2020

Characterizing Heterogeneous Behavior of Non-Point-Source Polluters in a Spatial Game under Alternate Sensing and Incentive Designs

Asim Zia

Shanshan Ding

Kent D. Messer

Haoran Miao

Jordan F. Suter

See next page for additional authors

Citation/Publisher Attribution

Zia, A., Ding, S., Messer, K. D., Miao, H., Suter, J. F., Fooks, J. R.,...Koliba, C. (2020). Characterizing Heterogeneous Behavior of Non-Point-Source Polluters in a Spatial Game under Alternate Sensing and Incentive Designs. *J. Water Resour. Plann. Manage., 146*(8), 04020054. doi: 10.1061/(ASCE)WR.1943-5452.0001242
Available at: http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0001242

ASCE

Characterizing Heterogeneous Behavior of Non-Point-Source Polluters in a Spatial Game under Alternate Sensing and Incentive Designs

Asim Zia[1]; Shanshan Ding[2]; Kent D. Messer[3]; Haoran Miao[4]; Jordan F. Suter[5]; Jacob R. Fooks[6]; Todd Guilfoos[7]; Simona Trandafir[8]; Emi Uchida[9]; Yushiou Tsai[10]; Scott Merrill[11]; Scott Turnbull[12]; and Christopher Koliba[13]

Abstract: Behavioral research on natural resource management has revealed a number of variables that can impact collective action. This research builds upon an interactive decision game using experimental economics methods with a focus on production decisions and the corresponding impact they have on ambient water quality. Using hierarchical clustering algorithms, four primary types of behavior are identified: competitive, hypercompetitive, cooperative, and hypercooperative. The results from the experiment are used to test the following three hypotheses: (1) financial incentives increase cooperative behavior, (2) increasing the number and frequency of water quality sensors increases cooperative behavior, and (3) the spatial location of the agents and sensors affect cooperative behavior. Mixed-effect multinomial logistic models reveal that policy incentives, sensor location, and frequency of sensing alter the behavioral strategies of decision makers in the experiment and that outcomes vary by spatial location. From a watershed planning perspective, minimal investments in advanced environmental monitoring/sensing systems can potentially have large effects in improving water quality; however, there is also some evidence of marginal diminishing returns associated with such investments. **DOI: 10.1061/(ASCE)WR.1943-5452.0001242.**

Author keywords: Behavioral change; Experimental economics; Hierarchical clustering; Incentive mechanisms; Nash equilibria; Policy designs; Sensor information; Watershed management; Cyberinfrastructure development.

Introduction

Policymakers and watershed managers now have the potential to invest in adopting advanced environmental monitoring systems and sensor networks as a way to respond to the risks posed by the impact of human activities. In assessing the benefits of such technologies, however, the role of human behavior, in particular strategic behavioral adaptation to the sensor locations, needs to be explicitly taken into account. Over the last decade, there has been tremendous progress in the development and application of in situ optical water quality sensors, a promising tool for advancing understanding of the behavior of aquatic systems at finer temporal and spatial scales (Downing et al. 2012). The critical premise underlying the application of advanced optical sensors to watershed studies is that such sensors can frequently and comparably measure a suite of biochemical and geochemical parameters (i.e., diesel oxidation catalysts,

[1]Professor of Public Policy and Computer Science, Dept. of Community Development and Applied Economics, Director, Institute for Environmental Diplomacy and Security, Gund Fellow, Gund Institute for Environment, Univ. of Vermont, 146 University Place, Burlington, VT 05405 (corresponding author). ORCID: https://orcid.org/0000-0001-8372-6090. Email: Asim.Zia@uvm.edu

[2]Assistant Professor, Dept. of Applied Economics and Statistics, Univ. of Delaware, 531 South College Ave., Newark, DE 19716. Email: sding@udel.edu

[3]Professor, Dept. of Applied Economics and Statistics, Univ. of Delaware, 531 South College Ave., Newark, DE 19807. Email: messer@udel.edu

[4]Postdoctoral Fellow, Dept. of Civil, Architectural and Environmental Engineering, North Carolina A&T State Univ., 1601 E. Market St., Greensboro, NC 27410. Email: haoran_miao@my.uri.edu

[5]Associate Professor, Dept. of Agricultural and Resource Economics, Colorado State Univ., B-306 Clark Bldg., Fort Collins, CO 80526. ORCID: https://orcid.org/0000-0003-0520-7150. Email: jordan.suter@colostate.edu

[6]Senior Economist, Cigna Group Insurance, 1601 Chestnut St., Philadelphia, PA 19192. Email: Jacob.Fooks@cigna.com

Note. This manuscript was submitted on March 6, 2019; approved on February 5, 2020; published online on May 21, 2020. Discussion period open until October 21, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

[7]Associate Professor, Dept. of Environmental and Natural Resource Economics, Univ. of Rhode Island, 1 Greenhouse Rd., Kingston, RI 02881. Email: guilfoos@uri.edu

[8]Assistant Professor, Dept. of Environmental and Natural Resource Economics, Univ. of Rhode Island, 1 Greenhouse Rd., Kingston, RI 02881. Email: simona@uri.edu

[9]Professor, Dept. of Environmental and Natural Resource Economics, Univ. of Rhode Island, 1 Greenhouse Rd., Kingston, RI 02881. ORCID: https://orcid.org/0000-0001-6818-2849. Email: euchida@uri.edu

[10]Associate Research Scientist, Future H_2O, Office of Enterprise Knowledge Development, School of Sustainable Engineering and the Built Environment, Arizona State Univ., 330 E University Dr. Tempe, AZ 85281. Email: yushiou.tsai@gmail.com

[11]Research Assistant Professor, Dept. of Plant and Soil Science, Gund Fellow, Gund Institute for Environment, Univ. of Vermont, 63 Carrigan Dr., Burlington, VT 05405. ORCID: https://orcid.org/0000-0001-5574-9703. Email: scmerril@uvm.edu

[12]Software Engineer, Vermont Established Program to Stimulate Competitive Research, Univ. of Vermont, 23 Mansfield Ave., Burlington, VT 05405. Email: Scott.Turnbull@uvm.edu

[13]Professor, Dept. of Community Development and Applied Economics, Gund Fellow, Gund Institute for Environment, Univ. of Vermont, 146 University Place, Burlington, VT 05405. Email: ckoliba@uvm.edu

nitrates, and particle loads) in concert over long periods of time. Consequently, their potential for use in some settings is limited only by the capability of the technology to function in the environment and measure relevant parameters accurately and consistently.

These sensors have demonstrated their potential to help increase understanding of water quality and aquatic ecosystem functions in the context of local land management (Pellerin et al. 2009; Florsheim et al. 2011), hydrochemical evolution in river networks during storms (Saraceno et al. 2009), mobility of atmospheric pollutants (Bergamaschi et al. 2012), downstream solute processing (Hernes et al. 2008; Kraus et al. 2010), and diurnal to seasonal biochemical and geochemical behaviors (Pellerin et al. 2012). Furthermore, development of larger-scale regional advanced sensor networks holds tremendous promise in enhancing understanding of the effects of regional drivers of aquatic ecosystems, which include climate disturbances, droughts, and even long-term regional environmental change (Groisman et al. 2004; Kaushal et al. 2010).

Despite these advances in the understanding of ecosystem behavior (Rode et al. 2016), critical knowledge gaps exist with respect to understanding the social and policy implications of instituting sensor networks, which need to be investigated with experimental and field research, ideally with a watershed management and policy design perspective. Improved understanding of the precision, location, and frequency of sensor data effects on human behavior may inform local and regional adaptive watershed management practices (Huang and Xia 2001; Lu et al. 2015; Waylen et al. 2019) and enable calibration and validation of regional integrated assessment models (Zia et al. 2016). This study is an initial attempt to identify the extent to which investments in environmental monitoring systems coupled with appropriate fine-tuning of environmental policy designs impact human behavior.

This paper builds upon the premise that the ability to understand, tune, and adjust incentives and environmental information available to decision makers can provide testable opportunities for governance of human–environmental systems. Improved understanding of human behavioral responses to such tunable conditions may enable iterative and adaptive adjustment of environmental policy designs because whether individuals and organizations cooperate or act selfishly affects the performance of environmental policy (Madani and Lund 2011; Giuliani et al. 2014; Norton 2015; Yu et al. 2019).

Specifically, this study assesses human behavior related to non-point-source pollution, an important contributor to water pollution, using the results of a laboratory economics experiment that combines sensor data information and financial incentive treatments. The data allow for exploring the extent to which an observational water quality sensor network combined with financial incentives determine whether agents behave cooperatively or non-cooperatively in the decisions that they make related to water pollution. Observational sensor networks are an important ingredient in adaptive management of social ecological systems (SES) (Staudinger et al. 2013). This study broadly seeks to contribute toward an improved understanding of the impacts of observational sensor networks on human behavior through economics experiments that get to the core of the water-resource-related decision-making process of non-point-source polluters.

