ESSAYS ON EXPERIMENTS AND AGENT-BASED MODELING

by

Shang Wu

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Summer 2017

ESSAYS ON EXPERIMENTS AND AGENT-BASED MODELING

by

Shang Wu

Approved: ____

James L. Butkiewicz, Ph.D.
Chair of the Department of Economics

Approved: ____

Bruce W. Weber, Ph.D.
Dean of the Alfred Lerner College of Business and Economics

Approved: ____

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Kent D. Messer, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

William R. Latham III, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

George R. Parsons, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Leah H. Palm-Forster, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Asim Zia, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

I would like to acknowledge funding support from NSF EPSCoR Rii3 Track 1 and NEWRnet for the first two chapters of my dissertation and Northeast Extension Risk Management Education Center for the third chapter of the dissertation. I am also grateful for the generous support from USDA Hatch, Department of Applied Economics & Statistics (APEC), Department of Economics, and DENIN.

I wish to thank my advisor, Kent Messer for continued encouragement and support throughout the years. I would also like to express my gratitude to my committee members, and numerous other collaborators, co-authors, colleagues, and friends both in the APEC and Economics Departments, and elsewhere.

Finally, I can never thank my family enough for their moral support and love through all these years, none of this is possible without them. May my love be with you, “fino alla fine”.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Heterogeneous agents and information nudges in non-point source water pollution management

Abstract: Non-point source (NPS) water pollution from agricultural runoff is a leading cause of impairment for many water bodies in the United States; however, sources of NPS pollution are difficult to identify because of hidden actions and asymmetric information. Theoretical and experimental research has shown that ambient pollution policies can induce groups to reduce pollution to socially efficient levels, but many of these studies have imposed restrictive assumptions about farmer homogeneity and management choices. In reality, agricultural firms differ in both size and location, and farmers make numerous management decisions that can affect runoff and nutrient loss, including decisions about production intensity and pollution abatement technologies. Researchers have shown that introducing either size or location heterogeneity affects the efficiency of ambient pollution policies, but no research has analyzed policy performance while considering several sources of heterogeneity and multiple management decisions. Furthermore, despite multiple examples in using non-pecuniary interventions to promote environmental conservation, little research has

examined how to use information nudges, like social comparisons or information about peer actions, to induce better NPS pollution abatement decisions.

In this study, we designed an economic experiment to test the effects of multiple layers of heterogeneity, information nudges, and an extended decision space on the performance of the classic ambient tax/subsidy policy. Experiment participants (n=192) were recruited from a large public university in the U.S. In the experiment, each individual was assigned a firm and asked to make individual decisions that affected the profitability of his/her firm and ambient water pollution of their group. In each round of the experiment, participants selected their production intensity and chose one of two production technologies—a conventional technology or a more expensive technology that generated less pollution.

Eight within-subject treatments were tested, including two policy variations (no policy and a tax/subsidy policy) and four size/location variations (homogeneous, location heterogeneity, size heterogeneity, and both location and size heterogeneity). Three between-subject information treatments were also tested, including a no information control. In information treatment 1, we tested how individual decisions were affected by information nudges about decisions that similar individuals had made in past sessions. In information treatment 2, participants were provided with information about the average production and technology adoption rate in their group during the last round. A unique dominant strategy Nash Equilibrium was calculated for both the adoption decision and production decision based on location and size.

Our results demonstrate that, without information nudges, more firm heterogeneity reduces the effectiveness of ambient tax/subsidy policies and target pollution levels are achieved less frequently. However, the tax/subsidy policy was effective under different heterogeneity scenarios when information is provided about peer and group decisions in past rounds. Furthermore, information treatment 1 and information treatment 2 generate higher policy efficiency than no information treatment. Lastly, participants are able to find and retain their dominant strategy better in the information 1 treatment, suggesting that providing individually targeted information is more effective than providing information about aggregate group-level decisions. Our findings suggest that traditional ambient pollution policies may be less effective when agents are heterogeneous and make multiple decisions that affect pollution, but information nudges can improve policy performance.

Simulating Heterogeneous Farmer Behaviors under Different Policy Schemes: Integrating Economic Experiments and Agent-Based Modeling

Abstract: In this paper, we develop an agent-based model that scales up results from economic experiments on technology diffusion and abatement of non-point source

water pollution under the conditions of an actual watershed. The results from the economic experiments provide the foundation for assumptions used in the agent-based model. Data from geographic information systems and the US Census of Agriculture initialize and parameterize the model. This integrated model enables the exploration of the effects of several policy interventions on technology diffusion and agricultural production and, hence, on agricultural non-point source pollution. Simulation results demonstrate that information 'nudges' based on social comparisons increase ambient based policy performance as well as efficiency, especially individual-level tailored information on what others like them have done in past similar situations.

Auctions versus Posted Price in Experiments: Comparisons of Mean and Marginal Effect

Abstract: Economic experiments have been widely used to elicit individuals' evaluation for various commodities and non-market goods. Common elicitation methods include auctions and posted price mechanisms. Experimental auctions are theoretically incentive compatible so are assumed to give an unbiased estimate of individuals' evaluation including willingness to pay (WTP). However, the vast majority of purchasing decisions are not made in auctions but in market settings, such as grocery stores, where consumers make yes/no decisions in response to a set price. In this research, we carefully design an experiment to compare homegrown-value WTP estimates between an auction and a posted price elicitation format. This design enables us to make both within- and between-subjects comparisons of the mean WTP and marginal effect estimates. Results from 115 adult consumers indicate that WTP estimates obtained from an auction are approximately 32% - 39% smaller than WTP estimates obtained from a posted price mechanism. In addition, we compare the statistical significance and conclude that auctions require a smaller sample size than posted price mechanisms in order to detect the same preference change. Nevertheless, the signs of marginal effects for different product characteristics are consistent in both mechanisms.

Chapter 1

HETEROGENEOUS AGENTS AND INFORMATION NUDGES IN NON-POINT SOURCE WATER POLLUTION MANAGEMENT

1.1 Introduction

Regulation of non-point source (NPS) water pollution is a difficult task since it involves hidden actions and asymmetric information from individual polluters, making it impossible or prohibitively costly to track and set up individual-based policies (Xepapadeas, 2011; Miao et al., 2016). Segerson (1988) showed that policy instruments could be designed to overcome these problems and reduce pollution to near an exogenously determined ambient pollution level. However, ambient-based policies have not been carried out in reality on a large scale due to obstacles such as political feasibility and fairness concerns (Cason and Gangadharan, 2013; Xepapadeas, 2011). Therefore, researchers primarily use experimental or theoretical methods to investigate how ambient pollution policies can be used to improve water quality (Xepapadeas, 1992; Spraggon, 2002; Poe et al., 2004).

In these ambient-based policy schemes, the regulator usually compares the pollution reading to a target level of pollution, and imposes monetary policy instruments (tax and/or subsidy) to everyone in the watershed. Researchers showed that ambient-based policies can induce groups to reduce pollution to socially efficient levels, but many of these studies are based on restrictive assumptions on farmer homogeneity and their management decisions. In reality, agricultural firms may differ in both production capacity and location relative to the sensor, which may result in different pollution behavior. Studies have shown that introducing either size or location heterogeneity affects the efficiency of ambient pollution policies, but no research has analyzed policy performance while considering multiple sources of heterogeneity. Besides, in recent years, with water pollution becoming increasingly severe in many watersheds, pollution abatement technologies to reduce individual nutrition runoffs have become more developed and available to households. For example, a technology (e.g., conservation buffers) could remove up to 50% or more of nutrients and pesticides in runoff (Conservation Technology Information Center, Purdue University, 2016). Unlike individual pollution levels, the adoption of a certain abatement technology is usually more visible to neighbors in the same community and shows a producer's commitment to conservation. Regulators may also be able to gather information on the status of adopting certain abatement technologies. However, such technology decisions have seldom been explicitly considered in the past (Palm-Forster, Suter and Messer, 2017). Along with size and location heterogeneity, management decisions including production intensity and pollution abatement technologies may affect runoff and nutrient loss.

Furthermore, in recent years, both the public and private sectors realize the benefits of using behavioral economic principles to influence people's behavior. Behavioral based policies are especially attractive to policy makers because they are more cost-effective compared to pecuniary policies. It has been shown in various domains that using behavioral insights, especially information nudges, can improve private as well as social welfare. But in NPS pollution management, most studies focus on various monetary policies, not much attention has been paid on using information nudges, such as social comparisons or peer actions, to affect people's pollution behavior. We explore how information nudges could be used to induce better behavior in a NPS pollution context.

In this study, we design an economic experiment to test the effects of multiple layers of heterogeneity, information nudges, and an extended decision space on the performance of the classic ambient tax/subsidy policy. We find that in general the policy becomes less effective as heterogeneity is introduced, but restores its effectiveness with the aid of information nudges.

1.2 Literature Review

Segerson (1988) showed that the non-point-source pollution problem could be solved by creating policy incentives based on the ambient level of pollution. Because of the collective nature of these ambient schemes, efficiency of the policy may become a concern (Xepapadeas, 2011). However, since ambient policies have not been carried out on a large scale in practice, the lack of empirical data leads to the use of experimental economics as test beds for these policy schemes. A stream of literature has shown both theoretically and experimentally that various types of ambient schemes could lead to effectively attaining the target level of pollution (Xepapadeas, 1992; Spraggon, 2002; Alpízar, Requate, and Schram, 2004; Poe et al. 2004).

Most of the research in this area has focused on homogenous agents partly for simplicity, and partly due to the suggestion that watershed settings that mostly consist of a small number of homogenous farmers would be most conducive to the application of ambient-based policies (Weersink et al., 1998; Suter, Vossler, and Poe, 2009). A few researchers have made efforts to add heterogeneity in different directions. Spraggon (2004, 2013) and Suter, Vossler, and Poe (2009) consider the heterogeneity in the size of the polluters. Spraggon (2004) concluded that ambient policies could be designed to induce target pollution levels for heterogeneous sized farmers at the cost of substantial inefficiency and inequality. Suter, Vossler, and Poe (2009) extended Spraggon (2004) by adding a watershed context and showed size heterogeneity has an impact on group decisions and may generate desirable or undesirable outcomes depending on specific conditions.

Another type of heterogeneity that has drawn more attention recently is spatial heterogeneity of agents. In reality, environmental monitoring is generally done at certain fixed spatial locations. The spatial location of a polluter relative to the monitoring point has significant impact to the tested environmental damage since pollutants will be diluted in the course of travel. A growing body of research has shown that spatial heterogeneity could influence agent decisions especially in common pool resource settings (e.g., Schnier 2009; Suter et al. 2012; Li et al. 2014; Liu et al. 2014). Cason and Gangadharan (2013) included spatial heterogeneity in terms of proximity to the monitoring station to study the effectiveness of informal neighbor punishment versus a formal ambient tax. In an ambient tax/subsidy experiment that included a realistic physical nutrient transport model to calculate the marginal damage of each spatially explicit polluter, Miao et al. (2016) tested the effect of increasing the frequency of water monitoring on firm decisions.

Informal ways to reduce non-point source pollution have also been investigated in laboratory experiments. Cason and Gangadharan (2013) reported that a formal ambient tax is more effective than empowering neighbors to be able to punish each other after observing their group members' emissions and the formal mechanism can be improved by adding peer punishment. Suter et al. (2008) showed communication would lower the emission level to below the social optimal level.

However, past research has not focused much on using information nudges to improve the performance of ambient based policies. Such information nudges usually use narrative messages – especially about how their behavior compare with others – to influence human behavior. These nudges originate from social comparison theory by Festinger (1954), which posits that people evaluate the appropriateness of their behavior by comparing with others. Past research has demonstrated that this principle could be used to promote environmental conservation, such as reducing power consumption (Allcott, 2001), reducing water usage (Ferraro and Price, 2013; Bernedo, Ferraro and Price, 2014), and environmental conservation behavior in hotels (Goldstein et al., 2008). It is reasonable to assume that such information nudges could be utilized in non-point source pollution management to induce better decisions, but few work focused on this topic. Spraggon (2013) varied the information the participants have on the number of other polluters and their payoffs. His study concluded that while information and heterogeneity do not affect aggregate level policy effectiveness, they both reduce policy efficiencies. In Spraggon and Oxoby (2010), they show that providing participants with a description of marginal decision making increases optimal strategy behavior, thus increases policy efficiency. This "recommended play" still focuses mainly on the private decision of the participants themselves. We are interested in exploring how information on others' or the group's behavior would influence participant's own decision making. We examine the effect of two types of information nudges on participants' behavior. Specifically, we explore if information on past group technology adoption rate and group average production would impact participant behavior, and if testimonial information on what others have done in the same situation would serve as a guideline on individual decisions. Corresponding policy schemes could be designed to improve the effectiveness and efficiency of existing policies.

By combining all the previously mentioned pieces together, our study contributes to the literature in several ways. First, we extend participants' decision spaces to include both production and technology decisions. Second, unlike past literature where at most one type of heterogeneity is taken into account, our setting includes size heterogeneity, spatial heterogeneity and also the combination of both types of heterogeneity simultaneously. Third, we examine if and how information nudges could be used to improve policy performance in the NPS context. By including an extended decision space (production and adoption decisions), multiple layers of heterogeneity (size and location), and information nudges (social comparison and peer actions), our experiment evaluates the performance of ambient-based policies under these interactive effects.

1.3 Model

1.3.1 Model Background

Following the classic set up of NPS pollution experiments, participants play the role of farmers that are adjunct to a watershed and make farm management decisions for a given year. The farmers are price-takers of an exogenously determined price for their products. The production generates a byproduct, which we refer to as emission (e.g., excessive fertilizers that run off the farm during rain), and incurs a social cost to the environment (e.g., pollution in downstream watersheds). The farmers have the option to choose to adopt a pollution abatement technology (e.g., buffers that could reduce runoff of excessive nutrients) at a fixed cost ratio relative to the size of the farm. The technology would reduce the pollution that farmer generates at a constant rate. Therefore, the farmers make two decisions, a production decision and an adoption decision. A regulator monitors the density of emission at downstream and has perfect information on the aggregate emission levels. We assume the regulator has no information on individual production/emission levels, but has knowledge on the average production and average adoption rate of people in the group. The regulator may impose an ambient tax or subsidy based on the observed downstream emission level.

1.3.2 Model Setup

We start the discussion with a homogenous case. Suppose there are N farmers along the river, the private income function for a farmer is identical among the participants, and the form is similar to the one used in Spraggon (2002) and subsequent literature (e.g., Suter, Vossler, and Poe 2009; Spraggon 2013; Cason and Gangadharan 2013; Miao et al. 2016):

$$B(x_i) = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2$$

Where γ_i are parameters and x is the decision variable. Individual profit is maximized when $x_i = \gamma_2$ and γ_2 can be regarded as firm's capacity.

By producing x_i the farmer also generates environmental damage. The damage function follows Spraggon (2002) and can be denoted as $D(x_i) = \beta_0 x_i$, therefore the total environmental damage is $TD = \sum_1^N D(x_i) = \sum_1^N \beta_0 x_i$

The social planner's problem is to maximize the social benefit (denoted as SP), where

$$SP = \sum_{i=1}^{N} B(x_i) - \sum_{i=1}^{N} D(x_i)$$

The first order condition indicates that the optimal level of production is at $x_i = \gamma_2 - \frac{\beta_0}{2\gamma_1}$, which is smaller than the private optimal for individual farmer γ_2 since β_0 and γ_1 are both positive parameters.

