Using Numerical Dynamic Programming to Compare Passive and Active Learning in the Adaptive Management of Nutrients in Shallow Lakes

Craig A. Bond*
Assistant Professor
Department of Agricultural and Resource Economics
Colorado State University
Fort Collins, CO 80523
Ph: (970) 491-6159
Fax: (970) 491-2067
craig.bond@colostate.edu

John B. Loomis
Professor
Department of Agricultural and Resource Economics
Colorado State University
Fort Collins, CO 80523
Ph: (970) 491-2485
Fax: (970) 491-2067
john.loomis@colostate.edu

DARE CSU Working Paper 003-08
December 09, 2008

Abstract: This paper illustrates the use of dual/adaptive control methods to compare passive and active adaptive management decisions in the context of an ecosystem with a threshold effect. Using discrete-time dynamic programming techniques, we model optimal phosphorus loadings under both uncertainty about natural loadings and uncertainty regarding the critical level of phosphorus concentrations beyond which nutrient recycling begins. Active management is modeled by including the anticipated value of information (or learning) in the structure of the problem, and thus the agent can perturb the system (experiment), update beliefs, and learn about the uncertain parameter. Using this formulation, we define and value optimal experimentation both ex ante and ex post. Our simulation results show that experimentation is optimal over a large range of phosphorus concentration and belief space, though ex ante benefits are small. Furthermore, realized benefits may critically depend on the true underlying parameters of the problem.

KEYWORDS: adaptive control, adaptive management, dynamic programming, value of experimentation, value of information, nonpoint source pollution, learning, decisions under uncertainty

* Corresponding author.

Using Numerical Dynamic Programming to Compare Passive and Active Learning in the Adaptive Management of Nutrients in Shallow Lakes

Introduction

A central fact of life for decision- and policy-makers engaged in ecosystem management is the need to make decisions in the presence of considerable uncertainty about system response to management actions. Although system complexity virtually ensures that such uncertainty will never be completely eliminated, any management action (or sequence thereof) theoretically provides at least some information that can be incorporated into subsequent decisions. Often, but not always, this partial resolution of uncertainty (learning) can lead to decisions that are "improved" in the sense that the linkage between decisions and management objectives is better understood.

The prospect of learning, however, introduces another margin over which to make decisions; namely, the tradeoff between collecting information about system response and managing the system to meet an objective. This has been realized by a growing number of ecologists and other natural scientists who have proposed the paradigm of "adaptive management" (AM), in which small-scale experiments are used to test hypotheses about larger-scale system responses, and the information gained is used in subsequent management decisions (Walters, 1986; Holling and Meffe, 1996; Thrower and Martinez, 2000; Prato, 2005). As such, under AM, management plans are path dependent and not *a priori* defined over the entire length of the planning horizon. Despite adoption of this management technique for several high-profile projects, including endangered species management in the Grand Canyon, Everglades restoration, and management efforts in the Missouri river and Klamath River basin (Bureau of Reclamation, 1995; Prato, 2003; USDA Natural Resources Conservation Service, 2004;

DeAngelis et al., 2000), natural resource economists have largely been silent in documenting the tradeoffs involved in AM (Milton et al, 1998; Loomis, Bond, and Harpman, forthcoming). Traditional benefit-cost analysis generally does not take the potential of learning into account, even though information collection may be implicitly valuable (see, e.g., Graham, 1981; Miller and Lad, 1984; Hanemann, 1989; Fisher and Hanemann, 1987; Chavas and Mullarkey, 2002), and thus is not particularly useful in evaluating AM strategies. More useful methodologies relating to ex ante AM analysis would help define the value and optimal extent of experimentation to collect information, and thus the trade-offs between short and long run management goals.

This paper provides one example of such a methodology using numerical dynamic programming. Our approach is based on previous research on adaptive (or dual) control (Bar-Shalom and Tse, 1976; Kendrick, 1981, 2005; Wieland, 2000), in which the state space of a dynamic optimization problem is augmented with parameters describing the extent of system uncertainty or beliefs about the system. These parameters are updated via structural rules (such as those based on Bayes' Theorem) that depend on realized, observed values of system variables, which in turn depend (in part) on management actions. Thus, the optimizing agent/manager in the dynamic model must trade-off between controls that are likely to contribute most directly to the primary management objective (such as the optimal level of pollution) and learning about system response. By varying the assumptions about the treatment of information in the problem and using the value function defined by the Bellman equation, approximations of the ex ante valuation of experimentation can be derived.

We illustrate the technique using a numerical dynamic model of management of nonpoint source pollution into a shallow lake, taken from Peterson, Carpenter, and Brock (2003). The key component of this model is a threshold effect of phosphorus concentrations which has the potential to switch the steady state of a lake from oligotrophic to eutrophic, though the precise concentration at which this switch occurs is unknown. Given a particular parameterization of this problem, we characterize the optimal management strategies assuming a) certainty over this threshold level; b) passive adaptive management under which there is uncertainty over the threshold level but learning is not anticipated; and c) active adaptive management under which there is uncertainty over the threshold level and learning is anticipated. We then calculate the optimal level of experimentation by comparing optimal paths, and estimate the value of experimentation by comparing the ex ante expected values of each strategy. This value (as well as the values of sub-optimal experimentation) can be used in ex ante policy analysis to plan and choose between experiments and to augment traditional benefit-cost analysis, and the overall solution can help simulate optimal AM paths (Bond, 2008).

This paper makes several contributions to the literature. First, the methodology provides a clear illustration of the conceptual linkages between dual control/dynamic learning and the real-world practice of active adaptive management in a theoretically-consistent manner. Second, the results document not only the circumstances under which experimentation for this problem is optimal from the point of view of the resource manager, but also the value of that experimentation. As such, we provide a means to determine the trade-offs between management goals and information gathering inherent in adaptive management. Finally, the project advances the study of the effects of learning on optimal decision making beyond that of passive information-gathering in the context of environmental management, incorporating techniques developed in the engineering and numerical sciences for the study of active experimentation by the relevant decision-maker.

Adaptive Management, Information Processing, and Dynamic Programming

The adaptive management paradigm is well-suited to a dynamic programming conceptualization in which at least one underlying parameter in a state-transition equation(s) is unknown to the controller, but information gathered over the course of the planning horizon can help this manager learn about the true nature of the parameter(s). In this context, an adaptive management problem is defined through an objective function and associated state equations which include not only the evolution of the physical system, but also the evolution of the controller's beliefs about the sufficient statistics of the uncertain parameter distribution(s).

In theory, the optimizing agent has several options in dealing with the uncertain parameter(s) (Cunha-e-Sá and Santos, 2008; Kendrick, 2005; Wieland, 2000). Most restrictive is to manage assuming that all parameters are fixed and known (at, say, the mean of the prior distribution), while a second option is to assume the prior distribution of the uncertain parameters is fixed and unchanging over time. In either case, there is no learning, and thus management is not "adaptive" in the sense that none of the information gained after the initial time period is used to update the sufficient statistics representing beliefs. In this sense, the optimal solutions are of the "open loop" variety.