The present study is motivated by the idea that designs for environmental policies and investments in the deployment of water quality sensors can be the most effective when they are informed by accurate information about human behavioral responses generated using spatially explicit marginal incentives. This study develops a non-point-source pollution control game to examine decisions that impact water quality and the effects of those decisions on nutrient fluxes. Different from existing works that study average treatment effects on production levels (e.g., Miao et al. 2016), this study investigates varying treatment effects associated with different sensing and incentive regimes on heterogeneous agent behavior, especially the effect on inducing cooperative behavior. Specific behavioral types, including cooperative behavior, are identified using cluster analysis. Mixed-effects multinomial logistic (MMNL) modeling is then used to analyze how specific treatment conditions impact the behavioral types, which allows specific behavioral hypotheses related to water quality to be addressed.

The study provides a more comprehensive understanding of heterogeneity of human behavior and new insight into how monitoring and policy differentially influence behavior types. Here, cooperative behavior is defined as the adoption of decisions that reduce non-point-source pollution below the predicted Nash equilibrium, which is based on strictly self-interested behavior. The placement and frequency of sensors may impact cooperative behavior in a watershed because they change the incentives of agents under a policy mechanism that uses the information from the sensors to determine taxes and subsidies. Spatial and location-based effects of water quality sensor information are tested by changing the effective costs of polluting through a penalty (tax) and reducing the costs of taking an action through a reward (subsidy) for inducing cooperative behavior via three hypotheses (formally defined in Supplemental Materials S1.1):

1. Incentives induce cooperative behavior among agents.
2. Increasing the number of water quality sensors as well as their sensing frequency increases cooperative behavior.
3. The spatial locations of the agents relative to the spatial locations of the sensors, i.e., upstream, midstream, or downstream, affect the induction of cooperative behavior.

Methods

Experiment Design

Non-point-source water pollution is characterized by individual agents whose behavior impacts downstream water quality conditions, but the pollution contribution of any one agent cannot be accurately determined. The design of the experiment is motivated by a situation where a regulator seeks to reduce damages caused by water pollution generated by a group of producers situated along a river reach, where the pollution contribution of each individual producer cannot be detected. This section presents important highlights of the basic design and the experimental treatments that define the implemented experimental game. The Supplemental Material presents the complete technical details of the game design (also discussed by Miao et al. 2016). The theoretical decision framework that underlies the experiment is based on Segerson (1988)'s non-point-source pollution model, and the basic payoff functions and parameters are similar to those of Spraggon (2002). Spatial differences and a nutrient transport model (QUAL2K) are incorporated into the experiment to simulate the physical environment (Chapra et al. 2008) (also discussed in Supplemental Materials S2.3).

A session of the experiment is composed of multiple groups. Each group is composed of six student participants who play the role of agents making production decisions in a common watershed. The production by agents occurs on specific parcels, which are located linearly along a river: Parcel 1 is the farthest upstream and Parcel 6 is the farthest downstream (Fig. 1). The spatially explicit positioning of the agents has parallels to lab and field experiments that have explored the differences in water-use decisions

J. Water Resour. Plann. Manage., 2020, 146(8): 04020054

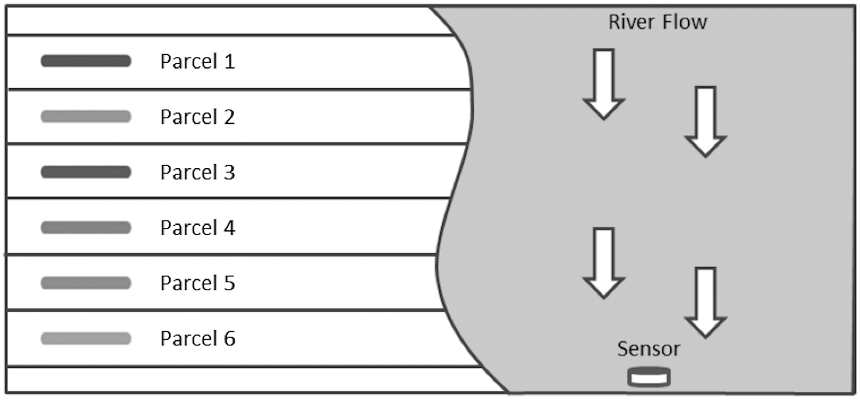

Fig. 1. Spatial location of experimental design.

of upstream and downstream irrigators (D'Exelle et al. 2012; Bell et al. 2015).

The agents in the experiment make decisions over a series of rounds, using a web-browser-based interface built into the Python package Willow (Weel and McCabe 2015). In every round, the six agents simultaneously make production decisions that determine their private incomes. Agents' production creates byproducts (pollution) that enter a river and cause damage to hypothetical downstream users. A principal, representing an environmental regulator, penalizes or rewards each agent based on the measured ambient (group) pollution level at a downstream monitoring point (sensor) that results from the production decisions of all upstream agents.

The measurement of the pollution level occurs at specific sensors and is assumed to be measured over an interval of time, referred to here as a sensing window. The sensing window in a given round occurs after the production decisions of each of the agents are complete. Specifically, the sensing window starts from the time the pollution enters the stream and lasts until pollution concentrations at the sensors reach a steady state. The effect of an individual production decision on the observed concentration at the sensors depends on all agents' production decisions, their locations, and the frequency of the sensing measurements. Given the non-point-source pollution motivation for the study, the agents cannot observe the behavior of other agents and the regulator cannot observe the individual agents' actions, just the ambient level of pollution measured at the sensors.

Each agent's profit in a round is composed of the income they earn from production, adjusted by any tax or subsidy imposed by the principal according to pre-established policy rules. These rules are communicated to the agents in the experiment instructions (Supplemental Material). The income functions of the six agents are identical (indexed by $n = 1, \ldots, 6$). The income function is quadratic in the agent's production level, which is the decision variable, x_n, and takes the form of Eq. (1)

$$I_n(x_n) = 35 - 0.0075(50 - x_n)^2, \quad \text{for } n = 1, \ldots, 6 \qquad (1)$$

where I_n = private income of agent n in a given round. The decision variable, x_n, can take on values between 0 and 50.

The amount of pollution byproduct generated by the agent is proportional to their production level (it is assumed that 1 unit of production leads to 1,000 μgN/L emissions of nitrite), and the pollution concentration over time at one point in the river is determined jointly by all agents' decisions. Each round is independent, and pollution concentrations do not transfer across rounds. The tax/subsidy that applies in each round is a linear function of the measured pollutant concentration and is therefore contingent on the production decisions of all agents. The tax/subsidy takes the form of Eq. (2a) if there is only one sensor placed just downstream of the last parcel

$$\mathrm{TS}(x_1, \ldots, x_6) = 10(\mathrm{Threshold} - C(x_1, \ldots, x_6)) \qquad (2a)$$

where TS $(x_1, \ldots, x_6)$ = tax/subsidy that is charged on an individual agent; and $C(x_1, \ldots, x_6)$ = measured concentration of pollution. However, if there is one additional sensor placed between Parcels 3 and 4, the tax/subsidy takes the form of Eq. (2b)

$$\mathrm{TS}_{\mathrm{up}}(x_1, x_2, x_3) = 10(\mathrm{Threshold}_{\mathrm{up}} - C(x_1, x_2, x_3))$$
$$\mathrm{TS}_{\mathrm{down}}(x_4, x_5, x_6) = 10(\mathrm{Threshold}_{\mathrm{down}} - C(x_4, x_5, x_6)) \qquad (2b)$$

where $\mathrm{TS}_{\mathrm{up}}(x_1, x_2, x_3)$ = tax/subsidy that is charged to an individual agent from Parcels 1 to 3; $C(x_1, x_2, x_3)$ = measured concentration of pollution from the sensor between Parcels 3 and 4; $\mathrm{TS}_{\mathrm{down}}(x_4 x_5, x_6)$ = tax/subsidy that is charged to an individual agent from Parcels 4 to 6; and $C(x_4, x_5, x_6)$ = measured concentration of pollution from the downstream sensor. The thresholds in Eqs. (2a) and (2b) are exogenous concentration levels determined by the principal (regulator). The thresholds are different across treatments and are designed such that average TS is approximately zero when each agent chooses a production level of 25. The target production level of 25 is assumed to be determined exogenously and is not a product of the economic model. When TS is negative, the agents pay a tax; otherwise, they receive a subsidy.

The profit earned by an individual participant in a given round is the sum of private income [Eq. (1)] and the tax/subsidy in that round [Eq. (2)]. When the measured pollution concentration is greater than the threshold, TS is negative and therefore the participants' realized profits in that round are lower than their private income.

The amount of pollution generated by each agent is assumed to enter the river simultaneously and then mix together, flow downstream, and dissipate. In other words, the pollution becomes less concentrated as it spreads across a larger area. Although based on the sum of pollution generated by each agent, there are nonlinear effects that can increase pollution concentrations as contaminates flow downstream through nitrification. The nutrient transport model QUAL2K is here calibrated to simulate water flow and the diffusion of the nitrite that is produced by agents and other background nutrients (details can be found in Supplemental Materials S2). In this specification, a nitrification process is observed

Table 1. Treatment descriptions

Treatment label	Number of sensors	Frequency of sensing	Ambient tax/ subsidy
One sensor, low frequency	1	Low	Yes
One sensor, high frequency	1	High	Yes
Two sensors, low frequency	2	Low	Yes
Two sensors, high frequency	2	High	Yes

that causes upstream parcels to have a large effect on the measured concentration, whereas downstream parcels have smaller effects. This specification simulates river ecosystems with larger residence times of nutrients, such as the Delaware River and lower Mississippi River systems.

