Title
Nonpoint source pollution control under incomplete and costly information

Journal
Environmental & Resource Economics, 28(4)

ISSN
0924-6460

Authors
Farzin, Y. Hossein
Kaplan, J D

Publication Date
2004-08-01

Nonpoint Source Pollution Control under Incomplete and Costly Information

Y. H. Farzin*

University of California, Davis
Department of Agricultural and Resource Economics

and

J. D. Kaplan

California State University, Sacramento
Department of Economics

Abstract

We analyze the efficient management of nonpoint source pollution (NPS) under a limited pollution control budget and incomplete information. We focus on the tradeoff between data collection and pollution abatement efforts by incorporating information acquisition into a NPS pollution control model. Comparative static results show conditions under which (*i*) a favorable change in the abatement costs at one source may lead to an *increase* in the treatment level at all sources, and vice versa, (*ii*) an increase in data collection cost leads to an *increase* in data collection level, and (*iii*) an increase in the efficiency of information acquisition leads to a *decrease* in the level of data collection. More importantly, the model simulations illustrate that acquiring and exploiting information on heterogeneity of sediment loading distributions across polluting sources leads to a more efficient budget allocation and hence a greater reduction in pollution damage than would be the case without such information.

***JEL Classification*:** D61, D81, D83, Q28

Key Words: nonpoint source pollution, uncertainty, costly information, constrained pollution control budget

****Correspondences to***: Y. H. Farzin, Department of Agricultural and Resource Economics, University of California, Davis, One Shield Avenue, Davis, CA 95616, U.S.A., Phone (530) 752-7610, Fax (530) 752-5614, Email: farzin@primal.ucdavis.edu

We thank the two anonymous referees of this journal for very helpful comments and suggestions and also the participants at annual conferences of the European Association of Environmental and Resource Economists (EAERE), Oslo, Norway, and the American Agricultural Economics Association (AAEA), Nashville, U.S.A. Farzin thankfully acknowledges a research grant from Giannini Foundation. Financial support for Kaplan was provided, in part, by a grant (Project No. W-887) from the University of California Water Resources Center, and the United States Environmental Protection Agency "Science to Achieve Results" Graduate Fellowship Program.

Nonpoint Source Pollution Control Under Incomplete and Costly Information

1. *Introduction*

This paper examines the role of information acquisition in efficient management of NPS pollution. By incorporating information acquisition into a NPS pollution control model, we focus on the tradeoff between abatement effort and abatement effectiveness, a question that has not received adequate attention in the literature. We explicitly consider the heterogeneity among the polluting sources in the manager's decision to reduce the pollution-related damage for a given expenditure on abatement activity. We analyze how abatement cost, data collection cost, and efficiency of extracting information from collected data affects the efficient budget allocation between information acquisition and abatement effort across polluting sources, and also elaborate on the policy implications.

In the analysis NPS pollution is defined as pollution from diffuse sources where the information on the linkage between polluting sources and ambient load is incomplete. The pollution manager observes total ambient load or the consequent damage but is unable to detect with certainty the pollution from individual sources. This incomplete information (uncertainty) about the pollution loading creates inefficiencies in allocation of abatement effort across the sources. The pollution control manager depicted in this model reduces pollution loading uncertainty by obtaining information through data collection.

The manager updates her subjective prior distribution about the pollution loading with the acquired information, resulting in a posterior distribution that improves abatement effectiveness by allowing the manager to reallocate abatement effort to sources with relatively larger expected pollution loading, all else the same. However, the manager eventually faces an explicit tradeoff between the scale of abatement effort and abatement effectiveness because data collection is costly and the manager is fiscally constrained.

We consider, for example, the case of sediment loading from forestland in Redwood National Park, located in northwestern California. Sediment loading mostly occurs during high storm events, (i.e., when rainfall intensity is high and storm duration is short), when storm run-off overflows stream channels at road crossings, causing sediment to enter tributaries as the run-off returns to the channel downstream.[1] The sediment that enters the waterways in Redwood National Park fills in salmon spawning pools, thus reducing the number of available spawning sites. This sediment also fills in the stream channel upstream and adjacent to the Tall Trees Grove, home of the world's tallest trees, increasing the incidence of flooding, bank erosion and saturation of the root zone, which all cause the tall trees to topple (Sprieter, Franke and Steensen, 1981).

If perfect information was available on sediment loading attributable to each pollution sources (the logging roads), park managers could allocate their entire sediment control budget to abatement effort. However, with incomplete information, the management of sediment loading requires an explicit allocation of resources between information collection and abatement. The results drawn from this analysis can, in general, shed light on the role of information and budgetary constraints on the efficient management of NPS pollution problems such as groundwater contamination, greenhouse gas emissions and acid rain, which are characterized as pollution generated from diffuse sources. This research has implications for other targeting programs as well. Babcock et al. (1996, 1997) consider targeting options under the USDA Conservation Reserve Program. However, in these analyses, there is no explicit provision of information that allows the budget managers to improve the allocation of limited resources, thereby increasing overall environmental benefits.