1.3.2.1 Tax/Subsidy Scheme

Consider the government imposes a tax/subsidy scheme in a manner similar to Segerson (1988) and other subsequent literature where the tax equals to the environmental damage minus the target level of pollution,

$$t(TD) = (TD - \overline{D})$$

where $\overline{D}$ is the environmental damage target that the regulator sets.

Now the individual payoff function under the tax/subsidy scheme becomes:

$$\pi_i = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - \left(\sum_1^N \beta_0 x_i - \overline{D}\right),$$

Solving for optimal x_i we get $x_i = \gamma_2 - \frac{\beta_0}{2\gamma_1}$, note that this is a unique, dominant strategy Nash Equilibrium.

Under the tax/subsidy scheme, the social planner's problem remains unchanged, and the optimal $x_i = \gamma_2 - \frac{\beta_0}{2\gamma_1}$, meaning that the farmers produce at the socially optimal level.

1.3.2.2 Technology

Now consider that we provide a technology that is available for adoption to the farmers at a fixed cost ratio τ relative to firm's capacity γ_2, the technology could reduce environmental damage to a rate of α <1 of the original level.

Specifically, by adopting the technology, the private income function of a farmer is now:

$$B(x_i) = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - \tau\gamma_2$$

and the environmental damage caused by each firm is reduced to $D(x_i) = \beta_0 \alpha x_i$

We find the equilibrium by backward induction. Consider firm i, given the pollution level of others in the group D_{-i}, its profit function from producing x_i and adopting the technology is:

$$\pi^A = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - (D_{-i} + \beta_0 \alpha x_i - \overline{D}) - \tau\gamma_2$$

$$\frac{\partial \pi^A}{\partial x_i} = 2\gamma_1(\gamma_2 - x_i) - \beta_0 \alpha = 0$$

$$x_i^A = \gamma_2 - \frac{\beta_0 \alpha}{2\gamma_1}$$

Plug in the optimal production level to get the maximum profit of adopting the technology:

$$\pi^A = \gamma_0 - \frac{(\beta_0 \alpha)^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_0 \alpha \gamma_2 - \frac{\beta_0^2 \alpha^2}{2\gamma_1}\right) - \tau\gamma_2$$

Consider firm i, given the same pollution level from others D_{-i}, firm's profit function of not adopting the technology and producing at x_i is:

$$\pi^N = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - (D_{-i} + \beta_0 x_i - \overline{D})$$

$$\frac{\partial \pi^N}{\partial x_i} = 2\gamma_1(\gamma_2 - x_i) - \beta_0 = 0$$

$$x_i^N = \gamma_2 - \frac{\beta_0}{2\gamma_1}$$

The maximized profit of not adopting is:

$$\pi^N = \gamma_0 - \frac{{\beta_0}^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_0 \gamma_2 - \frac{\beta_0^2}{2\gamma_1}\right)$$

In order for the farmer to prefer adopting the technology, it requires $\pi^N < \pi^A$.

Solving for these conditions, we get the following restrictions on the parameters for the farm to adopt the technology:

$$\frac{\beta_0^2}{4\gamma_1}(1-\alpha^2) - \beta_0\gamma_2(1-\alpha) + \tau\gamma_2 < 0$$

Under this condition, we solved for a unique, dominant strategy Nash Equilibrium for this homogeneous case. At this equilibrium:

(1) Firms adopt the technology

(2) Firms choose production level

$$x_i = \gamma_2 - \frac{\beta_0\alpha}{2\gamma_1}$$

1.3.3 Heterogeneity

The above sections were the homogeneous case, in this part we introduce both spatial and production heterogeneity.

1.3.3.1 Spatial Heterogeneity

We introduce spatial heterogeneity in a similar fashion as Cason and Gangadaran (2013). Specifically, the firms are positioned at different geographical proximity relative to the monitoring point, which is located at the downstream of the river. We assume the emissions from the firms closer to the monitoring point generate larger recorded environmental damage than firms further from the monitoring point. As explained in Cason and Gangadaran (2013), this is because the pollutants from the upstream firms are more diluted as they arrive at the monitoring point, while the emissions from the downstream firms are more concentrated. Miao et al. (2016) also introduces spatial heterogeneity by imposing a nutrient transport model to calculate the marginal damage of each farmer. The model includes two effects in determining pollutant concentration, the duration effect, which increases the marginal damage of upstream farmers, and the magnitude effect, which increases marginal damage of downstream farmers. Depending on parameterization of the model, either effect may dominate. We follow the heterogeneity introduced in Cason and Gangadaran (2013) since it would allow us to solve for a closed form solution and it creates less complexity for the participants.

Specifically, let β_i denote the marginal environmental damage from emissions generated by firm i. The environmental damage caused by firm i can thus be written as $D_i(x_i) = \beta_i x_i$, and the total environmental damage is $TD = \sum_{i=1}^{N} D_i(x_i) = \sum_{i=1}^{N} \beta_i x_i$.

Given the same private profit function as before, under no policy scheme, the profit maximizing firm would produce at $x_i = \gamma_2$.

The social planner's problem could be solved in a similar fashion, resulting in an optimal production level at $x_i = \gamma_2 - \frac{\beta_i}{2\gamma_1}$.

Imposing a tax/subsidy scheme as before, we would be able to solve for the private optimal production level under policy, which is $x_i = \gamma_2 - \frac{\beta_i}{2\gamma_1}$. The corresponding socially optimal level under tax/subsidy, which is the same as without tax, is $x_i = \gamma_2 - \frac{\beta_i}{2\gamma_1}$. Therefore, the private optimal agrees with the social optimal.

Now consider we apply a similar technology, which requires an installation cost $\tau\gamma_2$, but reduces environmental damage to a rate α of the original level.

Similar as before, by adopting the technology, the private income function of a farmer is:

$$B(x_i) = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - \tau\gamma_2$$

and the environmental damage caused by each firm is reduced to $D_i(x_i) = \beta_i \alpha x_i$

We again find the equilibrium by backward induction. Consider firm i, given the pollution level of others in the group D_{-i}, its profit function from producing x_i and adopting the technology is:

$$\pi_i^A = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - (D_{-i} + \beta_i \alpha x_i - \overline{D}) - \tau\gamma_2$$

$$\frac{\partial \pi^A}{\partial x_i} = 2\gamma_1(\gamma_2 - x_i) - \beta_i \alpha = 0$$

$$x_i^A = \gamma_2 - \frac{\beta_i \alpha}{2\gamma_1}$$

Plug in the optimal production level to get the maximum profit of adopting the technology:

$$\pi_i^A = \gamma_0 - \frac{(\beta_i \alpha)^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_i \alpha \gamma_2 - \frac{\beta_i^2 \alpha^2}{2\gamma_1}\right) - \tau\gamma_2$$

Consider firm i, given the same pollution level from others D_{-i}, firm's profit function of not adopting the technology and producing at x_i is:

$$\pi_i^N = \gamma_0 - \gamma_1(\gamma_2 - x_i)^2 - (D_{-i} + \beta_i x_i - \overline{D})$$

$$\frac{\partial \pi^N}{\partial x_i} = 2\gamma_1(\gamma_2 - x_i) - \beta_i = 0$$

$$x_i^N = \gamma_2 - \frac{\beta_i}{2\gamma_1}$$

The maximized profit of not adopting is:

$$\pi_i^N = \gamma_0 - \frac{{\beta_i}^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_i\gamma_2 - \frac{\beta_i^2}{2\gamma_1}\right)$$

We parameterize the heterogeneity treatment of this experiment so that depending on β_i, half of the farmers prefer to adopt, and half of the farmers prefer not to adopt. In order for the farmer to prefer adopting the technology, it requires $\pi^N < \pi^A$.

Solving for these conditions, we get the condition for a farmer to prefer adopting:

$$\frac{\beta_i^2}{4\gamma_1}(1-\alpha^2) - \beta_i\gamma_2(1-\alpha) + \tau\gamma_2 < 0$$

By setting different β_i, we can create a unique dominant strategy Nash Equilibrium so that it is optimal for some farmers to adopt, some not to adopt, depending on their proximity to the monitoring point.

1.3.3.2 Production Heterogeneity

We next introduce production heterogeneity by varying the size of the farmers, in a similar way as Spraggon (2002). Recall in the case of homogenous production functions under no policy schemes, the farmers maximize their own profit by setting production at $x_i = \gamma_2$. Now suppose the farms are of different sizes, meaning that their maximum capacities are different. The production function for farmer i who does not adopt the technology is $B_i(x_i) = \gamma_0 - \gamma_1(\gamma_{2i} - x_i)^2$, for farmer i who adopts the technology is $B_i(x_i) = \gamma_0 - \gamma_1(\gamma_{2i} - x_i)^2 - \tau\gamma_{2i}$. Farmers maximize their profit by producing at $x_i = \gamma_{2i}$.

When there is no location heterogeneity, the environmental damage caused by each farmer is $D(x_i) = \beta x_i$ without using technology, and $D(x_i) = \beta\alpha x_i$ with technology.

Solving for the social planner's problem, the socially optimal production level for each farmer is $x_i = \gamma_{2i} - \frac{\beta}{2\gamma_1}$.

With the same tax/subsidy policy scheme, the private optimal production level is $x_i = \gamma_{2i} - \frac{\beta}{2\gamma_1}$ and the socially optimal production level remains to be $x_i = \gamma_{2i} - \frac{\beta}{2\gamma_1}$.

Consider the decision of whether or not to adopt the technology, solving it in a similar fashion, we can get the optimal profit for a firm to adopt the technology is $\pi_i^A = \gamma_0 - \frac{(\beta\alpha)^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta\alpha\gamma_{2i} - \frac{\beta^2\alpha^2}{2\gamma_1}\right) - \tau\gamma_2$, which is reached by producing $x_i^A = \gamma_{2i} - \frac{\beta\alpha}{2\gamma_1}$, and for the farmer to not adopt the technology is $\pi_i^N = \gamma_0 - \frac{\beta^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta\gamma_{2i} - \frac{\beta^2}{2\gamma_1}\right)$, which can be reached by producing $x_i^N = \gamma_{2i} - \frac{\beta}{2\gamma_1}$. The condition for a farmer to prefer to adopt compared with not adopt is $\frac{\beta^2}{4\gamma_1}(1-\alpha^2) - \beta\gamma_{2i}(1-\alpha) + \tau\gamma_2 < 0$.

By setting different β and γ_i, we can create a unique dominant strategy Nash Equilibrium so that it is optimal for some farmers to adopt, some not to adopt, based on their farm size and spatial location.

1.3.3.3 Spatial and Production Heterogeneity

Now, we include production heterogeneity and spatial heterogeneity simultaneously. As before, the production function for farmer i who does not adopt the technology is $B_i(x_i) = \gamma_0 - \gamma_1(\gamma_{2i} - x_i)^2$, for farmer i who adopts the technology is $B_i(x_i) = \gamma_0 - \gamma_1(\gamma_{2i} - x_i)^2 - \tau\gamma_{2i}$. Farmers maximize their profit by producing at $x_i = \gamma_{2i}$.

Meanwhile, the environmental damage caused by each farmer is $D_i(x_i) = \beta_i x_i$ without using technology, and $D_i(x_i) = \beta_i \alpha x_i$ with technology.

Solving for the social planner's problem, the socially optimal production level for each farmer is $x_i = \gamma_{2i} - \frac{\beta_i}{2\gamma_1}$.

With the same tax/subsidy policy scheme, the private optimal production level is $x_i = \gamma_{2i} - \frac{\beta_i}{2\gamma_1}$ and the socially optimal production level remains to be $x_i = \gamma_{2i} - \frac{\beta_i}{2\gamma_1}$.

Consider the decision of whether or not to adopt the technology, solving it in a similar fashion, we can get the optimal profit for a firm to adopt the technology is $\pi_i^A = \gamma_0 - \frac{(\beta_i\alpha)^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_i\alpha\gamma_{2i} - \frac{\beta_i^2\alpha^2}{2\gamma_1}\right) - \tau\gamma_{2i}$, which is reached by producing $x_i^A = \gamma_{2i} - \frac{\beta_i\alpha}{2\gamma_1}$, and for the farmer to not adopt the technology is $\pi_i^N = \gamma_0 - \frac{{\beta_i}^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_i\gamma_{2i} - \frac{\beta_i^2}{2\gamma_1}\right)$, which can be reached by producing $x_i^N = \gamma_{2i} - \frac{\beta_i}{2\gamma_1}$. The condition for a farmer to prefer to adopt compared with not adopt is $\frac{\beta_i^2}{4\gamma_1}(1-\alpha^2) - \beta_i\gamma_{2i}(1-\alpha) + \tau\gamma_{2i} < 0$. Therefore, the optimal strategies of the farms depend on their farm size and spatial location.

We parameterize the experiment so that when there is no spatial or size heterogeneity, it is optimal for everyone in the same group to adopt; when at least one type of heterogeneity is introduced, it is optimal for half of the people to adopt the technology, and the other half to not adopt.

1.3.4 Information

Besides the baseline where no information is provided to the participants, we conduct two information treatments. Both information treatments include narrative messages on how the participant's decisions compare with others. In information treatment 1, we provide participants testimonial information on what production and technology adoption decisions people "like them" have made in the past. The idea is to use this information nudge to help people find their optimal strategies. To ensure that the information participants receive are truthful, the information we provide comes from the "no information" treatments. Conditioning on their size and location, we find the actual decisions made by participants that are closest to the Nash Equilibrium. Therefore, this information differs by the location and the size of the firm and approximates the actual Nash Equilibrium. This resembles some policy recommendation on what people should consider doing based on their location and size. In information treatment 2, we give participants information on the technology adoption rate and average production in their group in the last round. With this information, participants will have knowledge on their group members' peer actions and how they compare with others in the group. This is similar to a policy that provides information on what the majority of people in a neighborhood are doing and has a self-evolving nature.

Since our experiment features a dominant strategy Nash Equilibrium in each of the treatments, the optimal strategies are not influenced by the information treatments. Therefore, if participants are fully rational, neither of the information treatments should influence their behavior. However, as demonstrated by previous studies, people evaluate the appropriateness of their behavior by comparing to others and may change their behavior accordingly. We are interested in see if these nudges could be used to increase the performance of ambient based policy.

1.4 Experimental Design

1.4.1 Treatments

The basic setup of our experiment is a three by two design. As shown in Table 1.1, on the between subject level, we conduct three information treatments (including no information as the baseline). On the within subject level, we vary whether an ambient-based policy scheme is being implemented. Within each policy treatment, we further break up by heterogeneity treatments: homogeneous (Homo); heterogeneous type 1 (Hetero1) with only spatial heterogeneity; heterogeneous type 2 (Hetero2) with only production heterogeneity; and heterogeneous type 3 (Hetero3) with both spatial and production heterogeneity. We vary the order of the heterogeneity treatments that were presented. Four sessions were ran for each information treatments, making 12 sessions in total. Within each session, we have two groups of participants and each group consists of 8 people.

During the experiment, each participant makes 5 decisions in each policy/heterogeneity treatment. The groups are randomly reassigned after each policy/heterogeneity treatment. A five-round practice part is conducted at the beginning of each session to help participants familiarize themselves with the computer program. After the experiment, we gather a few quick survey questions on participants' basic demographics.

