Pollution Trading in Water Quality Limited Areas: Use of Benefits Assessment and Cost-Effective Trading Ratios

R. Scott Farrow, Martin T. Schultz, Pinar Celikkol, and George L. Van Houtven

ABSTRACT. *This paper proposes a water quality trading design that addresses common implementation problems. Trading ratios, which are calculated from damages integrated over each source's spatial zone of influence, drive the system to a socially cost-effective outcome. The design is applied to combined sewer overflow management in the Upper Ohio River Basin, where trading ratios can vary significantly among trading partners. The analysis shows that significant compliance cost savings are possible without incurring a penalty in terms of social damages or overall water quality despite a higher level of discharge relative to the command and control option.* (JEL Q53)

I. INTRODUCTION

Trading approaches to water pollution control have been advocated and implemented in a few watersheds but a successful design for general application remains elusive. Among those programs that have been implemented, the approaches vary widely to reflect the environmental conditions and regulatory objectives in watersheds across the country. However, a common feature of those programs is their failure to generate significant trading activity. Lack of trading has been attributed to a variety of causes including transaction costs, administrative complexity, lukewarm acceptance of trading among regulators, and program design (Hahn 1989; Atkinson and Tietenberg 1991; Malik, Letson, and Crutchfield 1993; Hoag and Hughes-Popp 1997). The search continues for a successful water pollution trading design that will help polluters achieve cost-effective compliance with regulations.

Trading ratios, the exchange rate at which pollution reductions from source j can be traded for pollution from source i, have been discussed in concept, used in practice, and are the central regulatory feature on which we focus. Trading ratios based on modeled estimates of relative damages are used as a mechanism to drive the system to a socially cost-effective outcome. We evaluate the damages integrated over each source's spatial zone of influence. This integration of impact is an important refocusing of regulation back to the source of pollution. Faced with a regulatory requirement to trade at the ratio of damages, the cost-minimizing behavior of independent, atomistic polluters leads to the same social cost outcome as a centrally planned allocation that minimizes social costs including both damages and control costs. A distinguishing feature of this framing is that the

Land Economics • May 2005 • 81 (2): 191–205
ISSN 0023-7639; E-ISSN 1543-8325

The authors are, respectively, chief economist, U.S. Government Accountability Office; research associate, Department of Engineering and Public Policy, Carnegie Mellon University; senior research associate, Duquesne University; and senior economist, RTI International. The authors wish to thank Tim Bondelid, Bob Hahn, Larry Lennon, James Shortle, Andy Solow, and Tom Tietenberg for information, comments, and advice. We also thank the Allegheny County Sanitation Authority and the Three Rivers Wet Weather Demonstration Program for access to data and information. This work was supported by U.S. Environmental Protection Agency (EPA)/National Science Foundation (NSF) Water and Watershed Science to Achieve Results (STAR) Grant # 4-82802101-0. The Center for Study and Improvement of Regulation at Carnegie Mellon University and the Andrew W. Mellon Foundation provided additional support. The work represents the author's views and does not necessarily represent the views of the funding agencies or of the U.S. Government Accountability Office.

commodity being traded is no longer a spatially distributed set of impacts at discrete measurement sites as is typical in the literature, but a single measure of monetary damages caused by each source within its zone of influence.

The proposed trading approach builds on existing regulatory structures, bureaucrats' revealed preferences for trading ratios, and utilizes the existing system of water pollution permits. It establishes, some would say re-establishes, the foundation of trading under certainty and linear damages functions to which it might be expected that uncertainty and higher order terms of damages offer variations. This trading approach is also consistent with quantity restrictions on emissions, such as total maximum daily loads (TMDLs) or combined sewer overflow (CSOs) control plans, which are currently being developed in many areas around the country. However, practical implementation of this approach requires a generally accepted means of estimating relative social damages from each source of pollution within the trading area, a procedure that has advanced in recent years but is still controversial.

In addition to presenting a trading program design in Section 2, this paper applies the trading approach to the regulation of CSOs in the Upper Ohio River Basin (UORB), a substantive problem that is expected to involve costs of over three billion dollars over the next 20 years. The application is also responsive to a recent national assessment of research needs in urban wet weather flow management that identified 69 priority research needs including flexible regulatory approaches (Heaney, Wright, and Sample 1999). In Section 3, trading ratios are calculated for the largest sources in the UORB, illustrative cost savings from trading are presented, and complexities relating to practical implementation are considered.

II. CONCEPTUAL STRUCTURE

In contrast to models of either uniformly mixed pollutants or pollution impact measured at a single site (Tietenberg 1985; Shortle and Abler 1998) usual models of non-uniformly mixed pollutants (e.g., Tietenberg 1985; Oates, Krupnick, and Van de Verg 1983) involve individual markets for environmental quality at multiple monitoring sites based on trading ratios reflecting the transfer coefficients from each source to each individual monitor. However, the location of monitoring sites can be arbitrary and coverage of the stream network in practice tends to be sparse. In addition, monitoring sites are not the natural unit of regulation. The power to control pollution is at the source. The trading approach developed below creates a regional market denominated in units of pollution from each source and adapts existing water quality permits and measures that are enforced at the source.

Theoretical Precursors

Undergraduate textbooks drill students on the efficient Pigouvian tax being the equality of the marginal social damage of pollution with the marginal social cost of control at the optimal level of output. Although typically discussed independently of spatial concerns, tax theory was extended by Tietenberg (1974) to spatially distinct damages where the optimal tax on emissions is the sum of spatial damages. At that time, prior to major developments in trading theory and contingent valuation, he was skeptical of a regulator's ability to satisfy the information requirements necessary to implement an optimal tax. He set derived-decision rules for fixed concentrations at selected monitoring sites to find the optimal tax. He notes that while the absolute level of the optimal tax may be difficult to fix, the ratio of taxes may be calculable; an insight that will be key to results here. Developments in trading theory since then have led to the expectation that trading solutions can replicate the tax solutions. Similarly, Hahn (1989) developed a generalized model for trading that was independent of spatial concerns, but focused on trade-offs in arbitrary policy concerns, such as safety and fuel efficiency. The trading regime involved a two-step process: firms exchanging permits with the government based on trading rates determined by political trade-offs, and then a market in permits that would generate permit prices equal to the government-defined exchange rate.