The pollution control literature typically looks at market-based approaches to controlling NPS pollution. However, market-based approaches are not relevant when a private individual or public manager must decide where on their property (or the property they manage) to concentrate pollution abatement efforts (i.e., target abatement resources). An individual polluter may not face an explicit budget constraint but a public manager is usually limited in her abatement decisions by a fixed annual budget.[2]

Previous research on NPS pollution control has also focused on the social welfare optimization problem without regard for fiscal constraints. For instance, Cabe and Herriges (1992) consider the unconstrained social planner problem of NPS pollution control, incorporating uncertainty and information acquisition, where information reduces the social cost of setting a control mechanism through an ambient tax, as proposed by Segerson (1988). Elsewhere, Xepapedeas (1995) examines the unconstrained social planner's use of an effluent tax, in conjunction with an ambient tax, as an incentive for individual polluters to reveal information useful in uncovering the connections between generator and ambient emissions. In the specific case of sediment loading, the budget-constrained manager cannot impose fees against nature to learn more about each source's contribution to the total ambient load. The manager must expend limited resources to obtain the information that otherwise could have been obtained by the social planner through the use of effluent fees. Furthermore, the optimal decision

[1] The primary focus of erosion control in Redwood National Park is preventing or reducing erosion from logging roads within the Park (DOI 1981). It is well known that logging roads are the main contributors to sediment loading (Mount 1995; GAO 1999; EPA 1999).

[2] The European Environmental Agency (2000, 2001) and the US-GAO (1999) report that a lack of financial resources limits the ability of public environmental managers to achieve their core objectives.

rules derived in a social welfare framework differ from those derived from the constrained management approach and lead to different policy prescriptions (Barrett and Segerson 1997).

Garvie and Keeler (1994) consider a fiscally constrained regulator who minimizes non-compliant pollution generation by allocating a limited budget between data collection and enforcement. Data collection provides evidence necessary to prosecute, and, as such, improves enforcement effectiveness. Our focus is on the role of acquired information in reducing pollution loading uncertainty, which improves abatement effectiveness. We also evaluate the minimization of expected cost of environmental damages whereas Garvie and Keeler consider minimizing pollution irrespective of the related damage. Unless the damage function is linear the result will vary between these two objectives.

The rest of the paper is structured as follows. Section 2 develops a model to analyze sediment control when information is incomplete and data collection is costly. Section 3 provides the comparative static results for key parameters of the model. Section 4 presents the results of simulating the model for a case resembling the public management of sediment loading in Redwood National Park, located in northwestern California. There, we compare the allocation decisions of a perfectly informed manager, a completely uninformed manager, and an imperfectly informed manager who acquires information through data collection. Section 6 concludes.

2. *A Model of NPS Pollution Control with Costly Information Acquisition*

Faced with budget and information constraints, the manager chooses between the level of abatement effort at each source and data collection, where the information acquired through data collection reduces pollution loading uncertainty. We model the decision to acquire information and abate as a sequential problem since each activity occurs during separate periods throughout the year. For the Redwood National Park case, data (stream flow and ambient sediment load measures) are collected during the rain season between October and April. The abatement projects begin in late summer and end before the rain season begins again. These decisions are linked by a single budget, which is allocated over both periods.[3] There is no discounting of the budget given the short duration of time that expires between allocating expenditures for information acquisition and abatement effort. In this formulation, we look directly at the tradeoff the manager faces between abatement effort and abatement effectiveness, where the latter depends on information about pollution loading across the sources.

The information acquisition or sequential updating process requires the manager to make a decision ex ante (prior to the realization) on the data collection frequency.[4] This ex ante decision is made using a prior expectation on the information content of a given data collection frequency. The expected information content is simply the expected reduction in uncertainty about the pollution generated by each source. [5]

Initially, the manager chooses the frequency of data collection on total sediment loading and stream flow and updates the prior subjective sediment loading distribution for each of the sources. In practice, the manager, prior to the beginning of the rain season, determines a fixed number of daily samples to collect at each data collection point throughout the rain season. We

---

[3] The separability of the information acquisition and abatement effort decisions is not unique to the example of sediment loading in Redwood National Park. Take for example, self-reporting provisions for the USDA EQIP policy (Cattaneo, 2001). Here, the government solicits information from agricultural producers on farming practices that conveys the environmental benefits derived from implementing such practices. This information acquisition cannot occur simultaneously with the decision to allocate resources toward environmental benefits.

[4] Sequential updating typically results in "sub-optimal" decision-making when observed ex post. That is, if we knew yesterday what we know today, then the "optimal" decision we made yesterday could have been improved upon.

[5] The model presented here examines a single *round* of decision-making process in a two-period sequence. Although information acquisition and learning are dynamic phenomena, the two-period model we present can capture the inter-temporal decision-making process by repeating the two-period model. The single *round* of decision making demonstrates the theoretical underpinnings of the problem.

denote the frequency of data collection by δ. An increase in the frequency of data collection implies that a greater number of samples are collected each day throughout the rain season. Since the manager is constrained by a fixed budget, *B*, the maximum number of possible data collection frequencies δ_Ω is equal to *B*/*m*, where *m* is the per frequency cost of data collection. This constraint $(\delta \leq B/m)$ never binds since damages cannot be controlled without abatement at some sources. Note, however, that the reverse is not true.

Next, having collected data, abatement effort levels $X = (x_1, x_2, \cdots, x_N)$ for the *N* polluting sources are determined so as to minimize the expected cost of environmental damage, given the updated posterior sediment loading distribution. The damage cost function $D(Q)$ is twice continuously differentiable, increasing and convex in *Q*, the ambient pollution (i.e., $\frac{\partial D}{\partial Q} > 0$ and $\frac{\partial^2 D}{\partial Q^2} > 0$).