1.4.2 Parameterization

The parameters of the experiment are shown in Table 1.2:

Most of the parameters follow previous experiments in the literature. The parameters for size heterogeneity stem from Spraggon (2002) and location heterogeneity are based on Cason and Gangadaran (2013). For the Homogeneous treatment, it is optimal for all the participants to choose to adopt; for the heterogeneous treatments, it is optimal for half of the participants to adopt, and the other half to not adopt. The social planner's optimal strategy to maximize social welfare agrees with the dominant strategies of each participant.

1.5 Hypotheses

We summarize the hypotheses in Table 1.3.

1.5.1 Hypothesis 1

Hypothesis 1 focuses on group level effect of the ambient based policy. Without the Tax/Subsidy policy, subjects will pollute at their maximum level. With Tax/Subsidy in place, group level pollution would be reduced to the target level despite heterogeneity or information treatments.

1.5.2 Hypothesis 2

Hypothesis 2 deals with group level policy efficiency. In information treatment 1, participants are given individual level information on what others like them have done, and such information relate to their optimal strategies. We posit this information nudge should improve policy efficiency. In information treatment 2, the average adoption and production levels in each group are provided to the participants. It is likely that participants would anchor their decisions to the group averages, but the direction that this nudge changes policy efficiency is ambiguous and we test it empirically.

1.5.3 Hypothesis 3

Hypothesis 3 aims at individual level decision making. Compared to no information baseline, we anticipate to observe an increase in optimal decision-making at the individual level for information treatment 1. However, for information treatment 2, the effect is unclear. Participants' decisions should be anchored towards the group average. If this anchoring is in the direction towards private optimal decisions, this information should increase the frequency of individual dominant strategies; however, if the anchoring effect biases decision making to a non-optimal direction, people should be further away from optimal.

1.6 Results

One pilot session and twelve real sessions were conducted in November and December, 2016 in a large public university in Northeastern United States. 192 participants took part in our real sessions.

We divide the results into three sections. The first section focuses on the aggregate group pollution levels and compare them to the target levels. The second section discusses what influences the efficiency of the policy. The third section deals with individual level decisions and how they compare to theoretical predictions.

1.6.1 Result 1

Without any information, group level pollution is not significantly different from the target level under homogeneous case with ambient tax/subsidy. As more heterogeneity is introduced, group level pollution exceeds the target. Both information treatments make the group level pollution closer to the target level.

Figure 1.1 and Figure 1.2 depict the group aggregate pollution levels for the no policy and policy treatments. The four segments in each figure means four size and location homogeneity/heterogeneity treatments. The x axis denotes round number (in total five) and y axis represents environmental damage level. The red, green and blue lines indicate the average group pollution level for no information baseline, information treatment 1, and information treatment 2 scenarios, respectively. The segregated dots represent outliers. In the no policy treatments, the theoretical predicted pollution level is 240, which happens when everyone produces at their maximum and not adopting the technology. In the no policy treatments, all groups are polluting close to their maximum level without much variation among the treatments. In Figure 1.2, when ambient policy is introduced, the black dotted lines represent the target group pollution level. In general, as more heterogeneity is introduced, the lines deviate more from the target pollution level. Besides, the green line (denoting information treatment 1) is generally closer to the target, especially compared to the red line.

We next compare group pollution levels quantitatively. Table 1.4 suggests that aggregate pollution levels are not significantly different from the target levels in homogeneous cases for all information scenarios. With location heterogeneity only (Hetero1), the policy would still induce group level pollution to meet the target in no information and information treatment 2; Under information treatment 1, the group marginally under pollutes. When instead size heterogeneity is introduced (Hetero2), group pollution marginally exceeds the target level under no information, and does not significantly differ from the target under either information treatments. When two layers of heterogeneity are combined together, group level pollution significantly exceeds the target level in the no information case. With information treatment 2, the group level is marginally significantly different from the target while with information treatment 1 it is not significantly different.

The aggregated results reinforce our findings. When all information treatments are combined, we find that the group level pollutions are not significantly different from the target in homogeneous and heterogeneity 1 treatments, but are significantly different from the target at 5% level in heterogeneity 2 and at 1% level in heterogeneity 3 treatments, as shown in the last column of Table 1.4. When we instead combine homogeneous/heterogeneous treatments and compare the effects of information treatments (as shown in the last row of Table 1.4), we find that overall the group pollution level is significantly different from the target level at 5% in no information treatment, but not significantly different from the target level in the other two information treatments.

1.6.2 Result 2

From a social planner's perspective, policy efficiency decreases as more heterogeneity is introduced, however both information treatments increase efficiency.

Similar to Spraggon (2013), efficiency is defined as the change in the value of the social planner's problem as a percentage of the optimal change in the social planner's problem. The social planner's problem for a group could be formulated as follows:

$$SP = \sum_{i=1}^{8}[\gamma_0 - \gamma_1 * (\gamma_2 - x_i)^2 - a_i * \tau * \gamma_2 - a_i * \alpha * \beta_i * x_i - (1 - a_i) * \beta_i * x_i]$$

Efficiency is calculated as:

$$E = \frac{SP_{Actual} - SP_{StatusQuo}}{SP_{Optimal} - SP_{StatusQuo}}$$

where SP_{Actual} is the actual value of the social planner's problem when calculated using the actual decisions of the participants; $SP_{Optimal}$ is the optimal value of the social planner's problem when the participants all choose their optimal production and technology decisions; $SP_{StatusQuo}$ is the value of the social planner's problem when all participants choose to produce at their maximum and do not adopt the technology. Theoretically, $SP_{Optimal}$ and $SP_{StatusQuo}$ should correspondingly be the upper and lower bounds of the social planner's problem. Therefore, efficiency is a value between 0 and 1.

Table 1.5 presents efficiency values by heterogeneity and information treatments. Only the treatments with policy are presented here since there is no policy efficiency in the no policy treatments.

We observe that overall, policy efficiency is highest for information treatment 1, then followed by information treatment 2, and the lowest is no information treatment. Meanwhile, as more heterogeneity is introduced, policy efficiency decreases. In the homogeneous treatments, on average the tax/subsidy policy achieved 88.13% efficiency, while in heterogeneous 3 treatments where both location and size heterogeneities were introduced, the average policy efficiency was only 72.77%.

We also conduct a random effects regression at the group level to understand how efficiency is influenced by treatments. The regression is written as:

$$\begin{aligned} Efficiency_i = \alpha &+ \beta_1 * Hetero1_{it} + \beta_2 * Hetero2_{it} + \beta_3 * Hetero3_{it} + \beta_4 \\ &* Info1_{it} + \beta_5 * Info2_{it} + \beta_6 * Hetero1\_info1_{it} + \beta_7 \\ &* Hetero1\_info2_{it} + \beta_8 * Hetero2\_info1_{it} + \beta_9 * Hetero2\_info2_{it} \\ &+ \beta_{10} * Hetero3\_info1_{it} + \beta_{11} * Hetero3\_info2_{it} + \beta_{12} * round_{it} \\ &+ \beta_{13} * round\_sq_{it} + v_i + e_{it} \end{aligned}$$

where v_i is individual level random effects, e_{it} is individual and time specific error term.

As shown in Table 1.6, heterogeneous treatment 1 increases efficiency by 0.002 percentage points, heterogeneous treatment 2 and 3 decrease efficiency by 14.14 and 13.39 percentage points, respectively. Meanwhile, information treatment 1 increases efficiency by 10.08 percentage points and information treatment 2 increases efficiency by 8.69 percentage points (marginally significant). However, a Wald test suggests that we cannot reject that these effects are statistically the same. The interaction terms of information and heterogeneity, round and round-squared controls are not significant at the 5% level.

This result suggests that policy efficiency decreases as more heterogeneity is introduced, and increases in either of the information treatments. Though it appears that information 1 generates higher policy efficiency compared to information 2, this effect is not statistically significant.

1.6.3 Result 3

At individual decision level, introducing more heterogeneity leads to larger deviations of pollution from theoretical pollution predictions, and both information treatments reduce deviations from theoretical values.As discussed in previous sections, there exists a unique dominant strategy Nash Equilibrium for each of the participant's decisions. We calculate the predicted pollution levels based on theoretical predicted production and adoption decisions. We are ultimately interested in how people's decisions deviate from theoretical predictions (which is also the socially optimal decisions) and how to induce better behavior by reducing this deviation from both directions (over and under pollute). Therefore, instead of simply taking the difference of the individual pollution level to the theoretical level, we calculate the absolute deviation of the two values. To standardize the deviation across all treatments, we calculate a percent absolute difference from the actual pollution level, predicted pollution level and the maximum pollution level, similar to the metric used in Spraggon (2013). Specifically,

$$\text{PerAbsDiff}_i = \left|\frac{p_i - p_i^*}{p_i^{max}}\right|$$

where p_i represents the actual pollution level by participant i; p_i^* stands for the theoretical predicted Nash Equilibrium pollution level of participant i; p_i^{max} is the maximum pollution level of participant i.

1.6.3.1 Individual Results by Treatment

We run a random effects model that includes indicators for treatments and their interactions, as well as round and round squared. The regression model is as follows:

$$\begin{aligned} PerAbsDiff_i = \alpha &+ \beta_1 * Hetero1_{it} + \beta_2 * Hetero2_{it} + \beta_3 * Hetero3_{it} + \beta_4 \\ &* Info1_{it} + \beta_5 * Info2_{it} + \beta_6 * Hetero1\_info1_{it} + \beta_7 \\ &* Hetero1\_info2_{it} + \beta_8 * Hetero2\_info1_{it} + \beta_9 * Hetero2\_info2_{it} \\ &+ \beta_{10} * Hetero3\_info1_{it} + \beta_{11} * Hetero3\_info2_{it} + \beta_{12} * round_{it} \\ &+ \beta_{13} * round\_sq_{it} + v_i + e_{it} \end{aligned}$$

where v_i is individual level random effects, e_{it} is individual and time specific error term.

We conduct separate regressions for the no policy treatments and policy treatments and the results are listed below:

In the no policy treatments, the heterogeneity and information treatments alone do not affect the deviation of actual pollution levels to the theoretical predictions. However, for information treatment 1, deviations decrease significantly for heterogeneity 2 and heterogeneity 3 treatments, meaning that in hetero2 and hetero3 treatments with individual information, participants are less likely to deviate from their Nash Equilibrium.

When we look at the data with the tax/subsidy policy instrument (the last three columns of Table 1.7), participants deviate more from the Nash predictions in hetero2 and hetero3 treatments compared to the homogeneous treatment, and the deviation decreases in both info 1 and info 2 treatments. The interaction terms suggest that in treatments under information 1, the deviation from Nash in hetero 1 treatment is significantly more than the homogeneous treatment. Similarly, under information treatment 2, hetero1, 2 and 3 treatments all have significantly higher deviation than the homogeneous treatment under information 2. This suggests that people are able to find and retain their Nash equilibrium better in both information treatment 1 and information treatment 2, compared to no information baseline. Also, heterogeneity has less effect on deviation from Nash in information treatment 1 than in information treatment 2, meaning that individual level more tailored information helps people better overcome the heterogeneity.

An explanation for the more robustness of information treatment 1 to heterogeneity compared to information treatment 2 is that in information treatment 1, the social comparisons are individual level, which takes into account heterogeneity by providing different information based on different sizes and locations. However, information 2 provides group level information about average peer actions. As more heterogeneity is introduced, the span of every individual's optimal strategy is wider. Anchoring to the group average values no longer accounts for heterogeneity, and therefore we observe more deviations from the theoretical predictions.

1.6.3.2 Individual Results by Information, Location and Size

Next, we test how information treatments affect deviations from theoretical predicted values across different sizes and locations of farms.

The first three columns of Table 1.8 denote results from the No Policy treatments. Almost all of the variables are insignificantly different from zero. This means that regardless of the information treatment, size or location, people are generally polluting at the maximum level, which is in line with the theoretical prediction.

The last three columns demonstrate results from treatments with the ambient based policy. Several results are worth pointing out: first, compared to medium sized farms, small farms deviate more from the target pollution level; second, farms at the most downstream marginally deviate more from the predicted level compared to a farm in mid-stream; third, both information treatments reduce deviations from the theoretical prediction; forth, interactions of information and size, interactions of information and location are widely insignificant, meaning that information treatments are equally effective to participants with different farm size or location.

Compared to results from the previous subsection, these results demonstrate that both types of information nudges have impacts on people's decision making. Moreover, the impact appears to have identical effects for farms at different locations or with different sizes, meaning that people's responses to information nudges are robust to their relative size and location. This result adds confidence in using information nudges based on social comparison theory as a policy intervention since it is equally effective to different subgroups of people.

1.7 Conclusions

In our study, we conduct an experiment on non-point source water pollution with location and size heterogeneity and an extended decision space that includes both a production and a technology decision. We find that as more heterogeneity is introduced, the ability for the tax/subsidy policy instrument to reduce group pollution to the target level decreases. However, the tax/subsidy policy increases its effectiveness with the introduction of two information nudges based on social comparison theory. In information treatment 1, people are provided with information on what others like them have done in the past, based on the size and location of their farm. In information treatment 2, we give people information on the mean production and adoption levels in their group in the past round. We further demonstrate that policy efficiency is negatively affected by heterogeneity but can be improved by either information treatment. Comparing individual pollution levels to Nash predictions, besides the findings that heterogeneity increases deviations from Nash and information decreases them, we observe that individual level information is more robust to heterogeneity compared to group level information. Furthermore, we find that both information treatments are equally effective to individuals possessing farms at different sizes or locations.

As a conclusion, the results suggest that introducing more heterogeneity and a more complex decision space result in ineffectiveness of the classic tax/subsidy ambient policy, but information nudges based on social comparison and peer actions are able to help the performance of the policy and more individually targeted information works better in terms of policy efficiency and individual level decision making. From a policy perspective, it is important to consider multiple layers of heterogeneity as well as a more complex decision space when designing ambient based policies, but information nudges have the potential to improve the performance of ambient based policies in a cost–effective way.