Uncertainty has been addressed in several articles as a basis for the use of trading ratios. In the case of point/nonpoint source trades, if loads can be characterized by skewed probability distributions such as the lognormal distribution, Shortle (1990) suggested that greater stochastic variability of pollutant loadings from nonpoint sources implies that reductions at the less-variable source yield greater environmental benefits. Since nonpoint sources typically have greater variance in their loading rates, greater emissions reductions are needed at nonpoint sources to compensate for this difference. In addition to stochastic variability of loadings, Malik, Letson, and Crutchfield (1993) incorporated uncertainty in the effectiveness of controls, accuracy of monitoring, and prediction of control costs. They argued that greater uncertainty in these factors interferes with participants' expectations of cost savings and regulators' expectations of compliance. However, there are alternative conceptual developments as are summarized by Rodriguez (2000) and some argue that trading ratios interfere with achieving cost effectiveness. Klaassen (1996) suggests that, in practice, agencies tend to determine trading ratios on an arbitrary basis.

We seek to identify a core basis for the existence of trading ratios and empirical approaches that might support wide-spread implementation. Like Tietenberg, we acknowledge spatially differentiated damages. Like Hahn, we frame the cost-minimization problem in terms of a government-defined exchange rate. We demonstrate that uncertainty is not required for trading ratios that are different than one and, although we do not incorporate uncertainty in the definition of the trading ratio, it becomes a factor subsequently through Monte Carlo simulation of damages and the choice of a design rainfall event.

Social Cost-Effectivenes

Define social cost-effectiveness as the least cost to achieve a given level of social damages from water pollution. The level of damage can be politically chosen, although one choice is an economically efficient level of damages that balances marginal benefits and costs. The geographic unit can be a watershed, a state, or the nation which is here undefined for generality. A linear marginal social damage from each source j, d_j is assumed, although a non-linear damage function could be constructed and provides a direction for future research. Consistent with this choice, we work with the first order approximations that we think are consistent with knowledge about economic values and social damages. Using notation similar to that of Tietenberg (1985), the social cost minimization problem[1] is to minimize the sum of costs from all sources j, c_j, from controlling pollution, R_j, subject to an aggregate social damage constraint $\bar{s}$.

Specifically

$$\begin{aligned} &\underset{R_j}{Min}\sum_{j=1}^{J} c_j(R_j) \\ &s.t.\sum d_j(\bar{e}_j - R_j) \le \bar{s} \\ &\quad R_j \ge 0, \end{aligned} \qquad [1]$$

where

d_j	damage (absolute value) per unit of pollution from source j
R_j	reduction in emissions (emissions controlled) from source j
$c_j(R_j)$	cost of controlling pollutant emissions at source j
$\bar{e}_j$	maximum uncontrolled emissions at source j
$(\bar{e}_j - R_j)$	residual, uncontrolled pollution, also denoted e_j as the loading from source j

[1] The minimization problems in this paper are presented using the more intuitive "less than or equal" constraint although the first order conditions are derived by multiplying by –1 and creating the appropriate "greater than or equal" constraint. Arrow-Enthoven sufficiency conditions are met by linearity of the constraints.

$\bar{s}$ cumulative social damage (absolute value) limit from pollution

The necessary conditions for an internal solution, where primes (′) indicate partial derivatives, imply

$$\frac{c'_{R_j}}{c'_{R_i}} = \frac{d_j}{d_i}. \quad [2]$$

The socially cost-effective solution is to equate the ratio of marginal costs to the ratio of unit damages. Note that this implies that marginal costs are not equal unless damages per unit are equal. This equality of ratios is similar to results by Malik, Letson, and Crutchfield (1993), although their focus is on physical transfer coefficients. It is also a special case of Hahn (1989) where his linear aggregator is replaced by social damages from each firm and the constraint is the level of social cost. Here, the level of social cost to be achieved changes the shadow price, λ, but not the ratio form of the solution. Hence whatever level of control the regulator seeks, its solution will be of this form.

Regulatory Design and the Firm's Problem

The regulatory design problem is how to create incentives so that firms achieve the social solution while pursuing their own self-interest. Consider a regulatory design, assuming that per unit social damages integrated over the zone of influence, d_j can be assigned to each source j, such that

- Each firm meets a firm specific quantity limit $\bar{x}_j$, either through on-site controls, through emissions reductions purchased elsewhere, or through a combination of the two.
- Firms may trade reductions at one source for pollution at another one source at any price as long as incremental social damages from the trade are offsetting, $e_i d_i - R_j\, d_j = 0$ indicating the reduction at source j to offset a unit pollution from source i is $\frac{R_j}{e_i} = \frac{d_i}{d_j}$. This ratio, the number of reductions from j to offset a unit of pollution from i, is the trading ratio for this direction, from j to i. Equivalently, the credits (allowable emission) source i obtains by purchasing reductions at source j is $\frac{d_j}{d_i} R_j$ (noting that later notation will distinguish between units of control that are retained and those that are sold).

Firm i's problem given this regulatory design involves meeting its permitted level of pollution by minimizing its control costs and the cost of permits purchased at the fixed-trading ratio set by regulation, less any arbitrage opportunities to sell its own controlled pollution (at a profit). The firm may either control units of pollution and keep them for regulatory purposes, r_{ki}, or control pollution and sell the units, r_{si}. The cost of control is a function of the total amount of pollution controlled, R, equal to the sum of r_{ki} and r_{si}, so that the marginal cost is unaffected by the ultimate use of the control. However, only units of pollution that are controlled and for which ownership is retained can be used to meet a quantity pollution constraint. Purchased units of control count at the credit rate to meet the emission constraint and can augment the retained units of control.

