Assessing Point Source Discharge Permit Trading: Case Study in Controlling Selenium Discharges to the San Francisco Bay Estuary*

Alexandra Teitz**

CONTENTS

Introduction .................................................... 81
I. Lessons From Theory and Practice .......................... 85
   A. Implemented Permit Trading Programs ................. 85
   B. Necessary or Beneficial Elements for Trading Program Success ........................................ 87
   C. Issues Specific to the Water Pollution Setting ...... 93
      1. Physical Context of Water Pollution ............... 93
      2. Legal Requirements of the Clean Water Act ....... 94
   D. Types of Transferable Discharge Permit Systems ... 96
      1. Allowance System ..................................... 96
      2. Group Permit System .................................. 96
      3. Tradable Limit System ................................ 97
II. Evaluating Discharge Permit Trading in a Specific Problem Context ........................................ 98
   A. Examination of Factors Relevant to Permit Trading ............................................................... 98
      1. Physical Context—Selenium Discharges to the San Francisco Bay Estuary .......................... 99

Copyright © 1994 by Ecology Law Quarterly
* An earlier version of this comment won the 1994 Thelen Marrin Prize for the best comment at Boalt Hall and the 1994 Ellis J. Harmon Environmental Law Writing Competition.
** Candidate for J.D. 1994, School of Law (Boalt Hall), University of California at Berkeley; M.E.S. 1991, School of Forestry and Environmental Studies, Yale University; B.A. 1988, Oberlin College. The author gratefully thanks Professor John Dwyer for his insightful suggestions, Dr. William S. Pease for his generous sharing of research materials, ideas, and critiques, and those interviewed for their time and thoughts. The author also thanks Anthony J. Gordon, Thomas Brooks, and John Palmisano for introducing her to permit trading.
2. Regulatory Context—Current System for Regulating Selenium Discharges to the San Francisco Bay Estuary ........................................... 101
3. Legal Context—State and Federal Laws and Regulations Affecting Permit Trading .......... 104
   b. Discharge Permit Trading .................... 107
      i. Requirement of Effluent Limitations in Permits ........................................ 108
      ii. Restrictions on Changing Permit Limitations ........................................... 110
4. Economic Context—Controls, Costs, and Market Conditions ........................................ 113
5. Institutional Context—Capacities for Developing, Implementing, and Carrying Out a Trading Program ....................................................... 119
6. Political Context—Alignments and Motivations ...................................................... 122
B. Design Options: Examination of Factors Relevant to Permit Trading Approaches ............. 129
   1. Design Options Common to All Permit Trading Systems ........................................ 129
      a. Structural Options .............................. 130
      b. Options in Allocating Discharge Rights ... 133
      c. Options in Market Form and Operation ... 136
      d. Monitoring and Enforcement Options ...... 139
   2. Design Options Differing Across Types of Trading Systems .................................. 142
      a. Allowance System ............................... 142
      b. Group Permit System ....................... 147
      c. Tradable Limits System ..................... 150
III. Implications for Designing a Discharge Permit Trading Program ........................................ 152
   A. Relative Strengths and Weaknesses of Different Trading Systems ......................... 152
      1. Legal Feasibility ............................... 152
      2. Economic Efficiency ............................ 153
      3. Regulatory Administrability and Environmental Effectiveness ............................ 155
      4. Political Acceptability ........................ 156
   B. Matching Trading System Design to a Problem Context ........................................ 157
      1. Likely Constraints on Trading in the San Francisco Bay Estuary .......................... 158
INTRODUCTION

Marketable permit systems for pollution control provide a means to meet environmental goals at lower cost than traditional approaches based on uniform standards. Lowered compliance costs may, in turn, reduce resistance to environmental regulation and allow greater environmental gains to be made. Most permit trading programs to date have been developed to reduce air pollution. Tradable permits have not been used much, but do have potential, in water pollution control. Point source discharges of conventional pollutants have been regulated fairly extensively since the passage of the Federal Water Pollution Control Act (Clean Water Act or CWA) in 1972.1 Regulators have recently begun shifting their focus to such water pollution issues as nonpoint source pollution, water quality and quantity relationships, toxic pollutants, and cost-effective regulation. This comment addresses the last two issues in examining whether, and under what circumstances, marketable permits could be a cost-effective means of controlling point source discharges of persistent contaminants.

While recent attention has focused more on marketable permits for point-nonpoint source water pollutant trading,2 this comment specifically examines point-point source trading for several reasons. It should be easier to establish a viable trading system for point sources, where all the sources are identified permitted entities and trading transaction costs may be lower. There is also less uncertainty about environmental outcomes for more easily monitored and enforced point source reductions. Finally, experience in the point source arena may help develop trading mechanisms, expertise, and confidence for extension to the point-nonpoint source arena.

Marketable permits can be used where the adverse environmental impact relates to the total quantity of pollutant loading within a given area and where that impact is largely independent of the source of the loading. The theory underlying market-based systems is that reductions should be made by the least-cost source. The theory assumes that individual facilities rather than government regulators have the best information regarding control costs and capabilities. Thus, environmental protection and cost-effective controls are most likely to be achieved through a system where the government deter-

mines the desired level of pollution loading to the environment, but allows individual firms to choose how and where to make reductions.

A marketable permit system is implemented by the government issuing pollution permits that allow net discharges equal to the regulatory level of allowable pollution. Individual firms can then trade (i.e., buy and sell) these permits. Where control costs differ between firms, facilities facing relatively higher control costs should be willing to buy permits from facilities that can reduce discharges for less than the offered price. Reductions are made where they are least expensive, while the total impact on the environment is kept to the predetermined level. Regulators do not try to determine the optimal technological control approaches, since this determination is left to the firms. Firms have an incentive to develop improved control technology, because they can realize savings by selling any excess reductions.\(^3\)

Regardless of the theoretical appeal of tradable permits, whether they work in practice to achieve cost-effective environmental benefits depends on how they are implemented in a given setting. Tradable permits will be cost effective only where firms face different marginal control costs, and where a market in permits develops. Transaction costs such as information costs and regulatory requirements will impose trading barriers and reduce realizable cost savings. Also, the environmental effectiveness of tradable permits depends on accurate permit tracking and emissions monitoring.\(^4\) For a marketable permit system to be worthwhile as a matter of policy, the costs of developing, implementing, and administering the system must be outweighed by actual savings in control costs, while the level and certainty of environmental protection must be maintained. Some of these factors relate to the setting of a given environmental problem, exogenous to the specific program contemplated, while others directly depend upon trading program design and implementation choices.

Whether or not permit trading is the optimal regulatory approach cannot be determined at a purely theoretical level, because the factors on which success will turn vary tremendously in practice. The theory of permit trading has been explored in an extensive literature.\(^5\) Crude

---


\(^4\) While a uniform standards approach will also require some kind of monitoring, less precision may be necessary with simple technology-based requirements.

modelling also has been done to estimate possible cost savings.\(^6\) To assess permit trading further it is now necessary to move from general theory to the specifics,\(^7\) since the details of problem context and program design ultimately will determine the success of using permit trading to address a particular environmental problem. The specifics, however, are difficult to weigh effectively in the abstract. Considering the options in relation to an illustrative environmental problem makes it easier to identify some of the implications of different design choices for a trading program.

This comment considers the potential for controlling water pollution through tradable discharge permits in a plausible context—selenium discharges from oil refineries into the San Francisco Bay Estuary.\(^8\) Selenium discharged by oil refineries is bioaccumulating in the San Francisco Bay Estuary foodchain to potentially toxic levels, threatening aquatic and avian wildlife populations.\(^9\) The detrimental ecological impacts of selenium correspond more closely to the mass loading of selenium in the system than to selenium concentration levels in the water column.\(^10\) Recognizing this, the San Francisco Bay Regional Water Quality Control Board (regional board or SFBRWQCB) staff has proposed regulations that focus on reducing mass emissions of selenium.\(^11\) The key issue under this proposed approach would be the total selenium output of all sources over a given time period, not the effluent selenium concentrations of individual sources.

Permit trading may be a desirable policy approach to the selenium problem for several reasons. First, the environmental context is appropriate for emissions trading. The element of concern is mass emissions, not individual source emissions levels, so trading emissions among sources within an appropriate, spatially defined, trading zone should have no detrimental environmental effects. Since the ambient concentration levels are very low, trading is unlikely to pose a toxic hot spot problem, where concentrating emissions in one area would

\(^6\) See, e.g., Scott Atkinson & Tom Tietenberg, Market Failure in Incentive-Based Regulation: The Cost of Emissions Trading, 21 J. ENVT. ECON. & MGMT. 17 (1991) (exploring the "trading process hypothesis").

\(^7\) While this has been done in a number of areas, such as global warming, there is little information available in the context of point source discharge trading.


\(^9\) See id at 1, 37. Selenium also is regulated as a human health risk due to concentration in duck tissues. Id. at 1, 8-11.

\(^10\) Id. at 6.

\(^11\) Id. at 6-8.
cause acute or chronic toxic effects. Second, permit trading also may respond well to the economic context of the problem. The refineries currently have widely varying concentrations and levels of selenium emissions, and they face different marginal costs of reductions. Thus, there may be substantial cost savings in allowing reallocation of uniform emission requirements. Moreover, the expected control costs are very high. This creates substantial resistance to selenium regulation and enhances the political value of a more cost-effective approach.

Examining whether permit trading is feasible in the San Francisco Bay Estuary involves two major inquiries. First, what is the best way to design a permit trading program for this setting? Second, what are the major constraints both on design options and on the general use of trading to address water pollution? The San Francisco Bay selenium problem is analogous to other point source discharges of toxic or persistent conservative pollutants, and it raises many of the issues likely to affect point source permit trading in other contexts. Thus, this analysis allows some broader lessons to be drawn for point source trading in general.

Part I of this comment examines what we have learned about marketable permits from our experiences in applying them to air pollution problems, considers how the water context may be different, and categorizes types of trading approaches. Part II explores how a tradable discharge permit system might work in practice, using the San Francisco Bay Estuary's selenium problem as a case study. This portion examines certain aspects of the problem context that affect the feasibility of using trading generally as a regulatory approach, and asks specifically how different types of trading programs might function here. This analysis also serves as a model for evaluating the possibility of applying permit trading in other settings. Part III then draws upon the case study to make some further generally applicable observations about designing a discharge permit trading system. This part notes the strengths and weaknesses of different types of trading systems, as well as the most likely constraints on using permit trading as revealed by the case study. It then suggests how choices in designing a permit trading program can be made to mitigate these likely constraints. The comment concludes with some recommendations for future actions to facilitate use of marketable permits to control point sources of water pollution.

12. Heavy metals are another example of a persistent pollutant. Interview with Dr. William S. Pease, Environmental Health Policy Program, School of Public Health, University of California at Berkeley, in Berkeley, Cal. (Apr. 21, 1993) (on file with author).
LESSONS FROM THEORY AND PRACTICE

The United States has now had almost twenty years of experience with a variety of permit trading approaches.\textsuperscript{13} Through empirical studies of different applications and theoretical analyses, a consensus has arisen about what may be necessary, or at least beneficial, for a successful permit trading program. This discussion provides an overview of the aspects of an environmental problem context that make it appropriate for permit trading and explains the options of program design that will facilitate an effective permit trading program. Since the vast majority of our experience has been with air pollution, this part next highlights issues specific to water pollution. Finally, this part delineates three broad categories of trading approaches: an allowance system, a group permit system, and a tradable limit system.

A. Implemented Permit Trading Programs

Only the trading systems for pollutants regulated under the national ambient air quality standards and the lead trading program are sufficiently well established to serve as examples of fully functioning permit trading programs. The bubble,\textsuperscript{14} netting,\textsuperscript{15} offset,\textsuperscript{16} and banking\textsuperscript{17} programs of the U.S. Environmental Protection Agency (EPA) were the first marketable permit systems. Bubbles and netting have yielded substantial cost savings, with neutral and slightly negative en-

\textsuperscript{13} EPA instituted netting in 1974, as its first foray into applying trading approaches to pollution control. Robert W. Hahn & Gordon L. Hester, \textit{Where Did All the Markets Go? An Analysis of EPA's Emissions Trading Program}, 6 \textit{Yale J. on Reg.} 109, 133 (1989); see also infra note 15.

\textsuperscript{14} Bubbles were introduced in 1979. An existing facility is regulated on the basis of an imaginary bubble placed over the complex; emission levels from individual points within the bubble can be freely "traded" as long as total emissions do not increase. Most bubbles have involved internal trades within one plant, but there have been a few multiplant bubbles. Robert W. Hahn & Gordon L. Hester, \textit{Marketable Permits: Lessons for Theory and Practice} 16 \textit{Ecology L.Q.} 361, 372 (1989); see also Emissions Trading Policy Statement, 51 Fed. Reg. 43,814, 43,814 (1986).

\textsuperscript{15} Netting allows an expanding facility to avoid the strict standards applied upon plant modification by using internal trading to keep emissions increases below a significant level. Because "insignificant" increases are allowed, netting causes some environmental degradation. Hahn & Hester, \textit{supra} note 14, at 371; 40 C.F.R. § 52.21 (1992).

\textsuperscript{16} Offsets have been used since 1976 to allow continued economic development in nonattainment areas, where all emission increases are prohibited. After applying the technology-based standards, a new or expanding source can purchase emissions reduction credits from other sources in the nonattainment area to offset its planned emission increases, resulting in no net degradation of air quality. Hahn & Hester, \textit{supra} note 14, at 371-72; 40 C.F.R. § 52.24(c).

\textsuperscript{17} Banking allows sources to save emissions reduction credits for future use under state administered programs. Banking has been allowed since 1979. 44 Fed. Reg. 3274, 3285 (1979).
Environmental effects respectively. Cost savings under the offset program cannot be estimated directly, but the program has allowed continued siting of new sources in nonattainment areas with no net negative environmental impact; without offsets, new economic development with any significant emissions would have been prohibited.

Trading between firms under these programs, however, has fallen far short of predicted cost-effective opportunities, indicating that much of the potential cost savings have not been achieved.

The only very successful program to date, in terms of trading activity, size of estimated cost savings, and environmental benefits, is the lead trading program. This program established lead credits to increase the flexibility of petroleum refineries and ease the phase-out of leaded gasoline. Each refiner was allocated a decreasing pool of annual rights to add lead to gasoline, and was allowed to trade these rights with other refiners or bank unused rights for a limited period of future use.

Two new trading programs embody market mechanisms more ambitious than any yet tried, but it is still too early to evaluate their success. The 1990 amendments to the acid rain provisions of the Clean Air Act (the CAA) created the sulphur dioxide (SO₂) allowance trading program, which is designed to reduce SO₂ emissions from utilities by roughly fifty percent by the year 2000. The market has been less active than expected, but trading for the second set of compliance deadlines is not yet fully underway. The second program is the South Coast Air Quality Management District’s (SCAQMD) recently approved RECLAIM, a comprehensive trading program targeting two types of urban air pollutants: nitrous oxides (NOₓ) and

18. Cost savings are estimated at $525 million to $12 billion for netting and $435 million for bubbles. Hahn & Hester, supra note 14, at 374.
19. Id. at 375.
20. Hahn & Hester, supra note 13, at 110.
21. Hahn & Hester, supra note 14, at 389. One measure of the level of trading activity is the number of lead rights traded as a percentage of all lead used. From 1983 to 1987 when the program was in place, this percentage ranged from nearly 10% to over 50% annually. Id. at 387. Hahn and Hester roughly estimated cost savings at several hundred million dollars. Id.
22. See 40 C.F.R. § 80.20(d); Hahn & Hester, supra note 14, at 380-91.
23. See Hahn & Hester, supra note 14, at 381-84.
CONTROLLING SELENIUM DISCHARGES

sulfur oxides (SOx). The program includes the largest permitted stationary sources of these pollutants in the South Coast air basin.

One of the few U.S. permit trading programs that have been implemented to address water pollution is the Wisconsin Fox River program, which allows companies to trade discharge rights for wastes that increase biological oxygen demand. As of 1992, however, no trades had occurred since the program's implementation in 1981. Other programs allow trading between point and nonpoint sources of pollutants to such water bodies as the Dillon Reservoir and Cherry Creek Reservoir in Colorado, and the Tar-Pamlico River Basin in North Carolina. Only one trade has taken place. It may be too early, however, to assess the ultimate success or failure of at least some of these programs.

B. Necessary or Beneficial Elements for Trading Program Success

The empirical studies of existing trading programs and the theoretical literature on designing permit trading programs make it possible to derive some general criteria for what will work and where it will work. Success depends not only upon how well the program functions once in place, but also upon whether the particular problem context allows a trading system to be developed, approved, and implemented. While it is impossible to identify conditions of a problem context and of program design that will guarantee success, certain elements are

27. RECLAIM encompasses all major stationary sources with NOx and SOx emissions generally greater than four tons per year. South Coast Air Quality Management District, 1 RECLAIM: The Regional Clean Air Incentives Market, Oct. 1993, at EX-2 [hereinafter RECLAIM Executive Summary]. This accounts for approximately 65% of NOx permitted emissions and 80% of SOx permitted emissions. Id. at EX-5.
30. Id. at 4; Bartfeld, supra note 2, at 83-88.
31. The trade occurred in the Dillon Reservoir program. WATER POLLUTION, supra note 29, at 4. Unanticipated declines in point source loads in Dillon—due to overestimation of point source loads in the initial allocation, lower than expected population growth rates, and improved operating efficiencies for treatment plants—have eliminated any incentive to trade. Bartfeld, supra note 2, at 84-85; WATER POLLUTION, supra note 29, at 4, 12.
32. While the Dillon Reservoir program was approved in 1984, Bartfeld, supra note 2, at 84, the Tar-Pamlico program only began operating in 1992. WATER POLLUTION, supra note 29, at 4.
either necessary for, or would facilitate, establishing a successful trading program.\footnote{See Industrial Economics, Inc., The Benefits and Feasibility of Effluent Trading Between Point Sources: An Analysis in Support of Clean Water Act Reauthorization 2-13 (May 1992) (unpublished draft, prepared for EPA, on file with author) (identifying many of the same necessary conditions for trading to be beneficial).}

First, the environmental problem must be physically amenable to a trading approach. The harm must relate to total mass loading of pollution to the environment and be independent of particular sources. Trading is not appropriate where concentrations of the pollutant in certain areas, or "hot spots," pose a concern, since a reallocation of pollutants could potentially defeat environmental protection goals.

Second, a variety of economic considerations influence market functioning. If trading is to reduce control costs, there must be potential cost savings in a redistribution of reduction efforts among firms. In other words, firms must face different marginal costs of reducing emissions to the required level.\footnote{In the lead trading program some refineries had greater difficulty than others in meeting the lower lead limits. Hahn & Hester, supra note 14, at 384.} Moreover, the costs must be of sufficient magnitude to make trading worthwhile.\footnote{While potential cost savings from trading on the Fox River were estimated at \$7 million per year, the Wisconsin paper industry's pollution control costs are estimated to be less than one percent of the cost of the product. Thus, the total savings from trading were insignificant relative to the overall production costs. Martin H. David & Erhard F. Joeres, Is a Viable Implementation of TDPs Transferable?, in Buying A Better Environment: Cost Effective Regulation through Permit Trading 233, 240 (Erhard F. Joeres & Martin H. David eds., 1983); Hahn & Hester, supra note 14, at 391.} Experience has shown, however, that the existence of large differential costs is far from sufficient. All permit programs instituted to date rested on the premise that reallocating reduction efforts would decrease control costs, but the low level of trading in many cases shows that other factors may prevent the predicted savings.\footnote{For example, Hahn and Hester found that firms use bubbles only where there are large cost savings (upwards of several million dollars per firm). Hahn & Hester, supra note 13, at 124 (Table 3), 125. Bubbles that would provide smaller cost savings are discouraged by the cost of submitting an application and the low likelihood of approval. Id. at 125.}

The incentive to trade provided by cost differentials must be accompanied by the opportunity to trade. The availability of excess tradable reductions is one key to trading opportunity. Lack of available credits was a frequent constraint on the market in the offset program.\footnote{John P. Dwyer, California's Tradable Emissions Policy and Greenhouse Gas Control, 118 J. Energy Engineering 59, 64-65 (1992).} Permit availability, in turn, depends upon the firms' technological capabilities to reduce emissions to different levels and upon the level of control required. If there is just one possible control technology, which can reduce emissions only to the required level,
generally no excess emissions reductions will be available to trade, since each firm can only just meet its own limit. Moreover, reduction options are likely to be increasingly constrained as emission standards tighten. Th.oretically, a trading system should encourage firms to develop innovative technology to decrease control costs; the costs of technological development can then be recouped through permit sales. In practice, however, it appears that this incentive has had relatively little impact.

In addition, there must be a sufficient number of market players and transactions to produce a clear price signal for a competitive market to function. The number of players often will depend upon the geographic scope of the market, which in turn should be determined according to the range within which pollution reductions are interchangeable. For the cost-effective competitive market price to be attained, none of the players must be influential enough to exercise market power.

Assuming control costs could be reduced through permit trading, transaction costs will significantly affect whether these potential sav-

38. Tight and tightening standards were a large factor in the lack of available credits in the SCAQMD offset market. See id. at 65-66.


40. This is certainly true for the SCAQMD. Dwyer, supra note 37, at 65. The lack of innovative technologies is probably due to factors such as uncertainty regarding market prices and demand, as well as regulators' tendency to require implementation of any new technology as a technology-based standard. Id. But see Dudek & Palmisano, supra note 3, at 234-36 (arguing against criticism that trading approaches do not produce technological innovation).


42. Cf. Thomas H. Tietenberg, Transferable Discharge Permits and the Control of Stationary Source Air Pollution: A Survey and Synthesis, 56 LAND ECON. 391, 396-98 (1980) (presenting a far more detailed discussion of considerations in defining the market's geographic scope). Pollution reductions are interchangeable where the emission of a given quantity of a pollutant would have the same impact on the environment, regardless of the location of the pollutant's source.

43. Thomas H. Tietenberg, Economic Instruments for Environmental Regulation, 6 OXFORD REV. ECON. POL'Y 17, 22 (1990); cf. Hahn & Hester, supra note 14, at 364 (stating that while the exercise of market power will "influence the distribution of gains in trading, total cost savings may still be significant"). Market power occurs when a firm has sufficient influence over the market that it can set prices above the level of marginal cost; monopolies are one example. EARL L. GRINOLS, MICROECONOMICS 702 (1994). While not a necessary factor, the presence of noncompetitive firms is beneficial, since competitor firms may be less willing to trade with each other. See Industrial Economics, Inc., supra note 33, at 2-5.
ings are realized. Transaction costs inhere to some extent in any market, but their magnitude can vary according to market design. These unavoidable costs include the information costs of finding interested buyers and sellers, and the costs of arranging deals. While permit availability is increased by a larger number of players, having more participants also raises the information and bargaining components of transaction costs.

Additional transaction costs may be imposed by regulators in the form of requirements—sometimes mandated by statute—for firms to conduct studies to quantify reductions gained or the amount of credits needed to offset a certain activity, or to acquire regulatory approval for a trade. Hahn and Hester have credited part of the success of the lead trading program to the fact that EPA did not insist on preapproving trades; rather, the refineries reported trades to the Agency at the end of each calendar quarter. Also, the required reporting of quantities of lead added was straightforward; it did not necessitate expensive air quality monitoring.

