Towards an understanding of the performance of ambient tax mechanisms in the field: Evidence from dairy farmers

Jordan F. Suter and Christian A. Vossler

Correspondence should be directed to Christian A. Vossler, 523 Stokely Management Center, Department of Economics, University of Tennessee, Knoxville, TN 37996; 865-974-1699 (phone); 865-974-4601 (fax); cvossler@utk.edu (email). Jordan Suter is assistant professor, Department of Economics, Oberlin College. Christian Vossler is associate professor, Department of Economics, and Research Fellow, Howard H. Baker Jr. Center for Public Policy, University of Tennessee. Funding for this research was provided by USEPA grant R830989-2002. We are indebted to Greg Poe, who was fundamental in implementing the experiments with dairy farmers, and who provided critical input on all facets of the project. We greatly appreciate Jane Maestro-Scherer's survey expertise and efforts in recruiting and coordinating the experiments with dairy farmers. We are also thankful to Cornell University and Cornell Cooperative Extension personnel Brain Aldrich, Fay Benson, Kerry Case, George Conneman, Janice Degni, Wayne Knoblauch, Monika Roth, Ynte Schukken, and Joan Taber for their help in identifying and recruiting dairy farmers to participate in the experiments. Of course this research could not have been completed without the patience and participation of dairy farmers and students in our experiments, for which we are very grateful. Bill Schulze and Kathy Segerson provided critical input in earlier stages of this research.

Abstract

Using a design characterized by heterogeneous firms and stochastic ambient pollution, this study explores how results from ambient tax experiments using student subjects translate to a richer field context with dairy farmers. Results suggest that the ambient tax induces group-level compliance among students and farmers. However, relative to students, farmers operating "small" firms pollute less and farmers operating "large" firms tend to pollute more. Deviations from theory among farmers are tied to beliefs over the impacts of farming on water pollution, and knowledge of neighbors' pollution. This study highlights the importance of framed field experiments in the policy test-bedding process.

Key words: ambient tax; dairy farmers; laboratory experiment; framed field experiment; firm heterogeneity; nonpoint source pollution

JEL codes: Q58, H23, Q53, Q52, C91, C92

Running head: Ambient tax experiments with dairy farmers

It is recognized that laboratory experiments with student subjects are useful for test-bedding potential policies for which naturally-occurring data is lacking. Such experiments can highlight the empirical tradeoffs among possible policy instruments. In designing an experiment, often out of necessity, the researcher must consider important characteristics of the policy setting that the theory is silent about. As a result, experimental findings often motivate important refinements to theory. As argued clearly by Loewenstein (1999), however, laboratory experiments are "particularly vulnerable" to problems of external validity. Among potential concerns are the "context in which the decision is embedded" and the "manner in which participants and tasks are selected" (Levitt and List 2007). Indeed, relative to students, participants selected to represent target populations may bring prevailing purviews, professional bias and confounding norms and conventions, despite the experimental rules articulated in the instructions. Considerations of policy relevance have contributed to a methodological trend combining lab and field experiments to investigate economic behavior, permitting sharper and more convincing inference (Harrison and List 2004).

In this study we report the results of policy test-bed experiments on an ambient tax mechanism designed to regulate nonpoint source water pollution.[1] The main focus is on whether desirable empirical results from experiments based on student subjects translate to a richer field environment with Upstate New York dairy farmers, a targeted group of nonpoint polluters. As we show in this paper, the particular mechanism we test – a pure ambient tax with a payment threshold below the pollution standard – leads to (near) optimal behavior among groups of student participants. Such a mechanism is a prime candidate for field testing. As Vernon Smith (1982, p. 397) writes, "If the [economic] theory is not falsified in several replications, then one can begin to ask whether the results generalize to different subject pools and field environments."

Experimental researchers have explored a variety of ambient-based policies for addressing nonpoint source pollution, initially motivated from the theoretical work of Segerson (1998) and Xepapedeas (1991) (e.g., Alpizar, Requate and Schram 2004; Cochard, Willinger and Xepapadeas 2005; Poe *et al.* 2004; Spraggon 2002, 2004; Suter *et al.* 2008, 2010; Suter, Vossler and Poe 2009; Vossler *et al.* 2006). These experiments have uncovered some important pitfalls associated with specific ambient-based policies, such as the potential for collusive behavior in policies that involve subsidies for over-compliance, and the potential for predatory behavior when there are a few large polluters in a heterogeneous polluter group. At the same time, some appropriately designed ambient-based tax policies have performed well in terms of achieving ambient pollution objectives. The mechanism we use in this study represents what has performed best in previous experiments where participants are not allowed to communicate. In particular, as shown in Suter *et al*. (2008, 2010), a linear ambient tax with a payment threshold (sufficiently) less than the ambient pollution standard can achieve the pollution standard (on average), and is highly efficient. In contrast to the Suter *et al.* (2008, 2010) experiments, the experimental design utilized in this study introduces firm-level heterogeneity.

Our experimental design places the decision task in a context congruent with the relevant policy setting. Parameters are chosen based on information gathered from the

Upstate NY dairy industry. Farmer participants are aware that the experiments are funded by the EPA, in part to highlight the potential policy relevance of the research. In the student experiments, students are exogenously assigned to a firm size. In our experiments with dairy farmers, each farmer's firm size is endogenously determined by the size of his actual dairy herd. While such context-rich experiments may allow only limited inference about behavior in other settings, they are employed in this work because they represent, in the view of some, the most appropriate way to draw inferences about behavior that are valid for investigating policy design (Lowenstein 1999; Plott and Porter 1996). Congruent with the advice of Herberich, Levitt and List (2009), cooperative extension personnel were integral to the implementation of the experiment and recruitment of dairy farmers. With an eye toward aiding future researchers planning experiments with farmers, we provide details on our recruitment process and experiment implementation.

There are several key experimental results. First, we confirm that the chosen ambient tax mechanism meets the ambient standard on average, even with parameter heterogeneity, using student participants. Second, we find that this desirable (average) result continues to hold for dairy farmers. Third, we do – however – find important differences at the individual level across participant pools. Relative to students, small farmers tend to choose significantly lower emissions, while operators of large farms emit significantly more. This behavioral heterogeneity has an adverse impact on industry composition, as it led to the bankruptcy of small farms in two of the eight farmer participant groups. Based on the results of a post-experiment survey, we find that farmers who believe they can determine their neighbors' pollution or believe that agriculture contributes to water pollution choose statistically lower emissions in the experiment. Although we do not find a statistical link between farmer decisions and beliefs regarding the policy usefulness of experiments, we acknowledge that some choices may have been motivated by a desire to send signals to regulators regarding the implementation of ambient-based policies.

Review of relevant experimental literature

Laboratory experiments in economics typically rely on subject pools composed exclusively of students primarily due to their convenience for academics (Harrison and List 2004). For purposes of testing economic theories, experiments with diverse or targeted subject pools are generally not seen as essential as economic theories are meant to apply to the entire population, and therefore students are not expected to respond any differently (Croson 2005). In cases where experiments are being used to potentially inform policy, however, issues of parallelism are paramount for establishing external validity. In this section we first highlight policy experiments that, as is the case with our study, compare results from student and nonstudent subjects. Then, we briefly review findings from experiments that utilized participants from agriculture.

