March 2017 APEC RR17-02

Machine Learning Based Policy to Ease Information Asymmetry in Non-Point Pollution Management

Jacob R. Fooks[1], Kent D. Messer[2], and Jordan F. Suter[3]

[1]USDA Economic Research Service, [2]University of Delaware, [3]Colorado State University

APPLIED ECONOMICS & STATISTICS

APEC Research Reports

Department of Applied Economics and Statistics College of Agriculture and Natural Resources • University of Delaware

ABSTRACT

Machine Learning Based Policy to Ease Information Asymmetry in Non-Point Pollution Management

Keywords: Nonpoint source pollution, experimental economics, neural network

This research examines how an artificial neural network incorporating high-frequency monitoring data and natural system dynamics can inform policies that regulate an environmental externality with inherent information asymmetry. Using an experiment with both students and agricultural producers we study strategic behavior under various policies and measure participants' relative values for different levels of information accuracy under such policies. First, we show that a neural-network-based recursive filter can be applied to monitoring data to estimate an individual polluter's contribution to the ambient level of pollution, in essence, turning nonpoint sources into estimated point sources. We then test the implications of this result using an economic experiment that explores the effects of spatial relationships and the information structure of policies on behavior and preferences. The results of the experiments show that participants change their emissions in response to both policy and information treatments and that there are no significant differences in behavior between professional and student participants. However, we find that the agricultural producers are more willing than student participants to pay for policies that more accurately target the individual sources of pollution. This latter result suggests a strong preference for polluter-pay policies instead of ambient-based policies amongst producers, even if they do not necessarily lead to higher total profits.

Research Highlights:

1. Incentives for polluters may be discriminated using machine learning based policy.
2. High frequency data, system dynamics can target non-point source (NPS) polluters.
3. Economic experiments test behavior in neural-network based pollution policies.
4. Pollution was reduced most efficiently with ambient-focused policies.
5. Agricultural producers showed a positive value for non-strategically useful information.

ACKNOLWEDGEMENTS

This research was supported by funding from the National Science Foundation Northeast Water Resource Network (NEWRnet) project. The views expressed are those of the authors and not necessarily those of the Economic Research Service or the U.S. Department of Agriculture.

Suggested Citation for APEC Research Reports

Wiest, W.A., K.D. Messer, and W.G. Shriver. 2013. "Incorporating Climate Change with Conservation Planning: a Case Study for Tidal Marsh Bird Conservation in Delaware, USA." *Applied Economics & Statistics Research Report*, University of Delaware, RR13-01.

JEL Codes: C45, Q25, Q58

Machine Learning Based Policy to Ease Information Asymmetry in Non-Point Pollution Management

1. Introduction

Implementing policy to correct market failure often presents challenges related to information asymmetry and behavioral uncertainty (Shogren and Taylor, 2008). Machine learning, in combination with increased availability of a variety of spatially explicit data, may offer opportunities to improve policy under these conditions through adaptive management and increased efficiency in the use of the information that is available to regulators. Indeed, many mechanism design problems can be reduced to machine learning problems (Balcan et al,, 2007; Mohri and Muñoz Medina, 2010). Previous work incorporating machine leaning into to the design of policy mechanisms has focused on using computational constructs to understand the behavior of regulated agents (Arthur, 1993; Terna, 2000; van der Hoog, 2016), or for applications in the computer science domain, such as allocating computational resources within a network (Demirci, 2015). Much of the application of machine learning in economics has been limited to its potential as an econometric tool (Varian, 2014; Athey and Imbens, 2016). This research applies machine learning directly to an environmental policy mechanism, and then tests that mechanism in a laboratory economics experiment (Tagiew, 2012; Tagiew et al.,). Specifically, it explores the potential introduction of artificial neural networks into nonpoint source (NPS) water pollution regulation.

NPS pollution is a classic problem of asymmetric information in environmental regulation. Typically, the total amount of pollution in a water resource (or a statistically noisy measurement of it) can be publicly monitored, but information about each individual's contribution to total pollution, is too costly, technologically infeasible, or politically impractical for a regulator to obtain. Thus, individuals have an incentive to shirk on pollution-reduction efforts since individual activity is unobservable. Water sampling to detect and measure pollutants, such as nitrogen, phosphorus, and bacteria, is typically done via manual "grab sampling" at intervals of weeks to months or by mechanical auto-samplers that collect samples every few days, providing data that are retrieved and analyzed periodically. Actual pollution levels, on the other hand, vary widely day to day based on when the pollution is generated and when and how it is transported into water bodies. Most contaminant flows occur during high-runoff storm events that create transient fluxes in pollution concentration over time (Inamdar et al., 2015; Dhillon and Inamdar, 2013).

The difficulty in obtaining information about contributions to NPS pollution has led to an assortment of proposed theoretical mechanisms and several empirical analyses (Shortle and Horan, 2001; Xepapadeas, 2011). A basic assumption underlying these efforts is a positive relationship between the production levels of agricultural landowners contributing to the pollution and the social damage caused by the pollution, such as increased eutrophication and hypoxia from nitrogen and phosphorus pollution. This tradeoff implies that reductions in emissions can improve social welfare. Mechanisms proposed by economists to achieve optimal pollution levels have included taxes and subsidies, pollution standards, water quality markets, contracts, and liability rules (Shortle and Horan, 2001). In practice, the primary policy tools used to regulate NPS water pollution have been focused on controlling inputs into the landowners' production processes (Xepapadeas, 2011). Evidence from laboratory economics experiments has suggested that input-focused policies can be highly effective (Cochard et al., 2005); however, the resulting provision of information about private costs can lead to extraction of large surplus rents by landowners (Kirwan et al., 2005; Arnold et al., 2013). Another type of mechanism recommended by economists is output-focused ambient tax and subsidy mechanisms that target an exogenously set pollution level (Segerson, 1988; Meran and Schwalbe, 1987; Segerson, 1998). These types of policies align producers' incentives in a cost-effective way by implementing financial penalties, such as taxes, when a producer deviates from the pollution level target. Such mechanisms have attractive theoretical properties and can lead to zero transfers in equilibrium. A number of studies (Alpizar et al., 2004; Camacho-Cuena and Requate, 2012; Cochard et al., 2005; Spraggon, 2002; 2013; Poe et al., 2004; Suter et al., 2008; Vossler et al., 2006) have found that these ambient mechanisms are highly effective in the laboratory under quite general assumptions. However, in the face of producer heterogeneity, they can lead to distributional inequalities that make practical implementation politically problematic (Suter et al., 2009).

Emerging technologies for water quality sampling have led to preliminary deployments of in-stream sensors that can collect and analyze samples frequently and transfer real-time data wirelessly to policy makers, researchers and stakeholders. These technologies provide a rich source of monitoring data, but a larger quantity of more-accurate data alone cannot directly improve water-quality and pollution-regulation policies. Such information could, however, potentially allow regulators to move from ambient measures to policies that regulate nonpoint sources essentially as point sources. Identifying those sources based on downstream pollution levels involves constructing a model based on knowledge of the dynamics of the system that governs pollution concentrations. Complete structural models, such as the QUAL2K model used in this paper (Chapra et al., 2008), incorporate advection, or downstream flow; diffusion, or transport of nutrients from high concentration areas to low concentration areas; and kinematics, or within stream chemical and biological processes. These have been shown to be non-invertible, so an attribution of pollution to a particular point source based on this data is necessarily estimated. One approach is to use semi-structured spatial statistical techniques such as the U.S. Geological Survey's popular Spatially Referenced Regression on Watershed Attributes (SPARROW) model (Smith et al., 1997), which works well for large-scale, regional pollution sources but does not provide highly accurate estimates at an individual parcel scale. Probabilistic Bayesian model inversions (Huang and McBean, 2007; Shen and Yuan, 2009) have also been used, but they can pose dimensionality problems that prohibit use over a large area.

The approach we explore for estimating individual parcel emissions using high-frequency monitoring data uses a neural filter (NF) that consists of a recursive series of artificial neural networks constructed from information about the downstream accumulative structure of the water system (Lo, 1994; Shtauss, 2008). The resulting decomposition allows estimation of individual contributions of nitrogen runoff to a "pulse" of nutrients moving downstream. This approach has the ability to turn non-point sources into point sources, albeit with some error in the estimation process. We explore the properties of this tool in a synthetic setting that incorporates realistic instream nutrient-routing dynamics and find that it is effective in offering reasonable estimates of differentiated pollution contributions from individual sources. This model is not uniquely invertible, so the estimates are necessarily subject to uncertainty that generally, though not exclusively, increases nonlinearly with distance upstream from the sensor.

Using the spatial attribution approach to estimate emissions from individual sources has the potential to improve NSP policy over approaches that rely solely on ambient pollution observations. The behavioral response of polluters to such policies, however, is not well understood since it is not possible to observe decision making under both spatially attributable policies and ambient-based policies. To test the behavioral response to NPS policies and information that are spatially targeted, we conducted a laboratory economics experiment using both student subjects and agricultural producers. The experiment includes treatments gauging both the strategic behavior under different policy mechanisms and information sets as wells as costly voting treatments to gauge participants' preferences over these policy types.

2. Background

In this section we begin by describing the decision setting that underlies our experiment, involving individual firms situated along a river. We then describe how a neural networks approach is used to attribute pollution emissions in this setting to each of the individual firms. This attribution process motivates a series of experimental treatments and in the final portion of this section, we discuss the specific experimental design and protocol that we followed.

2.1 Decision Setting

Consider a river that has a directional flow and six firms, $i = 1, \ldots, 6$, arranged linearly along the river with parcel 1 being furthest upstream and parcel 6 being furthest downstream (see Figure 2). Two additional dummy parcels, $i = 0$ and $i = 7$, respectively represent the headwater and outflow of the river. The firms produce a good and in the process generate pollution emissions that flow into the river and are transported downstream according to a physical model as described in appendix B. Each firm chooses a level of emissions over a series of rounds, $k = 1, \ldots, K$, generating revenue based on a functional relationship that is concave and increasing in emissions. The revenue function in equation 1 is based on Spraggon (2002) where the revenue, G, is based on emission, $x_{i,k}$, for firm i in round k:

$$G_{i,k}(x_{i,k}) = 35 - 0.0075 * (50 - x_{i,k})^2. \quad (1)$$

Within each round, the pollution emissions generated by each of the firms enter the river simultaneously at time $t = 0$ and flows downstream according to the physical model over iterations $t = 1, 2 \ldots, 200$ where t represents fifteen-minute intervals following the initial runoff of nutrients into the river.

In the model, the ambient level of pollution at time t at parcel i during round k is represented as $p_{i,k,t}$. The level of emission chosen by each firm determines the initial ($t = 0$) level of pollution at each parcel, $p_{i,k,0} = x_{i,k}$. The pollution level at each parcel in successive iterations is determined solely by the level in the prior iteration and the physical model.

The pollutant emissions released into the river impose a cost on users downstream of parcel 7. This external cost is represented by a quadratic social-damage function that is based on the total amount of pollution that reaches parcel 7:

$$D_k = \left[\sum_{t=1}^{200} p_{7,k,t}\right]^2 \approx \; 0.0053 * \left[\sum_{i=1}^{6} x_{i,k}\right]^2 . \quad (2)$$

This function implies that the marginal damage of any given firm is not spatially explicit.