Uncertainty and risk impose additional constraints, which are at least partially subject to regulatory control. Uncertainty about the permanence of an emissions reduction credit and its value under new regulations imposes a substantial disincentive to trading activity. The greater the regulatory authority to change the terms of the market, the greater the uncertainty for firms fearing exercise of regulatory discretion. Regulators may decrease market availability of credits through directly confiscating credits, tightening reduction standards, or imposing additional constraints. Overall, the expected savings

44. "High transactions costs are the single most important determinant of program performance." Hahn & Hester, supra note 14, at 376.
46. For example, in the Fox River program, dischargers could not trade simply to reduce compliance costs; rather they had to show either new or increasing production, or a need for additional credits based on the infeasibility of meeting the discharge standards through technological means. Wis. Admin. Code § NR 212.115(2)(c) (1988); Industrial Economics, Inc., supra note 33, at 2-5.
47. Dwyer, supra note 45, at 109-10; Hahn & Hester, supra note 14, at 377-78. Where regulators must approve each trade, uncertainty about obtaining regulatory approval imposes an additional transaction cost.
49. See id. at 388 n.146 (stating that EPA received data on lead use from refiners' self-reports on the quantities added and from manufacturers of lead additives' sales reports).
50. This has been particularly evident in programs that allow banking, where companies have been discouraged from banking credits for fear of confiscation by regulators. Hahn & Hester, supra note 13, at 131.
51. One such constraint is the use of a trading ratio, whereby credited reductions must more than compensate for increases elsewhere. Generally, offset ratios range from 1.1:1 to 8:1. AER*X, EMISSIONS TRADING AND THE NATURAL GAS INDUSTRY 10 (1990).
produced by a given trade must outweigh the transaction and risk costs in order for any trading to occur. Smaller transaction and risk costs relative to cost savings increase the likelihood that participants will develop a viable market and realize the savings.

Finally, legal, institutional, and political conditions must be appropriate for a workable permit trading system to be developed, approved, and implemented. The relevant statutory authority must either explicitly or implicitly approve a permit trading approach. Additionally, some political constituency must support implementing tradable permits. To date, most trading programs appear to have been initiated by regulators, affected local groups, or Congress, often as a compromise to break political deadlock over expensive reductions. In the absence of clear statutory authority for permit trading, as in the water context, regulators' support is particularly critical to developing and approving a trading system, since regulators alone can claim that implementing the statute through market mechanisms is within their mandate to exercise discretion. Regulatory support is even more necessary for effective implementation.

Attitudes of the regulated industries, public interest groups, and sometimes the general public also will be critical for trading programs, particularly in the approval stage. Since concentrated "losers" are the most likely source of substantial opposition, the absence of such losers is a key factor motivating decisions to use economic instruments. Indeed, positive external support may be necessary to overcome the hurdles of inertia and the status quo in policy approaches. The most ambitious permit trading programs pursued to date, the acid rain program and RECLAIM, both are supported by at least some environmentalists and industry

---

Dillon Reservoir program applies a 2:1 ratio, in part to account for the uncertainty in estimating the credited nonpoint source reductions. Bartfeld, supra note 2, at 68, 84.

52. Risk costs are the additional costs imposed by uncertainty in a transaction (i.e., a gain is reduced in value by the probability that it will not occur).

53. The netting, offset, and bubble programs were initiated by EPA. See supra notes 13-16. The acid rain program allowed action on the issue after years of deadlock, when the trading option was proposed in President Bush's bill. Michael Weisskopf, Clean Air Pact Ends Acid Rain Impasse; Longtime Adversaries Bridge Differences, WASHINGTON POST, Mar. 2, 1990, at A1. Trading was a compromise approach that balanced the widely varying regional impacts of significant sulfur dioxide reductions. Id. (While presidential support was key in that instance, due to divided party control of the Executive and Congress, strong political support from the executive branch does not seem to be necessary elsewhere.) RECLAIM also has been promoted by SCAQMD as a means of getting further reductions from point sources despite potentially high control costs. RECLAIM Executive Summary, supra note 27, at EX-3 to EX-5.

54. Hahn & Stavins, supra note 41, at 35.


56. Hahn & Stavins, supra note 41, at 31.
representatives. Not only did environmental groups push for approval of the acid rain trading system, but also they were instrumental in developing the proposals and working out the details. In general, overcoming the status quo is easiest in new issue areas, where either the pollutant of concern and/or the emitters have not previously been regulated. In a new area, regulators and regulatees have not invested resources in understanding and using an existing system, nor have expectations developed in reliance on that system.

The attitudes of industry, interest groups, and the public depend on their understanding of how permit trading theoretically works and their perceptions about how it actually performs. Thus, a better general understanding of tradable permit approaches, particularly among the most active public interest groups and industry representatives, may be important for program development and approval. Understanding and familiarity may not be sufficient to produce support, however. Instead, industry and environmentalists must view tradable permits as actually advancing their respective agendas if each is to form an active constituency for the program. Environmentalists are most likely to be supportive if they feel assured that the underlying environmental goals will be achieved; strict monitoring and enforcement provisions are probably necessary to provide such assurance. Industry’s support will depend on its perception of true cost savings, net of transaction costs (such as perceived regulatory “hassle”). Both groups are most likely to advocate trading where they believe it is the best outcome they can get; unfortunately, this may become apparent only after the groups have spent years battling each other to a standstill, as was the case with acid rain.

57. Id. at 34 n.181 (regarding RECLAIM); Telephone Interview with Terry Young, Senior Consulting Scientist, Environmental Defense Fund (May 5, 1994) (regarding acid rain) (on file with author). The Environmental Defense Fund (EDF) has been the environmental group most supportive of market incentive approaches, particularly permit trading. See Hahn & Stavins, supra note 41, at 33 & n.180.


60. See id. at 36 (discussing the impact of understanding on interest group attitudes).
61. See infra notes 262-76 and accompanying text.
62. Hahn & Stavins, supra note 41, at 33; see also, e.g., Telephone Interview with Trish Mulvey, Member, Environmental Action Committee, Santa Clara Valley Audubon Society (Mar. 17, 1993) (on file with author).
63. Hahn & Stavins, supra note 41, at 34.
64. See Weisskopf, supra note 53, at A1.
C. Issues Specific to the Water Pollution Setting

Discharges to water bodies create a unique policy challenge for permit trading. The physical differences between the water and air media, and the resulting behavior of pollutants, raise several issues specific to trading in the water pollution context. Differences between the legal settings established by the CWA$^{65}$ and the CAA$^{66}$ also make trading more difficult in the water context.

I. Physical Context of Water Pollution

Water and air media may tend to exhibit different mixing qualities.$^{67}$ Pollutants that are broadly dispersed, or well mixed, in a wide geographic area create regional pollution problems, while poorly mixed pollutants will concentrate in a given area, producing hot spots. Marketable permits are an appropriate tool for addressing regional pollutant problems, but should not be used where significant hot spots may result from trading.$^{68}$ The air and water media each may exhibit both regional pollution problems and the potential for hot spots.$^{69}$ Water pollution problems tend to be more restricted geographically than air problems, however, and pollutants tend to disperse less quickly and completely in a watershed than in an airshed. Unlike such transboundary air pollution problems as acid rain, chlorofluorocarbons, and the global warming gases, water quality problems receiving regulatory attention rarely concern total loading over a large geographic area.$^{70}$ Also, the relative slowness of pollutant dispersal in water makes ecological hot spots a more common concern with water pollution trading than with air trading.$^{71}$ The area for which permits can justifiably be freely traded is thus constrained to a given well-mixed water body or portion of a water body. This means far fewer firms will likely be able to participate in a water market than would in a comparable airshed.

In considering permit trading for controlling water pollution, regulators also must account for the differences between conservative

---

69. The Los Angeles and Denver Basins are two examples of airsheds with poor mixing and hot spot problems; PCB contamination in the Great Lakes is a regional water pollution problem. Interview with Dr. William S. Pease, *supra* note 12.
70. *Cf.* John W. Kindt, *Solid Wastes and Marine Pollution*, 34 CATH. U. L. REV. 37, 90 (noting that “[l]ittle attention has been given to the environmental problems created by the total loading of wastes into the ocean from all input sources”).
71. This is not as true for human health hot spots, where air toxics could present serious trading concerns. Monitoring is probably not significantly different for the two media, although it is easier to do multipollutant screens in water.
and nonconservative pollutants. Conservative pollutants like toxic metals tend to accumulate in a given ecosystem. The control issue thus becomes total ecosystem loading capacity over the timeframe of contemplated discharges. Nonconservative pollutants, such as those contributing to biological oxygen demand, are assimilated by the ecosystem over shorter timeframes. The rates of assimilation usually vary, as they depend on factors such as temperature and flow rates of a stream. Assimilation rates may also vary at different discharge points along the stream, as well as at different times. Thus, equal quantities of effluent will have different environmental impacts depending upon the time and location of the discharge. Designing a tradable permit system that accounts for variable impacts of a traded unit of pollution is considerably more complicated than designing a system where equal units have basically equal impacts. Thus, the viability of trading for water pollutants may be more pollutant dependent than for air pollutants.

2. Legal Requirements of the Clean Water Act

The CAA and the CWA both establish complex regulatory structures governing pollution releases, and many of their provisions will affect whether and how permit trading programs can be used. The CWA requirements relevant to permit trading are discussed in part II. This subpart notes more generally several characteristics unique to the CWA regulatory system that will affect permit trading in the water medium.

The CWA requires all point source dischargers to meet technology-based treatment standards established for each industry or process category. Each discharger must obtain and comply with the

72. See Donna Downing & Stuart Sessions, Innovative Water Quality-Based Permitting: A Policy Perspective 57 J. WATER POLLUTION CONTROL FED’N 358, 362 (1985) for a description of different types of water quality-based permitting, including both a variety of tradable permits, and permits that make greater use of the assimilative capacity of a water body, such as seasonal or variable permits.

73. A similar type of problem could be presented in the air context if emissions of smog-forming NO, or hydrocarbons were regulated on the basis of current weather conditions (e.g., wind speeds and temperature), which influence the degree of smog formation and consequent health risk.


75. Section 1311(a) prohibits the discharge of pollution except in compliance with the provisions in 33 U.S.C. §§ 1312, 1316, 1317, 1328, 1342, & 1344 (1988 & Supp. IV 1992). The EPA Administrator must publish guidelines for effluent limitations based on the “best practicable control technology currently available” (BPT) and the “best conventional pol-
terms of a National Pollutant Discharge Elimination System (NPDES) permit, which incorporates the technology-based standards.\textsuperscript{76} In addition, states set water quality standards consisting of designated beneficial uses for each water body and ambient water quality criteria set to protect those uses.\textsuperscript{77} Where the technology-based permit limits are insufficient to meet the water quality standards, a second level of water quality-based effluent permit limits must be established.\textsuperscript{78}

Unlike the CAA,\textsuperscript{79} the CWA does not explicitly authorize using market mechanisms to meet reduction goals.\textsuperscript{80} In addition, mandated technology-based standards probably foreclose the possibility of applying trading to meet the first-level effluent limits.\textsuperscript{81} Trading opportunities appear restricted to those water bodies that are water quality limited (i.e., need a second level of controls). Thus, probably only those reductions beyond what is required by technology-based standards could be traded.\textsuperscript{82} While it is likely that trading could be used in some circumstances, uncertain legal authorization of market mechanisms under the CWA has contributed to the small number of discharge trading programs.\textsuperscript{83}

One advantage of the water medium for permit trading is that point sources, which are the likely market participants, already hold permits and monitor their emissions.\textsuperscript{84} In contrast, while the CAA

\textit{luratant control technology” (BCT). }Id. \textsection 1314(b). The first tier BPT controls apply to existing sources, while the second tier BCT controls apply to conventional pollutants. Toxic and nonconventional pollutant effluent standards are based on the “best available technology economically achievable” under \textsection 1311(b)(2).

\textsuperscript{76} Id. \textsection 1342(a)(1); 40 C.F.R. \textsection 122 (1992). Either EPA or a state with an approved permit program issues the permit.

\textsuperscript{77} 33 U.S.C. \textsection 1313(a), (c)(2); 40 C.F.R. \textsection\textsection 130.3, 131.10, 131.11.

\textsuperscript{78} 33 U.S.C. \textsection\textsection 1311(b)(1)(C), 1312(a), 1342(b); 40 C.F.R. \textsection 122.44.

\textsuperscript{79} Congress first recognized market-based approaches in the form of offsets, in the 1977 CAA Amendments, although EPA had experimented with offsets even prior to statutory authorization. Hahn & Stavins, \textit{supra} note 41, at 15; 41 Fed. Reg. 55524 (1976); see \textit{also} Chevron U.S.A., Inc. v. NRDC, 467 U.S. 837, 847-48 (1984) (discussing EPA’s attempt to address the problem of nonattainment through offsets). Congress added more explicit directives to pursue market mechanisms in the 1990 CAA Amendments. See, e.g., 42 U.S.C. \textsection\textsection 7502(c)(6), 7502(c), 7511a(a)-(e), 7511a(g), 7651-76510.

\textsuperscript{80} It seems likely, however, that inclusion of such language will be an issue in the upcoming reauthorization, which is due this year. \textit{See, e.g., WATER POLLUTION, supra} note 29 (requested by the House Committee on Public Works and Transportation).

\textsuperscript{81} Industrial Economics, Inc., \textit{supra} note 33, at 1-2.

\textsuperscript{82} Id.\textsuperscript{ Also, EPA has not set technology-based standards for some water pollutants, and trading could be used to meet whatever water quality-based standard is set for these pollutants.

\textsuperscript{83} \textit{WATER POLLUTION, supra} note 29, at 5.

\textsuperscript{84} Interview with John P. Dwyer, Professor, School of Law (Boalt Hall), University of California at Berkeley, in Berkeley, Cal. (Jan. 19, 1993) (on file with author). Point sources of water pollution must hold NPDES permits specifying discharge limits, but nonpoint sources of water pollution are not individually permitted and often are altogether unregulated. \textit{See} 33 U.S.C. \textsection 1342(f); Bartfeld, \textit{supra} note 2, at 45, 48.
requires technology standards to be met, many small sources do not need to hold permits to emit air pollutants. Trading programs, which require careful tracking and monitoring of emissions from each source, essentially mandate a permit-based program. Thus, a shift to permit trading in the water context might impose a smaller additional regulatory burden than it would in the air context.

D. Types of Transferable Discharge Permit Systems

The design of a transferable discharge permit system for water pollution could range along a spectrum of different market incentive approaches, from a system constituting a small shift from current command-and-control approaches, to a relatively unfettered market in pollution discharges. For clarity of discussion, these approaches can be grouped into three discrete categories: an allowance system, a group permit system, and a tradable limits system.

1. Allowance System

The purest market system can be termed an allowance system, analogous to the acid rain allowance trading system. Under this approach, each firm receives a given quantity of allowances annually. One allowance represents the right to emit one unit of pollution. The total number of allowances distributed annually is equal to the maximum level of allowable emissions, as established by the regulatory agency. At the end of each year, each firm must show that it holds a quantity of allowances equal to the quantity of pollution it emitted. Apart from this accounting requirement, firms may freely buy, sell, or save excess allowances for use in subsequent years, with no regulatory approval required for any transaction. This approach can be modified to provide additional environmental assurances. Prudent constraints might include restricted trading zones or quantities (to protect against hot spots), quarterly or more frequent accounting (to reduce the risk of substantial year-end violations), and regulatory recording of each transaction (to reduce the risk of cheating through convoluted and confusing transfer arrangements).

2. Group Permit System

An alternative trading approach is the group permit, modelled after permit arrangements occasionally used to control nonpoint

---

85. Interview with John P. Dwyer, supra note 84. While all significant point sources of air pollution do hold permits in California, this is not necessarily the case in many other states.

86. See supra note 24 and accompanying text; see infra part II.B.2.a.
sources of water pollution. Under a group permit approach, as with allowances, the regulatory body determines the maximum level of allowable pollution discharged, in mass units of pollution. This quantity is not allocated among the dischargers, however. Instead, the dischargers are held jointly responsible as a group for their total quantity of emissions; they allocate allowable individual pollutant quantities among themselves. Regulators could enforce the overall limit on the basis of joint and several liability, which would allow the agency to recover fines for group violations from any one or all of the dischargers. A targeted entity could then attempt to recover from other parties through contribution actions.

3. Tradable Limit System

A tradable limit approach would allow permitholders to reallocate their current permit limitations amongst themselves through a formal permit modification process. Tradable limits could be instituted using some CWA provisions allowing redistribution of certain discharge quantities.

The CWA water quality provisions regulate discharges by setting a maximum total loading level of a pollutant into a water body. Allowable discharges up to this level are then apportioned—as “waste-load allocations” (WLA’s)—among all the dischargers to the water body. The statute allows WLA’s to be redistributed among dischargers, as long as the overall water quality standard is attained.

87. See, e.g., SFBRWQCB, Order No. 91-146, NPDES Permit No. CA0029831 (Oct. 16, 1991) [hereinafter NPDES Permit I]. In air terms, a group permit is an interfirm bubble.

88. This is similar to the joint and several liability approach applied under CERCLA to potentially responsible parties for a Superfund site. United States v. Chem-Dyne Corp., 572 F. Supp. 802, 805-11 (S.D. Ohio 1983). To date, this program has produced massive litigation and staggering costs. For example, in the fiscal year 1993, the Federal Government spent about $1.6 billion, and industry spent another $1.5 billion on Superfund work. Keith Schneider, Superfund at 13: A White Knight Tarnished, N.Y. TIMES, Sept. 6, 1993, at 7.

89. See infra part II.A.3.b.ii.


91. 40 C.F.R. § 130.2(h).

92. The CWA, 33 U.S.C. § 1342(o)(1), prohibits establishing less stringent water quality-based effluent limitation “except in compliance with” § 1313(d)(4). OFFICE OF WATER, U.S. ENVTL. PROTECTION AGENCY, EPA/505/2-90-001, TECHNICAL SUPPORT DOCUMENT FOR WATER QUALITY-BASED TOXICS CONTROL 113 (1991) [hereinafter TECHNICAL SUPPORT DOCUMENT]. The Act also allows less stringent water quality-based effluent limitations in a permit for discharge into nonattainment water only if (1) the existing permit limitation was based on a total maximum daily load (TMDL) or other WLA, and (2) the cumulative effect of all such revised effluent limitations will assure attainment of water quality standards. 33 U.S.C. § 1313(d)(4); see also TECHNICAL SUPPORT DOCUMENT, supra, at 113.
Redistribution occurs not through informal trades, but rather through the statutorily mandated permit modification process, which includes public notice and comment, and a hearing.93

While the mechanisms for reallocating WLA’s currently exist in the CWA, it does not appear that this process has been used, particularly not with monetary payments in the context of a formal market.94 One barrier to such use is that states have not implemented WLA’s for many water bodies,95 and they have implemented few, if any, for toxics control.96 The rigorous permit modification requirements act as an additional barrier.97

II
EVALUATING DISCHARGE PERMIT TRADING IN A SPECIFIC PROBLEM CONTEXT

To evaluate the potential of using permit trading to address a particular environmental problem, one must first identify characteristics of the problem context that are relevant to how trading will function, and then design a trading system appropriate for the setting. This part illustrates the evaluation process by considering how well trading might work to address the selenium problem in the San Francisco Bay Estuary. While the San Francisco Bay selenium problem is a unique factual situation, it raises many of the permit trading issues that would also arise in other point source pollution settings.

A. Examination of Factors Relevant to Permit Trading

For our purposes, the selenium pollution problem context can be divided into physical, regulatory, legal, economic, institutional, and political aspects. The issues identified in this subpart relate to the basic question of using a trading approach to address a given water pollution problem; they are independent of the specific type of trading system chosen or details of the system design.

93. 40 C.F.R. § 124.5.
95. The extent of implementation varies widely by state; some states have not even considered how to use WLA’s, while others have fully implemented WLA procedures. Id. at 1-3 n.6.
96. Interview with Dr. William S. Pease, Environmental Health Policy Program, School of Public Health, University of California at Berkeley, in Berkeley, Cal. (Apr. 4, 1993) (on file with author).
97. See 40 C.F.R. § 122.63 (allowing permit modifications only where one of the listed causes is present). See infra notes 182-86 and accompanying text.
1. Physical Context—Selenium Discharges to the San Francisco Bay Estuary

Selenium is now found in bay organisms at potentially toxic levels. It threatens both wildlife and human health and is driving an effort to reduce mass loading to the system. While selenium is an essential trace element for animals and humans, it is toxic in greater amounts. Selenium bioconcentrates in the foodchain and can cause hatching problems for waterfowl and, at higher levels, reproductive defects. Selenium levels currently found in the liver tissues of two species of diving ducks in the bay are comparable to or exceed the levels associated with the severe waterfowl reproductive problems at Kesterson Wildlife Refuge. Elevated selenium levels also have been found in other bay organisms, such as bivalves, fish, and other birds. While biologists have not observed any histopathological effects as yet, adverse effects could be expected at these levels of foodchain selenium enrichment. The rate and magnitude of adverse ecological effects is uncertain, yet it appears that quick action could avoid a pending problem.

The main pathway for selenium uptake is through bioconcentration by phytoplankton, which are subsequently consumed by other organisms, causing some further selenium biomagnification up the foodchain. The primary cause of adverse ecological effects is the total mass of selenium entering the foodchain, not, for the most part, ambient levels in the water column. This is significant because selenium is currently regulated on the basis of water column concentrations. Assessment and control of the selenium issue are further hindered by the complexity of the impacts of different chemical species of selenium and of the processes by which it is transformed.

---

99. Id. at 16-17; Staff Report, supra note 8, at 9-11.
100. Mean selenium levels in bay waterfowl range from 14 to 83 micrograms per gram dry weight (µg/g dw) and 40 to 209 µg/g dw, depending upon the species. At Kesterson, levels ranged from 46 to 82 µg/g dw. Staff Report, supra note 8, at 8. Publicity about the wave of deformed wildfowl hatchlings at Kesterson in 1983 brought attention to the potential problems associated with selenium in agricultural return flow. See, e.g., Lloyd Carter, Selenium Levels Rising at Kesterson Refuge: Poisons Still Found Despite Costly Cleanup, S.F. CHRON, Feb. 21, 1994, at A11.
101. These include the endangered California Clapper Rail, mussels, deposit-feeding clams, white sturgeon, and striped bass. Staff Report, supra note 8, at 9-10. The declines of striped bass in the bay may be due in part to selenium contamination. Id. at 10.
102. Id. at 8.
103. Id. at 11.
104. Id. at 4.
105. Id. at 6.
106. Id. at 14-15.
This analysis of the exposure pathway is driving local regulators to advocate a focus on mass loadings rather than on ambient concentrations of selenium. However, it is currently impossible to determine an ecologically “safe” mass selenium load. Therefore, the regional board staff has proposed a “mass loading reduction strategy.” This iterative approach would require mass emission loading levels to be reduced incrementally until selenium concentrations in sediment, algae, and indicator organisms meet proposed concentration guidelines.

The major sources of selenium to the bay estuary are oil refineries, municipal sewage treatment plants, and riverine loading from the Sacramento and San Joaquin Rivers. The refineries’ contribution varies from roughly fifty-eight percent of total selenium loading during periods of high riverine flow, to roughly seventy-four percent during low flow periods. At average riverine flows the refineries contribute approximately sixty-six percent of the total selenium. Selenium enters the refining process as a natural component of crude oil, but levels of selenium vary across different types of crudes. California crude oils from the San Joaquin Valley used by some of the bay refineries have particularly high levels of selenium. The refining process removes selenium from the oil to varying degrees, partially depending upon the type of product manufactured. Making lighter products such as gasoline, particularly the new “clean fuels,” will tend to result in higher selenium discharges. Riverine selenium principally comes from nonpoint sources, which include large volume, low concentra-

107. See id. at 37-38.
108. See id. at 37.
109. Id.
110. Id.
111. Id. at 13.
112. See SFBRWQCB, Proposed Amendment to the Water Quality Control Plan, San Francisco Bay Region: Mass Emission Reduction Strategy for Selenium 2 (June 21, 1993) (on file with author) [hereinafter Staff Report Amendment]. At average riverine flow, the refineries load an estimated 7 kilograms per day (kg/d) of selenium, while riverine sources and municipal sewage treatment plants produce an estimated 2 kg/d and 1.6 kg/d respectively. Id. Riverine sources’ contributions range from roughly 9% at low riverine flow to roughly 28% at high flow. The average flow is roughly 19%. See id. (percentage calculations based on quantity estimates in the source).
114. See Hearing, supra note 113, at 72-73 (testimony of Dan Glaze, Senior Environmental Engineer, Shell Oil Co.). Again, however, there is not necessarily a direct correspondence between the kind of product and the level of selenium discharges. The specific production processes employed and wastewater treatment methods also will affect discharges.
tion, surface runoff to the Sacramento River, and agricultural drain-age to the San Joaquin River.\textsuperscript{115} Nonpoint sources pose different and difficult control issues. Also, many of the riverine sources are outside the San Francisco regional board's jurisdiction.\textsuperscript{116}

2. Regulatory Context—Current System for Regulating Selenium Discharges to the San Francisco Bay Estuary

Dischargers to the San Francisco Bay Estuary are regulated by a complex structure of agencies, acting under both federal and state statutes and regulations. The main agencies involved in water pollution control are the Federal EPA,\textsuperscript{117} largely acting through the regional EPA branch (region 9), the State Water Resources Control Board (state board),\textsuperscript{118} and the regional state agency (the SFBRWQCB).\textsuperscript{119} The state board establishes statewide policies and standards in the state plan,\textsuperscript{120} and the regional board sets region-specific policies and standards in the basin plan.\textsuperscript{121} Amendments to either plan must be approved by EPA.\textsuperscript{122} Amendments to the basin plan must be approved by the state board as well. Individual permit provisions are set by the regional board.\textsuperscript{123} Permit issuances, reissuances,\textsuperscript{124} and modifications must be submitted to EPA, but unlike plan amendments, individual permits automatically become effective

\textsuperscript{115} Interview with Dr. William S. Pease, \textit{supra} note 96; Telephone Interview with Terry Young, Senior Consulting Scientist, Environmental Defense Fund (May 25, 1993) (on file with author).