A recent article by Normann and Ricciuti (2009) surveys dozens of controlled experiments specifically designed to inform policy. Among these, only one compares student and nonstudent decisions.[2] In particular, Krause, Chermak and Brookshire (2003) conducted common-pool resource experiments with students, working adults and retirees. They find significant responsiveness to price in all three pools, but differences in the magnitudes of the effect. The authors then demonstrate how this information can be used to design more effective water conservation policies. In more recent work, in experiments designed to test the effects of institutional change enacted on the wholesale electricity market during California's "electricity crisis," Vossler *et al.* (2009) find that comparative statics results are confirmed with both students and professional electricity traders. Also, Vossler and McKee (2012) find similar behavior across students and university staff in an income tax compliance experiment that explores the role of information services. The above examples are consistent with the broader (i.e. non-policy experiment) literature, which largely finds there to be minimal subject pool effects (see Ball and Cech 1996).

Turning to the agricultural context, there are a few policy experiments that compare students and agricultural professionals. Cummings, Holt, and Laury (2004) evaluate various auction designs for water acquisition in the State of Georgia using laboratory experiments with a set of student participants and then execute similar experiments in the field with a subject pool made up primarily of farmers. Although the results were not directly comparable, they find generally that individual behavior in the field experiments is very similar to that observed in the lab experiments. Interestingly, the lessons learned from their experimental treatments helped to tailor the design of the actual auction process used to purchase water rights from farmers later that year.

Duquette, Higgins and Horowitz (2012) measure discount rates amongst a group of farmers targeted by the American Farmland Trust (AFT) that had yet to adopt certain best management practices and a group of farmers that had already signed up to receive regular mailings from the AFT. The estimated discount rate across both groups was 34%, with higher discount rates observed for the group of farmers that had not yet adopted best management practices. These results allude to the fact that differences in time preferences may help to explain participation rates in conservation related programs and that conservation programs may improve their efficacy by altering the timing of compensation, given the high discount rates observed across the entire sample.

Herberich and List (2012) compare risk preferences of students in economics classes and farmers in the state of Illinois in an attempt to understand how these preferences might influence participation in carbon offset markets. The experimental evidence suggests that the sample of farmers was slightly more risk averse than the students, although the relatively small sample size (41 farmers and 27 students – of which a number of observations were not useable) makes further analysis difficult. A small subject pool (18 Ohio farmers and 15 Ohio State University students) also precluded direct comparison of the two groups in an experiment testing the use of auctions to reduce nonpoint source agricultural pollution reported in Taylor *et al.* (2004), although the authors characterize the results from the groups as being similar.

Studies in developing countries have frequently utilized subject pools composed of decision makers employed in agriculture. A review by Cardenas and Carpenter (2008) of social dilemma experiments finds that cooperation in widely-studied voluntary contribution mechanism (VCM) and common pool resource (CPR) games tends to be higher and more likely to be sustained with subject pools in developing economies than amongst college-aged participants in the United States. The authors attribute this outcome to the greater importance of social norms for the provision of public goods and maintenance of common pool resources in developing economies, which often lack formal institutions. They also note that variation in the magnitude of cooperation across studies is likely tied to the nature of social interactions that the participants engage in outside of the lab as well as demographic characteristics. For example, they specifically mention a study by Henrich *et al*. (2001) in which remote slash and burn farmers in Peru with low levels of social interaction exhibited low rates of cooperation in a VCM game. In a more recent VCM game conducted using participants from 50 different farmer associations in Uganda, Grossman and Baldassarri (2012) report further evidence that the contribution decisions of agricultural participants are correlated with the type of cooperative institutions they interact with in a naturally occurring setting.

The results from these studies suggest that the diversity of cooperative experiences that decision makers enter the lab with influence their behavior in the experiment. The behavior of the farmers that participated in our experiments was undoubtedly influenced by their social and professional experiences outside of the lab. The experiments that we conduct, however, are different from the typical social dilemma experiment investigated by previous researchers in that socially optimal behavior is individually rational under the ambient tax policy that is implemented in the experiment, as opposed to the typical social dilemma setting where free-riding is individually rational. Without the tax policy in place, behavior in our experiments should not be influenced by cooperative preferences since payoffs are determined completely independently, although it is possible that pro-social preferences may increase a participant's sensitivity to the environmental implications of their decisions.

In work most closely aligned with our study, Alpizar, Requate and Schram (2004) evaluate two ambient-based tax policies using both students and managers of coffee mills in Costa Rica. For both subject pools and mechanisms, the majority of participants play the Nash equilibrium (although there are many deviations). They do find some modest differences across subject pools, with participants under-abating relative to mill managers. In contrast to their study, we test a mechanism with farmers that has been shown to perform very well with students, incorporate important policy-relevant characteristics in the experimental design (e.g. stochastic pollution function; firm-level heterogeneity) and use six-player rather than two-player groups.

Experimental Design

In this section we describe the experimental design. We begin by presenting the theoretical model, followed by a description of the four experimental treatments. We conclude the section with a summary of the experimental protocol.

Decision setting and theoretical predictions

The basic theory underlying the experiment is as follows. Consider a polluter group composed of firms indexed by i =1,…, n. The sole decision task of each firm is to choose a level of emissions, r_i. Lower emissions levels are associated with lower private earnings, $PE_i(r_i)$. Ambient pollution, $x = x(\boldsymbol{r}, e)$, is a function of the vector of emissions, $\boldsymbol{r}$, and a random component, e, that is meant to capture stochastic elements, such as weather, that may influence observed pollution at the ambient monitoring point.

Suppose the regulator wishes to achieve a pollution standard, x^s, on average, which represents a reduction beyond baseline levels. Then, the regulator's problem can be written as:

(1) $\max_{r_i} \sum_{i=1}^{n} PE_i(r_i)$ subject to $\mathrm{E}[x(\boldsymbol{r}, e)] \leq x^s$.

Assuming an interior solution, the first-order conditions to this problem imply that the optimal emissions vector $\boldsymbol{r}^*$ must satisfy $PE_i{}'(r_i{}^*)/\mathrm{E}[x'(\boldsymbol{r}^*, e)] = \lambda^*$ for each firm in the group, where λ^* is the shadow value of the constraint. This condition simply states that the marginal private earnings of an additional unit of pollution for each firm must equal the expected marginal cost of ambient pollution.

To meet the pollution standard, the regulator imposes an ambient tax policy. Under the policy, all firms face a constant marginal tax, $\tau = \lambda^*$, that is levied on each unit of ambient pollution that exceeds a pre-defined tax threshold, x^t, with $x^t \leq x^s$. If ambient pollution is lower than or equal to the threshold, then tax liabilities are zero. Setting the tax threshold slightly below the pollution standard has been shown to engender highly efficient outcomes in past research (Suter *et al.* 2008, 2010). In addition, the tax threshold serves to moderate the magnitude of tax liabilities that firms face under the policy and reduces the range of potential collusive outcomes. Formally, the tax payment, *TP*, charged to *every* firm in the group is given by

(2) $TP = \tau(x - x^t) \quad if \quad x > x^t$

$TP = 0 \qquad\qquad if \quad x \leq x^t$.

It is important to note that the tax payment assessed to a firm is determined by the decisions of every firm in the group, not just its own. Firm earnings under the ambient tax are calculated simply as $PE_i - TP$.