This follows from the assumption that "damage" is based on the total amount of pollution traveling downstream rather than on the concentration of pollution at any particular point and that the stream flows fast enough that there is no significant reduction of biologically available nitrogen from settling or chemical reactions.[1]

The maximum private revenue that each firm can achieve occurs at a emission level of $x_{i,k} = 50$ while the socially optimal net benefits are achieved when total emission by all six firms is $\Sigma x_{i,k} = 150$, which corresponds to symmetric individual emission of $x_{i,k} = 25$, and a social damage $D_k = 120$. Note that the river system is dynamic but the firm's problem is not; in each round, the firm chooses a single level of emission.

2.2 Spatially Explicit River Dynamics

Similar to Miao et al. (2016), the nutrient-routing dynamics that define the physical model in this experiment were developed based on the QUAL2K model (Chapra et al., 2008). In the model, the total amount of pollution reaching the furthest downstream parcel is approximately equal to the total amount of pollution flowing in. Consequently, the river's dynamics do not *directly* affect the producer's problem. The function of the river dynamics in this experiment is to generate a characteristic pulse of pollutants that the regulator observes at the downstream monitoring point. From this pulse, the regulator attempts to infer each firm's individual pollution emissions, and the regulator may share information and/or assess penalties based on these measurements. Thus, the river's dynamics potentially *indirectly* affect the incentives facing individual polluters through the actions of the regulator.

---

[1] If we ease those assumptions, the marginal social damage from each parcel is heterogeneous, leading to an asymmetric social optimum, additional complications, and little if any additional insight.

The top panel of Figure 1 illustrates the characteristic "pulse" of pollution concentration measurements at one point over time. The effect of changes in parcel pollution on this pulse vary spatially. Conceptually, we leverage this difference in pulse contributions by individual parcels to estimate the amount of pollution emissions contributed by each parcel. If all of the parcels pollute, there will be a qualitatively similar pulse comprised of the aggregated individual pulses. We refer to the pollution emissions (which is a t-vector) from an individual parcel, i, as W_i.

2.3 Attribution Approach for the Estimated Information and Policy

Artificial neural networks are tools from machine learning (a field of computer science related to pattern recognition) used to generate approximations for arbitrary unknown nonlinear functions based on training data composed of observed inputs and outputs. These networks of compound linear and nonlinear functions and associated parameters minimize root mean squared error or other similar out-of-sample prediction errors.

Environmental modelers have successfully applied neural networks to several types of water pollution scenarios (Diamantopoulou et al., 2005; Singh et al., 2009; Kalin et al., 2010), but that work was focused on forecasting a single time series. We take a different approach that uses a series of feed-forward networks as a filter (Lo, 1994; Shtauss, 2008). We use our estimates of downstream polluters' emissions contributions and recursively filter out the contributions of parcels further upstream to decompose observed pollution fluxes from a single runoff event into emissions from specific upstream polluters.

We assume that the regulator has the capacity to collect high-frequency water-quality data; specifically, the regulator observes $p_{7,k,t,}$ in $t = 0, \dots, 200$. Based on the model parameters that we utilize, t is defined in 15-minute increments. We define N^i, the influence of parcel i on water quality, as a set of observations:

$$N^i = \{p_{7,k,t}\}_{t=\underline{t}_i}^{\bar{t}_i} \; s.t. \frac{\partial p_{7,k,t}}{\partial x_{i,k}} > \varepsilon \qquad (3)$$

in which $\underline{t}_i$ and $\bar{t}_i$ indicate the beginning and end periods during which a parcel's emissions have a significant impact on the pollution observed at the monitoring point. "Significant" in this case is defined by ε using a value of ε = 0.01 * max($\{p_{7,k,t}\}$). In other words, N^i is the set of end-of-stream total concentration measurements for which parcel i might have made a substantial contribution.

The network takes the observed pollution as an input and divides it into six subsets representing the emissions from each parcel. The structure of the network is displayed in Figure 3. The influence of the parcel closest to the sensor (N^6) feeds into a subnetwork that maps to the marginal contribution of that parcel to monitored pollution. Next, the output of that subnetwork and the next closest sensor (N^5) feed into a second subnetwork that maps to the marginal contribution of parcel five, and so forth. The subnetwork (S) associated with parcel i takes a weighted sum of the inputs (vector v_i) and passes it through several logistic transfer functions indexed by m:

$$S_{m,i}(v_i) = \frac{1}{1+e^{-v_i\beta}}. \qquad (4)$$

The weighted sum with weighting vector γ is mapped to an output vector representing the fitted value of parcel i's contribution to observed pollution over time,

$$\widehat{W_i} = \sum_m \gamma S_{m,i}, \qquad (5)$$

and the vector over time of errors from a subnetwork is defined as a standard residual:

$$e_i = W_i - \widehat{W_i} \quad . \qquad (6).$$

For the first subnetwork, the input vector is $v_1 = N^6$ and the target or output vector is W_i. Each subnetwork after the first takes both the coverage associated with that parcel and the fitted contribution from the prior subnetwork: $v_i = [N^{6-i+1}, \hat{w}_{l-1}]$ for $i = 2, \ldots, 6$. The number of logistic transfer functions ("hidden nodes"), m, is a key design parameter. For each subnetwork, a pruning algorithm is used that starts with twice as many hidden nodes as inputs to the layer. Nodes with the smallest marginal effect on prediction error are successively removed until the effect of the next node is greater than 0.01 of root mean squared error.

Training the artificial neural network consists of finding values for all βs and γs for all of the hidden nodes in all of the subnetworks that minimize prediction error based on a set of training data or on sets of input values (observed pollution) and known corresponding output values (W_i, the emissions from individual parcels). The network was synthetically trained (Lo, 1994) as the training data consisted of known inputs and target outputs generated from our hydrologic model rather than from empirical data. The training was implemented with a resilient backpropagation adaptive-learning algorithm (RPROP) (Riedmiller, 1994), using the Python package PyBrain (Shaul et al., 2010). To employ such a model in practice, this training would ideally be a multi-stage process involving both synthetic and empirical data.

The final output of the neural network is a set of individual curves representing the estimated emissions from each parcel. The first panel of Figure 1 provides an example of the output. The top curve is the measured pollution flux, the second set of curves is the actual marginal contribution of each parcel, and the final curve is the network's estimate of each parcel's emissions. In our case, the estimated emissions from each parcel is within 5–

10% of the actual value through the training process. Once the network is trained, a simple bootstrap is used to calculate error in the prediction for each parcel for the distributions of the emission decisions, as shown in Figure 4.[2]

3. Experimental Design

The experiments applied pollution tax policies that varied the extent to which parcel specific information on emissions was used. In each round, participants chose a production level between 0 and 50 and the pollution emissions entering the river from each parcel was equivalent to the production level[3]. For the purposes of the model, we assumed that emissions were in the form of ammonia, a major type of agricultural nitrogen runoff, and that the pollution concentration was measured in terms of total nitrogen. The distribution of the pollutant over time at any point in the river was determined by the joint decisions of all participants and excessive aggregate pollution levels triggered a tax. After participants experienced each of the treatments, described later in the section, they voted on the implementation of a costly policy that would increase the amount and accuracy of information available to both the regulator and the polluters in a final treatment. This portion of the experiment was designed to determine individual valuations for high-quality information.

The treatments interacted two types of information attributes; participant-

[2] When firm-specific emissions are highly symmetric (correlation of rho = 0.9), the error initially is quite small but increases for parcels further upstream. Interestingly, when correlation between emissions decreases (correlations of rho = 0.75, 0.50, and 0.35), the error increases more quickly for parcels downstream and decreases for parcels further upstream. The heterogeneity in the prediction error across firms suggests that under a policy that imposes a tax based on estimates such as we describe, a savvy emitter who is trying to extract rent can adjust emissions based on the parcel's location to manipulate the distribution of error.

[3] Note that individual participant decisions were framed in terms of "production." For consistency however we refer primarily to the "emissions" throughout this paper for consistency, since the two are equivalent.

information type (ambient, estimated, and exact) and policy-information type (none, ambient, estimated, and exact) (see Table 1) and were varied in a four-by-three experimental design in which two of the off-diagonal treatments (exact information / estimated policy and estimated information / exact policy) were dropped, resulting in ten total treatments. Those specific treatments were omitted because they combined conflicting information attributes (exact versus estimated information) for the participant and regulator, which seem unlikely to occur in practice.[4] To control for potential order effects, the order of the treatments (labeled A through J) was varied across experimental sessions using Latin squares that were blocked at the policy level.

The participant-information treatments consisted of ambient, estimated, and exact information. Under the ambient information treatment, the participants knew only about their own emission decision, revenue, and tax (if any) and the total ambient (group) level of pollution. They received no information about the decisions of others in the group.

Under the estimated information treatment, participants received all of the same information as in the ambient treatment, plus the estimated pollution emissions of all players according to the attribution approach previously described. Since they were informed of the amount of pollution produced by their emission decision, they could see any error in the regulator's estimate of their own emissions.

Under the exact information treatment, participants received all of the information from the ambient treatment plus the exact pollution output for each player, emulating a scenario with a fully dense sensor network. In other words, in the exact information treatment it was assumed that emissions could be attributed to individuals, similar to a

[4] In the first case, the regulator would have to intentionally limit its own information; in the latter, the regulator would choose to provide emitters with only some of the available information.

point-source setting.

The four policy-information treatments varied the type of tax policy that participants faced: no policy, an ambient policy, an estimated policy, and an exact policy. The threshold for excessive pollution in the three policy-information treatments that involved a tax was 120 units of damage for the group. Under the no policy treatment, taxes were not levied and all participants were expected to produce the revenue maximizing number of production/emissions units (50).

Under the ambient policy treatment, the tax charged to participants in each group was identical and was a function of total damage. Recall from equation 2 that damage, $D_k \approx 0.0053 * \left[\sum_{i=1}^{6} x_{i,k}\right]^2$, so that total emission of 150 corresponds to damage of 120:

$$tax_{Ambient}(D_k) = \begin{cases} 0, & D_k < 120 \\ 0.37 * (D_k - 120), & D_k \geq 120 \end{cases}. \quad (7)$$

Under the estimated policy treatment, the taxes were based on estimates of the individual participant's contribution to pollution, $\tilde{x}_{l,k}$, when total damage (D_k) exceeded the threshold of 120:

$$tax_{Estimated}(\tilde{x}_{l,k}, D_k) = \begin{cases} 0, & \tilde{x}_{l,k} < 25 \; OR \; D_k < 120 \\ 0.37 * (\tilde{x}_{l,k} - 25), & \tilde{x}_{l,k} \geq 25 \; AND \; D_k \geq 120 \end{cases}. \quad (8)$$

Under the exact policy treatment, the threshold for damage was the same but the tax was a function of the true value of the individual participant's emission decision, $x_{i,k}$:

$$tax_{Exact}(x_{i,k}, D_k) = \begin{cases} 0, & x_{i,k} < 25 \; OR \; D_k < 120 \\ 0.37 * (x_{i,k} - 25), & x_{i,k} \geq 25 \; AND \; D_k \geq 120 \end{cases}. \quad (9)$$

The marginal tax rate of 0.37 achieves the symmetric Nash equilibrium level of individual emission at the target 25 units for the individual taxes. We use the same formulation for the ambient tax to maintain consistency across treatments, however the marginal ambient tax rate would be non-constant with emissions, and is over-powered, potentially pushing emissions below the target level. This does not have any impact on the interpretation of information treatment effects in the results, but we do expect a negative average effect of the ambient treatment relative to the targeted treatments. If the marginal damage is differentiated across parcels based on their physical location, the tax rate in the pollution-targeting policies must be differentiated accordingly.

