefits or improved air quality are uncertain or because the relationship between emissions and air quality is poorly understood, regulatory policy will produce a more efficient result if this uncertainty is transmitted to businesses by the regulatory system. A system that does not specify the long-run emissions goal, but that makes some measured change in emissions over time (and that is ambiguous about the ultimate stopping point), will be more efficient if it encourages more flexible abatement. methods. Consequently, although the politics ot the situation may prove intractable, a design criteria for situations in which the amount of emissions to be allowed in an area is unknown ought to be to construct a permits system that conveys this uncertainty to polluting entities.

III. DIRECTIONS OF CHANGE IN U.S. POLICY

In the late 1970s, the Environmental Protection Agency began seriously to consider -- and then to encourage -- a policy of "controlled trading methods" in air pollution regulation. The idea was to introduce some elements of a market -- and its attendant flexibility -- into the standards-setting mode of regulation. Three such methods have been developed and, to a limited extent, implemented: "bubbles," "offsets," and "banks."

As discussed above, current regulatory practice is to set a standard for each source, with a source defined as a point from which emissions are released. For large, complex manufacturing processes, such as oil refining or steel making, a single plant can have several separate sources. The bubble policy is an attempt to introduce some relaxation of the exclusive focus on specific sources. It enables a firm to reduce emissions below the standard from one source and increase emissions from another at the same location if the effect is a net improvement in environmental quality. The term "bubble" was adopted to evoke the notion of placing a bubble around an entire plant and treating it as a single source, rather than as a series of independent sources.

Bubbles are intended to deal with cost-minimizing reallocations of emissions among sources at the same facility. A similar concept for trade-offs between two facilities underpins the second policy, called offsets. An offset is a reduction in emissions below the standard in one facility that more than compensates for an increase in emissions at another in the same general area. An offset is more like a normal market relationship, for the mechanism by which one firm induces another to reduce its emissions is to pay compensation. Because the offset policy is oriented towards trade-offs between two firms, the details of the transaction are expected to be the result of bilteral negotiations, rather than a transaction in a continuously operating centralized market.

Still closer to a normal market is an emissions bank. The bank policy provides a mechanism whereby a source can receive credit for an excess reduction in emissions without actually finding someone to whom it can sell the emissions reduction. If a firm beats its emissions standards, it receives a credit in the emissions bank. Normally the amount of the credit is some fraction (near to but less

than one) of the reduction in emissions below the standard, so that the process effectuates a net reduction in emissions. Other firms, seeking to receive emissions permits, then can purchase the banked credits by negotiating an agreement with one or more depositors.

All of these new policies are conceptually like tradable emissions permits in that they allow economic considerations to guide some reallocation of emissions permits. But all are part of the standard-setting process, rather than a substitute for it. Each trade must be approved by regulatory authorities in the normal permitting process. This normally requires information about the abatement technology to be ued at both the point where emissions are reduced and at the source increasing emissions. Moreover, except for the bubble policy, these methods take existing emissions standards as a baseline: neither trading partner in an offset or bank exchange can end up emitting more than the existing source-specific emissions the standard allows. Offsets and banks are seen as means for allowing entry and expansion of pollution sources when the source would cause a reduction in air quality even if it was in compliance with source-specific standards. Thus, new or expanding sources can use offsets and banks to purchase emissions permits for the emissions they would release after having adopted the best abatement technology available to them. These policies do not relax uneconomic differences among old source categories, or between new and old sources, that are present in the existing system of source-specific standards.

Offsets and banks also suffer from the absence of a formal method of organizing the market. Each depends on bilateral negotiations, and each can expect infrequent transactions because only new or expanded major sources are candidates to engage in trades. Consequently, the transactions costs for trades can be expected to be high, and problems of market imperfection are likely to emerge.

Finally, the status of traded emissions permits is secondary to permits inherent in existing source-specific standards. The EPA guidelines for the controlled trading methods include provisions for situations in which ambient air quality standards are not satisfied after all sources are controlled. In general, if environmental quality objectives are not attained, all emissions permits are to be regarded as provisional and subject to change. But traded emissions permits are to be the first to be examined for redefinition. For example, the guidelines governing the emissions banks — the most market-like of the three options -- state that local air pollution control agencies are to select among four alternatives if air quality objectives are not realized: (1) a moratorium on the use of permits obtained through the bank; (2) a revision on a source-by-source basis of the number of permits from the bank that are necessary to allow a given quantity of emissions (e.g., a depreciation of the value of a permit); (3) an across-the-board depreciation of the value of all traded permits; or (4) a forfeiture of all traded permits.[19] Thus traded permits are required to bear the brunt of the uncertainties in environmental policy. Obviously, polluting entities can be expected

---

19 Emission Reduction Banking Manual, Emission Reduction Banking and Trading Publication BG200, U.S. Environmental Protection Agency, September 1980.

to be reluctant to risk substantial capital investments on the validity of traded emissions permits.

The new trading methods offered by EPA should properly be regarded as a supplement to a continued primary reliance on source-specific standards. For the reasons given above, they are not appropriately characterized as a system of marketable emissions permits. Instead, they are a mechanism whereby sources can collaborate to obtain some changes in source-specific emissions standards from local regulators, subject to some important conditions about the kinds of changes that are feasible and the long-term status of the revised standards. They do not abandon the premise of the existing regulatory structure, which is that in order to control pollution one must directly control every polluter.

The argument for these changes in regulatory policy -- assuming that the problems with the existing system are major — must be primarily political, based upon the idea that controlled trading options are a necessary step in the transition from the old system to a new, market-oriented one. In this light, the controlled trading options have one major advantage over the designs discussed in the previous section: not everyone has to participate in the system and face the uncertainties inherent in it. Local regulators can continue to set standards for every source, even when trades take place. A polluting entity can choose to satisfy its current source-specific standards, avoid subjecting itself to a market process, and presumably live happily ever after in a blissful but expensive world of relative regulatory certainty. Assuming that regulators can avoid a policy catastrophe, such as observing continued failure to satisfy ambient air quality standards and then voiding all traded permits, the presence of a functioning, and profitable, permits trading system can induce more firms to enter offset or bank arrangements. And, as experience is gained, environmentalists and local regulators can develop more faith in decentralized decisionmaking, and exercise less and less scrutiny over the details of each transaction. Moreover, areas that do relax the source-by-source review process and essentially permit any reasonable trade will gradually be rewarded by the flexibility the system allows for economic growth, perhaps forcing more reluctant regulatory authorities to go along.

The major problem with this line of reasoning is that the idea of marketable permits may sink because the new trading methods are so procedurally freighted, so limited in applicability, and so burdened with uncertainties over the long-term value of the permits. Too timid a reform leads to few transactions and market imperfections that undermine the efficiency of the trades that take place. Even in the absence ot a policy catastrophe, the system could prove so cumbersome that it is uninteresting to polluting entities, and hence does not lead to the gradual transition to a true market system.

In any event, at some point it is likely to be desirable to make a major alteration in policy, regardless of which transition path to a market system is followed. That is the regularization of a market situation -- such as an auction -- that avoids source-by-source review and that facilitates trades by providing a thick market with low transactions costs. This is the key to substantially improving

the efficiency of the environmental regulatory process in the United States. Moreover, the method for implementing this change is the principal design issue that has thus far been ignored by policymakers and that must be addressed if an economically rational regulatory method is to be put in place.