Taking the tax policy as given, the firm's problem is to maximize expected earnings, formally:

(3) $max_{r_i} PE_i(r_i) - F(x^t, \boldsymbol{r}) \cdot \tau(E[x(\boldsymbol{r}, e)] - x^t)$,

where $F(\cdot)$ represents the probability that the tax threshold is exceeded, conditional on the emissions vector. Maximization of (3) leads to the following first-order conditions:

(4) $PE_i{}'(r_i) - F'(x^t, r_i, \boldsymbol{r_{-i}}) \cdot \tau(E[x(r_i, \boldsymbol{r_{-i}}, e)] - x^t) - F(\cdot) \cdot \tau \mathrm{E}[x'(\boldsymbol{r}, e)] = 0 \;\; \forall i$,

where $\boldsymbol{r_{-i}}$ represents the vector of emissions decisions from all other regulated firms. With $F(\cdot) = 1$ (expected ambient pollution will exceed the threshold with certainty) and $F'(\cdot) = 0$ (marginal changes in emissions will not impact the probability of experiencing the tax), the first-order conditions simplify to $PE_i{}'(r_i)/\mathrm{E}[x'(r_i, \boldsymbol{r_{-i}}, e)] = \tau \; \forall i$. With $\tau = \lambda^*$, the first-order conditions mirror those of the social planner's problem. Thus, conditional on beliefs that $\boldsymbol{r_{-i}} = \boldsymbol{r}^*_{-i}$, $r_i = r_i{}^*$ is firm i's best response. When all firms hold these beliefs, $\boldsymbol{r} = \boldsymbol{r}^*$ constitutes a Nash equilibrium of the compliance game. The conditions $F(\cdot) = 1$ and $F'(\cdot) = 0$ require that the tax threshold, x^t, be set sufficiently far below the ambient standard, x^s, so that there is no chance that realized ambient pollution will be less than tax threshold when total emissions by firms is equal to the standard.

Treatments and Parameterization

We use a 2 x 2 experimental design characterized by two subject pools and two types of abatement cost heterogeneity. Our first subject pool includes undergraduates at Cornell University, whereas the second subject pool consists of dairy farmers from nine Upstate NY counties. All experiments involve participants playing the role of firm managers with heterogeneous parameters.

The firms are organized into regulated groups of six, and ambient pollution in a given decision round is calculated as the sum of the individual emissions choices by each of the six group members plus the random term, e. Formally, ambient pollution is determined as

(5) $x = \sum_{i=1}^{6} r_i + e, \; e \sim Uniform[-4,4]$.

Within each group, two participants each are assigned to manage “small”, “medium” and “large” firms. The private earnings functions differ by firm size, and we use two sets of abatement cost functions implemented as distinct treatments. Given that there is likely to be considerable variation in abatement costs across firms in practice, implementing treatments with two different types of cost heterogeneity provides some feedback on the robustness of the ambient tax effects. In the “Heterogeneous I” treatments, small firms have the most elastic abatement cost functions, which theoretically leads to the greatest proportional increase in abatement under the ambient policy we study. In the “Heterogeneous II” treatments, large firms have the most elastic abatement costs and theoretically engage in the greatest proportional emissions reductions. The abatement cost functions do not differ across treatments for medium size firms.

The actual parameterization of the private earnings function used in the experiments is

$$(6) \qquad PE_{s,i}(r_i) = I_s - \left|\phi_s\left(\gamma_s - \frac{r_i}{\alpha_s}\right)\right|^3,$$

where the subscript *s* represents the size of the firm, and I, ϕ, γ, and α are firm size specific parameters, the values of which are provided in table 1 for the two forms of heterogeneity that we implement.

With no regulatory policy in place it is expected that participants will choose the emissions decision that maximizes their private earnings. This implies individual emissions of 10, 20 and 30 units for small, medium, and large firms respectively, leading to expected group emissions of 120 units.

Under the ambient tax policy, which is identical to that described in the last section, the ambient standard, x^s, is 72, the tax threshold, x^t, is 66 and the marginal tax rate, τ, is 2500. The ambient standard of 72 is also the sum of individual emissions under the Nash equilibrium, and represents a 40% reduction below the no policy scenario. Note that if aggregate emissions are in the neighborhood of the ambient standard, the probability of being above the tax threshold is one and marginal changes in emissions do not impact the probability of taxation (i.e., $F(\cdot) = 1$ and $F'(\cdot) = 0$ from equation (4)). The optimal emissions levels for each firm size under each type of heterogeneity are provided in table 1.

For undergraduate subjects, firm size in the experiment is randomly determined. In contrast, using information obtained during the recruitment process, farmers are assigned to the relevant small, medium, and large firm sizes based on the size of their own dairy herd, using information obtained in the recruitment process.

Common protocol

All sessions were conducted at the Cornell Laboratory for Experimental Economics and Decision Research. An experiment session includes participants from the same subject pool. Two groups of six participants, with one group corresponding to each type of firm heterogeneity, play simultaneously within a session. Participants read through a set of printed instructions and then observe a PowerPoint presentation given by the experiment moderator.[3] Given the applied nature of the experiments, the experimental instructions for both student and dairy farmer sessions provide deliberate context. Specifically, subjects are told that they are playing the role of firm managers operating along a common water resource and that the actions of their firm generate water pollution. The context included in the instructions represents an attempt to avoid the creation of an experimental environment "that is too sterile and too abstract from reality" (Shogren 2006), and to increase external validity.

Rather than a continuous array of emissions options, participants choose from a discrete set of emissions decisions derived from equation (6). The specific emissions options and associated "firm earnings" (congruent with *PE*) are provided to participants in an "Emissions Decision Sheet". All decisions in the experiment are made via personal computers outfitted with privacy shields, using a Microsoft Excel interface programmed with Visual Basic for Applications (VBA).

Payments in the experiment are denominated in “tokens,” which are exchanged for dollars according to a privately known exchange rate printed on each participant’s emissions decision sheet. The exchange rate was adjusted by firm size so that given expected emissions decisions all firms would earn approximately equal dollar amounts over the course of the experiment.

Participants complete at least 17 decision rounds in the experiment, divided into two scenarios. Each round is independent from one another, and feedback on ambient pollution and own earnings is provided at the end of each round. The first scenario, a no policy baseline, lasts five decision rounds. Since no regulatory policy is in place, and players are simply expected to maximize private earnings. After the first scenario is completed, participants receive written and verbal instructions on the second scenario, the ambient tax policy. As part of the instructions, participants receive a “Tax Calculation Sheet” that lists the tax payments associated with a range of possible ambient pollution outcomes. Participants then undertake at least 12 rounds under the ambient tax policy. Although the computer screen on which decisions are made lists 30 decision rounds, the terminal round for each session is determined randomly.

Farmer-specific protocol

In contrast to students, who were recruited via email from an existing database maintained by the experimental lab,[4] to recruit dairy farmers we adapted methods used in survey research. Our initial list was created from data on farms that had participated in milk testing programs, and then, working with extension and outreach personnel, the list was narrowed down to individual farmers who were not adversarial or likely to disrupt an experiment related to farm regulation. To those on the list, we mailed out a three-fold brochure and followed this with a phone contact. These efforts, in addition to high earnings expectations, fostered a high level of interest. In particular, very few declined, and nearly one-third of those contacted did actually participate in the study.