Following all ten treatments, participants were given the opportunity to vote to determine which treatment would be implemented in the final part of the experiment. In a series of six yes/no votes, the participants chose between just having the treatment with ambient participant information and ambient policy (Treatment D) or paying a set price ($0, $5, or $10)[5] and to instead face one of the following two policy treatments:

Estimated participant information and estimated policy (Treatment H); or

Exact participant information and exact policy (Treatment J).

We expected participants who preferred the ambient treatment to vote against policies that used more exact information regardless of the price.

After all votes were made, one of the six vote options was randomly selected to determine the outcome in the final treatment. For the option selected, when a simple majority (four or more) of the voters favored a given treatment, that treatment was adopted for the final rounds of the experiment and all participants in the group paid the associated price. When there was no majority of yes votes, the ambient treatment was implemented and the participants did not have to pay a price.

[5] No exchange rate was used for the voting so that the vote involved potential payments of US Dollars.

3.1 Experimental Protocol

We conducted seven experimental sessions, six with undergraduate students and one with agricultural producers. The undergraduate student sessions were conducted at a large public university on the East Coast of the United States and involved 60 students with 18 or 24 students participating in each session. The session with agricultural producers was conducted during an extension event at a regional agricultural meeting and involved 24 participants. All of the participants' decisions were made on networked computers. The experiment was implemented in Python and incorporated the QUAL2K stream dynamics using the numerical analysis package NumPy and EconWillow (McCabe and Weel, 2010) for the web-browser-based user interface.

Participants in the experiment were randomly assigned to a computer by drawing a numbered ball before entering the laboratory. Groups of six participants were assigned to an independent river system, and the participant's group and parcel location were randomly shuffled between treatments. Each treatment consisted of six decision rounds, and participants made one emission decision in each round that determined their revenue and pollution emissions. The information provided to participants at the end of the round varied by treatment, but all participants were informed about their own revenue, the group's total measured pollution, the tax imposed on them (if any), and their net profit from the round. The rounds were independent from the standpoint that the emissions decisions made in one round did not affect future rounds.

At the beginning of the experiment, subjects were given up to twenty minutes to read the instructions for the first treatment (see Appendix A). The administrator then gave an oral presentation that emphasized how participants' decisions affected their revenue and the ambient tax mechanism.

To help participants understand the experiment and formulate their emission decisions, the software included a calculator that allowed participants to enter hypothetical emission decisions for their parcel and for the other five parcels in their group. Using these inputs, the calculator would generate the average tax and profit and varied with the policy treatment. The written instructions included a guided training for using the calculator in which each participant entered several decisions and the participant and experiment administrator reviewed the outcomes.

The experiment consisted of five practice rounds of the baseline treatment and then six actual rounds for each of the ten treatments listed in table 1. Following this, participants completed six voting rounds, followed by one more treatment of six rounds based on the voting results, for a total of 77 rounds. On average, participants completed the experiment in one hour and 45 minutes. Their profits in each round consisted of their total private revenue minus any tax imposed by the regulator. Payments for the session were calculated based on the total profit earned across all rounds. The exchange rate for student sessions was $1 US for 45 experimental dollars, while the producer session was $1 US for 25 experimental dollars. Student participants earned an average of $29 at the end of the experiment, while the agricultural producers earned an average of $50.

4. Results and Discussion

In this section we analyze the results of the experimental sessions. The analysis focuses on assessing differences across treatments in the emissions decisions made by individual participants and on preferences for the various policies as revealed by the voting decisions.

In particular, we are interested in addressing three specific outcomes related to the effect of information on behavior under the policy treatments. First, we explore the extent to which more accurate emissions information influences participant decision-making and pollution outcomes. Next, we use the observed individual emissions decisions to estimate models that evaluate the mechanisms that may underlie the differences in behavior across treatments and spatial location. Finally, we use the results of the voting portion of the experiment to investigate differences in preferences across the information policy treatments by parcel location and participant type.

We first compare individual participants' emission decisions relative to the social target. The graphical and econometric models that we present below pool the decisions of students and agricultural producers. In terms of decision-making related to emissions, we expected the behavior of the students and agricultural producers to be similar – an assumption that we test and fail to reject in section 4.1. Since voting captures the effects of relatively subjective elements, such as information and social preferences, we did not have expectations related to differences in voting behavior across the two groups.

4.1 Strategic Emissions Decisions

The targeted, symmetric level of production/emissions was 25 units per individual (150 units per group). Under the individual policies carried out in the experiment, emission of 25 units was also the symmetric Nash equilibrium. Figure 5 provides a graphical depiction of the number of emission units associated with each policy treatment averaged across the policy-information treatments. As expected, participants in the no policy treatment choose to produce close to the maximum of 50. Emission under the ambient policy is closest to the target level while the estimated targeted policy and exact targeted policy lead to correspondingly higher levels of emission. While we expect that the ambient policy would be lower, we also expect that in that targeted treatments, improved information would lead to outcomes that were more efficient. This result suggests that this is not necessarily the case.

Figure 6 further breaks down the outcomes, showing emissions under each policy by information treatment. The emission levels that diverge the most from the symmetric NE tend to occur in the treatments that provide participants with parcel-specific emission information. Specifically, emission levels are highest when exact information about emission decisions are provided to other participants. Potential explanations for these inferior outcomes could include spite (e.g., Cason et al., 2002), where participants who saw that others were overproducing chose to overproduce themselves to "get back" at the others. Another possibility is strategic compensation; in response to under-emission by others, a participant may have chosen to increase emissions to take advantage of the "available" emission units below the target. This type of behavior would tend to adjust towards the equilibrium. If, however, there are multiple participants aggressively compensating for "slack" emission below the threshold created by another individual's under-emission this could lead to short-term over-emission.

To more formally evaluate differences in outcomes across treatments, Table 2 reports the results of two-limit Tobit models of the individual emission decisions with participant fixed effects. There are observations at both ends of the uncensored range. The upper end, with participants producing the maximum, is typical censoring; their desired emission was likely higher and they only emitted at the maximum because of the emission constraint. The lower level observations, where participants chose emissions of zero, are less clear-cut. It could be that participants are being taxed due to others' over-emission and would be willing to try the strategy of negative emissions to avoid taxation, which would be a case of censoring. They may also have "opted-out," the decision to emit zero being an exact strategy, regardless of other emission options available. If this were the case, a Cragg double hurdle model would be more appropriate. Since the number of zero censored observations is so small in this case (8 observations out of 2760 for the non-lagged model), we opted to apply the Tobit model for simplicity. A Chow composite likelihood ratio test between Model A, described in the following paragraphs, estimated on the full sample, and models estimated using only the farmer or student participants failed to reject a statistical difference in emission decisions ($\chi^2(8) = 6.69$, $p = 0.43$).

The primary variables reported in all models are "Tax" and "Info," each of which has levels "Exact" and "Estimated." The Tax variable corresponds to the policy-information treatments, with a value of "Estimated" indicating that the tax was based on the neural filter estimation of individual emission (i.e. treatments G and H in table 1), while "Exact" assumes that there is a tax based on the exact observation of damages (treatments I and J), with the ambient tax (treatments D, E, and F) being the omitted case. The Info variable corresponds to the participant-information treatments, indicating the level of information provided to individuals. "Estimated" Info indicates that participants observed noisy estimates of others' production/emissions (treatments E and H) and "Exact" Info indicates perfect information about other participants' emission (treatments F and J). The omitted Info value is ambient information, in which participants observe only their own decisions and the total damage (treatments D, G, and I). The interactions specifically give the marginal increase of both participant-information and policy-information above having either individually (treatments H and J). The constant represents the treatment in which both regulators and participants observe only the ambient damage measurement (treatment D). All coefficients are emission units in the latent uncensored variable, equivalent to a change in emission units in the observed variable weighted by the probability of censoring.[6]

Model A includes only indicator variables for the participant-information and policy-information treatments, as well as the interaction between the treatments. Based on the estimated model coefficients, we find that emission is significantly higher (2.85 emission units in the latent uncensored variable, $P < 0.001$) with the exact policy-information tax policy, regardless of the participant-information. We also find that emission is higher at the 5% significance level with the estimated policy-information, but only when combined with estimated participant-information (1.43 emission units, $P = 0.019$). In other words, participants always respond to the exact tax policy with higher emissions, but only respond to the estimated tax when additional information is available on others' emissions. The participant-information treatments on their own do not have a significant impact on emitting behavior when they are implemented in conjunction with the ambient policy-information treatments.

Models B include controls for prior round outcomes. This controls for persistence in decision-making, and also indicates differential dynamic effects of the participant-information treatment, potentially offering insight on the behavioral and strategic

[6] Note also that the No Tax treatments were excluded from this portion of the analysis. As one might expect, and as demonstrated in Figure 5, these had a large positive effect on emissions since the only potential repercussions of over-emissions were social. In a similar model looking just at the non-tax treatments (not reported) the increased information treatments showed a significant decrease in emissions of about 2 units.

motivation behind the effect of participant-information on emissions. The relative increase in over-emitting that results from increased information accuracy runs contrary to the expectation that this information would aid in improving coordination. We speculate above that this could be driven by spite or strategic heuristics. These models both include the set of "Tax" policy-information and "Info" participant-information variables included in Model A, but also interact these with two different variable specifications intended to capture the prior round's outcome.

Specifically, Model B includes the variable "Prior Gap" which is defined as the maximum emissions in the prior round over all group members, minus each individual's prior round emissions. If an individual was the highest producer in a round, Prior Gap is 0, while an individual whose emissions lagged far behind the high producer has a larger Prior Gap value. In this case, the un-interacted treatment variables indicate within-period treatment effects on the high producer in the prior round, while the interacted variables indicate relative changes in the rate of convergence towards (or away from, if negative) symmetric emissions.

The un-interacted (i.e., effect on the highest emitter) coefficient estimates for Model B show somewhat stronger significance of treatment effects on high emitters than we see across all emitters in Model A, although they are broadly consistent. We now see both estimated participant and tax information lead to a significant increase on emissions, while the exact tax information now also increases emissions when combined with exact participant information. To generalize a conclusion from this, it appears that the participants who are producing at high levels to start out with are using the available information to play even more aggressively – a story that supports that idea of strategic information use.

The Prior Gap variable on its own is negative and significant. This means that, on average, if a participant is producing lower in a round relative to the top emitters, they will produce *even lower* relatively speaking in the following round regardless of the information setting. This is offset somewhat, but not totally, in both of the estimated policy-information treatments, and in the exact policy- and participant-information treatment relative to the ambient policy treatments. Since they are lower emitters, the tax they face in the ambient case would likely be relatively higher than in the case where the tax is targeted based on the parcel specific information, so it may make sense that in those treatments they would more aggressively decrease emissions to avoid the tax. This also speaks to a strategic heuristic adjustment over spiteful behavior. For both high and low emitters we seem to see a short term effect away from the equilibrium, which is the behavior we expected to see from strategic heuristic overcompensation.