Let, q_n be the unobservable tonnage of pollution loading from the *n*th source, where $n = 1, 2, \cdots, N$. We define $q_n = q_n(x_n; w_n, \alpha_n(\theta_n))$ as a function of abatement effort, stochastic rainfall (w_n) and site-specific characteristics (α_n) that define the relationship between abatement and rainfall on pollution loading at that source.[6] Following Shortle and Alber (1997), uncertainty is introduced into the problem by taking these site-specific characteristics to be uncertain. The incomplete information about α_n, is explicitly incorporated into the definition of sediment loading by allowing site-specific characteristics to depend on θ_n, where low values of θ_n are less certain than high values. This incomplete information does not directly affect the sediment

---

[6] This characterization of pollution loading is similar to prior models of stochastic nonpoint source pollution control (see Beavis and Walker; Shortle and Dunn; Shortle; and Horan, Shortle and Abler among others). However, in our interpretation of the model θ_n is the manager's information or knowledge about site-specific characteristics for the nth source rather than the nth firm's private knowledge as depicted in Shortle and Alber (1997).

loading put does affect the marginal productivity of abatement effort (i.e., the marginal abatement effectiveness). The pollution loading function has the following properties $\frac{\partial q_n}{\partial x_n} < 0, \frac{\partial^2 q_n}{\partial x_n^2} > 0, \frac{\partial q_n}{\partial w_n} > 0, \frac{\partial q_n}{\partial \alpha_n} \gtreqless 0$ and $\frac{\partial \theta_n}{\partial \delta} > 0$. The observable ambient pollution load, Q, is defined such that $Q \equiv \sum_n q_n$.[7] When decisions are made on allocating resources to abatement effort, the manager uses $\pi_n(w_n | \theta_n)$ the post-data or posterior conditional distribution for stochastic rainfall given the uncertainty about pollution loading. We shall fully incorporate data collection and information acquisition into the management model shortly. For now, assume knowledge is fixed so that we can derive the manager's optimal abatement decision.

The total abatement cost expenditure, C, is defined as $C = \sum_n c_n x_n$. Recall that the per frequency cost of data collection, *m*, is also assumed to be constant so that $M = \delta m$ represents the total data collection expenditure. These linear cost specifications allow us to focus attention on the tradeoff between abatement effort and abatement effectiveness. The budget constraint is

$$\sum_n c_n x_n + \delta m \le B \qquad (1)$$

Given a fixed level of data collection, and subject to equation (1), the manager chooses abatement effort across the sources to

$$\underset{x_1, x_2, \cdots, x_N}{Minimize}\, ED(Q) \qquad (2)$$

where *E* is the expectation operator for the posterior conditional distribution for stochastic rainfall. The first order conditions for an interior abatement effort allocation are (1) and

[7] We assume a linear specification since interactions between sources are negligible in our example. However, this simplification cannot be maintained as a rule (Lintner and Weersink, 1999).

$$E\left[\frac{\partial D}{\partial Q}\frac{\partial q_n}{\partial x_n}\right]+\lambda c_n=0, \forall n=1,2,..,N \tag{3}$$

where λ is a Lagrangean multiplier (the shadow price of the budgeted resources).[8] From (1) and (3), we obtain the optimal abatement effort allocations $x_n=\tilde{x}_n(c_1,c_2,\cdots,c_N,m,\pi_n(w_n,\alpha_n(\theta_n)))$, where $\tilde{x}$ maps the parameters of the model into the optimal abatement effort.

Conditions (3) simply state that, at the optimum, the manager chooses abatement effort at each source such that the expected marginal reduction in the cost of damages (*i.e.*, the expected marginal benefit from abatement effort) is equal to the marginal cost of abatement effort. We can rewrite (3) as

$$\frac{E\left[\frac{\partial D}{\partial Q}\frac{\partial q_n}{\partial x_n}\right]}{E\left[\frac{\partial D}{\partial Q}\frac{\partial q_j}{\partial x_j}\right]}=\frac{c_n}{c_j}, \forall n\neq j=1,2,..,N \tag{4}$$

or

$$\frac{E\left[\frac{\partial D}{\partial Q}\right]E\left[\frac{\partial q_n}{\partial x_n}\right]+\mathrm{cov}\left[\frac{\partial D}{\partial Q},\frac{\partial q_n}{\partial x_n}\right]}{E\left[\frac{\partial D}{\partial Q}\right]E\left[\frac{\partial q_j}{\partial x_j}\right]+\mathrm{cov}\left[\frac{\partial D}{\partial Q},\frac{\partial q_j}{\partial x_j}\right]}=\frac{c_n}{c_j}, \forall n\neq j=1,2,..,N \tag{4'}$$

which recasts the optimality condition in the form of the familiar requirement that the expected marginal rates of transformation across any two sources should equal the relative marginal abatement costs (the point at which the budget constraint and the iso-expected damage curve are tangent). Equation 4’ illustrates that the optimality condition with nonlinear damage costs differs

[8] Given desirable curvature properties of the damage cost function, we assume the second order condition holds at the minimum.

from the linear damage function and the solution to minimizing an environmental goal such as pollution loading, where the optimality condition simplifies to

$$\frac{E\left[\frac{\partial q_n}{\partial x_n}\right]}{E\left[\frac{\partial q_j}{\partial x_j}\right]} = \frac{c_n}{c_j}, \forall n \neq j = 1,2,..,N \tag{5}$$

Before formally deriving the optimization problem when information is acquired through data collection, let us examine the mechanism by which abatement effort allocations are affected by data collection and acquired information. If, for example, source *n*'s actual contribution to the ambient load is greater than expected and *j*'s contribution is less than expected, given abatement effort and rainfall, then data collection changes $\pi_n(w_n|\theta_n)$ such that for a given abatement effort the expected loading from the *nth* source increases while it decreases at the *j*th source. In this example sources *n* and *j* were selected arbitrarily from among the *N* sources but in actuality, loadings from any two sources may be greater than or less than expected loadings, a priori. However, since the total ambient load is fixed for a given abatement allocation and rainfall event, it must be the case that if one source's load is greater than previously expected then at least one other source's load must be less than previously expected. We have simply characterized this example. A reexamination of equation 4' reveals that data collection increases the numerator and lowers the denominator on the LHS, so that the LHS of (4') rises. To restore the equilibrium associated with a larger value of δ, we must have $\frac{dx_n}{d\delta} > 0$ and $\frac{dx_j}{d\delta} < 0$ (these conditions follow from the assumption $\frac{\partial^2 q_n}{\partial x_n^2} > 0$). So, there will be a reallocation of the abatement efforts from source *j* to source *n*. But this is for an unchanged abatement budget. Since an increase in δ reduces the budget for abatement activity, one has to consider the net effect of a change in δ by differentiating both (1) and (3) with respect to δ.

