E.ON Energy Research Center
FCN | Institute for Future Energy Consumer Needs and Behavior

FCN Working Paper No. 10/2009

Water Quality Trading when Nonpoint Pollution Loads are Stochastic

Gaurav Ghosh and James Shortle

November 2009

Institute for Future Energy Consumer Needs and Behavior (FCN)

Faculty of Business and Economics / E.ON ERC

RWTH AACHEN UNIVERSITY

FCN Working Paper No. 10/2009

Water Quality Trading when Nonpoint Pollution Loads are Stochastic

November 2009

Authors' addresses:

Gaurav Ghosh
Institute for Future Energy Consumer Needs and Behavior (FCN)
Faculty of Business and Economics / E.ON Energy Research Center
RWTH Aachen University
Mathieustrasse 6
52074 Aachen, Germany
E-mail: gghosh@eonerc.rwth-aachen.de

James Shortle
Director, Environmental and Natural Resources Institute
Department of Agricultural Economics and Rural Sociology
Pennsylvania State University
112C Armsby,
University Park, PA 16802, U.S.A.
E-mail: jshortle@psu.edu

Publisher: Prof. Dr. Reinhard Madlener
Chair of Energy Economics and Management
Director, Institute for Future Energy Consumer Needs and Behavior (FCN)
E.ON Energy Research Center (E.ON ERC)
RWTH Aachen University
Mathieustrasse 6, 52074 Aachen, Germany
Phone: +49 (0) 241-80 49820
Fax: +49 (0) 241-80 49829
Web: www.eonerc.rwth-aachen.de/fcn
E-mail: post_fcn@eonerc.rwth-aachen.de

Water Quality Trading when Nonpoint Pollution Loads are Stochastic

Gaurav S. Ghosh* James S. Shortle†

November 2009

Abstract

We compare two tradable permit markets in their ability to meet a stated environmental target at least cost when some polluters have stochastic and non-measurable emissions. The environmental target is of the safety-first type, which requires probabilistic emissions control. One market is built around the trading ratio, which defines the substitution rate between stochastic and deterministic pollution, and is modeled on existing markets for water quality trading. The other market is built around a new definition of the pollution credit as a multi-attribute good, where the attributes supply information to the market on the environmental risks associated with stochastic pollution loads. The market with multi-attribute credits is found to out-perform the trading ratio market in its ability to satisfy the safety-first environmental target at least cost. This result comes about because polluters are able to directly price risk in this market. In the trading ratio market risk is not a factor in polluters' trading decisions and is only controlled, through the trading ratio, under highly restrictive conditions.

*Institute for Future Energy Consumer Needs and Behavior (FCN), Faculty of Business and Economics, E.ON Energy Research Center, RWTH Aachen University, Mathieustrasse 6, 52074 Aachen, Germany

†Department of Agricultural Economics and Rural Sociology, Pennsylvania State University, 112C Armsby, University Park, PA 16802, USA

Keywords: Water Quality Trading, Stochastic Pollution, Market

1 Introduction

Beginning with the seminal works of Crocker (1966), Dales (1968) and Montgomery (1972), a large literature has developed on the use of emissions trading to achieve environmental targets. The economic case for emissions trading is that it can achieve environmental targets at lower social cost than emissions taxes and traditional design and performance standards (Tietenberg, 1990). Real world success stories for air emissions trading, such as the US EPA's Acid Rain Control Program (Lobsenz, 1993), have spurred interest in expanding the scope of emissions trading to water-based pollutants (Shortle and Horan, 2008).

Water quality trading was the focus of Dales (1968), which recommended the market as a tool for environmental management. However, it has not until recently been a focus of applications of the tool. Interest is now high. Dozens of water quality trading initiatives are underway for nutrients, sediments and other pollutants in the US and other countries (Shortle and Horan, 2008). The economic appeal of water quality trading is, as is the appeal of emissions trading generally, that it offers a means for achieving a cost-effective allocation of allowable emissions among alternative pollution sources without environmental regulators knowing the abatement costs of individual agents. This promise, if markets can in fact be designed to realize it, is compelling for water pollution control in the US (Shortle and Horan, 2008). Water pollution there has largely been regulated through non-tradable, technology-based effluent permits applied to point sources of water pollution, while nonpoint sources have been largely unregulated (Ribaudo, Heimlich, and Peters, 2005). The economic consequence is that the pollution control that is achieved, because it does not consider the relative costs of alternative point and nonpoint sources, is overly costly (Davies and Mazurek, 1998; Ribaudo, Heimlich, and Peters, 2005). Studies show that water quality trading programs could result in annual savings of US$ 1 billion when compared to the traditional command-and-control type technology standards (US Environmental Protection Agency, 2001).

Yet, while the promise of water pollution trading may be compelling, significant challenges confront the design of markets that can realize it. Some key assumptions underpinning the standard emissions trading model are that the polluter has deterministic control over loadings, and that they are cheaply and accurately measurable (Malik, Letson, and Crutchfield, 1993). In the context of water quality trading, these assumptions are valid for point sources, but not for nonpoint sources (Shortle, 1987). Nonpoint water pollution is largely the result of runoff from agricultural land and other uses. Farmers and other types of land owners cannot control the quality and quantity of runoff with precision; their abatement activities are best characterized as influencing the probabilistic distribution of emissions from their land (Segerson, 1988; Horan and Shortle, 2001). Further, the dispersed nature of nonpoint pollution makes cheap and accurate metering of emissions impossible. Observational uncertainty compounds the inherent stochasticity of the process. These features of nonpoint pollution raise two significant and related issues for market design. The first is what to define as the nonpoint commodity. The second is how to design trading rules to manage nonpoint trading risk.

The optimal basis for nonpoint instruments, whether for trading or other incentive or regulatory mechanisms, has been much discussed in the economic literature on nonpoint pollution. The choices that have drawn attention are observable inputs or practices that affect nonpoint pollution (e.g. fertilizer applications to farm land, farming practices that affect sediment or nutrient runoff), emissions estimates constructed from observations on input and practices (e.g. estimates of pollution delivery from farm fields using observations of inputs and practices as data), and realized ambient pollution loads. Tax / subsidy schemes based on realized ambient loads have been of particular interest to economists as a means for nonpoint pollution control (e.g., Segerson, 1988; Cabe and Herriges, 1992; Horan, Shortle, and Abler, 2002), but do not offer a plausible basis for nonpoint pollution trading because trading involves an exchange of individual responsibilities, while ambient loads measure group performance. Thus, policy interest has largely focused on trading inputs and practices, or estimated pollution loads.

The approach in US water quality trading programs is to define the tradable nonpoint commodity in terms of estimated reductions in emissions or delivered loads (Shortle and Horan, 2008). An important aspect of this approach should be noted in comparison to conventional trading of actual, metered emissions. This is the enormous uncertainty about the actual water quality outcomes of individual trades based on modeled emissions. This uncertainty exists because the prediction errors for water quality models are known to be quite large (National Research Council, 2001; Reckhow, 1994, 2003). This uncertainty is addressed in US program design by the application of an "uncertainty" trading ratio to nonpoint trades. The ratio is typically defined in terms of the reduction in nonpoint emissions required to offset a unit of point source emissions. Designing uncertainty trading ratios to address nonpoint risk has been the focus of the limited economic research on point-nonpoint trading (e.g., Hennessy and Feng, 2008; Rabotyagov, Feng, and Kling, 2006; Malik, Letson, and Crutchfield, 1993; Shortle, 1990). Three essential insights have resulted (Shortle and Horan, 2008). One is that trading ratios are optimally differentiated across sources to address differences in relative risk. A second insight is that optimal trading ratios defining the exchange rate between nonpoint discharges and point source emissions may be less than one, reflecting the fact that nonpoint sources are often a greater source of risk than point sources. A third insight is that optimal trading ratios depend on the choices of other market design parameters, including the allocation of initial allowances and the cap placed on the aggregate supply of permits, and thus optimality requires selecting these parameters simultaneously. These findings are contrary to standard practice in which point-nonpoint trading ratios are almost always uniform across nonpoint sources or categories of nonpoint sources, often (substantially) more than one, and specified independently of other parameters (Horan, 2001; Shortle and Horan, 2008).

The economic literature thus points out flaws in the design of actual trading programs that will affect their capacity to minimize costs and optimally manage water quality risk. This paper will argue that those flaws aside, the use of modeled mean emissions as an approximation of nonpoint source pollution and the use of trading ratios to address nonpoint risk is suboptimal, and will offer an alternative market design that offers greater promise for efficiency in water quality management through trading.

As noted above, water quality planners are concerned with risk. Emissions markets are conventionally designed to meet an exogenously specified cap on a linear aggregation of polluting emissions. When emissions are stochastic, an alternative to the conventional design is to place a limit on the probability that the linear aggregation exceeds the exogenously specified cap. Probabilistic environmental targets of this type are labeled "safety-first" and are consistent with the Total Maximum Daily Load (TMDL) approach to water quality management mandated by the Clean Water Act (Beavis and Walker, 1983; Lichtenberg and Zilberman, 1988; Horan, Shortle, and Abler, 2002; Qiu, Prato, and McCamley, 2001; National Research Council, 2001).

In this paper we develop the concept of a safety-first environmental target in the context of water quality management and derive necessary conditions for the attainment of this target at least cost. We then propose a market design and show that at equilibrium, and under trivial conditions, the market outcome meets the safety-first target at least cost. The primary innovation in market rules, when compared to conventional environmental markets for nonstochastic pollutants in which polluters trade emissions allowances, or US water quality markets in which nonpoint sources trade estimated emissions-based credits, is in the definition of the nonpoint source credits. The nonpoint source credits traded in this market have multiple attributes: the mean and variance of the underlying abatement and its covariances with other abatements. We also study the performance of existing trading ratio-based water quality trading markets vis-à-vis their attainment of the safety-first environmental target and show that outcomes in this market are typically sub-optimal.

In the next section we formally derive the safety-first environmental target and then obtain conditions on polluter behavior required for its least cost attainment. The intent is to generate a clear picture of the optimal scenario. In the section after that, the market with the multi-attribute nonpoint abatement credits, referred to hereafter as the MANA market, is described. Market rules are defined and necessary conditions for equilibrium discussed. We examine the restrictions required for ensuring that the MANA market equilibrium is safety-first compliant at least cost. The next section focuses on the trading-ratio-based market, which is conceptually identical to water quality trading markets in current use. This market is referred to as the TR market in the ensuing discussion. The conditions required for the TR market equilibrium to be safety-first compliant at least cost are derived. In the final section, the MANA market and the TR market are compared, and the results put in a broader policy context.