Table 1.1 Treatment Orders

<table>
<tr><td rowspan="8">No Info</td><td rowspan="2">Session 1</td><td colspan="4">No Policy</td><td colspan="4">Policy</td></tr>
<tr><td>Homo</td><td>Hetero1</td><td>Hetero2</td><td>Hetero3</td><td>Homo</td><td>Hetero1</td><td>Hetero2</td><td>Hetero3</td></tr>
<tr><td rowspan="2">Session 2</td><td colspan="4">No Policy</td><td colspan="4">Policy</td></tr>
<tr><td>Hetero3</td><td>Hetero2</td><td>Hetero1</td><td>Homo</td><td>Hetero3</td><td>Hetero2</td><td>Hetero1</td><td>Homo</td></tr>
<tr><td rowspan="2">Session 3</td><td colspan="4">Policy</td><td colspan="4">No Policy</td></tr>
<tr><td>Homo</td><td>Hetero1</td><td>Hetero2</td><td>Hetero3</td><td>Homo</td><td>Hetero1</td><td>Hetero2</td><td>Hetero3</td></tr>
<tr><td rowspan="2">Session 4</td><td colspan="4">Policy</td><td colspan="4">No Policy</td></tr>
<tr><td>Hetero3</td><td>Hetero2</td><td>Hetero1</td><td>Homo</td><td>Hetero3</td><td>Hetero2</td><td>Hetero1</td><td>Homo</td></tr>
<tr><td>Info 1</td><td colspan="9">4 sessions identical to No Info but with Information Treatment 1</td></tr>
<tr><td>Info 2</td><td colspan="9">4 sessions identical to No Info but with Information Treatment 2</td></tr>
</table>

Table 1.2 Parameter Choice

Parameter	Value	Parameter	Value
γ_0	40	γ_1	0.0025
γ_2	75, 100, 125	τ	0.082
α	0.5	β_i	0.24,0.28,0.32,0.36
β	0.30		

Table 1.3 Hypotheses Table

Topic	Hypotheses	Results
Group Level Pollution	H_0:	
	1. $Pollution_{NP} = Pollution_{Max}$	1. Fail to reject H_0
	2. $Pollution_{P\_Homo} = Pollution_{Target\_Homo}$	2. Fail to reject H_0
	3. $Pollution_{P\_Hetero1} = Pollution_{Target\_Hetero1}$	3. Fail to reject H_0
	4. $Pollution_{P\_Hetero2} = Pollution_{Target\_Hetero2}$	4. Reject H_0
	5. $Pollution_{P\_Hetero3} = Pollution_{Target\_Hetero3}$	5. Reject H_0
	6. $Pollution_{P\_NoInfo} = Pollution_{Target}$	6. Reject H_0
	7. $Pollution_{P\_Info1} = Pollution_{Target}$	7. Fail to reject H_0
	8. $Pollution_{P\_Info2} = Pollution_{Target}$	8. Fail to reject H_0
Group Level Efficiency	H_0:	
	1. $Efficiency_{NoInfo} = Efficiency_{Info1}$	1. Reject H_0
	2. $Efficiency_{NoInfo} = Efficiency_{Info2}$	2. Reject H_0
	3. $Efficiency_{Info1} = Efficiency_{Info2}$	3. Fail to reject H_0
Individual Level Pollution	H0:	
	1. $Pollution_{NP} = Pollution_{Predicted}$	1. Fail to reject H_0
	2. $\beta_{Deviation\_P\_Hetero} = 0$	2. Reject H_0
	3. $\beta_{Deviation\_P\_Info1} = 0$	3. Reject H_0
	4. $\beta_{Deviation\_P\_Info2} = 0$	4. Reject H_0

Table 1.4 Mean Group Total by Treatment

	Target	No information	Information treatment 1	Information treatment 2	Total
Homo (no heterogeneity)	84	89.96 (5.13) [8]	83.90 (2.44) [8]	81.89 (3.54) [8]	85.25 (2.25) [24]
Hetero1 (location hetero)	94.4	93.41 (6.49) [8]	90.79* (2.35) [8]	89.73 (4.32) [8]	91.31 (2.61) [24]
Hetero 2 (size hetero)	75	81.42* (3.43) [8]	76.39 (3.76) [8]	80.60 (4.15) [8]	79.47** (2.14) [24]
Hetero 3 (location & size hetero)	73	85.93** (4.99) [8]	78.47 (4.66) [8]	76.19* (1.93) [8]	80.20*** (2.42) [24]
Total		87.68** (2.56) [32]	82.38 (1.92) [32]	82.10 (1.93) [32]	84.06** (1.26) [96]

Each cell contains mean, (standard error) and [frequency]. *, **, *** indicate significant at 10%, 5% and 1% level, respectively.

Table 1.5 Group Efficiency Level by Treatment

	No info	Info1	Info2	Total
Homo	81.87% (0.087) [8]	93.98% (0.046) [8]	88.53% (0.11) [8]	88.13% (0.096) [24]
Hetero1	79.61% (0.12) [8]	90.73% (0.038) [8]	83.41% (0.083) [8]	84.59% (0.094) [24]
Hetero2	67.91% (0.083) [8]	86.44% (0.066) [8]	71.47% (0.11) [8]	75.27% (0.12) [24]
Hetero3	66.81% (0.11) [8]	78.34% (0.13) [8]	73.18% (0.058) [8]	72.77% (0.11) [24]
Total	74.05% (0.12) [32]	87.37% (0.095) [32]	79.15% (0.11) [32]	80.19% (0.12) [72]

Each cell contains mean, (standard error) and [frequency].

Table 1.6 Random Effects Model on Group Efficiency and Treatment Variables

	Coefficient	Std. Err.	P-Value
Hetero1	0.002	0.046	0.962
Hetero2	-0.141	0.040	0.000
Hetero3	-0.134	0.047	0.005
Info1	0.101	0.035	0.004
Info2	0.087	0.048	0.072
Info1_hetero1	-0.040	0.055	0.468
Info1_hetero2	0.052	0.058	0.371
Info1_hetero3	0.018	0.061	0.774
Info2_hetero1	-0.017	0.063	0.791
Info2_hetero2	-0.017	0.069	0.805
Info2_hetero3	-0.015	0.063	0.816
Round	0.002	0.010	0.833
Round_sq	-0.001	0.002	0.705
Constant	0.819	0.033	0.000
Num. of Obs.	480		
Num. of groups	96		
Wald chi2	111.52		
Prob > chi2	0.000		

All standard errors are clustered as group level.

Table 1.7 Random Effects Model on Individual Pollution and Treatment Variables

	Without Policy			With Policy		
	Coeffi.	Std. Err.	P-value	Coeffi.	Std. Err.	P-value
Hetero1	-0.002	0.008	0.788	0.006	0.009	0.517
Hetero2	-0.002	0.008	0.762	0.032	0.009	0.000
Hetero3	-0.001	0.008	0.888	0.054	0.009	0.000
Info1	-0.008	0.018	0.671	-0.072	0.162	0.000
Info2	-0.018	0.018	0.310	-0.056	0.162	0.001
Info1_hetero1	-0.017	0.011	0.107	0.043	0.012	0.001
Info1_hetero2	-0.025	0.011	0.016	0.017	0.012	0.181
Info1_hetero3	-0.030	0.011	0.004	0.012	0.012	0.351
Info2_hetero1	0.006	0.011	0.580	0.028	0.012	0.027
Info2_hetero2	0.003	0.011	0.754	0.038	0.012	0.003
Info2_hetero3	0.002	0.011	0.824	0.028	0.012	0.026
Round	-0.006	0.006	0.250	-0.002	0.007	0.731
Round_sq	0.001	0.001	0.379	0.0004	0.001	0.708
Constant	0.063	0.015	0.000	0.137	0.014	0.000
Num. of Obs.	3840			3840		
Num. of groups	192			192		
Wald chi2	26.68			229.13		
Prob > chi2	0.0138			0.0000		

All standard errors are clustered at individual level.

Table 1.8 Random Effects Model on Individual Pollution, Size, Location and Information.

	Without Policy			With Policy		
	Coeffi.	Std. Err.	P-value	Coeffi.	Std. Err.	P-value
Info1	-0.0108	0.018	0.556	-0.060	0.017	0.001
Info2	-0.0166	0.020	0.407	-0.047	0.020	0.020
Large	0.0003	0.009	0.968	0.005	0.014	0.725
Small	-0.0017	0.011	0.876	0.075	0.017	0.000
Region1	0.0035	0.017	0.832	0.013	0.015	0.412
Region2	-0.0054	0.008	0.521	0.002	0.020	0.928
Region3	-0.0080	0.006	0.214	0.006	0.013	0.655
Region4	0.0082	0.013	0.525	0.034	0.018	0.065
Info1_large	-0.0193	0.011	0.072	-0.003	0.017	0.882
Info1_small	-0.0196	0.013	0.138	-0.012	0.024	0.610
Info1_region1	-0.0089	0.019	0.637	0.016	0.023	0.501
Info1_region2	-0.0148	0.012	0.220	0.024	0.026	0.361
Info1_region3	-0.0035	0.010	0.728	0.019	0.019	0.296
Info1_region4	-0.0164	0.016	0.315	0.017	0.026	0.507
Info2_large	-0.0074	0.014	0.584	0.016	0.018	0.387
Info2_small	0.0073	0.015	0.630	0.022	0.023	0.344
Info2_region1	-0.0123	0.018	0.488	0.023	0.022	0.286
Info2_region2	-0.0015	0.013	0.909	0.022	0.025	0.391
Info2_region3	0.0245	0.016	0.137	0.008	0.024	0.744
Info2_region4	-0.0008	0.021	0.968	-0.017	0.026	0.510
Round	-0.0065	0.006	0.290	-0.002	0.006	0.684
Round_sq	0.0008	0.001	0.366	0.0004	0.001	0.649
Constant	0.0618	0.018	0.000	0.133	0.018	0.000
Num. of Obs.	3840			3840		
Num. of groups	192			192		
Wald chi2	35.02			148.06		
Prob > chi2	0.0385			0.0000		

All standard errors are clustered at individual level.

Figure 1.1 Group Pollution Levels for No Policy, by Treatments and Round

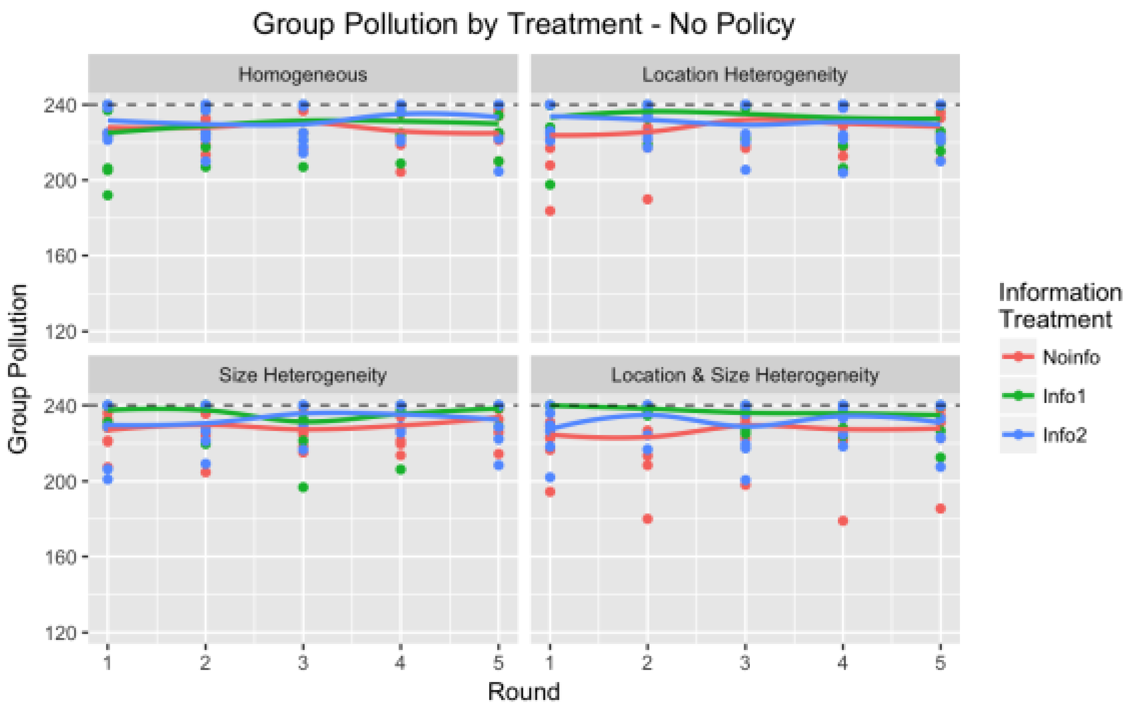
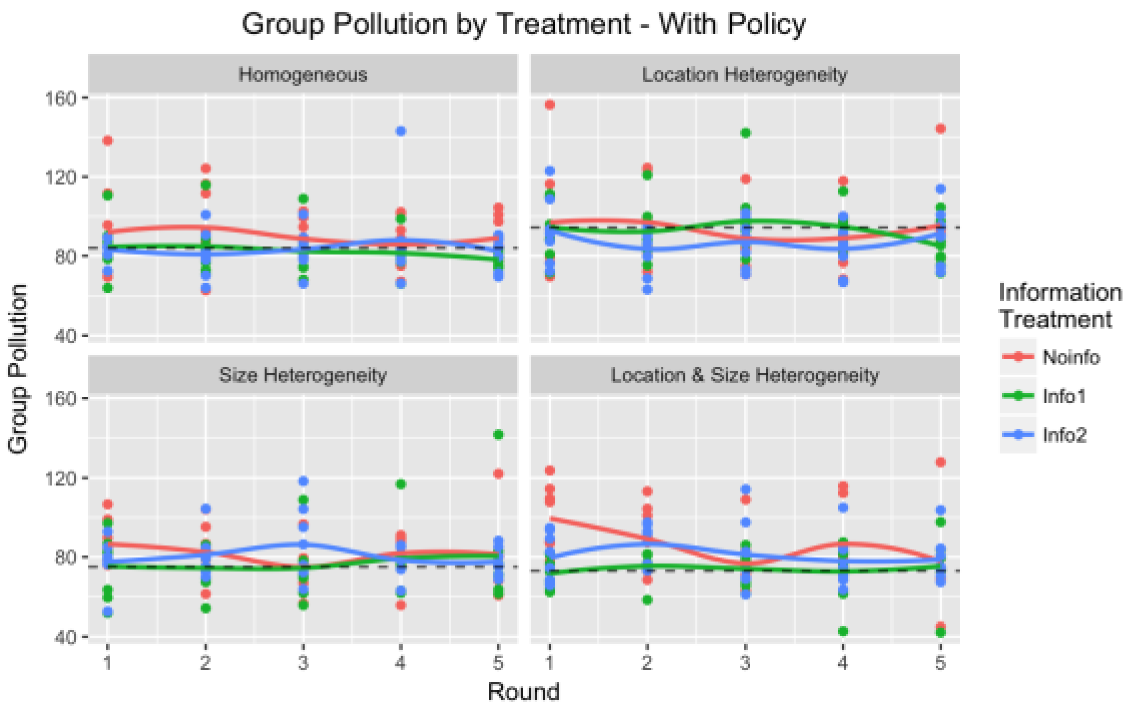


Figure 1.2 Group Pollution Levels for Policy, by Treatments and Round



Chapter 2

SIMULATING HETEROGENEOUS FARMER BEHAVIORS UNDER DIFFERENT POLICY SCHEMES: INTEGRATING ECONOMIC EXPERIMENTS AND AGENT-BASED MODELING

2.1 Introduction and Literature Review

Non-point source (NPS) pollution in water systems mainly comes from rainfall and snowmelt that move over and through the ground, bringing natural and human-made pollutants into waterbodies. NPS pollution, which comes mostly from nutrients and chemicals carried by agricultural runoff, is the primary cause of water pollution in the United States today. Unfortunately, regulation and remediation of NPS water pollution is a difficult task. It typically is hard and at times impossible to identify individual contributors to such pollution, and policies designed to address it must be designed to take polluters' hidden actions and asymmetric information into account. The cost of this type of individual monitoring and enforcement is often prohibitive (Xepapadeas, 2011).