To formally model information acquisition, we assume the information acquired through data collection allows the manager to update $\hat{\pi}_n, \forall n = 1,2,\cdots,N$, the prior subjective probability distribution for pollution loading at each source and derive $\pi_n(x_n|\theta_n(\delta;\phi);\hat{\pi}_\mathrm{n})$, the posterior probability distributions for pollution loading at each source that is closer to the true underlying distribution. In this context, "closer" refers to the notion that the information content of the posterior distribution is closer to the information contained in the true, yet unknown, distribution. The parameter ϕ reflects the efficiency of information acquisition. When data is collected, the rate at which the expected pollution loading at each source is updated toward the true underlying loading values increases as ϕ increases. In essence, ϕ represents the extent of the manager's skill and ability in utilizing the collected data to extract information. This notion of ability is in keeping with Arrow (1974, pp. 37) who states that each individual has the ability to receive a signal from natural and social environments. However, it is the limited capacity and scarcity of information-handling ability that sets individuals apart. Learning capacity may be enhanced through exogenous means such as technological advances in data collection or education programs and thus we are interested in how changes in capacity may affect the optimal data collection strategy.

Returning to the pre-data collection problem, the manager chooses the level of data collection that minimizes the expected damage cost from pollution loading. Substituting the optimal abatement functions ($\tilde{x}$) derived above into (1) and (2), the ex ante optimization problem is formally written as

$$\underset{\delta}{Minimize}\, ED\left[\sum_n q_n(\tilde{x}_n, w_n, \alpha(\theta_n(\delta;\phi)))\right] \tag{6}$$

s.t. $$\sum_n c_n \tilde{x}_n + \delta m \leq B \tag{7}$$

It should be noted again that implicit in Equation (6) is the assumption that the manager has an expectation about how data collection affects her subjective probability distributions about pollution loading, and uses this prior knowledge to choose the data collection frequency that minimize the expected cost of pollution related damage.

The first order conditions for the optimal level of data collection are (7) and

$$E\left[\frac{\partial D}{\partial Q}\sum_n \frac{\partial q_n}{\partial \tilde{x}_n}\frac{\partial \tilde{x}_n}{\partial \theta_n}\frac{\partial \theta_n}{\partial \delta}\right] = \eta\left(m + \sum_n c_n \frac{\partial \tilde{x}_n}{\partial \theta_n}\frac{\partial \theta_n}{\partial \delta}\right) \tag{8}$$

where η is the Lagrangean multiplier on the ex ante allocation of the pollution control budget. Noting that information acquisition is akin to provision of a collective good, condition (8) has a straightforward interpretation. That is, the manager optimally allocates resources to data collection so that the expected marginal reduction in the cost of damages over all sources (the LHS of (8)) equals the total marginal opportunity cost of acquiring information (the RHS of (8)). The total marginal opportunity cost of acquiring information consists of a direct cost of an additional unit of information, *m*, and an indirect cost (or benefit) given by the expression $\sum_n c_n \frac{\partial \tilde{x}_n}{\partial \theta_n}\frac{\partial \theta_n}{\partial \delta}$ which reflects the effect of an additional unit of information on the abatement expenditure by causing a reallocation of abatement effort among the various sources. Solving equation (8) yields $\delta = \hat{\delta}(m, \phi, B)$, the optimal allocation of data collection, where $\hat{\delta}$ maps the parameters into the optimal data collection allocation. Now substituting this optimal allocation of data collection into the optimal abatement allocation yields

$$x_n = \hat{x}_n(\hat{\delta}(m,\phi,B), c_n, c_j, m, B, \pi_n(w_n, \theta_n(\hat{\delta}(m,\phi,B));\phi))$$

where $\hat{x}$ is the optimal abatement effort that maps the parameters of the data collection problem into the optimal level of abatement effort.

3. *Comparative Static Analysis*

The comparative static results for the costly information acquisition model are derived in the standard manner. Because of the sequential nature of this problem, the timing of any cost change determines the relevant comparative static result. For example, if a change in the state of abatement costs occurs after data collection, then the manager can only change the optimal allocation of abatement across the sources.

In this section we first consider the effect of a change in abatement cost (c_n) on the optimal abatement levels. We then consider the effect of a change in the data collection cost (*m*) and the information efficiency (ϕ) on the optimal data collection frequency ($\hat{\delta}$). Note that $\hat{\delta}$ is unaffected by c_n as data collection precedes abatement activity.

The sign of the comparative static $\frac{\partial \tilde{x}_j}{\partial c_n}$ is ambiguous and depends on how abatement expenditure at source *n* changes in response to a change in c_n. The marginal cost of abatement is likely to experience favorable change with advances in technology and unfavorable change due to such events as landslides within the abatement area, or regulation on abatement effort to preserve endangered species for instance. The change in abatement expenditure at source *n* depends on the magnitude of the own cost elasticity of abatement effort ($\varepsilon_{x_n,c_n} = \frac{c_n}{x_n}\frac{\partial \tilde{x}_n}{\partial c_n}$).