2 The *Safety-First* Environmental Target

The problem faced by the Environmental Agency (EA) is to design a market that will achieve the safety-first environmental target at least cost. Following Shortle and Horan (2008), the EA has three integrated tasks when pursuing this objective. The first is to *define the commodities that will be traded in the market.* A key element of the definition is the specification of the observable indicators of environmental performance (e.g. pounds of total nitrogen annually discharged from a farm or sewage treatment facility in the Susquehanna River Basin) to which rights pertain. The indicators must be observable so that trading is enforceable, and under the control of the polluter if the polluter is to be held responsible for non-compliance. The second task is to *define specific rules for trading the commodities between alternate sources*. The defined commodities may not be homogenous across sources. For example, discharges of nitrogen at one point in a watershed may have different water quality impacts than discharges at another. The trading rules are designed such that after their application, all commodities become homogeneous. In current markets trading rules are generally implemented in the form of trading ratios that define the rates at which different commodities can be exchanged between different sources. From the perspective of water quality agencies, trading ratios are intended to assure that water quality outcomes resulting from commodity trades are at least equal to those that would occur without the trade (USEPA, 2007)[1]. The third task is to *cap the aggregate supply of the commodities such that feasible market allocations of polluting emissions, given trading rules, do not violate the environmental goal(s).*

We design an emissions trading market such that within its framework the EA is able to complete its three integrated tasks satisfactorily. The first task we consider is that of defining the environmental target and capping aggregate credit supply such that the target is not violated. Following the US EPA's watershed based approach to controlling water quality (see US Environmental Protection Agency, 2003), we situate the problem at the same geographic scale. Consider a representative watershed. Completing the task successfully requires that the EA design a policy that targets loadings such that all water bodies in the watershed meet their designated use. This requires that ambient aggregate loadings in the watershed be controlled. Let L be aggregate loadings and let $\bar{L}$ be the watershed-level loadings target. Following the safety-first literature (e.g., Lichtenberg and Zilberman, 1988; Qiu, Prato, and McCamley, 2001) the environmental target is expressed probabilistically. The EA is required to ensure that the probability that aggregate loadings exceed the target is less than α, i.e. $\Pr(L \geq \bar{L}) \leq \alpha$ where α is the sanctioned probability with which the environmental target may be violated. It may be interpreted as the socially acceptable probability of failure in

[1]need citation

meeting water quality policy goals. If $\alpha = 0.05$ then the safety-first environmental target implies that $\bar{L}$ is exceeded in 5% of all possible scenarios. We assume that α is a fixed coefficient set externally and not controlled by the EA.

Let the market consist of I nonpoint sources and J point sources, indexed by i and j respectively. Let r_i be i's runoff or loadings and let e_j be j's emission loadings into the watershed in a given trading period. Let $\mu_i \in \mathbb{R}_+$ be the mean value of nonpoint source i's runoff r_i and let $\boldsymbol{\sigma}_i = \{\sigma_{i1}, \cdots, \sigma_{iI}\} \in \mathbb{R}^I$ be the vector of covariances between i's pollution and the loadings of all other nonpoint sources, where $\{\sigma_{il} \in \mathbb{R} \mid \sigma_{il} \in \boldsymbol{\sigma}_i, i \neq l\}$ is the covariance between the loadings of nonpoint sources i and l, and $\sigma_{ii} \in \mathbb{R}_{++}$ is the variance of i's loadings. $\boldsymbol{\sigma}_i$ characterizes the risk that i's pollution poses vis-à-vis the attainment of the safety-first target. Taken together μ_i and $\boldsymbol{\sigma}_i$ provide a second order approximation of the probability distribution of i's emissions. These moments are approximated through statistics calculated from appropriate soil transport and hydrological models. Similar information is already used to define credits in extant water quality trading programs, such the Pennsylvanian nutrient trading program (e.g., Pennsylvania Department of Environmental Protection, 2007).

Summing over all i and j, true aggregate emissions loadings are $L = \sum_i r_i + \sum_j e_j$ and hence the EA is required to ensure that $\Pr(\sum_i r_i + \sum_j e_j \geq \bar{L}) \leq \alpha$. If the joint distribution of emissions is known, and if the distribution is such that a "deterministic equivalent" exists, then the probability statement above can be expressed exactly in terms of means, variances, covariances and other moments of the distribution of the emissions (Beavis and Walker, 1983; Wets, 1983; Kampas and White, 2003). But given the nonmeasurability of nonpoint emissions and the prediction errors in current water quality models, knowledge of the emissions distribution is not known with certainty and the conditions for a deterministic equivalent are not satisfied. A common nonparametric approximation of probability statements when the distribution of the random variable is unknown (or a deterministic equivalent does not exist) is based on Chebychev's inequality (Wets, 1983). By Chebychev's inequality the statement

$\Pr\left(x \geq E(x) + \sqrt{V(x)/\alpha}\right) \leq \alpha$ is always true, where x is a random variable, $E(x)$ and $V(x)$ are its mean and variance, and $\alpha \in [0, 1]$ is a probability.

Let $x = \sum_i r_i + \sum_j e_j$, which implies that $E(x) = \sum_i \mu_i + \sum_j e_j$ and $V(x) = \sum_i \sum_l \sigma_{il}$. By Chebychev's Inequality, $\Pr\left(\sum_i r_i + \sum_j e_j \geq \sum_i \mu_i + \sum_j e_j + \sqrt{\sum_i \sum_l \sigma_{il}/\alpha}\right) \leq \alpha$. Comparing the inequality to the probabilistic statement of the safety-first environmental target, $\Pr(\sum_i r_i + \sum_j e_j \geq \bar{L}) \leq \alpha$, it is apparent that the safety-first target is satisfied when

$$\sum_i \mu_i + \sum_j e_j + \sqrt{\frac{\sum_i \sum_l \sigma_{il}}{\alpha}} \leq \bar{L} \tag{1}$$

$\sum_i \mu_i$ is the mean loadings of all nonpoint sources and $\sum_j e_j$ is the measured loadings of all point sources in the watershed. $\sum_i \sum_l \sigma_{il}$ is the aggregate variance of all loadings. For convenience we sometimes use the substitution $V \equiv \sum_i \sum_l \sigma_{il}$ in the subsequent discussion. The safety-first environmental target in (1) is similar to that in Beavis and Walker (1983), except that we do not assume that pollution loads are independent. This assumption is unsuitable in the context of nonpoint source pollution. Contiguous polluters in a watershed have similar weather and topography, which imply that their loadings will be correlated. For example, a sudden storm in a spike in runoff from all neighboring farms.

From (1), satisfying the safety-first target requires that J point sources control their measured emissions and that I nonpoint sources control the mean and variance of their runoff, and the covariance of their runoff with the runoff of other nonpoint sources. Reductions in μ_i, σ_{il} and e_j weaken the constraint and are thus desirable actions from the EA's point of view. Indeed, μ_i, σ_{il} and e_j may be interpreted as substitutes in the satisfaction of the safety-first environmental target. When (1) binds then μ_i and e_j are perfect substitutes. The marginal rate of substitution between these variables and σ_{il} is $2\sqrt{\alpha V}$ when $l = i$ and $\sqrt{\alpha V}$ otherwise. $\sqrt{V}$ is the standard deviation of all loadings in the watershed. $\sqrt{\alpha V}$ may be interpreted as the socially acceptable standard deviation of aggregate loadings in the watershed since V

is the variance of aggregate loadings and α defines society's risk appetite. Alternatively, it may be thought of as the socially acceptable level of aggregate environmental risk. Since $\sqrt{\alpha V}$ units of σ_{il} (or $2\sqrt{\alpha V}$ units of σ_{ii}) are equivalent to one unit of μ_i or e_j at the margin, one may infer that control of the latter is suitable for rapid movement towards the safety-first outcome. Once in the vicinity of the safety-first outcome however, control of $\boldsymbol{\sigma}_i$ allows minute calibration towards the precise satisfaction of (1).

2.1 The Least Cost Safety-First Solution

We define pollution control as a reduction in its expected value, a reduction in its variance and a reduction in its impact on V through covariance effects. By the latter we imply that a change in nonpoint source i's pollution should result in a fall in the variance of aggregate pollution. We assume that pollution control is costly. Hence decreasing μ_i, σ_{ii} or $|\sigma_{il}|$ reduces i's profits and decreasing e_j reduces j's profits. We also assume that the marginal cost of pollution reduction is higher at lower absolute levels of the pollution variable. In other words, if $\pi_i(\mu_i, \boldsymbol{\sigma}_i)$ and $\pi_j(e_j)$ are the restricted profit functions[2] for i and j respectively, then they are increasing and concave in all μ_i, σ_{ii} and e_j, with (suppressing indices and arguments) $\pi' \geq 0$ and $\pi'' < 0$. $\pi_i(\mu_i, \boldsymbol{\sigma}_i)$ is increasing in the absolute value of σ_{il} and always concave, i.e. $|\partial \pi_i / \partial \sigma_{il}| \geq 0$ and $\partial^2 \pi_i / \partial \sigma_{il}^2 < 0$.

The aggregate restricted profit function for all polluters in the watershed is $\sum_i \pi_i(\mu_i, \boldsymbol{\sigma}_i) + \sum_j \pi_j(e_j)$. The optimal market outcome from the EA's point of view requires that the safety-first environmental target (1) be met at least cost. To achieve this outcome polluters must choose e_j, μ_i and $\boldsymbol{\sigma}_i$ such that aggregate profits are maximized subject to (1) being satisfied.

[2]Restricted profit functions define the maximum profit associated with any emissions level, given that the input mix and output levels are chosen optimally. These functions are derived from a standard profit function, where profits are a function of inputs, production technology and prices. We do not supply a formal derivation, but refer interested readers to Graff Zivin and Small (2003) for a deeper discussion. Restricted profit functions are useful because they allow modeling of profit as a function of the policy variables of interest – in our model the policy variables pertain to pollution load.

The necessary conditions for the EA's optimum outcome, obtained through manipulation of the Kuhn-Tucker conditions that characterize the aggregate profit-maximizing outcome, are given in (2).

$$\frac{\partial \pi_j}{\partial e_j} = \frac{\partial \pi_i}{\partial \mu_i} = \sqrt{\alpha V}\left(\frac{\partial \pi_i}{\partial \sigma_{il}} + \frac{\partial \pi_l}{\partial \sigma_{il}}\right) = \lambda \quad \forall\, i, l \in I,\ j \in J \tag{2}$$

In (2), λ is the Lagrange multiplier and shadow price associated with the safety-first environmental constraint (1). Given the concavity of π_i and π_j in all arguments, if the optimum vector $\{\mu_i^*, \boldsymbol{\sigma}_i^*, e_j^*\}_{i,j}$ forms a convex set then (2) is also sufficient for the attainment of the EA's optimum market outcome when the outcome is an interior solution.