Theoretical work (e.g., Segerson 1988, Xepapadeas, 1992) has shown that policies based on ambient levels of pollution can lead to reductions of NPS pollution to a regulator-specified target level. However, since no program has implemented an ambient-pollution-based policy on a large scale to provide empirical data, researchers have often turned to economic experiment lab settings as test beds for such policies

(Spraggon, 2002; Poe et al., 2004; Suter, Vossler and Poe, 2009, Miao et al., 2016). And since researchers must recruit and compensate participants in economic experiments, the experiments generally have been limited in scale and have restricted the ability to draw conclusions in contexts outside the lab. Thus, researchers have been interested in finding other ways to study the effects of these policies as part of efforts to improve their outcomes in terms of reducing NPS pollution.

Agent-based modeling (ABM) can help fill this gap by providing a mechanism for scaling up the findings in experiments to contexts that are closer to reality. With ABM, researchers can use findings from an experiment, create model agents that behave according to patterns identified in the experiment, and conduct simulations using an environment that better mimics a real-world setting. ABM also allows the researcher to observe the results of those agent interactions, which are extremely difficult to capture using other methods. Furthermore, we compared to traditional top-bottom methods such as econometric techniques, ABM imposes less distributional restrictions or assumptions.

ABM has been applied in various fields in recent years (Farmer and Foley, 2009), such as ecological modeling (Grimm and Railsback, 2005), population growth (Axtell et al., 2002), business strategies (Khouja, Hadzikadic, and Zaffar, 2008), land use policy (Tsai et al. 2015), transportation policy (Zia and Koliba 2015) and education (Johnson, Lemasters, and Bhattacharyya, 2017). In the context of agricultural and environmental applications, it has been used mainly for problems associated with changes in land cover to develop models that simulate land use decisions by farmers facing multiple constraints (Parker, 2014; Matthews et al., 2007; Veldkamp and Verburg, 2004), especially in studying coupled human and natural systems (An, 2012). In such systems, agent decisions generate environmental consequences, which could in turn affect human decisions and behavior. Recently, Tesfatsion et al. (2017) developed the Water and Climate Change Watershed (WACCShed) platform that allows the systematic study of interactions of hydrology, human and climate in a watershed over time. Ng et al. (2011) demonstrates an agent-based model of farmer decision making on water quality in the context of first and second generation biofuel crops and carbon trading. The ABM integrates a SWAT based hydrologic-agronomic model.

In the bottom-up construction of an ABM, modelers need to assign decision rules to agents under specific scenarios. A major challenge lies in constructing credible decision rules for ABM (Zenobia et al., 2009). Most of the previous work usually assumes perfect rationality, meaning that the agents could perfectly solve for utility maximizing problems in various and sometimes complex scenarios. However, behavioral economics have repeatedly shown that human behavior is often, at best, rationally bounded and that individuals often use heuristics instead of optimization when making decisions. As noted by Hechbert, Baynes, and Reeson (2010), combining economic experiments with ABM offers researchers many new opportunities. Experimental economics can be used to guide calibration of ABM so that the agents' behaviors and decisions reflect patterns identified by actions in experiments.

Some researchers have used survey methods to develop decision rules for ABM (Dia, 2002). Compared to using survey-based approaches to calibrate decision-making in ABM, we can use data collected through experiments that capture the “interpersonal” and “interplayer” dynamics that arise in experimental games (and are overlooked by surveys). Furthermore, Duffy (2006) pointed out that ABM projects also could facilitate researchers’ ability to interpret the aggregate findings of an experiment involving human subjects.

Not many studies have combined experimental economics and ABM. Evans, Sun, and Kelley (2006) compared results from a spatially explicit lab experiment to outputs of a simulation from a land-use ABM involving utility-maximizing agents. They concluded that the participants in the experiment deviated from revenue-maximizing actions and that it was thus valuable to use non-maximizing agents in ABM. Heckbert (2009) also acknowledged the value of combining experiments and ABM, reporting a study in which a participant replaces the role of an agent and the participant’s behavior under several treatments can be used to recalibrate the ABM.

A few studies have attempted to integrate economic experiments and ABM in NPS pollution management context. Zia et al. (2016, in review) constructed agent-based models using an economic experiment documented in Miao et al. (2016). The agents were categorized to pursue different behavioral strategies under alternate policy and sensor information regimes, and the agents’ type categories were predicted by a multi-level multinomial logistic regression model built from experimental data. Our research extends this idea by designing an experimental setting that includes technology adoption decisions and two layers of heterogeneity, meanwhile building a closer link between the experiment and the ABM.

We also include two information treatments to examine the ability of information ‘nudges’ to induce desired outcomes from the participants. Originating from the social comparison theory by Festinger (1954), it has been shown that information ‘nudges’ on social comparison and peer actions can promote environmental conservation behavior (e.g., Allcott, 2011; Ferraro and Price, 2013; Goldstein, 2008). These information ‘nudges’ are attractive from a policy design perspective since they are more cost-effective compared to traditional monetary based programs. However, not much research has considered incorporating information ‘nudges’ in NPS pollution management. We are interested in if information ‘nudges’ based on social comparison and peer action could help the performance of ambient based policies. In the first information treatment, participants are provided with information about what people “like them” have chosen in a similar situation in the past. In the second treatment, participants are provided with information regarding average production and average rate of adoption of technology by their group in the preceding round. Participants’ responses to the policy and the information treatments given the heterogeneity of production types are used to guide the agent’s behavior in the models under various scenarios.

In this study, we scale up findings from an economic experiment with ABM in a spatially explicit watershed setting to provide insight into the effects of different policy interventions addressing NPS pollution. The models capture interactions among heterogeneous agents in terms of diffusion of technology adoption by farmers, which is difficult to model using other techniques. Specifically, we test how tax/subsidy policies based on ambient levels of water pollution work in scenarios involving heterogeneous production and pollution schemes and focus on cases in which the decision space of the agents is extended from making a single production decision to making a production and a technology decision. We also investigate how information influences people's behavior and whether policies can be designed to incorporate information 'nudges' to induce more-desired outcomes.

Our study makes contribution to the literature in two main aspects. In environmental and resource economics, our experiment investigates the effect of information 'nudges' in an experimental setting that simultaneously incorporates an extended participant decision space and multiple layers of heterogeneity. Moreover, we use an ABM that features heterogeneous agents in a spatially explicit context to understand implications of the complex actions and interactions created based on experimental data. In the field of ABM, despite rising interest in using non-fully rational agents, not much work has actually done so. We are one of the first to introduce bounded rational agents into an ABM based on an economic experiment. The ABM agent decision rules are closely linked with human decisions in the economic experiment using an underlying game-theoretical model. Our research demonstrates that economic experiments can be useful to capture bounded rationality and guide ABM development. This study provides an example to incorporate human-based decision rules and a possible framework to integrate experiments and ABM in future research.

2.2 Experimental Design and Theoretical Foundation

In this part, we discuss the experimental design of our economic experiment. We first lay out the theoretical model, and describe the treatments in the experiment.

2.2.1 Theoretical Model

We build upon and extend the classic model framework in the environmental economics literature. Consider a group of agricultural producers indexed by i = 1 … N operate farms or ranches adjunct to a common watershed. The farmers' operations generate pollution as byproduct. A regulator monitors water quality by a sensor at the downstream of the watershed. The farms may differ both in their capacity and their distance to the sensor. The farmers may choose to adopt a pollution abatement technology (e.g., buffer, cover crop) at a cost (τ) proportional to farm size. Each year, the farmers make two decisions: a production decision x_i and a decision on whether to adopt an abatement technology a_i. $\frac{\partial PE_i(x_i,a_i)}{\partial x_i} > 0$, $\frac{\partial PE_i(x_i,a_i)}{\partial a_i} < 0$, indicating lower production and the adoption of the technology are associated with lower private earnings through $PE_i(x_i, a_i)$. The environmental damage generated by each farm is $D_i(x_i, a_i) = \alpha\beta_i x_i a_i + \beta_i x_i(1 - a_i)$, where $\frac{\partial PE_i(x_i,a_i)}{\partial x_i} > 0$, $\frac{\partial PE_i(x_i,a_i)}{\partial a_i} < 0$, and β_i depends on the location of the farm relative to the sensor and α denotes the effect of the technology. We assume that the total environmental damage is $TD = \sum_1^N D_i(x_i, a_i)$. Without any regulation, a profit maximizing farm will produce at their capacity level and not adopt the technology. The social planner's problem is to maximize social benefits (denoted as SP), where $SP = \sum_{i=1}^N PE_i(x_i, a_i) - \sum_{i=1}^N D_i(x_i, a_i)$. Suppose the regulator hopes to achieve a pollution standard D and imposes a tax/subsidy policy, where the tax/subsidy equals to the environmental damage minus the target level of pollution, $t(TD) = (TD - \overline{D})$. Following the literature, suppose $PE_i(x_i, a_i)$ takes a quadratic form $\gamma_0 - \gamma_1(\gamma_{2i} - x_i)^2 - \tau\gamma_{2i}a_i$, where $\tau\gamma_{2i}a_i$ takes into account whether the firm adopted the technology. Now the individual payoff function under the tax/subsidy scheme becomes: $\pi_i = PE_i(x_i, a_i) - (TD - \overline{D})$. We find the Nash strategy by backward induction. Consider firm i, given the pollution level of others in the group D_{-i}, its profit function from producing x_i and adopting the technology is: $\pi_i^A = \gamma_0 - \frac{(\beta_i\alpha)^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_i\alpha\gamma_{2i} - \frac{\beta_i^2\alpha^2}{2\gamma_1}\right) - \tau\gamma_{2i}$, taking first order condition, the maximum is reached at $x_i^A = \gamma_{2i} - \frac{\beta_i\alpha}{2\gamma_1}$. The profit for not adopting the technology is $\pi_i^N = \gamma_0 - \frac{\beta_i{}^2}{4\gamma_1} - \left(D_{-i} - \overline{D} + \beta_i\gamma_{2i} - \frac{\beta_i^2}{2\gamma_1}\right)$, and the maximum can be reached by producing $x_i^N = \gamma_{2i} - \frac{\beta_i}{2\gamma_1}$. The condition for a farmer to prefer to adopt compared with not adopt is therefore $C = \pi_i^N - \pi_i^A = \frac{\beta_i^2}{4\gamma_1}(1 - \alpha^2) - \beta_i\gamma_{2i}(1 - \alpha) + \tau\gamma_{2i} < 0$. Thus, a unique dominant Nash strategy for a farm is defined as $\{C < 0: x_i = \gamma_{2i} - \frac{\beta_i\alpha}{2\gamma_1}, a_i = 1;\ C \geq 0: x_i = \gamma_{2i} - \frac{\beta_i}{2\gamma_1}, a_i = 0\}$. This dominant Nash strategy is also the same as the social planner's optimal strategy.

2.2.2 Treatments

We consider two dimensions of treatments. On the within-subject level, we varied whether the tax/subsidy policy is in place and also the complexity of heterogeneity that is in the experiment. For each of the policy treatment, we conducted four heterogeneity treatments, namely,

(1) A homogeneous treatment where the locational impact on water quality and size of each farm is the same (Homo);

(2) A first heterogeneous treatment where the locational impact on water quality vary, but the size of each farm is the same (Hetero1)

(3) A second heterogeneous treatment where the size of the farms vary, but locational impact on water quality is the same (Hetero2);

(4) A third heterogeneous treatment where both size and locational impact on water quality of farms vary (Hetero3).

To control for potential order effects, we randomly varied the order of the within-subject treatments that are presented. On the between-subject level, we provided participants with three information treatments. No Info serves as the baseline. In the Info1 treatment, we provide testimonial information on what production and technology adoption decisions people "like them" have made in the past. The information comes from the "no information" treatments. We find true decisions participants made that are closest to the Nash optimal strategies conditioning on their size and location. Therefore, this information differs by the location and the size of the firm and approximates the actual Nash optimal strategies. This resembles some policy recommendation on what people should consider doing based on their location and size. In the information treatment 2, we give participants information on the technology adoption rate and average production in their group in the last decision. This is similar to a policy that provides information on what others in the neighborhood are doing and has a self-evolving nature. Since each decision is independent and each participant has a unique dominant Nash strategy, theoretically the information treatments should not change participants' decisions. However, as noted before, human decisions often demonstrate bounded rationality and may follow simple heuristics or *ad hoc* rules.

2.2.3 Experiment procedure

The economic experiment consists of twelve sessions conducted in late 2016, involving a total of 192 participants recruited at a large public university in the northeastern United States.

2.3 Agent-Based Model Setup

In this part, we discuss the ABM setup and initialization. We design the ABM to capture key elements of the economic experiment and an actual watershed while avoiding including unnecessary assumptions and processes. We first set the ABM to a spatially explicit context based on the Murderkill[1] watershed located in the southeast part of Kent County, Delaware (Figure 2.1). The Murderkill watershed is chosen

---

[1] Note that the origin of the name, Murderkill, has a Dutch origin as "moeder" means mother and "kill" means river or creek in Dutch. Thus, the rough translation of the name is "Mother River", and not a reference to a bloody past.

mainly because it consists primarily agricultural land use and it is a typical coastal plain. Besides, it has promulgated TMDL regulations and has research efforts on the estuarine portion of the watershed. Moreover, the watershed is comprised of 68,000 acres of land, which is large enough to generate meaningful conclusions, but not too large to create computational obstacles.

2.3.1 GIS Environment Setup

In our model, the agents are farmers operating farms in the watershed. However, since farm level data is not publicly available, we develop a method to simulate farm level agents from parcel level data. We obtain three sources of geographic information system (GIS) data for the Murderkill River watershed: (1) Parcel level size and location data for Delaware; (2) Watershed boundary data for Murderkill watershed; and (3) National Land Cover Database (NLCD, 2011). We combine these three data sources together to generate an estimate of the agricultural land for each parcel in the watershed.

2.3.2 Agent Initialization

By combining parcel-level GIS information with data on land cover for this watershed, we can estimate the amount of land used for agriculture within each parcel and the X-Y coordinates of the parcels. Since farms often consist of a constellation of parcels and we do not have data on the actual allocation of parcels to specific landowners, we initialize the size of each farm based on the probability density function from data from the 2012 Census of Agriculture (U.S. Department of Agriculture, 2012) for Kent County, Delaware. Using that information and the GIS information, we match a simulated landowner agent to various numbers of parcels. In this process, we first calculate a “distance matrix” that contains information on the geographic distance between the individual parcels and every other parcel in the watershed. We then create landowner agents by grouping the nearest neighboring parcels until they meet criteria identified by the probability density and average size of each category of farms in the Census of Agriculture. The result is that our agents constructed from neighboring parcels closely mimic the census data on farm size distributions. Figure 2.2 displays the farm size distribution of Kent County, Delaware, and Figure 2.3 shows our simulated farm size distribution.

2.3.3 Network and Layout

An agent is assumed to operate a farm that consists of a number of parcels. The agents are placed at the center points of their farms, which are determined using GIS data. Each agent is connected to a number of neighbors based on geographic proximity and influences those neighbors. The number of agents in one neighbor group is determined by the modeler at the beginning of each simulation.

2.3.4 ABM Model Framework

In the ABM, we adopt the modification of the classic model in environmental and resource economics as documented in our previous section. Each agent operates a farm and generates income by producing an agricultural product (e.g., corn) and simultaneously generate byproducts that cause NPS pollution. The agents may choose to adopt a technology at a cost proportional to its size that could reduce byproducts. As explained before, an underlying dominant Nash strategy could be solved for every agent in the watershed. Since the dominant Nash strategy is the same as the optimal strategy for the social planner's problem, we can treat the Nash strategy as the "Theoretical Target" level of participants' response. A pollution monitor (i.e., sensor) is placed at the downstream end of the watershed, and amount of pollution contributed by each farm is based on the farm's distance from the monitoring point (our experiment measured individual contributions of pollution in the same way). Different policy and regulatory scenarios influence the agents' production and technology-adoption decisions based on results drawn from the experiment. Table 2.1 summarizes the variables used in the ABM.

