Optimal Policies for the Control of Non Point-Source Pollution in a Second Best Environment

by

Marca **Weinberg**[1]
Catherine L. Kling
James E. **Wilen**[2]

May 7, 1991

---

[1] Weinberg is an economist with the Resources and Technology Division, Economic Research Service, USDA.
The opinions expressed in this paper are those of the author and do not necessarily represent those of the USDA.

[2] Kling and Wilen are Assistant and Full Professors of Agricultural Economics, University of California, Davis, respectively.

Optimal Policies for the Control of Non Point-Source Pollution in a Second Best Environment

Policy prescriptions for the control of non point-source pollution have been examined by a variety of authors (e.g. Griffen and Bromley, Shortle and Dunn, Segerson). Implicit in most models and analyses has been the assumption that the externality represented the only source of market failure in the sector being analyzed. However, in cases involving agricultural non point-source pollutants there are likely to be distortions other than those generated by the externality that effect the input or output markets for the primary product. Crop prices may be supported through commodity programs or marketing orders, input prices may be distorted or input use may be controlled through quantity restrictions and use regulations. Irrigation water supplied by the Bureau of Reclamation is an example of a case in which a primary input is provided at a subsidized price, but in limited quantities. Subsidized electric rates, nitrogen fertilizer taxes, cost sharing for soil conservation practices, and pesticide use regulations provide additional examples of input market distortions. The general theorem of the second best (Lipsey and Lancaster) suggests that the use of first-best policy tools to correct production externalities under these circumstances may actually move the economy away from a second-best optimum, rather than nearer to it.

Several authors have addressed the second-best conditions created by output market distortions (e.g. Lichtenberg and Zilberman, Buchanan), but relatively little attention has been paid to input market distortions. This paper develops a conceptual framework for examining non point-source control in a setting characterized by input market distortions that are particularly relevant to irrigated agriculture. Two conditions giving rise to a second-best problem are examined: input price distortions and institutionally set quantities of the input. These conditions describe surface-water supply institutions in much of the western United States. In addition to its importance as a primary input to crop production, irrigation water is also a primary input in the generation of agricultural externalities in many areas.

The implications of input market distortions on non point-source problems are examined first in a single firm context. The analysis is then extended to multiple firms to examine optimal policy tools when the input price and quantity distortions differ between firms. Finally, an empirical example is presented that explores policy options for addressing the agricultural drainage problem emanating from irrigation activities in California's San Joaquin Valley. Prices and quantities of water delivered to farmers in the study area are set by the Bureau of Reclamation and do not reflect market forces. Hence, non-point source emissions with distortions in the market for an input (water) characterizes this problem.

Conceptual Framework

In this analysis, profit maximizing farmers are assumed to jointly produce crops and a volume of effluent through the use of two inputs, a polluting input (x) and an abating one (z). For purposes of discussion, input x may be thought of as irrigation water while z may represent irrigation technology or management. Farms face a maximum constraint $(\bar{x})$ on the quantity of the polluting input used and receive that quantity at a distorted price (w^i). The optimization problem for a representative farm is to choose input levels to:

$$\begin{aligned} &Max\ \pi^i = pf^i(x_i, z_i) - w^i x_i - rz_i \\ &subject\ to:\ x_i \le \bar{x}_i, \end{aligned} \tag{1}$$

where π^i represents net returns to land and management for the i^{th} farm, p is the price received for crop output, $f^i(\cdot)$ describes crop production opportunities for the i^{th} farm and is assumed to be twice differentiable and concave, and r represents the marginal cost of the abating input (z). The social objective in this scenario is to maximize social welfare (S) defined as the sum of farm-level net returns less the social damages incurred as a result of the production externality:

$$Maximize\ S = \sum_i (pf^i(x_i, z_i) - w^s x_i - rz_i) - p_d D(g^1(x_1, z_1), \ldots, g^I(x_I, z_I)) \quad (2)$$
$$subject\ to: \ x_i \le \bar{x}_i \quad \forall i,$$

where w^s represents the "true," per unit cost of supplying the input x (the supply price), $D(\cdot)$ represents total effluent levels, p_d represents the marginal social cost of damages from emissions, and $g^i(\cdot)$ is a production function describing non-point source emissions from the i^{th} farm. The non-point source production function is assumed to be twice differentiable and convex, and to increase with input x and decrease with levels of z, i.e. $g_x > 0,\ g_z < 0,\ g_{xx} \ge 0,$, and $g_{xx} \ge 0$. Subscripts on functions denote partial derivatives. In addition, from here on it is assumed for simplicity that total damages are an additive function of farm-level emissions, i.e. that $D = \sum_i g^i(x^i, z^i)$.

Non point-source control studies often suggest the use of input taxes to motivate optimal behavior (Griffen and Bromley, Stevens, Pfeiffer and Whittlesey). Under the assumptions that the input constraint is strictly nonbinding and that all farms face the true cost of water (i.e. $w^i = w^s, \forall i$), it can easily be seen that a set of input taxes $t_x = p_d g_x,\ t_z = p_d g_z$ will motivate socially optimal behavior on the part of farms.

Modeling policy choice for markets with resource constraints is straightforward when it can be assumed that the nature of the restriction is invariant to the parameters of the problem, e.g. if a constraint is binding before a price change, then it will be binding for the regulated firm. However, the impact of a policy may be such that movement on to, or off of, the constraint is generated. A taxonomy of "regimes" can be defined to describe these alternatives, as shown in Table 1. The cases in which the constraint is always binding (regime 1), as is generally the case with fixed inputs, or is nonbinding (regime 2) lend themselves to traditional analytic methods. Two additional regimes are necessary to describe a switch in conditions: regime 3, in which the constraint is binding prior to government regulation of the externality, but not binding afterward, and the opposite case in which the constraint is initially nonbinding, but binds at the optimum as a result of a policy (regime 4). Comparative statics results cannot be used to analyze the case of inequality constraints or discrete changes in the values of policy instruments, as is necessary to address the conditions implied by the latter two regimes. Policy analysis should, however, be conducted with respect to all four regimes. A different approach is therefore required.

Table 1. Taxonomy of Policy Regimes

		Solution Incorporating Externality	
	Input constraint:	Binding	Not Binding
Pre-policy Solution	Binding	Regime 1	Regime 3
	Not Binding	Regime 4	Regime 2

Single Firm Analysis

Input price distortions can be analyzed by assuming that a representative firm faces a private input price of $\mathbf{w^p}$ $(\mathbf{w^p} \neq \mathbf{w^s})$. In the irrigated agriculture example, $\mathbf{w^s}$ may represent the cost to the Bureau of Reclamation of producing and delivering an acre-foot of water, while $\mathbf{w^p}$

represents the price paid by farmers to the Bureau. The definition of w^s does not incorporate the external costs associated with the use of the water (these are addressed explicitly in this analysis) though it may include external costs associated with development of water **supplies.**[1] Farmers often pay less than full cost for water deliveries due to the subsidies that are built into the Bureau's pricing structure, so that $w^p < w^s$. In addition, farmers are allotted a fixed quantity of water per acre per year.

The level of input prices plays an important role in determining whether the input constraint is binding in a given problem and, therefore, whether the constraint is relevant for analysis of non point-source problems. This relationship suggests an alternative approach in which a virtual price $w^v(p, w, r, \bar{x})$, defined as the price at which the constraint would be "just" binding, is specified. In this example, w^v represents the price at which the firm's demand for input x would equal $\bar{x}$ and is implicitly defined by $x(p, w^v, r) = \bar{x}$. The work of Neary and Roberts illustrates the use of virtual prices to model consumption when some commodities are rationed. Analogously, w^* is the level of the input price that would induce the firm to choose the socially optimal level of the input and is implicitly defined by $x(p, w^*, r) = x^*(p, w^s, r, p_d)$.

The taxonomy of regimes and the response to policy instruments that can be expected from the firm under each regime can be depicted graphically. The firm's demand for input x is labeled $VMP|_z$ (Figure 1). The social cost function is defined as the private marginal cost of the input plus marginal social damages, i.e. $MSC|_z = w^s + p_d g_x$. The inputs are assumed to be

[1] External costs of developing water supplies include degradation of wildlife habitat, declines in fisheries, and loss of white water recreation. For a more complete description of these costs see Willey.

technically independent ($f_{xz} = g_{xz} = 0$) since without this assumption a family of curves is necessary to depict social marginal costs and private marginal benefits in a two dimensional figure. The marginal product in the input demand function and the marginal effluent product in the social damage function are both evaluated at $z = z^*$. The unconstrained (first-best) social optimum for x is denoted by x^* and is defined by the intersection of the value of marginal product and marginal social cost curves, i.e. where $pf_x = w^s + p_d g_x$, as is required by the social first order conditions (derived from equation (2)). An unconstrained firm would choose $x = x^p$ in the absence of policy intervention. However, the input constraint is binding so the firm chooses $x = \bar{x}$.

The theory of the second-best suggests that attempts to apply first-best tax rules to correct an externality may be suboptimal when other distortions are present (Lipsey and Lancaster). A contrapositive is also suggested: if one can address all distortions, then optimal taxation rules *will* be effective. Therefore, it is equally important to consider the potential for policy tools to restore a first-best state as to examine the implications of first-best tools under second-best conditions. A distinction is made between first and second-best optima. In this paper, it is assumed that market distortions occur in two forms: (i) input market distortions including a price distortion and a quantity constraint, and (ii) an environmental externality. The first-best, or social, optimum is defined without regard to the market distortions while second-best optima result from maximization of social welfare subject to the condition defining the second-best nature of the problem.

Figure 1 depicts a constraint that is binding in both the social and private optimization problems (regime 1). Traditional prescriptions for attaining a social optimum suggest an input tax on x equal to $p_d g_x$, the value (cost) of the marginal effluent product. This is the distance $w^* - w^s$ in Figure 1. The second-best optimum is attained at $\bar{x}$ when the constraint is immutable. This is the point at which net social benefits are maximized subject to the input constraint.[2] The firm is already using the second-best optimal quantity of x in this case and, because the virtual price is greater than that implied by the true social cost (w^*) of using the input, it will not respond to a tax reflecting marginal social damages. Neither the constraint nor the externality are policy relevant in this case. The price distortion in the presence of a binding input constraint serves only as an income transfer and in the short run does not influence the firm's decision regarding the quantity of input used when w^p and w^s are both less than w^v, regardless of whether the private price reflects a tax or a subsidy.[3]

When the input constraint is not binding at the social optimum only policy instruments that address both the price distortion and the social damages associated with the use of the input will assure first-best optimality, regardless of whether the constraint is binding in the private problem (regimes 2 and 3). Figure 2 illustrates this result for the case that the constraint is binding at the private (pre-policy) optimum. The diagram is

---

[2] For diagrammatic purposes, the optimal level of the abating input (z) is implicitly assumed to be invariant to the level of the polluting input (x), implying that $f_{xz} = g_{xz} = 0$. However, it is likely that the optimal level of z will vary with the constraint on input x, i.e. $z^*(p, w^s, r, p_d) \neq z^*(p, w^s, r, p_d; \bar{x})$, requiring a policy instrument to induce the firm to use the optimal quantity of the abating input. The results regarding the constrained input remain valid when $f_{xz} \neq 0$ and $g_{xz} \neq 0$, provided that the cross input effects are small enough that the constraint remains binding after the shift in the value marginal product and marginal social cost curves.

[3] Though taxes and subsidies have identical implications on the margin, firm entry and exit decisions may be different under a price subsidy than with a tax. The long run effects may therefore also differ (see Spulber).

similar to Figure 1 but depicts a constraint at an input level that is greater than socially optimal. The unregulated firm chooses input level $x = \bar{x}$. Policy intervention is required to assure social optimality in this case.

The firm chooses x^*, the optimal level for input x, only when faced with a price of w^* (Figure 2). The policy instrument must therefore increase the cost to the firm of using input x from the private marginal cost (w^p) to the true shadow price (w^*). This distance is composed of two parts: a) the price distortion, illustrated as the distance from w^p to w^s, denoted t_w, and b) the distance from w^s to w^*, which represents the marginal cost of the externality at the optimum and is denoted t_x^*. The optimal tax on input x in this case is a composite one that includes a correction for the price distortion and a term equal to the marginal social cost of input use to correct for the externality: $t_x = (w^s - w^p) + p_d g_x$. An input tax that does not account for both problems (setting $t_x = t_x^*$ for example) will motivate the firm to move towards but not to the optimum level for input x. Furthermore, if the true private cost is artificially high, i.e. $w^p > w^s$, then ignoring the price distortion and setting $t_x = t_x^*$ will cause the firm to move beyond the social optimum to a level $x < x^*$.

An alternative to input taxes to address non point-source problems is a tax on estimated effluent levels. When input prices are distorted, it is not possible to specify an effluent tax that will, by itself, assure a first-best optimum. Two instruments are necessary to correct both the price distortion and the externality. A tax on the input x must be introduced in addition to a tax on estimated effluent levels. The form of these will be $t_w = w^s - w^p$ and $t_d^* = p_d$, respectively.

When the supply price is below the virtual price, but the private price is not ($w^s < w^v < w^p$) and the true cost, w^*, is less than w^v, the input constraint is not binding in the private problem but is binding in the social optimization framework (regime 4) and a second-best approach is required. This case occurs when the input is taxed, rather than subsidized, and is illustrated in Figure 1 where the firm faces an input price of $w^{p'}$. The second-best optimum could be achieved with a partial correction of the price distortion, such as subsidy equal to $w^{p'} - w^v$.