However, several "closed loop" solutions are available as well. A decoupling of the updating of prior probabilities and the optimization problem results in what might be termed “passive”, or myopic, adaptive management (Wieland, 2000). In this case, the optimizing agent acts according to the policy rules of a dynamic optimization problem in which it is assumed that the distribution of the uncertain parameters does not change (as above), but after each management decision is made and the results observed through monitoring, the distribution is updated (Bond, 2008). The manager then makes the next decision in accordance with the policy rules associated with those parameter values, and the sequence continues.

On the other hand, if the optimizing agent fully anticipates learning, the optimal “active” adaptive management control path is followed. In this case, the updating of the sufficient statistics is endogenous, and not separated from the optimization step. Functionally, this implies that the optimization problem includes not only the state-transition equations related to the natural system, but also those related to the updating of the sufficient statistics of the uncertain parameters. As shown in Bond (2008), the difference between the passive and active management problems is the possibility of deviating from the passive policy rule in order to (endogenously) gain information about the unknown distributions. Thus, any difference between the passive and active policy rules can be naturally be interpreted as experimentation, and the difference in the associated Bellman expected value functions at any point in the state space could thus be viewed as the expected benefits of active adaptive management, or alternatively the expected benefits of experimentation.

To date, these types of models have not been fully adopted by mainstream environmental and resource economics (Mercado and Kendrick, 2006; Kendrick, 2005), and few papers have attempted to link these models with the paradigm of adaptive management.[1] Examples in the environmental and natural resource economics literature of models which incorporate passive Bayesian updating include those related to nonpoint source pollution (Kaplan, et al., 2003), climate change (Kelly and Kolstad, 1999), and shallow lake management (Peterson, et al., 2003), while active learning is discussed in Cunha-e-Sá and Santos (2008) with respect to air pollution, Springborn (2008) with respect to invasive species management, and Brock and Carpenter (2007) in the context of general environmental policy. We believe this paper represents the first application of dynamic programming with endogenous learning and experiment valuation to the paradigm of adaptive management.

The paper proceeds as follows. First, we develop the dynamic model and discuss the treatment of information, including definitions of passive and active adaptive management/learning. We then briefly discuss the solution technique. Next, the results are discussed with a focus on the optimal solutions of the passive and active management problems, the extent and value (both ex ante and ex post) of optimal experimentation, and the differences in evolution of beliefs and state/control paths. A final section concludes.

Model

In order to investigate the potential benefits of active adaptive management, this paper naturally extends the passive approach documented in Peterson, Carpenter, and Brock (2003) in the context of shallow lakes. We add the component of anticipatory learning, or active adaptive management, in which expectations about future information regarding the unknown parameters is jointly considered with the optimization step.

The Shallow Lakes Model

[1] Springborn (2008) is a notable exception in the context of ship examination and invasive species.

Briefly, the model is characterized by a stochastic difference equation that describes evolution of the pollution concentration of a lake as a function of prior phosphorus pollution concentrations, baseline natural phosphorus loading, managed phosphorus loading, and stochastic deviations. There is a critical level of phosphorus concentrations (say, P_{crit}) above which nutrient cycling is initiated (contributing to eutrophication), though the controlling agent/manager may be uncertain about the true value of this parameter. Formally, the difference equation can be defined as:

$$P_{t+1} = \begin{cases} \gamma P_t + b + l_t + \varepsilon_t & \text{if } P_t < P_{crit} \\ \gamma P_t + b + l_t + r + \varepsilon_t & \text{if } P_t \geq P_{crit} \end{cases}, \tag{1}$$

where P_t is the concentration of phosphorus in the lake at time *t*, *b* is the baseline natural loading of P per time period, l_t is additional P loading, presumably from anthropocentric sources, into the lake at time *t*, γ is a decay parameter, and *r* is recycled and maintained P in the lake. We assume ε_t is a normally-distributed, mean-zero, constant variance (σ^2) stochastic term that represents deviations from mean loadings at time *t*, and is known by the manager.[2] Although this is a much-simplified representation of the ecosystem, it still maintains the properties of admitting potential (but not guaranteed) oligotrophic and eutrophic stable steady-states, depending on the relationship between P loadings and concentrations.

Instantaneous payoffs depend positively on both the level of pollution allowed (e.g., the benefits from surrounding agricultural production) and negatively on the concentration of phosphorus in the lake. As such, we assume a simple linear-quadratic utility function of the form:

---

[2] We maintain this assumption following Wieland (2000) to focus attention on the uncertainty of the critical P_{crit} value. Relaxing this assumption is straightforward, but adds little economic insight to the problem.

$$U\left(l_t, P_t; \beta\right) = \beta l_t - P_t^2, \tag{2}$$

which is clearly concave in l_t and P_t so long as the (known) parameter $\beta > 0$. The manager chooses anthropocentric loadings l_t in each time period. Parameter values used in the simulations are reported in Table 1, and generally follow Peterson, Carpenter, and Brock (2003).[3]

Treatment of Information

In the shallow lakes context, it is assumed that the manager is uncertain about the value of the parameter P_{crit}, but all other parameter values are known. For simplicity (and to restrict the dimensionality of the state space), we assume that P_{crit} can only take two discrete values: $P_{crit} = 0.2$ or $P_{crit} = 0.7$. Interpreted slightly differently, this suggest that the optimizing agent is making decisions on the basis of competing models of the ecosystem, originating perhaps from rival theories or empirical findings. At the beginning of the planning horizon, the manager has beliefs over the veracity of the two models, represented by the probability weight $0 \le \pi_0 = \Pr\left(P_{crit} = 0.2\right) \le 1$.

If we assume passive or active learning, these probabilities are updated between each decision stage. Following Wieland (2000), we use the Bayes operator for updating, though this is by no means the only possibility (Klumpp, 2006). Formally, the structural updating equation is

$$\pi_{t+1} = \frac{\pi_t \cdot L_{.2,t}}{\pi_t \cdot L_{.2,t} + (1-\pi_t) \cdot L_{.7,t}}, \tag{3}$$

[3] In Peterson, Carpenter, and Brock (2003), passive learning in the context of two competing models was considered: one in which recycling never occurs at any phosphorus pollution level, and one in which recycling is always present. We introduce the notion of the critical phosphorus level.

where $\pi_t = \Pr\left(P_{crit} = 0.2\right)$ and $L_{i,t}$, $i = \left(0.2, 0.7\right)$ is the likelihood of observing a particular phosphorus concentration in period $t+1$ $\left(P_{t+1}\right)$ under the hypothesis of model *i* (Walters and Ludwig, 1994). For the ecological model defined by equation (1), this likelihood is defined by

$$L_{i,t} = \frac{\exp\left[-\left(P_{t+1} - E(P_{t+1} \mid P_{crit} = i)\right)^2\right]}{2\sigma^2}. \tag{4}$$

Of course, at any time period *t*, $\Pr\left(P_{crit} = 0.7\right) = (1 - \pi_t)$ by the properties of a probability density function.