$$\underset{r_{ki},r_{si},r_{sj}}{Min}\; c(R_i) - p_i r_{si} + \sum_{j=1}^{J} p_j r_{sj}$$

$$s.t.\; (\bar{e}_i - r_{ki}) - \sum \frac{d_j}{d_i} r_{sj} \le \bar{x}_i \quad [3]$$

$$r_{si}, r_{ki}, r_{sj}, \lambda \ge 0,$$

where

$\bar{x}_i$ allowable emissions limit at source i

$\frac{d_j}{d_i}$ credit rate for source i obtaining an offset from source j, based on trading ratio

p_i, p_j market price for a permit to emit a unit of pollution from source i or j

r_{sj} units of pollution control purchased from source j (sold by source j) to offset units of pollution from i

r_{ki} units of pollution control from source i that is kept by source i to meet requirements

r_{si} units of pollution control sold by source i

$R_i = r_{ki} + r_{si}$ Total pollution controlled at source i

This optimization has the necessary conditions for an interior solution for r_{si}, r_{ki}, and r_{sj} (which are also sufficient given usual convexity conditions for the control cost function):

for units sold,

$$c'_{R_i} - p_i \geq 0; \qquad [4]$$

for units kept,

$$c'_{R_i} - \lambda \geq 0; \qquad [5]$$

for units purchased,

$$p_j - \lambda \frac{d_j}{d_i} \geq 0 \ \forall_j. \qquad [6]$$

A variety of special cases emerge from these equations in conjunction with the complementary slackness conditions. If the quantity constraint is not binding so that λ is zero and no market exists for pollution control (a preregulatory scenario), then the marginal cost is zero and there is no control. If the constraint is binding so that λ is non-zero, but the firm optimally meets its constraint by controlling its own pollution, then the marginal cost of control equals the shadow price. In that case, the cost of purchasing reductions from other sources is larger than the shadow price and does not enter the least cost solution.

When the firm meets its quantity constraint with a portfolio of its own pollution control and purchased control from other sources, then both equations [5] and [6] are met with equalities. Substituting for λ,

$$c'_{R_i} = p_j \frac{d_i}{d_j}, \qquad [7]$$

and also,

$$\frac{p_j}{p_k} = \frac{d_j}{d_k}. \qquad [8]$$

Consider that firm j is solving a similar minimization problem and if it is selling pollution control, its equivalent to equation [4] above must hold. Consequently the trade between firms i and j must satisfy

$$\frac{c'_{R_i}}{c'_{R_j}} = \frac{d_i}{d_j} = \frac{p_i}{p_j}. \qquad [9]$$

The equality of the ratio of marginal cost with the ratio of marginal damages is exactly what is required for social cost minimization from equation [2]. In addition, the regulatory requirement that trades take place based on the damage ratio yields a price ratio that is also equal to the damage ratio. Should firm i find itself facing a situation in which it both buys and sells pollution control (likely to be a disequilibrium situation) equation [9] follows more directly because equation [4] would hold as an equality for firm i as well. The result is that regulatory use of a specific trading ratio can be socially cost effective in the case of non-uniformly mixed pollutants and the ratio can be used to drive the firms to a socially cost-effective outcome. Note that while there have been advances in the estimation of social damages and preliminary estimates of damages by source exist, such damages can be incorrect by a factor of proportionality in this linear marginal damage model and still lead to optimality. We anticipate that the national regulatory agency would estimate the damages and publish pre-approved trading ratios, or publish trading ratios for local agency use that could differ by some agreed upon proportion to reflect local conditions.

Extensions and Caveat

The firm's problem in the linear damage function case can alternatively be written as constrained by its social cost, since multiplying both sides of the firm's constraint by d_i will not change the solution. While this structuring is more parallel to the social-cost problem and implicitly, the notion of damage trading, it distracts from emphasis on the trading ratio that is emphasized here for regulatory purposes.

The basic trading design presented above can also be modified to reflect various policy objectives such as concern about existing environmental quality, "hot spots," or equity issues. Several authors have modeled these concerns as constraints in air quality trading models, noting that as more constraints are imposed, it is quite possible to remove all gains from trade such that no trades would be expected to occur (e.g., Montgomery 1972; Atkinson and Tietenberg 1982; Krupnick, Oates, and Van De Verg 1983; McGartland and Oates 1985; Rodriguez 2000). We anticipate that similar results would occur in the water context although our case study provides an example where correlation between water quality and economic damages results in overall water quality improving compared to a uniform reduction requirement.

The recent literature on optimal taxation with pre-existing distortions also provides a caveat to a partial equilibrium analysis of trading suggesting that the welfare gains may not be as large as anticipated and could even be negative in the presence of pre-existing taxes and a fixed government budget constraint. The degree of caution is still evolving on the extent of welfare improvement (e.g., Fullerton and Metcalf 1998; Farrow 1999) but is not investigated here.

III. DAMAGE ESTIMATION

If a regulatory agency were to pre-set bilateral, firm-specific trading ratios, how would it proceed? Estimation of water quality impacts and monetary damages from a set of pollution sources requires a consistent, widely applicable, theoretically acceptable, and practical method of modeling water quality impacts and estimating the value of those changes in water quality. The Environmental Protection Agency (EPA) has recently undertaken an attempt to develop a national water quality model that could conceivably fill this gap, the National Water Pollution Control Assessment Model (NWPCAM), described in Bondelid, Unger, and Stoddard (2000).

In this section, we develop a method to estimate water quality impacts and economic damages over the zone of influence of a pollution source. We begin with water quality impacts. At the time of this study, NWPCAM was not sufficiently developed that it could be used effectively in this study, but the pollution loading rates and hydrologic assumptions used here are based on that model.[2] The contribution of each source to downstream pollutant concentrations over the zone of influence is estimated using a one-dimensional fate and transport model with a lumped-parameter first-order decay term following standard methods described in Thomann and Mueller (1987) and Chapra (1987). We focus on these contributions with background and other sources assumed into terms that do not affect the relevant derivatives. The plug flow model estimates pollutant concentration due to emissions at the source, C (mg/L), at distance, n (meters), downstream from a waste source, i, as a function of the waste emissions rate at that source, e_i (kg/day), equal to $(\bar{e}_i - R_i)$, stream flow, Q (meter^3/day), and an exponential decay term:

$$C_{ni} = \frac{e_i}{Q}\exp\left(-k\theta^{(T-20)}\frac{n}{U}\right). \qquad [10]$$