4.2 Willingness-to-Pay for Sensing Technology Upgrades

To assess the participants' preferences for the policy- and participant-information treatments, we analyzed the voting data using logit models in which willingness to pay (WTP) for increased information and policy treatments was calculated as the estimated parameter for the treatment variable divided by the estimated cost parameter. The point estimates for all participants' WTP by parcel as well as the 95% upper and lower bounds calculated by the Krinsky-Robb bootstrap method are reported in Table 3. The average WTP estimates are broken down by participant type. These results show that the students did not have significant willingness to pay for either the estimated or exact policy upgrades.

Interestingly, the agricultural producers did not have a significant willingness to pay for estimated information but had a significant and positive willingness to pay for the exact information and policy treatment. This is notable because producers also tended to produce the furthest away from the optimum under this treatment, so on average would have made less.[7] So the estimated positive cost to adopt the information technology would also imply lower earnings as a result of adopting the technology. Perhaps these results reflect the importance of "transparency," which has been cited as being particularly important to program managers (Messer et al., 2016).

The policy based on estimated policy-information is unbiased across parcels, but does have different levels of uncertainty. As we saw in Figure 4, parcels farther away from the sensor generally have higher RMS prediction error. The experiment was not designed specifically to focus on the effect of changes in uncertainty however we are able to see whether this changing uncertainty is represented in WTP for the estimated policies. If participants are risk averse we could expect that the estimated information would be less valuable for those further from the sensor.

Table 4 shows WTP estimates for all participants across the two treatments, differentiated by parcel. Indeed, there are differences in willingness to pay for the estimated participant- and policy-information treatment but not for the exact participant- and policy-information treatment. Specifically, parcels 4 and 6 have a positive and statistically significant WTP. Parcel 5 is not significantly different from zero, but also has one of the highest variances in WTP across the parcels. Since parcels 4 and 6 are among the

[7] A model of earnings is not included here because its conclusions are essentially analogous to the emissions model, and the emission model focuses more clearly on the strategic choice variable. However a linear regression of the treatments on earnings shows that the exact policy- and participant-information treatment is associated with a 4.9% decrease in average income (P-Value = 0.019).

closest to the sensor, their positive WTP estimates may represent participants' desire for a more-targeted policy so long as the emissions estimates are likely to be accurate.

6. **Conclusion**

This research explored integration of a high-frequency water-quality sensing technology into NPS pollution policies in a setting that used realistic nutrient-transport dynamics. The field of real-time water-sensing technologies is advancing rapidly and offers the potential to assuage some of the information asymmetry that lies at the heart of the NPS pollution problem. We used experiments involving both students and agricultural producers to test three types of information—ambient, estimated, and exact—related to water pollution. The ambient information represented the current state of water-quality sensing while the exact information represented the ideal of perfect monitoring at an individual farm level. The estimated water-pollution information was generated using an estimation technique based on an artificial neural network that acted as a recursive filter to spatially identify sources of pollution.

The results of the estimated information treatment are interesting not only because of the potential inaccuracy involved in the estimates but also because the precision of the estimates is asymmetric. Differences in information were incorporated into separate treatments that varied both the policy mechanism and the information provided to participants. After experiencing all of the treatments, participants voted on whether to pay for an upgrade from the baseline ambient information to the estimated and exact information treatments respectively. These voting data were used to estimate participants' willingness to pay for enhanced information.

In terms of the relative effectiveness of the policies, we find that more precise information does not reduce pollution relative to aggregate information; in fact, pollution under estimated and exact information exceeds the level of pollution under ambient information. This is not to say that the estimated and exact information approaches should be discounted. A spatially differentiated policy may be appealing for many reasons. Targeted "polluter pays" programs, for example, may be much more palatable politically than a flat ambient-based tax on all producers. In addition, the mechanisms applied here were structured specifically to avoid productivity losses from aggregate under-emission. A number of variations of the mechanism payout could be tested. The estimated and exact policies, as described in equations 8 and 9, involve a double threshold where both the individual and the group have to exceed pre-defined thresholds. It would be of particular interest to assess policy performance if only one or the other threshold were implemented.

The estimates of willingness to pay for more accurate information based on preferences for the final treatment are also interesting. The estimates of students' WTP generally were not significant, although those representing some downstream (and hence more precisely estimated) parcels had a positive value for the estimated information treatment. The students had no positive WTP for the exact information treatment. The agricultural producers did not significantly value the estimated information but had significant and positive WTP for the exact information. This result is particularly interesting since there is no evidence that they earned a higher profit or polluted less under the exact information treatment. They may have had professional experience that led them to an aversion to ambiguity or uncertainty or placed a social value on fairness and transparency that was not shared by the student population.

References

Alpizar, F., Requate, T., and Schram, A. 2004. Collective versus random fining: An experimental study on controlling ambient pollution. *Environmental & Resource Economics* 29(2): 231–252.

Arnold, M., J.M. Duke, and K.D. Messer. 2013. Adverse Selection in Reverse Auctions for Environmental Services. *Land Economics*. 89(3): 387-412.

Arthur, W.B. 1993. On designing economic agents that behave like human agents. Journal of Evolutionary Economics. 3: 1.

Athey S, G. Imbens. 2016. Recursive partitioning for heterogeneous causal effects. Proceedings of the National Academy of Science. 113:7353–7360.

Balcan, M.-F., Blum, A., Hartline, J., and Mansour, Y. 2007. Reducing Mechanism Design to Algorithm Design via Machine Learning." *Journal of Computer and System Sciences*, *special issue on Learning Theory*. Originally appeared in the 46th Annual IEEE Symposium on Foundations of Computer Science.

Camacho-Cuena, E., and Requate, T. 2012. The regulation of non-point source pollution and risk preferences: An experimental approach. *Ecological Economics* 73, 179–187.

Cason, T. N., Saijo, T., and Yamato, T. 2002. Voluntary participation and spite in public good provision experiments: an international comparison. *Experimental Economics*, 5(2), 133-153.

Cochard, F., Willinger, M., and Xepapadeas, A. 2005. Efficiency of Nonpoint Source Pollution Instruments: An Experimental Study. *Environmental and Resource Economics* 30(4): 393–422.

Dhillon, G. S., and S. Inamdar (2013), Extreme storms and changes in particulate and dissolved organic carbon in runoff: Entering uncharted waters?, Geophys. Res. Lett., 40, 1322–1327, doi:10.1002/grl.50306.

Diamantopoulou, M.J., Papamichail, D.M., and Antonopoulos, V.Z. 2005. The use of a Neural Network technique for the prediction of water quality parameters. *Operational Research*. 5(1):115–125.

Huang, J.J. and E.A. McBean. 2007. Using Bayesian statistics to estimate the coefficients of two-component second-order chlorine bulk decay model for a water distribution system. *Water Resources* 41(2): 287–294.

Inamdar, S., G. Dhillon, S. Singh, T. Parr, and Z. Qin (2015), Particulate nitrogen exports in stream runoff exceed dissolved nitrogen forms during large tropical storms in a temperate, headwater, forested watershed, J. Geophys. Res. Biogeosci., 120, 1548–1566, doi:10.1002/2015JG002909.

Kalin, L., Isik, S., Schoonover, J.E., and Lockaby, B.G. 2010. Predicting water quality in unmonitored watersheds using artificial neural networks. *Journal of Environmental Quality* 39(4):1429–1440.

Kirwan, B., R.N. Lubowski, and M.J. Roberts. 2005. "How cost-effective are land retirement auctions? Estimating the difference between payments and willingness to accept in the Conservation Reserve Program." *American Journal of Agricultural Economics* 87(5):1239–1247.

Kunwar P. S., Basant, A., Malik, A., Jain, G. 2009. Artificial neural network modeling of the river water quality—A case study. *Ecological Modelling* 220(6): 888–895.

Lo J. T.-H. 1994. Synthetic approach to optimal filtering. *IEEE Transactions on Neural Networks* 5: 803–811.

Logan, T.J. 1990. Agricultural Best Management Practices and Groundwater Protection.

Journal of Soil and Water Conservation 45(2): 201–206.

McCabe, K. and J. Weel (2010). Willow: Experiments in Python, Software (Version 1.0). Available from http://econwillow.sourceforge.net/.

Demirci, M. 2015 A Survey of Machine Learning Applications for Energy-Efficient Resource Management in Cloud Computing Environments, Machine Learning and Applications (ICMLA) 2015 IEEE 14th International Conference. 1185-1190.

Meran, G., and Schwalbe, U. 1987. Pollution control and collective penalties. *Journal of Institutional and Theoretical Economics* / Zeitschrift für die gesamte Staatswissenschaft, 616–629.

Messer K.D., W. Allen, M. Kecinski, and C. Chen. 2016. "Agricultural preservation professionals' perception and attitudes about cost-effective parcel selection methods." *Journal of Soil & Water Conservation.*

Miao, H., J. Fooks, T. Guilfoos, K.D. Messer, S. M. Pradhanang, J. Suter, S. Trandafir, E. Uchida. 2016. "The Impact of Information on Behavior Under an Ambient-based Policy for Regulating Nonpoint Source Pollution" *Water Resources Research*. 52: 3294-3308.

M. Mohri, M. and A. Munoz Medina. Learning theory and algorithms for revenue optimization in ˜ second-price auctions with reserve. In Proceedings of ICML, 2014.

Poe, G.L., W.D. Schulze, K. Segerson, J.F. Suter and C.A. Vossler. 2004. "Exploring the performance of ambient-based policy instruments when non-point source polluters can cooperate." *American Journal of Agricultural Economics* 86, 1203-1210.

Riedmiller, M. 1994. Advanced supervised learning in multi-layer perceptrons: From backpropagation to adaptive learning algorithms. *Computer Standards and Interfaces* 16(5): 265–278.

Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., Rückstieß, T., and Schmidhuber, J. 2010. Pybrain. *Journal of Machine Learning Research*. 743–746.

Segerson, K. 1988. Uncertainty and Incentives for Nonpoint Pollution Control. *Journal of Environmental Economics and Management* 15(1): 87–98.

Shen, J. and Z. Yuan. 2009. A Bayesian approach for estimating bacterial nonpoint source loading in an estuary with limited observations. *Journal of Environmental Science and Health, Part A* 44(14).

Shortle, J., and Horan, R.D. 2013. Policy Instruments for Water Quality Protection. *Annual Review of Resource Economics* 5(5): 111–138.

Shortle, J.S., and Horan, R.D. 2001. The economics of nonpoint pollution control. *Journal of Economic Surveys* 15(3): 255–289.

Shtrauss, V. 2008. Nonlinear decomposition filters with neural network elements *Proc. 12th WSEAS International Conference on Systems, Heraklion, Greece* 603–608.

Smith, R.A., Schwarz, G.E., and Alexander, R.B., 1997. Regional interpretation of water-quality monitoring data. *Water Resources Research* 33(12): 2781–2798.

Spraggon, J. 2002. Exogenous targeting instruments as a solution to group moral hazards. *Journal of Public Economics*. 84(3): 427–456.

Spraggon, J.M. 2013. The impact of information and cost heterogeneity on firm behavior under an ambient tax/subsidy instrument. *Journal of Environmental Management*, 122: 137–143.

Suter, J.F., Vossler, C.A., Poe, G.L., and Segerson, K. 2008. Experiments on damage-based ambient taxes for nonpoint source polluters. *American Journal of Agricultural*

Economics 90(1): 86–102.

Suter, J.F., and Vossler, C.A. 2013. Towards an understanding of the performance of ambient tax mechanisms in the field: Evidence from Upstate New York dairy farmers. *American Journal of Agricultural Economics* 96(1): 92–107.