Due to the model's complexity, experiment participants were not provided with the specific analytical relationship between the concentration of the pollutant in the river and their production decisions. However, the instructions provided numerical examples, and the computer screen included a calculator to help the participants understand the effects of their decisions. The participants were given ample time to use the calculator to enter hypothetical production levels for each parcel and then observe the resulting measured concentration, tax/subsidy, and profits.

The pollution damage (social cost) was assumed to happen immediately downstream from Parcel 6, and the degree of damage is determined by the maximum concentration of the pollutant that is observed during the sensing window. Specifically, the damage function takes the form of Eq. (3)

$$D(x_1, \dots, x_6) = -10C_{\max}(x_1, \dots, x_6) \quad (3)$$

where $D(x_1, \dots, x_6)$ = economic damage; and $C_{\max}(x_1, \dots, x_6)$ = maximum concentration, within the sensing window, just downstream of Parcel 6. The net social benefit takes the form of Eq. (4)

$$\text{NSB} = \sum_{n=1}^{6} [35 - 0.0075(50 - x_n)^2] - 10C_{\max}(x_1, \dots, x_6) \quad (4)$$

Table 1 describes the four treatments, which differ in terms of the number of sensors and frequency of sensing. In the first two treatments, there is one sensor located just downstream of Parcel 6; in last two treatments, there is a sensor located just downstream of Parcel 6 and another located just downstream of Parcel 3. In terms of sensing frequency, the two low-frequency treatments involve the least monitoring—one time during the sensing window—and the tax/subsidy is calculated based on the measured pollutant concentration. In the two treatments with high-frequency sensing, sensors measure the pollutant concentration four times randomly chosen during the sensing window, and the tax/subsidy is determined based on the maximum concentration measured. A status quo (control) treatment is also included in the experiment in which there is no policy or sensor intervention. Two additional treatments with continuous sensing were also played but are excluded from the analysis in this paper due to the possibility of multiple Nash equilibria.

The socially optimal production is a vector of production decisions at each parcel $x_1, \dots, x_6$ that maximizes the net social benefit in Eq. (4) and is displayed in the last row of Table 2. The Nash equilibria (NE) for each parcel for the control (status quo) and the four treatments are also presented in Table 2. The Nash equilibria are calculated based on the assumption that every agent will maximize their private expected payoff given the decisions of other agents. In every equilibrium, each agent has no better strategy (production decision) to maximize the expected payoff given the behaviors of other agents. Since a numeric nutrient transportation model is used in this study, it is not possible to get analytical Nash equilibria solutions, and therefore solve for the Nash equilibria numerically (Supplemental Materials S2.2 provides the method). As illustrated in Table 2, the NE primarily differ due to the frequency of sensing, which determines the expected magnitude of the tax/subsidy. By design, the number of sensors does not have a large impact on the NE, but the number of agents that can impact the tax/subsidy is influenced by the number of sensors. Under the calibrated stream and pollution dynamics, the optimal total emissions across all agents is 200, with heterogeneity across agents based on spatial location. Specifically, emissions from upstream parcels have a larger effect on the maximum concentration than downstream parcels. Optimally, the farthest downstream parcel (Parcel 6) has the highest emissions whereas the parcel farthest upstream (Parcel 1) has the lowest emissions.

Experiment Implementation

A total of 108 undergraduate students (18 groups of six) participated in six experimental sessions at an experimental laboratory at the University of Delaware. Each student was randomly assigned to a computer station before entering the lab. The assigned computer determines their initial group membership and their parcel location along the river.

In the experiment, each agent faces every treatment (a within-subject design) and each treatment consists of six rounds. In each round, agents make one production decision, and then they are shown a screen with the measured level of pollution, the tax/subsidy, and their total profit. Each subject's location (Parcels 1–6) changes randomly between treatments, and the groups are rearranged between treatments to reduce group experience effects. The order of the treatments is also varied across six experimental sessions to reduce potential order effects. At the beginning of the session, subjects are allowed 15 min to read the first part of the instructions (Supplemental Material). Next, they watch a PowerPoint presentation to ensure their comprehension of the tax/subsidy and underlying pollution-diffusion model. Additional instructions and presentations are offered at the beginning of each treatment. For the results reported in this study, sessions lasted approximately 90 min, and participant earnings were directly determined by the profits that they

Table 2. Predicted (Nash equilibrium) and socially optimal production levels for each parcel

Treatment	Parcel 1	Parcel 2	Parcel 3	Parcel 4	Parcel 5	Parcel 6
One sensor, low frequency	33.5	33.4	33.3	33.3	33.3	33.3
One sensor, high frequency	21.8	22.8	28.5	36.5	44.6	48.8
Two sensors, low frequency	33.3	33.3	33.3	33.3	33.3	33.3
Two sensors, high frequency	22	20.2	29	38.4	45	48.9
Status quo	50	50	50	50	50	50
Social optimum	20	24	30	33	44	49

J. Water Resour. Plann. Manage., 2020, 146(8): 04020054

earned over the course of the sessions. On average, participants received \$29 in earnings at the end of the experiment.

Cluster Analysis

The objective of the experiment is to observe behavioral responses to differences in the placement and frequency of sensors, when used in conjunction with financial incentives related to water quality. Cluster analysis is used to classify participants into distinct behavioral types. Then, using the behavioral types identified by the cluster analysis, MMNL models, described in the next subsection, are used to analyze how the treatment conditions impact the behavioral types. The results from the MMNL modeling allow the three introduced hypotheses to be empirically addressed.

Cluster analysis is a method for identifying heterogeneous classes/groups among subjects where subjects within each class/group display similar behavior and subjects between classes/groups exhibit distinct behavior (Anderberg 1973; Hastie et al. 2001; Johnson and Wichern 2002; Berry and Linoff 1997; Everitt 2011). It is particularly useful when a priori knowledge regarding the number of behavioral types or the type structure is not available. To explore heterogeneity of agent behavior, cluster analysis is applied to segment the agents into distinct behavioral types using the difference between an agent's actual production decision and the NE as a measure of the agent's behavior in each round, shown in Eq. (5)

$$\text{diff}_t = \text{production}_t - \text{Nash equilibrium} \tag{5}$$

where t = round in which the decision is made. Thus, six observations are obtained for each subject in each treatment, designated diff_1, diff_2, diff_3, diff_4, diff_5, and diff_6 representing the differences in each of the six periods. The cluster analysis is then implemented based on the six observations together to segment the subjects. That is, the distance matrix between two subjects used in the cluster analysis is calculated with all six observations. This allows for possibly capturing subjects' changing behavior over time, although the empirical results show that agent decisions are largely consistent over the six rounds.

Three common clustering methods are considered: the Gaussian mixture model, K-means clustering, and hierarchical clustering. Because there is no objectively correct clustering algorithm for a particular problem, the most appropriate clustering method is typically selected, experimentally, based on the specific structure of the data (Estivill-Castro 2002; Keritino et al. 2015). After model diagnostics, it was found that the Gaussian mixture model does not satisfy the required normality assumption. Several statistics were then computed, such as the Calinski and Harabasz index (or variance ratio criterion, Calinski and Harabasz 1974), Krzanowski-Lai (KL) index (Krzanowski and Lai 1988), and Scattering and Dispersion (SD) validity index (Halkidi et al. 2000), to evaluate the goodness of fit of the hierarchical and K-means methods. For example, the Calinski and Harabasz (CH) index measures the ratio of between-group variation to within-group variation. A good clustering method will create within-group subjects as similar as possible and between-group subjects as dissimilar as possible. Therefore, a large CH index is indicative of good clustering. The Calinski and Harabasz index of the hierarchical clustering in this study is 208.94, which is higher than 38.76 of the K-means method, indicating that the hierarchical method has a better goodness of fit. The other statistics in general suggest similar results. Henceforth, only the key results derived from hierarchical clustering are described.

Mixed-Effects Multinomial Logistic Regression Analysis

To study possible varying effects of sensor and incentive conditions on heterogeneous agent behavior and to test the Hypotheses 1–3, MMNL models are estimated using the results from the cluster analysis. The MMNL model was introduced by Boyd and Mellman (1980) and Cardell and Dunbar (1980) in the economic literature. Since then, various economic studies have improved or applied the MMNL model for data analysis (e.g., Beggs 1988; Enberg et al. 1990; Reader 1993; Jain et al. 1994; Revelt and Train 1998; Train 1998; Nevo 2001; Small and Verhoef 2007; Fiebig et al. 2010). The MMNL model is a generalization of the multinomial logistic model and thus it can handle multilevel categorical or discrete response variables. It measures the relationship between a multilevel categorical response and one or more independent variables (predictors) and can be used to predict the probabilities of the different possible outcomes of the response based on a logistic function of the linear combinations of the predictors. In addition, it can capture repeated measurement information and heterogeneity in the population by specifying random effects in model fitting.

In this study's analysis, the data are fit with a categorical response $Y = \{1, 2, \ldots, K\}$, representing the behavioral groups estimated from the hierarchical clustering algorithm, and then the conditional probabilities of the agents' type of behavior are predicted given the experimental treatments (policy, sensors, and frequency of sensing) and array of spatial locations. The clustering algorithm (results described subsequently) identifies four groups, e.g., $K = 4$. Interactions between the treatments and the locations (a full factorial design) are also considered when fitting the model. The experiment sessions and rounds are treated as random effects because there might be heterogeneous structures associated with the two variables.