\textsuperscript{116} See Interview with Kim Taylor, \textit{supra} note 113; Telephone Interview with Terry Young, \textit{supra} note 115.


\textsuperscript{120} \textbf{Cal. Water Code} § 13140 directs the state board to formulate and adopt state policy for water quality control.

\textsuperscript{121} \textbf{Cal. Water Code} § 13240 directs the regional boards to adopt regional water quality control plans that conform to any state policy for water quality control.

\textsuperscript{122} Telephone Interview with David Smith, TMDL Coordinator, Watershed Branch, Region 9, U.S. EPA (Mar. 18, 1993) (on file with author).

\textsuperscript{123} See \textbf{Cal. Water Code} § 13263(a).

\textsuperscript{124} Permits can be changed either by modification or revocation and reissuance. When a permit is modified, only the conditions subject to modification are reopened. If a
in the absence of EPA disapproval. Permits are not approved by the state board. Rather, the board functions as an initial forum for appeals. Subsequent appeals can be brought in the courts.

In the bay, selenium is regulated based on the violation of water quality standards and failure to protect designated uses. While states must regulate discharges to water segments using both technology-based requirements and water quality-based requirements, EPA has not set any technology-based standards for selenium, so only water quality-based requirements are applicable here. The CWA requires states to set water quality standards for each water segment. These standards include a designation of uses to be achieved and protected for each water segment, and numeric and/or narrative water quality criteria designed to protect those uses. NPDES permits must comply with applicable water quality standards.

The bay's designated beneficial uses include estuarine habitat, preservation of rare and endangered species, and noncontact water recreation, which includes hunting. The applicable narrative water quality standards are contained in the state plan:

Enclosed bay and estuarine communities and populations . . . shall not be degraded as a result of the discharge of waste. . . . Toxic pollutants shall not be discharged at levels that will bioaccumulate in aquatic resources to levels which are harmful to human health. . . . The concentrations of toxic pollutants in the water column, sediments, or biota shall not adversely affect beneficial uses.

The state plan also has a relatively high numeric water quality criterion for selenium, which is widely viewed as insufficiently protective. The regional board's basin plan includes salt and freshwater permit is revoked and reissued, the entire permit is reopened and subject to revision . . . ."

125. 40 C.F.R. § 123.44; Telephone Interview with Laurie Kermish, Regional Counsel, Region 9, U.S. EPA (May 27, 1993) (on file with author).
127. CAL. WATER CODE § 13330.
128. Technology-based standards are set as effluent limitations in individual permits under 33 U.S.C. §§ 1311 and 1314(b). EPA must publish guidelines for effluent limitations, identifying the degree of effluent reduction attainable through the best practicable control technology currently available. Id. § 1314(b).
129. Id. § 1313(a).
130. 40 C.F.R. §§ 130.2(d), 130.3, 131.10; CAL. WATER CODE § 13241 directs regional boards to establish water quality objectives in water quality control plans that will ensure reasonable protection of beneficial uses.
133. State Plan, supra note 118, at 3.
134. The state plan numeric water quality objectives for selenium require ambient saltwater concentrations below 71 µg/l, based on a four-day average. State Plan, supra note 118, at 4. As a saltwater limit, this standard is not directly applicable to the San
numeric criteria for selenium, based on EPA's national water quality criteria.\textsuperscript{135} While current ambient selenium concentrations in the bay meet the state and regional numeric criteria, they violate the narrative standards designed to protect beneficial uses.\textsuperscript{136}

The 1987 Amendments to the CWA added provisions specifically addressing water quality-based control of toxic pollutants.\textsuperscript{137} Selenium is one of the toxics of concern identified by the Act.\textsuperscript{138} Section 304(l) requires states to identify and list water segments that fail to meet applicable water quality standards due to discharges of toxics, despite the application of technology-based effluent limitations.\textsuperscript{139} States also must identify the point source dischargers of those toxics\textsuperscript{140} and must submit individual control strategies (ICS's) for the listed sources.\textsuperscript{141} Each ICS consists of an NPDES permit and documentation to show that the water quality standard will be met "as soon as possible," and at most within three years of establishing the ICS.\textsuperscript{142}

The EPA listed portions of the North Bay as section 304(l) water segments in 1990 because elevated selenium levels in duck tissues were preventing the segments' beneficial uses from being met.\textsuperscript{143} The

---

Francisco Bay Estuary. This limit is based on EPA’s water quality criteria, which failed to account for either biomagnification or possible differential impacts on wildlife other than aquatic organisms. Staff Report, supra note 8, at 4; 33 U.S.C. § 1314(a)-(b) (EPA shall publish water quality criteria reflecting the latest scientific knowledge on the effects and characteristics of pollutants). The state plan includes a footnote to the selenium objective stating that “[e]specially in estuarine waters, regional boards should assess whether more stringent objectives or control measures are warranted in order to protect fish and wildlife and their consumption.” State Plan, supra note 118, at 4.

135. Staff Report, supra note 8, at 3; Basin Plan, supra note 119, at III-7, Table III-2A, III-9, Table III-2B. The basin plan sets the same saltwater numeric criteria as the state plan and sets a four-day average limit in freshwater of five μg/l. Id. at III-9, Table III-2B.

136. Staff Report, supra note 8, at 1; see infra note 143 and accompanying text.


138. 33 U.S.C. § 1317(a)(1); see also 40 C.F.R. § 423, app. A (listing the 126 “priority pollutants”).

139. 33 U.S.C. § 1314(l)(1)(B); 40 C.F.R. 130.10(d)(2). State lists must be approved by EPA. Where a state fails to submit a list or its list is disapproved, EPA is to implement the listing provisions. 33 U.S.C. § 1314(l)(3); 40 C.F.R. § 130.10(d)(10).


142. Id.; 40 C.F.R. § 123.46(d).

143. The State of California submitted its final lists and ICS's to EPA in 1989, but did not include the northern portions of the San Francisco Bay Estuary of concern here (Suiseun Bay, San Pablo Bay, and the Carquinez Strait). U.S. Envtl. Protection Agency, Region 9, Decision of the United States Environmental Protection Agency on Listings Under Section 304(l) of the Clean Water Act Regarding the State of California 2, 19 (Sept. 28, 1990). Since elevated selenium levels in waterfowl tissue had prompted a state health advisory affecting sport hunting of waterfowl, these segments' beneficial uses were not being met. Staff Report, supra note 8, at 4. Consequently, EPA disapproved the state's list, and added the North Bay segments and the six refineries to the 304(l) lists. U.S. Envtl. Protection Agency, supra, at 16. In response, the refineries brought a lawsuit against the state and
regional board adopted ICS's for the refineries in 1991.144 The board modified the refineries' NPDES permits to set interim discharge limits;145 it established a December 1993 deadline for each to meet concentration limits of fifty µg/l and mass emission rates for selenium discharges, based on the concentration limit.146 These mass emission levels represent a forty-six percent reduction from the 1990 loading.147 Nevertheless, the regional board staff and environmentalists believe the fifty µg/l limit is too high and does not account for bioaccumulation or protect wildlife.148 Nor does it appear possible that the three refineries currently emitting above the final permit limits will be able to meet those limits by the December 1993 deadline.149

3. Legal Context—State and Federal Laws and Regulations Affecting Permit Trading

Two steps are necessary to shift to a permit trading approach from the current regulatory system for controlling selenium discharges regional boards challenging their amended permit limits, claiming that designated uses are not impaired and that the segments should not have been listed. Western States Petroleum Assoc. v. EPA, No. 91-70388 (Cal. Ct. App. filed June 19, 1991); Five Refineries, Oil, Trade Group File Suit Over Selenium Discharges, Cal. Env't Rep (BNA) No. 18, at 311 (July 8, 1991); Telephone Interview with Gary J. Grimm, Counsel to the SFBRWQCB (Feb. 25, 1993) (on file with author); Telephone Interview with Laurie Kermish, supra note 125. This suit was settled in January 1994. Erin Hallissy, Refineries Win Delay in Selenium Standards, S.F. CHRON, Jan. 20, 1994, at A21; see also infra note 166.


145. The interim limits were to control mass emissions of selenium over the following three years. See id.

146. SFBRWQCB, Order No. 91-026 (Feb. 20, 1991); SFBRWQCB, Order No. 91-099 (June 19, 1991); Staff Report, supra note 8, at 4-5. Both the refineries and the environmental groups appealed these limits to the state board. Industry, Environmentalists Fight San Francisco Bay Selenium Limits, Cal. Env't Rep. (BNA) No. 21, at 369 (Aug. 19, 1991). The state board did not rule on the permit limit challenge, noting that the regional board planned to revisit the issue shortly. Telephone Interview with Laurie Kermish, supra note 125.

147. Staff Report, supra note 8, at 37.

148. See id. at 3. The limit is based on EPA's five µg/l water quality criterion for selenium in fresh water, which is modified for deep water discharges with a 10:1 dilution ratio. Id. at 3, 5. EPA's freshwater criterion apparently attempts to account for bioaccumulation by dividing the initial value by two as a safety factor, but the regional board staff does not find this satisfactory. Id. at 3-4. Moreover, the criterion was not derived to protect wildlife using selenium-contaminated habitats. Id. at 3.

149. These are Shell, Union, and Exxon. See id. at 37. Over the period of the interim limits total selenium discharges in fact have been increasing. Since 1989 the quantity of selenium discharged has risen by 60%. San Francisco Regional Water Board Urged to Control Selenium Discharges, Cal. Env't Rep. (BNA) No. 20, at 444 (Aug. 17, 1992). One of the refineries, Shell, is currently being sued by environmentalists for exceeding its interim limit and the narrative standard for selenium. California Pub. Interest Research Group v. Shell Oil Co., 840 F. Supp. 712 (N.D. Cal. 1993); Citizen Action Group Sues Shell Oil Alleging Excess Selenium Discharges, Cal. Env't Rep. (BNA) No. 1, at 15 (Nov. 9, 1992). The National Environmental Law Center is also involved in the suit.
CONTROLLING SELENIUM DISCHARGES

into the San Francisco Bay Estuary. The first step is to shift the focus of regulation from concentration limits to mass emissions. Currently, water pollutants probably may be limited either on the basis of maximum pollutant concentrations in effluent, maximum pollutant mass discharges, or both, but concentration limits are most common in NPDES permits. Trading concentration limits makes little sense conceptually and probably would be more difficult to implement than trading mass quantities discharged. The second step is to create some form of a permit trading system. The legal authorities regulating discharges to water segments impose a variety of clear and possible constraints on both of these steps.

a. Mass Emissions Reduction Strategy

The CWA appears to allow regulators to impose water quality-based controls solely on the basis of mass loadings, rather than on effluent concentrations. For a water quality-limited segment, the CWA requires the state to determine the total maximum daily load (TMDL) of a pollutant that will allow narrative and numerical water quality standards to be attained. Regulators determine a waste load allocation (WLA) for each individual point source, and load allocations (LA's) account for loading from nonpoint sources and natural background sources. Together, the sum of individual WLA's and LA's cannot exceed the TMDL. TMDL's can be expressed in terms of mass per time, toxicity, or other appropriate measures. Thus, water quality controls are based on a mass approach.

In the NPDES permit a WLA is translated into an effluent limit. Permit limits are commonly, but not always, expressed as effluent concentration limits. The CWA defines “effluent limitation” as “any restriction . . . on quantities, rates, and concentrations. . . .” While this wording is ambiguous, it can be read to include a restriction on any one of the three measurements. In practice, permits do not always specify limitations on all three measurements. Federal law

be approved by EPA. Id. § 130.7(d)(1).
151. 40 C.F.R. §§ 130.2(g)-(h), 130.7(c)(1).
152. Id. § 130.2(i).
153. Id.
154. See TECHNICAL SUPPORT DOCUMENT, supra note 92, at 93, 111. Permits must
include effluent limits when the segment exceeds either a numeric or narrative water quality standard. Id. § 122.44(d)(1).
156. See, e.g., SFBRWQCB, Order No. 92-111, NPDES No. CA0005134 (Sept. 18, 1992) (setting only mass limits for some discharge constituents and only concentration limits for others). Apparently, early permits for the refineries were solely mass based, and later were amended to add concentration limits. See Telephone Interview with Tom Mumley, Associate Water Resource Control Engineer, SFBRWQCB (Feb. 3, 1993) (on file
does not appear to require permit concentration limits for selenium; this allows regulators to focus on mass emissions.157

State legal authorities actually encourage a mass-based approach. In addition to the state plan, the state board also has adopted a Pollutant Policy Document for the San Francisco Bay Estuary (the PPD), which explicitly requires the regional board to implement a "mass emissions strategy" to control selenium, among other substances.158 The mass emissions strategy resembles the TMDL process, and it requires the regional board to "regulate mass emissions based upon an assessment of reductions in loadings for principal sources."159 While the mass emissions strategy does not explicitly address the issue of concentration versus mass-based permit limits, it clearly directs the regional board to focus on controlling total loading to a water segment.160 The most straightforward approach to reducing total loading is through limiting individual mass discharges rather than concentrations.

In fact, the regional board has expressed the current interim limits in the refineries' permits solely as mass emissions rates, while the final limits are expressed as both concentration limits and equivalent

with author). Also, EPA guidance suggests, but does not require, that permits specify both mass and concentration limits for waters with less than 100-fold dilution. TECHNICAL SUPPORT DOCUMENT, supra note 92, at 111. Note that 40 C.F.R. § 122.45(f)(ii) does require pollutants limited in permits to have limitations expressed in terms of mass, except when applicable standards and limitations are expressed in terms of other units of measurement. There is no similar requirement for concentration limits, however.

157. Since the permits already specify concentration limits, however, removal might violate anti-backsliding provisions. See infra note 190.


159. Id. at 4-2. The mass emissions strategy "includes the following major elements: 1. Identify pollutants and locations of concern. 2. Identify sources of pollutants. 3. Develop and implement a program to regulate mass emissions based upon an assessment of reductions in loadings for principal sources. 4. Develop tissue alert levels and sediment quality objectives." Id.

160. For example, in adopting regulatory strategies the regional board should consider total loads on the water body; significant sources of the loads; the estimated load reductions that can be achieved by alternative control measures; and the economic, social, and environmental consequences of the measures. Id. at 4-5 to 4-6. While the PPD is a duly adopted regulatory policy, its practical weight is somewhat uncertain. It was adopted as part of the ongoing process of resolving the disputes over Bay-Delta water allocations and water quality. Since the regulatory process is ongoing, interim responses may receive less regulatory attention at both the state and regional levels. See Telephone Interview with Dr. William S. Pease, Environmental Health Policy Program, School of Public Health, University of California at Berkeley, in Berkeley, Cal. (Jan. 19, 1993) (on file with author). But see Telephone Interview with John Norton, Chief Nonpoint Source Unit, Division of Water Quality, State Water Resources Control Board (Mar. 17, 1993) (on file with author) (stating that while there has not been an out-and-out push to implement the PPD, regional boards are in the process of submitting work plans for mass emissions strategies, despite a lack of funding, and this process is being tracked by the state).
CONTROLLING SELENIUM DISCHARGES

1994

mass emissions rates.\textsuperscript{161} Moreover, as long as control levels are set on the basis of mass emissions, regulators could retain relatively high concentration limits in the permits to protect against hot spots, without necessarily interfering with either a mass loading or permit trading approach.\textsuperscript{162} The current final permit limit of fifty $\mu$g/l,\textsuperscript{163} however, will definitely act as an independent constraint on the three refineries now discharging at higher concentrations. Since this numeric standard is set in each refinery's permit, the refineries could not use trading as a means of complying with that limit. If an overall mass-based limit were set, it would not matter which source emitted how much selenium. By contrast, a concentration-based limit in each permit means each source must meet an identical standard. The regional board also has adopted a daily average of fifty $\mu$g/l as the deep water effluent limitation for selenium in the basin plan, so a pure mass-based control approach would require changing both the basin plan and individual permit limits.\textsuperscript{164} Nevertheless, there does not appear to be any state statutory or regulatory authority that requires using concentration-based limits.\textsuperscript{165} On these grounds, the regional board staff has proposed to implement a mass emissions reduction strategy approach.\textsuperscript{166}

b. Discharge Permit Trading

Neither federal nor state water pollution laws and regulations explicitly mention permit trading or other market-based approaches. In the absence of any statutory ban, regulators should have the discretion

\begin{itemize}
  \item \textsuperscript{161} SFBRWQCB, Order No. 91-099 (June 19, 1991).
  \item \textsuperscript{162} Retaining high concentration limits would be one way of ensuring that most discharge permits were not traded to a single discharger, which might then create a hot spot, since that discharger could not emit above its concentration limit.
  \item \textsuperscript{163} See supra note 146 and accompanying text.
  \item \textsuperscript{164} Basin Plan, supra note 119, at Table IV-1B.
  \item \textsuperscript{165} See Telephone Interview with Gary J. Grimm, supra note 143 (leaning towards the position that concentration limits could be removed if effective mass-based limits were imposed).
  \item \textsuperscript{166} Staff Report, supra note 8, at 37-41. In January 1994, however, the regional board adopted a settlement agreement in the litigation brought by the refineries challenging the listing of the bay segments as water quality impaired. See supra note 143. Under the settlement agreement, the refineries currently in violation (Shell, Exxon, and Unocal) have until July 31, 1998 to meet their current final permit limits of 50 ppb, which were supposed to be met by December 1993. Id.; Stipulation for Settlement and Dismissal of Action at 2-3, Western States Petroleum Assoc. v. California Regional Water Quality Control Bd., No. 121078 (Cal. Super. Ct. 1994). The refineries are also to pay $2 million ($1,800,000 of this goes into a “Selenium Mitigation Fund”). Id. at 3. Fund monies will be used to (a) fund studies of the fate and transport of pollutants in the estuary, (b) fund expert consultants to help the board review the progress and products of WSPA’s technology study, and (c) fund projects for the protection and/or enhancement of habitat in the estuary affected by selenium. Id. at 3-4. Any schedule for selenium reductions in an MES adopted by the board must be consistent with the schedule in the settlement. Id. at 5; SFBRWQCB, Tentative Order, Cease and Desist Order (Dec. 15, 1993).
\end{itemize}
to institute a permit trading approach to water quality-based controls,\(^{167}\) as long as no legal requirements are violated.\(^{168}\) Moreover, the CWA may give state regulators more flexibility than it gives EPA to implement innovative control approaches, such as permit trading.\(^{169}\) States must meet the criteria established in the CWA, but are free to implement programs that are more stringent or have a greater scope of coverage.\(^{170}\) Thus, there are two main inquiries for applying permit trading in the water context: What are the existing statutory requirements for control programs that would constrain the implementation and design of a permit trading system? And, if these are prohibitive, what statutory changes would be necessary to develop a viable trading approach? California does not appear to have any requirements stricter than federal standards that would impose additional burdens on a trading program, so the focus of inquiry here is on the federal statutes and regulations. The most significant potential statutory and regulatory constraints on trading are the requirement that an NPDES permit specify effluent limits and the restrictions on changing those permit limits.\(^{171}\)

i. Requirement of Effluent Limitations in Permits

The CWA requires that NPDES permits specify effluent limitations, which can be changed only through an elaborate modification procedure.\(^{172}\) Although this constraint would not necessarily doom all types of trading systems, it undermines, at least, the allowance and group permit systems.

One can argue that a permit provision barring discharges in excess of pollution allowances held constitutes a quantity-based effluent limitation. The pollution allowances would simply translate WLA's into tradable permits. Casting a permit mandate to hold allowances equal to emissions as an effluent limitation, however, may be viewed as a relatively drastic departure from current approaches. Even if using allowances could withstand the inevitable legal challenge, regulators are unlikely to be willing to pursue such an innovative and risky

\(^{167}\) Permit trading could not be used to satisfy technology-based control requirements because use of the technology is mandated by statute. 33 U.S.C. § 1311 (1988); Industrial Economics, Inc., *supra* note 33, at 1-2.

\(^{168}\) Industrial Economics, Inc., *supra* note 33, at 1-2; Marty Rothfelder, *Reducing the Cost of Water Pollution Control Under the Clean Water Act*, 22 NAT. RESOURCES J. 407, 418 (1982). See id. at 415-16 for a discussion of cases where courts deferred to EPA's interpretation and implementation of the CWA.


\(^{170}\) Id. at 418 n.79; 40 C.F.R. § 123.1(i) (1992).

\(^{171}\) 33 U.S.C. § 1312(a)-(b).

\(^{172}\) See *supra* notes 154-55 and accompanying text; see *infra* notes 181-86, 190 and accompanying text.
approach based on slim statutory authority. Explicit legislative language authorizing the use of market mechanisms generally or permit trading in particular would facilitate use of a strongly market-based approach, such as allowances.

The NPDES permit system provisions make it difficult to use a group permit approach and may foreclose this approach altogether. While a group permit would retain the overall form of existing permits by specifying numeric discharge limits, it would apply to multiple sources rather than to a single source. The structure of permitting and enforcement under the CWA appears to contemplate individual permitting, although the Act does not explicitly require that each point source hold its own NPDES permit. A group permit approach has been used by the regional board in the past, but only in the context of nonpoint source control. Accordingly, it does not provide a fully analogous precedent.

Further statutory requirements regarding the form of NPDES permit limits could reduce flexibility in designing a trading program. Unless "infeasible," permit limits must be expressed as both average monthly limitations and daily maximum limitations. Daily or monthly limits may be incompatible with a trading system because they reduce flexibility in redistributing discharge quantities. Such limits, however, are not necessarily inconsistent with a trading system based on quarterly or yearly accounting. Generally, discharges to waters are regulated on the basis of periodic composite samples' exceedances of statistical boundaries, rather than on continuous emission monitoring. The daily or monthly statistical limits could

---

173. See, e.g., Telephone Interview with Laurie Kermish, supra note 125. Regulatory approaches must meet the two criteria of enforceability and defensibility (in terms of statutory authority). These are less likely to be met the further one gets from clear statutory language and established regulatory mechanisms. Id.

174. The CWA requires effluent limitations (including alternative effluent control strategies) to be imposed for a "point source or group of point sources" when discharges from a point source or group of point sources interfere with attainment of water quality standards. 33 U.S.C. § 1312(a) (1988). A "point source" is "any discernible, confined and discrete conveyance, including but not limited to any pipe, ditch, channel, tunnel, conduit, well, discrete fissure, [or] container... from which pollutants are or may be discharged." Id. § 1362(14). The permit section does not specify the form of entities receiving permits. Id. § 1342. The relevant regulation, 40 C.F.R. § 122.41 (1992), refers to "the permittee" only.