We can now define the dynamic programming problems that define the certainty equivalent, passive adaptive management, and active adaptive management decision rules. Assuming that the manager's objective is to maximize the expected net present value of the infinite stream of utility from managing the lake, the objective function can be defined as

$$\max_{l_1, l_2, \ldots, l_\infty} \sum_{j=0}^{\infty} \delta^j E\left[U\left(l_{t+j}, P_{t+j}, \pi_{t+j}; \beta, \gamma, b, r, \sigma^2\right) \mid \pi_t\right], \tag{5}$$

where δ^t is the discount factor in time *t*, and the notation explicitly recognizes that the expected payoffs depend on the beliefs about the true values of P_{crit}. The *certainty equivalent problem* is defined by the objective in equation (5), subject to the biological equation (1), and initial condition on P_0, and the equation $\pi_{t+1} = \pi_t = 0$ or $\pi_{t+1} = \pi_t = 1$, depending on the manager's beliefs about the correct model. Note that in this specification, there is neither learning over time nor uncertainty over the correct model.

The *passive adaptive management* problem, or *passive learning* problem, is defined by the initial condition, equations (1) and (5), and the probability updating equation $\pi_{t+1} = \pi_t$. Passive adaptive management is defined as following the optimal policy rule from this problem, say $l^{pas}(P,\pi)$. As seen by the probability updating equation, $l^{pas}(P,\pi)$ does not anticipate the updating of probability beliefs, or equivalently, the rule does not incorporate the tradeoff between information collection and expected utility. However, unlike the certainty equivalent problem, we assume that after making decision l_t, the manager observes P_{t+1} and updates π_{t+1} using updating equation (3) before making decision l_{t+1}. This new value of π_{t+1} becomes the new prior, and thus forms the information set for decision l_{t+1}.

Finally, the *active management*, or *active learning problem*, is defined by the initial condition, the objective in equation (5), and the constraint set defined by both equations (1) and (3). We define the associated policy rule for this problem as $l^{act}(P,\pi)$. In contrast to the passive management problem, then, anticipatory learning is *not* decoupled from the optimization step, but rather an opportunity exists for the manager to deviate from $l^{pas}(P,\pi)$ in order to capitalize (in the future) on the information gained about the probabilities related to P_{crit}. Thus, it is natural to define *optimal experimentation* as any non-zero difference $l^{act}(P,\pi) - l^{pas}(P,\pi)$, as any deviation from the passive management strategy must be due to information effects (see Bond, 2008, for more details).

Solution Technique

We solve each of the three dynamic programming problems documented above using value function iteration with policy function acceleration over a discrete grid (with linear interpolation) in control and state space. Table 1 documents the solution algorithm parameters. We very briefly describe the method here; for more details, see Judd (1998).

The Bellman equation for each problem can be written as

$$V\left(P,\pi\right)=\max_{l}\left\{U\left(l,P;\beta\right)+\delta E\left[V\left(P^{+},\pi^{+}\right)\right]\right\}, \tag{6}$$

where P^{+} and π^{+} are the appropriate values of the phosphorus concentration state variable and probability state variable following the current period decision, and $V\left(P,\pi\right)$ is the (unknown) value function representing the expected net present value of utility along the optimal path as a function of the state variables of the problem. An iterative value function iteration procedure (with acceleration) is used to solve equation (6) for the unknown function $V\left(P,\pi\right)$ at each grid point by a) choosing an initial value for $V\left(P,\pi\right)$ at each grid point in the state space; b) calculating the right hand side of equation (6) for each point along the control and state grid space and choosing the maximum over the control dimension; c) updating the value of $V\left(P,\pi\right)$ using these values; and d) repeating b) through c) until convergence occurs. As the grid is discrete, we use linear interpolation for points between grid values.

Expectations on the right hand side of equation (6) are calculated using sixteen-point Hermite quadrature. Policy acceleration is implemented between steps b) and c) by using the estimated policy function from b) (the “optimal” control values) to iterate between the right-hand and left-hand sides of equation (6) at each grid point without the maximization step until convergence. Solutions and simulations were written using the GAMS 2.0.35.10 development language, and code is available from the senior author.

Results

We begin by characterizing the solutions to the certainty problem, the passive adaptive management problem, and the active adaptive management problem, and characterizing the extent of optimal experimentation. We then discuss the evolution of beliefs and provide an example of alternative management paths for passive and active learning, and show how to value optimal experimentation both ex ante and ex post.

Optimal Policy Functions

Figure 1 displays the estimated optimal anthropocentric loading (policy function) under differing prior beliefs. Given the nature of the problem, the optimal strategy for the manager under certainty ($\pi = 0$ and $\pi = 1$) is characterized by a most rapid approach path to the deterministic steady state level, which in this case implies loading in each period such that $E\left(P_{t+1}\right) = P_{\infty}^{i}$, where P_{∞}^{i} is the optimal steady-state level when $P_{crit} = 0.2$ ($P_{\infty} \approx 0.68$) or when $P_{crit} = 0.7$ ($P_{\infty} \approx 0.51$). The presence of the stochastic loading term, however, ensures that the long run equilibrium values of phosphorus concentrations comprise a distribution around the appropriate steady state.

As such, optimal loadings under certainty are non-continuous, downward-sloping functions of phosphorus levels, with exactly one discontinuity at the threshold level of pollution concentration.[4] Comparing the two policy functions yields the insight that as the prior belief that $P_{crit} = 0.2$ increases (from $\pi = 0$ to $\pi = 1$), optimal loadings increase at

[4] When $\Pr(P_{crit}=0.2)=0$ or $\Pr(P_{crit}=0.2)=1$, the solutions to the passive and active learning problems are identical.

low (less than 0.2) and high (greater than 0.7) phosphorus concentrations, but decrease at intermediate levels. This results from the interplay between the threshold effects and the fact that marginal disutility increases with phosphorus concentrations, resulting in differing optimal steady states.