The implementation of the experiments also had some unique features. First, participants met at a central location and were shuttled to the experimental lab. Second, participants gathered in a conference room for a short reception. This reception allowed participants to talk with one another and also to meet the experiment moderators. Refreshments were provided, along with a souvenir (a Cornell University hat). Third, the participants were directed to the experimental lab.

In addition to the usual experiment prologue, participants were informed that the experiments were part of an EPA-funded initiative charged with using experiments to better understand the performance of proposed tax, subsidy, and voluntary policies for controlling nonpoint source pollution. The mention of non-tax policies was deliberate (and truthful), and intended to thwart particular concerns over the ambient tax mechanism. Participants were made aware that firm sizes assigned in the experiment were based on the size of their dairy herd.

In the first session run, a few participants had serious difficulty negotiating the computer interface and as a result only eight rounds were completed.[5] In light of this, for those needing assistance in subsequent sessions, we recruited undergraduate agribusiness students to enter experiment decisions. The student helpers were instructed not to provide input/advice regarding experiment decisions, but were allowed to answer clarification questions. Participants wrote down the decisions to be entered by their helper so as not to influence the choices of other participants. This form of assistance expedited the experiment considerably.

Upon completion of the experiment, participants filled out a questionnaire that elicited demographics as well as information relevant to pollution regulation. The participants were then paid, and again were gathered in a conference room. This provided the opportunity for an experiment moderator to thank everyone for their participation, as well as address questions regarding the purpose of the experiment and potential water quality regulations.

Sample sizes and earnings

Forty-eight undergraduate students across four sessions and forty-eight farmers also across four sessions participated in the experiment. Thus, overall there are 96 participants and four replications (groups) for each treatment. In addition to the differences in protocol described above, exchange rates varied across the two subject pools to reflect differences in opportunity costs. On average, students earned $20 for a one-hour session and farmers earned $190 for a two-hour session.

Results

In this section we report on the results from the four experimental treatments described above. We begin with a graphical and econometric analysis of group-level emissions outcomes and then focus on individual-level decisions. The individual-level results highlight some important differences in decisions across firm sizes and subject pools. We conclude the section by presenting the survey results from the farmers and tie these results back to experiment decisions.

Group-level outcomes

Figure 1 presents the average group-level time-series for each treatment. As clearly illustrated, group emissions in the "no policy" scenario (i.e. rounds 1-5) are very close to the theoretical prediction of 120 for students, but farmers are considerably lower than the prediction. When the ambient tax is introduced (round 6), there is a sharp decline in group emissions for all treatments. With the exception of the earlier rounds of the Hetero I - Farmers treatment, group emissions hover around the prediction of 72 in ambient tax rounds. The downward trend in emissions (from under-compliance to compliance) for Hetero I - Farmers can be attributed in part to bankruptcies. In particular, in one session a small firm went bankrupt in round 11, and in a second session a small firm went bankrupt in round 14.[6]

To formally test whether observed outcomes are consistent with theory, as well as test for subject-pool effects, we undertake an econometric analysis of the panel data on group-level emissions. In particular, we estimate an ordinary least squares regression model that specifies group-level emissions as a linear function of a full set of indicator variables (i.e. no intercept is included) that allow mean emissions to vary across the four treatments, and, within each treatment, to vary across "no policy" rounds 1-5 and "ambient tax" rounds 6-17.[7, 8] Formally, the estimating equation is:

(7) $$x_{it} = \left(\sum_{m=1}^{4}\sum_{k=1}^{K}\beta_{mk} \cdot D_{it}^{m} I_{it}^{k}\right) + \varepsilon_{it},$$

where x_{it} is the sum of the emissions choices for group i in decision round t; the D_{it}^{m} are indicator variables that equal 1 for treatment m; the I_{it}^{k} are indicator variables that equal 1 for observations corresponding with period grouping k; the β_{mk} are estimable parameters; and ε_{it} is a mean-zero error term.[9] To account for unspecified heteroskedasticity and within-group serial correlation, cluster-robust standard errors are computed.

The estimated model is presented in table 2. Given that the model covariates are treatment/round-grouping dummy variables, the regression coefficients are simply interpreted as mean group emissions levels. For ease of presentation, the estimated coefficients are organized according to treatment and period grouping. The model results, as they relate to theoretical predictions, correspond with what was gleaned from figure 1: (i) for students, group emissions are not statistically different from theory in both the no policy and ambient tax scenarios; (ii) for farmers, group emissions are statistically different, and lower than, the theoretical prediction for no policy rounds; and (iii) with the exception of pre-bankruptcy rounds in Hetero I, group emissions in ambient tax rounds are not statistically different than the standard in the farmer treatments.

To explore subject pool effects, we test whether the mean outcomes over the two types of heterogeneity are equal across farmers and students. Not surprisingly, given the much lower group emissions among farmers in the no policy scenario, we do find that mean group emissions are statistically different in the no policy rounds across pools

($F_{(1,15)} = 37.60$; $p < 0.01$). The lower emissions observed amongst farmers in the no policy rounds may be related to the higher levels of cooperative behavior of farmers in a number of studies in developing economies (see Cardenas and Carpenter 2008), although it should be noted that emissions reductions in these rounds do not provide direct benefits to other participants. It may however, be the case that farmers are more likely to have pro-social motivations that lead them to be more concerned with the environmental implications of their decisions in the pre-policy setting. We also observe a marginally significant difference in mean emissions for ambient tax rounds based on the pre-bankruptcy outcome ($F_{(1,15)} = 3.42$; $p = 0.08$), but not the post-bankruptcy outcome ($F_{(1,15)} = 0.00$; $p = 0.95$).

Individual-level outcomes

The fact that both bankruptcies occurred among small firms managed by farmers points to the more general outcome that there are important differences in emissions decisions by firm size across subject pools. To evaluate firm-level differences in emissions, we estimate a regression that is similar to the one reported above, but uses individual emissions as the dependent variable and includes additional interaction variables so that mean emissions are further allowed to differ by firm size. Similar to the group emissions model, we control for heteroskedasticity and serial correlation through the estimation of cluster-robust standard errors, although the clustering in this case is at the participant level.

The results of the firm-level emissions model presented in table 3 illustrate two noteworthy outcomes. First, in the pre-policy rounds emissions outcomes are not significantly different from expectations across all firm sizes among students, but are significantly different and lower than expectations for all sizes in the treatments with farmer participants. This indicates that the lower emissions levels observed at the group level are a product of the systematic over-abatement by firms of all sizes in the farmer subject pool. We explore this issue in more detail below, when we evaluate the relationship between farmer characteristics and their experiment decisions.

The second primary outcome of interest from the individual-level analysis is that there are systematic differences in the decisions of small and large firms across subject pools. In particular, students assigned to small firms tend to emit significantly more than farmers assigned to small firms, while students assigned to large firms students choose significantly lower emissions levels than their farmer counterparts. Formally, based on the pre-bankruptcy rounds in Hetero I and all ambient tax rounds for Hetero II, the null hypothesis that the mean emissions for a particular firm size are equal across subject pools can be rejected for both small ($F_{(1,95)} = 6.30$; $p = 0.01$) and large firms ($F_{(1,95)} = 4.08$; $p = 0.04$) . The null hypothesis of equality cannot be rejected for medium-sized firms ($F_{(1,95)} = 0.49$; $p = 0.49$). The former set of differences goes away in the Hetero I, post-bankruptcy rounds which is attributable to the fact that the small firms that went bankrupt had relatively low emissions levels and are no longer in the sample and the large firms reduced emissions slightly in the post-bankruptcy rounds (although the difference is not statistically significant).