Regime 4 can also be examined with Figure 2 by allowing the constraint, $\bar{x}$, to represent a minimum rather than a maximum level for the use of input x. This situation may apply to farmers with appropriative water rights who also have an alternative water source such as groundwater. Appropriative water rights are often assigned on a "use it or lose it" basis and these farmers will use a minimum volume of water each year to maintain their rights. In the general case, the unregulated firm will select $x = x^p$, as it does under regime 3. An input tax $t_x^* = p_d g_x$ (illustrated as the distance $w^* = w^s$) will induce the firm to reduce the use of the input only to $\bar{x}$, the point that the constraint becomes binding. A tax set equal to marginal social damages ($p_d g_x$) will be unnecessarily large. The input level $\bar{x}$ represents the second-best optimum because the constraint is binding at the social optimum and a tax of only $t_x^* = w^v - w^s$ is sufficient to motivate optimal behavior.

In sum, an input tax imposed on the constrained input has no impact when the input constraint is binding at the social optimum (regimes 1 and 4). However, an input tax is an effective means of inducing the firm to consider the social damages associated with the input when the constraint is not binding at the social optimum (regimes 2 and 3). This will be true whenever the true shadow price (including the social cost) for the polluting input is greater than the virtual price, i.e. $w^* > w^v$. Policy intervention is necessary to induce optimal behavior when the constraint is not binding in the private optimum but is binding at the social optimum. However, the level of the policy instrument may be different from that specified as optimal under first-best conditions and found to be appropriate for regimes 2 and 3. These results suggest that the existence and location of an input constraint can be critically important for the policy maker and may have a direct effect on optimal policy choice.

Multiple-firm Analysis

The potential danger of ignoring the conditions that create second-best policy environments when addressing externalities is illustrated in the following example. Two firms or regions are assumed to contribute to a water pollution problem with emissions that arise from use of input (x). The firms are identical except that they face different institutional parameters related to the polluting input. Firm 1 receives a relatively low input allocation and pays a relatively high price per unit, while firm 2 receives a larger quantity of the input and a large price subsidy, i.e. $\bar{x}^2 > \bar{x}^1$, and $w^2 < w^1$. This example of heterogeneous institutional parameters reflects Bureau of Reclamation water supply policies.

The implication of variation in institutional parameters for externality control is examined in Figure 3, where firm 1 and firm 2 are depicted on the right and left sides of a back-to-back diagram. The net private marginal benefits ($PMB_j = pf_x^j - w^j,\ j=1,2$) and the social marginal benefits resulting from use of input x are illustrated for each firm. Social marginal benefits (SMB^j) equal net private marginal benefits minus the marginal social cost of using input x: $SMB_j = pf_x^j - w^j - p_d g_x^j,\ j=1,2$. Firm 1 does not receive a price subsidy in this example ($w^1 = w^s$), firm 2 faces a non-binding resource constraint, there is a one-to-one relationship between input use and effluent production ($g_x = 1$), and the inputs x and z are independent ($f_{xz} = g_{xz} = 0$). These assumptions are made for diagrammatic purposes only and the results remain valid for less restrictive assumptions.

The socially optimal level of input use by each firm occurs where the marginal (private) benefits are equal to marginal (social) costs, or when net social marginal benefits are zero. In Figure 3 this occurs at input levels x_1^* and x_2^*. However, firm 2 will chose the input level that sets net private marginal benefits equal to zero in the absence of policy intervention ($x_2 = x_2^o$). Firm 1 would like to do the same but may not use more than $\bar{x}_1$ units of x so that $x_1^o = \bar{x}_1$. The sum of inputs used in the private solution ($X^o = \bar{x}_1 + x_2^o$) generates negative net social marginal benefits.

The marginal effluent products of input use are the same for firms 1 and 2 in Figure 3 and the marginal effluent contributions are additive. The socially optimal level of total input use, given optimal levels for other inputs, is $X^* = x_1^* + x_2^*$. The policy maker must either mandate optimal input levels for each firm or must devise a tool that will reduce input use from the pre-policy level (X^o) to the optimal level (X^*).

As described in the previous section, an input tax or Pigouvian tax on estimated effluent can be effective in the presence of a resource constraint that is not binding at the social optimum, but neither will be an optimal policy choice when a persistent price distortion exists. For example, an effluent discharge tax set at p_d will motivate firm 1 to select x_1^* but firm 2 will reduce input use only to x_2' because its input price is subsidized. Total input use and effluent levels are higher than optimal in this case. This result is a consequence of the price subsidy and not merely due to the difference in input prices. If w^2 represented the true value of the input used by firm 2, x_2^* would coincide with x_2' and the Pigouvian charge p_d would achieve the optimal solution. Under this condition, optimal effluent levels are higher than those implied by X^*. When the price paid by firm 2 is subsidized and does not represent the true value of the resource, the social value of input use is lower than that implied by PMB_2 and the higher input use and effluent production by firm 2 are inefficient.

To achieve the optimal level of input use, and thus of effluent production, it is necessary to develop policy tools that will motivate firms to consider the social costs of input use given that they observe a distorted input price. Under these conditions, there are several regulatory schemes that policy makers might consider.

One set of options includes a tax, either on estimated effluent levels or on input use, that incorporates the shadow value of an effluent (or input) constraint set at desired levels. Suppose that achieving optimal effluent levels requires a fifty percent reduction in use of the polluting input ($X^* = .5X^o$), as illustrated in Figure 3. The first order condition for firm optimization under appropriate tax options requires that input use and effluent production be allocated among firms so that the net private marginal benefits per unit of externality are equal:

$$\frac{(pf_x^1-w^1)}{g_x^1} = \frac{(pf_x^2-w^2)}{g_x^2} = t_d. \tag{3}$$

The PMB_j curves in Figure 3, defined as net private marginal benefits per unit of *input,* also represent the net private marginal benefits per unit of externality under an assumption that $g_x = 1$ for both firms. This assumption is made for diagrammatic simplicity only.[4] An effluent tax level of $\hat{t}_d$ might be expected to achieve the fifty percent reduction objective at least cost, as follows from equation (3) and is illustrated in Figure 3. This is the familiar result that efficiency is achieved when marginal abatement costs are equilibrated among polluters (Baumol and Oates). The input allocations that arise from a policy of charging an effluent tax of $\hat{t}_d$ are denoted $\hat{x}_1$ and $\hat{x}_2$ in Figure 3 ($\hat{x}_1 + \hat{x}_2 = X^*$). An effluent tax of $\hat{t}_d$ does represent an efficient solution to the fifty percent reduction objective when there is no price distortion, i.e. when w^2 represents a true price, but may move the firms further away from the optimal solution when a price distortion is present.

Another possible method for attaining optimal input use is to require uniform reductions in input use among firms. In this scenario, both firms are required to reduce input use by fifty percent. The activity levels resulting from this uniform reduction scheme are denoted x_1^u and x_2^u.

The uniform reduction scheme causes firm 1 to use too little of the input and firm 2 to use too much, relative to optimal levels. However, as seen from Figure 3, a policy such as an effluent tax that equates net private marginal benefits (marginal abatement costs), rather than increasing efficiency relative to the uniform reduction, actually requires further reduction in input use by firm 1 and less reduction by firm 2. Total welfare is thus reduced under this policy.

The welfare changes associated with input allocations implied by a policy of equating net marginal benefits relative to a uniform allocation are illustrated in Figure 3. Area (acdf) represents the loss in welfare experienced by firm 1 as a result of the reduced input use, while area (ghkl)

[4] For example, a non-constant marginal effluent product can be incorporated in the diagram but will increase its complexity without changing the results.

is the true welfare gain to firm 2 from the greater input allocation. The net private welfare loss is area (acdf) - area (ghkl). The additional benefits that firm 2 would receive from the greater input allocation in the form of the price subsidy (area hijk) is an income transfer only, and is not included when measuring the efficiency gains or losses of a policy. The difference in net social welfare under the two allocation schemes is the sum of the areas between the social marginal benefit curves and the axis and between the input levels associated with each allocation, i.e. area (abef) + area (glmn). Areas (abef) and (glmn) represent negative values because social losses result from the reduced input use by firm 1 and from the increased input use by firm 2, Each input allocation represents the same level of total input use and effluent production in this example. As a result, the difference in (true) private net benefits (acdf-ghkl) is identical to the net social welfare loss (abef-glmn).[5]

Drainage Case Study

The presence and magnitude of welfare losses resulting from alternative policies introduced to address an externality in a second-best setting are examined with regard to an agricultural drainage problem in California. A brief description of the problem setting is presented next, followed by a summary of the model developed to simulate decision making in the area and results from simulations conducted under alternative drainage reduction policies.

Many of the West’s most valuable agricultural lands are naturally arid and have been made productive only through large-infrastructure water delivery systems. Developed water is typically sold to water districts under contracts

5 Area (acdf) = (abef) + (bcde) and (ghkl) = (nhkm) - (glmn). In addition, (bcde) = (nhkm) by symmetry. It follows that (acdf) - (ghkl) = (abef) + (bcde) - (nhkm) + (glmn) = (abef) - (glmn).

that specify the quantity of water to be delivered and the price per acre-foot. The terms of these contracts vary by district so that one farmer may receive a generous allotment at a relatively low price while a farmer in a neighboring district may be more limited in the quantity of water received and pay a higher price.

Increasingly, many regions are facing salinity and drainage problems. In these regions, as in arid regions throughout the world, irrigation water is applied in excess of crop water requirements to leach accumulated salts out of the root zone and to provide the minimum amount of water required by plants in all portions of non-uniform fields. In areas with limited natural drainage, this excess applied water contributes to regional saline high water tables that can cause crop yields to decline on overlying lands through upward capillary motion of salts and, in extreme cases, saturation of root zones. Artificial drain systems may be installed to maintain sufficient depth to the high water table and sustain agricultural productivity in these areas.

Much of the water collected in subsurface drain systems installed on the westside of California’s San Joaquin Valley (Valley) is high in dissolved solids and contains naturally occurring selenium, molybdenum, boron, and other elements. The 1983 discovery of toxic concentrations of selenium in waterfowl at Kesterson Reservoir, a holding pond for agricultural drainage located in the Valley, underscored the complex pathways through which water collected in drain systems can concentrate in ecosystems both near the source and far away.

As a result of events at Kesterson, the State of California has established a water quality standard for selenium in the San Joaquin River and is considering standards for other elements and salts (California, 1988). It has been estimated that the river quality standard could be met with approximately thirty percent decreases in drain water volumes discharged from a 94,000 acre drainage study area on the westside of the Valley, and that these decreases are feasible with water conservation through improved management of irrigation applications (California, 1987). The means by which growers might be encouraged to adopt the changes necessary to achieve the recommended drainage reductions have not yet been determined.

The implications of alternative policies for addressing the drainage problem are examined with an agricultural production model designed to simulate farmer decision making in the drainage problem area. The model describes economic, agronomic and hydrologic characteristics pertaining to the area and predicts changes in agricultural production decisions and drainage volumes in response to policy alternatives.

Water can be conserved from agriculture in three ways: (i) water applications can be reduced, allowing crop yields to decline as a result of water stress; (ii) irrigation application efficiency and uniformity can be increased as water applications are reduced to maintain crop yields; or (iii) cropping patterns can be changed to replace crops that have relatively high water requirements with those with lower water needs. All three possibilities are incorporated in the simulation model.

Siphon tube furrow irrigation systems with half mile runs are typically used to irrigate cotton, tomatoes, sugarbeets, and melons (cantaloupes) in the area, while wheat fields are generally irrigated with border check systems. These crops represent 80 to 90 percent of irrigated acreage in the study area.

Changes in irrigation practices can conserve water and may help to reduce drainage production, but will necessarily increase costs. To incorporate this aspect of the problem, crop specific irrigation technology cost-efficiency functions are estimated and included in the model. Irrigation efficiency is defined as the ratio of the depth of water beneficially used (plant needs plus minimum leaching fractions) to the average depth of water applied to a field.

Production of the principle crops is modeled with crop-water production functions. Water applications (x) are multiplied by irrigation efficiency (z) in the production functions so that yield is a function of effective applied water, i.e. the amount of water available for plant growth.

The objective in this problem is to chose cropping patterns and crop-specific water applications and irrigation efficiency levels to maximize net returns to land and management from crop production, subject to water allotments, land availability and constraints on acreage allocations for selected crops, and given the technological relationships specified for crop production and irrigation technology costs. Collected drain water volumes are predicted with a mass balance equation adapted from the Westside Agricultural Drainage Economics model (Hatchett, et al.). The model is specified as a non-linear programming problem and solved with an appropriate algorithm (see Weinberg for a more complete model description).

Prices charged to farmers in the area range from $0 to $36 per acre-foot and allocations from approximately 2.3 to more than 4 acre-feet per acre. Official estimates of the irrigation subsidies for the area range from $15 per acre-foot to nearly $50 per acre-foot (United States).

Three “farms” representing different water districts are selected in order to incorporate heterogeneous institutional parameters in the analysis. One farm (Farm 1) represents a district with a token charge of $1 per acre-foot of delivered water and an allotment of 4 acre-feet per acre. The other two farms face a price of $60 per acre-foot, reflecting “full-cost” water, with farms 2 and 3 receiving allocations of 3.7 and 3.3 acre-feet per acre, respectively.