Turning next to the optimal passive learning policy functions and assuming $0 < \pi < 1,$ one can gain insight into the effects of introducing parameter uncertainty (but not anticipatory learning) into the problem. Unsurprisingly, at relatively low and high phosphorus concentrations, optimal loadings at intermediate values of π are associated with values between the bounds established by the models with complete certainty. However, at phosphorus levels between 0.2 and 0.7, this relationship does not hold. In the upper graph of Figure 1, for example, optimal loadings when $\pi = 0.5$ and $\pi = 0.75$ are actually *lower* than when $\pi = 0$ or $\pi = 1.$ Furthermore, at these same intermediate probability levels, the entire optimal loading function is less (by a constant) than that same function when $\pi = 1.$ As a lower P_{crit} level is associated with a higher steady state phosphorus concentration level (and lower steady-state utility), these two effects can be considered manifestations of the "precautionary principle", in which agents tend to act conservatively in the face of uncertainty (see Immordino, 2003 for a recent review). This is not true, however, for $\pi = 0.25,$ in which the optimal policy is identical to that under certainty that $P_{crit} = 0.7.$

In contrast, the active learning policy function illustrated in the lower graph of Figure 1 introduces still more complexity into the optimal decision. While the optimal active loading response is identical to the passive rule when $\pi = 0.25,$ it is more *conservative* (lower loadings) at low and high phosphorus concentrations, but more *aggressive* (higher loadings) at intermediate levels, when $\pi = 0.5$. In other words, the effects of active, anticipatory learning are of the opposite sign and of greater magnitude than the precautionary effects at intermediate concentration levels, but are complementary to it at lower and higher levels. When $\pi = 0.75,$ however, this latter effect is not observed, though the learning effects are present for intermediate phosphorus concentrations.

Optimal Experimentation

The deviation in optimal loadings between the passive and active learning/adaptive management problems is interpreted as optimal experimentation, as it is solely the result of taking into account the effect of gathering information to reduce uncertainty about the unknown parameter. Figure 2 provides a graphical illustration of the patterns of deviation found over the belief space.

Some degree of experimentation over the feasible range of phosphorus concentrations (zero to one hundred percent) is optimal for all $0.4 < \pi < 1$. As such, experimentation is more generally more prevalent when $P_{crit} = 0.2$ is believed (and hence the steady state levels of pollution is higher and utility is lower than when $P_{crit} = 0.7$). However, the form of this experimentation differs with both beliefs and pollution concentration. For example, when $\pi = 0.425$ and $\pi = 0.975,$ it is optimal under active learning to experiment only at low and high concentration levels, and by setting optimal loadings *lower* than the passive case.

In contrast, when $\pi = 0.7,$ experimentation (active management) is optimal at intermediate concentration levels, and by setting loadings *higher* than the passive case. However, the solutions are identical at more extreme pollution levels. Finally, when the manger is completely uncertain about which model is correct $(\pi = 0.5)$, experimentation is optimal over the entire range of feasible concentration levels. As described in the previous subsection, loadings under active learning are lower at extreme concentration levels for this case, but higher at intermediate levels.

Thus, we conclude that optimal experimentation is state-dependent in both manager beliefs and pollution concentrations, and may involve loadings that are either more conservative or more aggressive than management under passive learning. For this problem, aggressive experimentation (higher loading) is optimal only at intermediate pollution concentrations for beliefs of $0.45 \le \pi \le 0.95$, while experimentation under relatively low and high concentration levels is optimal (and conservative) only at $0.425 \le \pi \le 0.525$ and $0.925 \le \pi \le 0.975$. Experimentation of any sort is not optimal for $\pi \le 0.4$. Note that these results are likely specific to this problem and the chosen parameter values, and may not generalize to other contexts, but illustrate the (endogenous) tradeoffs between information collection and current period utility.

Evolution of Beliefs and Management Simulation

As seen in equation (4), the likelihoods of each realization depend only on the error between actual realized phosphorus concentrations and their expectation regardless of the coupling or decoupling of the probability updating and the optimization steps. As this expectation is dependent on the pollution level P_t due to the threshold effect, difference in likelihoods between the passive and active learning models begin occurring only when P_t in each model are on either side of one of the P_{crit} values (i.e., $P_t^m < P_{crit}^i$ and $P_t^n > P_{crit}^i$ for models m and n, $i \in (0.2, 0.7), m \ne n)$. Thus, the evolution of beliefs does not often differ markedly between the active and passive solutions, though one instance of differing errors can be propagated over a number of periods. If, however, belief evolution is identical over some simulated path of management, it implies that differences in the path of pollution concentration between passive and active learning are due solely to experimentation, rather than differences in the evolution of beliefs.

Figure 3 displays several realized simulations of active and passive learning, each assuming identical stochastic shocks, initial pollution concentrations of 0.5 and initial beliefs of $\pi = 0.5,$ but varying the true P_{crit} level. In each case, beliefs converged to the true P_{crit} level fairly rapidly (19 periods for the $P_{crit} = 0.2$ simulation, and 31 periods for the $P_{crit} = 0.7$ simulation). Deviations between control and stock levels are greatest for the low threshold simulation, partially due to the fact that pollution levels in period 2 are 0.69 for the passive learning model and 0.78 for the active learning model. This does not occur in the high threshold simulation, and as such, beliefs are identical for each time period, and loadings and stock levels are identical after period 4 (when the manager is almost 99% certain that $P_{crit} = 0.7$). In each case, active learning outperformed passive learning in the sense of maximizing discounted utility, though the effects were small (0.4% when $P_{crit} = 0.2$ and 2.6% when $P_{crit} = 0.7$ over the first twenty years).

Value of Optimal Experimentation and Monte Carlo Results

In addition to using the obtained policy rules to illustrate when (and by how much) it is optimal to experiment and to simulate optimal management paths, the dynamic programming methodology can be used to obtain the ex ante expected value of experimentation by subtracting the expected value of following the passive learning rules from the expected value of following the active management rules. For the active learning paths, this expected value is given by $V^{act}\left(P,\pi\right)$ from the active adaptive management Bellman equation given by equation (6). However, the value function from the passive management problem is *not* the correct value for this model, as probabilities are, in fact, updated along any simulated path. This fact is not taken into account in the optimization step when the passive learning policy rule is derived.

Instead, the proper value is the value function associated with the equation

$$V^{pas}\left(P,\pi\right)=U\left(l^{pas}\left(P,\pi\right),P;\beta\right)+\delta E\left[V^{pas}\left(P^{+},\pi^{+}\right)\right], \tag{7}$$

where $l^{pas}\left(P,\pi\right)$ represents optimal loading for the passive learning model and all other variables are as defined previously (Bond, 2008). It can be shown that equation (7) remains a contraction mapping, and thus $V^{pas}\left(P,\pi\right)$ can be recovered via a process identical to the policy acceleration mentioned above (Judd, 1998).

Figure 4 documents the ex ante expected value of experimentation $\left(V^{act}\left(P,\pi\right)-V^{pas}\left(P,\pi\right)\right)$ as a function of beliefs for three pollution concentration levels (0.1, 0.5, and 0.9). Overall, the value of experimentation is small, which is unsurprising given that experimentation and uncertainty is concentrated in earlier time periods, the discount factor is assumed close to one $\left(\frac{1}{(1+r)}=.99\right)$, and we assume an infinite time horizon.[5] However, we observe a similar pattern of the gains regardless of starting pollution level.