Lower relative emissions among small firms are potentially problematic in a policy setting since the ambient tax, by its construction, imposes a higher relative financial burden on small firms when violations occur.[10] The potentially negative outcomes associated with lower emissions levels are evidenced in part by the two bankruptcies that occurred among small firms in the farmer subject pool. Note that although average emissions levels among small firms in the treatment where the bankruptcies occurred are still above the Nash equilibrium levels, the two participants that went bankrupt chose levels of emissions that were considerably below optimal levels. This, in tandem with the under-abatement by large firms – and corresponding high ambient tax payments –, lead to the bankruptcies.

In previous ambient tax experiments with heterogeneity and student participants (Spraggon 2004; Suter, Vossler and Poe 2009), small firms under-abated and large firms over-abated relative to expectations, as was observed in the student sessions of the experiment we report here. The reversal of this outcome amongst farmer participants is surprising and deserves additional attention. Recall that in the experimental sessions, farmers are assigned to a particular experimental firm size based on the size of their own operation. The potential exists therefore that the characteristics of the operators of smaller farms lead them to choose greater abatement levels than their large farm counterparts.[11] In the analysis that follows, we utilize the results of the post-experiment survey to identify whether experiment decisions correlate with farmer characteristics.

Survey results and additional analysis

Table 4 provides some descriptive statistics for the farmer participants, broken down by (experiment) firm size, collected from the post-experiment survey. As illustrated in the table, farmers assigned to the small firm size have an average of 52 milking cows, compared to 183 for medium firms, and 615 for large firms. The mean participant age among farmers is fairly consistent at 46-47 across firm sizes. Among the most notable differences is the amount of education received by the farmers: 88% of farmers operating large firms had bachelor's degrees, compared to only 44% of medium and 38% of small farmers.

There are other results that pertain to nonpoint source pollution regulation that are noteworthy. First, three-quarters of the small and medium farmers and 56% of the large farmers felt that experimental studies such as this one should be used to help inform government agencies. This indicates that not only do farmers see policy test-bedding as important, but that experiments may be helpful in generating the requisite buy-in from the agricultural community prior to regulatory policies being implemented. Additionally, the survey results reinforce the claim that most farmers are already influenced by agro-environmental policy in the form of conservation payments from state and federal authorities. Nearly three-quarters of the farmers in the sample receive conservation payments, with the highest proportion being among larger farms.

One concern with ambient-based policies is that in order to generate effective incentives, agricultural decision makers need to have an understanding of the relationship between their abatement activities and the reduction in pollution loads entering the watershed. They must also be able to estimate the costs of their abatement activities. The survey results are not particularly optimistic on these questions. When asked if they could estimate the load reduction associated with a change in practice, only slightly more than 50% responded positively. Further, slightly more than 60% agreed that they could estimate the cost of the actions that they would take in response to new agricultural regulations. Interestingly, it appears that in each case the medium sized farms have greater confidence in their ability to estimate the outcomes of their activities than operators of small or large farms. One explanation for this result could be that smaller farms do not have the technology in place to appropriately monitor changes and larger farms have too many variables to deal with to effectively predict changes in emissions and costs. While the overall results are not completely unexpected, they provide policy makers with a challenge in the efficient design of policy when a significant percentage of decision makers claim they cannot effectively determine the outcome or cost of their abatement activities.

Given that the incentives generated by ambient-based regulations are dependent on the actions of every firm in the watershed, the interactions that farmers have with other agricultural agents is important. As mentioned above, nearly three-quarters of the participants believe that they could determine the abatement activities of their neighbors. In addition, more than eight out of ten believe that if an ambient-based regulation were to be put into place, they would engage in communication with other local farmers to determine how best to address the new policy. Although not evaluated specifically in this experiment, communication has been shown in past research to significantly influence observed decision making (Suter *et al.* 2008, 2010; Vossler *et al.* 2006). Policy makers should therefore take into account the potential changes in incentives when firms engage in explicit collusion. It should be noted, however, that although the overwhelming majority of participants stated that they would engage in communication with their neighbors, fewer than half said that they would explicitly attempt to monitor the decisions made by their neighbors.

To formally analyze the relationship between farmer characteristics and decision making in the experiment, we utilize an OLS regression where the dependent variable is the difference between a participant's average emissions decision across rounds and the Nash equilibrium prediction. The set of independent variables include the variables described in table 4, with the exception of (given clear endogeneity concerns) herd size, self-assessment of experiment comprehension, and computer use.

Table 5 reports the results of separate models estimated for the Part A (no policy) and Part B (ambient tax) rounds. In the no policy model, farmers that are confident in their ability to estimate their own abatement costs, believe they can determine their neighbor's pollution and believe that their farm reduces water quality are associated with lower emissions levels. The intuition for this outcome appears to be that when no policy is in place, farmers that better understand the relationship between their actions and the actions of their neighbors on water quality are more likely to voluntarily choose lower levels of emissions despite the fact that this reduces their earnings in the experiment.

For ambient tax periods, two variables are found to significantly influence decision making. Farmers that believe they can determine their neighbors' contributions to pollution levels and believe that agriculture is the most significant polluter in their watershed tend to choose lower emissions relative to expectations. This again suggests that farmers that are conscious of agricultural pollution are more likely to take measures to reduce their own pollution loads. It may also be indicative of the fact that if a farmer believes he can monitor his neighbor's abatement then he may himself feel social pressure to reduce emissions. There appears to be a systematic relationship between a farmer's beliefs about whether agriculture is the leading source of water pollution and the size of the farm operation. Half of small farmers believe that agriculture is the leading pollution source, whereas fewer than 17% of large farmers believe agriculture to have the greatest impact on water pollution levels in their watershed. This systematic difference in beliefs contributes to the observed outcome that small firms choose emissions decisions that are relatively lower than expectations compared to the large firms.

Neither model establishes a statistical relationship between decisions and underlying beliefs regarding the usefulness of an experiment such as ours for informing policy. This provides some suggestive evidence that concerns over broader policy implications may not have heavily influenced decisions in the experiment.

Discussion

Informed by the large set of existing laboratory experiments on the performance of ambient pollution tax mechanisms, this study implements what is arguably the best performing of these mechanisms in the absence of communication on a pool of participants composed of a key policy target group. Specifically, we explore the performance of a linear ambient tax, with a tax threshold slightly below the pollution standard, amongst both undergraduate students and dairy farmers. In an effort to provide a relevant field context, the experimental design utilizes parameters based on data from the Upstate NY dairy industry, parameter heterogeneity (reflective of dairy herd size), and stochastic ambient pollution. Further, farmer participants are assigned to firm sizes based on the size of their actual dairy herd. Overall, the group-level results are generally consistent with theoretical predictions, as group emissions tend to approximate the pollution standard. This result is evidence that in aggregate, ambient-based policies can motivate groups of polluters to achieve pollution targets and that these results are robust to different forms of heterogeneity and subject pools.