In the decay term, k (day^{-1}), is the nominal decay rate; θ is a coefficient reflecting sensitivity of k to temperature, T (Celsius) is the mean summer temperature; and U (meters/day) is stream velocity. C_n is calculated assuming steady-state conditions at approximately one-mile intervals below the source. Characterizing the problem using first order decay asserts that changes in C_{ni} at any point in the stream network are proportional to the change in concentration from the source at the point of dis-

[2] The National Water Pollution Control Assessment Model (NWPCAM) Version 1.1 estimates the concentration of selected conventional pollutants in 600 thousand stream miles within the 48 contiguous states and classifies each stream mile as satisfying water quality criteria for swimmable, fishable, or boatable use (Bondelid, Unger, and Stoddard 2000). Changes in use-support can then be linked to estimates of the value of use-support improvements (Carson and Mitchell 1993).

charge C_{0i} (equal to e_i/Q_{0i}). Therefore, the impact of a change in emissions at some downstream location, n, can be calculated using a transfer coefficient. This dimensionless transfer coefficient, $\tau_i(n)$, is the ratio of concentrations at location n and the point of discharge for a given emissions level,

$$\tau_i(n) = \frac{C_{ni}}{C_{0i}} \qquad [11]$$

In order to link changes in water quality to economic damages per unit from source i, we specify a total benefits function and evaluate the absolute value of the loss in benefits, the damages, due to increased emissions from the source. Define total benefits for a unit length of a river as

$$B_n = V(W_n, H_n), \qquad [12]$$

where B is the benefit, W is water quality, and H are households. We assume that water quality is measured by the concentration of pollution in segment n, C_n, which as above, depends on the loading at the source and physical characteristics (abstracting from other sources due to assumed linearity of damages.) Hence, W_n is described by the following nested functions:

$$W_n = C_n(C_0(e_i)), \qquad [13]$$

$$\frac{\partial W_n}{\partial C_n} = -1; \frac{\partial C_n}{\partial C_0} = \tau_i(n); \frac{\partial C_0}{\partial e_i} = \frac{1}{Q_{0\mathrm{i}}}. \qquad [14]$$

Turning to values, the marginal value of a change in water quality is frequently estimated based on surveys of willingness-to-pay at the household level, *WTP*. These studies often attempt to estimate *WTP* on a per household basis from information about age, race, income, education, and personal preferences, as well as information about baseline water quality conditions and the distance of survey respondents to the resource (Desvousges, Smith, and Fisher 1987; Carson and Mitchell 1993; Farber and Griner 2000; Magat, Huber, and Viscusi 2000). However, these studies vary widely in their methods, assumptions, and characterization of water quality change. The relationships between *WTP* and socioeconomic variables that are described in these studies are frequently weak. Studies that address national- or regional-scale changes in water quality (e.g., Mitchell and Carson 1993; Magat, Huber, and Viscusi 2000) can only be adapted to value site-specific changes through numerous assumptions about benefits transfer. Given these practical difficulties, we assume that the household marginal willingness-to-pay for a small improvement in water quality, *WTP*, is constant over socioeconomic variables and over the range of water quality conditions considered in this study. This assumption seems most appropriate where changes in water quality are relatively small. The marginal value then depends on the number of households, H_n, affected and the marginal value for a change in water quality per household:

$$\frac{\partial V(H_n, W_n)}{\partial W_n} = WTP^*H_n. \qquad [15]$$

Damages from source i for segment n, d_{ni}, are then based on the value of the marginal reduction in benefits from an increment of pollution at the source. Through iterative use of the chain rule we obtain

$$d_{ni} = \left|\frac{\partial B_n}{\partial e_i}\right| = \left|\frac{\partial V_n}{\partial W_n}\frac{\partial W_n}{\partial C_n}\frac{\partial C_n}{\partial C_0}\frac{\partial C_0}{\partial e_i}\right|, \qquad [16]$$

substituting and integrating over stream segments yields

$$d_i = \int_{n=0}^{n}\left[(WTP^*H_n)^*\tau_i(n)^*\frac{1}{Q_{0i}}\right]dn. \qquad [17]$$

These marginal damages per unit of pollution from the source have units of dollars per kilogram per unit time.

If the transfer coefficient, $\tau_i(n)$, is continuous, smooth and decreasing, then the damages avoided by reducing the emissions of a single pollutant can be estimated from a continuous analytical form. However, in many cases, there is no reason for $\tau_i(n)$ to be smooth. For example, if the zone of influence extends to a different stream segment, with a greater flow rate, one would expect an instantaneous drop in $\tau_i(n)$ as a result of dilution at the confluence. Given the many confluences of streams in the network, we lack a smooth, explicit function for $\tau_i(n)$. In this case, damages from one unit of emissions can be estimated numerically using discrete intervals, ℓ_n, and substituting for $\tau_i(n)$ using a baseline measure, C_{0i}^b of the initial concentration due to pollutant loadings at the source i:

$$d_i \cong \sum_{n=0}^{N} (WTP \cdot H_n) \cdot \ell_n \frac{C_{ni}^b}{C_{0i}^b} \frac{1}{Q_{0i}}. \qquad [18]$$

A further simplification results from our decision to assume a marginal WTP for improved quality per household that is constant over socioeconomic variables and water quality conditions. In the ratio structure of damages, the unit WTP per household cancels in a ratio of damages from different sources. Consequently, a value for WTP is not required. However, it is necessary to define the affected number of households for each n,

$$\frac{d_i}{d_j} \cong \frac{WTP * \sum_{n=0}^{N_i} H_n \ell_n \frac{C_{ni}}{C_{0i}} \frac{1}{Q_{0i}}}{WTP \sum_{n=0}^{N_j} H_n \ell_n \frac{C_{nj}}{C_{0j}} \frac{1}{Q_{0j}}} = \frac{\sum_{n=0}^{N_i} H_n \ell_n \frac{C_{ni}}{C_{0i}} \frac{1}{Q_{0i}}}{\sum_{n=0}^{N_j} H_n \ell_n \frac{C_{nj}}{C_{0j}} \frac{1}{Q_{0j}}}. \qquad [19]$$

Empirically, we define the number of households affected by a water quality change, H_n, in a stream interval, l_{n}, as equal to the number of households in the county in which the stream segment is located. Although this approach acknowledges the concept of population-dependent damages, it does not address the issue of proximity to the resource within a county and assumes the willingness-to-pay for the resource drops to zero outside the county. Alternative household allocation approaches could be developed but are not investigated in this study.