Specifically, the gains from experimentation approximately double at the belief point where aggressive experimentation begins, and stays relatively constant or slightly decreasing over the range $0.45\leq\pi\leq0.75$, where uncertainty is fairly substantial. The

[5] This assumption follows Peterson, Carpenter, and Brock (2003), and reflects a relatively low discount rate (high weight on the future) relative to market rates.

value of optimal experimentation is gradually increasing as the belief that $P_{crit} = 0.2$ increases from zero until $\pi = 0.45$, but decreasing rapidly for $\pi > 0.75$. As such, for this problem, we conclude that experimentation is most valuable under conditions of considerable uncertainty about the pollution threshold level, especially if the manager believes that low thresholds are slightly more likely. In other words, in the face of uncertainty when it is believed the relatively negative outcome is more likely than the alternative, experimentation is more valuable ex ante.

We can perform a similar ex post analysis using Monte Carlo simulation. Table 2 displays the results of 1,000 simulations each when $P_{crit} = 0.2$ and $P_{crit} = 0.7$ for the passive and active adaptive management models, along with the mean and standard deviations of the net present value (NPV) of the utility stream over a 100 year time period and the percentage of the simulations where the sum of discounted utility is greater for the active model. While the mean NPV of utility is greater for the active model under both true threshold levels, it is not significantly different for either. Furthermore, we can rate the performance of the active versus passive learning policy rules by counting the number of simulations under which the NPV of utility is greater for the model that includes anticipatory learning. Active learning outperforms passive learning decisively at high threshold levels (89% of the simulations), but underperforms when the threshold level is low (45% of the simulations). It seems likely that this result is directly related to the quadratic term on pollution concentrations in the utility function, as aggressive experimentation when the threshold is low will tend to decrease utility as a result of increased pollution.

Conclusions

This paper has illustrated the use of numerical dynamic programming techniques to aid researchers, policy makers, and resource managers in defining the optimal extent and value of experimentation when underlying ecosystem parameters are unknown. Although greatly simplified relative to many complex ecosystems, the shallow lakes management model used here includes an uncertain parameter representing a threshold effect that may be fairly common in real world systems, and provides a natural way of defining experimentation through loading deviations from a passive learning path. It was shown that experimentation is optimal over a large segment of the state space, which included both pollution concentrations and beliefs about the threshold level, but that ex-ante gains were relatively small and that ex-post gains depended on the true (but unknown) threshold level.

Although many of our results are likely specific to the model and parameterization used in this paper, a few general conclusions can be deduced. First, these adaptive control/dynamic programming techniques can be used to identify optimal policy paths under models of active and passive adaptive management, or help evaluate potential (possibly non-optimal) experiment sets ex ante. This includes not only the level and extent of experimentation, but also the expected values of that experimentation. As such, application of these methods and models can aid in the implementation of adaptive management (AM) programs and help AM managers make better, more informed decisions in the presence of significant uncertainty. Specifically, the information from these models can be used to help choose between experiments, provide expected values of experiments, and augment more traditional benefit cost analysis to account for the value of future information, as well as simulate potential AM paths (Bond, 2008).

Second, regardless of the problem, the subjective beliefs of the decision maker will likely be a factor in the value and optimal extent of experimentation. While this may be troubling to those more comfortable with "objective" analysis, it seems unlikely that purely objective recommendations exist regarding optimal management in models of learning. Perhaps one benefit is that the relevant values can be calculated over the entire belief space, which can subsequently be used to make decisions when experts disagree.

Finally, our Monte Carlo analysis suggests that the realized gains from an active adaptive management/learning approach may depend critically on the (initially unknown) true parameter values, and despite expected ex ante gains from anticipating future information, passive learning may perform equally as well. In the case explored here, following the optimal experimentation path can exacerbate environmental problems when the underlying threshold value is low, and this can lead to relatively disappointing active learning results.

Clearly, this method will not be applicable for all AM applications, especially when the number of unknown parameters is large and continuous distributions are assumed. The "curse of dimensionality" is a very real problem when means, variances, and covariances must be modeled jointly, and despite recent advances in numerical techniques, these problems may remain intractable. However, as computing power advances and better solution algorithms are derived, dynamic programming techniques may prove to be a powerful tool for economists and others looking to improve adaptive environmental management.

References

Bar-Shalom, Y. and E. Tse. 1976. Caution, probing, and the value of information in the control of uncertain systems. Annals of Economic and Social Measurement 5 (3) : 323-337.

Bond, C.A. 2008. On the potential use of adaptive control methods for improving adaptive natural resource management. Working Paper. Fort Collins: Colorado State University.

Bond, C.A. and Y.H. Farzin. 2008. Alternative sustainability criteria, externalities, and welfare in a simple agroecosystem model: A numerical analysis. Environmental and Resource Economics 40 (3): 383-399.

Brock, W.A. and S.,R. Carpenter. 2007. Panaceas and diversification of environmental policy. *Proceedings of the National Academy of Sciences* 104 (39): 15206-15211.

Bureau of Reclamation. 1995. Operation of Glen Canyon Dam Final Environmental Impact Statement. Salt Lake City, Utah: Bureau of Reclamation.

Chavas, J.P. and D. Mullarkey. 2002. On the valuation of uncertainty in welfare analysis. American Journal of Agricultural Economics 84 (1): 23-28.

Cunha-e-Sá, M.A. and V. Santos. 2008. Experimentation with accumulation. Journal of Economic Dynamics and Control 32: 470-496.

DeAngelis, D., L. Gross, W. Wolf, D. Fleming, M. Nott and E. Comiskey. 2000. Individual based models on the landscape: Application to the Everglades. In Landscape Ecology: A Top-Down Approach., edited by J. Sanderson and L. Harris, 199-211. Boca Raton, FL: Lewis Publishers.

Ferrari, S. and R.F. Stengel. 2005. Smooth function approximation using neural networks. IEEE Transactions on Neural Networks 16 (1): 24-38.

Fisher, A.C. and W.M. Hanemann. 1987. Quasi-option value: Some misconceptions dispelled. Journal of Environmental Economics and Management 14: 183-190.

Fonnesbeck, C.J. 2005. Solving dynamic wildlife resource optimization problems using reinforcement learning. Natural Resource Modeling 18 (1): 1-40.

Graham, D.A. 1981. Cost-benefit analysis under uncertainty. American Economic Review 71(4): 715-725.

Hanemann, W.M. 1989. Information and the concept of option value. Journal of Environmental Economics and Management 16: 23-37.

Holling, C. and G. Meffe. 1996. Command and Control and the Pathology of Natural Resource Management. Conservation Biology 10 (20): 328-337.