Base case results are consistent with expectations. Water applications are higher and irrigation efficiencies are lower in Farm 1 than in Farm 2. Similarly, more water is applied and efficiencies are lower in Farm 2 than in

Farm 3. The higher efficiencies required in Farms 2 and 3 to obtain optimum crop yields result in increased irrigation technology and management costs. Higher irrigation costs result in net returns to land and management that are highest in Farm 1 and lowest in Farm 3, although results for Farms 2 and 3 are quite similar. Larger water applications are expected to generate larger volumes of drain water. The model predicts that Farm 1 will generate 1.25 acre-feet per acre of drain water, while Farm 2 will generate 1.02 acre-feet per acre and Farm 3 will generate .96 acre-feet per acre.

As noted above, it has been suggested that water quality objectives could be met with roughly 30 percent reductions in drain water collected in the drainage problem area. A number of alternatives exist for allocating the reduction objective among farms and water districts in the area. Two methods are considered here: an equilibrating scheme and a uniform one. The equilibrating scheme, so called because it equilibrates private marginal abatement costs among farms, specifies a 30 percent regional reduction objective and solves for the least cost means of achieving that objective given observed water prices. The uniform reduction scheme requires that each farm reduce drain water volumes by thirty percent from base levels.

Figure 4 illustrates the income and efficiency effects of alternative drainage allocation schemes. Comparison of average private returns under alternative allocation schemes provides an indication of the efficiency of each. However, in a second-best world this comparison must be made net of the price subsidy to determine true welfare costs of choosing between allocation schemes, these are denoted net returns in Figure 4. The water supply constraints are binding in the base results for all three farms, but are not binding in any case in which the thirty percent drainage reduction is met, This problem can thus be classified in regime 3. The input constraint is not policy relevant in this case, as illustrated in the previous section.

Results indicate that the equilibrating allocation scheme appears to be socially optimal when the price subsidy is not considered. Average private returns are $350.20 per acre under the equilibrating scheme and $349.70 per acre the uniform scheme when comparing Farms 1 and 2 (Figure 4a). The value of the $59 per acre-foot price subsidy is $218 per acre to Farm 1. Deducting this payment prior to comparison of the policy alternatives reveals that average returns are $10 per acre higher with the uniform allocation than with the equilibrating one.

The results demonstrate that application of a first-best policy prescription in a second-best environment can be welfare reducing. This result is not universal, however, and the advantage of the uniform allocation scheme is reduced when comparing farms with increasingly larger differences in initial water allotments. For example, a comparison of Farms 1 and 3 reveals that average private returns are essentially the same under the two drainage allocation schemes and that the equilibrating scheme results in net returns that are $.75 less than with the uniform scheme (Figure 4b).

The impact of the tighter water constraint for Farm 3 is to increase the private value marginal product for the input, in effect increasing marginal abatement costs relative to those for farms with higher water allotments. The equilibrating scheme gains, relative to the uniform one, by incorporating these factors in the final drainage reduction allocations. Nevertheless, the uniform allocation performs as well or better than the equilibrating one in both cases considered here. In addition, uniform reduction policies may require less information and involve lower implementation costs than policies that are generally considered to be efficient for achieving environmental objectives, particularly for cases involving non point-sources of emissions.

Summary and Conclusions

This paper addresses the problem of the second-best in non-point source control and describes conditions under which an effluent tax or set of input taxes will not assure a social objective at least cost. The second-best conditions examined include price distortions for a polluting input and input allocations that are institutionally determined. These conditions characterize Bureau of Reclamation irrigation water distribution policies.

Results indicate that even if an input constraint is binding at the private optimum, it is not policy relevant if it is not binding at the social optimum. The principle result that both effluent taxes and a set of input taxes define optimal policy choices remains valid in this case. However, if the constraint is binding at the social optimum (and is immutable) then: (i) a first-best solution is not attainable, and (ii) the input constraint defines the socially second-best optimal level of the input. No policy action with respect to the constrained input is required, though it may be necessary to introduce a policy tool to motivate optimal changes in the levels of other inputs.

Results indicate that the policy maker can ignore the input constraint if it is not binding at the social optimum but must correct for the input price distortion to achieve a social optimum. A set of input taxes is optimal if the input tax on the polluting input is a composite one including both the marginal social cost associated with input use and the price differential between the “true” and actual price. An effluent tax alone is not capable of assuring that the social optimum is realized.

The price distortion is not policy relevant in the short run when the constraint is binding at the social optimum, although the input price distortion acts as an income transfer with distributional consequences that the policy maker may want to address. The long run implications may be different than the results expressed here.

An agricultural drainage problem is examined to illustrate the implications of policy alternatives in a second-best setting. Equating marginal benefits among firms reduces welfare in this example. The pervasiveness of government intervention in agricultural input and output markets motivates incorporation of these results when designing policies to address the environmental problems associated with irrigated agriculture. Policy makers that do not examine the implications of second-best conditions before making policy recommendations may reduce social welfare in the process of addressing externalities associated with irrigated agriculture.

This paper has re-iterated the warning of Lipsey and Lancaster that society can be made worse off by the attempt to apply “first-best” policy rules in a second-best setting. The results suggest that alternative sources of market failure may have important implications for environmental policy makers. The optimality of policy instruments for externality control requires that these implications are considered.

References

Baumol, W. J. and W. E. Oates. *The Theory of Environmental Policy.* Cambridge: Cambridge University Press. 1988.

Buchanan, J. M. "External Diseconomies, Corrective Taxes, and Market Structure." *American Economic Review.* 59(1969):174-177.

California, Regional Water Quality Control Board, Central Valley Region. "Staff Report on the program of Implementation for the Control of Agricultural Subsurface Drainage Discharges in the San Joaquin Basin (5C)." Sacramento, California. August 1988.

California, State Water Resources Control Board. "Regulation of Agricultural Drainage to the San Joaquin River. Appendix G." Sacramento, California. August 1987.

Griffen, R. C. and D. W. Bromley. "Agricultural Runoff as a Non-point Externality: A Theoretical Development." *American Journal of Agricultural Economics.* 64(1982):547-52.

Hatchett, S. A., N. W. T. Quinn, G. L. Homer, and R. E. Hewitt, "A Drainage Economics Model to Evaluate Policy Options for Managing Selenium Contaminated Drainage." *In Toxic Substances in Agricultural Water Supply and Drainage-An International Perspective, proceedings of the Second Pan-American Regional Conference on Irrigation and Drainage.* USCID . June, 1989.

Lichtenberg, E. and D. Zilberman. "The Welfare Economics of Price Supports in U.S. Agriculture." *The American Economic Review.* 76(1986):1135-41.

Lipsey, R. G. and K. Lancaster. "The General Theory of Second Best." *Review of Economic Studies.* 24(1956):11-32.

Pfeiffer, G. H. and N. K. Whittlesey. "Controlling Nonpoint Externalities with Input Restrictions in an Irrigated River Basin." *Water Resources Bulletin.* 14(1978):1387-1403.

Neary, J.P. and K.U.S. Roberts. "The Theory of Household Behavior Under Rationing." *European Economic Review.* 13(1980):25-42.

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

Shortle, J. S. and J.W. Dunn. "The Relative Efficiency of Agricultural Source Water Pollution Control Policies." *American Journal of Agricultural Economics,* 68(1986):668-677.

Spulber, D. F. "Effluent Regulation and Long-Run Optimality." *Journal of Environmental Economics and Management.* 12(1985):103-116.

Stevens, B. K, "Fiscal Implications of Effluent Charges and Input Taxes." *Journal of Environmental Economics and Management.* 15(1988):285-96.

United States Department of Interior, Office of the Secretary. H.R. 1443, Irrigation Subsidy Legislation: Questions from the Subcommittee on Water and Power Resources of the Committee on Interior and Insular Affairs, memorandum. Washington, D.C. February 24, 1988.

Weinberg, M. "Economic Incentives for the Control of Agricultural Non-Point Source Water Pollution." Unpublished Ph.D. Dissertation, University of California, Davis. 1991.

Willey, Z. "Environmental Quality and Agricultural Irrigation--Reconciliation through Economic Incentives." *In Agricultural Salinity Assessment and Management,* K. Tanji, editor, American Society of Civil Engineers. New York. 1990.

Figure 1. Externality control in the case of input price and quantity distortions – Regime 1

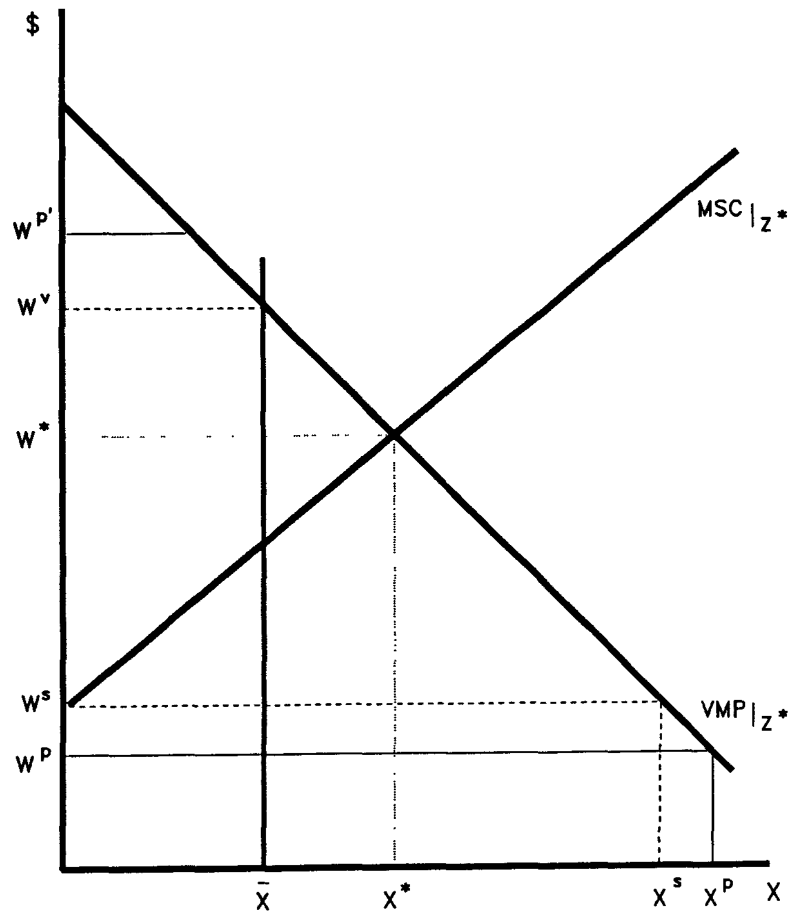


Figure 2. Externality control in the case of input price and quantity distortions – Regime 3