IV. COMBINED-SEWER OVERFLOWS IN THE UPPER OHIO RIVER BASIN

This section discusses prospective use of the damage trading approach to water quality management in the Upper Ohio River Basin (UORB), a region impacted by approximately 70 sewer systems with combined-sewer overflows (CSOs). Preliminary estimates suggest that the control of CSOs within the UORB will cost more than three billion dollars over the next 20 years. This application assesses whether a damage ratio trading program based on estimated damages would appear fundamentally different than the regulatory default of uniform percentage removal from each source. We estimate a matrix of both water quality impacts and of trading ratios based on economic damages for the eight largest CSOs in the region and discuss aspects of implementation in this case to confront specific issues that might impede putting this theory into practice.

A combined-sewer system receives surface runoff during rainfall events and contains built-in pressure relief points that discharge raw sewage directly to surface waters when service capacity is exceeded. CSOs are the portion of flow from these sewage systems that enter receiving waters without wastewater treatment. CSO impacts arise from the presence of bacteria, biological oxygen demand (BOD), and total suspended solids (TSS) in CSO discharge. In 1994, EPA issued a policy statement requiring combined-sewer system to characterize their sewer systems and implement technology-based effluent limitations in the form of nine minimum controls before 1997 (EPA 2001). Combined-sewer systems were further directed to prepare long-term control plans that evaluate a range of alternatives for Clean Water Act (CWA) compliance, including attainment of water quality criteria. Communities may adopt either a presumptive or a demonstrative approach to confirm their compliance with water quality goals. A presumptive approach is a limitation on the frequency and volume of overflows. Under this approach, each long-term control plan may propose a level of control that it expects will maintain the desired level of water quality, or may opt to reduce the frequency of CSO events to four per year (except that six may be allowed in some cases) and the volume of CSO emissions by 85%. A demonstration approach provides greater flexibility in establishing performance requirements for the system, but requires communities to monitor water quality and demonstrate that the chosen performance criteria achieve compliance with water quality goals (EPA 2001). EPA guidance does not specifically include pollution trading as a regional-management option. However, EPA's policy regarding CSO control has been evolving since the initial publication in 1994 and the progress to implement CSO controls has been slower than planned. EPA also encourages a watershed approach and has discussed trading approaches that may allow room to develop creative and cost-effective strategies for compliance with CSO regulations (EPA 1996, 2002).

Several types of pollutants in CSO effluent create public health and environmental concerns. However, analysis can often focus on one indicator pollutant given local priorities and concerns because the effects of control on water quality are highly correlated among the different pollutants. For example, bacteria, commonly measured in terms of fecal coliform (FC) concentration (most probable number (MPN)/100mL), may be the principal concern of residents located in proximity to outfalls, but relatively unimportant to residents further downstream since water temperatures are too cold for bacterial pathogens to persist in the water column and sediments much more than a day. BOD is a measure of the organic waste content in water, measured as the amount of oxygen demand (mg of O_2/L) exerted through bacterial decomposition of that waste over a specified period of time. Effects of BOD on water quality are both aesthetic and chemical. BOD is associated with generally foul-smelling and unpleasant conditions, which are aesthetic concerns, and low dissolved oxygen (DO) concentrations, which affect the integrity of the biotic community and, in extreme cases, can cause anoxic water quality conditions. The DO deficit caused by CSO discharges may be a more important issue for distant residents, because this effect is more pronounced at locations further downstream, after pathogenic bacteria cease to be an issue. This study uses BOD as an indicator pollutant because it has aesthetic and ecological impacts throughout the proposed trading region. Damage coefficients are calculated from estimates of BOD concentration using design flow analysis and a one-dimensional, steady-state plug-flow water quality model with a temperature-sensitive, lumped-parameter, first-order decay coefficient, as described above.[3] BOD emissions from each source were estimated for a five-year, one-hour storm event for each combined-sewer system generated in preparation for the 1992 Clean Water Act Needs Survey using information on system drainage, population served, and impervious area.

We analyze the eight largest CSOs in the UORB: Clairton, Greensburg, Pittsburgh, McKeesport, Morgantown, Steubenville, Uniontown, and Youngstown. Each CSO is located within a 60-mile radius of the confluence of the Allegheny and Monongahela Rivers to form the Ohio River in southwestern Pennsylvania. For the purpose of these estimates, the zone of influence extends approximately 260 river miles below Pittsburgh to the base of the Upper Ohio-Shade watershed, a distance that requires several days of travel time during summer months. Emissions enter receiving waters at the National Pollution Discharge Elimination System (NPDES) location of

[3] In agency guidance for monitoring and modeling of CSO impacts, EPA suggests several types of analysis. The simplest, design flow analysis, is based on a one-dimensional, steady-state model that determines whether CSO discharge causes a violation of water quality criteria. This approach requires no information on CSO overflow frequency; data that is not presently available for many CSOs, and often excludes concerns over the time and distance required for mixing of CSO effluent in large rivers below outfalls (EPA 1999). Elaborations of the basic modeling techniques determine the frequency of excursions by adding data on CSO overflow frequency and variability in stream flow, and time-variable watershed simulations.

TABLE 1
EXPECTED VALUE OF DAMAGE COEFFICIENTS FOR EIGHT CSOs IN THE UORB

	Damage Coefficients (kg/m^2)	
Combined-Sewer Systems	Expected Value	Standard Error
Clairton, Pennsylvania	0.15	0.05
Greensburg, Pennsylvania	2.51	1.45
McKeesport, Pennsylvania	0.09	0.03
Morgantown, West Virginia	0.35	0.15
Pittsburgh, Pennsylvania	0.09	0.03
Steubenville, Ohio	0.05	0.01
Uniontown, Pennsylvania	0.24	0.09
Youngstown, Ohio	0.51	0.18

record, which reflects the site of the associated wastewater treatment plant. When there are multiple CSO outfalls in a combined sewer system, it can be considered a diffuse source of emissions. However, in the context of large hydrologic systems, treating CSOs as point sources can be considered a reasonable approach.