Immordino, G. 2003. Looking for a guide to protect the environment: The development of the precautionary principle. Journal of Economic Surveys 17 (5): 629-643.

Judd, K.L. 1992. Projection methods for solving aggregate growth models. Journal of Economic Theory 58: 410-452.

Judd, K.L. 1998. *Numerical Methods in Economics*. Cambridge, MA: MIT Press.

Kaplan, J.D., R.E. Howitt, 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.

Kendrick, D.A. 1981. Stochastic Control for Economic Models. New York: McGraw-Hill, Inc.

Kendrick, D.A. 2005. Stochastic control for economic models: past, present, and the paths ahead. Journal of Economic Dynamics and Control 29: 3-30.

Kelly, D.L. and C.D. Kolstad. 1999. Bayesian learning, growth, and pollution. Journal of Economic Dynamics and Control 23: 491-518.

Klumpp, T. 2006. Linear learning in changing environments. Journal of Economic Dynamics and Control 30: 2577-2611.

Loomis, J.L., C.A. Bond, and D. Harpman. Forthcoming. The potential of agent based modeling for performing economic analysis of adaptive natural resource management. Journal of Natural Resources Policy Research.

Mercado, P.R. and D.A. Kendrick. 2006. Parameter uncertainty and policy intensity: Some extensions and suggestions for further work. Computational Economics 27: 483-496.

Miller, J.R. and F. Lad. 1984. Flexibility, learning, and irreversibility in environmental decisions: A Bayesian approach. Journal of Environmental Economics and Management 11: 161-172.

Milon, W., C. Kiker and D. Lee. 1998. Adaptive ecosystem management and the Florida everglades: More than trial-and-error? Water Resources Update 113: 37-46.

Peterson, G.D., S.R. Carpenter, and W.A. Brock. 2003. Uncertainty and the management of multistate ecosystems: An apparently rational route to collapse. Ecology 84 (6): 1403-1411.

Prato, A. 2003. Adaptive management of large rivers with special reference to the Missouri River. Journal of the American Water Resources Association 39: 935-946.

Prato, T. 2005. Bayesian adaptive management of ecosystems. Ecological Modelling 183: 147-156.

Springborn, M. 2008. Bayesian adaptive management with optimal learning. Working paper. Presented at Colorado State University, Fort Collins, CO., January 28, 2008.

Thrower, A and J.M. Martinez. 2000. Reconciling anthropocentrism and biocentrism through adaptive management. Journal of Environment and Development 9(10): 68-97.

USDA Natural Resources Conservation Service. 2004. Updated Work Plan on Adaptive Management. www.nrcs.usda.gov/feature/klmath.klamplan.html

Walters, C. 1986. Adaptive Management of Renewable Resources. New York: McGraw Hill, Inc.

Walters, C. and D. Ludwig. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Canadian Journal of Fisheries and Aquatic Science 51: 713-722.

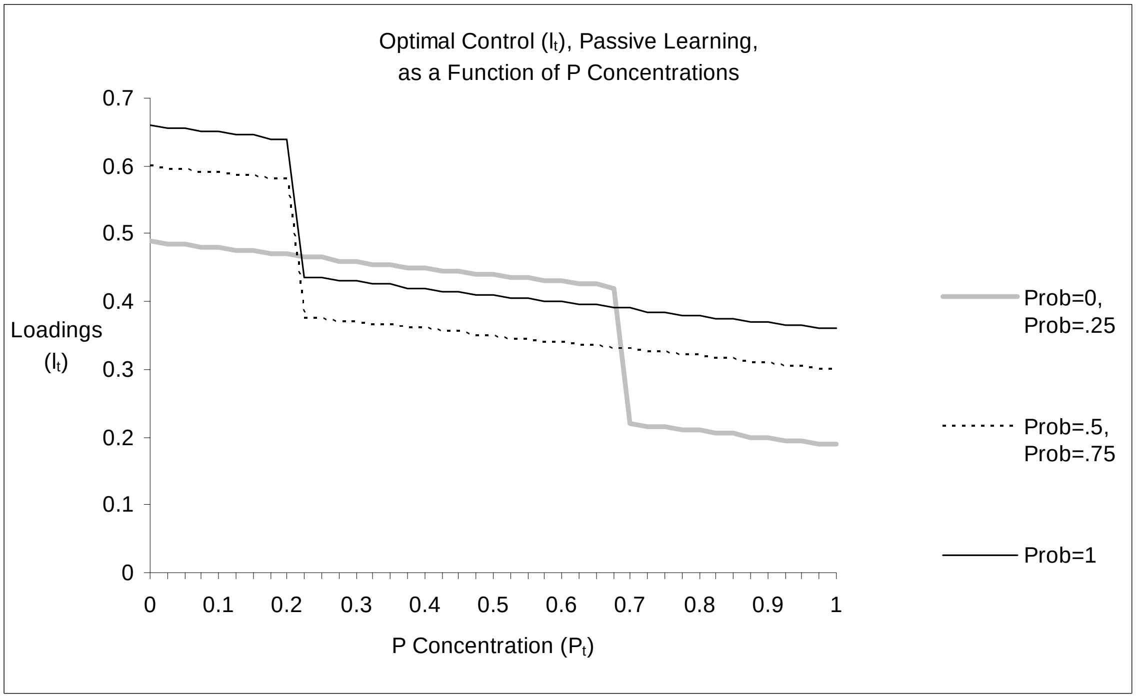
Wieland, V. 2000. Learning by doing and the value of optimal experimentation. Journal of Economic Dynamics and Control 24: 501-534.

Table 1. Maintained parameter and solution algorithm values used in numerical optimization and simulation.