The first equality in (2) indicates that μ_i and e_j are treated as perfect substitutes at the least cost safety-first solution since the marginal rate of substitution between them is $MRS_{\mu_i,e_j} = 1$. The second equality establishes the optimal marginal rate of technical substitution between μ_i and σ_{il}, which is $MRTS_{\mu_i,\sigma_{il}} = \sqrt{\alpha V}\left(1 + MRS^{\sigma_{il}}_{\pi_l,\pi_i}\right) \geq 0$. $MRTS_{\mu_i,\sigma_{il}} \geq 0$ because $\lambda \geq 0$ by the Kuhn-Tucker conditions. $MRS^{\sigma_{il}}_{\pi_l,\pi_i} = \frac{\partial \pi_l}{\partial \sigma_{il}} / \frac{\partial \pi_i}{\partial \sigma_{il}}$ indicates the trade-off between i's profits and l's profits at the optimal level of pollution correlation σ_{il}^*. Hence, the optimal $MRTS_{\mu_i,\sigma_{il}}$ is set to the socially-acceptable level of aggregate risk $\sqrt{\alpha V}$ after adjusting for the marginal impact of pollution correlation on nonpoint source profits. When $MRTS_{\mu_i,\sigma_{il}} \geq 1$ then control of μ_i is more important since more than one unit of σ_{il} trades for one unit of μ_i. This is because the distortion away from the optimum caused by a marginal change in σ_{il} is smaller than that caused by a marginal change in μ_i. A sufficient condition for $MRTS_{\mu_i,\sigma_{il}} \geq 1$ is that $\partial \pi_i / \partial \sigma_{il}$ and $\partial \pi_l / \partial \sigma_{il}$ share the same sign. Analogously, when $MRTS_{\mu_i,\sigma_{il}} \leq 1$ then controlling σ_{il} is more important. This occurs only when $\partial \pi_i / \partial \sigma_{il}$ and $\partial \pi_l / \partial \sigma_{il}$ have opposing signs *and* $\sqrt{\alpha V}$ is small. Typically, $MRTS_{\mu_i,\sigma_{il}} \geq 1$ and control of μ_i will be paramount.

We have three more observations to make regarding the least cost safety-first outcome. First, a corollary of the $MRTS_{\mu_i,\sigma_{il}}$ result is that the EA's optimum can only be achieved when nonpoint sources coordinate their pollution control strategies. Second, in the special case where $l = i$, $MRS^{\sigma_{ii}}_{\pi_i,\pi_i} = 1$, which implies that $MRTS_{\mu_i,\sigma_{ii}} = 2\sqrt{\alpha V}$. Finally, note that μ_i and σ_{ii} are the only two independent control variables available to i. To summarize, the safety-first environmental target (1) is achieved at least cost only when marginal profits with respect to μ_i and e_j and the risk-adjusted marginal profits with respect to σ_{il} are equated across all polluters. This is a standard marginality result, but achieving the outcome might be problematic given the need for coordination across polluters. In the next section we explore a market design that makes the need for explicit coordination redundant.

3 The MANA Market

We now tackle the other two tasks that must be completed successfully if the EA is to engineer its optimum. These tasks pertain to the definition of the traded commodity or credit and to the definition of trading rules. Since the EA's optimum requires explicit control of measured point source emissions and the modeled mean, variance and covariances of nonpoint source emissions, we define the traded commodities in terms of these variables. Let an allowance be a right to pollute. It defines the maximum emissions that an individual polluter may discharge into the watershed. It is allocated by the EA and known to the polluter at the start of the trading period. Let $\hat{r}_i$ and $\hat{e}_j$ be the allowances given to i and j. We assume, further, that the vector of allowances $\{\hat{e}_1, \cdots, \hat{e}_j, \cdots, \hat{e}_J, \hat{r}_1, \cdots, \hat{r}_i, \cdots, \hat{r}_I\}$ account for distance effects. Distance-related trading ratios or "delivery" ratios, which control for spatial impacts on ambient loadings, are not explicitly modeled. Implicit in our treatment of spatial effects is that they are known and non-stochastic. The allowances are allocated such that they sum up to the watershed level pollution target, i.e. $\sum_j \hat{e}_j + \sum_i \hat{r}_i = \bar{L}$. The allowances are parameters outside the control of the polluters.

3.1 Credits, Market Rules and Behavior

The difference between nonpoint source i's allowance and runoff is $a_i = \hat{r}_i - r_i$, which is unknown because of the stochasticity and non-measurability of r_i. The expected value of a_i is $E_i = \hat{r}_i - \mu_i$, the variance is $c_{ii} = \sigma_{ii}$ and the covariance between a_i and a_l for any $i, l \in I$ is $c_{il} = \sigma_{il}$. Let the vector of all variances and covariances associated with a_i be $\boldsymbol{c}_i(\boldsymbol{\sigma}_i) = \{c_{i1}(\sigma_{i1}), \cdots, c_{iI}(\sigma_{iI})\}$. Enforceability dictates that the definition of nonpoint source credits must depend on known and observable variables. Hence it cannot be based on the unknown a_i. Instead, we define i's credits as the multi-attribute good $\tilde{a}_i$, where the attributes are the mean, variance and covariances of a_i. This definition, $\tilde{a}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$, contains all the information needed to make a second order approximation of the probability distribution of a_i. Through this definition, information on the risk attributes of i's runoff is transmitted to the market. This is an important design consideration since controlling risk attributes is critical to the least cost attainment of the safety-first environmental target (1). The difference between point source j's allowance and emissions is $a_j = \hat{e}_j - e_j$, which is known because e_j is accurately measurable. Hence j's credit supply is simply defined as a_j. The credit supply of all polluters is negative when emissions exceed allowance and is positive otherwise. A polluter is in compliance when its credit supply, after adjustment for trade, is weakly positive. We develop this point in Section 3.3.

Let r_i^0 and e_j^0 be nonpoint source i's and point source j's profit-maximizing emissions prior to market entry. We assume that $r_i^0 \leq \hat{r}_i$ and $e_j^0 > \hat{e}_j$. Nonpoint source i faces a lenient allowance that does not prevent it from polluting up to its profit maximizing level under autarky. Point source j's allowance binds, which prevents it from polluting up to its profit maximizing level under autarky. The implication is that nonpoint sources have no incentive to buy credits from other polluters, but point sources do, if buying credits is less costly than restricting production.

These modeled relationships between allowances and profit-maximizing emissions reflect allocation rules in extant water quality trading markets. Although major contributors to water quality impairment, before the advent of markets nonpoint sources were outside the purview of the pollution control policy regime (US Environmental Protection Agency, 2002). Since their inception, water quality trading markets were treated as viable instruments to control nonpoint pollution without direct regulation. The aim was to use water quality trading as a carrot approach to nonpoint pollution control, whereby point sources give financial incentives to nonpoint sources to voluntarily reduce loadings.

The typical water quality trading scheme works as follows: allowances are allocated such that point sources are legally required to reduce their loadings. Given the cost of doing so, point sources are incentivized to contract nonpoint sources to reduce loadings. The nonpoint sources are incentivized to accept the contract when earnings from credit sales exceed the cost of reducing loadings. They have no incentive to buy credits because they are not legally required to control loadings. At all times point sources are allowed to trade abatements with each other. We allow all these transactions in the MANA market.

Since nonpoint source pollution is unobservable, in practice, point source j predicates its contract with nonpoint source i upon one or more observable pollution reducing technologies or inputs. For example, the contract may be based on i implementing a riparian buffer and a nutrient management plan. Each argument in i's credit supply $\tilde{a}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$ is then determined by the modeled reduction of runoff as a consequence of the buffer and the nutrient plan. Assuming the existence of a large set of abatement technologies and the feasibility of marginally adjusting the abatement-generating capability of each technology by varying the input mix the $\{E_i, \boldsymbol{c}_i\}$ space may be considered continuous.

Consider some i that implements a set of abatement technologies and generates $\tilde{a}_i(E_i, \boldsymbol{c}_i)$ credits. A point source j may buy a $\gamma_{ij} \in [0, 1]$ share of $\tilde{a}_i(E_i, \boldsymbol{c}_i)$ where $\gamma_{ij} = 0$ implies that j does not trade with i and $\gamma_{ij} = 1$ implies that j has bought i's entire credit supply. Buying a proportion of i's credit supply is equivalent to investing a γ_{ij} share in i's set of abatement technologies. Point source j can minimize the cost of complying with its allowance constraint by optimally adjusting its portfolio of shares in the abatement technologies implemented by other polluters. We assume that the market is competitive, which implies that each $i \in I$ sells its entire credit supply, i.e. $\sum_j \gamma_{ij} = 1 \; \forall \; i$. Point sources buy and sell shares in their own abatement technologies to other point sources. Let $\gamma_{jk} \in [0,1]$ be the proportion of j's generated credits sold to $k \in J$. $\sum_{k \neq j} \gamma_{jk}$ is the total proportion of credits sold by j to other point sources and γ_{jj} is the proportion of j's credits that remain unsold. The mutual exclusivity of these events dictates that $\sum_k \gamma_{jk} = \sum_{k \neq j} \gamma_{jk} + \gamma_{jj} \equiv 1$.

Given the continuity of the $\{E_i, \boldsymbol{c}_i\}$ space and of the e_j space, it is not necessary to model credit sales as investments in credit-producing projects. Alternatively, each transaction may be modeled as predicated upon a separate project. In other words, instead of assuming that j and k have bought 50% shares in i's project supplying 10 credits, we could have assumed that they invested in separate projects initiated by i, each supplying five credits. We made a conscious decision to conceptualize credit sales as investments to simulate the process through which credits are allocated in the real world, where credit purchases are analogous to offsets: the EA associates a number of credits with a project, which allows the point source to increase its emissions by the credit amount. Fractional investment in projects is allowed to facilitate pooling of risks.

3.2 The Nonpoint Source Problem

Prior to market entry i produced emissions at its unconstrained profit maximizing level. Let its unconstrained profits be π_i^0. After entering the market i produces $\tilde{a}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$ credits and its restricted profit function is $\pi_i(\mu_i, \boldsymbol{\sigma}_i)$. The cost of credit generation is the difference between unconstrained and restricted profits, i.e. $\pi_i^0 - \pi_i(\mu_i, \boldsymbol{\sigma}_i)$. Upon selling

$\tilde{a}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$ i earns $q_i(E_i, \boldsymbol{c}_i)$ in revenue. Hence its profits from production and trade of credits is $\Pi_i = q_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i)) - [\pi_i^0 - \pi_i(\mu_i, \boldsymbol{\sigma}_i)]$. Assuming that i is a profit maximizer, it faces an unconstrained optimization problem since its allowance is never binding with $r_i^0 \leq \hat{r}_i$. Through the implementation of credit generation technologies it controls μ_i and $\boldsymbol{\sigma}_i$. The necessary conditions for i to maximize profits are

$$\frac{\partial \pi_i}{\partial \mu_i} = -\frac{\partial q_i}{\partial \mu_i} \tag{3}$$

$$\frac{\partial \pi_i}{\partial \sigma_{il}} = -\frac{\partial q_i}{\partial \sigma_{il}} \quad \forall\, l \in I \tag{4}$$

As (3) and (4) indicate, i maximizes profits by equating marginal production profits with marginal revenue from credit sales. Since $\partial \pi_i / \partial \mu_i \geq 0$ and $\partial \pi_i / \partial \sigma_{ii} \geq 0$, it follows that sales revenue q_i is decreasing in μ_i and σ_{ii} at the profit maximum. And since $E_i = \hat{r}_i - \mu_i$ and $c_{ii} = \sigma_{ii}$, q_i is increasing in E_i and decreasing in c_{ii}. This makes intuitive sense because one expects point source j to prefer credits with high mean and low variance. Such credits reduce the likelihood of j exceeding its allowance. The optimal relationship between q_i and c_{il} when $l \neq i$ depends on the sign of $\partial \pi_i / \partial \sigma_{il}$. Revenue q_i decreases in c_{il} when the sign is positive and vice versa.