The model is mathematically formulated by introducing the following notation. Parcels, policy, sensors, and monitoring frequency are all categorical variables. Dummy (indicator) variables are used to represent each category of the variables. Let $z_1, z_2, \ldots, z_5$ denote the indicator variables for Parcels 1 to 5. Let z_6 denote the indicator variable for policy and z_7 denote the indicator variable for two sensors and measure the additional effect due to two sensors. Let z_8 denote the indicator variable for high frequency, let $z_9, z_{10}, \ldots, z_{13}$ denote the indicator variables for the interactions between parcel numbers and policy, let $z_{14}, z_{15}, \ldots, z_{18}$ denote the indicator variables for the interactions between parcel numbers and two sensors, let $z_{19}, z_{20}, \ldots, z_{23}$ denote the indicator variables for the interactions between parcel numbers and high frequency, and let z_{24} denote the interaction for two sensors and high frequency. Parcel 6 is not in the model formulation because it is treated as a baseline.

In addition, let $\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,24}$, $i = 2, 3, 4$, denote the coefficients of the corresponding variables in the multinomial logistic model. For convenience, let $[Z]$ be the predictor vector that contains elements $z_1, z_2, \ldots, z_{24}$, and β_i, $i = 2, 3, 4$, be the coefficient vectors that contain elements $\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,24}$, respectively. Let μ_i, $i = 2, 3, 4$, denote the intercepts. Eq. (6) describes the fixed-effect specifications. Random effects are omitted in the formulation because they mainly affect the variance but not the mean function

$$\log \frac{\Pr(Y = i)}{\Pr(Y = 1)} = \mu_i + \beta_{i,1} z_1 + \cdots + \beta_{i,24} z_{24} = \beta_i' Z, \quad i = 2, 3, 4 \tag{6}$$

Therefore, the predicted probabilities for the four behavioral group types can be calculated by Eq. (7)

J. Water Resour. Plann. Manage., 2020, 146(8): 04020054

Table 3. Mean differences in production levels relative to the NE, by clusters

Variable	Cluster 1: competitive (216 agents)	Cluster 2: cooperative (171 agents)	Cluster 3: hypercooperative (51 agents)	Cluster 4: hypercompetitive (102 agents)
Diff1	−0.23	−14.71	−26.06	6.91
Diff2	−2.02	−11.95	−28.01	7.56
Diff3	−1.34	−12.40	−28.42	7.32
Diff4	−1.97	−10.78	−27.83	8.67
Diff5	−2.31	−9.07	−25.26	8.49
Diff6	−3.36	−8.83	−24.22	7.05

$$\Pr(Y = i) = \frac{e^{\beta_i' Z}}{1 + e^{\beta_2' Z} + e^{\beta_3' Z} + e^{\beta_4' Z}}, \quad i = 2, 3, 4 \tag{7}$$

Results

This section describes the results of the cluster analysis and MMNL model estimated using the decisions made in the experiment. The cluster analysis is designed to separate the participants in each treatment into unique groups based on the production decisions they make. The MMNL model then describes how the treatment and other features of the experiment impact the behavioral types identified in the cluster analysis, which allows the three identified behavioral hypotheses to be addressed.

Cluster Analysis

The hierarchical clustering analysis provides a four-group partition, suggesting four agent behavior types. The group (cluster) mean difference in production relative to the NE are given in Table 3. An agent's behavior is defined relative to the NE. The NE for an agent is their privately optimal production decision in the game, provided all other agents are also playing their own optimal production decision. That is, agents are best responding to others' best response. The terms competitive and cooperative are used as labels to categorize agent behavior relative to the NE. As indicated in Table 3, Cluster 1 (competitive) behavior was, on average, closest to the NE. Clusters 2 (cooperative) and 3 (hypercooperative) chose lower than NE levels of production (and thus lower pollution), on average, and Cluster 3 tended to select the lowest production. Cluster 4 (hypercompetitive), on average, chose production decisions that exceeded the NE.

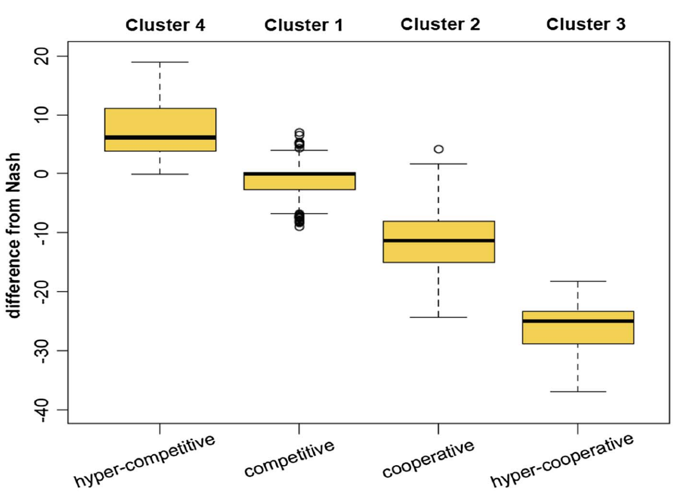

Fig. 2. Box plots of differences to Nash equilibria for the four behavioral groups.

Fig. 2 looks further at the data distribution within each cluster. Because the agent decisions (differences from NE) are largely consistent over the six rounds, the data distributions are very similar over time. To make the graphical illustration simple, the information was averaged over time for each agent and the distribution of the averaged differences for each cluster are summarized by a boxplot in Fig. 2. The bold line in the middle of each boxplot represents the median value of the (averaged) differences from the NE for the group, the horizontal lines on the top and bottom of each plot represent the maximum and minimum values of the cluster excluding outliers, and the circles represent outliers.

It can be seen from Fig. 2 that Cluster 4 generally produces more than the NE because the differences of the cluster are above zero. This class of agents is therefore characterized as hypercompetitive. Cluster 1 agents are closest to the NE because the median value of the cluster is closest to zero. These agents are characterized as competitive/selfish. In contrast, Cluster 2 and Cluster 3 are generally below zero, indicating that the agents in these two groups produce consistently below the NE, and the agents in Cluster 3 fall the farthest below. Therefore, Cluster 2 and Cluster 3 agents are characterized as cooperative/altruistic and hypercooperative, respectively. The information shown in these plots is consistent with the average per-cluster values presented in Table 3.

Fig. 3 shows the probability density functions of the four behavioral groups distributed under experimental control (status quo) and treatment conditions. Although approximately 83% of agents chose a competitive behavioral strategy predicted by the NE under the control status quo condition, the treatments induce increases in cooperative behaviors. The four behavioral groups derived from hierarchical cluster analysis are then used as response variables in the multinomial logistic regression analysis to investigate heterogeneous covariate effects and are further predicted by tunable conditions represented in the experimental treatments, parcel numbers, and their interactions.

Hypothesis Testing

Table 4 presents the estimated results from the MMNL model, which was selected based on the minimization of the Bayesian information criterion (BIC) across a suite of six different model specifications shown in the Supplemental Material. Models with different random effects specifications are compared across and with a null model (no random effects), as indicated in Table S1. It is found that the multilevel model that controls for sessions as random effects and includes a full factorial of interaction effects of experimental treatments with parcels is the model that best fits the data. The estimated coefficients from this best-fit model are provided in Table 4. The results indicate that the spatial location (whether a parcel is upstream or downstream of a sensor),

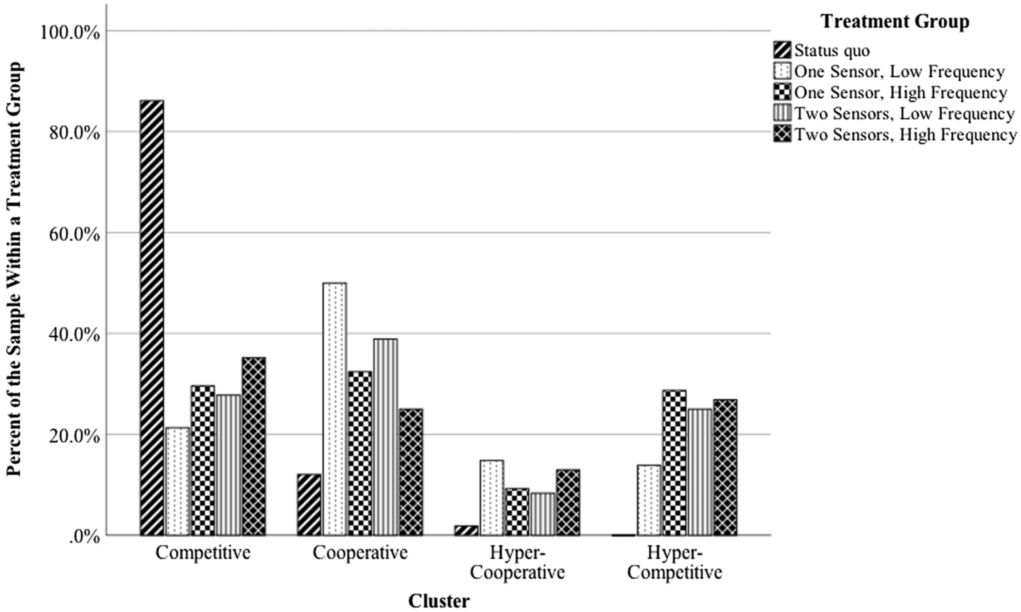

Fig. 3. Distribution of strategic behavioral clusters by treatment.