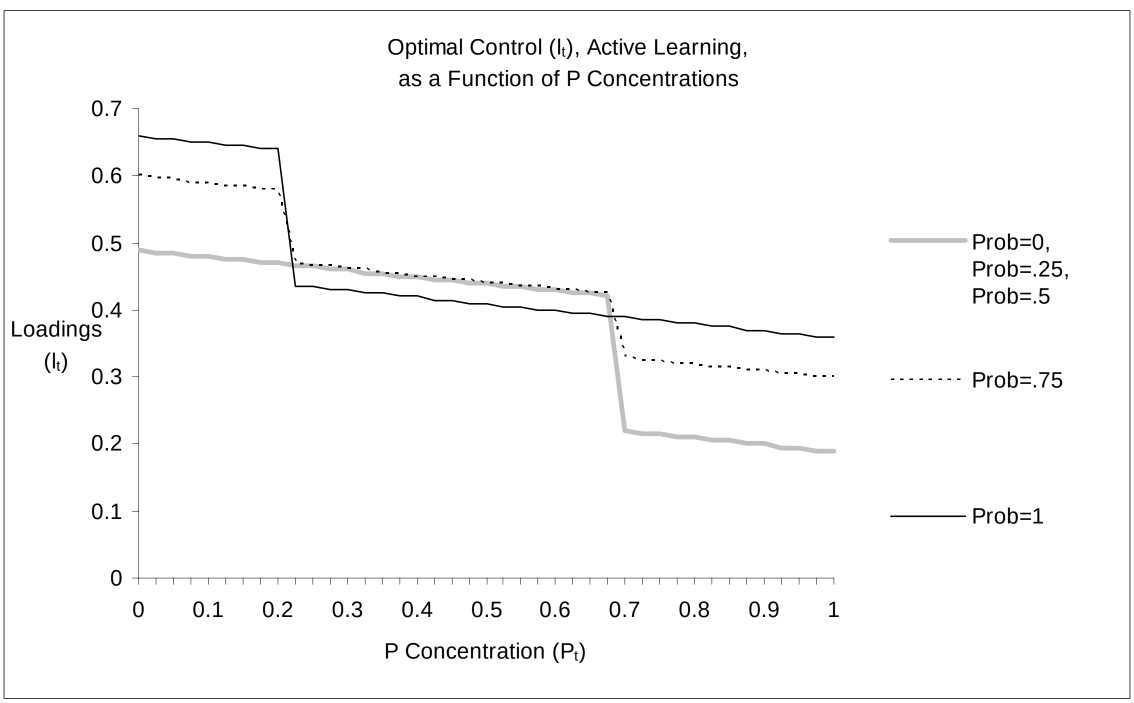
Description	*Variable*	*Value*	*Nature of Variable*
Model Parameters			
Threshold beyond which recycling occurs	P_{crit}	(0.2,0.7)	Fixed, unknown parameter
Decay rate of Phosphorus concentration	γ	0.1	Fixed, known parameter
Natural baseline loading	b	0.02	Fixed, known parameter
Phosphorus recycling parameter	r	0.2	Fixed, known parameter
Discount factor	δ	0.99	Fixed, known parameter
Relative marginal utility of loadings	β	1.5	Fixed, known parameter
Standard deviation of stochastic shock	σ	0.141421	Fixed, known parameter
Phosporus concentrations (range)	P_t	0.0-1.0	State variable
Belief that P_{crit}=0.2 (probability range)	π	0.0-1.0	State variable
Phosphorus loadings (range)	l_t	0.0-0.8	Control variable
Solution Parameters			
# grid points, phosphorus concentration	n/a	41	n/a
# grid points, belief probability	n/a	41	n/a
# grid points, phosphorus loadings	n/a	161	n/a
# Hermite nodes and weights	n/a	16	n/a
Value function error tolerance	n/a	0.001	n/a

Note: Where possible, model parameters closely followed Peterson, Carpenter, and Brock (2003).

Figure 1. Optimal loadings as a function of phosphorus concentrations for certainty, active, and passive learning under selected belief probabilities.

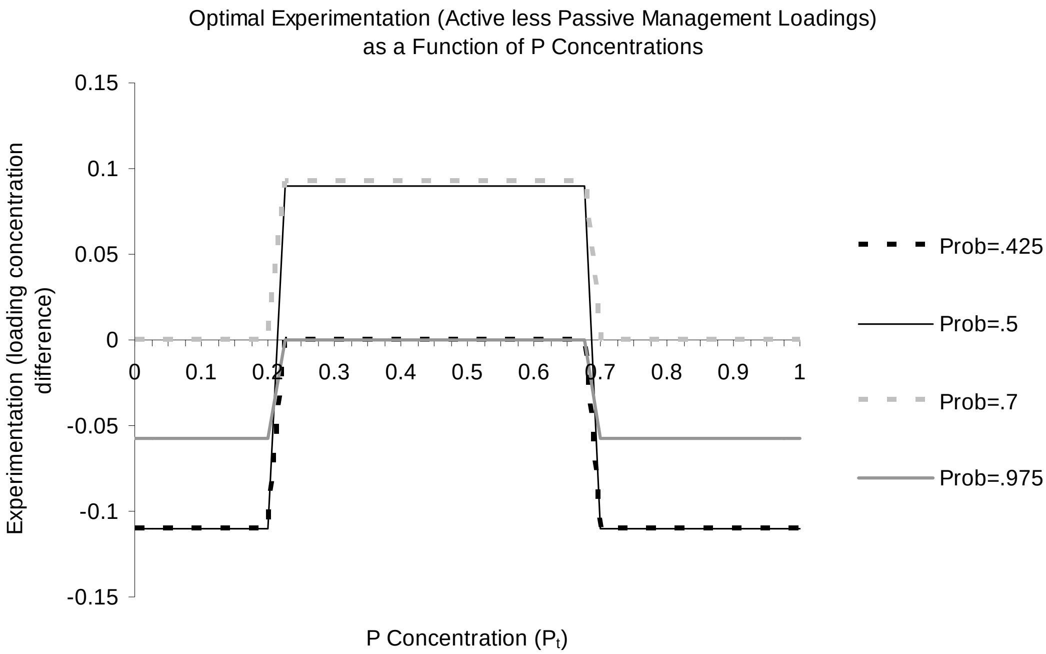


Note: Prob=0 denotes certainty that P_{crit}=0.7, and Prob=1 denotes certainty that P_{crit}=0.2.
Passive learning optimal loading functions (upper panel) identical for Prob=(0,.25) and Prob=(.5 ,.75).



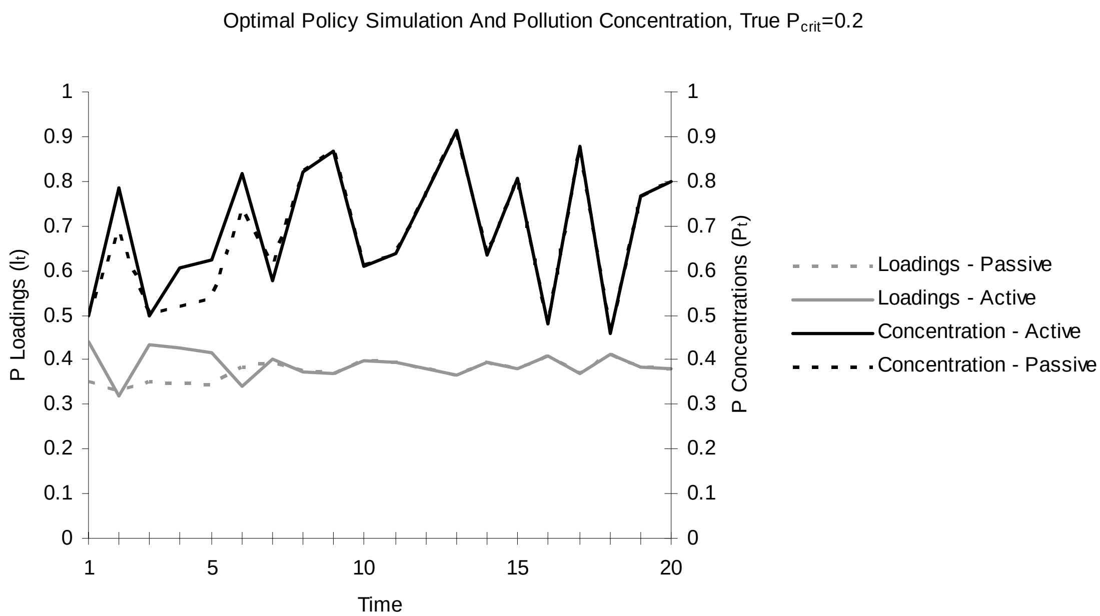
Note: Prob=0 denotes certainty that P_{crit}=0.7, and Prob=1 denotes certainty that P_{crit}=0.2.
Active learning optimal loading functions (lower panel) identical for Prob=(0,.25,.5).