Differentiating the budget constraint (7) with respect to a change in the per unit abatement cost at source *n* (c_n), we derive:

$$\operatorname{sgn}\left(\frac{\partial \tilde{x}_j}{\partial c_n}\right) = \operatorname{sgn}(\varepsilon_{x_n,c_n} - 1) \tag{9}$$

Proposition I: If $\varepsilon_{c_n,x_n} < 1$*, an* *<u>increase</u>* *in the per unit abatement cost at source n will result in an* *<u>increase</u>* *in abatement expenditure at that source and hence a* *<u>decrease</u>* *in the abatement expenditures at the remaining sources and vice versa.*

The intuition for Proposition I is straightforward and needs no further explanation. This response to a change in the per unit abatement cost at one source may occur at one other source or be distributed across multiple sources. Proposition I implies that regulation aimed at protecting a species in one area can, by raising the cost of pollution abatement in that area, also negatively affect control efforts in other areas within the watershed, with the unintended consequence of raising expected damage costs from pollution.

To capture the trade-off between data collection and abatement effort we turn to the comparative static result for a change in the data collection cost, *m*. The net effect on δ from a change in *m* can be decomposed into two separate effects. An increase in *m* causes a reduction in δ ex ante. In turn, this reduction in δ generates a chain of secondary effects by changing the expected sediment loading for given abatement effort levels at each source. Since this secondary effect occurs before abatement efforts are actually chosen the abatement effort levels can change to exploit the gains in abatement effectiveness generated by the increase in information. Changes in abatement effort in turn changes the abatement expenditure, and hence the resources available for information acquisition, and therefore δ. This chain of effects is depicted below, where Δ denotes change in a variable:

$$\begin{array}{lcr} \overset{(+)}{\Delta m} \to \overset{(-)}{\Delta \delta} \Big| (x_1, x_2) : \Delta[E(q_1), E(q_2)] \to \Delta(x_1, x_2) & & \\ \qquad \uparrow \qquad\qquad\qquad\qquad\qquad\qquad\quad \downarrow & & \\ \delta\Delta m + m\Delta\delta = \Delta(\delta m) = \qquad\qquad -\Delta(c_1 x_1 + c_2 x_2) & & \end{array}$$

We formally express this result by

$$\operatorname{sgn}\left(\frac{\partial \hat{\delta}}{\partial m}\right) = -\operatorname{sgn}(\delta + \sum_n c_n \frac{\partial \hat{x}_n}{\partial m})(m + \sum_n c_n \frac{\partial \hat{x}_n}{\partial \theta_n} \frac{\partial \theta_n}{\partial \delta}) \tag{10}$$

The sign of this comparative static derivative is ambiguous. The second parenthesized term on the right-hand side (RHS) of equation (10) is the marginal cost of data collection, which is positive from the first order condition (8). The sign of the first parenthesized term on the RHS of (10), which we term the "sequential effect," is however ambiguous. The sequential effect represents the tradeoff between higher data expenditure, through the change in data collection level δ, and lower abatement expenditure, through changes in the abatement effort levels, x_n's. In general, the effect of the change in data collection on each individual source's abatement level is ambiguous, but total abatement expenditure increases (decreases) with an increase in data collection cost, *m*, if period 1 data collection is cost elastic (inelastic). Thus, from (10) we can state the following proposition:

Proposition II: If data collection is sufficiently cost inelastic, so that with a higher cost of data collection (m), the <u>decrease</u> in abatement expenditure is larger than the <u>increase</u> in data collection expenditure, then the sequential effect is <u>negative</u>, thus inducing an <u>increase</u> in the level of data collection.

At first, this result seems counterintuitive. We would expect less data collection when the cost of data collection increases. But, because of the sequential effect of data collection, in the form of a more efficient abatement allocation, there will be efficiency gains in the form of net savings on abatement costs, which allow the manager to increase data collection efforts ex ante. To better appreciate the result stated in Proposition II, we should bear in mind that the change in abatement levels at various sources when data collection changes in response to a change in its cost (m), depends on the size and direction of changes respectively in the abatement budget and in abatement productivity effects. The budget effect is a decline in abatement levels at all sources because the resources spent on data collection must be taken from the same given budget. In other words, resources spent on data collection are unavailable for abatement. The abatement productivity effect is a decline in the productivity at sources with lower posterior *expected* sediment loading and a rise in the productivity at sources with higher posterior *expected* sediment loading. Thus, Proposition II is more likely to hold when, assuming linear costs, (*i*) the posterior distribution is highly sensitive to data collection (*i.e.*, farther away is the prior distribution from the underlying true distribution), and (*ii*) the marginal abatement productivity is very sensitive to sediment loading. Proposition II has an important policy implication. It cautions us, for example, that a policy subsidizing data collection to reduce the uncertainty about pollution flows from various sources, and thereby enhancing the efficiency of abatement programs, may lead to the opposite result by shifting resources away from data collection to more abatement activity.

We have also examined the effect of a change in the productivity of information acquisition (ϕ) on data collection. This is given by

$$\mathrm{sgn}\left(\frac{\partial\hat{\delta}}{\partial\phi}\right) = -\mathrm{sgn}(\sum_n c_n \frac{\partial\hat{x}_n}{\partial\theta_n}\frac{\partial\theta_n}{\partial\phi})(m + \sum_n c_n \frac{\partial\hat{x}_n}{\partial\theta_n}\frac{\partial\theta_n}{\partial\delta}) \tag{11}$$

We again obtain an ambiguous result. We refer to the first term on the RHS of (11) as the "information efficiency effect" since it reflects the change in abatement expenditure resulting from a change in the efficiency of information acquisition, given the frequency of data collection (δ). When the efficiency of information acquisition increases, the curvature conditions on the abatement functions ensure that, over an interval of abatement levels (x_n, x_j), the abatement expenditure at the *nth* source increases while it decreases at the *jth* source. Recall that we have assumed that the sediment loading from the nth source is greater than expected ex ante and vice versa for the jth source. This leads to the following proposition, which highlights the trade-off between the level and efficiency of information acquisition.