Table 4. Estimated fixed-effect coefficients (exponents of log-odds) from the MMNL model (base group = competitive, Cluster 1)

Variable	Cooperative (Cluster 2)	Hypercooperative (Cluster 3)	Hypercompetitive (Cluster 4)
Intercept	0.0597***	0.0001	4.33×10^{-11}***
Parcel 1	1.0194	463.1727	0.0698
Parcel 2	3.2563**	0.1430	0.0656***
Parcel 3	2.0763	0.1425	0.0705
Parcel 4	2.0628	0.1423	0.0702***
Parcel 5	5.0860***	599.9222	0.0501***
Policy	53.2394***	8738.554	4670.000
Two sensors	0.4036**	0.0472***	0.2277***
High frequency	0.6592	2.1316**	0.0008
Parcel 1 × policy	1.0173	0.0016	3.6096
Parcel 2 × policy	0.2776**	3.1205	0.6769
Parcel 3 × policy	0.3712**	8.0437	7.0202
Parcel 4 × policy	0.1315***	5.5118	2.5271
Parcel 5 × policy	0.1724**	1.12×10^{-7}	6.6352
Parcel 1 × two sensors	0.8738	6.7957**	11.4386***
Parcel 2 × two sensors	0.6647	6.0206**	32.3948***
Parcel 3 × two sensors	1.5380	15.1272***	1.6714
Parcel 4 × two sensors	4.8710***	13.1707***	14.8277***
Parcel 5 × two sensors	2.6195**	25.348***	13.8059***
Parcel 1 × high frequency	0.3418**	0.0338***	4685.838
Parcel 2 × high frequency	0.2266***	1.28×10^{-6}	11316.31
Parcel 3 × high frequency	0.4760**	0.0068***	1043.254
Parcel 4 × high frequency	2.5075**	0.1527***	5706.151
Parcel 5 × high frequency	1.0799	1254.51	819.4226
Two sensors × high frequency	0.8804	5.6598***	0.3081***

Note: *** $p < 0.001$; ** $p < 0.05$; and * $p < 0.01$. Sessions are controlled to account for random effects. Some large coefficients are due to exponential transformation.

frequency of sensing, number of sensors, and policy all affect the likelihood of an agent playing cooperatively or competitively.

Fig. 4 shows MMNL predicted marginal probabilities for all four behavioral groups in relation to four experimental predictors—policy (whether the tax/subsidy policy is in place), parcel ID, sensors, and frequency of sensing—while all other predictor variables in the MMNL are held constant at their mean values. The existence of the incentive-based policy increases the probability of cooperative behavior. Agents in upstream Parcels 1, 2, and 3 are more likely to be competitive and less likely to be cooperative compared with the agents in downstream parcels. The addition of one sensor (with policy) increases the marginal likelihood of cooperative behavior but diminishes this likelihood with another sensor. Finally, low frequency of sensing induces a higher marginal likelihood of cooperative behavior than higher frequency of sensing.

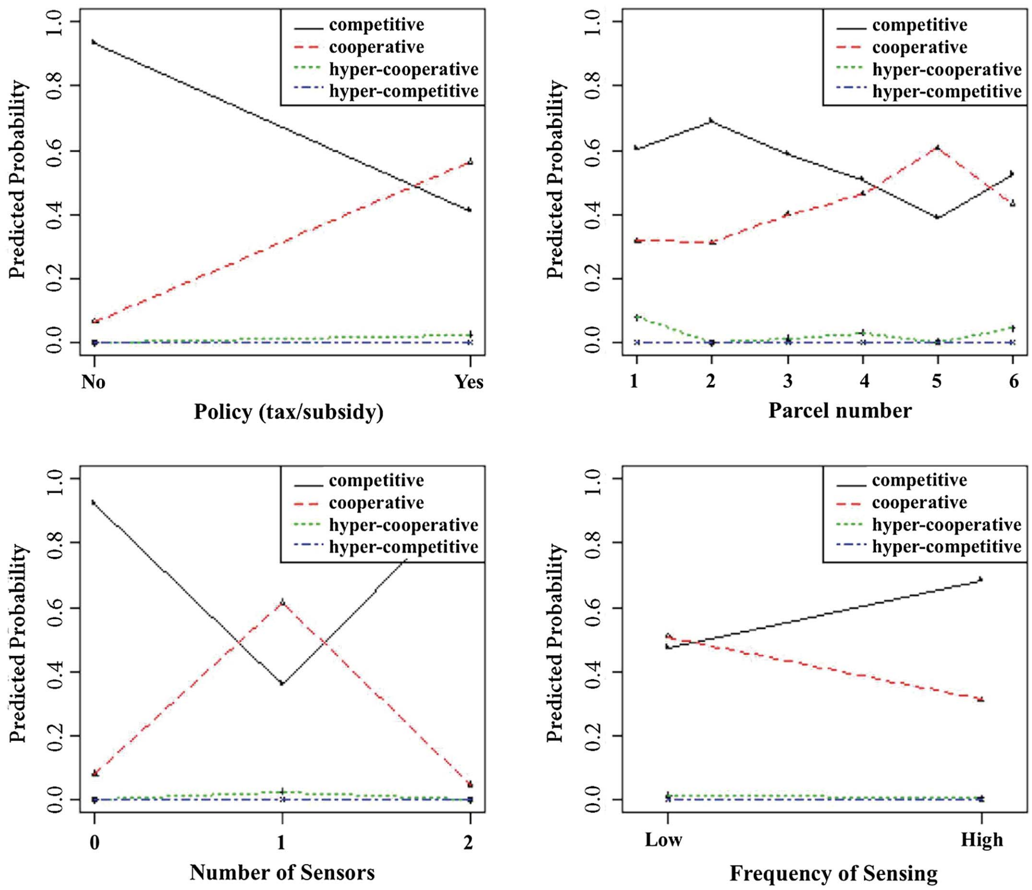

Fig. 4. Predicted marginal probabilities of four behavioral groups for four experimental predictors: policy (tax/subsidy), parcel ID, sensors, and frequency of sensing.

Hypothesis 1: Incentives Induce Cooperative Behavior among Agents

The results from the MMNL reveal that existence of policy (incentive manipulation) increases the likelihood of cooperative behavior. For example, when policy exists, the odds of cooperative behavior relative to competitive behavior are about 53 times higher as the odds without policy ($p < 0.001$) for agents at Parcel 6 and about 7–53 times higher for agents at other parcels (multiplication of the coefficient of policy by the coefficient of interaction). But policy alone (without increasing the number of sensors and frequency of sensing) does not significantly affect the likelihood of extreme behavior (hypercooperative or hypercompetitive).

Most significantly, the effects of interactions of policy with the parcel dummy variables (Table 4) are significant for the cooperative group, indicating that policy has heterogonous influence among parcels. In particular, the odds of cooperative behavior relative to competitive behavior tend to be smaller for agents of Parcels 2, 3, 4, and 5 ($p < 0.001$) compared with Parcel 6. However, the interaction terms are not significant for the extreme types of behavior (both hypercooperative or hypercompetitive).

These results support the conclusion that policy incentives generally induce cooperative behavior, but the effect also depends on the location of the agent's parcel in the river network. Policy can induce competitive behavior in upstream agents, but the farthest downstream agents tend to display a greater likelihood of cooperative behavior.

Hypothesis 2: Number of Water Quality Sensors as Well as Their Sensing Frequency Influence Cooperative Behavior

After controlling for the location of agents, it is found that both the number of sensors and the frequency of sensing generally affect the likelihood of cooperative behavior. Adding one sensor (with a policy) has a relatively large effect on inducing cooperative behavior (the odds of cooperative behavior relative to competitive behavior are about 7–53 times higher than those without a policy across the six parcels). The size of the effect diminishes with the addition of a second sensor (the odds are reduced to about 40% for Parcel 6 and more or less for other parcels compared with the one-sensor case, although the odds are still higher than those without a policy). It is also found that one sensor (with a policy) in the river network does not significantly affect the behavior of the extreme types of agents (hypercooperative and hypercompetitive). However, adding a second sensor (low frequency) reduces the odds of hypercooperative behavior relative to competitive behavior for most agents except those at Parcel 5, whose odds are not changed much. Meanwhile, the additional sensor decreases the odds of hypercompetitive behavior for agents spatially close to the sensors (Parcels 3 and 6, $p < 0.001$) but increases the odds for agents farther from the sensors.

Changes in the frequency of sensing produce some surprising results. A high frequency of sensing (four times) reduces the odds of cooperative behavior versus competitive behavior compared with a low frequency of sensing (once), especially for upstream agents (further reduced to about 22.7%–47.6% in comparison with

Parcel 6), although it still increases the odds of cooperative behavior in comparison to the control group (no policy). Moreover, sensing frequency affects the behavior of the extreme types of agents. A high frequency of sensing increases the odds of hypercooperative behavior versus competitive behavior for farther downstream agents (e.g., a 2.1 times increase for Parcel 6) but decreases the odds of hypercooperative behavior for upstream agents. When high-frequency sensing is coupled with two sensors, the odds of hypercooperative behavior for downstream agents are further increased (e.g., about 5.7 times for Parcel 6), but the odds of the upstream agents are still lower than those under two sensors and a low frequency of sensing. On the other hand, a high frequency of sensing with two sensors reduces the odds of hypercompetitive behavior versus competitive behavior for agents spatially closer to the sensors (Parcels 3, 5, and 6) but increases the odds for agents farther from the sensors, again compared with two sensors and a low frequency of sensing.