3.3 The Point Source Problem

As defined in Section 3, j faces an emissions allowance $\hat{e}_j$. Under autarky it meets its allowance by choosing emissions level e_j such that $e_j \leq \hat{e}_j$. Or, in terms of generated credits, $a_j = \hat{e}_j - e_j$, it meets its allowance when $a_j \geq 0$. Since the allowance $\hat{e}_j$ constrains profit-maximizing behavior j will augment its allowance by buying credits from other polluters in the market. Or, in other words, by creating a portfolio of shares in the credit-generating projects of other polluters.

Point source j's purchase of shares from other point sources is $\sum_{k \neq j} \gamma_{kj} a_k$ where k indexes the other point sources. Point source j's purchase of shares from nonpoint sources is $\sum_i \gamma_{ij} \tilde{a}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$. The credits sold by j to other point sources is $\sum_{k \neq j} \gamma_{jk} a_j$ and the unsold credits are $\gamma_{jj} a_j$. Summing over these trades and using the tautology that $\sum_k \gamma_{jk} = \sum_{k \neq j} \gamma_{jk} + \gamma_{jj} \equiv 1$, j's total credit stock is $\sum_k \gamma_{kj} a_k + \sum_i \gamma_{ij} \tilde{a}_i$. The true difference between j's allowance (augmented by credit purchases) and emission loadings is $\sum_k \gamma_{kj} a_k + \sum_i \gamma_{ij} a_i$. The mean difference is $\sum_k \gamma_{kj} a_k + \sum_i \gamma_{ij} \mu_i$ and the variance of the difference is $\sum_i \sum_l \gamma_{ij} \gamma_{lj} \sigma_{il}$.

Even though j's emissions are deterministic we define market rules with respect to its trades probabilistically. This is because meeting the safety-first environmental target at least cost [see (2)] requires probabilistic control of stochastic emissions and defining j's responsibilities probabilistically is a convenient way to do so. This specification – of probabilistic regulation of individual polluters – is inconsistent with current US water quality trading regulations, which are deterministic. However, it makes intuitive sense when developing a policy that allows offsets from stochastic sources and watershed level environmental targets are defined probabilistically (as with the TMDL). Also, by offsetting its deterministic emissions through purchase of stochastic nonpoint emissions, j's pollution liabilities become probabilistic, indicating the need for probabilistic control.

Let the socially sanctioned probability with which j may exceed its allowance be β. It is not necessary that $\beta = \alpha$, the socially sanctioned probability with which the aggregate environmental target may be violated. The EA (and by proxy society) may have different preferences over aggregate and individual emissions. Intuitively this makes sense: there is no reason for the EA to care as much about the likelihood of an individual polluter violating its cap as it does about the aggregate cap being violated. An individual cap violation is not necessarily in conflict with attainment of the safety-first environmental target. Under-abatement by some polluters may be accompanied by equal over-abatement by other polluters. Fundamentally, the choice of β must be consistent with the EA's goal of meeting the safety-first environmental target at least cost. As is shown subsequently, setting $\beta = \alpha$ is generally not consistent with this goal.

When offsets from only deterministic sources are allowed j does not exceed its allowance when *Emissions - Offsets* $\leq$ *Allowance*, or in terms of credit stock, $\sum_k \gamma_{kj} a_k \geq 0$. When offsets from deterministic and stochastic sources are allowed, safety-first control of j's emissions and credit purchases requires that Pr*(Emissions - Offsets* $\leq$ *Allowance)* $\geq 1 - \beta$, or in terms of credit stock, $\Pr\left(\sum_k \gamma_{kj} a_k + \sum_i \gamma_{ij} a_i \geq 0\right) \geq 1 - \beta$. Applying Chebychev's inequality, j's allowance constraint expressed as

$$\sum_k \gamma_{kj} a_k + \sum_i \gamma_{ij} E_i - \sqrt{\frac{\sum_i \sum_l \gamma_{ij} \gamma_{lj} c_{il}}{\beta}} \geq 0 \tag{5}$$

where $\sum_k \gamma_{kj} a_k$ is the net abatements bought by j from other point sources, $\sum_i \gamma_{ij} E_i$ is the mean abatement bought from nonpoint sources and $\sum_i \sum_l \gamma_{ij} \gamma_{lj} c_{il}$ is the total variance of abatement purchased. Since $a_j = \hat{e}_j - e_j$, $E_i = \hat{r}_i - \mu_i$ and $c_{il} = \sigma_{il}$, (5) may be rewritten in terms of emissions as $\sum_j \gamma_{kj} e_k + \sum_i \gamma_{ij} \mu_i + \sqrt{\sum_i \sum_l \gamma_{ij} \gamma_{lj} \sigma_{il} / \beta} \leq \sum_k \gamma_{kj} \hat{e}_k + \sum_i \gamma_{ij} \hat{r}_i$, where $\sum_k \gamma_{kj} \hat{e}_k + \sum_i \gamma_{ij} \hat{r}_i$ is the proportion of the watershed level loadings target $\bar{L}$ that j is responsible for. For convenience, we often use the substitution $V_j \equiv \sum_i \sum_l \gamma_{ij} \gamma_{lj} \sigma_{il} = \sum_i \sum_l \gamma_{ij} \gamma_{lj} c_{il}$ in the ensuing discussion. Comparing j's allowance constraint to the EA's safety-first environmental constraint (1) reveals that the constraints are identical except for the γ weights in (5). These weights, which measure j's preferences for the risks associated with different credit-producing projects, distort marginal rates of substitution away from what is required to satisfy (1). The impacts of these distortions on the market equilibrium are shortly analyzed.

The marginal trade-offs between variables as revealed in (5) are as follows: Just as e_j and μ_i are perfect substitutes in the satisfaction of (1), $\gamma_{kj} a_k$ and $\gamma_{ij} E_i$ are perfect substitutes in the satisfaction of (5) where $\gamma_{kj}a_k$ is the quantity of credits bought by j from k and $\gamma_{ij}E_i$ is j's purchased share of the mean abatement offered by i. These quantities are controlled by varying the corresponding γs. The marginal rates of substitution between these commodities and $\gamma_{ij}\gamma_{lj}c_{il}$ is $-\sqrt{\beta V_j}$ where $\gamma_{ij}\gamma_{lj}c_{il}$ is the covariance between j's credit purchases from nonpoint sources i and l. The negative marginal rate of substitution implies that $\gamma_{ij}\gamma_{lj}c_{il}$ is a bad and j prefers to reduce its holdings of c_{il}.

The point source allowance constraint (5) implies enforcement contingent on *ex ante* actions. A point source is not liable for actual emissions. It is compliant if it buys enough offsets such that it satisfies (5). The point source is not culpable for adverse random events that increase actual emissions beyond permitted levels. This is an attractive design feature for two reasons. First, it is not fair to penalize the point source for events beyond its control. Second, the non-measurability of a_i makes proving culpability problematic.

Let us now look at j's cost structure. Let unconstrained production profits be π_j^0 and the restricted profit function be $\pi_j(a_j(e_j))$. The cost of credit generation is $\pi_i^0 - \pi_j(a_j)$. Let p be the unit price of a point source credit. It does not vary across transactions because the point source credit is homogeneous and the market is competitive. The payment that j makes to other point sources for credit purchase is $p\sum_{k\neq j}\gamma_{kj}a_k$. The payment received from credit sale is $p\sum_{k\neq j}\gamma_{jk}a_j$. Using the tautology that $\sum_k \gamma_{jk} \equiv 1$, net payments to other point sources are $p\big[\sum_k \gamma_{kj}a_k - a_j\big]$. Finally j's total payments to nonpoint sources is $\sum_i \gamma_{ij}q_i(E_i, \boldsymbol{c}_i)$. Summing, j's profit function is $\Pi_j = p[a_j - \sum_k \gamma_{kj}a_k] - [\pi_i^0 - \pi_j(a_j)] - \sum_i \gamma_{ij}q_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$.

Point source j optimizes by maximizing Π_j while satisfying (5) through control of credit generation, i.e. a_j, and ownership of other polluters' credits, i.e. $\{\gamma_{kj}, \gamma_{ij}\}_{i,k}$. The necessary conditions for profit maximization are (6) and (7). The first necessary condition (6) indicates that j maximizes profits by choosing its emissions level such that marginal profits are equated with the unit price of point source credits.

$$\frac{\partial \pi_j}{\partial e_j} = p \tag{6}$$

$$q_i^* = p\left(E_i(\mu_i) - \frac{\sum_l \gamma_{lj} c_{il}(\sigma_{il})}{\sqrt{\beta V_j}} \right) \quad \forall\, i \in I \tag{7}$$

The second set of conditions (7) indicates i's payment function when j optimizes. $\sum_l \gamma_{ij} c_{il}$ in (7) is the sum of all covariances associated with j's purchases of i's credits. It may be interpreted as the risk associated with j's investment in i's abatement technologies. Hence $\sum_l \gamma_{ij} c_{il}/\sqrt{\beta V_j}$ may be interpreted as the risk associated with j's investment in i's abatement relative to j's aggregate risk exposure. The RHS of (7) can be decomposed into additively separable functions of E_i and $\sum_l \gamma_{lj} c_{il}$, i.e. $q_i^* = pE_i - p\sum_l \gamma_{lj} c_{il}/\sqrt{\beta V_j}$. Since $p > 0$, q_i^* increases in E_i, which means that j is willing to pay more for nonpoint projects with higher expected levels of abatement. As comparison to (6) shows, E_i is treated equivalently to a_j, with both being (implicitly) sold for p at the margin. Since the pricing term for j's relative risk exposure to i, i.e. $-p\sum_l \gamma_{lj} c_{il}/\sqrt{\beta V_j}$, is negative, j penalizes i for stochasticity when maximizing its profits. Comparative statics indicates that the penalty increases in c_{il}, but at a decreasing rate.