175. An enforceable group permit has been implemented by the regional board for nonpoint source control measures to be taken by municipalities. EPA did not disapprove the permit. See NPDES Permit I, supra note 87. Since the permit applies to municipalities, not industrial users, however, and targets nonpoint stormwater runoff from a multitude of small sources, rather than discrete discharges from industrial sources, the applicability of this example is questionable. See id.

176. 40 C.F.R. § 122.45(f).

177. See TECHNICAL SUPPORT DOCUMENT, supra note 92, at 95-96. A periodic composite sample adds discharges over a given time period (e.g., samples once an hour for 24
be set so high that they would not constrain trading as a practical matter; only the annual limits would have "bite" in most situations. This would reflect the real regulatory concern with annual mass loading, rather than daily discharges.\textsuperscript{178}

There are both advantages and disadvantages to using multiple permit limitations. Daily and monthly values would help to reduce the possibility of hot spots.\textsuperscript{179} Daily and monthly values would also allow regulators to prevent the rather unlikely possibility of large amounts of dumping in a short time span (e.g., an entire year's mass loading discharged in a few weeks). Nevertheless, multiple limits are bound to complicate and confuse the regulatory process. To the extent multiple limits add uncertainty regarding the sufficiency of allowances for compliance, they further discourage trading.\textsuperscript{180}

\textbf{ii. Restrictions on Changing Permit Limitations}

The other major potential legal constraints on trading are the restrictions on changing effluent limits in NPDES permits. If a trading system must depend upon changing NPDES permit limits to reflect transactions, as would a tradable limit approach, restrictions on changing limits could eliminate trading. The CWA does allow NPDES permit limits to be altered under certain circumstances. Permits may be

\begin{footnotesize}
\begin{itemize}
  \item See, e.g., SFBRWQCB, Order No. No. 89-002, NPDES No. CA0005053 (Jan. 18, 1989) (setting statistical daily and monthly limits for Union Oil Co. facility); SFBRWQCB, Order No. 92-111, NPDES No. CA0005134 (Sept. 18, 1992) (setting statistical daily and monthly limits for Chevron refinery); Telephone Interview with Tom Mumley, Staff, SFBRWQCB (Feb. 22, 1993) (on file with author). The composite value is limited based on a statistical analysis of the frequency and variability of the samples. \textit{TECHNICAL SUPPORT DOCUMENT, supra} note 92, at 95. Thus a limit will be set that, if met by 90\% of the composite samples, for example, will result in the maximum allowable total pollutant load. Continuous emissions monitoring, by contrast, measures ongoing total emissions.

  \item See Telephone Interview with Tom Mumley, \textit{supra} note 177.

  \item Alternatively, hot spots could be avoided by retaining acute toxicity-based concentration limits.

  \item See \textit{supra} notes 50-52 and accompanying text. Another potentially inhibitory "limit" issue is the inclusion of whole-effluent toxicity limitations in permits, although this is unlikely to affect selenium control. EPA has developed policies encouraging the use of biological monitoring and assessment methods, specifically whole-effluent toxicity, as an additional tool for addressing water quality problems from toxics. \textit{See} 40 C.F.R. §§ 122.2, 122.44(d) (1992). The whole-effluent approach uses acute and chronic toxicity tests to measure the combined impact of all the toxics in the wastestream on aquatic life, although most tests currently focus largely on acute effects. \textit{TECHNICAL SUPPORT DOCUMENT, supra} note 92, at 4; Interview with Dr. William S. Pease, \textit{supra} note 12. Whole-effluent toxicity permit standards might inhibit trading where a load transfer of one pollutant could cause violation of such a standard. This problem is less likely to arise where the focus of concern is bioaccumulation, not acute or chronic toxicity. \textit{Id.} For selenium, because the levels of concern are sufficiently below those causing acute effects, it is highly unlikely that a transfer would implicate any whole-effluent concerns. \textit{Id.}
\end{itemize}
\end{footnotesize}
altered upon renewal, may be modified, or may be revoked and reissued.\textsuperscript{181} Permits can be modified or revoked and reissued only if the regulator finds that a listed cause is present.\textsuperscript{182} The list of causes includes facility expansion,\textsuperscript{183} new information, new regulations changing the standard on which the permit was based,\textsuperscript{184} need to modify compliance schedules,\textsuperscript{185} and inability to meet technology-based effluent limitations despite proper installation and operation of the treatment technology.\textsuperscript{186} None of the listed causes encompasses increased cost-effectiveness of discharge reduction. An argument could be made that reopening the TMDL/WLA process is a sufficient change to allow modification, but this would be a fairly strained reading of the regulations.\textsuperscript{187} The one example of a point source trading program, Fox River, appears to have been based on these CWA provisions for permit modifications, as it incorporates restrictions that meet the regulatory constraints on permit modifications. Specifically, an applicant for a transfer must demonstrate that "the increase is needed due to: (a) New production by a new discharger, (b) Increased production which cannot be accommodated by the current treatment facility, or (c) The inability of the current waste treatment facility to meet current wasteload allocations despite optimal operation and maintenance of the treatment facility."\textsuperscript{188} Analysts have attributed the absence of any trading under this program in large part to the inhibiting effects of meeting these legal requirements.\textsuperscript{189}

The changes that can be made in the NPDES permit limitations are further constrained by anti-backsliding provisions, which prohibit "less stringent" water quality-based effluent limitations in a renewed or modified permit.\textsuperscript{190} It may be argued that removing or loosening concentration-based limits should not be considered "less stringent" where tightened mass-based limits reduce total discharges to the water segment. Similarly, an increase in one source's effluent limit balanced

\textsuperscript{181} 40 C.F.R. § 122.62 (1992).
\textsuperscript{182} Id.
\textsuperscript{183} Id. § 122.62(a)(1).
\textsuperscript{184} Id. § 122.62(a)(3).
\textsuperscript{185} Id. § 122.62(a)(4).
\textsuperscript{186} Id. § 122.62(a)(16).
\textsuperscript{187} Telephone Interview with Laurie Kermish, supra note 125 (noting the existence of the argument).
\textsuperscript{188} Wis. Admin. Code § NR 212.115(2) (1988). This last reason is not explicitly listed as a cause for modification in the CWA, although failure to attain technology-based standards despite proper installation, operation, and maintenance of the treatment technology is a listed cause. 40 C.F.R. § 122.62(a)(16).
\textsuperscript{189} See Industrial Economics, Inc., supra note 33, at 2-5.
\textsuperscript{190} 33 U.S.C. § 1342(o)(1) (1988). Limits can be changed if the "circumstances on which the previous permit was based have materially and substantially changed." 40 C.F.R. § 122.44(f). A cost-effective reduction reallocation would not necessarily be linked to any change of circumstances.
by a decrease elsewhere should not be deemed "less stringent," particularly when done under the auspices of a trading program designed to reduce significantly total mass loading to the water segment.

The CWA also allows reallocations of WLA's between sources only where the cumulative effect of all revised effluent limitations assures attainment of the water quality standard.191 Assured attainment, as opposed to simply not causing any net discharge increase, may be a prohibitively strict standard for trading in nonattainment waters, where trading is most likely to be instituted. This provision would effectively bar individual trades that reduced total loading but were insufficient in and of themselves to bring the segment into attainment. In contrast, California's PPD explicitly authorizes increased loading from one source if there is a corresponding equal or greater reduction from other sources.192

In addition, CWA procedures for changing permits impose substantial transaction costs and create a disincentive to trading. Permit issuances, reissuances, and modifications must comply with procedural requirements. These requirements include public notice, comment and hearing, and consideration of concerns of neighboring states.193 Changes also are subject to disapproval by EPA.194 The resulting paperwork, need to respond to objections, and delay all impose transaction costs on the parties.195 Short-term transfers are precluded alto-

191. 33 U.S.C. § 1313(d)(4)(A). This requirement applies to waters that have not attained the applicable water quality standard. Id. For waters that have attained the relevant standard, effluent limits can only be revised "subject to and consistent with the antidegradation policy established under" that section. Id. § 1313(d)(4)(B). The standard is attained where "the quality of such waters equals or exceeds levels necessary to protect the designated use for such waters or otherwise required by applicable water quality standards. . ." Id. For selenium, the bay's waters comply with the inadequate numeric criteria for ambient concentrations, but do not comply with the narrative criteria, and are not protecting the designated use. See supra notes 134-43 and accompanying text. Thus, the bay cannot be considered to have attained the standard.

192. Pollutant Policy Document, supra note 158, at 4-8. The PPD also states that where overall loading reductions are required for the water body, credited reductions from other sources must be "over and above any reductions" that are "necessary" under the program. Id. It gives as an example reductions from nonpoint sources, which presumably would be outside the regulated scope of the program. Id. It seems likely, however, that a point source that reduces below its allocated discharge quantity would be over and above necessary reductions. Id.

In addition to the CWA's anti-backsliding provisions for individual sources, the regulations require states to adopt antidegradation policies to maintain existing uses of a water body and the level of water quality necessary to protect those uses. 40 C.F.R. § 131.12. These policies should not limit trading in any way, however, since trading approaches are designed to maintain or improve overall water quality.

193. 33 U.S.C. § 1342(b) (permit issuances); 40 C.F.R. § 124.5(f).
194. 33 U.S.C. § 1342(9d); 40 C.F.R. § 123.44 (EPA review of and objections to state permits).
195. See Rothfelder, supra note 168, at 418.
The procedures also increase the administrative costs of implementing transfers that the local regulatory agency must bear. The possibility that EPA may veto a change heightens uncertainty, further raising transaction costs. An agency might be able to streamline these procedures in some manner, for example, by making findings that trades within certain parameters would have negligible environmental impacts. The statutory procedural requirements are fairly specific, however, leaving little room for agency adjustments.

In sum, the CWA should not prevent a shift to a mass emissions reduction strategy, and California laws tend to support such a regulatory focus. The further move to a discharge permit trading approach is more likely to be constrained or barred by current CWA provisions and regulations. The requirement that permits set an effluent limitation could constrain the types of possible trading approaches, while specifications for the form of the effluent limitation could decrease the efficacy of a trading program. The provisions governing whether and how permit limits can be modified could also potentially bar some trading approaches and may increase transaction costs. Finally, while anti-backsliding provisions arguably should be interpreted to support overall emissions reductions, regardless of individual increases, an alternate interpretation could hinder a trading program. Since the greatest constraints on trading approaches stem from the CWA and its regulations, much of this analysis of the legal context will apply to point source permit trading in other settings.

4. Economic Context—Controls, Costs, and Market Conditions

Instituting a permit trading approach to pollution control makes sense only where significant cost savings may be achieved through periodic reallocations of reduction responsibilities. The potential for such savings relates to the absolute size of the control costs and the differences in marginal control costs among dischargers. Differences in when control costs are incurred are another potential source of cost savings through trading. Whether potential savings can be realized

196. Id.
197. Id. at 418 n.83.
198. Id. at 418.
199. While 40 C.F.R. § 122.63 provides an exception for minor modifications, the specific categories listed do not encompass transfers of control responsibilities. Agency attitudes, however, may make a significant difference in the perceived degree of regulatory hassle accompanying a transaction. Given the potential for hostile regulators to use innumerable procedural hoops as barriers to trading, regulators supportive of trading might be a necessary condition for a viable trading program, particularly one based on existing legal authorities.
through permit trading depends on whether conditions in the problem setting are conducive to market development.

In the San Francisco Bay Estuary, selenium control costs are unquestionably substantial. Three broad control approaches have been identified: changing crude sources, changing production processes and products, and developing and implementing end-of-pipe technological controls to remove selenium from the wastestream for disposal as solid waste. The feasibility and results of each of these approaches are unproven and highly contentious.\(^{200}\) Cost estimates are correspondingly uncertain.

While the relative amount of selenium in the crude appears to have some effect on the quantities of selenium discharged, studies do not establish any direct correlation that would allow prediction of the magnitude of reductions that might be gained by switching crude sources.\(^{201}\) The regional board roughly estimates that crude switching could potentially reduce selenium discharges by a maximum of fifty percent; this assumes that all San Joaquin Valley (SJV) crude would be replaced by Alaska North Slope (ANS) crude.\(^{202}\) Actual reductions probably would be smaller than estimated, due to forthcoming production changes to comply with the 1995 clean fuels requirements.\(^{203}\)

The cost to a refinery of switching crudes would include the higher price of the substitute crude\(^{204}\) and the cost of plant and process changes that would be required to refine a substantially different type of crude.\(^{205}\) Although crude switching might nonetheless be cost effective for the refineries, additional costs should be taken into account and internalized, if possible, to any compliance decision. One cost to the environment would be the increased tanker traffic in and out of the San Francisco Bay (to import ANS crude and export SJV

\(^{200}\) See, e.g., Letter from Todd P. Royer, Supervisor, Exxon Co., U.S.A., to Steven R. Ritchie, Executive Officer, SFBRWQCB, Attachment I, Attachment II-2 (Nov. 16, 1992) [hereinafter Letter from Exxon to SFBRWQCB] (concluding that the iron co-precipitation process is the most promising control strategy, and casting doubt on a strategy of switching crude sources); Letter from Greg Karras, Director, Clean Bays and Coastal Waters Program, Citizens for a Better Environment, to Steven R. Ritchie, Executive Officer, SFBRWQCB 5 (Nov. 5, 1992) (estimating that adoption of process changes already used at the cleaner refineries by the others would cut more than 80% of present refinery discharges).

\(^{201}\) Letter from Exxon to SFBRWQCB, supra note 200, at Attachment II-2.

\(^{202}\) Staff Report Amendment, supra note 112, at 4.


\(^{204}\) ANS costs more to transport and commands a higher well-head price than the heavy, high nitrogen content SJV crude, although presumably there are also some cost savings from using the higher quality ANS crude. Letter from Exxon to SFBRWQCB, supra note 200, at Attachment II-1.

\(^{205}\) Hearing, supra note 113, at 21 (testimony of John Vautrain, Manager, Purvin & Gertz).
crude to other markets).\textsuperscript{206} Costs would also result from the disruption to SJV crude production.\textsuperscript{207} Currently, the six bay area refineries comprise a substantial share of the market for this crude.\textsuperscript{208} Limited transportation capacity and the higher cost of tanker transport would reduce both the absolute quantity of SJV crude that could be shifted to other markets in the short-term, and the demand over the long-term.\textsuperscript{209} Exxon estimates that crude switching would cost the state economy about $150 million in addition to the costs to the refineries themselves.\textsuperscript{210}

The option of changing production processes and products has been considered even less seriously than has crude switching. There is little information on the reduction potential or costs of this approach. Some environmentalists charge that the refineries are taking advantage of the clean fuels program to institute process changes that allow heavier (and less expensive) crude to be processed into lighter products, producing higher waste discharges.\textsuperscript{211} If the refineries produced heavier end-products (such as industrial fuel oils and diesel fuels), selenium discharges presumably would be reduced.\textsuperscript{212} The bulk of the market for at least some of the refineries, however, is for the lighter products such as gasoline and jet fuels,\textsuperscript{213} and the clean fuels program will require more, rather than less, processing.\textsuperscript{214} While process changes might be used to minimize the selenium increases expected to result from the installation of new processing capability, substantially changing existing production processes does not appear economically

\textsuperscript{206} Letter from Exxon to SFBRWQCB, supra note 200, at Attachment II-2.
\textsuperscript{207} Hearing, supra note 113, at 53 (testimony of Stowe Killingsworth, Director of Environmental Affairs, California Independent Petroleum Association).
\textsuperscript{208} One estimate is that the refineries together process a total of around 200,000 barrels/day of SJV crude. Letter from Exxon to SFBRWQCB, supra note 200, at Attachment II-1. The total production of SJV crude in 1991 was 673,000 barrels/day. Hearing, supra note 113, at 52 (testimony of Stowe Killingsworth). Thus, the bay area refineries probably comprise somewhat less than one-third of the current demand for SJV crude.
\textsuperscript{209} See Letter from Exxon to SFBRWQCB, supra note 200, at Attachment II-1. There is no other available pipeline capacity, so all of the SJV crude would have to be moved by tanker. Id. The refineries, however, also claim there is insufficient spare terminal capacity to transport significantly more SJV crude by tanker. Id. They assert that “the near term impact of reducing SJV processing in the Bay area would be to shut in California crude production.” Id.
\textsuperscript{210} Id. at Attachment II-2.
\textsuperscript{211} See Telephone Interview with Greg Karras, Director, Clean Bays and Coastal Waters Program, Citizens for a Better Environment (Mar. 18, 1993) (on file with author) (stating proposition). For example, the refineries are proposing to do more hydrotreating, which refines a heavier crude into lighter end-products. Id.
\textsuperscript{212} This may shift much of the problem to the air medium. Currently, selenium is removed with sulfur; if sulfur is not removed in the processing, it is released upon combustion, potentially contributing to urban air pollution and acid rain.
\textsuperscript{214} See 42 U.S.C. § 7545(k).
viable. This approach probably could be used only for incremental control of selenium discharges. Cost estimates are impossible without identifying the process and product changes and resulting reductions more specifically. Yet given the long lead time that the refineries are predicting for developing an end-of-pipe control technology (mid-1997 at the earliest), if the regional board were to enforce current limits sooner, this alternative might appear more feasible.

The refineries view end-of-pipe technological controls as the most likely route for reducing selenium discharges. The only viable existing technology for controlling selenium in refinery wastewater streams is the wastewater treatment pond, which has substantial feasibility, cost, and environmental drawbacks. It might become possible to remove some selenium through an iron co-precipitation process that is currently under development. At this point, however, the process is expensive, unproven on a large-scale, may not even meet the fifty μg/l final limit in the current permits, and generates large quantities of hazardous wastes. Examining this process as the best potential control technology gives an indication of the possible scale of control costs. These include capital costs, operating costs, and hazardous waste disposal costs, which constitute a large portion of the total control costs. Several of the refineries have estimated their operating and waste disposal costs to be millions of dollars annually, while they

215. Changing processes and products usually will require major new capital investment. Also, there may be insufficient demand for heavier products.

216. See Letter from Scott Folwarkow, Bay Area Coordinator, Western States Petroleum Association, to Steven Ritchie, Executive Officer, SFBRWQCB, Attachment 2 (Nov. 13, 1992) [hereinafter Letter from WSPA to SFBRWQCB].

217. See Letter from Richard F. Hallford, Corporate Director, Environmental Affairs, Tosco Refining Co., to Steven Ritchie, Executive Officer, SFBRWQCB 1-4 (Nov. 16, 1992) [hereinafter Letter from Tosco to SFBRWQCB] (on file with author). The ponds require a large quantity of land, which either is unavailable or prohibitively expensive at most of the refineries. Existing ponds are the subject of concern and investigations over air emissions and possible ground water contamination. Since the selenium removal mechanism is not well understood, there is some uncertainty regarding the achievable levels of control. Telephone Interview with John Staton, Manager of Wastewater and Hazardous Waste, Tosco Refining Co. (May 6, 1994) (on file with author). Tosco has rough estimates of the amount of selenium removed by their ponds, but is currently conducting a study to determine where the removed selenium actually goes. Id. It is also uncertain how well ponds would act to reduce selenium in other refineries' wastestreams. Id. However, existing ponds do provide an effective, lower cost treatment method. See Letter from Tosco to SFBRWQCB, supra. Tosco and, to some extent, Chevron, currently operate treatment ponds. Hearing, supra note 113, at 73 (testimony of Dan Glaze, Senior Environmental Engineer, Shell Oil Co.).

218. Hearing, supra note 113, at 53 (testimony of Charles Meyor, Senior Research Engineer, Shell Development Co.); see also Telephone Interview with John Staton, Manager of Wastewater and Hazardous Waste, Tosco Refining Co. (Mar. 17, 1993) (on file with author) (on the difficulty of meeting the 50 μg/l limit); Letter from Exxon to SFBRWQCB, supra note 200, at Attachment 1-1, 1-3, 1-4.
predict capital costs to be at least $10 million per refinery. Even in comparison to overall environmental compliance budgets of $20 to $100 million dollars per year, selenium control costs could be substantial.

Marginal cost per pound of selenium removed will vary across the companies. The refineries currently emit selenium at widely different mass quantities, concentration levels, and rates per barrels of crude input. Even if one technological control method were implemented by all of the refineries, it would have different costs and removal efficiencies when operating on different waste streams. These differences could be significant. Moreover, if a process could be developed to operate at high enough removal rates at some refineries, smaller emitters might be able to avoid installing the process altogether by purchasing permits.

Requiring incremental reductions over time also could create an opportunity for cost savings. If the regional board imposed a ratcheted reduction schedule, some refineries facing greater development problems might be able to buy time through permit purchases from refineries that achieved high reduction levels earlier. Ratcheted reductions provided most of the cost savings in the lead trading pro-

219. E.g., Letter from Exxon to SFBRWQCB, supra note 200, at Attachment I-3, I-4 (estimating $4,200,000 in annual operating cost, including $2,159,000 in annual waste disposal, with total selenium removal costs of $4000-$17,000/lb of selenium); Letter from Rand H. Swenson, Chief Refinery Engineer, Unocal, to Steven R. Ritchie, Executive Officer, SFBRWQCB 5 (Nov. 23, 1992) (on file with author) (estimating $760,000-$15,000,000 annual operating and maintenance cost); Letter from J.C. Harmon, Manager, Environmental Conservation, Martinez Manufacturing Complex, Shell Oil Co., to Steve Ritchie, Executive Officer, SFBRWQCB 5 (Nov. 13, 1992) [hereinafter Letter from Shell to SFBRWQCB] (on file with author) (estimating $4,560,000 in annual operating and disposal cost; $3,260,000 for disposal and $1,300,000 for operations). Letter from Tosco to SFBRWQCB, supra note 217, at 4 (estimating $10,960,000-$13,010,000 in capital costs and $1900-$2200/lb for selenium removal).

220. See Letter from Shell to SFBRWQCB, supra note 219, at 8 ($130 million); Letter from Exxon to SFBRWQCB, supra note 200, at Attachment I-4 ($20 million).

221. Staff Report, supra note 8, at 37.

222. See id. at 38. The refineries with the lowest emissions are those with wastewater treatment ponds. This may not have much impact for trading purposes unless allocations for these refineries are increased from current final permit levels, because the ponds now may be operating at maximum treatment levels. See Hearing, supra note 113, at 53 (testimony of Max Boone, Senior Environmental Engineer, Tosco Refinery Co.).

223. Telephone Interview with Warren Smith, Superintendent of Environmental Affairs, Unocal (Mar. 15, 1993) (on file with author); Telephone Interview with Dan Glaze, Senior Environmental Engineer, Shell Oil Co. (Mar. 15, 1993) (on file with author). The same technology could operate differently if the refineries use different processes and emit different quantities of various species of selenium, as seems likely. Id.

224. See Telephone Interview with Dan Glaze, supra note 223. On the other hand, treatment costs tend to escalate substantially at higher removal efficiencies, so there might not be any cost savings even if this were technologically feasible.
gram.\textsuperscript{225} They appear to be an attractive option here too, if any of the refineries could cost-effectively reduce early, producing excess permits available for trading.\textsuperscript{226}

Trading presents large potential cost savings in the bay, for control costs are high and the refineries will likely experience substantially different marginal removal costs. Realizing these potential cost savings, however, depends upon the availability of excess reductions; at least some firms must be operating at less than their maximum removal rates. The refineries fear that effluent limits are likely to be set on the basis of the maximum removal that can be achieved by the best control technology developed.\textsuperscript{227} In that case, all the refineries would be forced to install that technology and operate it at, or close to, maximum removal capacity, thereby generating very few, if any, excess reductions.\textsuperscript{228} Even if maximum removal were eventually required, however, a phased reduction approach would provide an opportunity for trading.