Proposition III: If a higher information acquisition efficiency (ϕ) raises the abatement expenditure at source n by less than it reduces the abatement expenditure at source j, then the information efficiency effect is <u>negative</u>, thus inducing an <u>increase</u> in the level of data collection.

One might normally expect, given diminishing marginal productivity of information, a manager who is more efficient in extracting information from collected data would take advantage of that skill and, everything else equal, opt for less data collection. Proposition III indicates the condition under which the opposite occurs. In such cases there is a tradeoff between information acquisition efficiency and the intensity of data collection, with a possible consequence of shifting resources from abatement activity to data collection. A policy implication of Proposition III is that a program aiming to improve pollution managers' knowledge and skills in information acquisition may come at the cost of reduced abatement activity.

4. *Model Simulation: Heterogeneity of Polluting Sources and the Value of Information*

To provide a numerical illustration, we simulate a simplified pollution control model based on Redwood National Park's sediment control program for Redwood Creek. Overall three separate models are simulated. Model I presents the case of a perfectly informed manager (PI), Model II is that of an uninformed manager (UI) who is assumed to believe that the sediment loading is *uniform* across all polluting sources, and Model III is the case of a data-collecting, imperfectly informed manager. These simulations highlight the value of information and the economic tradeoff between abatement effort and abatement effectiveness when information is acquired.

To evaluate the value of information we compare the optimal abatement decision in model I and II over a range of scenarios about the heterogeneity of sediment loading across sources. In each model two sources generate sediment loading. For each heterogeneity scenario, the damage costs attributable to the hypothetical assumption of a prior uniform distribution on sediment loading is compared with the damage costs under a perfect information assumption in order to derive the value of perfect information.

The second set of simulations evaluates Model III to determine the optimal level of data collection and abatement and reveals the economic tradeoff between abatement effort and abatement effectiveness when information is acquired. In these latter simulations we solve the model for the optimal data collection level that minimizes expected cost of damages and illustrate the response of optimal data collection to a change in the cost of data collection and the information acquisition efficiency. Given the assumption of diminishing marginal returns to information, complete resolution of uncertainty about the sediment loading generation by source may entail an infinite amount of data collection, which is prohibitively costly. Thus with data collection, it must be the case that damages are less than those resulting from an uninformed manager's abatement decision but greater than those arising from the decision of a perfectly informed manager.

To facilitate the simulation, the model is calibrated using estimates derived in Kaplan (1999). First, the damage cost function is

$$\ln(\text{Damage}(\$)) = -5.8667 + 1.674\ln(\text{Q}) \tag{12}$$

The functional form for sediment loading is adapted from a known physical relationship relating ambient sediment loading with stream flow such that[9]

$$\text{Q} = \sum_n q_n \equiv \sum_n \bar{q}_n = \sum_n \left(flw_n\right)^{s_n} \tag{13}$$

where (flw_n) is the average stream flow measure generated from within each of the polluting sources, which is a proxy for stochastic rainfall, and s_n is the sediment loading parameter for the nth source. For illustrative purposes and based on empirical evidence from Kaplan, Howitt and Farzin (2003), we define the relationship between the sediment loading parameter and abatement effort as follows: $s_n = 1 - 0.0015x_n$. To construct the underlying stream flow values for each source for each heterogeneity scenario, we applied the following formulas:

$flw_1 = (1+0.1*h)*140{,}768$, and $flw_2 = (1-0.1*h)*140{,}768$

where $h = 0,1,\ldots,9$, and given that the average stream flow for Redwood Creek is approximately 140,768 cubic yards. For computational convenience and without altering the qualitative results, we use h to be the scenario number. We vary the heterogeneity of stream flow from each source to reflect a few of the possible sediment loading distributions that nature imposes on the system.

In constructing q_n, the unobservable sediment loading at each source and the ambient sediment load Q we substitute the source specific sediment loading parameter and stream flow equations into (13).

Next, the abatement cost function was obtained from Kaplan (1999), where the abatement cost function for each source is estimated as a linear function of abatement level $C_n = c_n x_n + e$, and abatement is measured as the number of haul roads removed. To focus attention on the role of data collection in increasing the efficiency of abatement effort and thus the trade-off between data collection and abatement expenditures we assume the abatement cost coefficient is identical across sources and estimated the coefficient with least squares without an intercept term for obvious reasons. The value of the estimated cost coefficient is 2178.1 with t-statistic of 13.68 and an R^2 of 0.34. The annual budget is fixed at $200,000, which is the average annual abatement budget for removing haul roads in Redwood National Park.

Table 1 presents the optimal abatement levels for the perfectly informed manager under the various heterogeneity scenarios. In this case the manager knows q_n, the actual sediment loading from each source. To derive the optimal level of abatement under perfect information we minimize (12) subject to (13) and the abatement cost function defined above. For the *uninformed* manager (UI) who assumes a prior uniform distribution over all heterogeneity scenarios, the optimal abatement level for each source, for all scenarios is (x_1 = 57.7, x_2= 57.7). This corresponds to the optimal abatement levels chosen by the perfectly informed manager if the polluting sources were in fact homogeneous. This is because, given all the same information on costs and damages except the true heterogeneity of the sources, the manager will make the same decision when the true sediment loading distribution is uniform.

---

[9] This functional relationship can be found in Kaplan and Howitt (2002), and Singh and Krstanovic (1987).