2.3.5 ABM Model Process Flow

Figure 2.4 demonstrates the process flow of our ABM. Each agent makes two decisions, a production decision and a technology decision. Both decisions influence the income received and the pollution generated by the agent. Combined with pollution generated by other agents, the total pollution is calculated. Depending on whether an ambient based policy is in place, the agent's income may be affected by a tax or subsidy based on the target level and the total environmental damage. This influence on income further affects agent decisions in the next year. An agent's production and adoption decisions are modeled based on the production and adoption deviations from the target levels. These deviations are modeled in two phases as demonstrated in the next section.

2.4 Experimental Data Analysis

We conducted statistical data analysis on data from the experiment as documented in Wu, Palm-Forster, and Messer (2017). The analysis was done in two phases. First, we are interested in classifying people into different behavior groups. The idea is to capture the inherent behavioral difference among people (e.g., some people are more environmentally friendly, some are more self-oriented, etc.) Second, after we classify participants into behavior groups, we estimate how agent production and adoption decisions are influenced by their location, size, information treatment and type. We use the results to calibrate agent decision rules in the ABM model.

2.4.1 Cluster Analysis

Since we do not have any pre-defined knowledge or want to impose any assumption on how many groups participant behavior should be clustered into, the goal of this analysis is to identify the number of behavior types and cluster agents into that number of groups. With no pre-determined grouping structure, meaning that we do not observe the response variables, cluster analysis is suitable for this purpose. As a popular unsupervised statistical learning method, cluster analysis could generate grouping structures based on patterns in predictors. The first key question is to determine how many clusters the agents should be grouped into.

2.4.1.1 Clustering Metric

To account for the fixed effects of different treatments, the difference between an agent's actual pollution level and the Nash optimal strategy level in that treatment was considered as a measure of the agent's behavior at each round. Therefore, clustering analysis was implemented based on five variables (diff1, diff2, diff3, diff4, diff5), the agents' differences to Nash over five rounds. These variables are defined as:

$$Diff_{ijt} = Pollution_{ijt} - TargetPollution_{ijt}.$$

Where $Diff_{ijt}$ denotes the difference of participant i's pollution level to the target pollution level in treatment j, round t.

There are a number of clustering methods available, the most popular ones include K-means clustering, hierarchical clustering and Gaussian mixture models. There is no definite right or wrong for each of the clustering methods. We selected to use K-means clustering because it generated the most informative grouping structure.

For K-means clustering, the most important task is to determine how many groups to cluster into. This depends on both statistical criterion and knowledge on what a sensible grouping structure is. We perform various statistical procedures to determine the number of clusters.

2.4.1.2 The Elbow Method

The most intuitive way to determine the number of groups is the "Elbow Method". Figure 2.5 depicts the within groups sum of squares versus the number of clusters. We can see that there is a sharp turn when the number of cluster is equal to three. Therefore, three appeared to be a reasonable number of clusters to divide the agent types into.

2.4.1.3 Calinski Criterion

Another popular method for this purpose is the Calinski Criterion (also known as the Pseudo F statistics). Figure 2.6 shows the results of applying Calinski Criterion to our data. The Calinski Criterion suggests that we should also use three clusters.

2.4.1.4 Majority Rule

Third, we applied 26 other indices on the same problem and use the majority rule to select the number of clusters. We consider up to ten clusters as the possible number of clusters that we could group into. As shown in Figure 2.7, the Y-axis means the frequency that a number is selected as the best number of clusters chosen by the indices, and the X-axis is the possible best number of clusters. Eleven out of the 26 indices selected three as the best number of clusters. Therefore, according to the majority rule, we will assume three is the number of clusters we should use in the K-means clustering.

2.4.1.5 Separation Examination

We examine if the three clusters generated by K-means clustering provide reasonable separation for the data. We perform K means clustering, assigning three as the number of clusters, and setting the seed to 20 to ensure reproducibility.

The results of the K-means clustering are summarized in Table 2.2. As we can see, the median values for Group 3 in all five rounds are equal to zero, meaning that group 3 is the group that tend to behave in accordance with the theoretical prediction. Group 1 and Group 2 have median values that are lower and higher than the target pollution, respectively. This means that Group 1 is the group that tend to generate less pollution than theoretically predicted and Group 2 is the group that tend to generate more pollution than theoretically predicted. We do not see obvious skewness or scarcity of any groups and the magnitude of the separation seems reasonable. Next, we assign agents in the ABM into behavior groups using a multinomial logit model.

2.4.1.6 Mixed-effects Multinomial Logit model to assign group probabilities

Based on cluster analysis, agents' behavior could be clustered into three categories. Cluster 3 corresponds to agents that tend to agree with theoretical predictions, and cluster 1 and 2 correspond to agents that tend to under and over pollute, respectively. In this part, we use a mixed effects multinomial logit model to estimate the cluster distributions among agents conditioning on the policy, heterogeneity and information treatments.

The multinomial logit model could be formulated as follows:

$$\log\left(\frac{Pr(cluster = 1)}{Pr(cluster = 3)}\right)\Bigg| (Policy = j)$$

$$= f(u_{1,i}, HeteroTreatments, InfoTreatments, HeteroInfo\_Interactions)$$

$$= X_i B_{1i}$$

$$\log\left(\frac{Pr(cluster = 2)}{Pr(cluster = 3)}\right)\Bigg| (Policy = j)$$

$$= f(u_{2,i}, HeteroTreatments, InfoTreatments, HeteroInfo\_Interactions)$$

$$= X_i B_{2i}$$

where $u_{1,i}, u_{2,i}$ are the random effects on the intercept and are assumed to follow a normal distribution. J equals 1 or 0 and denotes whether the policy treatment is in place or not, respectively.

Therefore, the predicted probabilities for the three clusters could be calculated as:

$$\Pr(cluster_i = 1) = \frac{\exp(X_i B_{1i})}{1 + \exp(X_i B_{1i}) + \exp(X_i B_{2i})}$$

$$\Pr(cluster_i = 2) = \frac{\exp(X_i B_{2i})}{1 + \exp(X_i B_{1i}) + \exp(X_i B_{2i})}$$

$$\Pr(cluster_i = 3) = \frac{1}{1 + \exp(X_i B_{1i}) + \exp(X_i B_{2i})}$$

The results of the mixed effects multinomial logit model for both policy and no policy treatments are presented in Table 2.3.

In the no policy treatments, it is always in the agents' best interest to produce at the maximum and not adopt the technology, therefore, the theoretical optimal strategy is the upper bound of the pollution level. As a result, only two clusters exist in the no policy treatments, as reflected by having one intercept value in Table 2.3. Based on the above regressions, we calculate the cluster probabilities for each of the treatment cases to initialize the model.

2.4.2 Modeling agent production and adoption behavior.

For production decisions, we calculate the percentage deviations from the target production decisions, taking into account the size of the farm. The metric is defined as:

$$PerProdDiff = \frac{Production - TargetProd}{Size}.$$

In this case, we run a random-effects OLS model for each policy and information segment of data with standard errors clustered at the individual level.

For adoption decisions, we calculate the probability that an agent deviates from its target adoption decision, which means the probability that an agent changes its adoption decision away from the theoretical prediction. The metric is defined as the absolute difference of the actual adoption decision and the target adoption decision:

$$AdoptChange = |ActualAdopt - TargetAdopt|$$

Since the variable AdoptChange is binary, we run a random effects logit model for each policy and information treatment segments with individual clustered standard errors. The result of the model is shown below:

Based on the above regressions, we parameterize the agent's actual production and adoption decisions relative to their Nash optimal strategy levels.

2.5 Calibration

2.5.1 Prices

Agents are assumed to produce an agricultural good (corn) and act as price-takers. Given the constant fluctuation of corn prices in the US, we conduct an OLS regression for the mean corn price from 1996 to 2016 on logarithm of year to capture the general price trend, and use a triangular distribution with maximum and minimum defined by the predicted mean and standard deviation of the prices to reflect the fluctuation.

2.5.2 Yield

In order to determine how much agricultural product (corn) is produced by each agent, we calculate the average yield of each unit of land. Similarly, we conduct an OLS regression for mean corn yield from 1996 to 2016 on logarithm of year to capture the general trend in corn production, and use a triangular distribution to reflect the fluctuation, with maximum and minimum defined by the standard deviation of the average yield.

2.5.3 Pollution

Following Zia et al. (2016a), we provide an estimate for the average Phosphorus leakage of corn fields based on the maximum and minimum Phosphorus loss estimates. During each simulation, the modeler has the option to modify the mean and standard deviation of average Phosphorus leakage. However, this value affects all simulation cases equally and therefore does not influence any relative comparison conclusions we draw.

2.6 Simulation Results

We present the results of the simulation experiment, and discuss the sensitivity of the results.

2.6.1 The Effect of Information Treatments

We compare how different information treatments would affect the performance of the ambient based policy. Both location and size heterogeneity are included in this simulation. In Figure 2.8, the red line indicates the target level of pollution, and the blue line indicates the experiment simulation results. From left to right, the three subfigures indicate the results for the no information case, individual level information case, and group level information case. As shown in Figure 2.8, a gap exists between the simulated pollution level (blue line) and the target pollution level (red line) in the no information case; however, the gap is much smaller in either of the information treatments. This suggests that under this simulation scenario, both information treatments decrease the deviation between the target pollution level and the simulated level, which indicates that the effect of policy is stronger when ambient based policy is coupled with information ‘nudges’.

2.6.2 Comparing Individual Decisions

In this section, we break up pollution by production decisions, adoption decisions and size of the farms. We look at each information treatment separately.

2.6.2.1 No Information

Recall that under no information baseline, the aggregate simulated pollution level is mostly over the target level, but the deviation does not appear to be large. However, when we break up pollution into production and adoption decisions by farm size, we observe huge deviations in these decisions (Figure 2.9). The small farms are significantly over adopting the technology (blue lines), and the large farms are widely under adopting, even though it is not in their best interest to do so (as depicted by the blue lines). Similarly, the small farms are also over producing and large farms are under producing.

2.6.2.2 Individual Level Information

When we provide participants with individual level information on what people like them have done in the past, we observe that the deviations from participants' behavior to the target levels are much smaller (Figure 2.10). This clearly demonstrates that the individual level information induces participants to make better decisions, and improves policy efficiency.

2.6.2.3 Group Level Information

When participants are informed with group level information on average adoption and production decisions in their group in the last round, we find that this information helps participants make better decisions than the no information baseline, but the policy efficiency is lower compared to individual level information scenario (Figure 2.11). Furthermore, small farms tend to over adopt and over produce, and the large farms tend to under adopt and under produce. However, compared to the target levels, the deviations between the target and the simulated results are smaller compared to the no information scenario, but larger than when people were given individual level information. Therefore, group level information helps the policy performance and efficiency on an aggregate level, but the policy efficiency is lower than if people were given individual level information.

2.6.2.4 Possible Explanations

Finally, we want to provide some discussion of the potential reasons for the patterns that were demonstrated in the simulation results. Under the no information treatment, the adoption of the technology is largely negatively related to the size of the farm. A possible explanation for this observation is that since the cost of adopting the technology is proportionally related to the size of the farm, participants may follow some heuristic decision rules that attribute significant weight to the cost of adopting in the processes. This clearly demonstrates that as opposed to always following profit maximizing decision rules, human behavior is often limited in their calculating ability and may be affected by various cognitive reasons and therefore demonstrate bounded rationality in terms of forming some rather heuristic decision rules. Furthermore, both information treatments seem to provide anchors for the participants. Knowing what people like them have done in the past and what others in their group have done provide people with a reference point in their decision process. Since individual level information provides people with tailored information, it helps people make better decisions compared to the myopic baseline case. Under the group level information where a group average is provided, we can observe that the absolute adoption and production decisions for farms with different sizes tend to be very close. This suggests that people might be anchored to the group level averages, or peer actions, even though it might not be in their best interest to do so.

2.7 Sensitivity Analysis

In this part we discuss how our results would be affected by uncertain parameters in our ABM. Ideally, the result of an ABM should come from complex agent interactions and adaptions in a concise model rather than from complex assumptions about individual behavior and free parameters (Axelrod, 1997). Most of the parameters that influence the observed results in the ABM are calibrated and validated based on experimental data. Therefore the uncertainty only results from realization of the randomness in each simulation experiment, which is stochastic in nature and should not generate any systematic biases. Meanwhile, if an uncertain variable affects each scenario of the simulation in an equal magnitude, the relative comparisons between the scenarios will not be affected. Therefore, one uncertain parameter that would possibly affect the result is how many farms the participants consider part of their group. This parameter affects the grouping structure and the group level information that is shown to the participants. In our baseline scenario presented before, we assume five people are considered to be in one group. We increase this parameter to ten, fifteen and twenty in this part and the result is shown in Figure 2.12.

As shown in Figure 2.12, as the number of people that the participants consider themselves to be in the same group with increase, the deviation from the target pollution level and the simulated pollution level is not largely affected under individual level information treatment, but increases under the group level information treatment. This suggests that individual level information not only generates highest policy efficiency, but also is more robust to participant perceptions on their group size.

2.8 Conclusions and Discussions

Our study is one of the first that integrates economic experiments with agent based modeling in a nonpoint source pollution setting. The ABM extends and scales up the findings from the economic experiment by providing a spatially explicit simulation environment based on an actual watershed. Instead of assuming full rationality, the economic experiment calibrates and validates the ABM by defining human-based bounded rational decision rules for the agents. We apply a modification of a classic game theoretical model from the environmental economics literature to the ABM and the experiment as the core underlying model in both scenarios. We define the target level (fully rational theoretical level) by solving for unique dominant Nash strategy. Using experimental data, we first identify the number of behavioral groups using exploratory cluster analysis and then group agents into the three identified groups by multinomial logistic model; second, we define agent decision rules by estimating adoption and production deviations from the target levels based on the information treatment, type, size and location of each agent. The result of our simulation experiment demonstrates that both information 'nudges' help the performance of the ambient-based policy. Individual level information induces higher policy efficiency compared to group level information, where the individual decisions tend to be anchored to the group averages, even though it may not be in their best interest. Our results show in a spatially explicit watershed setting that ambient-based policies, coupled with information ‘nudges’ to provide guidance to people’s behavior, have the ability to induce group level compliance, and the policy efficiency is higher when individual level information is being provided. Therefore, it is important to use informational ‘nudges’ to help people make better decisions, especially under complex heterogeneous scenarios.

There are a number of limitations and directions of future work based on our research. First, a more complicated hydrological model may be developed and incorporated in the ABM and the experiment. Examples of such models include the WWACShed model by Tesfation (2017) and the SWAT model used in Ng et al., (2011). However, if one attempts to also include bounded rationality in the agent decision processes and use economic experiments to capture these irrationalities, it is crucial to ensure that the conclusions from the experiment could be safely carried over to the ABM. In our experiment, this link was built by adopting the same underlying model and therefore the same incentives around the dominant Nash strategies. If a more complicated model is in place, it would be hard to solve for a perfect rational utility maximization prediction, and therefore would be difficult to have a baseline to compare with actual human behavior. Additionally, the more complicated a model is, the more information burden is introduced to the participants and the harder for the participants to generate informed decisions. So one needs to think carefully about the tradeoff.