Tagiew, R. "Mining determinism in human strategic behavior", EEML. KU-Leuven, pp. 86-91, 2012.

Tagiew, R., D. Ignatov, F. Amroush, 2013.: Social learning in networks: Extraction of deterministic rules. In: Experimental Economics and Machine Learning. ICDM Workshops, IEEE Computer Society (2013)

Terna, P. 2000a. Economic Experiments with Swarm: a Neural Network Approach to the Self-Development of Consistency in Agents' Behavior, F. Luna and B. Stefansson (eds.)Economic Simulations in Swarm: Agent-Based Modelling and Object Oriented Programming. Dordrecht and London: Kluwer Academic.

van der Hoog, Sander. 2016. Deep Learning in Agent-Based Models: A Prospectus. Bielefeld Working Papers in Economics and Management No. 02-2016. Available at SSRN: https://ssrn.com/abstract=2711216

Varian, Hal R. 2014. Big Data: New Tricks for Econometrics. Journal of Economic Perspectives. 28(2): 3–28.

Vossler, C.A., Poe, G.L., Schulze, W.D., and Segerson, K. 2006. Communication and incentive mechanisms based on group performance: An experimental study of nonpoint pollution control. *Economic Inquiry* 44: 599–613.

Xepapadeas, A. 2011. The economics of non-point-source pollution. *Annual Review of Resource Economics* 3: 355–373.

Table 1. Within Subject Policy Treatments

	Ambient Information	Estimated Information	Exact Information
No Policy	A	B	C
Ambient Policy	D	E	F
Estimated Policy	G	H	—
Exact Policy	I	—	J

Table 2. Fixed Effects Tobit Model of Emission Decisions

	Model A		Model B	
	Coef	P-Val	Coef	P-Val
C	36.78	0.000	41.97	0.000
Tax				
Estimated	0.53	0.251	3.13	0.001
Exact	2.85	0.000	1.27	0.041
Info				
Estimated	-0.29	0.476	1.32	0.043
Exact	0.06	0.872	0.37	0.508
Tax#Info				
Estimated	1.43	0.019	-2.23	0.590
Exact	0.17	0.808	3.42	0.003
Prior Gap			-0.36	0.000
Prior Gap#Tax				
Estimated			0.00	0.982
Exact			0.17	0.000
Prior Gap#Info				
Estimated			-0.05	0.437
Exact			0.06	0.264
Prior Gap#Tax#Info				
Estimated			0.17	0.051
Exact			-0.11	0.098
N	2760		2208	
Left Censored	8		6	
Right Censored	140		112	
Pseudo-R2	0.068		0.100	

Table 3. Willingness to Pay for Information Upgrades Over Ambient by Type of Information and Participant

	Students	Ag. Producers
Upper 95% Bound	4.36	2.61
Estimated Policy and Info	-0.14	0.37
Lower 95% Bound	-6.10	-2.43
Upper 95% Bound	4.74	5.22
Exact Policy and Info	2.01	3.00
Lower 95% Bound	-2.08	1.01
Participants	60	24
Observations	360	144

Table 4. All Participants' Willingness to Pay for Information Treatment Upgrades Over Ambient by Parcel

	Parcel 1	Parcel 2	Parcel 3	Parcel 4	Parcel 5	Parcel 6	Overall
Upper 95% Bound	11.27	17.84	8.05	**16.40**	14.41	**9.91**	**5.34**
Estimated Policy and Info	3.12	-0.03	2.93	**5.22**	3.44	**4.92**	**3.00**
Lower 95% Bound	-8.56	-12.85	-5.46	**0.37**	-10.13	**1.70**	**1.07**
Upper 95% Bound	13.11	8.95	11.15	1.62	8.71	1.28	2.61
Exact Policy and Info	2.54	2.32	2.07	-3.76	-0.97	-1.61	0.37
Lower 95% Bound	-3.85	-3.07	-3.78	-17.58	-12.64	-6.70	-2.40