From these results, it can be concluded that a minimal amount of sensing in addition to incentive manipulation can have large and statistically significant effects in increasing cooperative behavior. An increased number of sensors and frequency of sensing can reduce the likelihood of hypercompetitive behavior for agents closer to the sensors but is not as effective for agents farther from the sensors.

Hypothesis 3: Spatial Locations of the Agents Relative to the Spatial Locations of the Sensors Affect the Induction of Cooperative Behavior

The effects of the interaction between the sensing variables and the parcel dummy variables demonstrate that the spatial location of a non-point-source polluter relative to the sensors in a river network matters. Generally, it was found that when one or two sensors are added, agents of all six parcels are more likely to behave cooperatively, but the effect on inducing cooperative behavior is smaller for agents of upstream parcels than for agents of the parcels farthest downstream. Additionally, the interaction of a high frequency of sensing and the parcel location demonstrates the complex heterogeneity of the location effects for cooperative and hypercooperative agents. A high frequency of sensing tends to decrease the odds of cooperative behavior and hypercooperative behavior (versus competitive behavior) for upstream agents, but may increase or reduce less the odds for downstream agents (e.g., a 2.1 times increase of hypercooperative behavior for Parcel 6). This is likely because the NE level of production is significantly higher for downstream parcels in the high-frequency sensing treatments.

Overall, the authors conclude that sensors paired with incentives increase the likelihood of cooperative behavior in all parcels, but the size of the effect on inducing cooperation is slightly larger in downstream parcels compared with upstream parcels. Moreover, a high frequency of sensing is likely to reduce cooperative and hypercooperative behavior in upstream parcels although it might increase or reduce these behaviors in downstream parcels.

Discussion

There are broad (1) watershed management and policy, and (2) decision behavioral and game theoretical implications of the results presented in this paper that are discussed in this section. Methodological limitations are also discussed.

Watershed Management and Policy Implications

Overall, the incentive-based policy was found to induce cooperative behavior, and a low frequency of sensing with the incentives in place can have large effects on increasing cooperative behaviors. Minimal investments in setting up advanced environmental monitoring/sensing systems can potentially have large effects in improving water quality. The experimental data, however, also show that increasing the number of sensors (from one to two sensors) and frequency of sensing (from low to high) does not necessarily increase cooperative behavior. This finding suggests that additional sensors and frequency of sensing may diminish the likelihood of cooperative behavior in maintaining water quality; however, they may reduce the likelihood of extreme behavior.

This implication needs to be further tested in a more elaborate game design and with more levels of the frequency of sensing. More elaborate watershed-scale field experiments and agent-based models can be used in future studies to ascertain the extent and scale of these sensing effects, both with and without incentive manipulation. Such studies may enable identification of optimal environmental monitoring and sensor network designs in specific watershed management and policy design contexts. For enhanced ability to draw general conclusions, future research may change assumptions in the nutrient dynamics model that is used to simulate pollution concentration (and hence damages from pollution). For example, the model in this study assumes large constant background loadings into the surface water system. It may be interesting to modify the parameters to reduce the background loading; this may enhance the diffusion effect and alter how pollutants from different parcels diffuse over time, thereby changing polluters' incentives across space.

Behavioral Implications

Spatially explicit social dilemma and principal agent games, a variant of which has been presented in this study, can generate useful insights about the design of incentive mechanisms and sensor monitoring systems in the management and governance of human–environmental systems. Incorporation of spatially sensitive upstream and downstream dynamics in the management of non-point-source polluters, in particular design of environmental monitoring systems and incentive mechanisms, can be adapted to specific watershed topologies and stream networks. For example, the authors found that when one or two sensors are added (coupled with policy), agents of all six parcels were more likely to behave cooperatively, but the effect on inducing cooperative behavior was smaller for agents of upstream parcels than for agents of the parcel farthest downstream agents. Further, under a high frequency of sensing, agents of downstream parcels tended to perform more cooperatively than agents of upstream parcels. Conversely, a lack of sensing and absence of incentive manipulation likely leads to more predictable Nash/competitive behaviors.

Location of players vis-a-vis sensor locations matters significantly in inducing cooperative versus competitive strategies. More specific watershed-scale studies need to be designed to tease out the location-specific effects of installing environmental monitoring sensor systems across upstream and downstream gradients. Experimental results from this study confirm the value of installing cyberinfrastructure through monitoring and sensor systems, but more studies, ideally with spatially explicit dynamic social dilemma game structures, could improve current understanding of the location effects of sensors on non-point-source polluters in the management of watersheds.

The power of advanced in situ sensor networks to provide water quality data to support science-based decisions by stakeholders and resource managers appears intuitive and compelling. As discussed in terms of the results of this experimental study, other similar behavioral and game theoretical studies (mentioned in the "Introduction") have suggested that the responses of private individuals and businesses to more detailed and accurate environmental information is variable and complex. Sensor-based water quality information can be gathered more often, accommodate a greater number of dimensions, have a higher resolution and/or accuracy, and cover a larger geographic and temporal environmental or social scope. Each of these factors can change human behavior, and the underlying premise may be that such changes in behavior will lead to a more sustainable use of natural resources. However, a behavioral change across the complex spatial terrains of watersheds in response to a given change in information is not guaranteed. Prior research, for example, has found that social information that can be interpreted as normative (such as choices of peers) can have strong and lasting effects on inducing cooperative behaviors (e.g., environmental conservation), perhaps stronger and longer lasting than the frequency or dimensional richness of scientifically measured environmental information (Dennis et al. 1990; Nolan et al. 2008; Peschiera and Taylor 2012). Location also matters, because the interaction effect of incentive mechanisms of resource users with sensor locations used by watershed managers has a strong effect in steering the multiscale environmental conditions.

Other studies in groundwater management, irrigation, and energy conservation domains investigating the effects of sensor information on strategic behaviors have found similar effects. For instance, greater availability of information regarding the condition of an aquifer could lead farmers to increase their water withdrawals in the short run at the expense of future water availability (Li et al. 2014). That could lead to lower seasonal water levels and overcapitalized farmers (e.g., Molden et al. 2003; Cosgrove and Rijsberman 2014). Similarly, provision of information that the level of pollution in a water body is low could encourage some farmers to forgo expensive pollution-abatement activities. In irrigation games, Bell et al. (2015) found that upstream players in the Indus Basin, whose water supply was unaffected by scarcity, did not change their competitive behavior in response to information that downstream players were not receiving adequate water. In contrast, D'Exelle et al. (2012) found that upstream irrigation players in Tanzania played cooperatively for enabling equitable sharing of water resources between upstream and downstream players. Studies of residential energy consumption have shown that more frequent provisions of usage and cost information can lead to declines in energy use (Petersen et al. 2007), but more frequent information often does not improve behavior or improves it only temporarily (Abrahamse et al. 2005).

Moreover, information that covers a larger spatial scale could have unintended spatial effects. For instance, information regarding high pollution levels in upper reaches of a watershed could discourage farmers in lower reaches from adopting conservation practices. Likewise, owners of lakefront homes could forgo adopting best management practices to control runoff or maintain septic systems if they view actions of their neighbors as making their individual actions irrelevant. People also sometimes misinterpret the risks associated with such information. For example, information that fosters new media reports regarding high levels of *E. coli* could lead to visitors shunning previously popular tourist destinations because they believe they are no longer safe even though the objective degree of risk has not changed. Broader decision behavior and game theoretic research in the future across a variety of watershed management conditions can explain the amplifying or dampening effects of strategic behavioral responses to incentive mechanisms and the placement of sensor monitoring systems.

Methodological Limitations

The proposed clustering method to isolate behavioral strategies as a function of NE does not work well for situations that involve multiple Nash equilibria. In such cases, distance from the NE may yield nonunique solutions. In addition, both the NE and the production decisions can change across treatment settings. The cluster analysis is based on the distance between production choices and relevant NE, and therefore the observed changes in behavioral types across treatments and scenarios are due to both because of changes in production decisions and changes in the NE. Future research should evaluate the relative importance of each of these avenues. Another important limitation of the game concerns the theoretical assignment of cooperative behaviors to agents despite the fact that agents do not see each other's moves or production levels. Treatments that enable communication between agents were not tested in this experiment but represent a fruitful area for future research (e.g., Guilfoos et al. 2019).

In addition, although logistic regression is a popular and effective approach for modeling categorical outcomes in classification problems, the method still relies on certain assumptions. For example, the model has a parametric formulation, and a linear relationship between the log odds and the predictors is assumed in the study. If these parametric assumptions are violated, the model could be misspecified. The model was diagnosed based on Pearson (chi-square) residuals, and severe violations were not observed. Because the dimension of the predictors in this study is relatively high, the authors consider that the linear assumption is relatively mild. Nevertheless, the conclusion is data-driven and model-based. There are other nonparametric techniques for classification problems, such as decision trees and random forests, which are worthy of future investigation. These methods are flexible and useful for the purpose of prediction but can be less effective for interpretation.