Realizing potential cost savings also depends upon market viability, which is partially related to the number of players in and geographic scope of the market. Here, the market is likely to consist of only six oil refineries. Having so few players raises the problem of noncompetitive market pricing, for a certain number of transactions are necessary to generate a strong price signal. A problem of market power may also arise, if only one or two refineries generate all the available excess permits. These potential infirmities will be exacerbated if the participants have relatively few permits to trade, as seems likely here. Potential additional players are municipal sewage treatment plants in the North Bay and nonpoint sources in the Sacramento and San Joaquin watersheds, but neither group is very promising.\textsuperscript{229}

The nature of the market also will determine the transaction costs that exist independently of constraints imposed by regulators. One source of transaction costs is the information cost of identifying and implementing trades. Such information costs should be very low in this market. Since there is a small number of players, who are all known to each other, there would be no need for outside brokers to

\begin{footnotes}
\item 226. See Telephone Interview with John Staton, supra note 218 (questioning whether such excess permits would be available).
\item 227. Telephone Interview with Dan Glaze, supra note 223.
\item 228. There would be excess reductions only if at some facilities the technology achieved much better removal rates, which were not taken into account in setting effluent limitations.
\item 229. See infra notes 314-22 and accompanying text. While municipal sewage treatment plants in the South Bay are a substantial source of selenium, trading between the North and South Bays cannot be justified on environmental grounds, because impacts are sitespecific within those zones.
\end{footnotes}
match buyers and sellers. The fact that the players are competitors in a single industry could inhibit trading, however, especially in a non-competitive market. If a company believes that the benefit of imposing higher compliance costs on its competitors outweighs the foregone profits of selling allowances, efficient trades probably will not occur and the associated cost savings will be lost.

In sum, the economic features of a pollution control problem will determine whether there is any reason to develop a market-based control program. Here, the magnitude of total control costs and the differences in marginal control costs faced by the refineries indicate that trading has significant cost saving potential, but the necessary economic and technological conditions for developing a market and realizing the cost savings may not exist. While the size of the market might pose some problems in achieving competitive pricing, the most significant barrier to market functioning is likely to be a lack of excess permits available for trading. Whether permit unavailability necessarily would undermine a trading approach is currently unknown, however, and depends on several factors, some of which are not amenable to regulators' influence. The availability of permits could depend on the level and timing of mandated reductions, and on the development and functioning of control technology or other control approaches.

5. Institutional Context—Capacities for Developing, Implementing, and Carrying Out a Trading Program

The institutional context of a pollution problem will significantly affect whether a permit trading solution is developed and adopted, how it is implemented, and whether it achieves its goals. Institutional players include the regulatory bodies relevant to the problem and the regulated industries. An institution's capabilities are affected by its organization, resources, and responsibilities, as well as its knowledge and experience in the environmental problem area.

The regulatory structure here is comprised of three agencies with overlapping authorities, which complicates the development and approval process for any innovative approach. Any willingness the

230. The small number of participants has been identified as one of the problems contributing to the failure of the Fox River program, where, as here, there were only about six or seven firms in each cluster that could trade with each other. Robert W. Hahn, *Economic Prescriptions for Environmental Problems: How the Patient Followed the Doctor's Orders*, J. ECON. PERSP., Spring 1989, at 95, 98.

231. *See supra* part II.A.2. The internal structure of individual agencies could have an influence, although this does not appear to be the case here. For example, a mass-based program that regulated total toxics loading to the bay, regardless of source, would be incompatible with the current programmatic divisions organizing the state board's activities. In the absence of regulatory reorganization, reductions in such a Bay loading program
The regional board might have to "push the boundaries" in amending the basin plan is probably markedly reduced by the possibility of a veto (or the lack of an active approval) by either the state board or the regional EPA. The possibility of EPA disapproval similarly may inhibit actions by the state board. Even the regional EPA is subject to an override by the EPA Headquarters, although this is unlikely to happen in a local permitting action. Finally, the decisions of all of these actors can be overruled by the courts for nonconformance with statutory or regulatory provisions. The multiplicity of checkpoints inhibits regulatory risk-taking and provides ample opportunity for those opposed to a regulatory action to derail the program. A regulatory setting with multiple checkpoints encourages regulators to gather consensus from all the major players before embarking on a new initiative. Perceived "win-win" programs that can appeal to all affected interests may thus receive regulatory support despite their innovative character. Proponents often present maximizing environmental dollars through trading as such a win-win proposal.

Regulatory resources and responsibilities further affect development, approval, and implementation of a permit trading program. Developing a new regulatory program is highly time- and resource-intensive, particularly when there is little precedent on which to model an approach. In other settings, environmental groups and industry were instrumental in working out program details and gathering political support. Here, however, no nongovernmental actor has a sufficient stake in promoting a trading approach to commit significant resources. Also, while the state board staff is interested in the development of a mass emissions approach, they would rely on the regional board to work out the details. Thus, a trading initiative for the bay selenium problem is most likely to come, if anywhere, from the regional board staff.

would have to be divided on source-specific grounds in order for the regulatory programs to assess the results that are the responsibility of each program. Telephone Interview with John Norton, supra note 160. This issue could arise in a point-nonpoint selenium source trading program.

232. While the regional board can issue permits with slightly less threat of interference, instituting a permit trading scheme would probably require amending the basin plan. Even permit issuances are subject to EPA veto, and may be overturned by the state board on appeal.


234. For example, EDF was very involved in the development of the CAA acid rain program and worked on the RECLAIM initiative. See supra note 57 and accompanying text.

235. Telephone Interview with John Norton, supra note 160.
The regional board currently has devoted limited staff resources to the selenium issue.\textsuperscript{236} It is unclear whether the problem is an overall lack of resources relative to responsibilities, a matter of priorities, or both. The present level of staffing probably is insufficient to build on the proposed mass emissions reduction strategy to create a new regulatory approach such as trading. Yet, the quality of the regional board's human resources in terms of education,\textsuperscript{237} experience, and knowledge is quite high.\textsuperscript{238} This seems an important precondition for successful development and implementation of innovative approaches.\textsuperscript{239} The staff's mass emissions reduction strategy proposals for both selenium and copper are being watched closely by the state board as the first effort to move to this new regulatory approach.\textsuperscript{240} Nevertheless, with a few notable exceptions, it does not appear that many of the regulators involved have thought seriously about a market trading approach. While understandable, since it is not clear that such a strategy is consistent with the current CWA, this lack of regulatory support diminishes the likelihood of a market approach being developed, adopted, or successfully implemented.

The institutional structure of the regulated industry is primarily relevant for assessing an industry's response to a permit trading program proposal and implementation. The locus of environmental compliance decisionmaking is an important factor influencing a facility's willingness to endorse this relatively unknown control approach. Upper level financial managers could generally be expected to support trading as a cost-effective compliance option, but these managers are rarely involved in compliance decisions. Instead, compliance decisions are made by environmental engineers and managers, who are most familiar with technological solutions, and who tend to value certainty of compliance over potential cost savings. As a whole, more aggressive risk-taking industries are expected to support market in-

\textsuperscript{236} Only one part-time staff member works solely on selenium. Interview with Kim Taylor, supra note 113.
\textsuperscript{237} Almost all of the staff members have graduate degrees. Susan Pantell, Lecture at the Energy Resources Group Colloquium (Mar. 10, 1993).
\textsuperscript{238} Susan Pantell characterizes the regional board's staff as highly motivated, in comparison to the Los Angeles Regional Water Quality Control Board staff. Id.
\textsuperscript{239} Meidinger has identified the presence of regulatory entrepreneurs in EPA's Office of Policy, Planning and Evaluation (OPPE) as an important force driving EPA's initial adoption of permit trading approaches in the mid- to late 1970's. Errol Meidinger, Discussion: The Politics of 'Market Mechanisms' in US Air Pollution Regulation: Social Structure and Regulatory Culture, in DISTRIBUTIONAL CONFLICTS IN ENVIRONMENTAL-RESOURCE POLICY 150, 159, 164 (Allen Schnaiberg et al. eds., 1986).
\textsuperscript{240} Telephone Interview with John Norton, supra note 160.
centive approaches more enthusiastically than relatively cautious industries.\textsuperscript{241}

Here, the environmental engineers and managers dealing with the selenium issue are aware and supportive of market incentive approaches in general, although dubious of their usefulness to address the selenium issue.\textsuperscript{242} Some engineers and managers are also familiar with specific permit trading concepts from the air context. Selenium compliance decisions generally are made at the refinery manager level, although the environmental managers would work with the regional board on issues such as specific allocations.\textsuperscript{243} This structure appears to allow substantial flexibility in decisionmaking, relatively unburdened by bureaucratic constraints. Thus, the refineries do not appear to have any institutional barriers to realizing the cost savings offered by a permit trading approach. They possess sufficient institutional knowledge and are unhampered by the structure of compliance decision processes. The refineries' failure to push for such an approach is probably better explained on economic and political grounds, rather than by lack of institutional capabilities.

In sum, as with the economic context, analysis to determine the appropriateness of permit trading in a given institutional setting will be largely setting-specific. However, the national, state, and regional or local regulatory structures and interactions will have at least some, and perhaps many, common elements across different problem contexts. The existence of multiple interacting regulatory bodies itself may be the single most important institutional factor affecting the viability of permit trading proposals. Regional or local entities may be freer to experiment and to craft solutions appropriate for setting-specific problems. Also, while the existence of multiple checkpoints often inhibits regulatory innovation, it may promote such consensus driven approaches as permit trading.

6. Political Context—Alignments and Motivations

The political context of an environmental problem will affect both whether a trading program is approved and what form the program takes, since political bargains necessary for agreement will dictate some program design choices. Important aspects of the political context include the identities of key players, the motivations for their

\textsuperscript{241} The characteristic caution (and possibly backwardness) of the highly regulated utility industry has been cited as a reason for utilities' reluctance to embrace potentially profitable trading opportunities under the acid rain program. Matthew L. Wald, Risk-Shy Utilities Avoid Trading Emission Credits, N.Y.\textsuperscript{\textsc{times}}, Jan. 25, 1993, at C2.

\textsuperscript{242} See Telephone Interview with Dan Glaze, supra note 223; Telephone Interview with Warren Smith, supra note 223; Telephone Interview with John Staton, supra note 218.

\textsuperscript{243} Telephone Interview with Dan Glaze, supra note 223; Telephone Interview with John Staton, supra note 218.
positions on trading, the intensity of their reactions, and their influence in the policy process. Together these elements determine the specific objectives and objections of each group that would need to be addressed in designing a policy response to the selenium problem.

Regulators' attitudes towards permit trading vary from agency to agency. The regional board staff is probably the most supportive of the trading concept. The regional board is faced with the prospect of trying to obtain substantial and costly discharge reductions, largely on the basis of anticipated future environmental impacts. As the group directly responsible for working out a solution acceptable to the affected interests, or else taking the political heat, the board wants to find ways to reduce the economic burden on the refineries while still achieving the environmental goals. Yet, since the regional board would implement and enforce a trading program, it might be expected that the staff also would be concerned about the administrative burden of this potentially more complex system. In discussions with the author, however, staff members did not express this as a major source of their concerns regarding market approaches.244

The state board and its staff support a mass emission approach and are looking to selenium regulation as a first attempt to implement it.245 Trading, however, is a further step that the state regulators have not seriously considered, and may not be all that familiar with.246 State regulators could oppose trading on the basis that adding complexity and innovation might jeopardize a mass approach. Alternatively, trading might appear to be a way to make mass approaches more workable and attractive. Generally, the state board finds market incentive approaches conceptually attractive as a means of "getting a bigger bang for the buck,"247 but it may have concerns regarding the legality and workability of specific measures, such as allowance-type trading.248

EPA is also interested in market incentive approaches,249 but is under greater pressure to fit any such approaches into the existing regulatory system. An incremental shift towards trading would probably be far more acceptable to EPA than a full-fledged allowance mar-

---

244. See, e.g., Interview with Kim Taylor, supra note 113; Telephone Interview with John Norton, supra note 160.
245. Telephone Interview with John Norton, supra note 160.
246. See id.
247. Id.
248. Id.
ket, even given statutory authorization.\textsuperscript{250} Assuming EPA took this position, a contributing motivation might be that EPA is somewhat better insulated from industry pressure to minimize costs than is the regional board. EPA's primary focus is to achieve environmental goals in a readily enforceable system, so to obtain EPA approval a trading system would have to ensure that these goals would be met.\textsuperscript{251}

The refineries' official position is that there is insufficient data to support listing the North Bay segments as water quality impaired.\textsuperscript{252} Nevertheless, most of the personnel involved recognize that some kind of control will be necessary (whether for ecological or solely regulatory purposes).\textsuperscript{253} The main areas of contention are the timing of reductions and the total price. If the regional board pushes non-technology-based controls, there will be a dispute over methods as well. While the refineries oppose changing crude sources or processes on the grounds of ineffectiveness and cost, regulatory interference with basic production decisions would seem to be an additional strong reason for industry opposition. The refineries strongly object to the regional board's current mass emissions reduction strategy proposal as setting impossible deadlines and excessive reductions.\textsuperscript{254} Their opposition appears to carry considerable weight with the regional board.\textsuperscript{255} Externally, the refineries have presented a united front of opposition, and they are jointly represented by the Western States Petroleum Association (WSPA).\textsuperscript{256} Nevertheless, the refineries can be roughly categorized into two groups with potential for different political positions.

\textsuperscript{250} In the acid rain context, EPA regulators were more comfortable with a higher degree of regulatory control over the market than the degree of control advocated by those concerned about market viability. (This and several other observations about issues in permit trading are derived from the author's experience as a Policy Analyst for Acid Rain, at AER*X, in the summer of 1990.) Here, assuming it would be legal within CWA antibacksliding provisions, one EPA official expressed support for the concept of periodically reopening all WLA's in a TMDL to allow cost-effective reallocations. Telephone Interview with David Smith, \textit{supra} note 122.

\textsuperscript{251} It is important to compare a trading approach to the alternative, a command-and-control type of regulatory approach, so as to make sure that trading is not being held to a higher standard of certainty of environmental outcomes than is achieved under the existing system.

\textsuperscript{252} Letter from WSPA to SFBRWQCB, \textit{supra} note 216, at 1. This listing is driving the selenium control effort.

\textsuperscript{253} Interview with anonymous refinery employee (on file with author).

\textsuperscript{254} \textit{See}, e.g., \textit{Hearing}, \textit{supra} note 113 (testimony of various refinery representatives).

\textsuperscript{255} During most of 1993 the regional board was largely unable to convene a quorum on the selenium issue, due to board members recusing themselves on conflict of interest grounds and a tendency to place selenium at the end of the agenda where it repeatedly was postponed to the next meeting.

\textsuperscript{256} WSPA has to take relatively cautious, least common denominator positions that are mutually acceptable to all the refineries. A major joint research effort on selenium control technology is being conducted under WSPA auspices, although individual refineries also are pursuing independent research efforts. Telephone Interview with Dan Glaze, \textit{supra} note 223.
Three refineries (Shell, Union, and Exxon) are currently unable to meet the final permit limits, either in terms of mass or concentration, while the other three (Chevron, Tosco, and Pacific) are already complying with the limits. The noncomplying refineries must be more immediately concerned by the board’s proposals, and they are devoting more resources to the problem, at least in research.

Despite their opposition to the current mass emissions reduction strategy, in the abstract the refineries tend to prefer mass-based over concentration-based regulation, because the former provides greater flexibility for controlling discharges. Refinery personnel also generally support the concept of market approaches, because such approaches have the potential to provide more flexibility and lower cost controls. Refinery personnel do not expect permit trading to provide significant savings here, however, principally because proposed control levels appear too stringent to produce any excess tradable reductions. Thus, while they would not oppose permit trading on principle, the refineries also will not push for a trading proposal.

Another plausible reason for the refineries’ disinterest is that opposing all proposals may buy time, which may be valued more highly than the opportunity to promote the most cost-effective regulatory approach. If trading is presented as part of an overall reduction package, the refineries almost certainly will fight the entire proposal. If reduction goals are approved with a believable threat of enforcement, however, the refineries might not oppose a separate implementation proposal based on trading.

Some environmental groups may fight trading proposals, although one, at least, is enthusiastic about trading approaches.

---

258. Id.
259. For example, Shell has an independent research effort beyond the WSPA joint approach. Telephone Interview with Dan Glaze, supra note 223.
260. See, e.g., Id.; Telephone Interview with John Staton, supra note 218.
261. See, e.g., Telephone interview with Dan Glaze, supra note 223; Telephone Interview with John Staton, supra note 218.
262. See supra notes 227-28 and accompanying text.
263. Telephone Interview with Dan Glaze, supra note 223; Telephone Interview with Todd Royer, Supervisor, Environmental Section, Exxon (May 25, 1993) (on file with author); Telephone Interview with Warren Smith, supra note 223; Telephone Interview with John Staton, supra note 218.
264. See supra note 57. Another interest group aligned with the environmentalists on this issue is the Northern California and Northern Nevada Pipe Trades District Council No. 5. The group’s members recognize that long-term sustainable growth in the bay area depends upon a healthy environment, and they view the San Francisco Bay as a particularly valuable resource. This group does not currently have a firm position on trading, but strongly advocates immediate reductions wherever possible. It has expressed some skepticism concerning the viability of trading in the San Francisco Bay Estuary. Letter from Marc Joseph, Attorney, Adams & Broadwell, to Steven R. Ritchie, Executive Officer,
Environmental opposition rests on both practical and moral grounds. One major objection is: "How do you do pollution trading and reduce as soon as possible?" This argument is based on the premise that every reduction made is legally required, and so there are no extra reductions to trade. The CWA requires water quality-based toxics standards to be met as soon as possible, and at most within three years. Yet, the final permit limitations are probably inadequate to meet the water quality standards. Assuming the three noncomplying refineries fail to meet the December 1993 deadline, any schedule that would allow them to come into compliance gradually would not comply with the "as soon as possible" standard, and thus none of those three would have any legitimately excess emissions available to trade. Moreover, any reductions below current permit limitations made by the complying refineries should not be considered excess because the permit limitations are too high, given that the water body is in nonattainment.

Environmentalists also oppose trading because of its expected effect on the pollution control regulatory process. Introducing a new, complicated, and contentious proposal will tend to consume regulatory resources that could be better spent enforcing current standards. Additionally, trading may "set up different control strategies for different pollutants, multiply[] regulatory complexities and hinder[] public interest enforcement campaigns." Environmentalists are also edgy about cheating, which some fear might be easier under a "freer" system. Since there are only six highly visible refineries discharging selenium, environmentalists acknowledge it might be possible to manage and monitor a trading system adequately in this context. Yet this creates what they see as another potential problem: the successful application of trading here might be used to sup-

SFRBWQCB 4-5 (Nov. 6, 1992) (on file with author); Telephone Interview with Marc Joseph, Attorney, Adams & Broadwell, (May 27, 1993) (on file with author).
265. Telephone Interview with Trish Mulvey, supra note 62; Telephone Interview with Greg Karras, supra note 211.
266. See supra note 141 and accompanying text.
267. Reducing selenium discharges to the levels specified in the final permit limits is not expected to result in the mass levels necessary to protect beneficial uses. See Staff Report, supra note 8, at 37-40; see also supra notes 98-110 and accompanying text.
268. One response to this argument is that regulators should establish a reduction schedule for selenium control based on their best assessment of what is possible. Given such a schedule, a trading system can be viewed as a challenge to the regulated parties to do even better and benefit from their efforts. Telephone Interview with Terry Young, supra note 115.
269. Telephone Interview with Greg Karras, supra note 211.
270. Interview with Dr. William S. Pease, supra note 12 (noting legitimate environmental objections).
271. Telephone Interview with Trish Mulvey, supra note 62.
272. Id.; see also Telephone Interview with Greg Karras, supra note 211.
port trading in other places, where it could be environmentally harmful.273

The issue of secondary environmental impacts raises another concern. Even if a trade poses no risk to the aquatic environment, redistribution of pollution control efforts might conceivably create human health hot spots in the communities near the refineries.274 Such an increase in impacts on certain communities would raise distributional and environmental justice issues. While human health hot spots appear highly unlikely here,275 the objection might be plausible in other contexts, such as the RECLAIM program. Finally, some environmentalists flatly oppose trading toxics on moral grounds, arguing that it is "marketable death."276

To the extent that environmentalists' objections to trading rest on practical concerns, conditions necessary to secure their acceptance probably include strict monitoring and reporting requirements, adequate regulatory resources dedicated to tracking and enforcement, and public access to discharge and trading records.277 Provisions for a tightly controlled periodic reallocation based on the CWA provisions for modifying WLA's probably would generate little objection.278

The Environmental Defense Fund (EDF) is probably the only environmental group that enthusiastically endorses using a trading approach to the selenium problem.279 EDF cites the potential for cost savings, flexibility, and incentives for pollution control innovation as potential benefits from a trading program.280 EDF views effective monitoring and enforcement as critical, but otherwise supports trading with few restrictions, in order to promote market viability.281 EDF might be the only non-refinery player actively to advocate reducing transaction costs by such means as eliminating regulatory approval of individual trades.282 To some extent, EDF's support counterbalances other environmentalists' opposition. Given EDF's position, trading

--

273. Telephone Interview with Greg Karras, supra note 211.
274. According to Greg Karras, concentrating selenium reductions at a few refineries could increase accident risks from toxic gases and hazardous waste transport in the nearby communities, as H2S levels rise with higher removal rates for sulfur and selenium. Id.
275. Interview with Dr. William S. Pease, supra note 12 (disputing human health hot spot risk of trading here).
276. Telephone Interview with Greg Karras, supra note 211 (explaining other environmentalists' opposition on this basis).
277. Telephone Interview with Trish Mulvey, supra note 62.
278. Telephone Interview with Greg Karras, supra note 211.
279. See supra note 57.
280. Letter from Terry Young, Senior Consulting Scientist, Environmental Defense Fund, to Steven R. Ritchie, Executive Officer, SFBRWQCB 5 (Nov. 18, 1992) [hereinafter EDF Comments] (on file with author).
281. Id.
282. Id. This would require that trades be conducted within a framework of rules designed to prevent detrimental environmental effects.
proposals cannot be portrayed as intrinsically anti-environmental, although trading can still be attacked as a wrong approach to environmental goals. EDF also might be able to provide some technical support in working out the details of the trading program. Nevertheless, as a single environmental group which is not even the most active of the groups working on refinery selenium discharges, EDF probably carries fairly limited political weight with the regional board.

Other groups involved in the selenium debate include the California crude oil industry (represented by the California Independent Petroleum Association), the California Energy Commission in the State Resources Agency, and North Bay cities with publicly owned treatment works. None of these interests has expressed a position on trading approaches in their comment letters or testimony before the regional board. The crude oil industry and the Energy Commission generally oppose any regulatory approach that would encourage crude source management. These interests would probably oppose trading if linked to a strategy of early reduction, rather than to delays in regulation until researchers develop an end-of-pipe control technology. The cities would like to avoid making selenium reductions, and thus prefer a strategy focused on the refineries. They would probably oppose any broad-based trading scheme that included municipalities, except on a purely voluntary basis. Yet, to the extent that trading makes it more politically palatable to impose reductions on the refineries, the cities might support the approach. In any case, none of these interests is likely to play a major role in the trading debate.

Some of the major political characteristics exhibited here are likely to crop up in other problem areas as well. The strongest source of political support for trading often will be a regional or local regulatory agency that is searching for a way to reduce the political heat of costly regulations. Whether industry advocates a trading approach will depend largely on economic factors specific to the setting, although industry tends to support trading concepts on principle. Environmental groups are split, with some generally opposing trading, some generally supporting trading, and others deciding on a problem-by-problem basis.

283. The Energy Commission expressed this more delicately as a hope that the regional board will consider the impacts on California petroleum supplies. Letter from B.B. Blevins, Executive Director, California Energy Commission, to Steven R. Ritchie, SFBRWQCB 2 (Nov. 18, 1992).

284. Cities could be included as “opt-ins” to an allowance system (similar to smelters under the acid rain provisions). See 42 U.S.C. § 7651i (Supp. IV 1992). This is unlikely, however, because the POTW’s probably face higher control costs than the refineries. Interview with Kim Taylor, supra note 113.
B. Design Options: Examination of Factors Relevant to Permit Trading Approaches

The viability of a trading approach will certainly depend on whether the characteristics of the problem context make it amenable to trading, as discussed above. Yet, success also will depend on the specifics of the trading program attempted. Trading proponents must choose which kind of trading system to use and select from a myriad of design options in adapting that system to a given problem setting. These decisions cannot be made in the abstract—there is no perfect trading program for all environmental problems or even for all point source water pollution problems. Therefore, it is critical to consider the design options in relation to the specifics of the given problem setting. This decisionmaking process is illustrated by considering the implications of different trading system design choices for regulating the San Francisco Bay Estuary selenium problem.