Note that (6) and (7) are necessary conditions only when j trades with both point and nonpoint sources. If j only traded with other point sources then (6) is the only relevant necessary condition. However, such behavior is sub-optimal when nonpoint source credit generation is cheaper than point source credit generation, as is typically the case. If j traded with only nonpoint sources then (6) and (7) are necessary for profit maximization under the caveat that p exists. The existence of p is predicated upon the existence of trade among point sources. Given that the market is competitive, such trade will typically exist. If such trade did not exist then p may be set by reference to prices in other tradable permit markets. Other possible scenarios, such as j not trading at all or only selling credits, are sub-optimal given that (5) is a binding constraint and point source abatement is expensive.

Consider some $k \in J$ where $k \neq j$. The profit maximizing necessary conditions for k are also (6) and (7). It follows that if both k and j maximize profits then $q_i^* = p\big(E_i - \sum_l \gamma_{lj} c_{il}/\sqrt{\beta V_j}\big) = p\big(E_i - \sum_l \gamma_{lk} c_{il}/\sqrt{\beta V_k}\big)$. But the previous statement is only true when $\sum_l \gamma_{lj} c_{il}/\sum_l \gamma_{lk} c_{il} = \sqrt{V_j}/\sqrt{V_k}$. Or, a necessary condition for an arbitrary pair of point sources, j and k, to simultaneously maximize profits is that the ratio of their risk exposure to any $i \in I$ must equal the ratio of their aggregate risk exposure in the market. Now, j's aggregate risk exposure $\sqrt{V_j}$ is the standard deviation of all his credit purchases. Since $\sqrt{V}$ is the standard deviation of all credits in the market, $\sqrt{V_j} = \gamma_j \sqrt{V}$ for all $j \in J$, where $\gamma_j \in [0, 1]$. If j buys all nonpoint source credits then $\gamma_j = 1$ and if it buys none then $\gamma_j = 0$. Given the relationship between purchased risk and total risk, the ratio of j's and k's aggregate risk exposures is γ_j/γ_k. It follows that a necessary condition for any pair of point sources to simultaneously maximize profits is (8).

$$\frac{\sum_i \gamma_{ij} c_{il}}{\sum_i \gamma_{ik} c_{il}} = \frac{\sum_l \gamma_{lj} c_{il}}{\sum_l \gamma_{lk} c_{il}} = \frac{\gamma_j}{\gamma_k} \quad \forall\, i, j, k, l \in I \times J \tag{8}$$

The first equality in (8) indicates that the relative risk exposure of j and k to one nonpoint source must equal to their relative risk exposure to another arbitrary nonpoint source. If j's relative risk exposure is twice as much as k's with regard to i then its relative risk exposure will also be twice as much with regard to l. The second equality indicates that the ratio in which two point sources share risk with respect to an arbitrary nonpoint source's abatement is the same as the ratio between their aggregate risk positions.

Point source j must allocate risks across his portfolio of nonpoint source credits by choosing $\{\gamma_{1j}, \cdots, \gamma_{Ij}\}$ such that his aggregate risk is γ_j of the total market risk. Attaining this desired risk position is easily accomplished by taking up identical risk positions in all nonpoint source credit generating projects, i.e. setting $\gamma_{ij} = \gamma_j$ for all $i \in I$. If every point source behaves thus, then (8) holds trivially as $\sum_i \gamma_{ij} c_{il} / \sum_i \gamma_{ik} c_{il}$ simplifies to γ_j/γ_k for all $i, j, k, l \in I \times J$.

Result 1. *In a competitive market point sources jointly maximize profits when each point source j controls its exposure to risk by buying the same share $\gamma_j \in [0,1]$ in all nonpoint source credit-generating projects. This share is equal to j's share of the aggregate risk in the market. This behavior may be described as* Equi-Proportional Risk Sharing *across the different nonpoint credits.*

Under Equi-Proportional Risk Sharing j's probabilistic allowance constraint (5) amends to $\sum_k \gamma_{kj} a_k + \gamma_j \big(\sum_i E_i - \sqrt{V/\beta} \big) \geq 0$ through the substitution of γ_j for γ_{ij} and γ_{lj}, but its profit function remains unchanged. The necessary conditions for j's profit maximization transform to (6) and (9).

$$\sum_i \tilde{q}_i = p \left(\sum_i E_i - \sqrt{\frac{V}{\beta}} \right) \tag{9}$$

$\tilde{q}_i$ is the payment that i must receive upon supplying $\tilde{a}_i$ to the market if j is to maximize profits under equi-proportional risk sharing. $\sum_i \tilde{q}_i$ is the total payment from all j to all i. Similar to q_i^*, the total payment $\sum_i \tilde{q}_i$ is additively separable into functions of E_i and c_{il} for all $l \in I$. By differentiating the RHS of (9) the optimal marginal payments to i associated with E_i and c_{il} are calculated. The marginal payment for E_i is p. Since p is also the marginal payment for e_j, E_i and e_j are treated equivalently at the joint profit maximizing margin.

The marginal payment for c_{il} is $-p/\sqrt{\beta V}$ when $l \neq i$. Hence, i is penalized $p/\sqrt{\beta V}$ for increasing $c_{il} \geq 0$ by one unit and rewarded $p/\sqrt{\beta V}$ when increasing (in the absolute sense) $c_{il} < 0$ by one unit. When $c_{il} \geq 0$ ($c_{il} < 0$) then there is positive (negative) co-movement in i's and l's actual loadings, i.e. r_i and r_l. In other words, when $c_{il} \geq 0$ ($c_{il} < 0$) an increase in r_i is accompanied by an increase (decrease) in r_l. As a consequence, when $c_{il} \geq 0$ and it increases, an increase in r_i causes a greater than proportional increase in aggregate load. This reduces the *ex ante* probability that j satisfies its loadings constraint. Since this is undesirable, j penalizes such behavior. Conversely, when $c_{il} < 0$ increases (in the absolute sense), then an increase in r_i prompts a larger negative response in r_l. r_i and r_l counteract each other, which reduces j's risk of constraint violation. Since this is desirable, j rewards such behavior.

The marginal penalty for increasing c_{ii} is $p/2\sqrt{\beta V}$. Since an increase in c_{ii} increases the probability of constraint violation, it too is penalized by point sources. However, note that the penalty is only half that associated with an increase in $c_{il} \geq 0$. This is because *ceteris paribus* changes in c_{ii} do not have positive feedback effects on the loadings of other polluters. As a consequence, changes in c_{ii} do not cause greater than proportional changes in pollution loads.

The reward / penalty for deviating from the optimal c_{il} is decreasing in β. This is because a higher β implies that j has social sanction to invest in riskier abatement projects, which weakens j's allowance constraint, reducing its demand for low risk projects. Consequently, it will pay less for reductions in $|c_{il}|$. The reward / penalty is also falling in V. When V is high then c_{il}'s impact on j's total risk and constraint violation is low. As a result j will value it less and pay less for it.

The optimal relationship between $\tilde{q}_i$, E_i and $\boldsymbol{c}_i$ is not specified in (9). As long as (9) holds, $\tilde{q}_i$ could be described by any function. Each nonpoint source could face a different payment function. We propose a payment function of the form $\tilde{q}_i = p(E_i(\mu_i) - \sum_l c_{il}(\sigma_{il})/\sqrt{\beta V})$ where $\sum_l c_{il} = \sum_l \sigma_{il} \equiv V_i$ is the aggregate variance or risk associated with r_i. The Cauchy-Schwarz inequality shows that summing over $\tilde{q}_i$ yields (9) and hence maximizes j's profits given equi-proportional risk sharing.[3] The inferences on the separability of $\sum_i \tilde{q}_i$ hold for $\tilde{q}_i$.

The payment function $\tilde{q}_i$ imposes conditions on the characteristics of the credits $\tilde{a}_i(E_i, \boldsymbol{c}_i)$. From the definition of $\tilde{q}_i$ it follows that i gets a positive payment if and only if $V_i/E_i \leq \sqrt{\beta V}$,

[3]The Cauchy-Schwarz inequality states that $\sum_i x_i y_i \leq \sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}$ with strict equality when $x_i = w \cdot y_i$ for all i and w is a scalar. Setting $x_i = \sqrt{c_{il}}$, $y_i = w \cdot \sqrt{c_{il}}$ and $w = 1$, the Cauchy-Schwarz inequality indicates that $\sum_i \sum_l c_{il} = \sqrt{\sum_i \sum_l c_{il}}\sqrt{\sum_i \sum_l c_{il}}$.

where V_i/E_i is i's risk-to-mean ratio. $\sqrt{\beta V}$ is the socially sanctioned aggregate risk that point sources jointly bear. Hence, the payment function is constructed such that i only gets paid if its risk-to-mean ratio is less than the aggregate socially sanctioned risk level. From (5), when β is reduced or when V falls, j's allowance constraint tightens. To satisfy this tightened constraint, j must reduce the risk in its credit portfolio, which it does by purchasing credits with low risk-to-mean ratios.

3.4 The Market Equilibrium

Given that nonpoint source i earns $\tilde{q}_i(\mu_i, \boldsymbol{\sigma}_i)$ from the market, the necessary conditions for i to maximize profits are obtained by differentiating $\tilde{q}_i$ and substituting the results into (3) and (4). These conditions are shown in (10) and (11) below.

$$\frac{\partial \pi_i}{\partial \mu_i} = \frac{\partial \pi_i}{\partial \sigma_{il}} \sqrt{\beta V} \left(\frac{V}{V - V_i} \right) = p \qquad \forall\, i, l \neq i \tag{10}$$

$$\frac{\partial \pi_i}{\partial \sigma_{ii}} = \frac{\partial \pi_i}{\partial \sigma_{il}} \left(1 - \frac{V_i}{2\,(V - V_i)} \right) \qquad \forall\, i, l \neq i \tag{11}$$

$\sqrt{\beta V} \cdot V/(V - V_i)$ in (10) is a risk adjustment factor applied to marginal profits associated with σ_{il} when $l \neq i$. This factor is positive and increasing in V_i. Since this factor is constant for all $l \neq i$ it follows that $\partial \pi_i/\partial \sigma_{il} = \partial \pi_i/\partial \sigma_{im}\ \forall\ l, m \neq i$. According to (10) i maximizes profits by equating p to marginal profits with respect to μ_i and adjusted marginal profits with respect to all σ_{il} when $l \neq i$. Given that $\partial \pi_i/\partial \mu_i \geq 0$, (10) indicates that it is optimal for i to set $\partial \pi_i/\partial \sigma_{il} \geq 0$. We know that V and β are always positive. Hence, if $\partial \pi_i/\partial \sigma_{il} \leq 0$, it must be that $V < \sum_l \sigma_{il}$, which is impossible from the definition of V. It follows that at i's optimum, $\partial \pi_i/\partial \sigma_{il} \geq 0$ for all $l \neq i$. By prior assumption, we know that $\partial \pi_i/\partial \sigma_{il} \geq 0$ if and only if $\sigma_{il} \geq 0$. Hence, a necessary condition for i to maximize profits is that its loadings be positively correlated with the loadings of all other polluters.