Figure 1. Flux Measurement, Actual Decomposition, and Estimated Decomposition

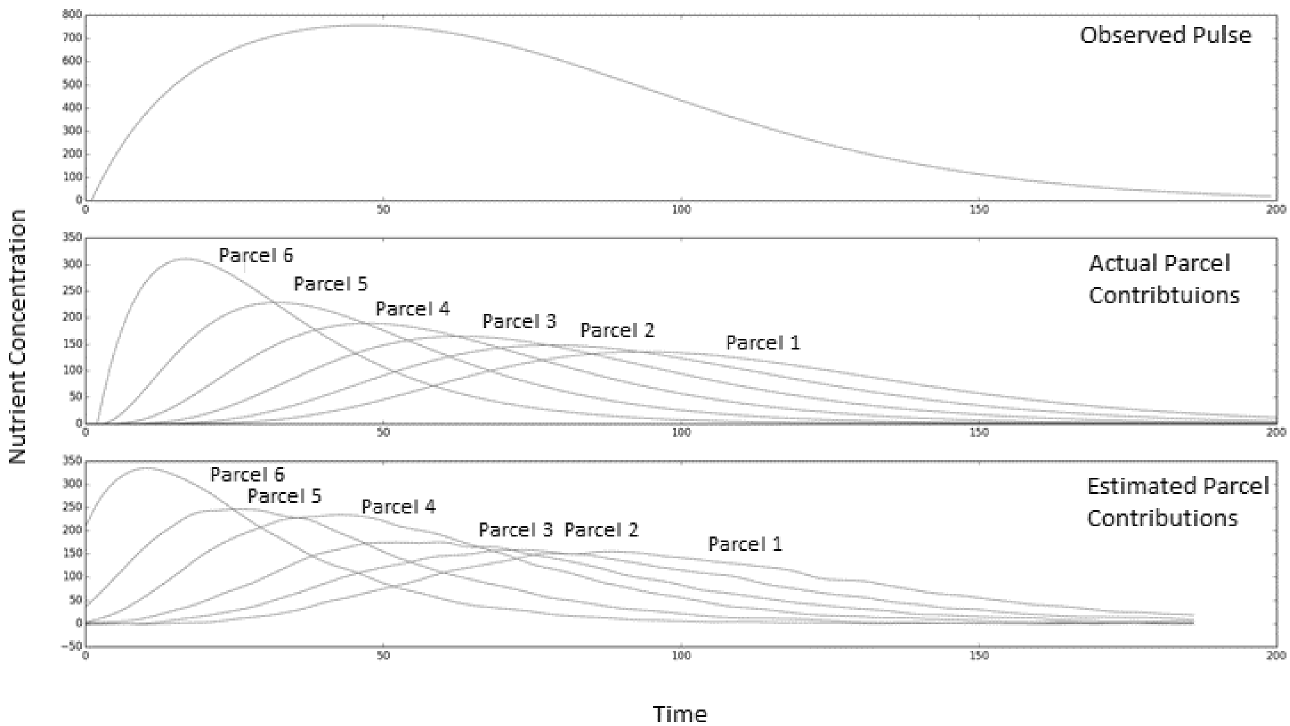


Figure 2. Parcel Map

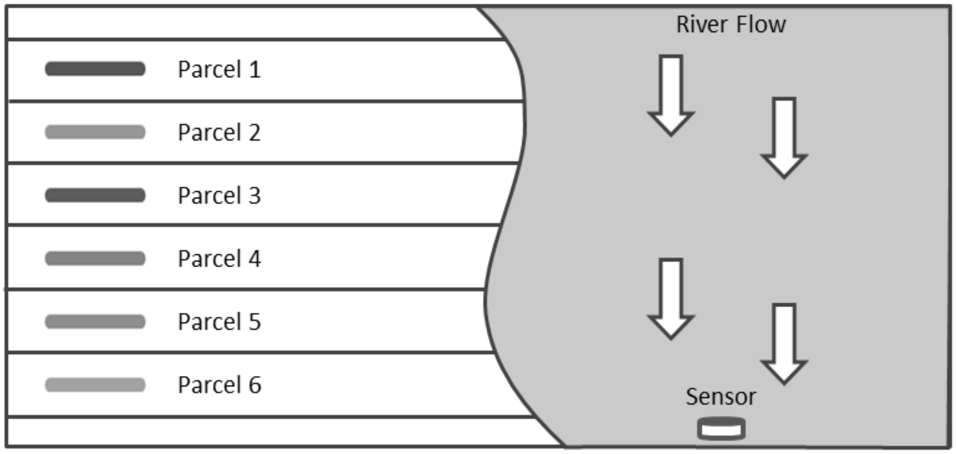


Figure 3. Artificial Neural Network Decomposition Filter Structure

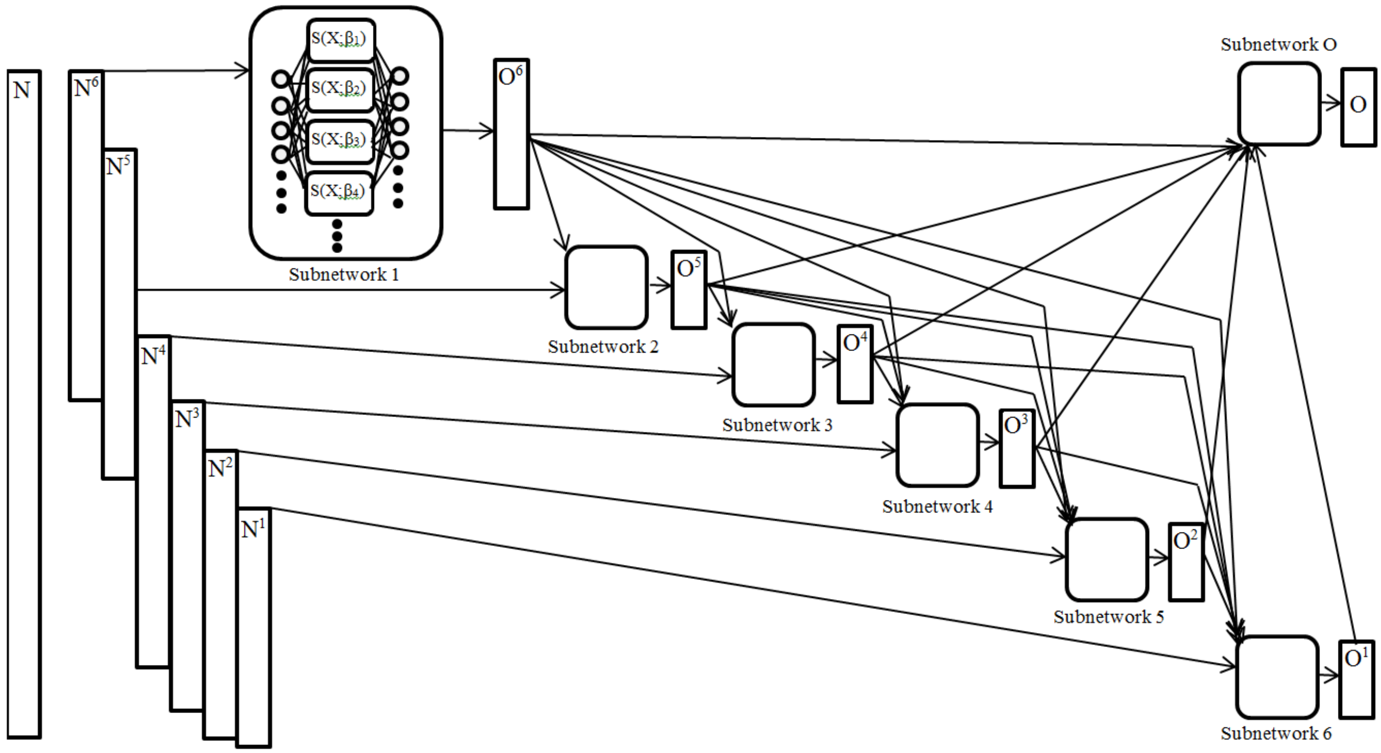


Figure 4. Root Mean Square Prediction Error by Correlation across Parcels

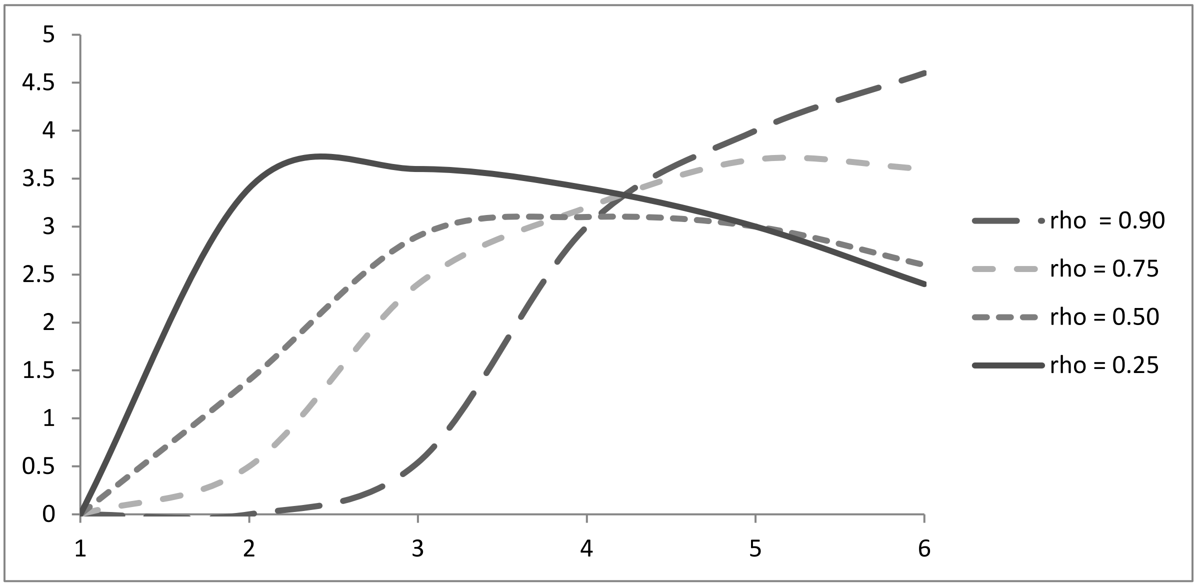


Figure 5. Observed Production/Emission Decisions by Policy Treatments

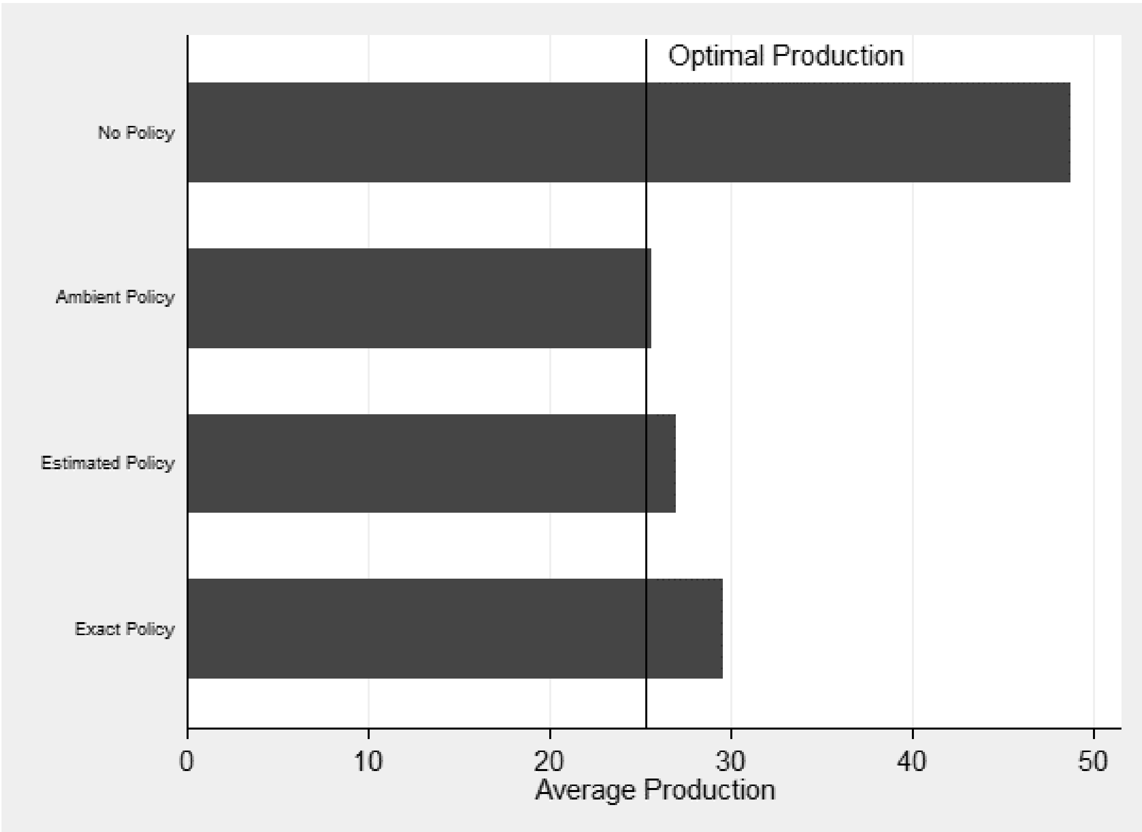
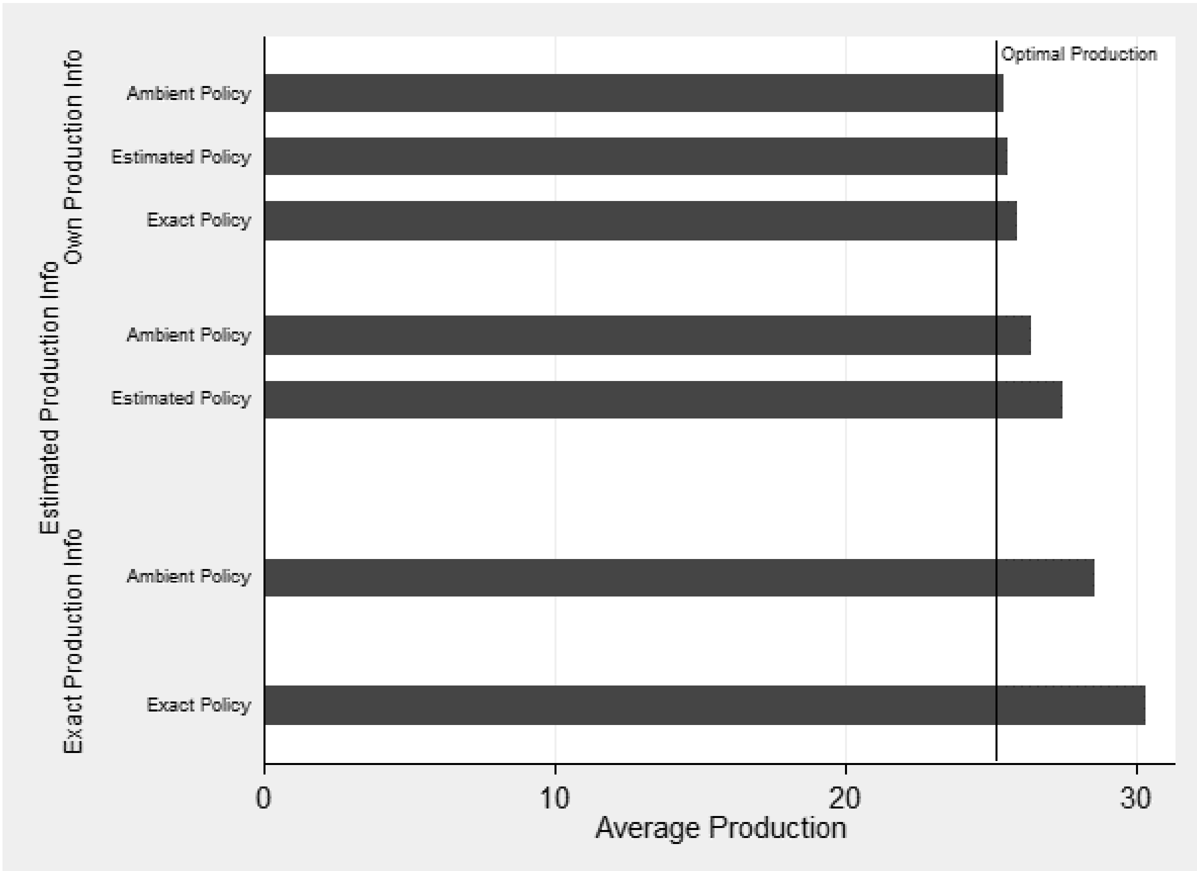


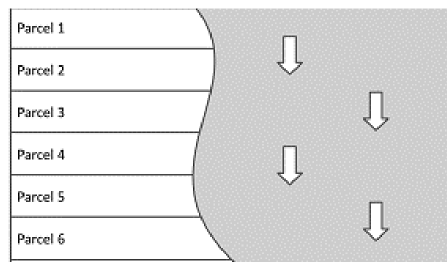
Figure 6. Observed Production/Emission Decisions by Policy and Information Treatments



Appendix A: Instructions

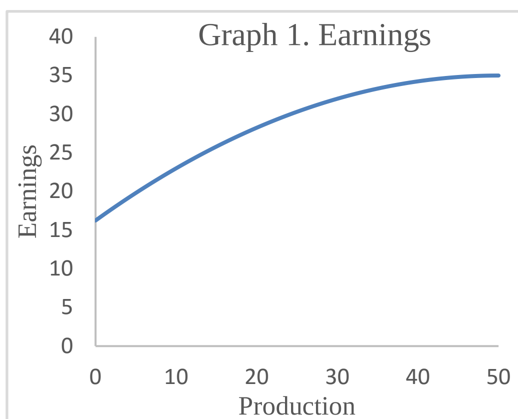
Welcome to an experiment about the economics of decision making. In the course of the experiment, you have the opportunity to earn money. Any money earned during this experiment will initially be recorded as **experimental dollars.** At the end of this experiment, we will convert your experimental dollars into actual **US dollars** at a rate of $1 US dollar for 45 experimental dollars. This amount will be given to you as you leave. The more experimental dollars you earn the more actual US dollars you will receive at the end of the experiment. Please read these instructions carefully and do not communicate with any other participant, unless informed by the administrator that communication is permissible.

In today's experiment you will be in a group of six players. Each player assumes the role of a business owner operating on a parcel of land along a river. The parcels are labeled Parcel 1 through 6, as displayed on the map to the right. Parcel 1 is the furthest upstream and Parcel 6 is the furthest downstream. The actual parcel that you operate will be indicated to you on your computer screen.