The first estimates, reported as expected values in Table 1, are the water quality effect of a unit of emissions. These damage coefficients are the biological oxygen demand (mg/L) resulting from the discharge of 1 kg of BOD load integrated over the zone of influence and evaluated in terms of BOD per unit of stream cross-sectional area at the source (kg/m^2). These damage coefficients are not weighted by population. Expected values are from a Monte Carlo simulation that accounts for uncertainties in the hydrologic conditions and water quality modeling coefficients as described in Schultz et al. (2004). Receiving water characteristics can have a large effect on these damage coefficients. For example, high impacts for Greensburg may be attributed to the relatively low flow in the receiving stream segment at Greensburg.

Water quality impacts reflect the assimilative capacities of receiving waters and are independent of the amount of loadings from other sources. Greensburg has a high damage coefficient because assimilative capacity in the first 25 kilometers of receiving waters is low. Pittsburgh and McKeesport, both located on the lower Monongahela River, have similar damage coefficients because there is considerable overlap in their zones of influence. The smallest damage coefficient is associated with Steubenville, the southern-most of these combined sewer systems on the Ohio River. Table 2 lists the *ratio* of water quality changes (using unrounded data) still unweighted by population. The ratio of expected value of water changes for Morgantown, West Virginia, relative to McKeesport, Pennsylvania, is 3.74 (cell 4,3). This ratio indicates Morgantown would have to purchase 3.74 units of emissions reduction at McKeesport to offset one unit of its own emissions if the trading ratio were based on unweighted changes in water quality. A unit load trade in the opposite direction is governed by the inverse of this ratio, 0.28. The highest ratio is 53.50 (cell 2,6) for Greensburg, Pennsylvania, obtaining offsets at Steubenville, Ohio. Every unit of emissions offset at Greensburg by emissions reductions at Steubenville would require 53.50 units at Steubenville.

Economic damages are incorporated into the ratio of water quality damages by multiplying the number of households in each of the counties intersected by a particular stream segment. The 29 stream segments impacted by the CSOs considered in this study intersect 14 counties. The largest number of counties intersected by any one stream segment is four. The effect of including economic values on the matrix of trading ratios is reported in Table 3. Trading ratios can be higher or lower relative to those in Table 2 depending upon the

TABLE 2
Ratio of Expected Value of Damage Coefficients

Regulated Source i	Source j: Source of Pollution Offsets 1	2	3	4	5	6	7	8
1 Clairton	1.00	0.06	1.57	0.42	1.60	3.17	0.61	0.29
2 Greensburg	16.87	1.00	26.56	7.10	26.99	53.50	10.32	4.89
3 McKeesport	0.64	0.04	1.00	0.27	1.02	2.01	0.39	0.18
4 Morgantown	2.38	0.14	3.74	1.00	3.80	7.54	1.45	0.69
5 Pittsburgh	0.63	0.04	0.98	0.26	1.00	1.98	0.38	0.18
6 Steubenville	0.32	0.02	0.50	0.13	0.50	1.00	0.19	0.09
7 Uniontown	1.63	0.10	2.57	0.69	2.61	5.18	1.00	0.47
8 Youngstown	3.45	0.20	5.44	1.45	5.52	10.95	2.11	1.00

number of households affected by changes in water quality. For example, there is a noticeable increase in the trading ratio for obtaining credits from Steubenville due to the relatively small number of households affected by this source.

Trading can also be facilitated by the existence of differential control costs among sources (see equation 2), although it is not a requirement for trading. Lower-control costs can compensate for high-damage coefficients and make these systems more attractive sources of load reductions. While control costs are largely unknown, some studies suggest that differential CSO control costs do exist in the Pittsburgh region. The Draft Long Term Control Plan (ALCOSAN 1999) investigated various combinations of individual capital improvements that exhibit economies of scale up to some capacity size including centralized treatment, storage basins, deep tunnels, swirl concentrators, and screens. One or more of these technologies is applicable at each of the combined sewer systems throughout the UORB, so technologies and capital improvements could be selected and sited to minimize control costs. It may be expected that the total cost of control, aggregated over control alternatives could exhibit constant or increasing economies of scale. The plan further notes that approximately 40% of dry weather flows resulting from inflow or infiltration could be reduced, alleviating the need for end-of-pipe treatment and control investments. An engineering cost estimate in the Pittsburgh region (Lennon 2001) suggests strongly increasing incremental control costs for controlling inflow or infiltration.

The potential impact on water quality, economic damages, and cost of control as

TABLE 3
Trading Ratios as the Ratio of the Expected Value of Damage Coefficients Weighted by the Number of Affected Households

Regulated Source i	Source j: Source of Pollution Offsets 1	2	3	4	5	6	7	8
1 Clairton	1.00	0.25	2.64	0.62	2.80	35.24	0.45	0.72
2 Greensburg	3.99	1.00	10.53	2.46	11.17	140.78	1.81	2.86
3 McKeesport	0.38	0.09	1.00	0.23	1.06	13.37	0.17	0.27
4 Morgantown	1.62	0.41	4.28	1.00	4.54	57.22	0.73	1.16
5 Pittsburgh	0.36	0.09	0.94	0.22	1.00	12.60	0.16	0.26
6 Steubenville	0.03	0.01	0.07	0.02	0.08	1.00	0.01	0.02
7 Uniontown	2.21	0.55	5.82	1.36	6.18	77.85	1.00	1.58
8 Youngstown	1.40	0.35	3.68	0.86	3.90	49.19	0.63	1.00

a result of trading is compared below to the current default technological control standard of 85% removal at all sites. Illustrative linear and quadratic total cost of control functions are used that are the same across all sites. The trading solution is obtained for each type of cost function based on evaluating c_i'/d_i with sites having the lowest marginal cost per unit damage entering the solution until social damages are reduced to the same level achieved from uniform 85% reduction. When a site reaches full control, its marginal cost is assumed infinite. With a linear total cost function, implying constant marginal costs that here are normalized to 1, the low marginal cost per unit damage sites control all their pollution. With the quadratic total cost function and its linear marginal cost function, some sites control all pollution but other sites have an internal solution resulting in equality of the marginal cost per unit damage. As with similar studies of air pollution, the least cost solution should be considered an upper bound on cost savings.