Result 2. *A necessary condition for nonpoint source i to maximize profits is that positive* *marginal profits with respect to all covariances are equal. This requires that correlation between polluters must always be positive.*

The optimal relationship between σ_{ii} and σ_{il}, where $l \neq i$, is given in (11). From the definition of variance, $\sigma_{ii} \geq 0$, and from (10), $\sigma_{il} \geq 0$ for all $l \neq i$. It follows that $V_i \geq 0$, which implies that $\partial\pi_i/\partial\sigma_{il} \geq \partial\pi_i/\partial\sigma_{ii}$. Marginal returns with respect to σ_{il} are always greater than marginal returns with respect to σ_{ii} at i's profit maximum. Given that i's profits are concave in σ_{il} and $\sigma_{il} \geq 0$ for all $l \in I$, (11) indicates that σ_{ii} is relatively greater than σ_{il}, $l \neq i$. This difference increases as V_i increases. In other words, as the risk associated with i's loadings increases, i adjusts by focusing on reduction of σ_{il}. This is because its profits are more sensitive to changes in σ_{il} than to changes in σ_{ii}.

Result 3. *The profit maximizing nonpoint source exerts stricter control over σ_{il}, $l \neq i$, than over σ_{ii}. This is because his profits are more sensitive to changes in the former variable.*

The market equilibrium is reached when all participants in the market have exhausted gains from trade and maximized their profits. This implies that the necessary conditions for profit maximization must hold for all i and for all j. In other words, at the market equilibrium (6) and (9) must hold for all j and (10) and (11) must hold for all i. Note that (9) is a necessary condition for (10) and (11) to hold since the latter are derived from the former.

Result 4. *The MANA market is in equilibrium when the profit-maximizing condition* (6) *holds for all point sources and the profit-maximizing conditions* (10) *and* (11) *hold for all nonpoint sources. The satisfaction of the latter conditions is predicated upon each nonpoint source i being paid according to the payment function $\tilde{q}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$ upon supply of $\tilde{a}_i(E_i(\mu_i), \boldsymbol{c}_i(\boldsymbol{\sigma}_i))$ credits such that* (9) *holds.*

In a competitive market, since each polluter is atomistic, $V - V_i \to V$ and $V_i/(V - V_i) \to 0$, in which case (10) is rewritten as $\sqrt{\beta V} \cdot \partial\pi_i/\partial\sigma_{il} \to \partial\pi_i/\partial\mu_i = p$ and (11) is rewritten as

$\partial\pi_i/\partial\sigma_{il} \rightarrow \partial\pi_i/\partial\sigma_{ii}$ for all $i, l \in I \times I, i \neq l$. Hence at the competitive market equilibrium profit making nonpoint sources equate marginal profits with respect to all variances and covariances. Also, $\partial\pi_i/\partial\sigma_{il} \rightarrow p/\sqrt{\beta V}$, which implies that the optimum emissions variances and covariances, i.e. $\boldsymbol{\sigma}_i^*$, are functions of total variance in the market, V. Since $\pi_i(\mu_i, \boldsymbol{\sigma}_i)$ is concave in $\boldsymbol{\sigma}_i$, when V is high then i maximizes profits by setting $\boldsymbol{\sigma}_i^*$ high and vice versa. The intuition might be as follows: Consider an i that sets σ_{il} low when V is high. By reducing σ_{il} marginally i earns $p/\sqrt{\beta V}$, which is small when V is high. The concavity of $\pi_i(\mu_i, \boldsymbol{\sigma}_i)$ in σ_{il} implies that the marginal reduction in σ_{il} is costly when σ_{il} is low. An assessment of costs and benefits would convince the nonpoint to keep σ_{il} high when V is high.

3.5 The Least Cost Safety-First Outcome and the MANA Market Equilibrium

The pertinent issue for the EA is whether conditions can be imposed on the market such that the safety-first target (1) is met at least cost at the market equilibrium. This requires that the market equilibrium satisfy (2). In general, this condition will be unsatisfied because polluters evaluate $\boldsymbol{\sigma}_i$ differently to the EA. The EA's optimal marginal rates of technical substitution typically differs to that of nonpoint source i's: $MRTS^{LCSF}_{\mu_i,\boldsymbol{\sigma}_i}/MRTS^{MME}_{\mu_i,\boldsymbol{\sigma}_i} = 2(1-V_i/V)\sqrt{\alpha/\beta}$. $MRTS^{LCSF}_{\mu_i,\boldsymbol{\sigma}_i}$ is the optimal $MRTS$ at the least cost safety-first outcome, which is the EA's optimum outcome. $MRTS^{MME}_{\mu_i,\boldsymbol{\sigma}_i}$ is the optimal $MRTS$ at the MANA market equilibrium, which is i's optimum outcome. When this ratio is greater than one then the EA values μ_i more than the polluters do. When this ratio equals one then the EA and polluters have identical valuations of μ_i and $\boldsymbol{\sigma}_i$ and the market equilibrium is the least cost safety-first outcome.

The EA can adjust α and β, which are society's tolerances for aggregate and individual target violation, to ensure that the market equilibrium will satisfy (2). Since marginal returns with respect to σ_{il} are equal for all combinations of i and l at the market equilibrium, comparison of (2) and (10) indicates that a necessary condition for the market equilibrium to satisfy (2) is $\sqrt{\beta/\alpha} = 4(1 - V_i/V)$. From this relationship, $\beta = \alpha$ if and only if $V_i = V/2$ for all i. In other words, having identical tolerances for individual and aggregate pollution target violation is optimal only when there are two nonpoint sources of pollution, each supplying exactly half the aggregate pollution risk. When the market is competitive, the other hand, $V_i/V \to 0$ and $\beta/\alpha \to 4$. The sanctioned probability that individual point sources violate their allowances should be four times greater than the sanctioned probability with which the aggregate safety-first target may be violated. If the sanctioned probability at the aggregate level is 5% then individual point sources should be sanctioned a 20% probability of violation. It is trivial to ensure that the competitive equilibrium outcome in the MANA market is safety-first compliant at least cost given the simple relationship between β and α.

Result 5. *The equilibrium in the MANA market satisfies the least cost safety-first outcome under a single condition: that the ratio between the sanctioned probabilities of individual and aggregate violation be chosen such that* $\sqrt{\beta/\alpha} = 2(1 - V_i/V)$. *When the market is competitive and* $V_i/V \to 1$ *then* $\beta/\alpha = 4$.

4 The Trading Ratio Market

Extant water quality trading markets do not use a multi-attribute definition of the nonpoint source credit. Instead the credit is defined as the expected value E_i of the true reduction in emissions below the allowance, a_i. Since $\mu_i \equiv E(r_i)$ it follows that the credit is $E_i(\mu_i) = \hat{r}_i - \mu_i$. Unlike in the MANA market, i's credit is a first order approximation of a_i. This approximation prevents transmission of information about the risk $\boldsymbol{\sigma}_i$ associated with i's runoff to the marketplace. Also, in extant water quality trading markets, point source constraints are not based on safety-first environmental targets, which imply that $\boldsymbol{\sigma}_i$ will not affect trading decisions. However, $\boldsymbol{\sigma}_i$ is a factor in the satisfaction of the safety-first environmental target (1) and it must be controlled if (1) is to be satisfied at least cost. We analyze conditions under which extant water quality trading markets are safety-first compliant at least cost in spite of not supplying price incentives for the control of $\boldsymbol{\sigma}_i$ for all i.

Instead of using the market mechanism to control $\boldsymbol{\sigma}_i$, an *ad hoc* adjustment called the 'trading ratio' is used to control emissions risk in current water quality trading markets. The trading ratio or TR is defined as the number of units of mean nonpoint source abatements required to allow a unit increase in point source loadings (Shortle, 1990). It is calculated by the EA and applied to all trades where nonpoint sources participate. Proponents of the TR argue that when it is correctly designed it controls the stochasticity of nonpoint emissions. Through its application point and nonpoint source credits are converted into perfect substitutes so that in effect apples will trade for apples (Woodward, 2000). There is no consensus on the design of the optimal TR. For example, Horan and Shortle (2005) find that the optimal TR must reflect the marginal environmental impacts of the trading polluters, the riskiness of their abatements and the transactions costs of the trade. In contrast Malik, Letson, and Crutchfield (1993) find that the optimal TR must also reflect abatement and enforcement costs.

4.1 The Trading Ratio Market Equilibrium

As in the MANA market, j earns revenues from credit sales, and faces credit generation and credit purchase costs. Its profit function is $\Pi_j = p[a_j - \sum_k \gamma_{kj} a_k] - [\pi_i^0 - \pi_j(a_j)] - \sum_i \gamma_{ij} q_i^t(E_i(\mu_i))$ where $p[a_j - \sum_k \gamma_{kj} a_k]$ is net income from trade with other point sources, $[\pi_i^0 - \pi_j(a_j)]$ is the cost of credit generation and $\sum_i \gamma_{ij} q_i^t(E_i(\mu_i))$ is total payment by j to all $i \in I$ upon procurement of $\sum_i \gamma_{ij} E_i$ credits. Let t_{ij} be the TR applied to a trade between i

and j. If $t_{ij} = 2$ then buying two units of abatement from i entitles j to increase its emissions by one unit. Analogous to the design in Horan and Shortle (2001) the emissions allowance for j is $\sum_k \gamma_{kj} a_k + \sum_i (\gamma_{ij}/t_{ij}) E_i(\mu_i) \geq 0$. Point source j optimizes by maximizing Π_j subject to this allowance. The necessary conditions, obtained by manipulation of the Kuhn-Tucker conditions, are (6) and (12).

$$q_i^t = p \frac{E_i(\mu_i)}{t_{ij}} \quad \forall\, i \in I \tag{12}$$

Point source j maximizes profits when i is paid q_i^t for supplying $\gamma_{ij} E_i$ credits. As (12) indicates, these payments decrease as t_{ij} increases. Also, since $E_i = \hat{r}_i - \mu_i$ it follows that q_i^t is inversely related to μ_i. Nonpoint sources with high mean emissions are paid less, which is consistent with a policy geared towards reducing mean emissions.

Since (12) is a necessary profit maximizing condition for all j it follows that $q_i^t = pE_i/t_{ij} = pE_i/t_{ik}$ for all $j, k \in J$, which in turn implies that $t_{ij} = t_{ik} = t_i$ for all $j, k \in J$. The TR must be chosen such that it varies across nonpoint sources but not across individual trades. This result differs from previous results in the literature where the TR is trade-specific, affected as it is by point source and nonpoint source characteristics (Malik, Letson, and Crutchfield, 1993; Horan and Shortle, 2005). The differences stem from the fact that earlier models only studied interaction between a single point source and a single nonpoint source, whereas we model the entire market.