Many design options can be discussed independently of the specific type of trading system proposed. The first subpart discusses such general design options in detail. Other design options, however, are specific to the different types of trading systems, such as an allowance system or a tradable limits system. The second subpart explores the implications of applying each of the trading systems to the case-study problem context and considers design choices specific to each trading system. This information is used subsequently to clarify some of the tradeoffs between using different trading systems in the selenium problem context.

1. Design Options Common to All Permit Trading Systems

A first category of choices regards the regulatory structure of the trading system. Allowances, group permits, and tradable limits are three possible kinds of systems. Additional structural options include the form of permit limitations, the method of integrating the new system into the existing regulatory structure, and feedback mechanisms allowing the system to respond to new information or changed circumstances. Structural choices will significantly affect the role of regulators and the administrative burden of the system, although these concerns also will be affected by choices about almost every design option. Secondly, choices must be made about permit allocation, including how rights to discharge pollutants should be divided between interested parties, and what method of distributing those rights should be used. Allocation issues are often key to the perceived fairness and political acceptability of the system. A third category involves choices governing the form and operation of the market. The
main issues are who can trade, and under what conditions (e.g., timing, geographic extent, and degree of regulatory oversight). These decisions will strongly affect transaction costs, the viability of the market, and the opportunity to achieve cost savings. Finally, choices must be made concerning monitoring and enforcement, to ensure that regulated entities comply with the requirements of the trading program. These options involve questions about the method and level of monitoring, enforcement, and the availability and efficacy of sanctions.

a. Structural Options

Generally, regulators' options in establishing the structure of a trading system tend to be unique to the particular type of trading system. For example, the discharge limitations that allowances, group permits, and tradable limits employ each has a different basis and form. Reconciling each of these trading systems with the existing regulatory scheme also raises issues specific to that system. Nevertheless, several issues regarding regulatory structure can be considered generally for all of the trading systems; these are the scope of the regulated parties, the timing of reductions, and feedback mechanisms that allow the system to respond to changed circumstances.

Where many different sources emit at varying levels, it can be difficult to determine which sources it would be cost effective to regulate. Here, however, the refineries comprise the only major selenium sources under the jurisdiction of the regional board. There is little question that the selenium discharges of all six refineries should be regulated. North Bay municipal sewage treatment plants may be regulated in the future, after easier reductions have been made and their emissions are better quantified. Most of the sources of riverine selenium cannot be regulated directly by the San Francisco Regional Board because they are located in another jurisdiction. These sources would have to be included in a trading scheme either on a voluntary basis, or under some agreement with the Central Valley Regional Board. In neither case would they be included as directly regulated parties. Thus, at least initially, the scope of the regulated parties under any trading scheme would probably be restricted to the six refineries.

The timing of mandated reductions would probably be based on the expected rate and intensity of damage to the bay ecology and on the feasibility of reducing discharges. Reduction feasibility depends

286. See supra notes 111-16 and accompanying text.
287. EDF is currently developing a tradable permit system for agricultural selenium sources. Telephone Interview with Terry Young, supra note 115. If implemented, this would create the possibility of integrating the two permit trading systems.
on the cost and timing of production change options and of developing and implementing technological solutions. The more timing decisions depend on cost concerns, the greater the potential influence of a trading approach that would affect those expected costs.288

The current staff proposal suggests phased reductions.289 This would allow ecological effects to be assessed over time, possibly reducing the cost burden on refineries. The refineries object to phased reductions on the grounds that reductions are a one-step process; they claim the only feasible method of control is to develop control technology that, once implemented, would reduce to a fairly constant level.290 If this is correct, any individual refinery could only reduce to a given level after a one-time installation. Yet, even if this is the case, different dischargers could implement the technology at different times if phased reductions were coupled with trading. Individual refineries would gain some flexibility in their development and installation schedules, and those refineries that wait would presumably realize cost savings. Thus, to the extent that the reduction schedule is not dictated by purely ecological factors, it makes sense to consider how the timing and quantity of mandated reductions might allow cost-effective trades to be made. The size of reduction steps, whether they are even for each time period or weighted towards the beginning or end of it, and the length of the phase-in time period are all factors that might influence the cost-effectiveness of trades, and hence the overall cost of the regulation.

Any trading system must be reconciled with other control requirements, either by conforming the new system to existing requirements or by changing those requirements. A critical issue in reconciling either allowances or group permits with the existing regulatory system is whether to retain the December 1993 permit concentration limits. Presumably, the concentration limits themselves would not be tradable.291 Complying with both concentration and mass limits would greatly complicate both a trading system and individual transactions. Yet permits routinely contain multiple limits now, so to some extent this is an issue of degree, not of kind.292 Usually, one

288. In other words, if high costs are discouraging regulators from implementing reductions as quickly as needed, the prospect of lowering costs through a trading program might allow a more rapid reduction schedule.
289. Staff Report, supra note 8, at 40, Table 7.
290. See Telephone Interview with John Staton, supra note 218.
291. There seems to be little point in making concentration limits tradable. The main reason to use concentration limits is to protect against overloading the local ecosystem with concentrated discharges from a single source. This purpose would be negated by allowing trading.
292. See, e.g., SFBRWQCB, Order No. 89-002 (Jan. 18, 1989) (Union’s NPDES permit). For one wastestream (not including the selenium limits), this permit contains monthly average and maximum daily mass limits for 3 pollutants, monthly average and
limit will be the most restrictive, allowing regulators to focus on detecting and enforcing violations of that limit. Under a trading system designed to reduce mass emissions, the total quantity limit should generally be the most restrictive limit. Retaining the current concentration limits also might substantially restrict the tradable mass, since only those refineries with relatively dilute discharges would be able to comply with the standards by purchasing emission rights.\textsuperscript{293}

Attempting to loosen or eliminate the concentration limits may create legal problems. Because the limits are already in the permits, loosening the standards may violate the CWA anti-backsliding provisions.\textsuperscript{294} Yet, it is arguable that loosening concentration limits in the context of tightened mass limitations is not setting less stringent limitations. Removing concentration limits is not simply a matter of changing permit limitations. While the state plan concentration limits would not impose a constraint,\textsuperscript{295} the basin plan limits are low enough to interfere with trading,\textsuperscript{296} and hence the basin plan must be amended as well.

Feedback mechanisms are a final structural concern. Any regulatory program requires the capacity to respond to increased knowledge, unexpected problems, and changed circumstances. A regulatory program is particularly likely to need adjustments where the program must be instituted on the basis of incomplete knowledge. For example, any regulatory system must be able to adjust if a different estimate is made of what maximum emissions load will avoid ecological damages. This adjustment requires that regulators have the capability to modify individual discharge allocations. With trading, however, this imperative must be balanced against the need to provide a measure of certainty regarding future allocations. Certainty is necessary to encourage long-term compliance strategies that depend upon trading. Companies are unlikely to invest in producing excess reductions

maximum daily mass and concentration limits for 6 pollutants, monthly average and maximum daily concentration limits for 1 pollutant, and maximum daily concentration limits for 12 pollutants. \textit{Id}.

\textsuperscript{293} There might be room for some trading even with the 50 $\mu g/l$ concentration limits, since the reductions in the staff proposal go substantially beyond the present requirement of 1212 kg/yr (a 44% reduction from baseline) to 541 kg/yr in 1998 (a 75% reduction from baseline) to 216 kg/yr in 2001 (a 90% reduction from baseline). Staff Report, \textit{supra} note 8, at 40, Table 7.

\textsuperscript{294} See \textit{supra} note 190 and accompanying text.

\textsuperscript{295} The state plan sets a saltwater ambient selenium concentration standard of only 71 $\mu g/l$. State Plan, \textit{supra} note 118, at 4, Table 1. Since this standard is for the ambient concentration in the receiving waters, and is relatively high, it seems unlikely that this might be violated under a trading program designed to reduce discharges sharply.

\textsuperscript{296} The basin plan limits deep water fresh and marine selenium effluent concentrations to 50 $\mu g/l$, and shallow water fresh and marine selenium effluent concentrations to 5 $\mu g/l$. Basin Plan, \textit{supra} note 119, at Table IV-1A, Table IV-1B. These standards apply to the refineries. See Staff Report, \textit{supra} note 8, at 5.
for trade or plan to comply by purchasing emission credits, if the future quantities of permitted discharges are uncertain. A phased approach that allows ecological reassessment mid-course, before requiring further specified reductions, may be a good compromise. Such an approach stops short of setting future mass emission levels in stone, but it does establish the parameters for any changes and so allows firms to make compliance decisions on the basis of some bounded scope of risk. It also identifies the factors that will affect future changes, giving dischargers a basis for predicting future outcomes. In lieu of a phased approach, there would need to be some provision allowing the regional board to lower the future mass emissions cap. The mechanism could vary depending upon the trading system, however.

Unanticipated problems in the regulatory system, as opposed to the environment, might also require adjustments. The most likely problem for an implemented trading system is market failure, in which the trading system would achieve targeted discharge reductions, but would not realize anticipated cost-savings. Adjustments to deal with such an eventuality could be facilitated by designing appropriate flexibility into the regulatory system. Given the significant legal and political constraints on designing a trading system, however, there may not be many options available in making initial design choices, let alone an opportunity to incorporate the flexibility to make subsequent modifications to the system.

One way to maximize flexibility is to authorize the system on the regulatory level most accessible to later adjustments. Permit provisions may be more easily modified than basin plan provisions, and both are more amenable to change than are the state plan or federal laws and regulations. Mid-course adjustments would be easier under a system designed to enhance regulatory discretion than under a system with more specific requirements binding regulators. The tradeoff, however, is that regulatory discretion can be used to undermine a program. Even where regulators are favorably disposed to trading, increased discretion enhances uncertainty, which itself is a barrier to trading. More specific feedback mechanisms will be addressed in relation to the different trading approaches.

b. Options in Allocating Discharge Rights

Allocation issues encompass both the mechanism used to distribute emission rights (i.e., are they given away or purchased?) and

297. See supra parts II.A.3, II.A.6.
298. See supra notes 120-25 and accompanying text.
299. See supra notes 50-52 and accompanying text.
the basis for the initial division of emission rights (i.e., which source gets how many rights and why?). One proposal that addresses both of these issues is to auction emission rights periodically to dischargers. An auction theoretically allows the optimally efficient initial and ongoing distribution of rights to be achieved without the distorting transaction costs of reallocating between dischargers in a market. The auction could be revenue-neutral, in which case the proceeds would be returned to the dischargers, preferably in some manner that did not affect, and thus distort, their bidding decisions. Alternatively, the government could retain the proceeds, although since this approach would raise revenues, it probably would not be legal without explicit legislative authorization.

Probably motivated largely by political and fairness concerns, most trading systems to date have distributed discharge rights free of charge, or “grandfathered” in some portion of existing emissions. Since trading systems are usually coupled with an emission cap that represents some level of reduction from current emissions, auctioning rights would make polluters pay for both required reductions and retained emissions. Not surprisingly, these proposals tend to incite virulent political opposition, and the selenium situation would be no exception. One economic argument for revenue-raising auctions (or pollutant-fee systems) is that paying for all emissions creates a strong incentive to control at efficient levels where marginal costs of control equal the costs of purchasing discharge rights. A trading system theoretically incorporates incentives for efficient control, even without requiring that all emission rights be purchased, because firms bear


301. The government might want to retain revenues raised through selling discharge rights either solely as a new source of funds, or on a theory of internalizing the social externalities of pollution and applying the proceeds to offset the social damages. A revenue-raising auction system could be characterized as a hybrid between tax and trading approaches, where all units of pollution are paid for, as under a tax system, but the overall quantity is capped, as under a trading system.

302. While it would be illegal for EPA to raise revenue under the CWA without statutory authority, states might be able to sell discharge rights. Rothfelder, *supra* note 168, at 417-18; 40 C.F.R. § 123.1(i) (1992). Even state-imposed revenue-raising discharge rights sales probably would have to be authorized through state legislation, however.

303. Fairness concerns are implicated by the possibility of disappointing investment-backed expectations. In other words, industry has developed in the expectation that the water would be available for disposal capacity. Changing the rules of the game always provokes protests of unfairness, despite the fact that a multitude of other factors in the business environment change without compensation to disappointed interests. Moreover, grandfathering in existing firms is arguably more unfair to new sources, which must commence operations at a competitive disadvantage.

304. *See*, e.g., 42 U.S.C. § 7651g(c).


an opportunity cost when they retain rights that could be sold for profit. Yet market failures in the permit trading market can undercut the effectiveness of such incentives.\textsuperscript{307}

Auctions have the advantage that the dischargers select the initial distribution. By contrast, under a grandfathering approach the regulators will have to develop an “equitable” allocation method. Given that every affected interest can devise some neutral principle that operates to its own benefit, equity can be a highly contentious issue and a regulatory nightmare. Group permits present a possible exception. Under the most extreme group permit approach, the entire group would be given a mass emission allocation, leaving it to the dischargers to allocate individual allotments. This approach would almost certainly require regulatory groundrules to have any chance of success. Stalemate appears inevitable without a tool to coerce agreement, for each discharger could hold out for the distribution method most advantageous to itself. If there is a means of forcing agreement, one or several of the dischargers are likely to feel themselves railroaded into an inequitable solution.

Wasteload allocations performed under water quality-based permitting also require regulators to distribute discharge rights among polluters.\textsuperscript{308} Since regulators have broad discretion in choosing a method for calculating WLA’s,\textsuperscript{309} they have developed a variety of allocation methods equally applicable to trading. Three of the most common methods are equal effluent concentrations, equal percent removal, or equal portions of the segment’s assimilative capacity (this would translate roughly to equal mass emissions, if discharges across sources have the same impact on assimilative capacity).\textsuperscript{310} Another method, more analogous to that used in the acid rain program, is equal percentage reductions from baseline emissions, sufficient to achieve the total mass emission objective.\textsuperscript{311}

Each of these methods can and probably will be contested on some equity ground. The last method requires a determination of a

\textsuperscript{307} See Tietenberg, supra note 43, at 22 (setting out conditions for achieving a cost-effective allocation of control responsibility regardless of initial distribution).

\textsuperscript{308} See supra note 151 and accompanying text.

\textsuperscript{309} Regulators “may use any reasonable allocation scheme that meets the antidegradation provisions and other requirements of State water quality standards.” Technical Support Document, supra note 92, at 69.

\textsuperscript{310} Industrial Economics, Inc., supra note 33, at 1-5 (citing William McLoud, Wasteload Allocation Methodologies for Multiple Dischargers, Masters’ Degree Study, Univ. of Colo., (1990)); Technical Support Document, supra note 92, at 69, Table 4-1.

fair emissions baseline, frequently based on average emissions over a time period of several years. Where, as here, some refineries have recently increased their selenium discharges, while others have reduced their discharges, the selection of a baseline is likely to be highly contentious. The equal effluent concentration method will disadvantage dischargers committed to water conservation and wastewater reclamation. Relating allocations to units of production will hurt those producing relatively cleaner petroleum products with relatively higher selenium discharges. It may be useful for regulators to determine initially a neutral principle to direct their choice of allocation methods. Thus, rewarding past environmentally good behavior might be the primary allocation objective. Alternatively, a goal of spreading the costs widely to minimize economic dislocation, regardless of merit, might support a different basis for allocation.

While the allocation method will significantly affect the distribution of reduction costs between the dischargers, the impact allocation has on the overall cost-effectiveness of the program will depend on how well the market functions. Under a tradable limits approach, where high transaction costs are likely to discourage all but the most cost-effective trades, the initial allocation method could substantially influence the overall cost of the program. On the other hand, under a successful allowance approach, the initial distribution will mainly affect distributive, not efficiency outcomes. To some extent, then, allocation decisions should take into account the type of trading approach envisioned.

c. Options in Market Form and Operation

Defining the market for any type of trading program includes identifying the market players and its spatial area. The participants in a market need not be identical to the regulated parties. All regulated parties might not be allowed to trade, while some unregulated parties might be allowed to trade. Here, at least all six of the refineries should be allowed to participate. Public interest groups could be included as buyers only, who would purchase and retire selenium allowances for environmental benefit. Additional potential participants are the North Bay municipal sewage treatment plants and


313. The high cost of selenium reductions and public interest groups’ inability to capture as funding anything near the true social benefit of reductions (the free-rider problem) makes it highly unlikely that public interest parties would be active to any significant extent. See Matthew L. Wald, Lung Association Getting A Donation of Cleaner Air, N.Y. TIMES, Mar. 20, 1993, at 26 (mentioning that one small environmental group is soliciting donations to purchase and retire allowances). Also, the premise of a mass cap is that it
the nonpoint sources of riverine selenium. Either of these groups could be included on a voluntary basis, analogous to an offset approach, which allows “new” allowances or credits to be created through reductions by unregulated parties.

The sewage treatment plants, however, are unlikely to be the source of any excess reductions. They emit very small quantities of selenium at low concentration levels, frequently below the analytical detection limits. Hence, estimates of their discharge quantities are uncertain, contrary to the trading prerequisite of reasonably accurate quantification. Moreover, the sewage treatment plants are likely to face even higher reduction costs than the refineries, due to their much lower effluent concentration levels and the difficulty of tracking their selenium sources. Also, the profit motivation to find the least-cost solution through trading may not drive municipalities to the same extent as private industry. Thus, the treatment plants might be inactive market players even if they were included as market participants.

Including riverine sources raises other issues. Selenium discharges from nonpoint and agricultural sources of selenium are significant, but intermittent, and quantification presents a significant problem. It is relatively difficult to measure accurately selenium discharges produced by different irrigation and land management practices on a given area of agricultural land. This creates serious problems in quantifying both baseline discharges and changes from that baseline produced by certain actions. Also, while discharges to the San Joaquin River are the more plausible target of tradable reduc-

would be set at a level sufficient to protect the environment, so there should be less need for public interest groups to purchase further reductions.

314. In an allowance system, brokers also could be included either as market participants, or simply as market facilitators who match up buyers and sellers. Here, however, the size of the market and familiarity of all the parties probably leave little room for brokers.

315. These parties would not have to be direct market participants; rather the refineries could generate credits through funding, quantifying, and monitoring agricultural and other nonpoint source reductions. This approach resembles the point-nonpoint trading programs established in Colorado and North Carolina. See Bartfeld, supra note 2, at 83-88.

316. Interview with Kim Taylor, supra note 113.

317. Staff Report, supra note 8, at 40.

318. Interview with Kim Taylor, supra note 113.

319. Experience with the acid rain program indicates the great difficulty that even a highly regulated private industry, let alone a public institution, has in single-mindedly pursuing the most profitable option. Public utility commissions (PUC’s) have discouraged or prevented utilities from pursuing least-cost compliance through allowance purchases, in deference to political pressures, such as an in-state high-sulfur coal industry, which prefers technology-based compliance (scrubbers) compatible with continued use of its product. See Wald, supra note 241, at C2 (discussing market inhibitions created by PUC regulation).
tions, San Joaquin River flows into the bay vary substantially by season. In crediting reductions, it would be necessary to credit selenium discharges avoided only if they actually would have reached the relevant ecological areas.

Moreover, it is important to ensure that the reductions achieved are "real" and "permanent," as well as quantifiable. Real gains are those that not only would not have been achieved absent the trading incentive, but also are not offset by a compensating increase in discharges elsewhere. Agricultural sources are regulated by a different regional board and are located in an area where there is substantial pressure to reduce both water usage and selenium contamination in the runoff. Regulators must avoid improperly crediting actions taken in response to such pressures, because they would have been taken without the incentive of trading reductions. Regulators also should not grant credit for reductions offset by moving damaging agricultural practices to another piece of land that otherwise would not have been a source of selenium.

The permanence of a reduction only matters for the time period of the emission credit produced. If allowances are being created and transferred on an annual basis, the reduction need only be in place for one year. If, however, a stream of credits or a right to discharge a given quantity annually is being transferred, it is important to ensure that the seller continues to achieve the reduction. This raises significant monitoring and enforcement concerns, and it creates greater difficulties in determining what emissions "otherwise" would have been over time.

Another major market design option is the delineation of the spatial area in which trading is allowed. To maximize market viability, the trading area should be made as broad as possible, subject to the overriding environmental imperative. The spatial extent of the market must be consistent with the area of source-neutral ecological impacts in order to comply with the basic environmental justification for trading: the focus of concern is mass loading, not the individual source's emissions. Here, the market should not encompass dis-

320. Sacramento River selenium sources are less well identified and, as much of the selenium loading is from general runoff, control of sufficient specific sources is likely to be difficult. See Telephone Interview with Terry Young, supra note 115; Interview with Dr. William S. Pease, supra note 12.

321. Interview with Dr. William S. Pease supra note 12. Moreover, future flow regimes cannot now be predicted and are likely to change as a result of the ongoing conflict over Bay-Delta water quantity and quality issues. Id.

322. Id.

323. See AER*X, supra note 51, at 8.

324. If the reduction would have been made without the trading incentive, it is double counting to allow that reduction to offset an increase in refinery discharges.
CONTROLLING SELENIUM DISCHARGES

charges that do not enter the North Bay, because the primary ecological impacts of selenium discharges to the North and South Bays are largely limited to each of those water bodies. The biogeochemical cycling of selenium, however, is not fully understood. If further study were to indicate that the impacts of selenium are more closely correlated with the site of discharges, it might be necessary to implement further spatial restrictions on trading. These could take the form either of restricting transfers to within the individual water segments (which essentially would eliminate the market), or limiting the mass quantity that could be discharged in a given segment (basically setting a maximum transfer quantity for each segment).

Another issue in defining the market is how to treat changing players over time (i.e., the addition of new sources and shutdown of existing sources). New sources could be allocated credits or could be required to purchase credits from existing sources. Upon shutdown, regulators must choose whether to allow the source to retain and transfer the credits that would be created by its ceased operations. To the extent that the value of the tradable reductions causes a shutdown that would not have occurred otherwise, there is no net environmental detriment to allowing transfers. However, environmental interest groups strongly resist allowing the net potential environmental gains upon shutdown to be transferred to another source. Also, it is difficult to determine whether the potential transfer of tradable permits has contributed to a decision to cease operations. Regulators may be spared making these decisions here. While most trading systems must provide for emissions from new sources, the development of major new selenium sources in the bay is unlikely. Shutdowns in the near future are also unlikely given the existing investment in these refineries.

d. Monitoring and Enforcement Options

The importance of monitoring and of the techniques used is not unique to market-based regulatory approaches. Accurate monitoring of actual emissions is important for implementing any standard that is

325. See Staff Report, supra note 8, at 3 (distinguishing between selenium enrichment in the North and South Bays, but not placing limits on individual water body segments within the North Bay). Given the current understanding of the action and transport of selenium in the receiving waters, the regional board staff have not found it necessary to develop limits for the individual water segments of Suisun Bay, Carquinez Straight, and San Pablo Bay. Thus, transfers would not need to be restricted to within each of these segments.

326. See generally id. at 14-16 (discussing the biogeochemical process).

327. There are only one or two refineries discharging into each of these water bodies.

328. See supra note 250 (regarding the author's experience).
not based on technology or management practices, but rather directly limits emission quantities. When the entities trading are the same as those that would be regulated under a traditional water quality-based WLA approach, the effects of inadequate monitoring would be no more detrimental under trading than under the traditional approach. Trading does raise additional monitoring concerns if unregulated sources are included in the market, either as direct participants, or as a source of offsets.

The monitoring of selenium discharges from the bay refineries presents little problem. An accurate monitoring system is already in place. The refineries measure discharges on a weekly sample basis, rather than by continuous emission monitoring. This is standard practice for controlling toxics in effluents and further precision seems unnecessary to track discharge quantities adequately. Nor does it appear that there would be any greater incentive or opportunity to cheat under a trading system than under any other type of actual emission-based regulation.