There are important differences in individual-level decisions across subject pools. Among students, small firms have a tendency to under-abate relative to the Nash equilibrium, while large firms tend to over-abate. This corroborates the findings of Suter, Vossler, and Poe (2009) and Spraggon (2004), who uncover similar firm size related tendencies under different types of heterogeneity. Relative to the students, farmers operating small firms tend to choose significantly lower emissions decisions, while operators of large farms choose emissions decisions that are significantly higher. The higher levels of abatement imply greater costs to the small firms that are already subject to the greatest relative ambient tax liability given their small size. This is borne out in our experimental results as small firms in two separate experimental groups experienced bankruptcy during the course of the session. This result suggests some caution in implementing the ambient tax in the field. One possible solution is to combine the ambient tax with a system of lump-sum subsidies that favors small firms. Alternatively, the regulator can consider notably different ambient policies, such as the system of subsidies and fines proposed by Xepapadeas (1991), which has the theoretical advantage of a lower (average) financial burden when a violation occurs and could be designed in a way that is more favorable to small firms.

Based on the results of a post-experiment survey, farmers that generally understand the relationship between agriculture and water quality tended to choose lower emissions in the experiment. Operators of large dairy farms are the least likely to view agriculture as the most significant polluter (12.5% compared to nearly 50% amongst small and medium farms) and this result helps to explain a portion of the differences in individual-level experiment decisions.

One interesting finding is that farmers significantly over-abated (by 20%) relative to the profit-maximizing outcome in our "no policy" baseline, a result which is not paralleled with student participants in our experiment or in previous experiments that included a similar baseline setting (see, for example, Vossler *et al*., 2006). This may be tied to self-image, or broader environmental concerns that were "brought into the lab" from the farmers' relevant field experience. This is in line with Henrich *et al*.'s (2001) notion that behavior in experiments can be closely tied to the structure of everyday life. Although abatement in the baseline rounds of our experiment does not directly benefit other participants, it may be indicative of the heightened sensitivity that farmers have for the environmental implications of their actions.

As an additional indication of differences in the preferences of farmers and students, we note that for half of the farmer groups we had time to run a short, additional experiment that mirrors the voluntary-threat mechanism proposed by Segerson and Wu (2006). Using a similar design, Suter *et al.* (2010) find that (in the absence of communication) the majority of student groups consistently violated the voluntary standard. In contrast we found that farmers in fact met the standard voluntarily in all groups. As the voluntary-threat mechanism may have political appeal relative to a pure ambient tax, this represents another candidate mechanism that may be worthwhile to test in a more systematic fashion with farmers.

As lower farmer emissions in the no-policy scenario of the experiment were tied to those who better understand the relationship between farm operations and water quality, this suggests that policy makers should also consider devoting resources to appropriately educate farmers in these matters. Indeed, the results suggest information programs are likely to lead to improvements in water quality without explicit regulation. The similar tie between beliefs and decisions in the ambient tax scenario further suggests that information programs may enhance the potential cost-effectiveness of an ambient-based tax policy.

By utilizing instructions that frame the decision task as an emission choice with ties to water quality, and by assigning farmer participants to an experimental firm size based on their actual herd size, the experiment was designed to enhance external validity. With an eye instead on internal validity, one could augment the design to include treatments that use generic framing for both student and farmer participants and to include treatments where farmers are exogenously assigned to a firm size. We do not expect framing to matter for students, however, given that meaningful differences are not apparent when comparing the experimental results of parallel ambient-based mechanisms conducted with students in Spraggon (2002) with generic framing and Vossler *et al*. (2006) with a water quality context. We leave as questions for future research the extent to which framing and endogenous selection into player type influences the decisions of farmers.

The use of context, the endogenous selection of experimental firm size, and acknowledgement of EPA funding leaves open the possibility that farmers' decisions may have been at least partially driven by beliefs regarding the potential link between the experiment and future regulations. At the extremes, farmers may have wished to excessively "pollute" in the experiment in attempt to convince the regulator that the ambient tax would not work, or excessively abate in the experiment in attempt to convince the regulator that a much lower ambient tax rate would achieve compliance. Alternatively, farmers may have wanted to be seen by the regulator as cooperative and thus endeavored to meet the ambient standard. However, such behavior would presumably have the undesirable effect (for the farmer) of making an ambient tax policy more likely. Although the direct financial incentives (i.e. experiment payoffs) were large, and we have some limited statistical evidence that suggests otherwise, potential influence over future policy may nevertheless have been a behavioral driver. In policy test-bed experiments with targeted participant groups, the influence of beliefs over broader policy considerations remains an important area for future research.

References

Alpízar, F., T. Requate, and A. Schram. 2004. "Collective versus random fining: an experimental study on controlling ambient pollution." *Environmental and Resource Economics* 29:231-252.

Ball, S.B., and P. Cech. 1996. "Subject pool choice and treatment effects in economic laboratory research." In R.M. Isaac, ed. *Research in Experimental Economics*, Volume 6. Greenwich, CT: JAI Press, pp. 239-92.

Cardenas, J.C., and J. Carpenter. 2008. "Behavioral development economics: Lessons from field labs in the developing world." *The Journal of Development Studies* 44:311-338.

Cochard, F., M. Willinger, and A. Xepapadeas, 2005. "Efficiency of nonpoint source pollution instruments: an experimental study." *Environmental and Resource Economics* 29:231-252.

Croson, R. 2005. "The method of experimental economics." *International Negotiation* 10:131-148.

Cummings, R., C. Holt, and S. Laury. 2004. "Using laboratory experiments for policymaking: An example from the Georgia Irrigation Reduction Auction." *Journal of Policy Analysis and Management* 23:341-363.

Duquette, E., N. Higgins, and J. Horowitz. 2012. "Farmer discount rates: Experimental evidence." *American Journal of Agricultural Economics* 94:451-456.

Grossman, G., and D. Baldassarri. 2012. "The impact of elections on cooperation: evidence from a lab-in-the-field experiment in Uganda." *American Journal of Political Science* 56:964-985.

Harrison, G.W., and J.A. List. 2004. "Field experiments." *Journal of Economic Literature* 42:1009-1055.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath. 2001. "In search of homo economicus: behavioral experiments in 15 small-scale societies." *American Economic Review* 91:73-78.

Herberich, D.H., and J.A. List. 2012. "Digging into background risk: experiments with farmers and students." *American Journal of Agricultural Economics* 94:457-463.

Herberich, D.H., S.D. Levitt, and J.A. List. 2009. "Can field experiments return agricultural economics to the glory days?" *American Journal of Agricultural Economics* 91:1259-1265.

Hong, J., and C.R. Plott. 1982. "Rate filing policies for inland water transportation: an experimental approach." *Bell Journal of Economics* 13:1-19.

Krause, K., J.M. Chermak, and D.S. Brookshire. 2003. "The demand for water: consumer response to scarcity." *Journal of Regulatory Economics* 23:167-191.

Levitt, S.D., and J.A. List. 2007. "What do laboratory experiments measuring social preferences reveal about the real world?" *Journal of Economic Perspectives* 21:153-174.

Loewenstein, G. 1999. "Experimental economics from the vantage-point of behavioural economics." *The Economic Journal* 109:F25-F34.