Conclusion

This study investigated how monitoring and policy influence heterogeneous agent behavior and found that policy coupled with sensing information induces cooperative behavior, but the effect is conditional on the location of the agent's property in the river network. Upstream agents with relatively more influence on downstream ambient pollution display marginally greater likelihood of behaving competitively, whereas downstream agents with relatively less influence on ambient pollution display marginally greater likelihood of behaving cooperatively. However, policy alone (without increasing number of sensors and frequency of sensing) does not significantly affect the likelihood of extreme behavior (hypercooperative or hypercompetitive). Further, introduction of a minimal amount of sensing with incentive manipulation can have large significant effects in increasing cooperative behaviors. The addition of another sensor or higher frequency of sensing does not necessarily increase cooperative behavior. The implication here is straightforward. Initial investments in sensor technologies that make it feasible to provide financial incentives to producers that are conditional on ambient pollution outcomes can generate significant reductions in water pollution. Additional investments in additional sensor locations or high-frequency sensors, however, may not yield benefits in terms of improved water quality and therefore are unlikely to be cost-effective.

In addition, there are important network effects to inducing cooperative behaviors, and those linkages through water and transport flows need to be considered when designing water quality monitoring systems and incentive mechanisms. The complexities faced by agents in a watershed suggest that policy impacts on the type of behavioral responses vary by spatial location, i.e., not just that incentives change across spatial location but that the responses to those incentives are location-dependent as well. This implies that a deeper understanding of the design of environmental monitoring systems in relation to decision behavioral responses under specific incentive mechanisms is needed for watershed management and planning.

Data Availability Statement

All data, models, and code generated or used during the study are available from the corresponding author by request.

Acknowledgments

Support was provided by EPSCoR with funds from the National Science Foundation Grants IIA-1330446 and OIA-1556770.

Notation

The following symbols are used in this paper:

$C(x_1, \ldots, x_5, x_6)$ = measured concentration of pollution;
$C(x_1, x_2, x_3)$ = measured concentration of pollution from the sensor between Parcels 3 and 4;
$C(x_4, \ldots, x_6)$ = measured concentration of pollution from the downstream sensor;
$C_{\max}(x_1, \ldots, x_6)$ = maximum concentration;
$D(x_1, \ldots, x_6)$ = economic damage;
diff_t = difference between an agent's actual production decision and Nash equilibrium;
I_n = private income of agent n;
k = homogenous parcels;
n = agent;
Pr = probability;
$\mathrm{production}_t$ = production decision;
R_i = reaches;
R_k = state of the river downstream from the last parcel;
R_0 = headwater;
t = round in which the decision is made;
TB = net social benefit;
$\mathrm{TS}(x_1, \ldots, x_6)$ = tax/subsidy charged on an individual agent;
$\mathrm{TS}_{\mathrm{down}}(x_4, \ldots, x_6)$ = tax/subsidy charged on an individual agent from Parcels 4 to 6;
$\mathrm{TS}_{\mathrm{up}}(x_1, x_2, x_3))$ = tax/subsidy charged on an individual agent from Parcels 1 to 3;
x_n = decision variable;
$x_1, \ldots, x_6$ = socially optimal production as a vector of production decisions;
$Y = \{1, 2, 3, 4\}$ = behavioral groups estimated from the clustering algorithm;
Z = predictor vector that contains elements;
$z_1, z_2, \ldots, z_5$ = Parcels 1 to 5;
z_6 = policy;
z_7 = two sensors and measuring the additional effect due to two sensors;
z_8 = high frequency;
$z_9, z_{10}, \ldots, z_{13}$ = interactions between parcel numbers and policy;
$z_{14}, z_{15}, \ldots, z_{18}$ = interactions between parcel numbers and two sensors;
$z_{19}, z_{20}, \ldots, z_{23}$ = interactions between parcel numbers and high frequency;
z_{24} = interaction for two sensors and high frequency;
$\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,24}$ = coefficients of the variables in the multinomial logistic model; and
μ = intercepts.

Supplemental Materials

Figs. S1–S6 and Tables S1–S26 are available online in the ASCE Library (www.ascelibrary.org). Supplemental Materials also present information on formal hypotheses, MMLM selection procedures, game design and calculation of NE, and experiment instructions as presented to the study participants.

References

Abrahamse, W., L. Steg, C. Vlek, and T. Rothengatter. 2005. “A review of intervention studies aimed at household energy conservation.” *J. Environ. Psychol.* 25 (3): 273–291. https://doi.org/10.1016/j.jenvp.2005.08.002.

Anderberg, M. R. 1973. *Cluster analysis for applications: Monographs and textbooks on probability and mathematical statistics*. New York: Academic Press.

Beggs, J. 1988. “A simple model for heterogeneity in binary logit models.” *Econ. Lett.* 27 (3): 245–249. https://doi.org/10.1016/0165-1765(88)90178-4.

Bell, A. R., M. A. A. Shah, A. Anwar, and C. Ringler. 2015. “What role can information play in improved equity in Pakistan’s irrigation system? Evidence from an experimental game in Punjab.” *Ecol. Soc.* 20 (1): 51. https://doi.org/10.5751/ES-07368-200151.

Bergamaschi, B. A., et al. 2012. “Mercury dynamics in a San Francisco estuary tidal wetland: Assessing dynamics using in situ measurements.” *Estuaries Coast* 35 (4): 1036–1048. https://doi.org/10.1007/s12237-012-9501-3.

Berry, M. J., and G. Linoff. 1997. *Data mining techniques: For marketing, sales, and customer support*. Indianapolis: Wiley.

Boyd, J., and J. Mellman. 1980. “The effect of fuel economy standards on the U.S. automotive market: A hedonic demand analysis.” *Transp. Res.* 14 (5–6): 367–378. https://doi.org/10.1016/0191-2607(80)90055-2.

Calinski, T., and J. Harabasz. 1974. “A dendrite method for cluster analysis.” *Commun Stat* 3 (1): 1–27. https://doi.org/10.1080/03610927408827101.

Cardell, N., and F. Dunbar. 1980. “Measuring the societal impacts of automobile downsizing.” *Transp. Res.* 14 (5–6): 423–434. https://doi.org/10.1016/0191-2607(80)90060-6.

Chapra, S. C., G. J. Pelletier, and H. Tao. 2008. *QUAL2K: A modeling framework for simulating river and stream water quality, Version 2.11: Documentation and users manual*. Medford, MA: Tufts Univ.

Cosgrove, W. J., and F. R. Rijsberman. 2014. *World water vision: Making water everybody’s business*. New York: Routledge.

Dennis, M. L., E. J. Soderstrom, W. S. Koncinski, and B. Cavanaugh. 1990. “Effective dissemination of energy-related information: Applying social psychology and evaluation research.” *Am. Psychol.* 45 (10): 1109–1117. https://doi.org/10.1037/0003-066X.45.10.1109.

D’Exelle, B., E. Lecoutere, and B. Van Campenhout. 2012. “Equity-efficiency trade-offs in irrigation water sharing: Evidence from a field lab in rural Tanzania.” *World Dev.* 40 (12): 2537–2551. https://doi.org/10.1016/j.worlddev.2012.05.026.

Downing, B. D., B. A. Pellerin, B. A. Bergamaschi, J. F. Saraceno, and T. E. Kraus. 2012. "Seeing the light: The effects of particles, dissolved materials, and temperature on in situ measurements of DOM fluorescence in rivers and streams." *Limnol. Oceanogr. Methods* 10 (10): 767–775. https://doi.org/10.4319/lom.2012.10.767.

Enberg, J., P. Gottschalk, and D. Wolf. 1990. "A random effects logit model of work welfare transitions." *J. Econom.* 43 (1–2): 63–75. https://doi.org/10.1016/0304-4076(90)90107-5.

Estivill-Castro, V. 2002. "Why so many clustering algorithms—A position paper." *SIGKDD Explor. Newsl.* 4 (1): 65–75. https://doi.org/10.1145/568574.568575.

Everitt, B. 2011. *Cluster analysis*. Chichester, UK: Wiley.

Fiebig, D. G., M. P. Keane, J. Louviere, and N. Wasi. 2010. "The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity." *Marketing Sci.* 29 (3): 393–421. https://doi.org/10.1287/mksc.1090.0508.

Florsheim, J. L., B. A. Pellerin, N. H. Oh, N. Ohara, P. A. M. Bachand, S. M. Bachand, B. A. Bergamaschi, P. J. Hernes, and M. L. Kavvas. 2011. "From deposition to erosion: Spatial and temporal variability of sediment sources, storage, and transport in a small agricultural watershed." *Geomorphology* 132 (3): 272–286. https://doi.org/10.1016/j.geomorph.2011.04.037.

Giuliani, M., A. Castelletti, F. Amigoni, and X. Cai. 2014. "Multiagent systems and distributed constraint reasoning for regulatory mechanism design in water management." *J. Water Resour. Plan. Manage.* 141 (4): 04014068. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000463.

Groisman, P. Y., R. W. Knight, T. R. Karl, D. R. Easterling, B. Sun, and J. H. Lawrimore. 2004. "Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations." *J. Hydrometeorol.* 5 (1): 64–85. https://doi.org/10.1175/1525-7541(2004)005<0064:CCOTHC>2.0.CO;2.