Second, another extension of this research is to use farmer sample instead of a sample from university students in the experiment, aiming to increase externality validity of the experiment. The majority of research comparing samples from students and professionals generally find the two samples demonstrate similar responses in both agricultural (e.g., Cummings, Holt, and Laury, 2004; Messer et al. 2008; Fooks et al. 2016) and non-agricultural (Vossler et al., 2009) contexts, it may still be a valid extension since the decision process of farmers is likely different from that of students. However, one also needs to note that the farmers may treat the experiment as a pre-policy evaluation and therefore behave strategically in hopes to potentially influence policy makers (Suter and Vossler, 2013).

There has been very few articles in the literature trying to integrate experiments and agent-based modeling even though the integration would benefit both fields. This could probably largely be attributed to the interdisciplinary nature of the field, and the challenge to build a credible link between the two. The method in our study could be used as a framework to combine these two fields and help motivate future research in this area.

Table 2.1 Summary of Fixed, Variable, and Uncertain Parameters

Parameter Name	**Value**	**Description**
Fixed Parameters		
Number_of_agents	About 174, depending on each realization	Number of agents in simulation
Simulation_horizon	25	Length of each simulation in years
Probability_of_farm_type	Based on cluster analysis and multinomial logistic regression	The probability that each agent will fall into each behavioral type
Target_adoption_rate	Depends on farm grouping results	The target probability that each agent will adopt the technology
Target_production_rate	Depends on farm grouping results	The target production rate for each agent
Number_of_connections	User defined	Number of neighbors of each agent
Factor_technology	0.5	Percent of pollution relative to original level if technology is adopted
Variable Parameters		
Unit_corn_production	User defined	Weight of corn produced on one unit of farm size
Unit_pollution_generated	User defined	Average phosphorus generated by one unit of production
Uncertain Parameters		
Percent_prod_deviation	Depends on experiment data	Adjusts amount of corn produced per unit of land based on agent type
Adoption_change_prob	Depends on experiment data	Adjusts amount of pollution per unit of production based on agent type
Adopted	Binary, value depends on each realization	Indicates whether the farm adopted the technology
Agent Type	One of several types depending on the cluster analysis of the experiment data	Different types of agents determine different production, pollution, and adoption probabilities

Table 2.2 Group Frequencies and Median Value by Grouping Variables and Groups.

Cluster	Frequency	Diff1	Diff2	Diff3	Diff4	Diff5
1	171	-7.500	-5.60	-6.00	-6.16	-5.44
2	150	6.615	5.28	5.11	5.94	5.25
3	1215	0.000	0.00	0.00	0.00	0.00

Table 2.3 Mixed Effects Multinomial Logit Model to Assign Agent Behavioral Types.

	Policy		No Policy	
	Coefficient	Std. Err.	Coefficient	Std. Err.
intercept 1	-2.0416***	0.1119	-10.5036***	1.3820
intercept 2	-0.8200***	0.1127		
hetero1	0.3183*	0.1691	0.5156	0.4567
hetero2	0.5703***	0.1562	0.5156	0.4567
hetero3	0.4495***	0.1688	-0.5666	0.4796
info1	-1.2637***	0.2073	0.8735	1.3800
info2	-1.1185***	0.2123	-1.3688	1.4154
info1_hetero1	0.7143***	0.2746	-1.5310***	0.6499
info1_hetero2	0.6048**	0.2647	-4.2738***	0.8099
info1_hetero3	1.2097***	0.2745	-4.4023***	0.9440
info2_hetero1	0.3696	0.2631	0.1780	0.7027
info2_hetero2	1.3090***	0.2661	0.7511	0.6957
info2_hetero3	1.4883***	0.2566	2.8191***	0.7137
Number of Observations	3840		3840	
Number of groups	192		192	

***, **, * denote significant as 1%, 5% and 10% level. All standard errors are clustered as individual level.

Table 2.4 Deviations from Target Production Levels

	With Policy			Without Policy		
	No Info	Info1	Info2	No Info	Info1	Info2
constant	1.32*** (0.11)	0.16 (0.16)	1.41*** (0.13)	0.0015 (0.051)	-0.0039 (0.015)	-0.11 (0.14)
size	- 0.0095*** (0.00067)	-0.00086 (0.0031)	- 0.0089*** (0.00081)	-0.00019 (0.00016)	-0.000073 (0.000074)	0.000083 (0.00012)
region	-1.19*** (0.30)	0.25 (0.42)	-1.41*** (0.30)	0.032 (0.15)	0.023 (0.038)	-0.015 (0.026)
cluster1	-0.20*** (0.040)	-0.19*** (0.04)	-0.13*** (0.038)	-0.12*** (0.042)	-0.068** (0.032)	-0.057** (0.024)
cluster2	0.26*** (0.031)	0.21*** (0.039)	0.19*** (0.036)			
info1_adopt		-0.21** (0.085)				
info1_prod		- 0.000028 (0.0029)				
info2_adopt			0.064 (0.063)			0.00040 (0.013)
info2_prod			-0.0020* (0.0011)			0.00098 (0.0014)

***, **, * denote significant as 1%, 5% and 10% level. All standard errors are clustered as individual level.

Table 2.5 Deviations from Target Adoption Decisions

	With Policy			Without Policy		
	No Info	Info1	Info2	No Info	Info1	Info2
constant	6.76*** (2.17)	5.69** (2.33)	9.67 *** (2.50)	-2.86 (2.05)	-5.17*** (1.09)	-14.96 (10.73)
size	-0.051*** (0.012)	-0.018 (0.042)	-0.050*** (0.013)	-.0096 (0.012)	-0.0041 (0.011)	-0.016 (0.017)
region	-7.09 (4.66)	-18.09*** (6.73)	-18.44*** (5.85)	-6.90 (5.96)	3.48 (3.35)	11.01* (6.61)
cluster1	-0.39 (0.44)	0.63 (0.57)	-0.27 (0.59)	2.13** (0.91)	3.48*** (0.54)	3.79*** (0.71)
cluster2	0.74** (0.33)	1.22*** (0.44)	1.10*** (0.38)			
info1_adopt		-2.29* (1.23)				
info1_prod		0.013 (0.040)				
info2_adopt			-0.59 (0.90)			4.48** (2.07)
info2_prod			-0.0012 (0.018)			0.079 (0.11)

***, **, * denote significant as 1%, 5% and 10% level. All standard errors are clustered as individual level.

Figure 2.1 Murderkill River Watershed, Delaware, United States. Figure source: delawarewatersheds.org

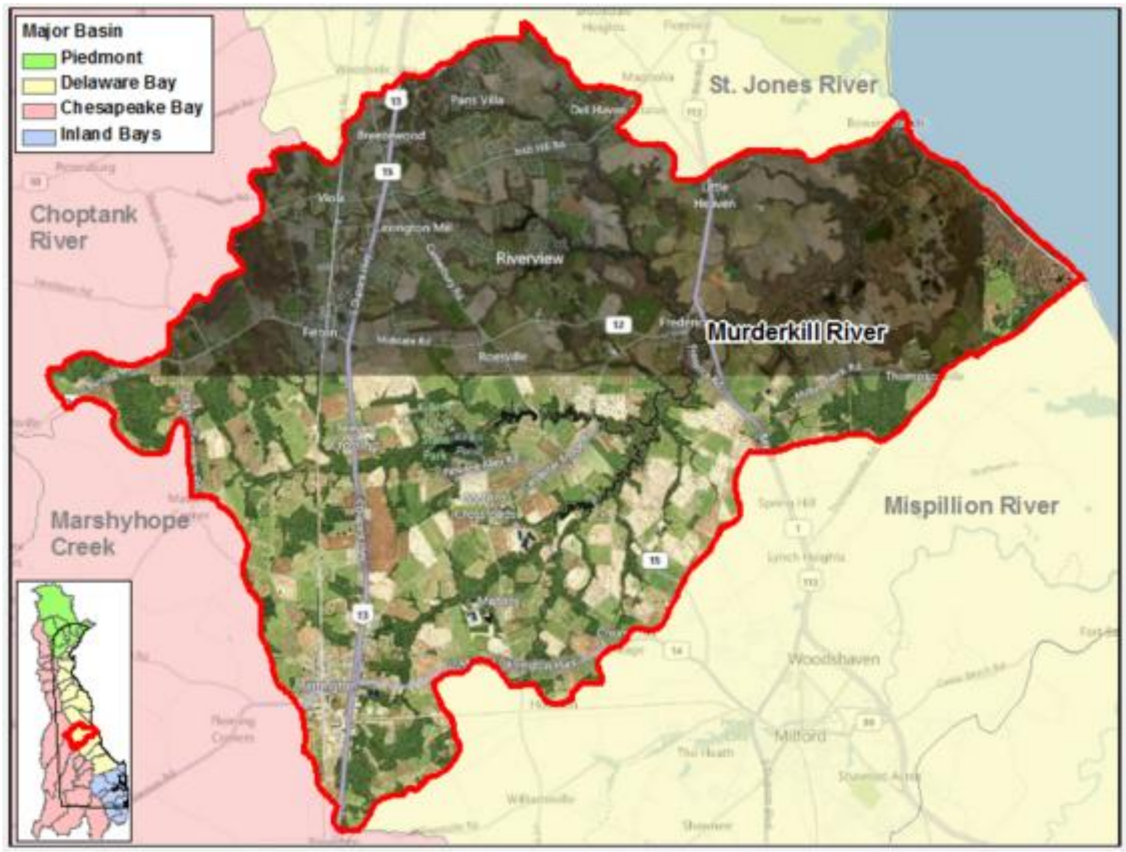


Figure 2.2 2012 Ag Census Farm Size Distribution of Kent County, Delaware, United States