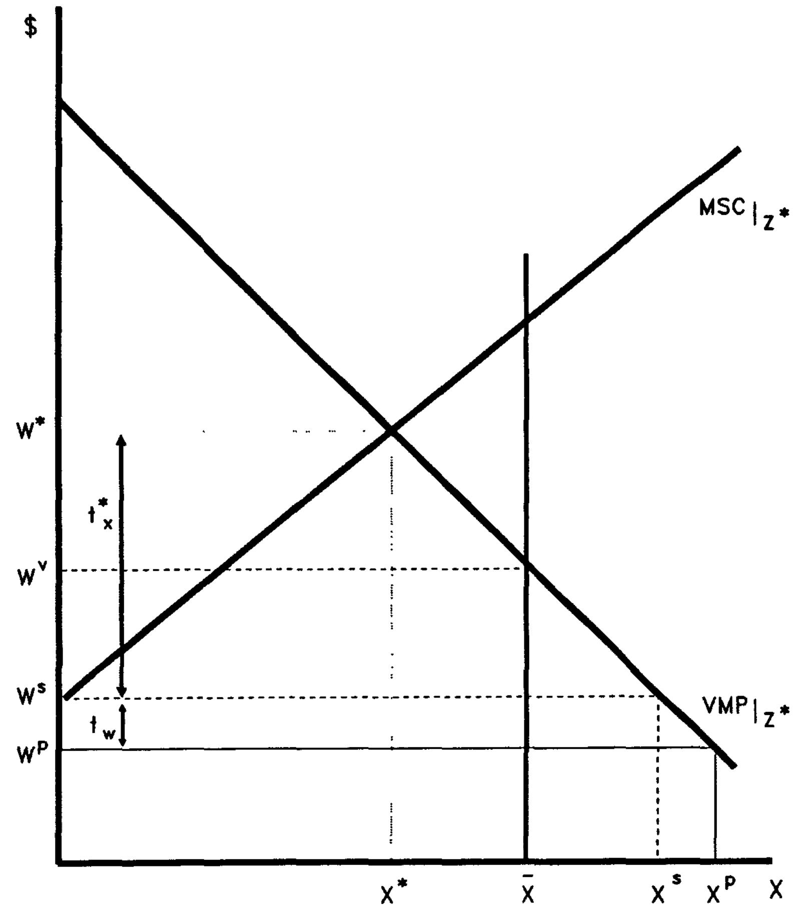


Figure 3. Welfare effects of policy alternatives under second best conditions

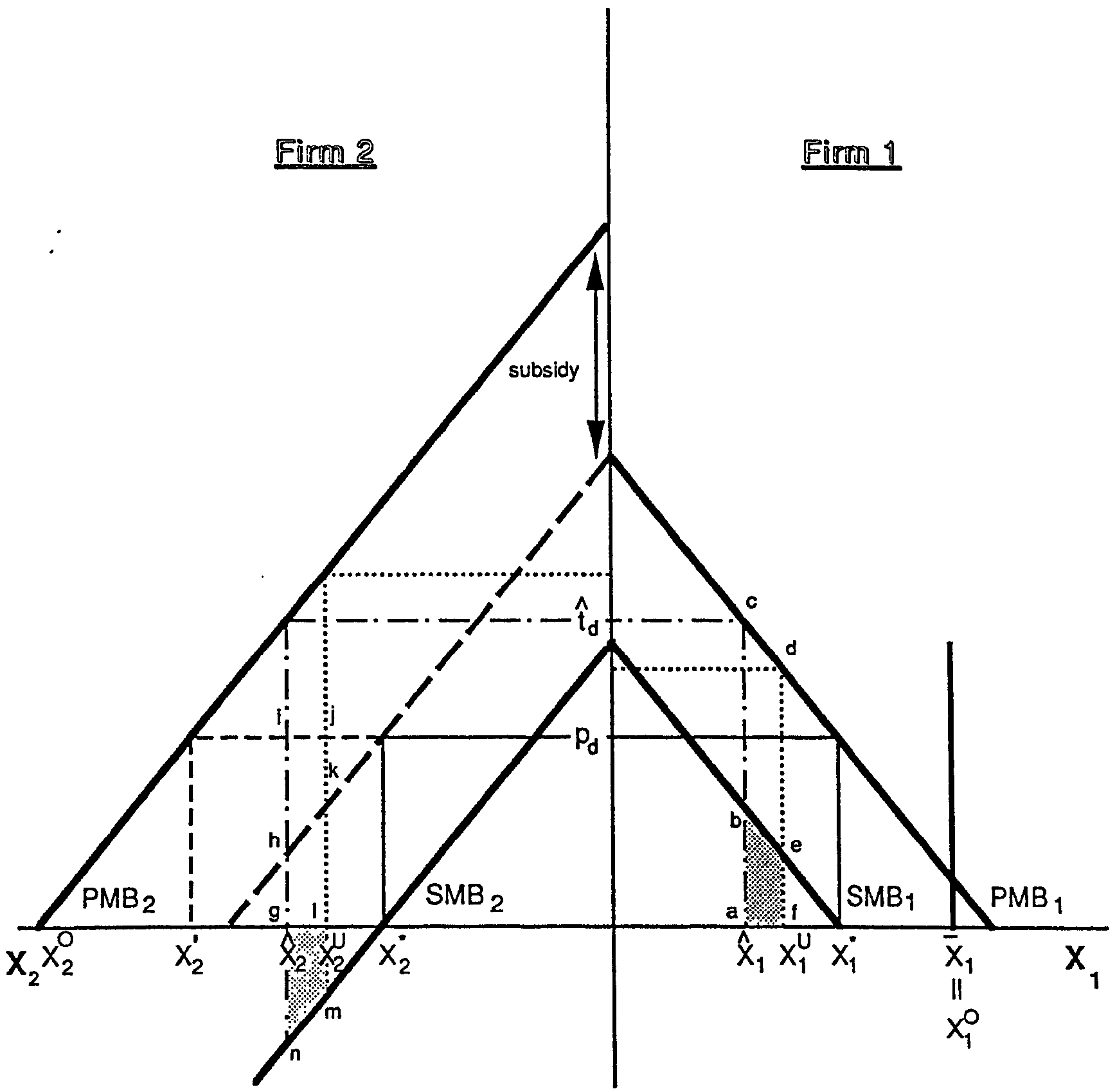
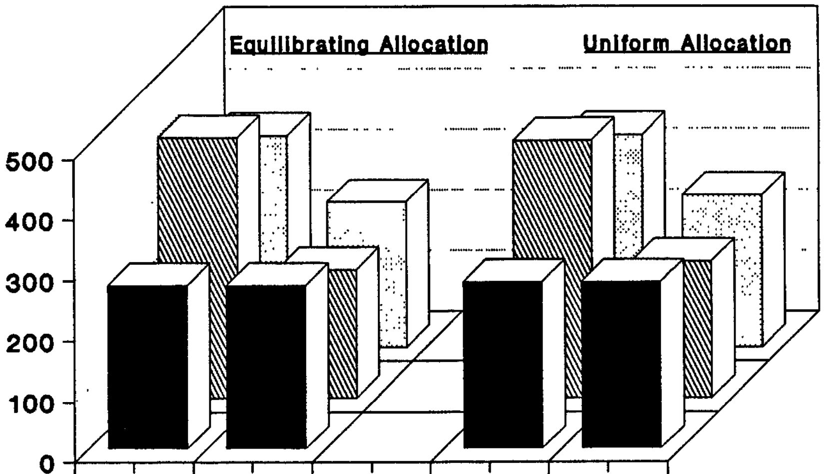
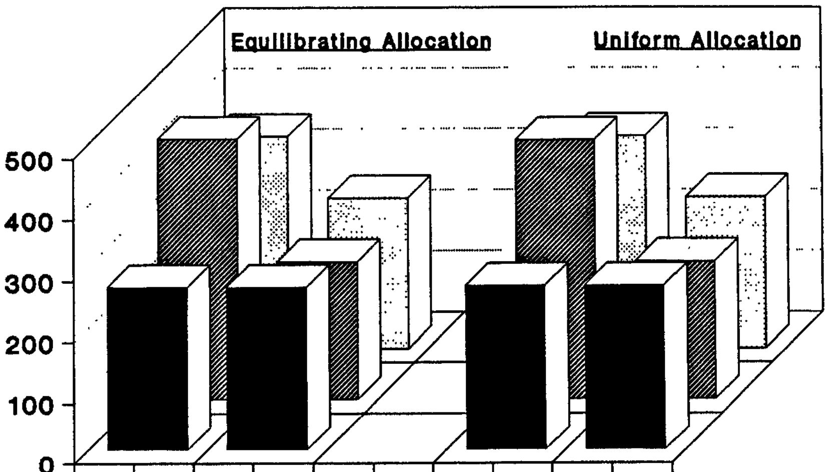


Figure **4. Income effects of policy alternatives**



	Equilibrating Allocation: Private Returns	Equilibrating Allocation: Net Returns		Uniform Allocation: Private Returns	Uniform Allocation: Net Returns
Average	350	241		350	251
Farm 2	430	212		424	226
Farm 1	270	270		275	275

a. Comparison of Farms 1 and 2



	Equilibrating Allocation: Private Returns	Equilibrating Allocation: Net Returns		Uniform Allocation: Private Returns	Uniform Allocation: Net Returns
Average	347	247		347	248
Farm 3	426	226		424	226
Farm 1	268	268		270	270

b. Comparisons of Farms 1 and 3

Point/Nonpoint Source Trading for Controlling Nutrient Loadings to Coastal Waters: A Feasibility Study

David Letson, Stephen Crutchfield, and Arun Malik
Resources and Technology Division, Economic Research Service, USDA

ABSTRACT: This paper attempts to make an initial assessment of the feasibility of point/nonpoint source trading in coastal watersheds. A theoretical model finds that the relatively greater uncertainty and monitoring costs associated with nonpoint source loadings make the setting of a trading ratio difficult a priori. A set of simple screening rules reveals that ten percent of coastal watersheds have significant contributions of loadings from both point and nonpoint sources. These results suggest that point/nonpoint source trading is more likely to work in a small number of coastal watersheds than as a means from bringing nonpoint sources in coastal watersheds under control nationally.

I. Introduction

While the reduction of point source (PS) discharges since 1972 has yielded some improvements in the nation's water quality (e.g. in lower bacterial contamination and higher dissolved oxygen levels), discharges from nonpoint sources (NPSs) remain and have increased as a share of the water quality problem. Impairments from sedimentation, nutrient enrichment, runoff from farmlands, and toxic contamination of fish tissue and sediments have become more evident (USEPA 1990). Extending regulatory controls to include NPSs of water pollution may be necessary for the objectives of Federal water pollution control legislation to be met.

Federal authorities have stepped up efforts to control NPS pollution in recent years. The 1987 Water Quality Act (WQA) authorizes the expenditure of up to $400 million by the EPA to help control water pollution from NPSs States are also required to file management plans under Section 319 of the Act, identifying steps for reducing loadings from NPSs. In addition, the 1990 Amendments to the Coastal Zone Management Act (CZMA) empower EPA and NOAA to manage land use in coastal areas. The President’s Water Quality Initiative, while aimed primarily at ground water protection, is also intended to support education and technical assistance efforts to promote voluntary adoption of farm management practices, which reduce agriculture-related impairments of surface water quality, in watersheds identified by 319 reports as having NPS pollution problems.

The principle behind these efforts, and the upcoming reauthorization of the Clean Water Act in 1992, is to treat ecosystems as a whole, rather than focusing on single sources of contamination or single pollutants. The recently enacted amendments to the CZMA serve as an example. While they stress the importance of managing land use in coastal areas to protect surface water quality, the amendments are coordinated with existing water quality management efforts. The most important provision, Section 6217 (“Protecting Coastal Waters”), requires each State to develop a new Coastal Nonpoint Pollution Control Program. The programs will seine as amendments to the State’s existing Coastal Zone Management (CZMA) and Nonpoint Source Management (Section 319 of the WQA) Programs. The central purpose is to strengthen links between coastal zone management, water quality programs, and land use restrictions.

A holistic approach may be especially appropriate for bringing NPSs under control. NPSs have grown as a share of the water quality problem because they are harder to identify and control than PSs. NPS contributions depend upon localized factors such as land use, climate, and a host of geomorphological characteristics. Therefore, control of NPSs is more likely to be cost effective if approached in a way that allows for their site-specific nature.

One approach for dealing with these problems would be to allow PSs to sponsor implementation of NPS controls rather than install controls of their own. PS operators and local environmental officials may be better situated to identify and manage localized water quality problems than regional and national regulators, and a PS/NPS trading program may give them both the means and the motivation to do so. Moreover, if the NPS component of the overall water quality problem is significant, and the relative costs of NPS reductions are lower than the costs of additional PS controls, then water quality goals could be met at a lower cost by substituting NPS reductions for PS ones. Two PS/NPS trading programs presently exist in Colorado, and a third has recently been approved for the Tar-Pamlico Basin of North Carolina.

The purpose of this paper is to offer an initial assessment of the feasibility of the PS/NPS trading option for coastal water quality management. Our approach has two parts. First, we explore some of the conceptual and practical issues involved in designing and implementing PS/NPS trading programs. Second, we consider the number of coastal watersheds that satisfy simple but necessary conditions for implementation of PS/NPS trading. Our focus on coastal waters is motivated by the recent reauthorization of the CZMA, which calls for increased efforts by States to improve the condition of coastal waters. We emphasize agricultural nonpoint sources because they are recognized as the single largest (USEPA 1990), the means for their control are less capital intensive than urban runoff controls, and they are more controllable than runoff from forestland or barren lands.

Given this objective, we develop and analyze a simple model of PS/NPS trading in Section 11. This model shows how the random loadings and higher monitoring costs generally associated with NPSs pose difficulties for those who would attempt to set up a trading program. In Section III we discuss some the real world complexities omitted from the conceptual model and consider their possible effects on implementing trading programs. We find no shortage of potential obstacles, but many of the problems also would apply to any attempt to control NPS pollution. In Section IV we conduct a screening study as an initial assessment of the feasibility of trading Programs for managing coastal water quality. Section V offers our conclusion that PS/NPS trading does not have broad applicability for coastal water quality management but might work in a few locations.

II. The Economics of Point/Nonpoint Source Trading

1. Limitations of the Standard Model of Trading

The standard economic argument in favor of PS/NPS trading is a simple one: by allowing point sources with high abatement costs to trade pollution reductions with nonpoint sources that have lower abatement costs, the total costs of achieving a given level of water quality can be reduced. This argument has long been used to support proposals for establishing trading programs among point sources of pollution that have different abatement costs. However, the economics of trading are not as simple when nonpoint sources are involved.

To begin, the standard argument assumes that pollutant loadings are deterministic. This may be a reasonable assumption for point sources, but it certainly is not for nonpoint sources. A characteristic feature of nonpoint loadings is that they are stochastic: the loadings are influenced by a variety of climatic and gemorphological factors. As such, they are more difficult to control than point source loadings, and the effectiveness of nonpoint controls is more difficult to predict.

Nonpoint loadings are also more difficult to measure (Harrington *et al.*, 1985 and Segerson, 1988). This raises questions about how trades involving nonpoint sources would be monitored and enforced. Even if we assume that estimates of, say, average loadings are obtainable, the estimates are likely to be costly. As a result, enforcement costs for nonpoint sources are likely to be high.

Although researchers examining PS/NPS trading have recognized these problems (e.g. see Kashmanian *et al.*, 1986), little attention has been given to what they imply for the economics of trading.[1] Discussions of the problems typically conclude by recommending that uncertainty about the magnitude and effect of nonpoint loadings and the difficulty of measuring them should be accommodated by making the terms of trade less favorable for nonpoint sources (Kashmanian, 1986 and EPA, 1991). The usual recommendation is that the "trading ratio" should be set above one: a unit reduction in loadings by a nonpoint source should count for less than a unit reduction by a point source. Although this recommendation may be appropriate, by itself, it says little about how the trading ratio should actually be set. In this section we develop a simple model that allows us to examine the economics of PS/NPS trading more formally and to identify the factors that determine the magnitude of the "correct" trading ratio.