Let us now look at the nonpoint source profit maximization problem in the TR market. Like in the MANA market i has one cost and one revenue stream. The cost is that of generating credits and is $\pi_i^0 - \pi_i(\mu_i)$, where π_i^0 is the unconstrained maximized profit and μ_i is mean emissions associated with E_i credits. The revenue $q_i^t(E_i)$ comes from the sale of E_i. When point sources are profit maximizers then $q_i^t = pE_i/t_i$ by (12). Hence i's total profits are $\Pi_i = pE_i(\mu_i)/t_i - [\pi_i^0 - \pi_i(\mu_i)]$. Like in the MANA market i does not face a binding constraint. It optimizes by choosing μ_i such that profits are maximized. The necessary condition is

$$\frac{\partial \pi_i}{\partial \mu_i} = \frac{p}{t_i} \tag{13}$$

Equation (13) indicates that i maximizes profits only if the EA chooses the TR, $\tilde{t}_i$, such that $\tilde{t}_i = p \cdot \partial\mu_i/\partial\pi_i$. Since $\partial\pi_i/\partial\mu_i \geq 0$ it follows that $\tilde{t}_i \geq 0$. The concavity of $\pi_i(\mu_i)$ indicates $\tilde{t}_i$ increases in μ_i. When μ_i increases then $\partial\pi_i/\partial\mu_i$ decreases (by concavity), which implies that $\tilde{t}_i$ increases (by (13)).

Substituting $\tilde{t}_i$ into (12) yields the payment that maximizes i's profits, $\tilde{q}_i^t = (\hat{r}_i - \mu_i)\partial\pi_i/\partial\mu_i$. Note that $\tilde{q}_i^t \rightarrow 0$ as $\mu_i \rightarrow \hat{r}_i$ and $\max(\tilde{q}_i^t) = \hat{r}_i \partial\pi_i/\partial\mu_i$ when $\mu_i = 0$. The concavity of π_i with respect to μ_i implies that $\tilde{q}_i^t$ is more that proportionately responsive to changes in μ_i when μ_i is low and vice versa. This implies strong market preferences for credits with low μ_i.

At the market equilibrium all polluters are profit maximizers, which implies that (6) and (13) must hold simultaneously for all i and j. The necessary condition for market equilibrium in the TR market is

$$\frac{\partial \pi_j}{\partial e_j} = \tilde{t}_i \frac{\partial \pi_i}{\partial \mu_i} \quad \forall \, i, j \in I \times J \tag{14}$$

At the TR market equilibrium, point and nonpoint sources have identical marginal costs of credit generation after adjusting for the TR. The results are similar to the equi-marginality results that characterize standard pollution trading models, except for the role played by the TR.

To optimally choose the TR the EA must know the market equilibrium outcome, which implies perfect information. If information were perfect however, the market is unnecessary because the EA could directly allocate allowances such that (1) is satisfied at least cost. In reality, since information is not perfect, the TR is suboptimal and so are market outcomes, characterized by low market participation and low earnings (Breetz et al., 2004).

4.2 The Least Cost Safety First Outcome and the TR Market Equilibrium

The least cost safety-first outcome requires control of μ_i and $\boldsymbol{\sigma}_i$ for all i and e_j for all j. Since $\sigma_{il} = \sigma_{li}$ by the definition of covariance, this outcome requires control of $\sum_{i=1}^{I} i + I + J$ variables. However, in the TR market there are only $I + J$ control variables since market rules imply that control of $\boldsymbol{\sigma}_i$ is not required. The TR market will generally never achieve the least cost safety-first outcome because the market supplies no signals to control environmental risk.

The only scenario where the TR market equilibrium does achieve the least cost safety-first outcome is when control of the I μ_is and J e_js also results in control of the $\sum_i i\ \sigma_{il}$s. Since point source abatement activity is independent of nonpoint source abatement activity, this requires $\sum_{i=1}^{I} i\ \sigma_{il}$s in the EA's problem to become deterministic functions of the I μ_is. In its most general form, this requires that $\sigma_{il} = \sigma_{il}(\mu_1, \cdots, \mu_i)$ for all $i, l \in I$. Under this scenario the safety-first environmental constraint amends from (1) to $\sum_j e_j + \sum_i \mu_i + \sqrt{\sum_i \sum_l \sigma_{il}(\mu_1, \cdots, \mu_i)/\alpha} \leq \bar{L}$ and the necessary conditions for the least cost safety-first solution transform from (2) to

$$\frac{\partial \pi_j}{\partial e_j} = \frac{\partial \pi_i}{\partial \mu_i}\left(1 + \frac{1}{2\sqrt{\alpha V}}\left(\sum_i \sum_l \frac{\partial \sigma_{il}}{\partial \mu_i}\right)\right)^{-1} \quad \forall\, i, l \in I, j \in J \tag{15}$$

$\sum_i \sum_l \partial \sigma_{il}/\partial \mu_i$ is the marginal change in variability of all pollution in response to a marginal change in μ_i. It may be interpreted as the marginal risk (MR_i) associated with μ_i at the market equilibrium and can be positive or negative. A positive (negative) MR_i implies that i's contribution to aggregate variance V (also interpreted as total market risk) increases (decreases) in μ_i. Hence (15) implies that when $\sigma_{il} = \sigma_{il}(\mu_i, \cdots, \mu_I)$, then at the least cost safety-first outcome risk-adjusted nonpoint source marginal profits are equated with point source marginal profits. The TR market will satisfy the safety-first environmental target at least cost when the necessary conditions for the TR market equilibrium (14) are identical to

(15), which requires that

$$\frac{1}{\tilde{t}_i} = 1 + \frac{1}{2\sqrt{\alpha V}} \left(\sum_i \sum_l \frac{\partial \sigma_{il}}{\partial \mu_i} \right) \quad \forall\, i, l \in I \tag{16}$$

Equation (16) indicates that $\tilde{t}_i < 1$ when $MR_i \geq 0$ and vice versa. The lower the TR the more favorable the terms of trade for the nonpoint source and, by (12), the higher the earnings from credit sale. Hence, by (16), optimality of the TR market requires that more favorable terms of trade be given to abatement projects where total market risk V is positively correlated to μ_i. To see why this is so, we study the behavioral effects induced by adjusting $\tilde{t}_i$. When $\tilde{t}_i$ is reduced i compensates by increasing $\partial \pi_i / \partial \mu_i$ [see (13)]. Since $\pi_i(\mu_i)$ is concave in μ_i, this requires a reduction in μ_i. When V trends in the same direction as μ_i, a reduction in μ_i causes a reduction in V. From (1) μ_i and V are substitutes in the attainment of the EA's safety-first environmental target and reduction in both are sought. Simultaneous reductions in both are attained when $MR_i > 0$ and μ_i is reduced. Since this is aligned with the EA's preferences, it favors such projects by giving them advantageous terms of trade.

Conversely, the EA gives disadvantageous terms of trade to i $\tilde{t}_i > 1$ when $MR_i < 0$ and μ_i and V move in opposite directions. This is because such projects cannot satisfy the requirement that both mean emissions and the risk associated with emissions reduce simultaneously. When $MR_i < 0$ a reduction in μ_i is accompanied by an increase in V and vice versa.

The TR market always fails when $MR_i = 0$ and μ_i and V are independent because then pollution risk is not controlled. When $MR_i \approx 0$ then $\tilde{t}_i \approx 1$ for all i, the marginal costs of (expected) abatement are equated across all polluters and the TR market is a poor instrument for the control of V. The EA cannot ensure attainment of the safety-first environmental target at least cost through the TR market. However, whatever the aggregate pollution level may be, it will be achieved at least cost.

Result 6. *Equilibrium in the trading ratio market satisfies the safety-first environmental target at least cost under three conditions. First, the EA must have perfect information because it must have knowledge of the market equilibrium to optimally allocate trading ratios. Second, the covariances of all nonpoint emissions must be deterministic functions of the mean:* $\sigma_{il} = \sigma_i(\mu_1, \cdots, \mu_I)\ \forall\ i \in I$. *Finally,* $\tilde{t}_i$, σ_{il} *and* μ_i *must satisfy* (16).

In most existing markets the TR is uniform and $\tilde{t}_i = 2\ \forall\ i$, (Breetz et al., 2004; Morgan and Wolverton, 2005). This is optimal only if μ_i and V move in opposite directions *and* $-\sum_i \sum_l \partial\sigma_{il}/\partial\mu_i = \sqrt{\alpha V}$ for all i [see (16)]. Also, when V is large then it is *extremely* responsive to changes in μ_i. Determination of whether such responsiveness is feasible, indeed possible, is beyond the scope of this paper, but it does seem unlikely. Instead, if a high response of V to μ_i is infeasible, (16) indicates that $\tilde{t}_i$ will be closer to one than two. When V and μ_i move in the same direction, optimal trading ratios will also be close to one. When $\tilde{t}_i = 2/3$, (16) indicates that $\sum_i \sum_l \partial\sigma_{il}/\partial\mu_i = \sqrt{\alpha V}$. V must be extremely responsive to the mean for a TR of 2 / 3 to be optimal. If such responsiveness is infeasible then $\tilde{t}_i$ will be closer to one.

In a scenario where allowances are externally chosen, if the TR market equilibrium is to satisfy the safety-first market outcome at least cost, then $\tilde{t}_i$ is close to one for all i. The uniform TR of two, which is commonly used in existing markets, is sub-optimal. These results imply that trading ratios are very sensitive instruments for achieving safety-first environmental targets and must be calibrated very precisely. But such precision is unlikely in a stochastic environment where the true relationship between μ_i and σ_{il} is unknown. In the real world the assumed relationship is merely a working hypothesis. We infer is that the TR is a poor instrument for attaining the safety-first environmental target at least cost.

5 Discussion and Conclusions

As stated in Result 6, there are three necessary conditions for the equilibrium in the TR market to satisfy the safety-first target at least cost. It is unlikely that any of these conditions are satisfied in a real world trading environment. The EA does not have information on the private costs of polluters and at the current state of the science the true probability distribution of nonpoint runoff is unknown. Imposing a deterministic relationship between μ_i and σ_{il} requires that a parametric distribution be fitted to the data. The validity is such an imposition cannot be ascertained. A wide variety of distributions can typically be fitted to a data set, each specifying different marginal relationships between μ_i and σ_{il}. The choice between these distributions is arbitrary, which implies that the assumed relationship between μ_i and σ_{il} is arbitrarily assigned, which in turn implies that the choice of TR is arbitrary. Precise calibration of the TR is impossible in the real world. But as discussed in Section 4.2 TRs are extremely sensitive market instrument and the market fails if they are not correctly calibrated. In the real world the TR market will not achieve the least cost safety-first outcome.

The MANA market does not require a deterministic relationship between E_i and $\boldsymbol{c}_i$ and does not require sensitive calibration. Instead it relies on two rules and the market mechanism to generate the least cost safety-first outcome. The first rule defines the tradable nonpoint source credit as a second order approximation of the underlying stochastic emissions load. The second is that the sanctioned probability with which a point source may violate its allowance should be approximately four times greater than the sanctioned probability with which the safety-first environmental target may be violated. When these rules are satisfied then the market equilibrium always satisfies the safety-first target at least cost.