Table 1. Optimal Abatement under Perfect Information

Scenario	x_1	x_2
0	57.0	57.0
1	62.6	51.3
2	68.4	45.6
3	74.4	39.5
4	81.0	33.0
5	88.3	25.7
6	96.7	17.2
7	107.3	6.7
8	114.0	0.0
9	114.0	0.0

Table 2 presents the resulting damage costs under PI and UI cases. Column 2 in Table 2 shows the damage cost corresponding to the optimal abatement levels when the manager has perfect information about the distribution of ambient load across sources. Column 3 shows the damage cost resulting from an uninformed budget allocation. In both cases the reduction in damage costs when abatement is undertaken, compared to the no abatement case, exceeds the $200,000 spent to control sediment loading. The greater damage cost for the uninformed model is a result of the manager's lack of information with respect to true sediment loading. When the manager has perfect information less damage costs results because the manager exploits the knowledge about the degree of heterogeneity to allocate the budget more efficiently. This case of a perfectly informed manager is analogous to point source pollution control since there is no uncertainty about the pollution generated from each source. Column 4 shows the value of perfect information (VPI) for each scenario as the difference between damage costs under UI and PI. These costs are respectively the upper and lower bounds for possible damage costs when information is optimally acquired. Furthermore, the last column of Table 2 shows that the marginal value of perfect information increases as the difference between the true distribution and the prior (uniform) distribution grows, *i.e.*, as heterogeneity rises, except in the last two scenarios due to the corner solution result.[10]

Table 2. Damages for Model I and II

Scenario	PI	UI	VPI=UI-PI	Marginal VPI
0	657,615	657,615	0	0
1	651,702	657,185	5,483	5,483
2	633,904	655,886	21,982	16,499
3	604,037	653,689	49,652	27,670
4	561,775	650,541	88,765	39,114
5	506,614	646,356	139,742	50,977
6	437,803	641,000	203,196	63,454
7	354,233	634,249	280,016	76,820
8	257,926	625,700	367,774	87,758
9	171,247	614,452	443,205	75,431

In the second set of simulations, where costly information acquisition is evaluated, we limit the heterogeneity to scenario 9; that is, we assume that this scenario represents the true sediment loading heterogeneity at the two sources. Selection of any other scenario does not alter the results. In this set of simulations, we stipulate different values for *m*, the data collection cost, and for ϕ, the information efficiency parameter (see Table 3 for the assumed values). Any data collection expenditure also comes from the initial budget of \$200,000 thereby reducing the resources available for abatement effort. To incorporate information acquisition into the manager's objective function, we substitute the sediment loading parameter in equation (13) with the expected sediment loading parameter function $E(s_n) = 1-(0.001+0.0005A(\delta))x_n$, where $A(\delta)=1-(1-\phi)^{\delta}$, $0 \le A(\delta) < 1$ and $0 < \phi < 1$. Absent any prior functional forms for $A(\delta)$ we constructed this correspondence from the desired curvature properties (i.e.,

[10] When the pollution load reaches higher levels of heterogeneity the abatement budget is allocated to abatement at one source only. Above this level of heterogeneity the marginal returns to perfect information decline because the manager, can no longer reallocate abatement effort and thus cannot take advantage of the “better” information.

$\frac{\partial A(\delta)}{\partial \delta} > 0, \frac{\partial^2 A(\delta)}{\partial \delta^2} < 0, \frac{\partial A(\delta)}{\partial \phi} > 0$ and $\frac{\partial^2 A(\delta)}{\partial \phi^2} < 0$). This new expected sediment loading parameter equation asymptotically approaches certainty (i.e., $A(\delta) \to 1$ as $\delta \to \infty$). This expected sediment loading parameter encapsulates both the abatement effort and the uncertainty about the relationship between stream flow and abatement effort as discussed in the theoretical section.

Table 3 presents the optimal data collection, and expected and actual damages for the low and high (ϕ=0.6, 0.8) information efficiency scenarios. The divergence between the expected and actual damage, reported in Table 3, is a result of the manager not having perfect information when choosing abatement expenditures across sources. If the manager were perfectly informed, then the expected and actual damages would coincide. As Column 4 and 7 show, when it is optimal to acquire information, the actual damage lies in between the extreme cases of UI and PI abatement, where actual damages are \$614,452 and \$171,247 respectively (see scenario 9 in Table 2). These results show the important role information acquisition can play in improving the budget allocation and hence reducing the expected damage when compared with the case of the ex ante, uniform prior distribution. When, under heterogeneity scenario 9, data is optimally collected, the actual damage cost is always *lower* than the actual damage (D = \$614,452) that would result from the allocation of the budget *only to abatement efforts* ($\delta = 0$) under the uninformed management. It is should be noted that this reduction in damage costs understates the actual benefits of information acquisition since, in practice, the benefit of information spills over to a much longer time period, a consideration not accounted for in this example.

Comparing the optimal levels of data collection (δ) under the two information acquisition efficiency scenarios, we see that an increase in the information acquisition efficiency has an ambiguous effect, which is consistent with *Proposition III*. In particular, we see that when m = \$90,000, an increase in ϕ from 0.6 to 0.8 leads to a higher level of data collection (δ increases from 0 to 1). This can be explained intuitively by noting that the effect of higher information efficiency is like lowering the cost of information. Coupled with a high marginal return on the first unit of data collection, this renders it optimal to collect data. On the other hand, when cost of data collection is as low as m = \$15,000, so that it is optimal to collect information with relatively high frequency (δ = 3 for ϕ = 0.6), the marginal return on information acquisition is relatively low. Together with high efficiency of information acquisition, this makes it optimal to reduce the frequency of data collection to δ = 2 when ϕ = 0.8. Reducing data collection to δ = 2 allows the manager to spend more on abatement effort where the marginal return (in terms of lowering the expected damage cost) is relatively high.