Normann, H.T., and R. Ricciuti. 2009. “Laboratory experiments for economic policy making.” *Journal of Economic Surveys* 23:407-432.

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.

Plott, C.R., and D.P. Porter. 1996. “Market architectures and institutional testbedding: An experiment with space station pricing policies.” *Journal of Economic Behavior and Organization* 31:237-272.

Segerson, K. 1988. “Uncertainty and incentives for non-point source pollution.” *Journal of Environmental Economics and Management* 15:87-98.

Segerson, K., and J. Wu. 2006. “Voluntary approaches to nonpoint pollution control: Inducing first-best outcomes through the use of threats.” *Journal of Environmental Economics and Management* 51:165-184.

Shogren, J. 2006. “Valuation in the lab.” *Environmental and Resource Economics* 34: 163- 172.

Smith, V. 1982. “Microeconomic systems as an experimental science.” *The American Economic Review* 72:923-955.

Spraggon, J.M. 2002. “Exogenous targeting instructions as a solution to group moral hazards.” *Journal of Public Economics* 84: 427-456.

Spraggon, J.M. 2004. “Testing ambient pollution instruments with heterogeneous agents.” *Journal of Environmental Economics and Management* 48:837-856.

Suter, J.F., C.A. Vossler, G.L. Poe, and K. Segerson. 2008. "Experiments on damage-based ambient taxes for nonpoint source polluters." *American Journal of Agricultural Economics* 90:86-102.

Suter, J.F., C.A. Vossler, and G.L. Poe. 2009. "Ambient-based pollution mechanisms: a comparison of homogeneous and heterogeneous groups of emitters." *Ecological Economics* 48:1146-1158.

Suter, J.F., K. Segerson, C.A. Vossler, and G.L. Poe. 2010. "Voluntary-threat approaches to reduce ambient water pollution." *American Journal of Agricultural Economics* 92:1195-1213.

Taylor, M.A., B. Sohngen, A. Randall, and H. Pushkarskaya. 2004. "Group contracts for voluntary nonpoint source pollution reductions: Evidence from experimental auctions." *American Journal of Agricultural Economics* 86:1196-1202.

Vossler, C.A., and M. McKee. 2012. "Efficient tax reporting: The effects of taxpayer information services." Working paper, Dept. of Econ., University of Tennessee.

Vossler, C.A., T.D. Mount, R.J. Thomas, and R.D. Zimmerman. 2009. "An experimental investigation of soft price caps in uniform price auction markets for wholesale electricity." *Journal of Regulatory Economics* 36:44-59.

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

Xepapadeas, A. 1991. “Environmental policy under imperfect information: incentives and moral hazard.” *Journal of Environmental Economics and Management* 20:113-126.

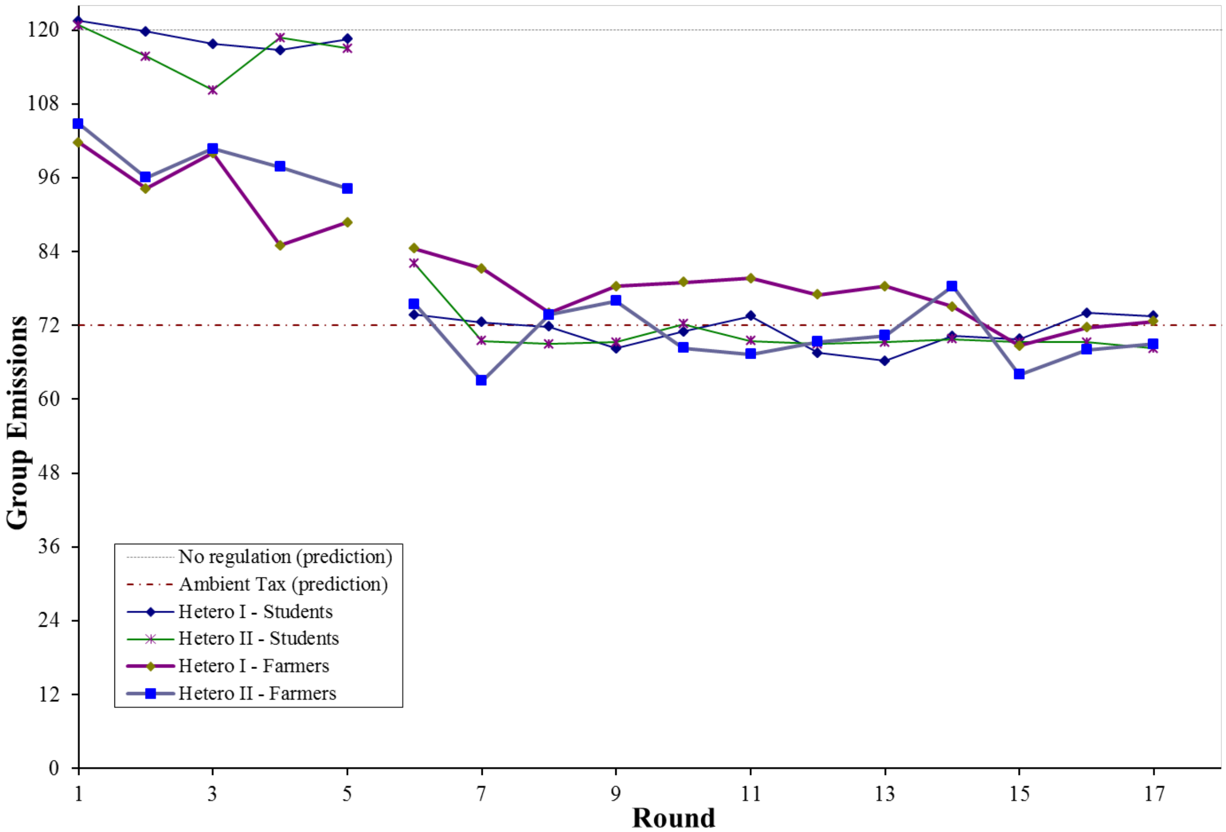

Figure 1: Group emissions by round

Table 1. Model Parameters and Theoretical Predictions

	Small	**Medium**	**Large**
I	30,000	75,000	120,000
ϕ (Hetero I)	227.94	470.77	810.93
ϕ (Hetero II)	414.08	470.77	648.02
γ	0.125	0.100	0.075
α	80	200	400
Theoretical Predictions (Emissions)			
No regulation	10	20	30
Ambient tax (Hetero I)	4	12	20
Ambient tax (Hetero II)	8	12	16

Table 2. Group Emissions Model

Treatment	Rounds 1-5: no policy	Rounds 6-17: ambient tax	Rounds 6-14: ambient tax	Rounds 15-17: ambient tax
Hetero I - Students	118.85 (0.95)	71.00 (3.12)		
Hetero II - Students	116.50 (2.25)	70.52 (0.92)		
Hetero I - Farmers	93.95** (2.96)		78.70** (1.56)	71.00 (1.81)
Hetero II - Farmers	98.70** (5.81)	70.28 (1.81)		
$N = 254$, $R^2 = 0.99$				

*Note: Theoretical predictions are 120 and 72 for rounds without and with the ambient tax policy, respectively. * and ** indicate parameter estimate is different from the theoretical prediction at the 10% and 5% significance levels, respectively. Cluster-robust standard errors are in parentheses.*