The degree of EA intervention is much lower in the MANA market. The EA's role is purely informational: it is required to provide information on μ_i and σ_{il} for all $i \in I$. In the TR market, by contrast, apart from providing information on μ_i the regulator is also required to calculate the optimal trading ratio $\tilde{t}_i$ for all $i \in I$. Although not modeled, transactions costs will be higher in real world trading ratio markets.

The reader may observe similarities between the MANA market and the CAPM market in the finance literature. Credits are similar to securities and both markets are designed to control variance and risk. However, as always, the devil is in the details and here differences are revealed. First, the markets exist for different reasons. There is no broad policy intent behind the CAPM. The market exists *a priori* and the CAPM analysis focusses on mean-variance efficiency as a solution concept. The MANA market on the other hand, like environmental markets in general, is brought into existence by policy dictate and the solution concept centers around the efficient attainment of the environmental target. Second, our model is static while the CAPM is typically a multi-period model. Third, there are differences in model development. We do not impose distributional assumptions on nonpoint source credits and there is no space for interest rate differentials in the model. Also, analysis of the constraint is the starting point in our model. Fourth, behavioral constraints are organic in the CAPM, springing from preferences and budget constraints. In the MANA market, on the other hand, constraints are generated through policy fiat. Overall, there are similarities in results across the two models, arising from a congruence of interests related to risk control, but model development and inferences differ.

Numerous extensions to the model are possible. An interesting possibility is to conduct a series of economic experiments to test the robustness of the results to violations in underlying assumptions. The assumption of competitiveness in the market can be removed. Real world markets are oligopsonistic with few large point sources and many small nonpoint sources. Another removable assumption pertains to the continuity of abatement production function. In reality decisions have discrete consequences and are not made using the marginality analysis that underpins our results. The experiments will be relevant to policy design because they represent an intermediate step in the transfer of policy ideas from the academic to the real world.

References

Beavis, B., and M. Walker. 1983. "Achieving Environmental Standards with Stochastic Discharges." *Journal of Environmental Economics and Management* 10:103–111.

Breetz, H., K. Fisher-Vanden, L. Garzon, H. Jacobs, K. Kroetz, and R. Terry. 2004. "Water Quality Trading and Offset Initiatives in the US: A Comprehensive Survey." Unpublished, Available online at http://www.dartmouth.edu/%7Ekfv/waterqualitytradingdatabase.pdf.

Cabe, R., and J. Herriges. 1992. "The Regulation of Non-Point-Source Pollution under Imperfect and Asymmetric Information." *Journal of Environmental Economics and Management* 22(2):134–146.

Crocker, T. 1966. *The Economics of Air Pollution*, W.W. Norton & Co., chap. The Structuring of Atmospheric Pollution Control Systems.

Dales, J. 1968. *Pollution, Property and Prices*. Toronto, Canada: University of Toronto Press.

Davies, J., and J. Mazurek. 1998. *Pollution Control in the United States: Evaluation the System*, illustrated ed. RFF Press.

Graff Zivin, J., and A. Small. 2003. "Risk Sharing in Coasean Contracts." *Journal of Environmental Economics and Management* 45:394–415.

Hennessy, D., and H. Feng. 2008. "When should Uncertain Nonpoint Emissions be Penalized in a Trading Program?" *American Journal of Agricultural Economics* 90:249–255.

Horan, R., and J. Shortle. 2001. *Environmental Policies for Agricultural Pollution Control*, Oxfordshire, U.K.: CAB International, chap. Environmental Instruments for Agriculture.

—. 2005. "When Two Wrongs make a Right: Second-Best Point-Nonpoint Trading Ratios." *American Journal of Agricultural Economics* 87:340–352.

Horan, R., J. Shortle, and D. Abler. 2002. "Point-Nonpoint Nutrient Trading in the Susquehanna River Basin." *Water Resources Research* 38:8–1–8–13.

Horan, R.D. 2001. "Differences in Social and Public Risk Perceptions and Conflicting Impacts on Point / Nonpoint Trading Ratios." *American Journal of Agricultural Economics* 83:934–941.

Kampas, A., and B. White. 2003. "Selecting Permit Allocation Rules for Agricultural Pollution Control: A Bargaining Solution." *Ecological Economics* 47:135–147.

Lichtenberg, E., and D. Zilberman. 1988. "Efficient Regulation of Environmental Health Risks." *The Quarterly Journal of Economics* 103:167–178.

Lobsenz, G. 1993. "SO_2 Auction a Success – EPA, CBOT." *Energy Daily* 21.

Malik, A., D. Letson, and S.R. Crutchfield. 1993. "Point/Nonpoint Source Trading of Pollution Abatement: Choosing the Right Trading Ratio." *American Journal of Agricultural Economics* 75:959–967.

Montgomery, W. 1972. "Markets in Licences and Efficient Pollution Control Regimes." *Journal of Economic Theory* 5:395–418.

Morgan, C., and A. Wolverton. 2005. "Water Quality Trading in the United States." Unpublished, Working Paper #05-07, National Center for Environmental Economics.

National Research Council. 2001. "Assessing the TMDL Approach to Water Quality Management." Working paper, National Research Council, Washington, DC.

Pennsylvania Department of Environmental Protection. 2007. “Agricultural Credit Calculations 12-4.” Working paper, PADEP.

Qiu, Z., T. Prato, and F. McCamley. 2001. “Evaluating Environmental Risks using Safety-First Constraints.” *American Journal of Agricultural Economics* 83:402–413.

Rabotyagov, S., H. Feng, and C. Kling. 2006. “Optimal Design of Permit Markets with an *Ex Ante* Pollution Target.” Unpublished, Working Paper 06-WP 430, Center for Agricultural and Rural Development, Iowa State University, Ames, IA.

Reckhow, K. 1994. “Importance of Scientific Uncertainty in Decision Making.” *Environmental Management* 18:161–166.

—. 2003. “On the Need for Uncertainty Assessment for TMDL Modeling and Implementation.” *Journal of Water Resources, Planning and Management* 129:245–246.

Ribaudo, M., R. Heimlich, and M. Peters. 2005. “Nitrogen Sources and Gulf Hypoxia: Potential for Environmental Credit Trading.” *Ecological Economics* 52:159–168.

Segerson, K. 1988. “Uncertainty and Incentives for Nonpoint Pollution Control.” *Journal of Environmental Economics and Management* 15:87–98.

Shortle, J. 1990. “The Allocative Efficiency Implication of Water Pollution Abatement Cost Comparisons.” *Water Resources Research* 26:793–797.

—. 1987. “Allocative Implications of Comparisons between the Marginal Costs of Point and Nonpoint Source Pollution Abatement.” *Northeastern Journal of Agricultural and Resource Economics* 16:17–23.

Shortle, J., and R. Horan. 2008. “The Economics of Water Quality Trading.” *International Review of Environmental and Resource Economics* 2:101–133.

Tietenberg, T. 1990. “Economic Instruments for Environmental Regulation.” *Oxford Review of Economic Policy* 6:17–33.

US Environmental Protection Agency. 2001. “The National Costs of the Total Maximum Daily Load Program (Draft Report).” Working paper, Office of Water, USEPA, Washington DC, 841-D-01-003, August.

—. 2002. “National Water Quality Inventory, 2000 Report.” Working paper, Office of Water, USEPA, Washington DC, EPA-841-R-02-001, August.

—. 2003. “Water Quality Trading Policy.” Working paper, Office of Water, USEPA, Washington DC, January.

Wets, R. 1983. *Mathematical Programming: The State of the Art*, Berlin: Springer, chap. Stochastic Programming: Solution Techniques and Approximation Schemes. pp. 566–603.

Woodward, R. 2000. “Discussant’s Comments on ”Market Based Solutions to Environmental Problems.” Unpublished, Invited Paper Session at the Annual Meeting of the Southern Agricultural Economics Association.

E.ON Energy Research Center

fcn Institute for Future Energy Consumer Needs and Behavior

RWTH AACHEN UNIVERSITY

List of FCN Working Papers

2009

Madlener R., Mathar T. (2009). Development Trends and Economics of Concentrating Solar Power Generation Technologies: A Comparative Analysis, FCN Working Paper No. 1/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Madlener R., Latz J. (2009). Centralized and Integrated Decentralized Compressed Air Energy Storage for Enhanced Grid Integration of Wind Power, FCN Working Paper No. 2/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Kraemer C., Madlener R. (2009). Using Fuzzy Real Options Valuation for Assessing Investments in NGCC and CCS Energy Conversion Technology, FCN Working Paper No. 3/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Westner G., Madlener R. (2009). Development of Cogeneration in Germany: A Dynamic Portfolio Analysis Based on the New Regulatory Framework, FCN Working Paper No. 4/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Westner G., Madlener R. (2009). The Benefit of Regional Diversification of Cogeneration Investments in Europe: A Mean-Variance Portfolio Analysis, FCN Working Paper No. 5/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Lohwasser R., Madlener R. (2009). Simulation of the European Electricity Market and CCS Development with the HECTOR Model, FCN Working Paper No. 6/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Lohwasser R., Madlener R. (2009). Impact of CCS on the Economics of Coal-Fired Power Plants – Why Investment Costs Do and Efficiency Doesn't Matter, FCN Working Paper No. 7/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Holtermann T., Madlener R. (2009). Assessment of the Technological Development and Economic Potential of Photobioreactors, FCN Working Paper No. 8/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Ghosh G., Carriazo F. (2009). A Comparison of Three Methods of Estimation in the Context of Spatial Modeling, FCN Working Paper No. 9/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

Ghosh G., Shortle J. (2009). Water Quality Trading when Nonpoint Pollution Loads are Stochastic, FCN Working Paper No. 10/2009, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, November.

2008

Madlener R., Gao W., Neustadt I., Zweifel P. (2008). Promoting Renewable Electricity Generation in Imperfect Markets: Price vs. Quantity Policies, FCN Working Paper No. 1/2008, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, July (revised May 2009).

Madlener R., Wenk C. (2008). Efficient Investment Portfolios for the Swiss Electricity Supply Sector, FCN Working Paper No. 2/2008, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, August.

Omann I., Kowalski K., Bohunovsky L., Madlener R., Stagl S. (2008). The Influence of Social Preferences on Multi-Criteria Evaluation of Energy Scenarios, FCN Working Paper No. 3/2008, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, August.

Bernstein R., Madlener R. (2008). The Impact of Disaggregated ICT Capital on Electricity Intensity of Production: Econometric Analysis of Major European Industries, FCN Working Paper No. 4/2008, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, September.

Erber G., Madlener R. (2008). Impact of ICT and Human Skills on the European Financial Intermediation Sector, FCN Working Paper No. 5/2008, Institute for Future Energy Consumer Needs and Behavior, RWTH Aachen University, September.