Table 3. Optimal Data Collection, and Expected and Actual Damages (\$)

	ϕ=0.6			ϕ=0.8		
m	δ	E(Damage)	Damage	δ	E(Damage)	Damage
\$90,000	0	\$657,615	\$614,453	1	\$842,576	\$599,297
\$50,000	1	\$816,532	\$332,650	1	\$515,209	\$332,650
\$15,000	3	\$356,754	\$310,063	2	\$276,776	\$252,335

5. *Concluding Remarks*

This paper has examined the problem of NPS pollution control under incomplete and costly information. We have analyzed the problem within a constrained management framework to bring to light a more realistic setting for studying NPS pollution control. The comparative static results showed the conditions under which (*i*) the manager may lower abatement efforts at *all* sources when an unfavorable change (*e.g.*, stricter environmental regulations or adverse natural events) causes the abatement cost to go up at one specific source, and vice versa (Proposition I), (*ii*) data collection effort may *increase* despite a rise in data collection cost (Proposition II), and (*iii*) a higher information efficiency can lead to *less* data collection (Proposition III). The model simulation results showed that by exploiting the knowledge of sediment loading heterogeneity across the polluting sources, the manager can improve the overall efficiency of budget allocation to abatement efforts and thereby further reduce pollution damages.

Of many possible extensions of the present study, we believe that studying the problem in a dynamic setting can be particularly insightful. Over a finite time horizon, the manager chooses investment paths for both information acquisition and abatement efforts. During this time horizon, several factors will influence the dynamics of each path. Principal among these factors is the decline of the productivity of information acquisition as uncertainty about the degree of heterogeneity is reduced. This suggests that the manager may find it optimal to decrease data collection and information acquisition over time. Secondly, as abatement at the source with the largest sediment load occurs early in the time horizon, the system will become increasingly less heterogeneous. With a decreasing heterogeneity of sediment loading over time, we expect that at some future date the abatement policy will change from a heterogeneous abatement strategy to a homogeneous one. Future research should shed light on these issues.

References

Arrow, K.J., *The Limits of Organization*. New York: W.W. Norton and Co., 1974.

Babcock, B.A., Lakshminarayan, P.G., Wu, J., and D. Zilberman (1996), "The Economics of Public Fund for Environmental Amenities: A Study of CRP Contracts," *American Journal of Agricultural Economics* **78**: 961-971.

_____ (1997), "Targeting Tools for the Purchase of Environmental Amenities," *Land Economics* **73(3)**: 325-339.

Barrett, J., and K. Segerson (1997), "Prevention and Treatment in Environmental Policy Design", *Journal of Environmental Economics and Management* **33**: 196-213.

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

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

Cattaneo, A. 2001, "EQIP: Conserving While Farming," *Agricultural Outlook* Spetember/AO-284, 26-27.

European Environmental Agency (2000), *EEA Annual Work Programme 2000*, European Environmental Agency, Doc EEA/053/99final.

European Environmental Agency (2001), "Chapter 2. Current Environmental Policy" in *The State of Action to Protect the Environment in Europe: Expert Corner Report. No 1*, http:themes.eea.eu.int/showpage.php/improvement/policy?pg=37486

Garvie, D., and A. Keeler (1994), "Incomplete Enforcement with Endogenous Regulatory Choice", *Journal of Public Economics* **55**: 141-62.

Horan, R.D., Shortle, J.S., and D.G. Abler (1998), "Ambient Taxes When Polluters Have Multiple Choices," *Journal of Environmental Economics and Management* **36**: 186-199.

Kaplan, J.D. (1999), "Nonpoint Source Pollution, Incomplete Information and Learning: An Entropy Approach," Ph.D. Dissertation, Department of Agricultural and Resource Economics, University of California at Davis.

Kaplan, J.D. and R.E. Howitt (2002) "Estimating Nonpoint Source Pollution: An Application of a Sequential Entropy Filter," *Water Resources Research* **38**(3): 10.1029/2000WR000088.

Kaplan, J.D., Howitt, R.E. and Y.H. Farzin (2003), "An Information-Theoretical Analysis of Budget-Constrained Nonpoint Source Pollution Control," *Journal of Environmental Economics and Management* **46**: 106-130.

Lintner, A.M. and A. Weersink (1999), “Endogenous Transport Coefficients: Implications for Improving Water Quality from Multi-Contaminants,” Environmental and Resource Economics **14**: 269-296.

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

Shortle, J.S., and D.G. Alber (1997), “Nonpoint Pollution,” in *The International Yearbook of Environmental and Resource Economics 1997/1998* (H. Folmer and T. Tietenberg, Eds.), Edward Elger, Cheltenham, UK, and Northhampton, MA, USA; pp114-155.

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

Shortle, J.S. (1990), “The Allocative Efficiency Implications of Water Pollution Abatement Cost Comparisons,” *Water Resources Research* **26(5)**: 793-797.

Singh, V. P. and P.F. Krstanovic (1987), "A Stochastic Model for Sediment Yield Using the Principle of Maximum Entropy." *Water Resource Research* **23**:5781-793.

Spreiter, T.A., Franke, J.F. and D.L. Steensen (1995), 'Disturbed Lands Restoration: The Redwood experience. Paper Presentation at the Eight Biennial Conference on Research and Resource Management in Parks and Public Lands, April 17-21, 1995, George Wright Society, Hancock, Michigan.

United States General Accounting Office (1999), *Water Quality: Federal Role in Addressing-and-Contributing to-Nonpoint Source Pollution*, GAO/RCED-99-45, US Government Press, Washington DC.

United States Department of the Interior (1981), *Watershed Rehabilitation Plan: Redwood National Park*, United States National Park Service, April, US Government Press, Washington DC.

Xepapedeas, A.P. (1995), “Observability and Choice of Instrument Mix in the Control of Externalities”, *Journal of Public Economics* **56**:485-498.