Table 3. Firm Emissions Model

		Theoretical Predictions		Students		Farmers			
		Rounds 1-5	**Rounds 6-17**	**Rounds 1-5**	**Rounds 6-17**	**Rounds 1-5**	**Rounds 6-17**	**Rounds 6-14**	**Rounds 15-17**
	Small	10	4	9.95	6.84**	6.23**		4.91	6.50**
				(0.05)	(0.94)	(1.11)		(0.90)	(0.81)
Hetero I	Med.	20	12	19.78	9.61**	15.43**		11.75	10.44
				(0.59)	(1.20)	(1.07)		(1.17)	(1.60)
	Large	30	20	29.70	19.04	25.33**		22.93*	20.72
				(0.28)	(1.44)	(1.39)		(1.70)	(2.23)
	Small	10	8	9.68	7.82	8.15**	6.12**		
				(0.21)	(0.51)	(0.36)	(0.39)		
Hetero II	Med.	20	12	19.73	13.10	15.30**	12.77		
				(0.65)	(1.01)	(1.70)	(1.68)		
	Large	30	16	28.85	14.33	25.90**	16.26		
				(0.96)	(1.26)	(1.99)	(1.31)		
$N = 1515$, $R^2 = 0.94$									

*Note: * and ** indicate parameter estimate is different from the theoretical prediction at the 10% and 5% significance levels, respectively. Cluster-robust standard errors are in parentheses. Rounds 1-5 correspond with a "no policy" scenario whereas rounds 6-17 correspond with the ambient tax.*

Table 4. Farmer Characteristics and Beliefs (N = 48)

	Small	Medium	Large
Number of dairy cows	52.2 (29.7)	183.1 (188.5)	615.2 (446.1)
Understood experiment instructions (1-5 scale)	4.0 (0.8)	4.2 (0.7)	3.9 (1.3)
Studies like this are useful for policy (%)	75.0 (44.7)	75.0 (44.7)	56.3 (51.2)
Uses computer in operation (%)	56.3 (51.2)	75.0 (44.7)	87.5 (34.2)
Age (years)	47.1 (11.1)	47.3 (14.0)	46.1 (13.3)
College educated (%)	37.5 (50.0)	43.8 (51.2)	87.5 (34.2)
Receives payments for conservation (%)	62.5 (50.0)	75.0 (44.7)	81.3 (40.3)
Can estimate abatement costs (%)	56.3 (51.2)	87.5 (34.2)	56.3 (51.2)
Can estimate load reductions (%)	43.8 (51.2)	62.5 (50.0)	37.5 (50.0)
Would communicate with neighbors (%)	87.5 (34.2)	87.5 (34.2)	81.3 (40.3)
Would monitor the actions of neighbors (%)	56.3 (51.2)	50.0 (51.6)	31.3 (47.9)
Can determine neighbors' pollution (%)	62.5 (50.0)	68.8 (47.9)	75.0 (44.7)
Farm reduces water quality (%)	43.8 (51.2)	31.3 (47.9)	50.0 (51.6)
Agriculture most significant polluter (%)	43.8 (51.2)	50.0 (51.6)	12.5 (34.2)

Note: Numbers in parentheses are standard deviations.

Table 5. Farmer Decision Making Models

Dependent variable:	Deviation of emissions choice from theory: no policy	Deviation of emissions choice from theory: ambient tax
Age	−0.07	−0.02
	(0.04)	(0.05)
College educated	0.26	1.08
	(0.97)	(1.22)
Receives payments for conservation	0.68	1.25
	(1.57)	(1.76)
Can estimate abatement costs	−2.90**	0.92
	(1.22)	(1.36)
Can estimate load reductions	−0.20	−0.57
	(1.03)	(1.13)
Would communicate with neighbors	−0.89	0.35
	(1.13)	(1.23)
Would monitor the actions of neighbors	−1.68*	−0.59
	(1.05)	(1.05)
Can determine neighbors' pollution	−2.94**	−2.54*
	(1.14)	(1.43)
Farm reduces water quality	−2.52**	−0.57
	(0.94)	(1.14)
Agriculture most significant polluter	−0.49	−1.99**
	(0.93)	(0.95)
Studies like this are useful for policy	1.73	0.01
	(1.32)	(1.29)
Intercept	4.24	2.01
	(3.04)	(3.51)
N	48	48
R^2	0.37	0.23

*Note: * and ** indicates parameter estimate is different from zero at the 10% and 5% significance levels, respectively. Robust standard errors are in parentheses.*

[1] Following the nomenclature of Harrison and List (2004), our experiment with farmers is best described as a "framed field experiment" given that it utilizes a nonstandard subject pool and introduces field context. Our experiment with students is a "laboratory experiment".

[2] Hong and Plott (1982) used a subject pool composed of "engineers, secretaries, housewives and university faculty members" as well as students in experiments related to the determination of rates for the transportation of bulk commodities in inland waterways. However, no comparisons were made across pools.

[3] Example instructions are included in a supplementary online appendix.

[4] To be eligible, students had to have completed at least one course in economics and could not have participated in a similar experiment in the past.

[5] We include data from this session in the analysis that follows. The main conclusions we draw are nevertheless robust to inclusion/exclusion of the data from this session.

[6] When a participant in the experiment went bankrupt, they received a small fixed payment in each subsequent round, but did not contribute to ambient pollution levels. Bankruptcies were not announced to other group members.

[7] As the second bankruptcy occurred in period 14, we define the pre-bankruptcy rounds as 6-14. The same qualitative results are obtained if we place this structural break at round 11 (when the first bankruptcy occurred), 12 or 13.

[8] We estimated analogous models to those reported in table 2 and table 3 using the random effects estimator. The estimates of mean emissions and standard errors, and conclusions drawn from hypothesis tests, are very similar. One exception is that estimated mean emissions for

"Small" firms in rounds 15-17 of the Hetero I – farmers treatment is no longer statistically different from the theoretical prediction.

[9] Given there is no constant term in the group and individual-level emissions models, the R^2 should be interpreted with caution. This diagnostic is reported only for replication purposes.

[10] To induce efficient pollution reductions, ambient-based regulatory policies must be designed such that the expected marginal cost to a firm of emitting more than the optimal quantity is at least as great as the marginal benefit. For the ambient tax mechanism we explore this necessarily implies a higher relative tax burden on smaller firms. To account for differences in the relative costs of the policy, the ambient tax could be accompanied by a lump-sum subsidy that varies by firm size. Alternatively, as proposed by Xepapadeas (1991), the regulator can use a system of subsidies and fines under which only one randomly chosen firm is fined when a violation occurs. This mechanism has the theoretical advantage of lower (average) sanctions when violations occur. Further, the regulator could adjust the mechanism to account for firm size by specifying a relatively lower probability that a small firm is selected to be fined should a violation occur.

[11] An alternative explanation for the differences in relative emissions decisions is that the behavior of farmers was motivated at least in part by attempts to influence policy. For example it is possible that the higher emissions of the farmers operating large firms may have resulted from attempts to signal to policy makers that an ambient-based policy would not effectively reduce emissions. Conversely, lower relative emissions for the small farmers could have been due in part to farmers over-abating in an effort to signal to policy makers that a low tax rate is sufficient to induce pollution reductions.