Figure 2. Optimal experimentation (active less passive optimal loadings) as a function of phosphorus concentrations under selected belief probabilities.



Note: Optimal experimentation equals zero (equivalent optimal policy functions for passive and active problem) for 0≤P<0.2 and 0.7<P≤1 when Prob=0.7 and for 0.2<P<0.7 when Prob=0.425 and Prob=.975.

Figure 3. Realized optimal loadings and pollution concentrations for simulations under passive and active learning.



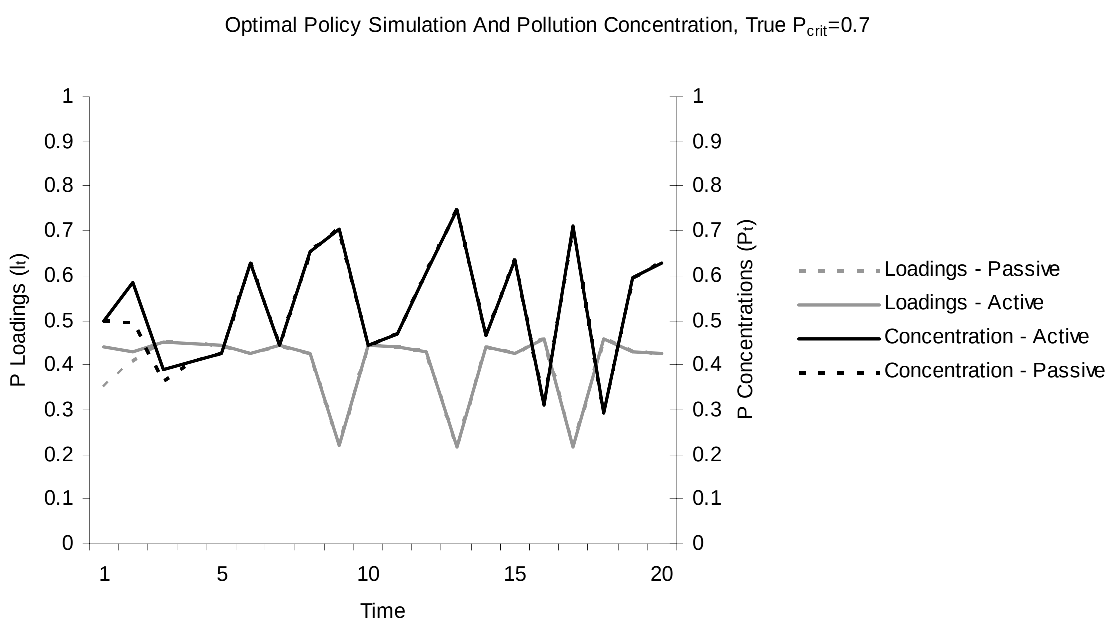


Figure 4. Ex ante expected value of experimentation (active vs. passive control rule) for selected phosphorus concentration levels.

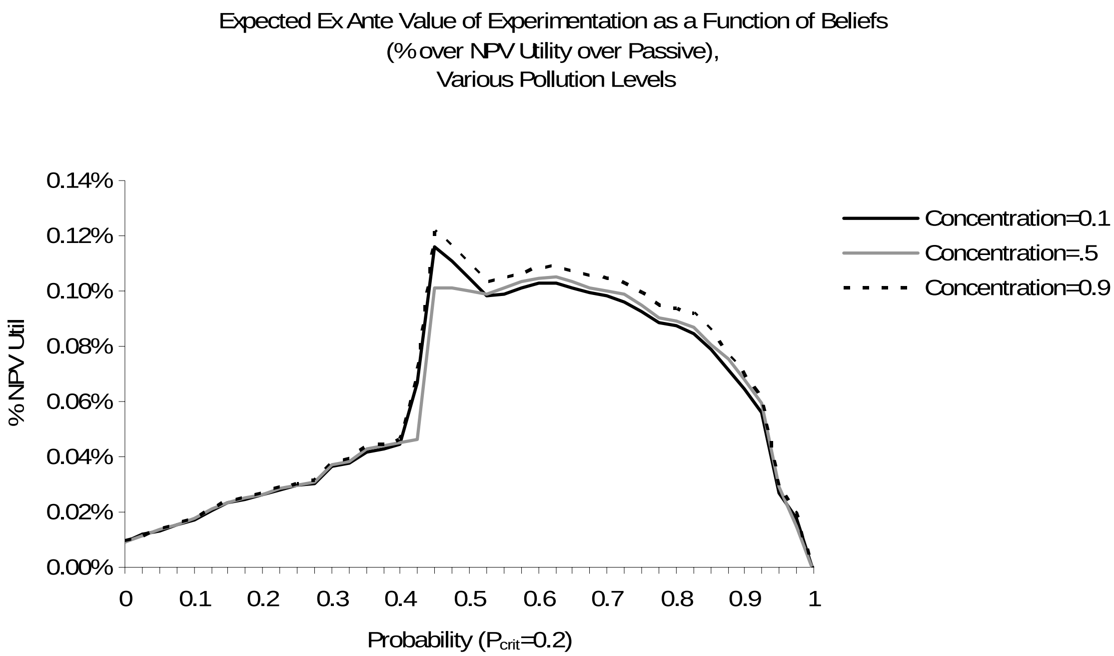


Table 2. Ex post Monte Carlo simulations under alternative values of P_{crit} under passive and active learning.

	NPV Utility (100 yrs)		*% of Simulations where NPV Utility*
True P_{crit}	**Passive**	**Active**	***Greater under Active Learning***
0.2	6.870	6.872	45%
	(1.43)	(1.44)	
0.7	22.021	22.050	89%
	(1.54)	(1.55)	

Std. errors of NPV Utility in parentheses.

Discount factor = 0.99.