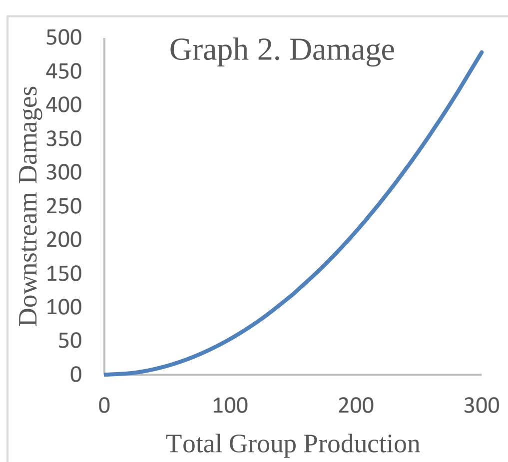


Each player must make decisions for their parcel over a number of decision **rounds**. Each round is independent from all other rounds, meaning that decisions made during a round do not affect future rounds in any way. A group of five decision rounds will be called a **part**. You will receive new instructions at the beginning of each part. The only value that is tracked across rounds is your cumulative **profit**, which will be used to calculate your cash payment at the end of the experiment.

In each round, your task is to decide the level of **production** between 0 and 50 units of output on your parcel. Your earnings depend only on how much output you produce. Your earnings will vary between 16.25 and 35.00 experimental dollars, as shown in Graph 1. The relationship between production and earnings generated are the same for everyone in your group.



Your production also generates **byproduct** that goes into the river. The higher the output that you produce, the more earnings you receive, and the more byproduct you generate. This byproduct does not affect you or others in your group, however too much byproduct can cause **damage** further downstream. The amount of downstream damage depends on the byproduct emitted by all six parcels, and varies between $0 and $480, as



shown on Graph 2. The dollar value here represents losses to some downstream firm in terms of lost productivity or increased costs due to the byproduct in the water.

To help you formulate your strategy, there is a calculator on your computer which will allow you to calculate profit for each parcel and the downstream damage for hypothetical sets of production for all six parcels. This calculator will be available to you throughout the experiment so that you can explore different strategies without it affecting your actual earnings. You can enter production decisions for each parcel by typing them directly into the column labeled "Production" or you can change production by using the slider for each parcel.

To help you understand how the calculator works, please complete the table below using the calculator provided. Note that for these examples all parcels produce the same amount. This need not generally be the case. The administrator will review the information featured in these tables with you.

A round is complete after all players submit their production decisions. The computer will automatically calculate the results for your group, and report your earnings from production, the downstream damage, and your **profit** for the round. At the end of the experiment, your earnings will be the sum of the profits you earned from all of the rounds.

The first part (five rounds) will be strictly for practice and will give you an opportunity to familiarize yourself with the software. These first five rounds <u>will not</u> result in any cash earnings. You will receive no information beyond your own earnings and the total downstream damage.

Example	Production for all Parcels	Profit for all Parcels	Downstream Damage
A	0	$16.25	$0.00
B	15		
C	35	$33.31	$234.61
D	50		

<u>Instructions – Part A</u>

<u>Information:</u> In this part, you will have information on your production, earnings, and profit, as well as the total amount of downstream damage. You will have no further information on others' production decisions.

<u>Fine:</u> There is no fine. Your profit will be based only on your production, and is not affected by the amount of byproduct.

Instructions – Part B

Information: In this part, you will have information on your production, and the total amount of downstream damage. You will also be provided with an estimate of the amount of production from all other parcels. These estimates are obtained using measurements of the byproduct just downstream of Parcel 6. Since the measurements are taken close to Parcel 6, they will be very accurate for that parcels, but may be less accurate for parcels higher upstream. The estimate for Parcel 1 will likely be correct within 15 unit, the estimate for Parcel 3 production will likely be correct within 5 units of production, while the estimate for Parcel 4 will usually be correct within one units of the correct value.

Fine: There is no fine. Your profit will be based only on your production, and is not affected by the amount of byproduct.

Instructions – Part C

Information: In this part, you will have information on your production, and the total amount of downstream damage. You will also be told the amount of production from all other parcels. This information will be an exact measurement, not an estimate.

Fine: There is no fine. Your profit will be based only on your production, and is not affected by the amount of byproduct.

Instructions – Part D

Information: In this part, you will have information on your production, earnings, and profit, as well as the total amount of downstream damage. You will have no further information on others' production decisions.

Fine: In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will be the same for everyone in your group, and will be based on the total downstream damage. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120:	Fine = 0.37*(Total Damage – 120)

In other words, if the total downstream damage is below 120 (which approximately corresponds to total group production of 150) there will be no fine. If damage is above 120, everyone in the group will pay a fine of 0.37 for every unit of damage above 120. Recall that the downstream damage is based on the decisions of everyone in your group, not just your production. Therefore your production decision will influence the profits of everyone in your group.

<u>Instructions – Part E</u>

<u>Information:</u> In this part, you will have information on your production, and the total amount of downstream damage. You will also be provided with an estimate of the amount of production from all other parcels. These estimates are obtained using measurements of the byproduct just downstream of Parcel 6. Since the measurements are taken close to Parcel 6, they will be very accurate for that parcels, but may be less accurate for parcels higher upstream. The estimate for Parcel 1 will likely be correct within 15 unit, the estimate for Parcel 3 production will likely be correct within 5 units of production, while the estimate for Parcel 4 will usually be correct within one units of the correct value.

<u>Fine:</u> In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will be the same for everyone in your group, and will be based off of the total downstream damage. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120:	Fine = 0.37*(Total Damage – 120)

In other words, if the total downstream damage is below 120 (which approximately corresponds to total group production of 150) there will be no fine. If damage is above 120, everyone in the group will pay a fine of 0.37 for every unit of damage above 120. Recall that the downstream damage is based on the decisions of everyone in your group, not just your production. Therefore your production decision will influence the profits of everyone in your group.

Instructions – Part F

Information: In this part, you will have information on your production, and the total amount of damage. You will also be told the amount of production from all other parcels. This information will be an exact measurement, not an estimate.

Fine: In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will be the same for everyone in your group, and will be based off of the total downstream damage. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120:	Fine = 0.37*(Total Damage – 120)

In other words, if the total downstream damage is below 120 (which approximately corresponds to total group production of 150) there will be no fine. If damage is above 120, everyone in the group will pay a fine of 0.37 for every unit of damage above 120. Recall that the downstream damage is based on the decisions of everyone in your group, not just your production. Therefore your production decision will influence the profits of everyone in your group.

<u>Instructions – Part G</u>

<u>Information:</u> In this part, you will have information on your production, earnings, and profit, as well as the total amount of downstream damage. You will have no further information on others' production decisions.

<u>Fine:</u> In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will NOT be the same for everyone in your group, and will be based off of the total downstream damage. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120 AND Your Estimated Production is above 25:	Fine = 0.37*(Estimated Production – 25)

In other words, if the damage is below 120, or your estimated total production is below 25 there will be no fine. If your estimated production is above 25, and total damage exceeds 120 you will pay 0.37 for every unit by which your production exceeds 25. Your profit in each round will be your earnings minus the fine. Note that the fine is based only on an estimate of your individual production. These estimates are obtained using measurements of the byproduct just downstream of Parcel 6. Since the measurements are taken close to Parcel 6, they can be most inaccurate. The estimate for Parcel 6 will likely be correct within 15 or more units. In contrast for Parcels 1-5, the estimate will be more accurate. Furthermore, the estimate will be more accurate for the higher number parcels (more downstream) and less accurate for the lower number parcels (more upstream). For example, the estimate for Parcel 5 production will likely be correct within one unit of production, while the estimate for Parcel 3 will usually be correct within 5 units of the correct value.

Instructions – Part H

Information: In this part, you will have information on your production, and the total amount of downstream damage. You will also be provided with an estimate of the amount of production from all other parcels. These estimates are obtained using measurements of the byproduct just downstream of Parcel 6. Since the measurements are taken close to Parcel 6, they will be very accurate for that parcels, but may be less accurate for parcels higher upstream. The estimate for Parcel 1 will likely be correct within 15 unit, the estimate for Parcel 3 production will likely be correct within 5 units of production, while the estimate for Parcel 4 will usually be correct within one units of the correct value.

Fine: In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will NOT be the same for everyone in your group, and will be based off of your individual production. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120 AND Your Estimated Production is above 25:	Fine = 0.37*(Estimated Production – 25)

In other words, if the damage is below 120, or your estimated total production is below 25 there will be no fine. If your estimated production is above 25, and total damage exceeds 120 you will pay 0.37 for every unit by which your production exceeds 25. Your profit in each round will be your earnings minus the fine. Note that the fine is based only on an estimate of your production. These estimates are obtained using measurements of the byproduct just downstream of Parcel 6. Since the measurements are taken close to Parcel 6, they will be very accurate for that parcels, but may be less accurate for parcels higher upstream. The estimate for Parcel 1 will likely be correct within 15 unit, the estimate for Parcel 3 production will likely be correct within 5 units of production, while the estimate for Parcel 4 will usually be correct within one units of the correct value.

<u>Instructions – Part I</u>

<u>Information:</u> In this part, you will have information on your production, earnings, and profit, as well as the total amount of downstream damage. You will have no further information on others' production decisions.

<u>Fine:</u> In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will NOT be the same for everyone in your group, and will be based off of your individual production. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120 AND Your Production is above 25:	Fine = 0.37*(Production – 25)

In other words, if the damage is below 120, or your total production is below 25 there will be no fine. If your production is above 25, and total damage exceeds 120 you will pay 0.37 for every unit by which your production exceeds 25. Your profit in each round will be your earnings minus this fine.

Instructions – Part J

Information: In this part, you will have information on your production, and the total amount of damage. You will also be told the amount of production from all other parcels. This information will be an exact measurement, not an estimate.

Fine: In this part, all players in your group may receive a fine. Your profit in each round will be your earnings minus this fine. This fine will NOT be the same for everyone in your group, and will be based off of your individual production. The fine will be determined as follows:

If Total Damage is 120 or less:	Fine = 0
If Total Damage is above 120 AND Your Production is above 25:	Fine = 0.37*(Production – 25)

In other words, if the damage is below 120, or your total production is below 25 there will be no fine. If your production is above 25, and total damage exceeds 120 you will pay 0.37 for every unit by which your production exceeds 25. Your profit in each round will be your earnings minus this fine.

<u>Instructions – Part K</u>

In this part, by default, everyone in your group will have information only on their own production and total damage, and will face a fine based on total damage, exactly the same as in **Part D**. We will call this Policy 1. However, the group has the opportunity to vote to "upgrade" to a production-based policy at an additional cost which will be paid by each person in the group. The possible policies in this part are:

Policy 1: Own Production Information and Damage Based Fine (as in **Part D**)
Policy 2: Estimated Production Information and Estimated Production Based Fine (as in **Part H)**
Policy 3: Exact Production Information and an Exact Production Based Fine (as in **Part J)**.

You will be given a series of choices on your computer. For each of these you can vote to keep Policy 1, or to upgrade to a different policy at some cost. Everyone in your group will have to face the same set of choices. After everyone has made decisions for each of these choices one of them will be chosen at random as the decision that is implemented. For the decision that is implemented, if more than three parcels in your stream group vote to adopt the new production-based policy, then in this part that policy will be in effect and everyone in the group will have to make a one-time pay payment of the associated cost. This cost will only be paid if the production-based policy is adopted. Note that while you are voting, the parcel that you will control for this part is listed on your computer. Once the voting has concluded you will see the policy implemented displayed on your computer below your parcel number.