Guilfoos, T., H. Miao, S. Trandafir, and E. Uchida. 2019. "Social learning and communication with threshold uncertainty." *Resour. Energy Econ.* 55 (Feb): 81–101. https://doi.org/10.1016/j.reseneeco.2018.11.003.

Halkidi, M., M. Vazirgiannis, and I. Batistakis. 2000. "Quality scheme assessment in the clustering process." In Vol. 1910 of *Proc., 4th European Conf., PKDD Principles of Data Mining and Knowledge Discovery, Lecture Notes in Computer Science*, edited by D. A. Zighed, J. Komorowski, and J. Żytkow. Berlin: Springer.

Hastie, T., R. Tibshirani, and J. Friedman. 2001. *The elements of statistical learning*. New York: Springer.

Hernes, P. J., R. G. Spencer, R. Y. Dyda, B. A. Pellerin, P. A. Bachand, and B. A. Bergamaschi. 2008. "The role of hydrologic regimes on dissolved organic carbon composition in an agricultural watershed." *Geochim. Cosmochim. Acta* 72 (21): 5266–5277. https://doi.org/10.1016/j.gca.2008.07.031.

Huang, G. H., and J. Xia. 2001. "Barriers to sustainable water-quality management." *J. Environ. Manage.* 61 (1): 1–23. https://doi.org/10.1006/jema.2000.0394.

Jain, D., N. Vilcassim, and P. Chintagunta. 1994. "A random coefficients logit brand choice model applied to panel data." *J. Bus. Econ. Stat.* 12 (3): 317–328. https://doi.org/10.1080/07350015.1994.10524547.

Johnson, R. A., and D. W. Wichern. 2002. *Applied multivariate statistical analysis*. 5th ed. Englewood Cliffs, NJ: Prentice Hall.

Kaushal, S. S., G. E. Likens, N. A. Jaworski, M. L. Pace, A. M. Sides, D. Seekell, K. T. Belt, D. H. Secor, and R. L. Wingate. 2010. "Rising stream and river temperatures in the United States." *Front. Ecol. Environ.* 8 (9): 461–466. https://doi.org/10.1890/090037.

Keritino, B., O. Feri, and L. Cheler. 2015. "Proposing a graph mining method on medical networks." *Int. J. Bioinf. Biomed. Eng.* 1 (2): 93–100.

Kraus, T. E., C. A. Anderson, K. Morgenstern, B. D. Downing, B. A. Pellerin, and B. A. Bergamaschi. 2010. "Determining sources of dissolved organic carbon and disinfection byproduct precursors to the McKenzie River, Oregon." *J. Environ. Qual.* 39 (6): 2100–2112. https://doi.org/10.2134/jeq2010.0030.

Krzanowski, W. J., and Y. T. Lai. 1988. "A criterion for determining the number of groups in a data set using sum-of-squares clustering." *Biometrics* 44 (1): 23–34. https://doi.org/10.2307/2531893.

Li, J., H. A. Michael, J. M. Duke, K. D. Messer, and J. F. Suter. 2014. "Impact of risk information in a spatially explicit groundwater environment with contamination risk: Experimental evidence." *Water Resour. Res.* 50: 6390–6405. https://doi.org/10.1002/2013WR015230.

Lu, Y., N. Nakicenovic, M. Visbeck, and A. S. Stevance. 2015. "Policy: Five priorities for the UN sustainable development goals." *Nature News* 520 (7548): 432. https://doi.org/10.1038/520432a.

Madani, K., and J. R. Lund. 2011. "California's Sacramento–San Joaquin delta conflict: From cooperation to chicken." *J. Water Resour. Plan. Manage.* 138 (2): 90–99. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000164.

Miao, H., J. R. Fooks, T. Guilfoos, K. D. Messer, S. M. Pradhanang, J. F. Suter, S. Trandafir, and E. Uchida. 2016. "The impact of information on behavior under an ambient-based policy for regulating nonpoint source pollution." *Water Resour. Res.* 52 (5): 3294–3308. https://doi.org/10.1002/2015WR018142.

Molden, D., H. Murray-Rust, R. Sakthivadivel, and I. Makin. 2003. "A water-productivity framework for understanding and action." In *Water productivity in agriculture: Limits and opportunities for improvement*, edited by J. W. Kijne, R. Barker, and D. Molden, 1–18. Wallingford, UK: CABI International Water Management Institute.

Nevo, A. 2001. "Measuring market power in the ready-to-eat cereal industry." *Econometrica* 69 (2): 307–342. https://doi.org/10.1111/1468-0262.00194.

Nolan, J. M., P. W. Schultz, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius. 2008. "Normative social influence is underdetected." *Pers. Soc. Psychol. Bull.* 34 (7): 913–923. https://doi.org/10.1177/0146167208316691.

Norton, B. G. 2015. *Sustainable values, sustainable change: A guide to environmental decision making*. Chicago: University of Chicago Press.

Pellerin, B. A., B. D. Downing, C. Kendall, R. A. Dahlgren, T. E. Kraus, J. Saraceno, R. G. Spencer, and B. A. Bergamaschi. 2009. "Assessing the sources and magnitude of diurnal nitrate variability in the San Joaquin River (California) with an in situ optical nitrate sensor and dual nitrate isotopes." *Freshwater Biol.* 54 (2): 376–387. https://doi.org/10.1111/j.1365-2427.2008.02111.x.

Pellerin, B. A., J. F. Saraceno, J. B. Shanley, S. D. Sebestyen, G. R. Aiken, W. M. Wollheim, and B. A. Bergamaschi. 2012. "Taking the pulse of snowmelt: In situ sensors reveal seasonal, event and diurnal patterns of nitrate and dissolved organic matter variability in an upland forest stream." *Biogeochemistry* 108 (1–3): 183–198. https://doi.org/10.1007/s10533-011-9589-8.

Peschiera, G., and J. E. Taylor. 2012. "The impact of peer network position on electricity consumption in building occupant networks utilizing energy feedback systems." *Energy Build.* 49 (Jun): 584–590. https://doi.org/10.1016/j.enbuild.2012.03.011.

Petersen, J. E., V. Shunturov, K. Janda, G. Platt, and K. Weinberger. 2007. "Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives." *Int. J. Sustainability Higher Edu.* 8 (1): 16–33. https://doi.org/10.1108/14676370710717562.

Reader, S. 1993. "Unobserved heterogeneity in dynamic discrete choice models." *Environ. Plann. A* 25 (4): 495–519. https://doi.org/10.1068/a250495.

Revelt, D., and K. Train. 1998. "Mixed logit with repeated choices: Households' choices of appliance efficiency level." *Rev. Econ. Stat.* 80 (4): 647–657. https://doi.org/10.1162/003465398557735.

Rode, M., et al. 2016. "Sensors in the stream: The high-frequency wave of the present." *Environ. Sci. Technol.* 50: 10297–10307. https://doi.org/10.1021/acs.est.6b02155.

Saraceno, J. F., B. A. Pellerin, B. D. Downing, E. Boss, P. A. Bachand, and B. A. Bergamaschi. 2009. "High-frequency in situ optical measurements during a storm event: Assessing relationships between dissolved organic matter, sediment concentrations, and hydrologic processes." *J. Geophys. Res. Biogeosci.* 114 (11): G00F09. https://doi.org/10.1029/2009JG000989.

Segerson, K. 1988. "Uncertainty and incentives for nonpoint pollution-control." *J. Environ. Econ. Manage.* 15 (1): 87–98. https://doi.org/10.1016/0095-0696(88)90030-7.

Small, K. A., and E. T. Verhoef. 2007. *The economics of urban transportation*. New York: Routledge.

Spraggon, J. M. 2002. "Exogenous targeting instruments as a solution to group moral hazards." *J Public Econ.* 84 (3): 427–456. https://doi.org/10.1016/S0047-2727(01)00088-3.

Staudinger, M. D. G., S. L. Carter, F. S. Chapin, P. Kareiva, M. Ruckelshaus, and B. A. Stein. 2013. "Impacts of climate change on biodiversity, ecosystems, and ecosystem services: Technical input to the 2013 National Climate Assessment, cooperative report to the 2013 National Climate Assessment." Accessed March 19, 2020. https://pubs.er.usgs.gov/publication/70039460.

Train, K. 1998. "Recreation demand models with taste differences over people." *Land Econ.* 74 (2): 230–239. https://doi.org/10.2307/3147053.

Waylen, K. A., et al. 2019. "Policy-driven monitoring and evaluation: Does it support adaptive management of socio-ecological systems?" *Sci. Total Environ.* 662 (Apr): 373–384. https://doi.org/10.1016/j.scitotenv.2018.12.462.

Weel, J., and K. McCabe. 2015. "Willow: Experiments in Python." Accessed March 19, 2020. http://econwillow.sourceforge.net/.

Yu, Y., J. Zhao, D. Li, and Z. Wang. 2019. "Effects of hydrologic conditions and reservoir operation on transboundary cooperation in the Lancang–Mekong River Basin." *J. Water Resour. Plan. Manage.* 145 (6): 04019020. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001075.

Zia, A., et al. 2016. "Coupled impacts of climate and land use change across a river–lake continuum: Insights from an integrated assessment model of Lake Champlain's Missisquoi Basin, 2000–2040." *Environ. Res. Lett.* 11 (11): 114026. https://doi.org/10.1088/1748-9326/11/11/114026.