2. An Alternative Model

For simplicity, let us suppose that our hypothetical watershed contains just one point source and one non-point source (the model can be generalized to multiple sources quite easily). The differences between the two sources are: (i) the pollutant loadings from the point source are deterministic, while the loadings from the nonpoint source are stochastic; and (ii) it is cheaper to measure (or monitor) average loadings from the point source than

[1] A notable exception is the work by Milon (1987), which explicitly takes into account the stochasticity of loadings and their effects on water quality.

the nonpoint source. These differences are obviously stylized. Point source loadings are likely to also be stochastic, albeit with less variability than nonpoint loadings. Furthermore, there are likely to be differences in the accuracy with which point and nonpoint loadings can be estimated, and not just in the costs of estimating them. (We assume that average loadings from the two sources are measured without error,)

Production, Technology, and Pollutant Loadings

For concreteness, suppose that the nonpoint source is a farm and the point source is an industrial plant and that both are risk-neutral profit-maximizers. The variable and capital inputs used by each are represented by the vectors $\mathbf{x}_i$ and $\mathbf{k}_i$; the vectors include both production inputs and abatement inputs. The capital input vector $\mathbf{k}_i$ characterizes each source's technology.

The inputs used by the point source (i = 1) determine its product output $q_1 = f_1(x_1,k_1)$, and its pollutant loadings $e_1 = g_1(x_1,k_1)$. The same is true for the nonpoint source (i = 2), except that its pollutant loadings also depend on a random variable ω: $e_2 = \tilde{g}_2(x_2,k_2;\omega)$, ω is intended to capture both the inherent stochasticity of nonpoint loadings (due to climatic and geomorphological factors), as well as uncertainty about the relationship between input use and the magnitude of loadings. We shall assume that larger values of ω imply higher loadings.

For our purposes, it is convenient to write nonpoint loadings as the sum of the average loading and a stochastic deviation term with mean zero:

(1) $$e_2 = g_2(x_2,k_2) + \epsilon_2(x_2,k_2;\omega),$$

where $g_2 = E[\tilde{g}_2]$ and $\epsilon_2 = \tilde{g}_2 - g_2$.[2] Note that the above specification does not, per se, impose any restrictions on the source's ability to control the distribution of its loadings. For instance, by choosing its inputs appropr-iately, the source could independently vary the mean and variance of its loadings.[3]

Environmental Damages

The pollutant loadings result in damages to the waterbody. The monetary value of these damages is given by $D(\theta_1 e_1 + \theta_2 e_2)$. We assume the damage function is "smooth" and that marginal damages are positive and non-decreasing: $D' > 0$ and $D'' \geq 0$. The constants θ_i allow for the possibility that loadings from the two sources have different effects on the waterbody (perhaps due to location or to the chemical composition of the loadings).[4]

Abatement Costs

To define abatement costs, we need to specify the loading parameter that sources would trade. An obvious choice is the average (or mean) loading over some period of time. We shall assume this is the parameter traded, not only because it is convenient analytically, but also because it is the quantity

---

[2] Shortle (1990) uses a similar approach.

[3] Whether or not it could actually do this would depend on the form of the loading function $\tilde{g}(\cdot)$ If $\tilde{g}(\cdot) \equiv h(x,k) + \omega$, the source would only be able to influence mean loadings. However, if $\tilde{g}(\cdot)$ has the more general form $\tilde{g}(\cdot) \equiv h(x,k) + t(x,k;\omega)$, it would be able to influence both the mean and variance of loadings.

[4] Location would be relevant for what Tietenberg (1985, p. 22) terms nonuniformly mixed assimilative pollutants. Chemical composition would be important for pollutants such as phosphorus, whose effect on the environment, depends on its exact form, which can vary across point and nonpoint sources (see Krupnick, 1989, p. 16).

traded in existing programs. It should be noted, though, that depending on the pollutant and the characteristics of the watershed, it may be preferable to consider some other parameter, such as maximum loadings.[5]

Each source's abatement costs can now be defined in terms of its average loadings. Two types of abatement cost functions are defined: a restricted cost function $C_i(\bar{e}_i,k_i)$ and an unrestricted cost function $C_i(\bar{e}_i)$. The restricted cost function gives the source's costs, in terms of foregone profits, of achieving an average loading of $\bar{e}_i$ using a prescribed technology k_i. $C_i(\bar{e}_i,k_i)$ can be derived from the source's profit-maximization problem.[6]

The unrestricted cost function, $C_i(\bar{e}_i)$, gives a source's abatement costs when it is free to choose the technology it uses. We can define $C_i(\bar{e}_i)$ in terms of the restricted cost function:

$$(3) \qquad C_i(\bar{e}_i) \equiv \min_{k_i} \; C_i(\bar{e}_i,k_i).$$

For both the restricted and unrestricted cost functions, marginal abatement costs are positive over the relevant range: $-\partial C_i/\partial \bar{e}_i > 0$ and $dC_i/d\bar{e}_i > 0$. Furthermore, assuming the production and loadings functions have

---

[5] It may even be desirable to trade more than one parameter, for instance, average monthly loadings and maximum daily loadings. Milon (1987) discusses the shortcomings of using average loadings alone.

[6] Letting $f_i(x_i,k_i)$ represent a source's production function,

$$C_i(\bar{e}_i,k_i) \equiv \pi_i^* - \left[\max_{x_i} \{p_i f_i(x_i,k_i) - w_i'x_i - r_i'k_i\} \quad \text{s.t. } g_i(x_i,k_i) = \bar{e}_i\right].$$

The first term above (π_i^*) represents the source's profits when its actions are unrestricted; the second term represents its profits when it must restrict average loadings to $\bar{e}_i$ using the prescribed technology k_i. w_i and r_i are simply the prices of the variable and capital inputs; and p_i is the price source receives for its output.

the appropriate properties, the marginal abatement cost schedules will have the usual downward sloping shape: $-\partial^2 C_i/\partial \bar{e}_i^2 < 0$ and $-d^2 C_i/d\bar{e}_i^2 < 0$.

Monitoring and Enforcement Costs

Given the importance of monitoring and enforcement costs in PS/NPS trading programs, we incorporate them in our model. However, to keep the model tractable, we restrict attention to enforcement policies that achieve full compliance, i.e., they ensure that neither source exceeds its allowed average loadings. The regulator accomplishes this by auditing the sources at random and fining them if they are found exceeding their allowed average loadings.

Let α_i denote the probability a source is audited, and $F(\bar{e}_i - \mu_i)$ the fine the source faces for exceeding its allowed average loading μ_i. For the source to be compliant, the marginal expected fine it faces when $\bar{e}_i = \mu_i$ must be no smaller than its marginal cost of abatement:

$$\alpha_i F'(0) \geq -\partial C_i(\mu_i, k_i)/\partial \bar{e}_i, \tag{4}$$

Note that this condition is appropriate even when the source chooses its technology, because technology is presumably fixed when the source makes its day-to-day compliance decisions; these decisions would be based only on the source's variable costs.[7]

Taking the fine schedule as exogenous, condition (4) implies that by setting an audit probability of

$$\alpha_i^c = \phi[-\partial C_i(\mu_i, k_i)/\partial \bar{e}_i], \tag{5}$$

---

[7] Thus we are modeling continuing compliance rather than initial compliance (Russell, *et al.,* 1986, p. 8).

where $\phi = [F'(0)]^{-1}$, the regulator could ensure a source's compliance, Enforcement costs in this situation would consist solely of audit costs.[8] Total expected enforcement costs would be $\alpha_1^c A_1 + \alpha_2^c A_2$, where A_i denotes the cost of an audit. Given our premise that it is more costly to monitor the loadings of nonpoint source than a point source, A_2 would be greater than A_1.

3. The Benchmark Optimum

We are now ready to specify a benchmark social cost minimization problem. The solution to this problem is intended to provide a realistic reference against which to compare trading programs; it does not represent the first-best solution. The benchmark problem is one where the regulator can dictate a source's average loadings (μ_i) and the technology it uses $(\mathbf{k}_i)$, but not its variable input use $(\mathbf{x}_i)$. (Enforcing variable input use is assumed to be prohibitively costly.)

The regulator ensures the source's compliance with μ_i, by auditing it with probability α_i^c. Ensuring the source adopts the prescribed technology is assumed to be costless at the margin -- it simply requires a one time check of the technology the source is using.

Formally, the benchmark problem is to find the technologies and average loadings that minimize the sum of abatement and enforcement costs, plus the expected damages from pollutant loadings:

$$(6) \quad \min_{\mu,\mathbf{k}} \sum_{i=1}^{2} \left[C_i(\mu_i, \mathbf{k}_i) + \alpha_i^c A_i \right] + E\left[D(\theta_1\mu_1 + \theta_2(\mu_2 + \epsilon_2^*(\mu_2, \mathbf{k}_2; \omega))) \right],$$

[8] There would be no fine-related costs, since fines would never be levied.

where $\epsilon_2^*(\mu_2, k_2; \omega) \equiv \epsilon_2(x_2^*(\mu_2, k_2), k_2; \omega)$ is the indirect loading deviation function. g In writing the expression for damages we have made use of the equalities $e_1 = \mu_1$ and $e_2 = \mu_2 + \epsilon_2^*$.[10]

Substituting for α_i^c from (5), and denoting the covariance operator by COV, the first-order conditions for the benchmark problem can be written as:11

$$(7) \qquad \frac{-\partial C_1}{\partial \mu_1} + A_1 \frac{\partial^2 C_1}{\partial \mu_1^2} \phi = \theta_1 E[D'(\cdot)],$$

$$(8) \qquad \frac{-\partial C_2}{\partial \mu_2} + A_2 \frac{\partial^2 C_2}{\partial \mu_2^2} \phi = \theta_2 E[D'(\cdot)] + \theta_2 \mathrm{COV}\left[D'(\cdot), \frac{\partial \epsilon_2^*}{\partial \mu_2}\right],$$

$$(9) \qquad \frac{\partial C_1}{\partial k_1} = A_1 \frac{\partial^2 C_1}{\partial k_1 \partial \mu_i} \phi, \qquad \forall\ k_1 \in \mathbf{k}_1,$$

$$(10) \qquad \frac{\partial C_2}{\partial k_2} = A_2 \frac{\partial^2 C_i}{\partial k_2 \partial \mu_2} \phi - \theta_2 \mathrm{COV}\left[D'(\cdot), \frac{\partial \epsilon_2^*}{\partial k_2}\right], \qquad \forall\ k_2 \in \mathbf{k}_2.$$

Optimal Allocation of Average Loadings

The first condition (7) calls for μ_1 to be set so that the sum of the marginal cost of abatement and the marginal cost of enforcement for the point

---

9 The choice function $x_i^*(\mu_i, k_i)$ is the solution to the profit maximization problem in footnote 6.

10 The above specification assumes that the regulator is risk-neutral. Allowing for risk aversion yields qualitatively similar results to those obtained below for the case of a strictly convex damage function ($D'' > 0$).

12 We have made use of the relationship $E(ab) = E(a)E(b) + COV(a,b)$. Here, $a = D'(\cdot)$ and $b = \partial\epsilon_2^*/\partial h$, where $h = \mu_2$ or k_2. Since ϵ_2^* has mean zero, $E[\partial\epsilon_2^*/\partial h] = \partial E[\epsilon_2^*]/\partial h = 0$.

source is equal to the expected marginal damage it causes. This is a natural extension of the usual rule that marginal costs should be equated to marginal damages.

The condition for the average loading from the nonpoint source (8) is somewhat more complicated: it contains an extra covariance term that reflects the uncertainty about nonpoint loadings. The sign and magnitude of the term depend, roughly, on the curvature of the damage function and on the relationship between the mean and variance of the source's loadings.

The term vanishes if damages are linear (since D'(.) is then a constant), o if the mean and variance of loadings are unrelated $(\partial\epsilon_2^*/\partial\mu_2 = 0)$. But if the damage function is convex (D" > 0) and larger average loadings imply a larger variability in loadings $(\partial\epsilon_2^*/\partial\mu_2\partial\omega > 0)$, the term is positive. In this case, the covariance term can be be thought of as representing the damage premium associated with the uncertainty about loadings.

Optimal Choice of Technology

Turning to the conditions for the capital inputs, (9) and (10), we can establish that the regulator's choice of technology is not the same as a source's. From (3), we can verify that a source would choose its technology so that $\partial C_i/\partial k_i = 0,\ \forall k_i$. Conditions (9) and (10) are more complicated than this: the regulator takes into account the effect of technology on enforcement costs and, in the case of the nonpoint source, on the damage premium.[12] The sources ignore these costs when choosing k_i, since they do not bear them,

[12] That technology may influence enforcement costs and/or the variability of loadings should not be surprising. Suppose, for example, that the farm can lower its nitrogen loadings by either building a retention pond or by reducing the amount of fertilizer it applies. Both enforcement costs and the variability of loadings are likely to differ for these two technologies.

Once again, the damage premium in (10) is positive only if damages are non-linear. If damages are linear, the variability of loadings is irrelevant, and the regulator will only worry about average loadings. However, if the damage function is convex, the regulator will exploit opportunities to reduce the variability of loadings by prescribing the appropriate technology.

The above results suggest the following two broad conclusions: (i) uncertainty about nonpoint source loadings is of concern only if the damage function is nonlinear; and (ii) allowing the regulator to prescribe the technology sources should adopt can reduce social costs, to the extent that the choice of technology influences the magnitude of enforcement costs and the damage premium.

4. Implications for the Design of a Trading Program

Let us examine the implications of the above analysis for the design of a PS/NPS trading program. We begin by considering the issue of the appropriate trading ratio. We shall assume, for the moment, that the regulator can dictate the technology a source adopts, and can thereby ensure that conditions (9) and (10) are satisfied. The question, then, is whether the regulator can set the trading ratio so that conditions (7) and (8) hold.