Trading requires tracking transfers in addition to individual quantities discharged. Currently, the refineries report discharge quantities on a monthly basis, allowing calculation of annual totals. Regulators also could readily require regulated entities to report all discharge transfers during the reporting period. Depending upon the timeframe chosen, compliance could be assessed by comparing actual emissions with allowances held or permit limits on a monthly, quarterly, or annual basis. With only six refineries, the accounting required should not be substantially more complex than under the current system, where regulators must determine compliance with maximum daily concentration limits and maximum annual average mass emissions rates. Measures to protect against hot spots, such as retaining higher concentration limits or instituting rolling averages to measure compliance with annual loads, could further complicate administra-

329. Technology-based standards tend to be easier to monitor, as it is fairly simple for regulators to ascertain whether the control technology has been installed; even so, there are issues of constancy of operation, breakdowns, etc. Management practices may be an easily assessed one-time change in operating procedures or may be ongoing practices that could be allowed to lapse on a short-term basis. Hence they can be difficult to monitor.

330. See supra notes 313-19 and accompanying text.

331. Telephone Interview with Tom Mumley, supra note 177.

332. Id.

333. See SFBRWQCB, Order No. 91-026 (Feb. 20, 1991); SFBRWQCB, Order No. 91-099 (June 19, 1991) (setting interim mass emission rates); see supra note 176 and accompanying text.

334. EDF Comments, supra note 280, at 3; Telephone Interview with Terry Young, supra note 115. A rolling average would continuously recalculate average emissions over a given time period, such as four or twelve weeks. For example, to calculate a rolling four week average, the previous four weeks' discharge quantities would be averaged each week.
tive tracking, but these too do not seem to impose a significant additional burden.

Strong enforcement is important to ensure the continued viability of the market under any trading approach. The integrity of the commodity traded depends upon accurate monitoring and reporting. Cheating effectively inflates the currency, thereby debasing its value. The CWA provides a variety of enforcement mechanisms, all of which can be applied to violations under a trading system. These include administrative penalties of up to $10,000 per day, with a maximum of $125,000,335 and civil penalties of up to $25,000 per day for each violation.336 Criminal penalties are also available.337 Administrative penalties may be imposed by EPA,338 or by the state under comparable state law.339 Civil actions can be brought by either state or federal regulators, or by individuals under the citizen suit provisions.340 The stringent criminal penalties for false reporting are a fine of up to $10,000 and/or two years imprisonment.341 These penalties provide additional assurance against cheating in a trading program.

The effectiveness of enforcement depends to a large extent upon regulatory vigilance in detecting and prosecuting violations and upon regulatory willingness to apply large sanctions. There is no reason to expect regulators to exercise less stringent enforcement practices under a trading approach. Indeed, as a new and somewhat controversial program will attract increased public scrutiny, regulators may feel even more pressure to ensure that trading meets environmental goals by vigorous enforcement. Independent of regulators' willingness to enforce, the threat of citizen suits is another major enforcement tool in the water context.342 Thus, any trading system would have to be

335. 33 U.S.C. § 1319(g) (1988); Lawrence R. Liebesman & Elliott P. Laws, The Water Quality Act of 1987: A Major Step in Assuring the Quality of the Nation's Waters, in CLEAN WATER DESKBOOK 23, 29-30 (Environmental Law Reporter, Environmental Law Institute ed., 1988). In determining the amount of a penalty, the regulator "shall take into account the nature, circumstances, extent and gravity of the violation, or violations, and, with respect to the violator, ability to pay, any prior history of such violations, the degree of culpability, economic benefit or savings (if any) resulting from the violation..." and other factors. 33 U.S.C. § 1319(g)(3).
340. 33 U.S.C. § 1319(b) (EPA authority for civil suits); id. § 1342(b)(7) (state permit programs must include adequate authority to "abate violations of the permit or the permit program, including civil and criminal penalties and other ways and means of enforcement ..."); id. § 1365 (citizen suits).
341. Id. § 1319(c)(4).
designed to be at least as easy to enforce through citizen suits as are current standards.343

2. Design Options Differing Across Types of Trading Systems

Although some design options are common to all three types of trading systems, several issues are unique to each of the individual systems. Since point source trading is more theory than practice, it is useful to survey additional options encountered in designing each of these types of trading systems and to note how different design choices might work in the specific context of the San Francisco Bay Estuary selenium problem. Again, both the general framework of analysis and many of the issues identified are equally relevant to other scenarios.

a. Allowance System

An allowance system comes the closest to “commodifying” pollution rights and allowing them to be transferred between dischargers. The environment is protected by capping the total number of available allowances and appropriately defining the scope of the market. One allowance permits the discharge of one unit of selenium. Each discharger receives a given number of allowances each year (or other accounting period), which are used up as units of pollution are discharged.344 The discharger’s NPDES permit would prohibit the discharge of selenium in excess of the number of allowances held, but the actual number of allowances would not be specified in the permit. While this prohibition clearly limits the quantity of selenium discharged, it does not set a single numeric permit limitation, as the CWA might be read to require.345 Moreover, an allowance system allows a discharger to “use up” its allocation fully.346 This differs from the current limitations, which require a discharger to keep emissions within a statistical margin of safety,347 meaning that discharges usually will be well below the maximum limit. Environmentalists may oppose allowances on this ground. The difference can be corrected, however,

343. Telephone Interview with Marc Joseph, supra note 264.
344. Alternatively, each discharger could receive a “permit limit” or “emission right” authorizing the discharge of X units of pollution annually (or over some other time period). Real, quantified, permanent reductions from this baseline could be sold. The transaction would be registered by amending the permit limits of each party to reflect the new allocations for all future periods. This approach, taken in RECLAIM, is addressed here under the general heading of “tradable limits.” See RECLAIM Executive Summary, supra note 27, at EX-12 to EX-13.
345. See 33 U.S.C. § 1312(a); see also supra part II.A.3.b.i.
346. Telephone interview with Tom Mumley, supra note 126.
347. See supra note 177.
by setting the total quantity of allowances based on the statistical mean emissions level, rather than on the maximum permit level.

One major design choice unique to allowances is establishing the accounting timeframe. Actual mass discharges could be balanced against allowances held on a monthly, quarterly, annual, or other basis, in order to determine whether the discharger has violated the permit limit. Annual accounting would maximize compliance flexibility and promote a more vigorous allowance market. Firms cannot predict shorter-term discharges with as much certainty and are less likely to make the small-scale transactions that would be needed to comply with monthly mass limits through trading. A short accounting period would encourage dischargers to achieve compliance through potentially more expensive means that are fully under their control, rather than depending on an allowance market. Environmentalists and regulators may be concerned, however, that a long accounting period could produce large year-end overruns, which could be avoided with more frequent accounting and earlier enforcement. Quarterly accounting is a possible compromise.\(^{348}\)

In addition, legal requirements and/or hot spot concerns are likely to compel regulators to include daily and monthly maximum permit limitations, either mass- or concentration-based.\(^{349}\) Within the bounds of providing necessary environmental protection, these limitations probably should be set as high as possible. For a market to function, daily and monthly limits must be sufficiently loose to allow dischargers to achieve compliance with annual mass limits by purchasing allowances.\(^{350}\)

It is unclear whether an allowance system incorporates feedback mechanisms that allow regulators to be more or less responsive to changed circumstances than a nontreading approach. One common objection to marketable permits is that they confer a “right” to pollute, which may be difficult to reduce or take away. In response to

\(^{348}\) This has been proposed for the SCAQMD’s RECLAIM program. RECLAIM Executive Summary, supra note 27, at EX-14.

\(^{349}\) See supra notes 155-56, 176 and accompanying text.

\(^{350}\) A hypothetical example illustrates the interaction of multiple permit limits. Suppose a firm receives 100 selenium allowances for a 100-day accounting period, which would allow it to emit on average, one unit of selenium per day. The permit limitations require that (1) the firm hold allowances equal to the total quantity of selenium emitted, (2) the maximum daily limit is six units per day, and (3) the average monthly limit is three units per day (alternatively, the maximum daily and average monthly limitations could be concentration limits). To discharge a total of 200 units of selenium over the period, the firm could buy an additional 100 allowances. If discharges were absolutely uniform, this would produce discharges of two units per day. Such uniformity is highly unlikely, so the secondary permit limits would allow discharges to vary from this mean by up to a factor of three on a given day, or a factor of two over a month’s average. Assuming that discharge quantities varied only within this range, on their face the additional limits should not restrict trading.
this concern, legal authorizations usually state explicitly that allowances or other forms of marketable permits are not property rights.\textsuperscript{351} If anything, the permanence of tradable permits has been too uncertain in practice. Regulators have tended to devalue or confiscate permits,\textsuperscript{352} which discourages compliance strategies based on banking or producing excess reductions for sale.\textsuperscript{353} Attempts to enhance the security of allowance values or allocations would rise to the level of a constitutional "taking"\textsuperscript{354} unless some kind of property right were statutorily conferred. More likely, regulators' capability to reduce future allocations will be determined largely by political and environmental considerations.\textsuperscript{355} These considerations probably would not have a substantially different influence on the capacity to change nontradable permit limits than to change allowance allocations.

Market design will be critical to whether a market in allowances actually develops, and thus whether the firms realize the potential cost savings.\textsuperscript{356} Regulators should design the market to minimize transaction costs to the extent possible, given necessary precautions to ensure meeting environmental goals. Regulators in other trading systems have commonly required regulatory preapproval of trades.\textsuperscript{357} However, the effort of preparing and submitting each proposal, the delay necessitated by waiting for regulatory approval, and the heightened uncertainty due to the possibility of a veto all impose substantial transaction costs. Proponents justify regulatory pre-approval as necessary to evaluate the possible detrimental environmental effects of each transfer.\textsuperscript{358} An alternative is to establish uniform groundrules to

\textsuperscript{351} See, e.g., 42 U.S.C. § 7671b(f) ("An allowance allocated under this subchapter is a limited authorization to emit sulphur dioxide in accordance with the provisions of this subchapter. Such an allowance does not constitute a property right. Nothing in this subchapter or in any other provision of law shall be construed to limit the authority of the United States to terminate or limit such authorization.") (emphasis added).

\textsuperscript{352} See supra notes 50-51 and accompanying text.

\textsuperscript{353} Dwyer, supra note 45, at 110 & n.44.

\textsuperscript{354} U.S. Const. amend. V.

\textsuperscript{355} For example, banked permits have been discounted by the SCAQMD, Dwyer, supra note 45, at 110, which has jurisdiction over the areas with the worst air pollution problem in the country and is under the greatest pressure from EPA to improve air quality.

\textsuperscript{356} Because an allowance approach establishes a more open market than group permits or tradable limits, several of the market design issues arise most prominently under this approach.

\textsuperscript{357} One example is the offset program. See supra notes 16, 19-20 and accompanying text. Preapproval also raises issues of what criteria regulators will use to decide and, if this is a discretionary decision, whether the California Environmental Quality Act (CEQA) will apply. See Cal. Pub. Res. Code §§ 21000-21177 (West 1986 & Supp. 1994). Telephone Interview with Marc Joseph, supra note 264.

\textsuperscript{358} See supra note 250 (regarding the author's experience).
protect against hot spots or other environmental concerns, without re-
quiring approval of each individual trade. The rules adopted under
this approach risk being overinclusive, underinclusive, or both, for a
given situation, thereby imposing some unnecessary restrictions and/
or allowing some environmental damage. Yet, groundrules would
save regulatory effort (assuming a significant number of trades are
proposed) and reduce transaction costs. Past experiences have
shown that a preapproval requirement imposes a substantial burden
upon the market and depresses trading.

Regardless of whether preapproval is required, all trades must be
reported to the regulators in order to allow tracking and accounting.
The extent of these reporting requirements has been another conten-
tious issue in developing trading systems. The main areas of poten-
tial controversy are whether the trading parties and transferred
quantities should be made public, and whether prices must be re-
ported. NPDES permits are currently available to the public and may
be the basis of citizen enforcement actions. Thus, it seems that in-
formation on allowances held and quantities discharged must be pub-
licly available, both to comply with the CWA and for practical
enforcement purposes. Nor is there any justification for the discharg-
ers to suppress this information. Whether to require public reporting
of prices for transactions is more debatable. Public price information
is necessary to establish a price signal for the market. Yet dischargers
may reasonably oppose revealing this information on the ground that
it constitutes a trade secret.

359. The regulatory burden under each of these options will depend upon the difficulty
of establishing groundrules, weighed against the number of trades proposed and the diffi-
culty of reevaluating the environmental impacts for each trade. It seems plausible that it
would be easier to establish acceptable trading parameters once, rather than repeatedly.
This decision should depend in part on how widely environmental impacts are expected to
vary from trade to trade.

360. The offset program provides one example. See Hahn & Hester, supra note 14, at
378-79 (discussing the impact of regulatory oversight of trades on trading activity).

361. This was true for the acid rain trading program, for example. See supra note 250
(regarding the author's experience).


363. Another potential restriction on the market is a requirement to show that al-
lowances offered for sale really are excess, due to identifiable emissions reductions. Such a
showing could be required either prior to transfer or simply upon regulatory request. This
requirement is clearly unnecessary under continuous emission monitoring, since a seller
without excess allowances will be found in violation at the end of the accounting period
and subject to sanctions. The same objection also should apply for any system where com-
pliance is monitored on the basis of measuring actual emissions, rather than on technology-
based standards. Here, there seems to be no reason to require regulatory review or ap-
proval of discharge reductions or allowance availability; actual discharge monitoring seems
sufficiently precise to assess compliance solely by comparing the level of reported dis-
charges to the quantity of allowances held.
A timing issue is whether allowances should be valid only for the accounting period in which they are issued, or whether they may be saved and "banked" for future use.\textsuperscript{364} Banking provides an incentive for early reductions in a phase-down program.\textsuperscript{365} It also allows dischargers to smooth non-stepwise control strategies to comply with stepwise mandated reductions.\textsuperscript{366} Banking, however, could interfere with ecological assessments premised on the mandated stepwise reductions.\textsuperscript{367} While banking might provide some valued additional flexibility to dischargers, neither its ecological nor economic effects are clear on the basis of current knowledge.\textsuperscript{368}

A difference between implementing an allowance system versus other trading approaches is the need to track allowances and enforce on the basis of allowances held. Actual enforcement mechanisms should remain unchanged.\textsuperscript{369} Assessing sanctions for violations under an annual or other long accounting period may raise new issues, such as how to calculate the magnitude of a violation.\textsuperscript{370} Whatever the

\textsuperscript{364} See supra note 17.

\textsuperscript{365} If refineries develop one control technology that reduces to a given level, however, there is little incentive to bank allowances, except as insurance against breakdowns; once the technology is in place, there would be little cost savings to operating it at a level allowing higher discharges.

\textsuperscript{366} For example, assume dischargers have only two control levels, at 30\% and 75\%, but regulators mandate reductions of 10\% per year. Dischargers could use excess reductions from the first few years of 30\% control to delay implementing the 75\% control measure. This allows dischargers to avoid implementing the 75\% control measure in the fourth year to achieve 40\% control. While this simplistic formulation begs the question of why regulators would not mandate 75\% control earlier, a smoothing function brings flexibility to a system where regulators make general rules that apply to a wide variety of sources with different control capabilities at different times.

\textsuperscript{367} See generally Staff Report, supra note 8, at 40 (staff proposal timeline, providing for a review of ecological monitoring data to determine if further reductions are needed after a 75\% reduction from baseline). If refineries were to overreduce early, ecological effects might not reflect the impacts of an ongoing 75\% reduction loading level. On the other hand, this will not be a problem if ecological changes do not quickly reflect loading changes, as is possible. See Letter from Royer to SFBRWQCB, supra note 312, at 2.

\textsuperscript{368} Although selenium is a conservative pollutant, it may become less accessible to the ecological system over time through deposit in the sediments. If so, the time of discharge does matter for detrimental effects on the ecosystem.

\textsuperscript{369} One potential problem for citizen enforcement under an allowance system is the Gwaltney doctrine, which limits citizen suits to continuing, not past violations. See Gwaltney of Smithfield v. Chesapeake Bay Found., 484 U.S. 49, 56-63 (1987). To establish a continuing violation the party must show either that (1) the violation exists on the day the complaint is filed, or (2) there is a reasonable likelihood the discharger will be in violation in the future. Id. at 56-59. Given that violation of an allowance-dependent limit is assessed only periodically and that the discharger might be able to purchase allowances to make up the difference after the fact, it might be difficult ever to show a discharger to be in continuing violation. Telephone Interview with Marc Joseph, supra note 264.

\textsuperscript{370} See Telephone Interview with Laurie Kermish, supra note 125. Currently, sanctions for violating a monthly maximum average limit may be calculated by assuming that if the monthly sample is in violation, the limit has been violated every day of the 30-day period. Telephone Interview with Tom Mumley, supra note 177. For a small exceedance of
method of calculating sanctions, it is critical that the sanction for discharging a unit of pollution without an allowance at least exceed the market price of an allowance. Otherwise, there would be no incentive to comply through allowance purchases.

b. Group Permit System

A group permit system would allow cost-effective reallocation of discharge reductions by placing the discharge limit on the dischargers as a group, rather than on individual dischargers. Since group permits would largely use the existing NPDES permitting approach, they could be seen as less of a change from the existing regulatory approach. All the individual sources would retain permits under this approach, but an additional permit would be issued for the entire group of dischargers, limiting the maximum mass discharge of selenium for the permit period. As with allowances, regulators would need to define the appropriate accounting period and resolve the issue of whether to impose any additional permit limits on either individual refineries or the group.

A group permit may be issued to some kind of formal entity composed of the dischargers. The permit may incorporate a plan to address the pollution problem and the agreement between the dischargers for how the plan will be implemented. This approach has been applied to nonpoint source pollution control by municipalities, such as storm water control, where the dischargers need to agree to implement certain control measures, but the impacts of specific control actions are difficult to quantify. Here, with high control costs and relatively precise pollutant quantification, it may be harder to negotiate an agreement for individual responsibilities.

If the agreement is incorporated directly into the permit (creating enforceable limits), it essentially becomes a mechanism for setting in-

---

371. See supra part I.D.2.
372. See, e.g., NPDES Permit I, supra note 87.
373. See id.
dividual permit limits. This would effectively use a market approach only for the initial distribution, as opposed to an allowance system, which uses markets forever after the initial distribution by the regulatory agency. Assuming the optimal cost-effective distribution of discharge quantities does not remain constant, such a group permit would achieve lower cost savings over time. Even the initial allocation may not be cost effective, due to distortions such as uneven bargaining power or unwillingness to trade. Thus, incorporating negotiated limits into individual permits would undermine the main point of using a group permit. A better approach would be for individual permits to reference the agreement in its existing form while leaving dischargers free to modify the agreement easily. This would allow enforcement against individuals on the terms of the agreement, but the dischargers would control those terms, subject to the aggregate mass limits. Alternatively, the system might simply provide for a periodic reopening of the group permit to negotiate new limits.

A group permit approach is not substantially different from the other approaches in its flexibility to respond to changed circumstances that require different levels of control or other adjustments. With any trading approach, requiring greater discharge reductions would necessitate determining a new mass limit and new individual allocations. Here, the new mass limit would be incorporated as the group permit limit, and new individual allocations would be determined in the same manner as initial allocations were determined. It seems unlikely that these changes would be either more or less legally and politically burdensome with a group permit than with an allowance or tradable limits approach.

A group permit is unlike the other trading-based approaches in that it could alleviate the allocation problem for regulators. The mechanism allows the mass discharge quantity to be set for the group as a whole, and it is then up to the dischargers to reach some agreement regarding individual allocations. The refineries tend to view this proposal as unworkable, however.\textsuperscript{374} It might be necessary for regulators to establish groundrules for negotiating responsibilities, set initial allocations, or somehow mediate the allocation process. A group permit approach probably requires that there be a fairly small number of market participants. The argument for limiting the trading system to the six refineries seems even stronger than with the other approaches. It would be difficult enough for these few similarly situated dischargers to coordinate reduction strategies; expansion to other sources could make implementation substantially more problematic.

\textsuperscript{374} See Telephone Interview with John Staton, supra note 218; Telephone Interview with Dan Glaze, supra note 223.
A group permit system raises unique enforcement issues because joint and several responsibility allows regulators to impose sanctions on either individual dischargers or the group as a whole for mass limitation violations. This raises the question of who in practice will be held responsible if the group exceeds the mass discharge limit. The regional board could take enforcement action against the refineries that are exceeding limits established under the dischargers’ contractual agreement. Yet, if that agreement is not incorporated by reference into the permit, this approach could be contested in court as an abuse of regulatory discretion, for it is not clear that the individual refinery would have violated either its permit or the statutory requirements by not complying with its agreement with the other dischargers. Alternatively, the board could proceed against the group as a whole, leaving it to the individual dischargers to determine liability among themselves in lawsuits for contribution. While this strategy may appear simpler from the regulator’s standpoint, it invokes the specter of wasteful, expensive, and time-consuming litigation. Economic theory identifies additional flaws in this approach, based on the free-rider problems inherent in collective action and provision of public goods. If the nonviolating dischargers do not fully recover from the responsible party through subsequent litigation, each discharger would then have an incentive to violate, because each would receive the full benefit of its exceedance, but pay only a portion of the fine. The uncertainty over who is responsible for violations and how enforcement would proceed is seen by virtually every party interested in the selenium issue as the fatal flaw in group permits with no individual limits. In addition, the incentive to develop innovative control strategies might be reduced because an individual discharger might receive only a portion of the benefits.

Yet another potential problem with group permits is that dischargers’ collaboration for the purpose of pollution control might be considered to violate antitrust laws. If collaboration is required by statute, however, a court should assume that Congress implicitly repealed the antitrust provisions with respect to those explicitly authorized activities. Thus, if the CWA explicitly set up a mechanism for

375. The dischargers could agree to levy some sort of fee amongst themselves to pay for sanctions and/or control programs.
376. This would be analogous to the Superfund approach. 42 U.S.C. § 9613 (1988).
377. See generally Grinols, supra note 43, at 569-70 (discussing free riders and public goods).
378. See Telephone Interview with Dan Glaze, supra note 223; Telephone Interview with John Staton, supra note 218; Telephone Interview with Laurie Kermish, supra note 125; Telephone Interview with John Norton, supra note 160; cf Telephone Interview with Tom Mumley, supra note 156 (noting it is possible to enforce against multiple parties).
group pollution control, compliance with those provisions should not violate antitrust laws. Where the statute is ambiguous in authorizing a group permit, as is the CWA currently, the question whether anticompetitive activities have been authorized for pollution control purposes remains. This legal uncertainty could raise a further barrier to effective discharger collaboration.

c. Tradable Limits System

A tradable limits approach would retain the current NPDES permit system and allow transfers pursuant to the CWA provisions that authorize reallocation of WLA's so long as overall water quality standards are met. Permit modifications require notice to interested parties, opportunity for public comment, and a hearing and approval by the regional board. All of the transaction costs of requiring regulatory preapproval are exacerbated here, including costs stemming from delay, studies, and paperwork. In addition, transfers under this approach may draw unfavorable publicity, for there is a public forum for environmental or other interest groups to express their concerns. It might be possible for regulators to reduce these transaction costs by somehow expediting the approval process. For example, an agency could make findings that certain trades are presumed “safe” trades, either because of their small quantity or limited area. These findings would be similar to the additional permit limitations used to protect against hot spots under other trading approaches.

For trades under these limits, a lesser showing would be required for regulatory approval. The statutory hearing provision requires the discharger or the regulatory agency to respond to objections raised in public comments, however, and the hearing process itself imposes a great burden. Thus, unless the hearing process can be made less burdensome or avoided altogether, regulatory approval will not be expedited. Even making findings of presumptive safety may not be allowed under the current CWA provisions, while avoiding a hearing under all or certain circumstances would almost certainly be possible only with statutory changes.