Pollution, water quality, social damage indices, and costs for the several scenarios are presented in Table 4. The four scenarios are (1) no control, (2) uniform 85% removal, (3) linear total cost, and (4) quadratic total cost. The data on BOD loadings, water quality damage, and weighted damage are from the study area as above with impacts measured by BOD loading times the appropriate impact coefficient. The uniform removal option reduces pollution by 85% or 222 kilograms, leading to an overall water quality impact after control of 9, social damage index of 1.7 million, and control costs of \$222. Note that the damage is an index and should not be interpreted in dollar terms and the level of cost is arbitrary due to normalizing marginal costs to \$1. When trading is allowed and sites have linear control costs to remove BOD load, the residual pollution is higher, but water quality is improved to 8, social damages are equivalent to the uniform removal case by construction, and costs are reduced to 167, a 25% cost saving over the uniform reduction scenario. The percentage saving is invariant to the cost normalization. The results are driven by removing pollution where the social damages are highest (which is positively correlated with water quality damages). Although more pollution is released, it is released in areas that have relatively less impact on social damages and in this case, on water quality as well. In the solution, there is less control in Pittsburgh and Steubenville, and more at the other sites than the uniform reduction scenario. The results using a quadratic cost function, reported in the last row of Table 4, are even more pronounced. Compared to the uniform reduction scenario, there is more pollution, better water quality, equal social damage, and a 61% saving in the cost of control. The nature of the solution is similar with Pittsburgh controlling slightly less than in the linear cost case and Steubenville controlling significantly more (results not shown in table). The illustrative results of Table 4 indicate that the large differences in social damages create a significant opportunity for trading. Differential control costs at the sites could further alter the cost-minimizing solution.

With regard to this regional case study, further research could improve our estimates of marginal damage functions from changes in ambient water quality, the length of impacted streams, the geographic extent of the trading zone, and potentially, a more specific representation of control costs. This is a multifaceted issue involving political, hydrologic, social, economic, and environmental constraints. In the Pittsburgh region, locks and dams, which are a common feature in the region, are capable of altering the fate and transport of pollutants like TSS and BOD and the water quality model used here could be expanded to account for these effects. In the case of CSOs, emissions are stochastic and depend upon individual system characteristics, and storm intensity and duration. While baseline water quality conditions may be relatively consistent during dry periods, they may degrade markedly during storm events due to increased surface runoff from nonpoint sources. Under these conditions, the marginal improvements expected by controlling CSOs could be hard

TABLE 4

Pollution, Water Quality, Damage, and Cost Comparisons for Illustrative Scenarios

Case	BOD Load Kg.	BOD Load Removed Kg.	Water Quality Impact Kg/M^2	Water Quality Change Kg/M^2	Economic Damage Index Million	Economic Damage Index Change Million	Cost of Control-Normalized Dollars	Cost Change Normalized Dollars
No control	261	0	62	0	11.3	0	0	– –
Uniform 85% removal	39	222	9	53	1.7	9.6	\$222 linear \$9,565 quadratic	\$222 \$9,565
Trading w/ linear cost (TC = R)	94	167	8	54	1.7	9.6	\$167	\$55 (25% saving to uniform)
Trading w/ increasing cost (TC = $.5R^2$)	88	173	8	54	1.7	9.6	\$3,701	\$5,864 (61% saving to uniform)

to recognize in the field although we have here used a design storm in low flow periods. Finally, the value of water quality improvements likely varies with human exposure patterns and damage estimates could potentially account for these differences among the different zones of influence.

V. CONCLUSION AND DIRECTIONS FOR FURTHER RESEARCH

This paper was motivated by a search for a more tractable water pollution trading design that could facilitate cost-effective compliance with water pollution control regulations. A social cost-minimization model was developed suggesting that a cost-effective solution could be achieved if allowable quantities from individual sources are traded using trading ratios based on relative damages between sources. The trading ratio is not a bureaucratic hindrance in this case, but a necessary element to achieve the socially cost-effective solution. The proposed trading framework is potentially suitable in situations where total maximum daily loads are being implemented or large investments to control combined sewer overflows are being considered. The CSO problem in the southwestern Pennsylvania was investigated in light of the theory. Results show that there appear to be substantive differences across sources in terms of environmental damages suggesting the existence of potential gains from trading activity even if control costs are similar. Generalized linear and quadratic cost functions were used to demonstrate that significant cost savings could exist while achieving the same level of social damages as a uniform reduction requirement that is the default condition in the law. In the illustrative estimates and this particular case, water quality is also improved compared to uniform removal.

Directions for further research beyond the case study are many. The structure of the economic trading model can be extended to include non-linearity and uncertainty that is likely to complicate the tractability of the trading ratio. Recent development of EPA's NWPCAM suggests that a consistent approach to estimate water quality effects from many different sources over a large geographic region could become available. While that model is still in development and was not used in this study, a fully developed national model could provide an acceptable basis for estimating the impact of individual point and nonpoint sources and CSOs within a geographic area. An environmental regulatory agency could use such a model to estimate damages per unit of emissions from each source. To facilitate the market, the EPA could then publish the entire matrix of potential damage (exchange) ratios with the intention of accepting trades at the pre-approved ratio. Firms could then "look up" trading ratios in this matrix to evaluate potential trading partners with the understanding that these exchange rates had been previously approved. This system could reasonably be expected to reduce transaction costs and uncertainties in the outcome of regulatory decisions regarding proposed trades. Many valuation issues also remain, beginning with improved algorithms to allocate households to stream segments and methods to better condition results on local water quality uses and characteristics. Although trade-offs in complexity and transparency remain, the apparent success of trading for air pollution control holds the promise of both environmental and economic improvements from carefully designed water pollution trading systems.