Rearranging (7) and (8), and dividing one by the other, we find

$$\text{(11)} \qquad \frac{\partial C_2/\partial\mu_2}{\partial C_1/\partial\mu_1} = \frac{-A_2\phi\cdot\partial^2 C_2/\partial\mu_2^2 + \theta_2 E[D'(\cdot)] + \theta_2 COV[D', \partial\epsilon_2^*/\partial\mu_2]}{-A_1\phi\cdot\partial^2 C_1/\partial\mu_1^2 + \theta_1 E[D'(\cdot)]} .$$

The LHS of this equation represents the ratio of marginal abatement costs. Therefore, the RHS is the trading ratio required for the benchmark conditions in (7) and (8) to hold.

In the simple setting where loadings are deterministic and enforcement is costless, the RHS of (11) reduces to θ_2/θ_1; this is the familiar result that the optimal trading ratio should equal the relative environmental impacts of the loadings from the two sources.

The optimal trading ratio is considerably more complicated when loadings are uncertain and enforcement is costly. The ratio now depends, in part, on the magnitude of the damage premium. The argument that the trading ratio should be increased to compensate for the uncertainty about nonpoint loadings is corroborated by (11), provided damages are convex (or the regulator is risk averse) and higher average loadings imply greater variability. If either of these conditions does not hold, uncertainty about nonpoint loadings is irrelevant in determining the proper trading ratio. Thus, in recommending that trading ratios be set above one, it is being implicitly assumed that both these conditions do hold.

The relative magnitude of marginal enforcement costs also influences the optimal trading ratio. Enforcement costs are a function of the audit costs A_i and the slopes of the marginal abatement cost curves $\partial^2 C_i/\partial \mu_i^2$. If the cost curves have similar slopes, the relative magnitude of marginal enforcement costs is just a function of A_2 and A_1. In this case, the higher audit cost for the nonpoint source ($A_2 > A_1$) has the effect of lowering the optimal trading ratio. This is not surprising: if it is more expensive to ensure the nonpoint source's compliance, abatement burden should be shifted toward the point source, which requires lowering the trading ratio. This effect is reduced to the extent that the the marginal abatement cost curve for the point source is likely to be more steeply sloped than the curve for the nonpoint source. A priori, it is difficult to specify whether this difference would more than compensate for the difference in audit costs.

The above discussion reveals that calculating the appropriate trading ratio requires a substantial amount of information. Equation (11) shows us that the trading ratio must reflect all relevant social costs other than the direct costs of abatement. Since sources just bear the latter costs, the only means of forcing them to take into account the other social costs incurred is by adjusting the trading ratio appropriately.

We have assumed thus far that the regulator can dictate the technology that sources use. The form of equation (11) does not change drastically when sources are free to choose their technologies. Analytically, the primary difference is that the $\mathbf{k}_i$ must be replaced by each source's choice functions $\mathbf{k}_i^*(\mu_i)$ (see eq. (3)). This change does, however, have important implications for the regulator's ability to attain the benchmark optimum. The regulator now has only one policy instrument for each source, namely the allowed average loading, μ_i. Although the regulator will take into account the effect of the allowed loading on a source's choice of technology, it will not be able to costlessly influence the source's technology choice. As a result, the benchmark will no longer be attainable, and social costs will be higher.

Regardless of whether the regulator can prescribe technologies, the above analysis makes clear that setting the trading ratio is no simple matter. In particular, it shows that the two distinguishing features of nonpoint loadings -- their uncertainty and the higher costs of monitoring them -- may have opposing influences on the optimal ratio. Therefore it is questionable whether one can recommend *a priori* that trading ratios should favor point sources and thus be set above one. The analysis shows that one has to to consider the nature of the damage function and the relative costs of enforcement.

III. Technical and Practical Aspects

1. Real World Complexities

Implementing a successful PS/NPS trading program is of course much more difficult than the above analysis suggests. Many simplifying assumptions of the model depart from reality in important ways. These departures represent technical and practical complications in implementing trading. Significant among them are the model's represention of regulators' objectives, their monitoring and enforcement capabilities, their knowledge of the costs and effectiveness of NPS control methods, and the fate and transport of target pollutants. Many of these problems would encumber any attempt to bring NPSs under control, but some are unique to the trading option. If NPS control in some form is deemed necessary for achievement of legislated water quality goals, then the latter set of problems is more germane to our discussion. To avoid the all too common mistake of assigning all of these problems to trading programs alone, we shall discuss each of them in turn for the purpose of classification.

2. Problems Unique to the PS/NPS Trading Option

The trading option differs from other approaches to NPS control because it relies on a market to coordinate the actions of relevant economic agents. A key aspect in successfully coordinating them is having them in the appropriate number. Failure can come from having too many or too few participants in a pollution rights market.

The first side of the coordination problem pertains to measurement: we cannot measure pollutant loadings from individual NPSs therefore we do not know the previous loadings of individual NPSs. Without historical records for NPS loadings, we cannot calculate the pollutant reductions to be traded. PSs are unlikely to enter into what would essentially be a purchase of an ambiguous property right. We are able to estimate gross NPS loading and even to classify them by origin: urban, forestland, and agricultural runoff, etc. Because of our ability to measure gross but not individual loadings it might be necessary to involve all or most of a watershed's farmers in a given trade.

The coordination problem also has its abstract side. While often described conceptually as a perfectly competitive market, PS/NPS trading more closely resembles a private subsidy scheme: PSs avoid costly abatement by paying farmers to alter their practices. Unfortunately lower marginal abatement costs for NPSs alone may not make trading a reality. Coasian transactions costs are likely to eliminate some trades that would lower total control costs. A "stick" provision might be necessary to encourage farmers to participate in what might otherwise seem to others a profitable trade but to them is costly to arrange.

These two problems could exacerbate one another. To get an accurate estimate of the potential pollutant reduction, a large number of farmers might need to participate. Bargaining costs, though, might prevent transactions with numerous participants. In any case, a sufficient number of NPSs will have to exist to create a loading reduction the PS can use. The fact that failure can come from having too many or too few participants for the pollution rights market is part of what makes this coordination problem difficult. Other problems affecting NPS control in general also encumber the PS/NPS trading coordinator.

3. Four Problems with NPS Controls in General

Modeling Regional Objectives

The conceptual model has a single objective while the region's water quality managers generally will have more than one. Cooperation among participating regulators is necessary for any type of NPS control to achieve its cost and environmental goals, so this problem would likely confront any NPS control plan. The single coastal water quality regulator of the model in reality is probably several cooperating watershed authorities. Within any portion of the watershed, trading is likely to be unworkable since the entire area of influence of trading must be included if its water quality is to be protected. The entire system is likely to fall into several political jurisdictions, and a single authority would have to be empowered to run a trading program:

The absence of such an authority would mean further institutional change is necessary and make implementation of trading more difficult. Regardless, the smaller jurisdictions will not be quick to relinquish their powers to a regional authority. Two of the existing examples of PS/NPS trading are suggestive here. For the PS/NPS program for Dillon Reservoir, a threatened growth moratorium provided the motivation behind the formation of the Northwest Colorado Council of Governments. On the other hand, the Tar-Pamlico program in North Carolina faces no such crisis and its Basin Association is having more difficulty setting rules for trading (Anderson 1991). In sum, wedding the objectives of individual jurisdictions with regional cost minimization can prove difficult.

Monitoring and Enforcement

Any approach to NPS control would encounter monitoring and enforcement problems to some degree, but a trading program may be more susceptible to them. If a market for pollution reductions is to be established, the regional authority must be able to enforce trades and detect violations. Two problems may exist. First, many states simply do not have standards for nutrients. These states could establish them or link nutrient discharges to dissolved oxygen standards, but either approach would require the use of water quality models and is not costless. For a number of reasons such models are far from simple. (a) Estuarine models must include the effects of tidal incursions, normal surface flows, groundwater inflows of nitrates, and benthic sediments. (b) Because of the presence of both saline and fresh water, more than one pollutant can be limiting. (c) Phosphorus and nitrogen must be in dissolved and in inorganic forms to be available to phytoplankton for growth. Chemical and biological activity can convert other P and N forms to these forms, and vice versa. Second, trades may be difficult to enforce and violations difficult to detect with the present monitoring capacity and more difficult still to attribute to individual sources (Segerson 1988). PSs do not present so much of a problem here, but NPS controls themselves would probably have to be monitored rather than the resulting loadings.

Uncertain Performance of NPS Controls

NPS controls, as mandatory measures undertaken for water quality improvements, are virtually untried compared to our history with PS controls. This uncertainty would affect any attempt at NPS control and not just trading. The cost and effectiveness of NPS control methods are not known to regulators or dischargers with certainty. The level and number of acceptable violations under PS/NPS trading in reality is a stochastic decision problem.[16] Pollutant loadings and the physio-chemical reactions to them are uncertain. The added risk could make PS/NPS trading difficult to defend, both politically and to potential participants. Explaining the parameters of a risk management problem might prove difficult since the public is used to the relative certainty of PS controls. Reduction of the likelihood of a violation may be politically desirable, but too high a required likelihood might have costs exceeding the possible benefits (Milon 1987). Also, to encourage program participation, farmers uncertain of the efficacy of NPS controls might need a “stick” provision to go along with the “carrot” (i.e. the subsidy from the PS) that trading would provide.

Fate and Transport of Pollutants

We do not know enough about the fate and transport of target pollutants. NPS controls that reduce pollutant loadings to surface water (e.g. grassed waterways and animal waste treatment lagoons for the Tar-Pamlico Basin

[16] Point source control is also a stochastic decision problem, although, to a lesser extent.

program) may actually increase loadings of these or other pollutants to groundwater. While predictive models are available, large areas of uncertainty remain.[17] Obviously, the regulatory authority should consider the consequences of these problems for trading as well and would have to for any other type of NPS control.

4. The Appropriate Question

PS/NPS trading may pose higher administrative costs and greater environmental risks than technology standards for PSs. If some type of NPS control is soon in coming, however, the appropriate question relates to the additional problems the trading option creates relative to other approaches to NPS control. Arguably, our ignorance and uncertainty related to monitoring and enforcement, untried NPS controls, and the fate and transport of pollutants extends beyond the trading option to any attempt to bring NPSs under control. The additional problems unique to trading pertain to our lack of historical information that would enable calculation of actual loading reductions and to transactions costs. These latter problems appear considerable but are distinct from the more general difficulties with NPS control. Below, one of these issues (size of contribution of PSs and NPSs to loadings) is the basis for a screening study that serves as a conservative means for assessing the number of coastal areas nationally for which a PS/NPS trading program might be feasible.

---

[17] USEPA (1989) surveys the types of models available for rivers and lakes.

IV. Feasibility of Trading for Managing Coastal Water Quality

1. Background

In this section, we focus on the question of whether PS/NPS trading of pollution reductions is feasible in America's coastal water systems. We examine coastal water systems for several reasons. First, it limits the scope of our analysis to a manageable level: instead of analyzing all pollutant sources and water quality conditions nationwide we can look at a subset of watersheds in coastal states. Second, coastal water quality issues are highly policy relevant, given the recent Coastal Zone Management Act amendments and the renewed interest in protecting coastal water quality. Finally, as we discuss below, there exist several detailed data sources on sources and types of pollutant flows into coastal waters that facilitate a screening analysis.

Below, we develop simple screening rules to identify water systems which may be potential candidates for PS/NPS trading. We apply these rules using the data on coastal water systems to get an initial assessment of how many water systems could potentially be managed with PS/NPS trading.

2. Data Sources

Data used in this study come from three basic sources: the National Coastal Pollutant Discharge Inventory, the National Resources Inventory, and the EPA's AGTRAK database.

- The National Coastal Pollutant Discharge Inventory (NCPDI) has been developed by the National Oceanic and Atmospheric Administration (NOAA). The NCPDI contains pollutant loading estimates for all major types of pollutant sources located within coastal counties in the continental US (excluding the Great Lakes). Data are calculated on a base line of 1985 conditions. The pollution estimates are drawn from a variety of sources and based on many different methodologies (See Basta et. al. for details).

  Data were obtained from the NCPDI on pollutant loadings for four types of pollutants: Nitrogen, phosphorus, suspended sediments, and 5-day biochemical oxygen demand (BOD5). Pollutant sources were broken down into eight categories: Wastewater treatment plants, powerplants, industrial sources, urban runoff, cropland runoff, pastureland runoff, runoff from barren land, and upstream sources (pollutant loadings from inland regions).

- The National Resources Inventory (NRI) is conducted every five years by the USDA's Soil Conservation Service. The NRI is designed to obtain natural resource data usable for analysis at substate (multi–county) level, such as watersheds. The NRI records a variety of land use and resource conditions, including agricultural uses, cropping history, soil condition, conservation need and practices, and estimated soil and wind erosion.

  Data were obtained from the NRI to augment the pollutant loading data from the NCPDI. The NRI data were obtained for sample points in coastal watersheds (USGS cataloging units). For each coastal cataloging unit, estimates were obtained of total soil erosion, soil erosion from cropland, average rates of soil loss (tons/acre/year), and the number of acres which were identified in the survey as in need of some form of soil conservation treatment.

- The EPA’s AGTRAK database records citings of water quality impairments related to agricultural sources. The citings are taken from state inventories of nonpoint source pollution problems filed with EPA under Section 319(h) of the Water Quality Act of 1987 and Section 305(b) of the Federal Water Pollution Control Act. The AGTRAK database records the number of identified impairments related to pesticides, nutrients (nitrogen or phosphorus) and sediment in each county.