Appendix B: Nutrient Transport Model Details

The process of calculating final nutrient levels in each reach is done in three steps. First the hydraulic characteristics (water flow, volume, and velocity) are calculated based on flow conservation and Manning's equation. Then in-stream kinematics (chemical and biophysical processes) are calculated based on the current flow and nutrient levels. Finally, nutrient advection (downstream flow) and dispersion (high concentration to low concentration flow) dynamics, results of kinematic changes, and changes is external load are used in a mass-balance equation to calculate new nutrient levels. We iterate over these three steps for each parcel over 200 iterations, by which point the nutrient levels return to their baseline levels. Appendix B details the kinematic equations and parameters used in the model.

Hydraulic Flow

The basic hydraulic unit is flow, measured in m^3/day, denoted Q_i. The inflow from the headwater, Q_0, is specified exogenously, and then conserved across the remaining reaches, $Q_i = Qi\text{-}1 + Q_{in,\,i} - Q_{out,\,i}$. $Q_{in,\,i}$ and $Q_{out,\,I}$ represent the total inflow or outflow from any sources across parcel i (for instance, pumping for irrigation or runoff from a storm event). Once Q_i is determined, the water velocity (U_i) and depth (H_i) across R_i can be calculated using Manning's Equation which equates the flow with the geometry of the reach's channel:

$$Q = \frac{S_0^{1/2} A_c^{5/3}}{n\, P^{2/3}}$$

where Q is flow (m^3/second), S_0 is the slope of the channel (m/m), A_c is the area of a channel cross section (m^2), and P is the "wetted perimeter," or length of the channel cross section that is in contact with water (m). The Manning roughness coefficient, n, takes values between 0.015 and 0.15, depending on the smoothness of the channel surface, with larger numbers indicating rough or weedy river bottoms. Assuming a trapezoid shaped channel and making a few geometric substitutions, this can be solved for water depth by recursing on the following series until it reaches a convergence rate below 0.001:

$$H_l = \frac{Q^{3/5}\left[B_0 + \ H_{l-1}\sqrt{s_{s1}^2 + 1} + H_{l-1}\sqrt{s_{s2}^2 + 1}\right]^{2/5}}{S_0^{3/10}[B_0 + 0.5(s_{s1} + s_{s2})H_{l-1}]}$$

where Q and S_0 are as described above, H_l is the water depth in iteration *l* (m), s_{s1} and s_{s2} are the slopes of the banks (m/m), and B_0 is the width of the bottom (m). Once H_i is determined, the velocity, U_i can be determined by diving the flow by the cross sectional area. Finally, the lateral dispersion rate is calculated as:

$$E_i = \max\left[0, (0.011\frac{U_i^2 B_i^2}{H_i\sqrt{9.81 S_{0,i} H_i}} - \frac{U_i \Delta x_i}{2})\right].$$

Kinematics

The kinematics represents the biochemical process in the system. There are three basic sets of processes: algae grow based on available light, nitrogen and phosphorus. Over time algae decays and releases nutrients or settles and remove the nutrients from the system. In the nitrogen cycle, algae decay releases organic nitrogen. This may settle or be converted to ammonium, which can be then be taken up by algae or converted to nitrite. Nitrite can be converted to nitrate, which can also be taken up by algae. The phosphorus cycle takes organic phosphorus from algae, mineralizes it to inorganic phosphorus, which then may be taken back up by algae. The equations that govern this are detailed in Appendix B.

Mass-Balance

The final nutrient levels are calculated for each nutrient using the following mass-balance equation:

$$\Delta c_i = \frac{Q_{i-1}}{Vi} c_{i-1} - \frac{Q_i}{Vi} c_i + \frac{E_{i-1}}{Vi} (c_{i-1} - c_i) - \frac{E_i}{Vi} (c_i - c_{i+1}) + \frac{W_i}{Vi} + K_i$$

Where Q_i is the flow, V_i the volume, c_i the nutrient level, E_i the dispersion rate, and W_i the new nutrient load for reach i. K_i is the net kinematic production of the nutrient.

Table B1. Parameters and Variables

	Parameter	Value
K_L	Michaelis-Menton Half Saturation Constant - Light	0.75
K_N	Michaelis-Menton Half Saturation Constant– Nitrogen	0.02
K_P	Michaelis-Menton Half Saturation Constant – Phosphorus	0.025
N_d	Hours of Daylight	12
Fr_{Ph}	Fraction of Photosynthetically active Daylight	0.3
H_{day}	Average Daily Solar Radiation	10.0
$k_{l,0}$	Algal Self-shading Intercept	1.0
$k_{l,1}$	Algal Self-shading Linear Coefficient	0.03
$k_{l,2}$	Algal Self-shading Quadratic Coefficient	0.054
α_0	Chlorophyll to Biomass Ratio	50.0
α_1	Nitrogen as Fraction of Biomass	0.08
α_2	Phosphorus as Fraction of Biomass	0.015
β_{N1}	Rate of Biological Oxidation of Ammonia to Nitrogen	0.55
β_{N2}	Rate of Biological Oxidation of Nitrite to Nitrate	1.1
β_{N3}	Nitrogen to Ammonium Hydrolysis Rate	0.21
β_{P4}	Organic P Mineralization Rate	0.35
μ_{max}	Maximum Algal Growth Rate	2.0
ρ_a	Algal Respiration Rate	0.3
σ_1	Algae Settling Rate	1.3
σ_2	Benthos Source Rate for Soluble P	0.05
σ_3	Benthos Source Rate For Ammonia	0.5
σ_4	Organic N Settling Rate	0.2
σ_5	Organic P Settling Rate	0.15
f_{NH4}	Algal Preference for Ammonia Nitrogen	0.5
DOX	Dissolved Oxygen	9.2
Endogenous Variables		
oN_t	Organic Nitrogen	
$NO2_t$	Nitrite	
$NO3_t$	Nitrate	
$NH4_t$	Ammonium	
oP_t	Organic Phosphorus	
iP_t	Inorganic Phosphorus	
$algae_t$	Algae	

Table B2. Kinematic Processes

Process	Function
Algal Growth	
Active Daylight	$\bar{I} = \dfrac{Fr_{Ph} * H_{day}}{N_d/24}$
Algal Self Shading	$k_l = k_{l,0} + k_{l,1} * \alpha_0 * algae_{t-1} + k_{l,2} * (\alpha_0 * algae_{t-1})^{2/3}$
Light Growth Limiting Factor	$F_L = \dfrac{0.92 * \frac{1.0}{k_{l,0}} * \log(K_L + \bar{I})}{K_L + \bar{I} * e^{-1.0 * k_l * depth}}$
Nitrogen Growth Limiting Factor	$F_N = \dfrac{NO3_{t-1} + NH4_{t-1}}{NO3_{t-1} + NH4_{t-1} + K_N}$
Phosphorus Growth Limiting Factor	$F_P = \dfrac{InorganicP_{t-1}}{InorganicP_{t-1} + K_P}$
Algal Growth Rate	$\mu_T = \mu_{Max} * F_L * min(F_N, F_P)$
Temp Adjusted Algal Growth Rate	$\mu = \mu_T * 1.047^{temp-20.0}$
Temp Adjusted Respiration Rate	$r_a = \rho_a * 1.047^{temp-20.0}$
Temp Adjusted Algae Settling Rate	$s_1 = \sigma_1 * 1.024^{temp-20.0}$
New Algae Level	$algae_t = \mu * algae_{t-1} - r_a * algae_{t-1} - \frac{s_1}{H} * algae_{t-1}$
Nitrogen Cycle	
Temp Adjusted Nitrite Oxidation Rate	$b_{N3} = \beta_{N3} * 1.047^{temp-20.0}$
Temp Adjusted Organic N Settling Rate	$s_4 = \sigma_4 * 1.024^{temp-20.0}$
New Organic Nitrogen Level	$oN_t = \alpha_1 r_a * algae_{t-1} - b_{N3} * oN_{t-1} + s_4 * oN_{t-1}$

Algal Ammonium Uptake Rate	$fr_{NH4} = \frac{f_{NH4} * \; NH4_{t-1}}{f_{NH4} * \; NH4_{t-1} + (1 - \; f_{NH4}) * \; NO3_{t-1}}$
Temp Adjusted Ammonia Benthos Sourcing	$s_3 = \; \sigma_3 * 1.074^{temp-20.0}$
Adjusted Ammonia Oxidation Rate	$b_{N1} = \; \beta_{N1} * (1 - e^{-0.6*DOX}) * 1.083^{temp-20.0}$
New Ammonium Level	
Adjusted Nitrite Oxidation Rate	$b_{N2} = \; \beta_{N2} * (1 - e^{-0.6*DOX}) * 1.047^{temp-20.0}$
New Nitrite Level	$NO2_t = b_{N1} * NH4_{t-1} - \; b_{N2} * NO_2$
New Nitrate Level	$NO3_t = b_{N2} * NO2_{t-1}$ $-(1 - fr_{NH4}) * \; \alpha_1 algae_{t-1}$
Phosphorus Cycle	
Temp Adjusted Organic P Mineralization	$b_{P4} = \; \beta_{P4} * 1.047^{temp-20.0}$
Temp Adjusted Organic P Settling Rate	$s_5 = \; \sigma_5 * 1.024^{temp-20.0}$
New Organic P Level	$oP_t = \alpha_2 r_a * algae_{t-1} - \; b_{P4} * oP_{t-1} + s_5 * oP_{t-1}$
Temp Adjusted P Benthos Sourcing	$s_2 = \; \sigma_2 * 1.074^{temp-20.0}$
New Inorganic P Level	$oP_t = b_{P4} * oP_{t-1} + \frac{s_2}{1000 * depth}$ $- \alpha_2 \mu_a * algae_{t-1}$

The Department of Applied Economics and Statistics
College of Agriculture and Natural Resources
University of Delaware

The Department of Applied Economics and Statistics carries on an extensive and coordinated program of teaching, organized research, and public service in a wide variety of the following professional subject matter areas:

Subject Matter Areas

Agricultural Policy	Environmental and Resource Economics
Food and Agribusiness Management and Marketing	International Agricultural Trade
Natural Resource Management	Price and Demand Analysis
Rural and Community Development	Statistical Analysis and Research Methods

The department's research in these areas is part of the organized research program of the Delaware Agricultural Experiment Station, College of Agriculture and Natural Resources. Much of the research is in cooperation with industry partners, the USDA, and other State and Federal agencies. The combination of teaching, research, and service provides an efficient, effective, and productive use of resources invested in higher education and service to the public. Emphasis in research is on solving practical problems important to various segments of the economy.

The mission and goals of our department are to provide quality education to undergraduate and graduate students, foster free exchange of ideas, and engage in scholarly and outreach activities that generate new knowledge capital that could help inform policy and business decisions in the public and private sectors of the society. APEC has a strong record and tradition of productive programs and personnel who are engaged in innovative teaching, cutting-edge social science research, and public service in a wide variety of professional areas. The areas of expertise include: agricultural policy; environmental and resource economics; food and agribusiness marketing and management; international agricultural trade; natural resource management; operations research and decision analysis; rural and community development; and statistical analysis and research methods.

APEC Research Reports are published by the Department of Applied Economics and Statistics, College of Agriculture and Natural Resources of the University of Delaware.

University of Delaware