380. See supra notes 167-68 and accompanying text.
381. Interview with Einer Elhauge, Professor, School of Law (Boalt Hall), University of California, Berkeley, in Berkeley, Cal. (Mar. 17, 1993) (on file with author).
382. 33 U.S.C. § 1313(d)(4) (1988). This assumes that the requirement to show cause for modification is either deemed to be met or removed. 40 C.F.R. § 122.62 (1992). See supra notes 182-87 and accompanying text.
383. See supra note 193 and accompanying text.
384. See supra note 195 and accompanying text.
385. Instead of presuming that a trade which does not meet the limits is not acceptable, this presumption would assume that a trade which does meet the limits is acceptable.
Assuming the environmental impacts of every proposed trade are evaluated individually, there is less need for additional permit limits to protect against hot spots under a tradable limits approach. Legally required limits can be set relatively high and concentration limits arguably are unnecessary. In defining the market for tradable limits, as with allowances, tradable discharge reductions either could be limited to those made by refineries, or an offset type approach could be used. With offsets, reductions in load allocations from nonpoint sources (or reductions in WLA’s from nonrefinery point sources) could be obtained by a refinery and used to justify increasing its permit limit. All the same issues with regard to quantification, approval, and monitoring arise here as with including nonrefinery discharges under other trading approaches.386

The discussion of each of the trading approaches has thus far assumed the market would function continuously. As noted, however, a hearing process ensuring that individual trades will attain water quality standards imposes a substantial administrative burden on regulators. This burden could be reduced by reopening the WLA’s on a periodic basis only.387 At the time of reopening, all of the WLA’s would be open to modification, and all the proposals would be considered together to determine net environmental impacts. The frequency of the market could be set according to the refineries’ planning period, keyed to changes in economic or environmental conditions, or determined on the basis of both factors. Such a periodic market would significantly constrain trading and would essentially eliminate any short-term trading. Given the substantial transaction costs in a tradable limit approach, however, it is unlikely that smaller-scale, short-term trades would be worthwhile anyway, so this additional constraint may be relatively unimportant. Moreover, a periodic reopening of the WLA’s might provide a convenient opportunity for reevaluating environmental conditions and adjusting the total annual mass emissions permitted. Otherwise, this system’s capacity to respond to changing conditions should be no different from that of a mass emissions approach.

Tradable limits can be viewed as an intermediate step towards a full-fledged trading system, which would allow the assessment of potential cost savings, regulatory burdens, and environmental impacts of trading, without relinquishing much regulatory control. Yet it is likely “that a half-hearted use of market incentives will not achieve the goals of more cost-effective regulation and technological innovation.”388 If a real market is not established, trading will not occur, and the experi-

386. See supra parts II.B.1.c, II.B.1.d.
387. Telephone Interview with David Smith, supra note 122.
388. Dwyer, supra note 45, at 112.
ment cannot serve as a proxy to determine the probable success of regulating discharges through a fully developed trading system.

III
IMPLICATIONS FOR DESIGNING A DISCHARGE PERMIT TRADING PROGRAM

The previous part laid out a detailed structure for analyzing the possibility of using discharge permit trading in a particular problem context. Based on earlier identified necessary conditions for using a trading approach, the analysis illustrated how to evaluate likely constraints on trading. This part takes the analysis a step further to link trading program design explicitly with identified characteristics of the problem context. Part III.A presents some general observations about the strengths and weaknesses of the different types of trading systems. Part III.B then discusses the interplay between the characteristics of the problem context and the selection and design of a trading system, seeking to provide a simple model for decisionmaking. Again, this model is illustrated by discussing choices about trading in the San Francisco Bay Estuary. The first subpart briefly restates the previously identified probable sources of constraints on using trading to control selenium, as well as some more generally applicable likely constraints on discharge permit trading. The next subpart considers variations in trading systems that would best respond to the constraints exhibited in the bay problem setting.

A. Relative Strengths and Weaknesses of Different Trading Systems

Market mechanisms are often discussed at a fairly high level of generality, with potential cost savings weighed against the administrative burden of implementing a new regulatory system. The success of an attempted trading program in a given problem context is far more likely to turn on the way specific design details of the program interact with facets of the problem setting. Key criteria for evaluating different types of trading programs are their legal feasibility, economic efficiency, regulatory administrability, and political acceptability. This subpart evaluates the strengths and weaknesses of the different trading systems using these criteria.

1. Legal Feasibility

The CWA statutory and regulatory requirements for water pollution regulation may significantly constrain the trading options currently available. Uncertainty as to what is legally authorized dampens regulatory enthusiasm even for arguably legal experimental
approaches. However, the current legal context is less of a concern in assessing whether tradable discharge permits would be a useful regulatory tool in general, since this context can change and is likely to do so. Tradable limits is the only trading approach explicitly authorized by the CWA, although with potentially crippling constraints. An allowance approach and group permit approach are arguably within regulatory discretion, but would be open to legal challenge in light of the statutory ambiguity. New regulatory approaches may present grounds for a legal challenge by environmental groups or industry, and the weaker the statutory support, the greater the disincentive for experimentation. Yet, experimental permit designs supported by higher-level state regulators, and more importantly, approved by EPA, are likely to receive judicial deference. Local regulators may be more willing to innovate, and are likely to be more successful in withstanding legal challenge, where there is strong regulatory consensus in favor of the policy.

2. Economic Efficiency

The major justification for using market incentive approaches for pollution control is to achieve the same or more environmental benefits as under traditional regulatory approaches, at substantially lower cost. Unless these cost savings are realized, there is little reason to use a tradable permit system. The realization of cost savings depends on the existence of different marginal control costs across dischargers and the development of a competitive market in the tradable permits. The economic effectiveness of different trading approaches can be

390. See supra part II.A.3.b.i.
391. See supra notes 172-73 and accompanying text.
392. These legal ambiguities become less of a problem where there is broad political support for the program. This support reduces the likelihood of legal challenge regardless of how well the statutory authorities support the program.
393. EPA approval is required for permit or basin plan amendments. See supra note 125 and accompanying text.
394. The selenium problem presented here may be somewhat unusual in the sense that numeric ambient water quality standards for selenium are being met, but additional controls are proposed to protect beneficial uses of the water segment. See supra note 136 and accompanying text. Courts might accord regulators greater discretion for imposing these additional controls than they would require for a segment that is clearly in nonattainment for both numeric and narrative water quality standards. Not surprisingly, most of the existing water discharge trading programs address hitherto unregulated nonpoint source pollution problems, where statutory requirements are more open-ended. See supra notes 30-32 and accompanying text.
395. Tradable permits also allow dischargers greater flexibility. This may be valued for its own sake as well as for the cost savings it produces, but it is not widely regarded as the incentive driving interest in trading systems.
compared on the basis of their ability to create and sustain the market in pollution reductions.\textsuperscript{396}

Using this criterion, allowances and group permits are clearly preferable to tradable limits. Regulatory restraints on trading and the resulting transaction costs are a significant barrier to market development.\textsuperscript{397} To the extent that a tradable limits approach retains the most regulatory control over transactions, it is likely to achieve the least, if any, cost savings. Whether a group permit or an allowance system would result in lower barriers to trading could turn on specific design elements of each system, although allowances probably have the potential to create a more competitive market.

Trading activity under a group permit will relate to whether there are enforceable limits and how easily these limits can be changed. A group permit system would hold individual facilities responsible for complying with the terms of an agreement on discharge allocations incorporated into the permit, but leave development and amendment of the agreement up to the dischargers. In order for the discharger-negotiated allocations to be cost effective, however, no single discharger should be able to exercise market power in setting the terms of the agreement. While to some extent the regulated entities' bargaining power is a function of the existing market conditions, regulators may be able to promote balance in bargaining power among the regulated entities.\textsuperscript{398} Whether a group permit approach achieves the most cost-effective pollution control will also depend on enforcement. Regulators should structure enforcement under a group permit program so that the individual discharger responsible for a violation bears the full cost of the sanctions and any associated legal costs of determining responsibility.\textsuperscript{399}

By commodifying discharges, allowances create the most straightforward market in emissions reductions. The efficiency of this form of trading system could be undermined by regulatory constraints such as transaction preapproval, overly restrictive trading areas, or additional permit limits. A relatively unfettered allowance trading system, how-

\textsuperscript{396} To compare the cost-effectiveness of different trading systems, one must balance the costs of developing and administering a trading system, the cost savings it produces, and any differences in environmental benefits obtained; here it is assumed that environmental outcomes would be substantially identical. Administrative costs are discussed in the next section. See infra part III.A.3.

\textsuperscript{397} See supra notes 50-52 and accompanying text.

\textsuperscript{398} For efficiency purposes it is not necessary that an initial allocation or the ground-rules for bargaining be equitable, only that they preclude market distortions. Whether the initial playing field is "level," however, will affect distributional impacts and who wins and loses under this approach.

\textsuperscript{399} Otherwise, if a portion of these costs is borne by the group, the full price of excess emissions is not internalized to the discharger. This gives each discharger an incentive to emit above its allocation.
ever, should create optimal conditions for cost-effective reallocation of control efforts. The potential for successful application of allowances is supported by experience; the most successful trading program was the relatively free market in lead rights, which are analogous to pollution allowances.400

3. Regulatory Administrability and Environmental Effectiveness

One administrability concern is the regulatory effort involved in developing and implementing a new program. A second concern is the degree of environmental protection achieved. Prospective assessment of these issues will be a major determinant of regulators’ and environmentalists’ attitudes towards trading proposals, since environmental effectiveness, not cost efficiency, is the primary goal of both of these interests.401

The administrative cost of initially developing a trading program roughly corresponds to how much the trading program departs from the existing system and the complexity of the replacement. Allowing some transfers of WLA’s should not impose much of a regulatory burden in terms of developing the rules for trading, although doing the initial WLA is a major regulatory effort.402 A group permit represents more of a change from the current approach; it would require greater development efforts. Yet, the group permit approach also uses existing and familiar regulatory instruments, applied in a new context. Allowance trading, however, would require more regulatory effort, particularly since there are few precedents; there is no easily adaptable model for a point source trading program. The administrative burden is likely to be a major reason for regulatory resistance to an allowance system.

The relative burdens of the trading systems shift when the systems are put into operation. Once the groundrules are in place, an allowance system does require tracking trades. Since the monitoring and reporting systems are already operating, however, this additional regulatory responsibility may not require much effort. The day-to-day operations of group permits might be even easier to administer than

400. It is too soon to declare the acid rain program a success or failure, given that much of the trading is expected to occur in the second round of reductions, which begins in the year 2000.

401. While the administrative burden of developing a new program might not appear to be a primary concern of environmentalists, they are well aware that regulatory resources are limited; resources devoted to a trading program are lost to other regulatory efforts, such as enforcement, deemed vital to environmental protection. See, e.g., Telephone Interview with Greg Karras, supra note 211.

402. Some sort of WLA would have to be done under an allowance system as well, in order to distribute allowances among the dischargers (unless allowances are auctioned), but allowance allocations might be derived on a less rigorous basis, short of establishing precise loading estimates for all inputs to the water body.
allowances, depending upon the degree to which regulators would police individual discharges as well as aggregate mass emissions. Tradable limits, by contrast, would be burdensome to operate, since each trade requires extensive involvement by regulators in the hearing and approval process. Indeed, the more successful the tradable limits market, the higher the demand on regulators' time and effort. The administrative burden could be lessened by opening the market only on a periodic basis, but this would further reduce the potential cost savings.

Ease of enforcement under either allowances or tradable limits should be comparable to the current system. Reported discharges are compared to permit limitations (either numeric or reported number of allowances held), and sanctions are imposed for violations. Group permits may be somewhat more difficult to enforce, due to uncertainties regarding individual control responsibilities.

The environmental effects of the approaches should be similar. In the absence of a requirement for regulatory preapproval of trading under group permit or allowance approaches, some hot spot problems could result. These could be guarded against by establishing additional permit limits or trading rules. Particularly given a small number of players and strong existing regulatory oversight, there seems little risk of a market going "out of control" and causing detrimental environmental impacts. Nonetheless, a perception that the reduced regulatory control under an allowance approach risks environmental degradation motivates some environmentalists' opposition.

4. Political Acceptability

Political acceptability will be key in determining whether a trading system will actually be developed and approved. Where trading is considered, it is often proposed as a compromise solution when high control costs and a serious environmental problem have led to a political stalemate over the best regulatory response. Years of political battles and extreme stances have often polarized the debate and engendered distrust in such situations, however, making it difficult to achieve agreement on an unproven compromise approach. Under these circumstances, who proposes trading, as well as what they propose, may affect the reactions different groups have to the proposal.

403. While regulators would have to evaluate trades, trading entities would be responsible for gathering any necessary data.
404. These costs could be defrayed by imposing some sort of fee to cover administrative expenses, but this would further discourage trading by raising transaction costs.
405. See supra notes 55-64 and accompanying text.
406. Examples include the acid rain program and RECLAIM.
CONTROLLING SELENIUM DISCHARGES

A tradable limits approach probably generates the least political opposition overall, although it also lacks a strong constituency. Allowing tightly controlled trading risks little and gains little. Nevertheless, the familiarity of WLA's and the appeal of the concept of using economic incentives may generate some support for this approach.

The political acceptability of group permits may strongly depend on how they are implemented. Environmentalists, industry, and regulators are all likely to oppose a group permit that lacks a mechanism for enforcing individual limits. Environmentalists and regulators fear ineffectual enforcement, while industry hates the prospect of joint and several liability, since this means sharing the costs of violations by other firms or engaging in expensive litigation to recover contribution. Moreover, competing entities in a single industry may question the prospects of negotiating allocations among themselves. To the extent that a group permit provides a mechanism for individual limits to be reallocated without regulatory preapproval, however, this may be viewed relatively favorably by industry, if not by environmentalists.

Allowance trading is likely the most politically controversial approach. Allowances are viewed as the most experimental, with the least certain environmental and economic results. Insofar as political opposition stems from the uncertainty of an innovative approach, as opposed to a fundamental dislike of the concept, allowance trading should become more politically acceptable after some successful applications. If allowance trading in other areas is perceived as a failure, however, as might happen with the acid rain program, the political fallout may affect the acceptability of using allowances in the water context.

B. Matching Trading System Design to a Problem Context

The approach to trading in a particular problem setting should take advantage of the strengths and weaknesses of different trading systems. Given an identified set of constraints on trading, both general across possible applications of discharge permit trading and specific to the particular problem, the trading system should be designed to minimize those constraints. The complexity of regulating any environmental problem through trading is such that there will rarely be an obvious design choice, for different factors cut in different ways. Nev-

407. See, e.g., James Dao, A New, Unregulated Market: Selling the Right to Pollute, N.Y. TIMES, Feb. 6, 1993, at A1, A25. This article criticized the allowance trading program for the potential negative environmental impact of individual trades, but scarcely acknowledged the overall 50% SO₂ reduction made possible by the legislative compromise allowing trading. See id.

408. Environmentalists dealing with the selenium issue already cite their experiences with the air medium as a reason for opposing permit trading here. Telephone Interview with Greg Karras, supra note 211; Telephone Interview with Trish Mulvey, supra note 62.
Nevertheless, clearly identifying the possible options and the reasons for certain choices enhances informed decisionmaking.

1. Likely Constraints on Trading in the San Francisco Bay Estuary

The selenium problem in the San Francisco Bay Estuary exhibits two of the necessary characteristics for successful use of permit trading: (1) an environmental problem susceptible to trading, and (2) the possibility of different marginal control costs across dischargers. Nevertheless, trading may produce only modest cost savings in this context. Moreover, it may be difficult for a trading program to be developed and approved here. The greatest underlying barriers to using trading in the San Francisco Bay Estuary probably are insufficient economic incentives, legal ambiguity, and lack of a strong political constituency. These are likely to be the most common constraints on point source discharge permit trading in many problem settings.

First, although marginal control costs differ across dischargers, the cost savings from reallocating selenium discharge quantities may be insufficient to motivate trading. Regulators are likely to require such high reductions that the individual refineries may be either technologically unable to make further reductions for trading or else face such exorbitant costs that it is unprofitable to do so. Thus, a lack of reductions available for trading is probable. If regulators implement a phased reduction schedule there may be real cost savings to reallocating reductions in the short-term. Trading might also be cost effective if a super-efficient technological solution is developed. Yet, the small number of dischargers also limits the savings available, since this restricts the quantity of discharges available to reallocate, particularly in a given time period, and reduces the competitiveness of the discharge permit market.

Another potential problem here, and for water pollution trading generally, is the ambiguity of statutory authority. Regulators are understandably reluctant to risk costly and time-consuming litigation, as well as the possibility of a humiliating judicial reversal. Although some form of trading is probably within regulators’ discretion under the CWA, specific statutory and regulatory requirements for control programs may substantially restrict flexibility in designing a trading system. The restrictions may even prevent a market from developing, due to transaction costs imposed by regulators.

Finally, there may not be any party who wants trading enough to get it approved. Industry is the logical source of support for cost-

409. See supra part II.A.4.
410. See supra part II.A.3.
411. See supra part II.A.6.
effective regulation, but the refineries are largely apathetic. The pro-
fessed reason is minimal cost savings, which if true, is sufficient to
explain their indifference. Another motivation may be that the one
thing better than cost-effective regulation is no regulation, so the re-
fineries are focusing their resources on stalling stricter regulation.
Moreover, it is easier to achieve delay if control costs can be por-
trayed as prohibitive. Regulators appear the most likely source of
support for trading in the San Francisco Bay Estuary, because trading
has the potential to achieve environmental goals at lower economic,
and hence political, cost. Yet regulators also have possibly conflicting
goals. The desire to retain areas of professional expertise under the
existing regulatory system, to minimize administrative burdens and
maximize regulatory resources, and to achieve environmental goals
with certainty militate against the adoption of an experimental regula-
tory approach. Without substantial external support trading may not
advance regulators' goal of minimizing political upheaval. Finally,
while EDF will support trading approaches (and where EDF is fully
involved, its support can be substantial), other environmental groups
tend to be dubious or downright hostile. Environmental opponents
consider trading to be a waste of scarce regulatory resources, environ-
mentally risky, and morally wrong. While opposition may not be so
strenuous as to doom a trading proposal, a lack of strong support
could be equally fatal.

Most of these constraints will arise in water pollution contexts
other than the San Francisco Bay Estuary. Lack of an appropriate
physical environmental problem or different marginal reduction costs
will winnow out a number of settings where point sources cause water
quality problems. Legal ambiguity in the CWA and state laws and
regulations is a universal constraint on any attempt to implement per-
mit trading for point sources, but it need not bar trading initiatives
altogether. The absence of a strong constituency and the presence of
active opposition are likely to be fairly pervasive problems, particu-
ardly where economic incentives are insufficient to attract much indus-
try support. Another factor not exhibited in the selenium context, but
likely to arise elsewhere, is failure to realize potential cost savings due
to an unwillingness of firms to trade excess credits.412 While recogniz-
ing that these will be the most common negative factors for applying

412. See Dwyer, supra note 37, at 65. A major incentive for hoarding is lack of trust in
the ongoing viability of the market (i.e., a fear that credits will not be available, or will be
exorbitantly expensive in the future), coupled with a distrust of regulators. Future unavail-
ability of credits is most likely to result from regulators tightening the supply by decreasing
allocations, or devaluing credits, or from regulators increasing demand by setting stricter
standards.
point source permit trading, none of them need necessarily prohibit trading in every context.

2. Responding to Constraints Through Trading Program Design

Assuming for the sake of illustration that none of the constraints identified above were considered sufficient to bar a trading approach, their effects nonetheless should be mitigated to the extent possible, through the choice of a trading system. The weak economic incentives for trading here support using an allowance system, since this approach would minimize any further disincentives. Generally, the system should be designed to enhance market viability through the variety of options identified earlier. In particular, if the market is too small, the trading system might be designed to incorporate additional selenium sources into the market. If regulatory resistance and legal challenges make legal ambiguity the largest obstacle to trading, however, then tradable limits would be preferable since this approach has a better statutory foundation. Finally, if political support is considered a vital missing element, trading system design choices should be made to mitigate the concerns of important constituencies. Here, for example, creating a very flexible system makes sense if that would make the refineries active supporters. If the refineries would remain apathetic regardless of such efforts, it might be preferable to focus on building in regulatory safeguards to assuage the concerns of environmental opponents.

CONCLUSION: FOCUS OF FUTURE EFFORTS

The constraints on trading identified throughout this comment are not immutable. Efforts can be made to create a problem context more conducive to implementing a discharge permit trading approach to pollution control. Also, trading tools can be further developed to expand their capabilities. Action to promote permit trading could be taken at a variety of levels, including specific statutory authorization and regulatory reduction of the requirements burdening a trading program. Various activities such as public education or developing model trading regulations would increase the political viability of trading by raising understanding, reducing uncertainty, and alleviating the regulatory effort required to initiate a trading approach. These activities could be taken by any interested party: governmental, nonprofit, or industry.

To the extent that legal uncertainties discourage regulators from developing trading programs, and specific statutory and regulatory requirements constrain the design of those programs, statutory clarification and amendment is vital to encourage trading. The CWA is
currently up for reauthorization and there is a continuing strong governmental interest in market incentive approaches to environmental policy.\textsuperscript{413} There certainly will be an opportunity to amend the CWA either to authorize explicitly some form of point source discharge trading or to provide sufficient regulatory discretion for state and local regulators to experiment with different regulatory approaches.

Congress should make several statutory changes to remove the constraints that the CWA currently may impose on point source trading programs.\textsuperscript{414} It should clarify that permit limitations need not be expressed as concentration limits. This disclaimer is probably legally unnecessary, but it would increase regulators' confidence about removing such limits to implement mass emissions-based regulatory approaches. Other statutory changes that would facilitate trading include allowing permit limitations to be expressed in terms of allowances held, and allowing daily and monthly limitations to be waived where they are inappropriate to the regulatory control strategy applied and the character of the pollution threat. The statute also could explicitly authorize the use of bubbles (group permits) and offsets, subject to necessary environmental safeguards, in regulatory areas such as attainment of water quality standards and nonpoint source pollution control. Anti-backsliding regulations should be amended to allow changes in individual permit standards where overall water quality impacts are neutral or beneficial. Finally, the CWA could allow permits to be modified to reflect transfers without requiring a full notice and comment hearing, subject to a specific or general finding that transfers meeting certain parameters would have no detrimental environmental impact.

Some of these changes could be made through regulatory action under existing statutory authority by modifying permitting regulations. Thus, EPA could act to facilitate discharge trading prior to congressional reauthorization of the CWA. All of these modifications would increase regulators’ support for trading and would encourage the design of trading systems more likely to realize cost savings.

A nonprofit or industry trade association could facilitate the use of trading approaches by developing model trading regulations that address regulatory concerns. Model regulations would reduce the agency resources required to consider and develop trading options. Also, they would indicate external political support and, depending upon the source, could provide intellectual authority for this policy approach. Trading proponents could also encourage the use of trading approaches by helping to change public attitudes about permit trad-

\textsuperscript{413} Hahn & Stavins, supra note 41, at 20-26; WATER POLLUTION, supra note 29.
\textsuperscript{414} See supra part II.A.3.
ing. Public education efforts should address objections raised against permit trading, as well as unfamiliarity with trading. Concerns will vary according to the interests and objectives of a given group. At a minimum, public education efforts must address the general public's ignorance of the concept of permit trading, environmentalists' concerns about operation and enforcement, and industry's suspicion that new regulations mean more regulatory hassle, which is likely to offset any potential cost savings.

The use of discharge permit trading for point sources to achieve water quality standards would apply a proven policy tool in a new setting. Trading is by no means appropriate for every situation, even in the limited context of water quality nonattainment. Trading is not even clearly authorized under the current CWA statute and regulations. Nevertheless, permit trading is one of few opportunities available to achieve environmental goals at a lower economic and, thus, political cost. Resources for environmental protection are not growing, at least not in relation to the rising number and increasing scope of environmental problems demanding our attention. Against this background, economic incentive approaches are receiving heightened and deserved attention. Point source discharge permit trading is one application of this concept with the potential, given further development, to achieve equal or greater environmental protection at reduced cost.