It should be emphasized that the AGTRAK data system only gives a qualitative assessment of water quality. Impairments are simply reported by the number of identified water quality problems in each county. There is no indication of the geographic extent of the reported problems (such as number of river miles or acres of lakes impaired. Also, “impairments” are rather loosely defined; impaired bodies are those the states have determined to be of insufficient quality to meet “designated uses.” Standards as to what constitutes "meeting designated uses” vary from state to state. Accordingly, the AGTRAK data should simply be used as an indicator that somewhere within a given county there have been identified impairments of surface water quality related to agricultural sources of pollution.

Data from these three sources have been combined to give an overall characterization of the sources and types of coastal water pollution and related resource conditions. Given to the amount of data, complete descriptive statistics are not presented here; summary tables are presented in the Appendix. Complete details on data sources and estimation procedures are available from the authors on request.

3. Data Analysis

Characteristics of Coastal Pollutant Flows

Data from the NCPDI covered 350 USGS cataloging units and 415 counties. Table 1 summarizes on a regional basis the relative shares of pollutant flows provided by point and nonpoint sources.

Overall, agricultural sources supply about forty percent of all nitrogen loadings, about thirty percent of phosphorus loadings, about 45 percent of sediment loadings, and about 28 percent of BOD5 loadings.[18] Agricultural loadings of nitrogen and sediment generally exceed point source loadings of these pollutants. (The figures reported in Table 1 are, of course, regional averages. Substantial variation is found among individual watersheds — see Appendix A).

We single out agricultural nonpoint sources (as opposed to all nonpoint sources) for several reasons. First, the EPA has identified agricultural nonpoint sources as the largest single component of nonpoint source pollutant

[18] Agricultural sources are defined here as pollutant loadings from harvested cropland, non-harvested cropland, pastureland, and rangeland. Other non-point sources in the NCPDI which are considered non-agricultural sources for our purposes include forestland, barren land, and urban non-point runoff.

loadings for the nation as a whole (USEPA 1990). Second, nonpoint source pollution controls on agricultural lands are less capital intensive (and thereby less expensive) than urban nonpoint controls (which involve installation of stormwater runoff control systems). Finally, other non-urban, non-agricultural sources of pollutant flows (such as runoff from forestland or barren land) are not readily controllable, at least to the extent that runoff from harvested cropland may be.

Table 1 also reports erosion conditions and soil conservation needs obtained from the NRI. Average erosion rates are highest in the East and Gulf regions. Also, the percentage of agricultural lands identified as needing some form of conservation treatment is highest in the Gulf and the East. Significantly, soil erosion is less severe in the coastal watersheds in the West: erosion rates in 1987 were less than 2 tons/acre/year: agricultural lands accounted for less than one-fourth of all erosion, and slightly less than 30 percent of agricultural lands were thought to need some form of conservation treatment.

Table 1

Sources of Coastal Pollution Loadings
And Related Erosion Data,
By Region

	Shares of Nitrogen Loadings		Shares of Phosphorus Loadings		Shares of Sediment Loadings		Shares of BOD5 Loadings		Average Cropland Erosion	Agland Share of all Erosion	Pct of Ag. Lands Needing Conservation Treatment
Region	Ag.	Point	Ag.	Point	Ag.	Point	Ag.	Point			
East	44.9	14.1	32.2	23.2	44.0	13.7	21.2	22.3	3.99	70.5	43.4
Gulf	38.7	19.3	10.2	42.5	57.8	3.5	33.5	8.1	2.90	63.7	65.7
west	44.1	13.6	33.4	25.3	43.3	6.6	28.5	23.4	1.99	24.7	28.9

Data based on cataloging unit-level estimates of pollutant loadings from NCPDI. Data on erosion and lands identified as needing conservation treatments from NRI. Agricultural loadings are loadings from cropland and pastureland. Point source loadings are loadings from wastewater treatment plants, powerplants, and industrial sources. Erosion estimates are tons/acre/year. Total number of coastal cataloging units assessed: 350.

Agricultural Impairments of Coastal Water Quality

Data from the AGTRAK database were examined to determine the extent of agricultural nonpoint source pollution in coastal counties[19]. State reports to the EPA indicate that half of the coastal counties contained at least one water body which did not meet designated uses due to pollution from agricultural nutrients (nitrogen or phosphorus). About a third of the coastal counties had at least one water body impaired by agricultural sources of

[19]The AGTRAK database is recorded on a county-by-county basis, rather than by USGS Cataloging Unit. The NCPDI data are available on either basis. Although it is technically feasible to construct county-level estimates of erosion and conservation needs from the NRI, the NRI was not designed to give statistically meaningful estimates at the county level. Accordingly, county-level data were not drawn from the NRI.

sediment. To get some idea of the relationship between agricultural pollutant loadings and impairments, the data from AGTRAK were merged with data from NCPDI, and screened for “significant” agricultural pollutant loadings. We define “significant” agricultural loadings in three ways: 20 percent, 25 percent, or 30 percent of total nutrient or sediment loadings. Between 114 and 132 counties out of 415 showed both identified agricultural impairments and significant agricultural sources of nutrients. Between 84 and 96 counties showed both significant agricultural sediment loadings and identified impairments from agricultural sediment.

Table 2
Agriculture’s Contribution Water Quality Impairments
in Coastal Counties

	Nutrients (No. of Counties)	Sediment (No. of Counties)
Counties with identified impairments from agricultural sources	225	155
Counties with identified impairments and agriculture contributes at least 20 percent of pollutant loadings	132	96
Counties with identified impairments and agriculture contributes at least 25 percent of pollutant loadings	122	87
Counties with identified impairments and agriculture contributes at least 30 percent of pollutant loadings	114	84

“Impairments” are defined as indication by state authorities that a county has a water body which does not meet designated uses in their 319 reports” to EPA. Data on impairments from AGTRAK database. Data on sediment and nutrient loadings from NCPDI, Total number of coastal counties assessed: 415.

Application of Screening Criteria for Potential Point-Nonpoint Trading

In order for point-nonpoint trading to contribute to overall water quality improvements in a watershed, several conditions have to be met. The data were examined to find how many watersheds satisfy some simple screening rules. The objective is to see how many coastal watersheds might pass a conservative test of their potential as sites for PS/NPS trading. One such test is to see in how many coastal watersheds both point and agricultural nonpoint loadings contribute “significantly” to total pollutant loadings. Simply put, if either the agricultural share of total pollutant flows or the point share of total loadings is small, then trading point and nonpoint reductions is unlikely to be feasible or to contribute much to water quality improvement.

Table 3 reports the number of coastal watersheds which satisfy some simple criteria of this sort. The data were examined to identify coastal watersheds where both point and agricultural nonpoint sources of pollutant loadings exceeded 20, 25, or 30 percent of total loadings each. We chose 20 percent of loadings by both point and nonpoint sources as a minimum criterion to ensure that there is enough potential for changes in loadings from point and nonpoint sources to affect overall water quality. Thirty percent of loadings from each class of pollutant source was the most conservative criterion; only a handful of water systems in the database had more than 35 percent of loadings coming from both pollutant sources.

Looking at the least strict criterion first, if we require that point and agricultural nonpoint sources both must account for al least 20 percent of total loadings, then out of 350 coastal watersheds 32 meet this requirement for nitrogen, 37 for phosphorus, 17 for sediment, and 32 for BOD5. If the requirement is that both point and nonpoint sources account for 30 percent each of total loadings, the numbers are considerably smaller: 16 watersheds meet this criterion for nitrogen and phosphorus, 13 for BOD5, and 8 for sediment. (See also Table 4 and Figures 1 - 3.)

Table 3

Coastal Watersheds Meeting Screening Criteria
For Potential Point-Nonpoint Trading
(Number of Watersheds)

	Nitrogen	Phosphorus	Sediment	BOD5
Cataloging Units with both point and agricultural sources supplying at least 20 percent of pollutant loadings	32	37	17	34
Cataloging Units with both point and agricultural sources supplying at least 25 percent of pollutant loadings	25	23	13	22
Cataloging Units with both point and agricultural sources supplying at least 30 percent of pollutant loadings	16	16	8	13

Data based on cataloging unit-level estimates of pollutant loadings from NCPDI . Agricultural loadings are loadings from cropland and pastureland. Point source loadings are loadings from wastewater treatment plants, powerplants, and industrial sources. Total number of coastal cataloging units assessed: 350.

Table 4
Distribution of Cataloging Units Meeting 30 Percent
Point, Nonpoint Pollutant Loading Shares
By Pollutant Category and Region

Region	Nitrogen Only	Phosphorus Only	Sediment Only	BOD5 Only	Multiple Pollutants
East (19)	4	6	3	3	BOD5, P:1 N,P:1 P, s: 1
Gulf Coast (13)	5	2	0	2	BOD5, N:4
west (11)	1	4	3	2	BOD5, N,P,S:1

Data based on cataloging unit-level estimates of pollutant loadings from NCPDI . Agricultural loadings are loadings from cropland and pastureland. Point source loadings are loadings from wastewater treatment plants, powerplants, and industrial sources. Total number of coastal cataloging units assessed: 350.

As revealed by Table 3 and 4 and Figures 1 - 3, it would appear that our initial screening does not show any particularly widespread potential for PS/NPS trading possibilities. Only at most 10 percent of the total number of coastal watersheds examined meet the most optimistic criteria we have established. Our analysis suggests PS/NPS trading might work in a few locations, but is unlikely to bring about NPS control in coastal regions nationally by itself.

Our results mirror those of a recent EPA study of all water bodies. The study examined information on waterbody impairments in 37 states, the District of Columbia, and two U.S. possessions for a count of the number of rivers, lakes, and estuary segments which a) do not meet designated uses from nutrient enrichment, and b) contain industrial point sources, municipal point sources, or both along with nonpoint sources (agriculture, silviculture, construction, resource extraction, land disposal, or hydro/habitat modification).

Their study showed that out of about 10,000 water bodies in their database not fully supporting designated uses due to nutrient loads, about 6 percent (618) impaired rivers, lakes, or estuaries could be considered for nutrient load trading.[20]

[20]Personal communication and memorandum supplied to the authors by Chris Faulkner, US EPA, Office of Water, Assessments and Protection Division.

Figure 1:

Eastern Cataloging Units Meeting Criteria

For Potential Point-Nonpoint Trading

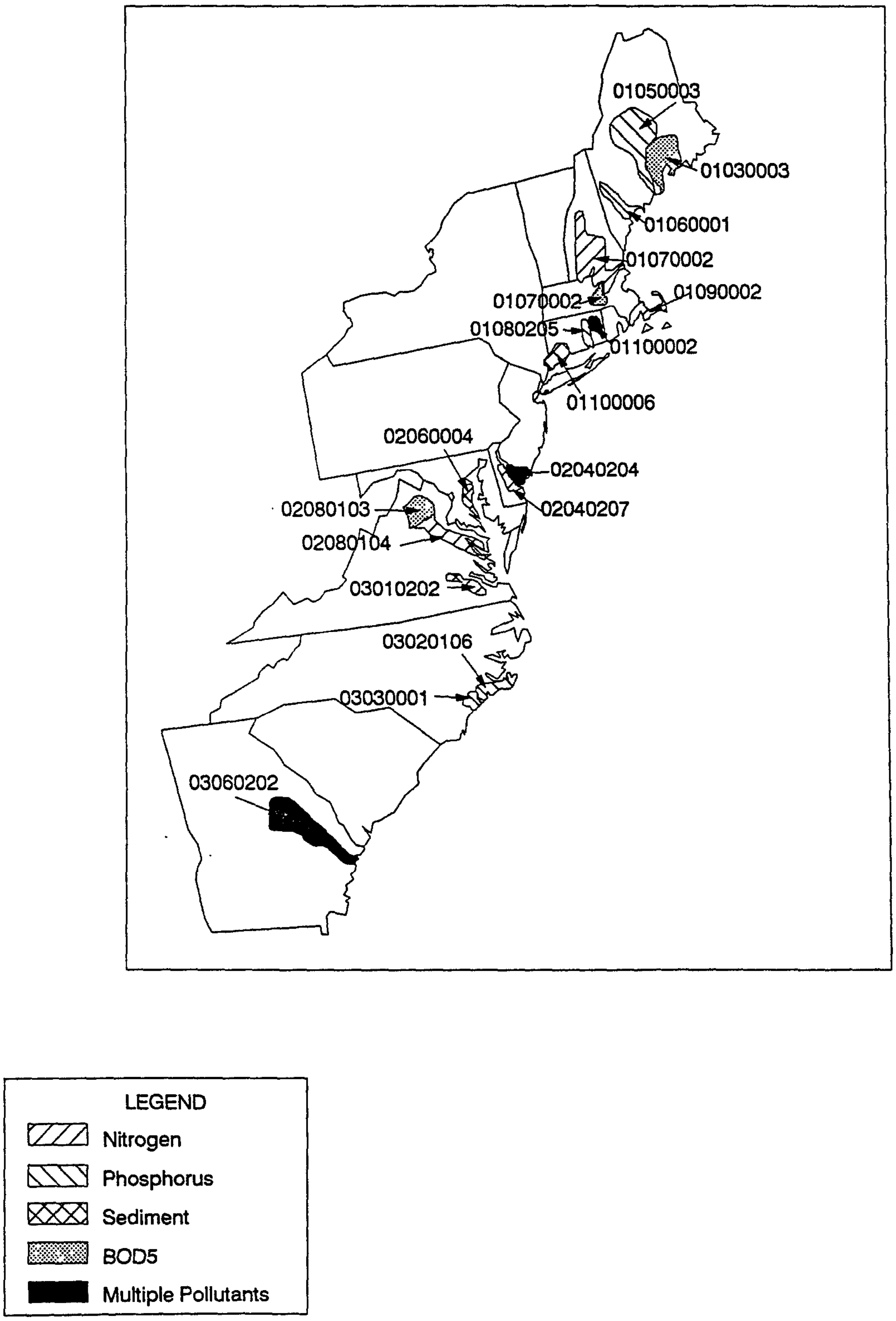


Figure 2:

Southern Cataloging Units Meeting Criteria
For Potential Point-Nonpoint Trading

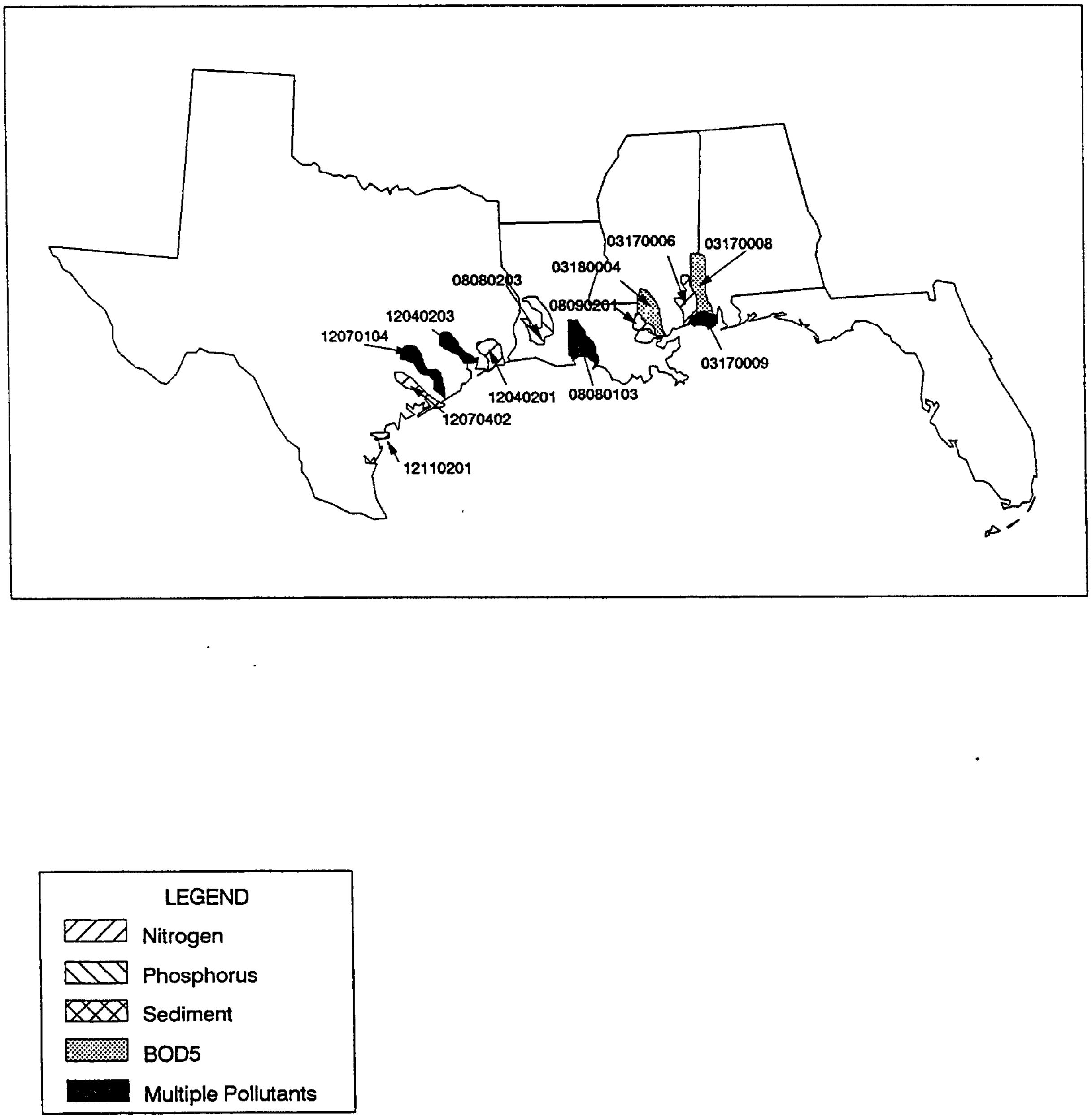
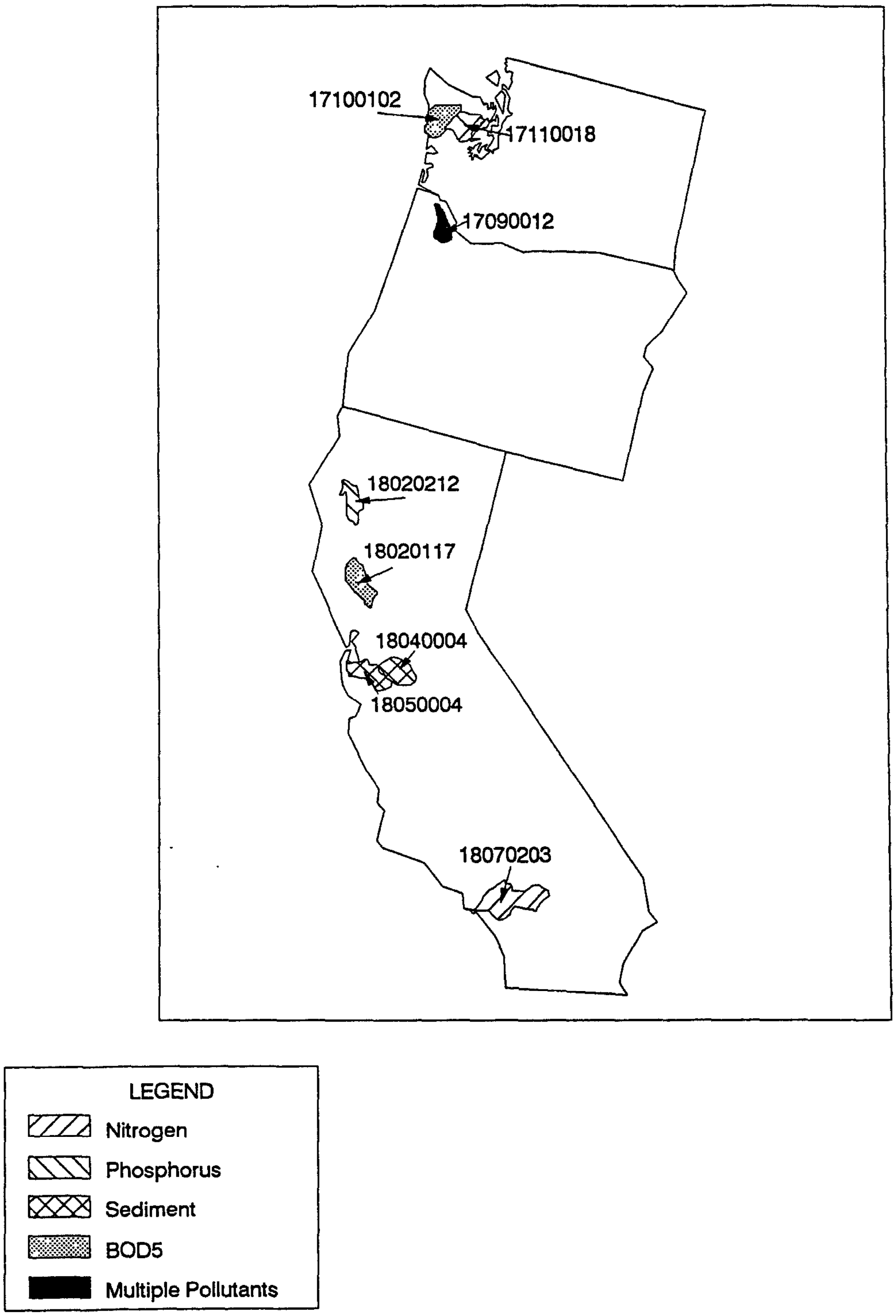


Figure 3:

Western Cataloging Units Meeting Criteria
For Potential Point-Nonpoint Trading



V. Conclusions

The purpose of this paper was to offer a national perspective on the feasibility of the PS/NPS trading option for coastal water quality management. We have used a conceptual model and a screening analysis to explore this question analytically and empirically. Our analysis indicates that it is unlikely that PS/NPS trading can form the basis of a national coastal water quality management program. It may, however, be a viable and attractive policy instrument in a few locations.

Our conceptual model showed that more attention needs to be paid to the issue of setting the appropriate trading ratio. Although it is often recommended that the ratio be set above one to allow for the randomness in nonpoint source loadings (as it has been for the Dillon Reservoir program), the greater difficulty in monitoring NPS loadings may call for a smaller ratio. Setting trading ratios above one may not result in cost effective outcomes.

PS/NPS trading does introduce some new problems because it approaches water quality management by attempting to create a market for pollution rights. A key element in properly coordinating such a market is having an appropriate number of participants. The number must be large enough so that the potential loadings reduction is of use to a PS and can be measured with some accuracy, yet small enough so that bargaining costs are not prohibitive. On the other hand, many of the obstacles to implementing PS/NPS trading programs apply to other approaches to NPS control also. The problems of implementing trading programs do not seem to be much greater than those associated with other approaches to controlling nonpoint source pollution.

We conducted a conservative screening analysis based on the consideration that the proportions of loadings contributed to a watershed by PSs and NPSs at least be consistent with the possibility of trading. Approximately ten percent of the coastal watersheds were reported by states as having "significant" contributions (twenty percent) of sediment, nutrient, or BOD5 loadings from both point and agricultural nonpoint sources. This nationwide screening analysis cannot locate "good" candidates for trading programs but does allow us to rule out many coastal watersheds, so researchers and planners can better focus their water quality efforts.

Acknowledgements: We would like to thank Dan Farrow, Chris Faulkner, Rich Kashmanian, Barry Korb, and Marc Ribaudo for their timely assistance in providing data and insights for this research. The usual rejoinder applies.

References

Anderson, G. Telephone conversation of 10 April 1991.

Basta, D., et. al. 1985. "The National Coastal Pollutant Discharge Inventory." Paper prepared for "Coastal Zone 85: the Fourth Symposium on Coastal and Ocean Management," Baltimore, MD July 30 - August 2, 1985.

Cease, R. "The Problem of Social Cost," Journal of Law and Economics 3 (1960): 1-44.

"Coastal Zone Reauthorization Contains New State Requirements for Coastal NPS Management; EPA-NOAA Share Federal Coordination," *EPA Nonpoint Source News* (December 1990): 4-6.

Coffey, S .; W. Berryhill; M. Smolen; and D. Miller. *Watershed Screening for Nonpoint Source Impacts and Controls.* Final Report for cooperative agreement 87-EXCA-3-8030, U.S. Environmental Protection Agency, Office of Policy, Planning, and Evaluation. (Draft). June 1989.

Crutchfield, S. "The Off-Farm Impacts of Agricultural Non-Point Source Pollution: Damages to Coastal Commercial Fisheries." Paper presented at the Western Economic Association meetings. July 1987.

Faulkner, C. "Number of Waterbodies in WBS not Fully Supporting Designated Uses for Nutrient Load Trading Consideration." Memorandum dated March 13, 1991. U.S. Environmental Protection Agency, Office of Water, Assessments and Water Protection Division.

Harding, D. "Tar-Pamlico Nutrient Sensitive Water Strategy," memorandum, North Carolina Department of Environment, Health, and Natural Resources/Division of Environmental Management. December 27, 1990.

Harrington, W., A. Krupnick, and H. Peskin. "Policies for Nonpoint-Source Water Pollution "Control." *Journal of Soil and Water Conservation,* January/February, 1985.

Kashmanian, R., J. Jaksch, D. Niedzialkowski, and M. Podar. "Beyond Categorical Limits: The Case for Pollution Reduction through Trading." Paper presented at the 59th Annual Water Pollution Control Federation Conference, 1986.

Krupnick, A. "Tradable Nutrient Permits and the Chesapeake Bay Compact." Resources for the Future Discussion Paper QE89-07. January 1989.

Letson, D. "Point/Nonpoint Source Pollution Reduction Trading: An Interpretive Survey." Forthcoming in *Natural Resources Journal.*

Milon, W. "Optimizing Nonpoint Source Controls in Water Quality Regulation" *Water Resources Bulletin* 23 (1987): 387-396.

National Oceanic and Atmospheric Administration, National Ocean Service. *Estuaries of the United States: Vital Statistics of a National Resource Base.* October 1990,

Russell, C. , W. Harrington and W. Vaughan. *Enforcing Pollution Control Laws.* Washington, D.C.: Resources for the Future, 1986.

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

Shabman, L. and P. Norris. "Coordinating Point and Nonpoint Control of Nutrient Pollution: Prospects for a Virginia Case Application." Working Paper SP-87-1O, VPI&SU Department of Agricultural Economics. November 1987.

Shortle, J. "The Allocative Efficiency Implications of Water pollution Abatement Cost Comparisons. "*Water Resources Research* 26(1990): 793-797.

Tietenberg, T. *Emissions Trading: An Exercise in Reforming Pollution Policy.* Washington, D.C. :Resources for the Future, 1985.

U.S. Environmental Protection Agency, Office of Water. *National Water Quality Inventory: 1988 Report to Congress.* EPA 440-4-90-003. April 1990.

U.S. Environmental Protection Agency, "The Cost of Clean Water." Notes from a Symposium on Market Forces and Alternative Financing held at Princeton, New Jersey, March, 1991.