References

Alleghany County Sanitation Authority (ALCOSAN). 1999. "Regional Long Term Wet Weather Control Concept Plan." Draft. Pittsburgh: Alleghany County Sanitation Authority.

Atkinson, Scott E., and Tom H. Tietenberg. 1982. "The Empirical Properties of Two Classes of Designs for Transferable Discharge Permit Markets." *Journal of Environmental Economics and Management* 9 (2): 101–21.

———. 1991. "Market Failure in Incentive-Based Regulation: The Case of Emissions Trading." *Journal of Environmental Economics and Management* 21 (1): 17–31.

Bondelid, Timothy R., Suzanne J. Unger, and Andrew Stoddard. 2000. "National Water Pollution Control Assessment Model (NWPCAM) Version 1.1: Final Report." Prepared by Research Triangle Institute (RTI) for the U.S. Environmental Protection Agency, Office of Water and Office of Policy, Economics and Innovation, Washington D.C.

Carson, Richard T., and Robert C. Mitchell. 1993. "The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Water Quality." *Water Resources Research* 29 (7): 2445–54.

Chapra, Steven C. 1997. *Surface Water Quality Modeling*. New York: McGraw-Hill Co.

Desvousges, William H., V. Kerry Smith, and Anthony Fisher. 1987. "Option Price Estimates of Water Quality Improvements: A Contingent Valuation Study for the Monongahela River." *Journal of Environmental Economics and Management* 14 (3): 248–67.

Farber, Stephen, and Brian Griner. 2000. "Using Conjoint Analysis to Value Ecosystem Change." *Environmental Science and Technology* 34 (8):1407–12.

Farrow, Scott. 1999. "The Duality of Taxes and Tradable Permits: A Survey with Applications in Central and Eastern Europe." *Journal of Environment and Development Economics* 4 (Oct.): 519–35

Fullerton, Donald, and Gilbert Metcalf. 1998. "Environmental Taxes and the Double Dividend Hypothesis: Did You Really Expect Something for Nothing?" *Chicago-Kent Law Review* 73 (1): 221–56.

Hahn, Robert W. 1989. "A New Approach to the Design of Regulation in the Presence of Multiple Objectives." *Journal of Environmental Economics and Management* 17 (3): 195–211.

Heaney, James P., Leonard Wright, and D. Sample. 1999. "Research Needs in Urban Wet Weather Flows." *Water Environment Research* 71 (Mar.): 241–50.

Hoag, Dana L., and Jennie S. Hughes-Popp. 1997. "Theory and Practice in Pollution Credit Trading in Water Quality Management." *Review of Agricultural Economics* 19 (2): 252–62.

Klaassen, Ger. 1996. *Acid Rain and Environmental Degradation: The Economics of Emission Trading*. Cheltenham, U.K.: Edward Elgar.

Krupnick, Alan J., Wallace E. Oates, and Eric Van De Verg. 1983. "On Marketable Air-Pollution Permits: The Case for a System of Pollution Offsets." *Journal of Environmental Economics and Management* 10 (3): 233–47.

Lennon, Larry 2001. Personal communication. Coraopolis, Lennon, Smith, Souleret Engineering.

Magat, Wesley, Joel Huber, and W.Kip Viscusi. 2000. "An Iterative Choice Approach to Valuing Clean Lakes, Rivers, and Streams." Discussion paper 295. John M. Olin Center for Law, Economics, and Business, Harvard University.

Malik, Arun S., David Letson, and Stephen R. Crutchfield. 1993. "Point/Nonpoint Source Trading of Pollution Abatement: Choosing the Right Trading Ratio." *American Journal of Agricultural Economics* 75 (5): 959–67.

McGartland, Albert, and Wallace E. Oates. 1985. "Marketable Permits for the Prevention of Environmental Deterioration." *Journal of Environmental Economics and Management* 12 (4): 207–28.

Montgomery, David. 1972. "Markets in Licenses and Efficient Pollution Control Programs." *Journal of Economic Theory* 5 (3): 395–418.

Rodriguez, Fernando. 2000. "On the Use of Exchange Rates as Trading Rules in a Bilateral System of Transferable Discharge Permits." *Environmental and Resource Economics* 15 (4): 379–95.

Shortle, James S. 1990. "The Allocative Efficiency Implications of Water Pollution Abatement Cost Comparisons." *Water Resources Research* 26 (5): 793–97.

Shortle, James S., and David G. Abler. 1998. "Nonpoint Pollution." In *The International Yearbook of Environmental and Resource Economics 1997/1998,* ed. H. Folmer and T. Tietenberg. Cheltenham, U.K.: Edward Elgar.

Schultz, Martin T., Mitchell J. Small, R. Scott Farrow, and Paul S. Fischbeck, 2004. "State Water Pollution Control Policy Insights from a Reduced Form Model," *Journal of Water Resources Planning and Management* 130 (2): 150–59

Thomann, Robert V., and John A. Mueller. 1987. *Principles of Surface Water Quality Modeling and Control*. New York: Harper and Row Pub.

Tietenberg, Tom H. 1974. "Derived Decision Rules for Pollution Control in a General Equilibrium Space Economy." *Journal of Environmental Economics and Control* 1 (May): 3–16.

———. 1985. *Emission Trading: An Exercise in Reforming Pollution Policy.* Washington, D.C.: Resources for the Future.

Environmental Protection Agency (EPA). 1996. "*Draft Framework for Watershed Based Trading*." EPA 800-R-96-001. Washington D.C.: Office of Water.

———. 1999. "*Combined Sewer Overflows: Guidance for Monitoring and Modeling."* EPA/832/B-99/002. Washington D.C.: Office of Water.

———. 2001. "*Guidance: Coordinating CSO Long-term Planning with Water Quality Standards Reviews."* Washington, D.C.: Office of Water.

———. 2002. "*Water Quality Trading Policy; Proposed Policy*." Federal Register 67 (May): 34709–10.