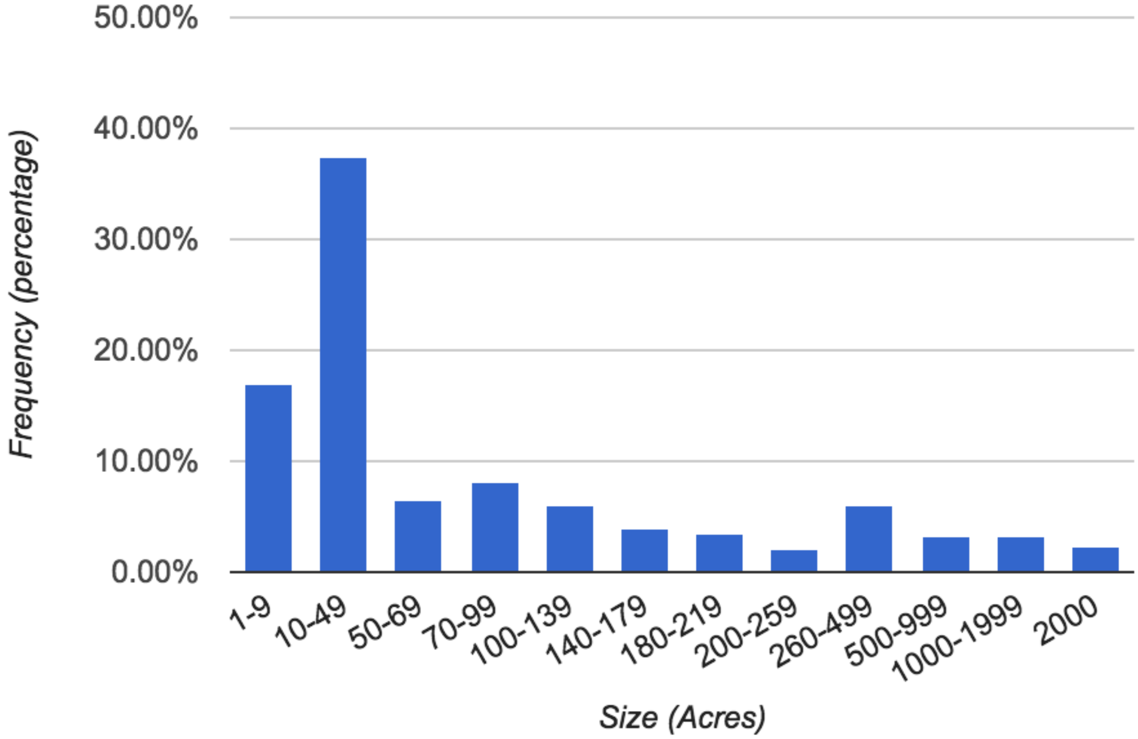


Figure 2.3 Simulated Farm Size Distribution

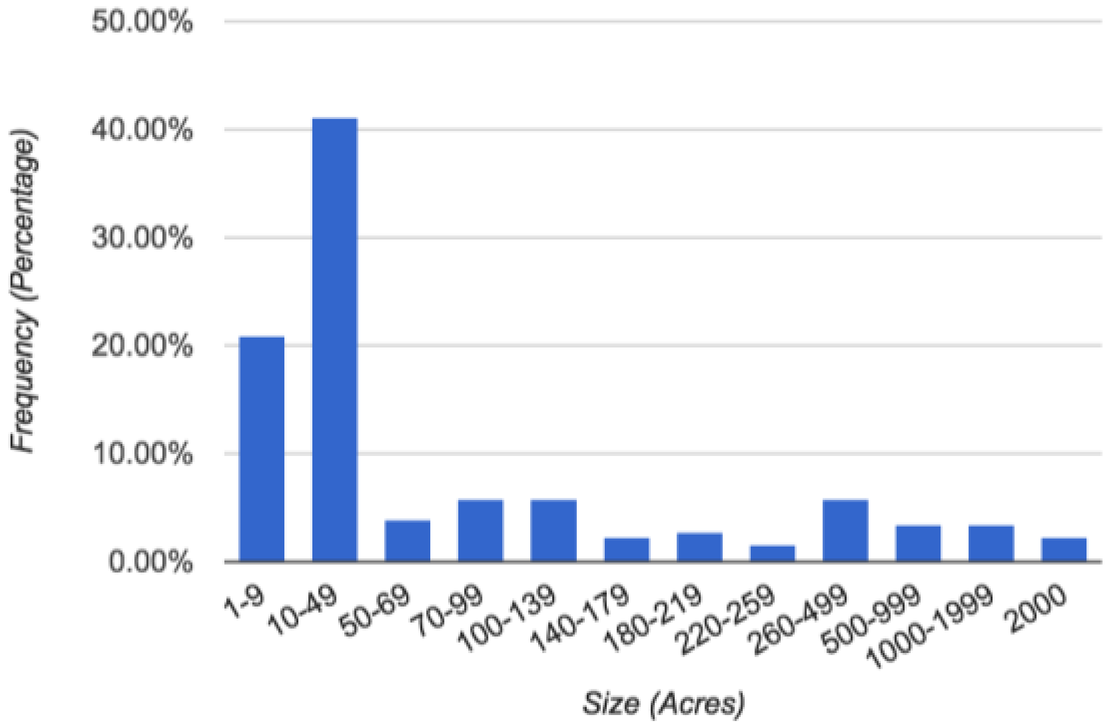


Figure 2.4 ABM Model Process Flow

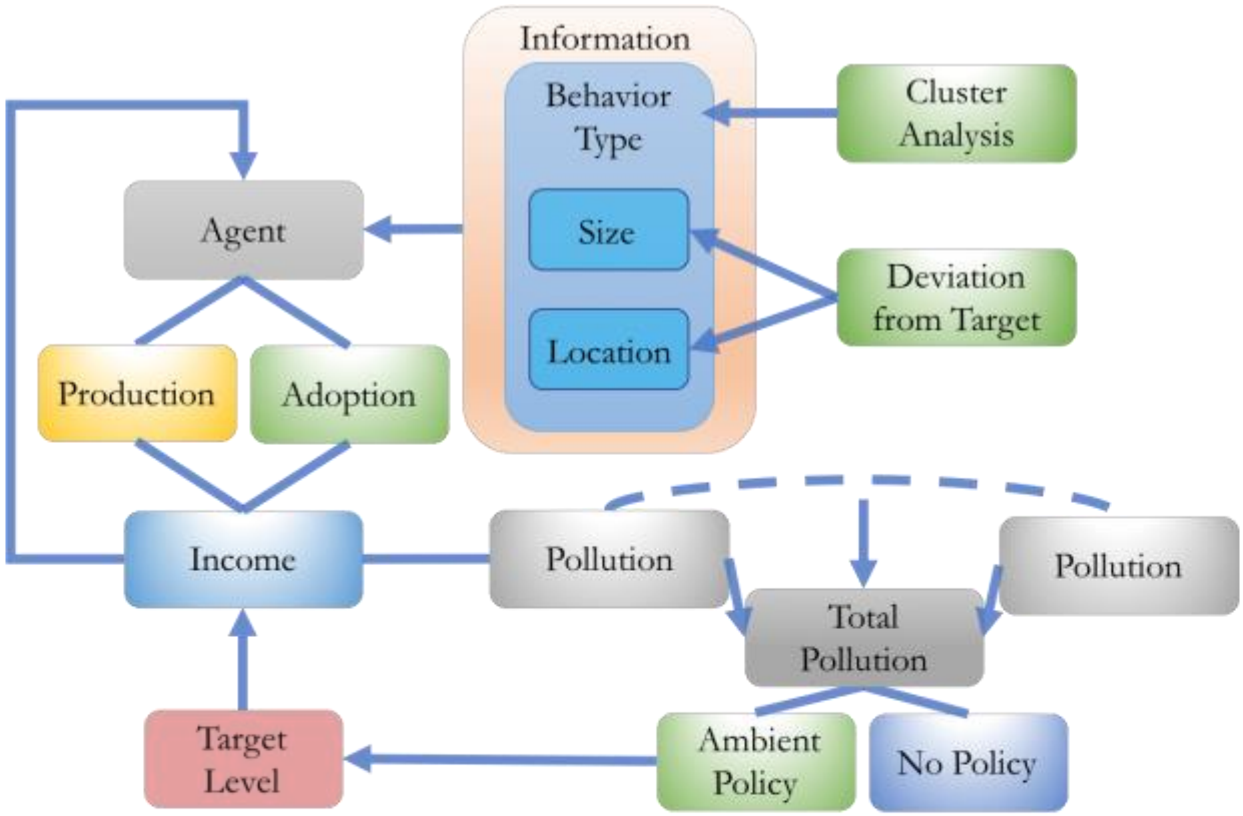


Figure 2.5 Within Groups Sum of squares versus the Number of Clusters.

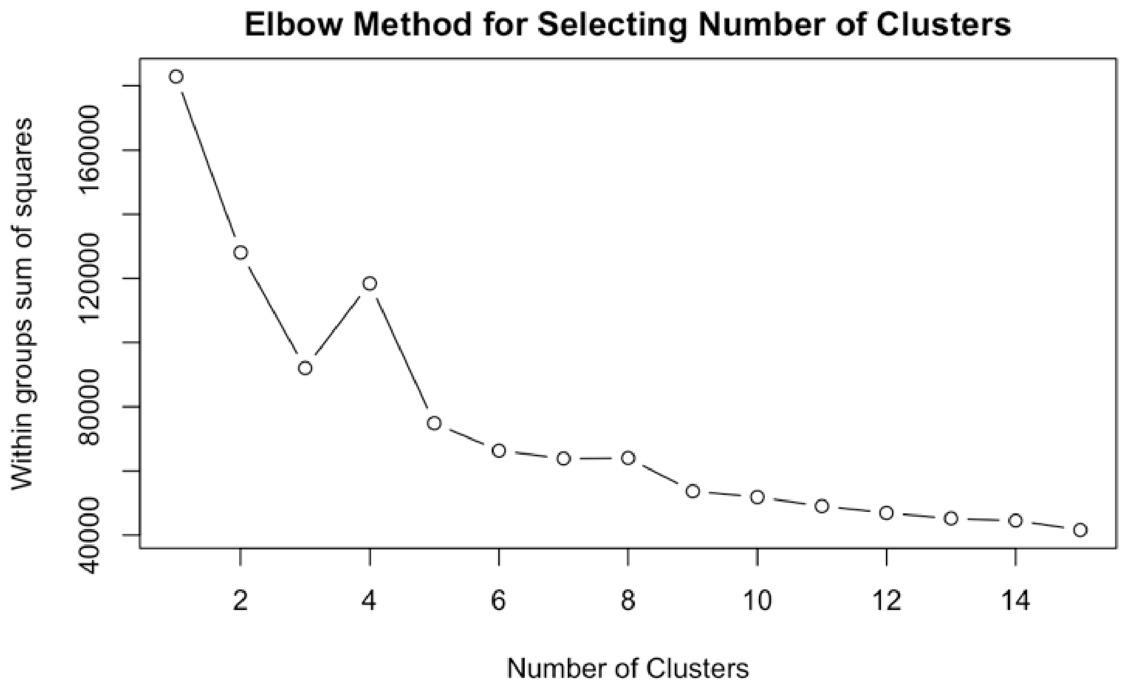


Figure 2.6 Calinski Criterion results

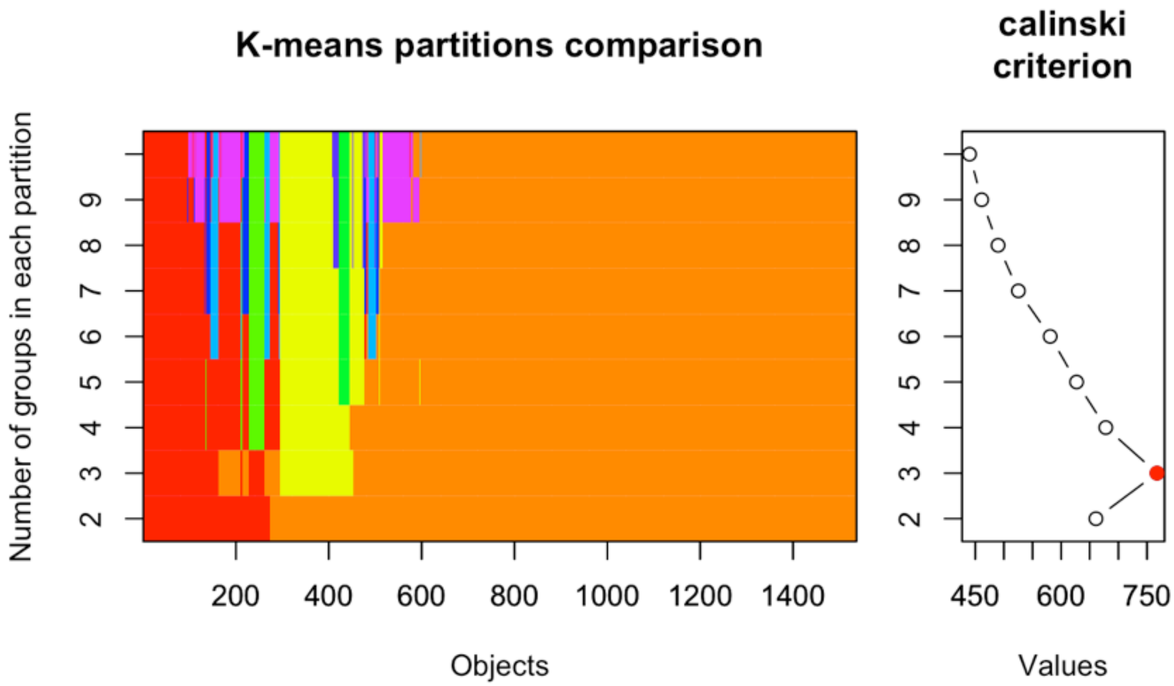


Figure 2.7 Results of using a majority rule with 26 grouping indices

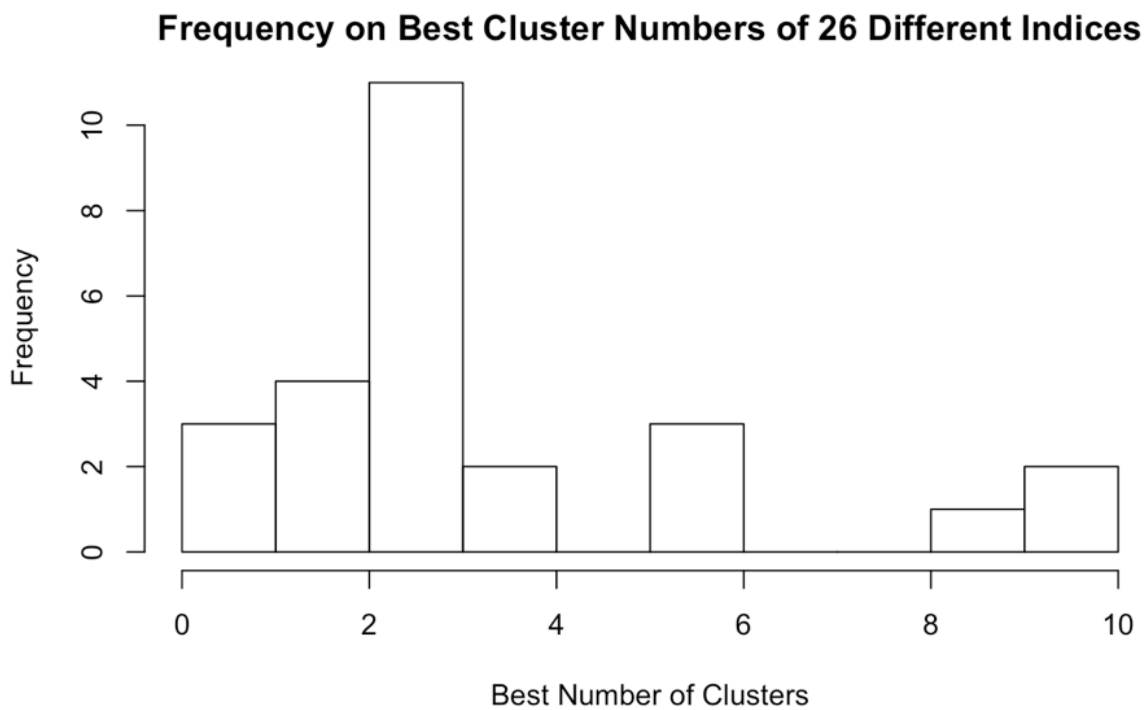


Figure 2.8 Effects of Information on Pollution Level. The red lines indicate target levels and the blue lines indicate experiment simulated results

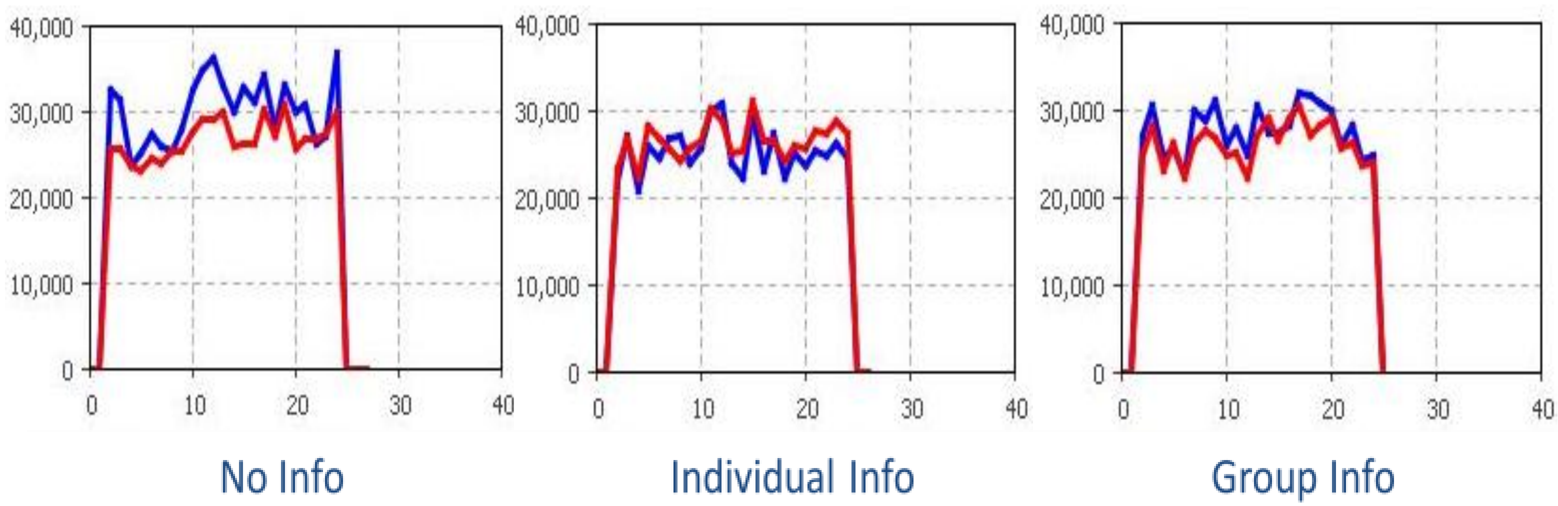
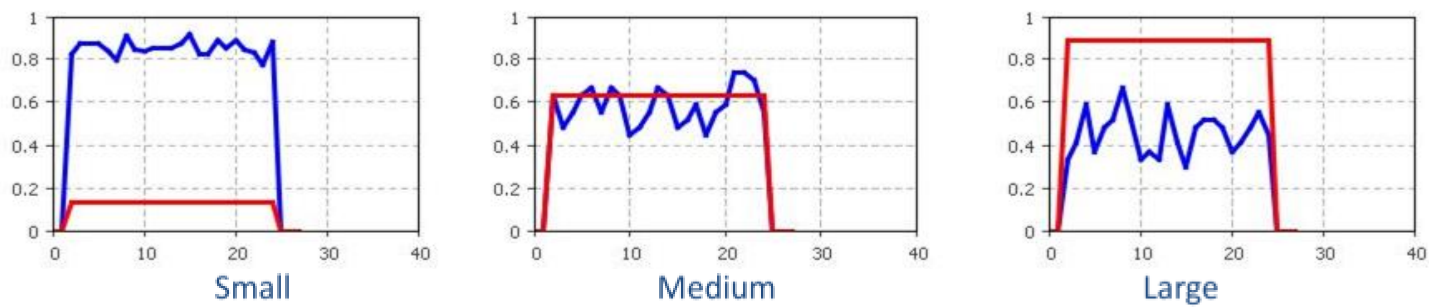
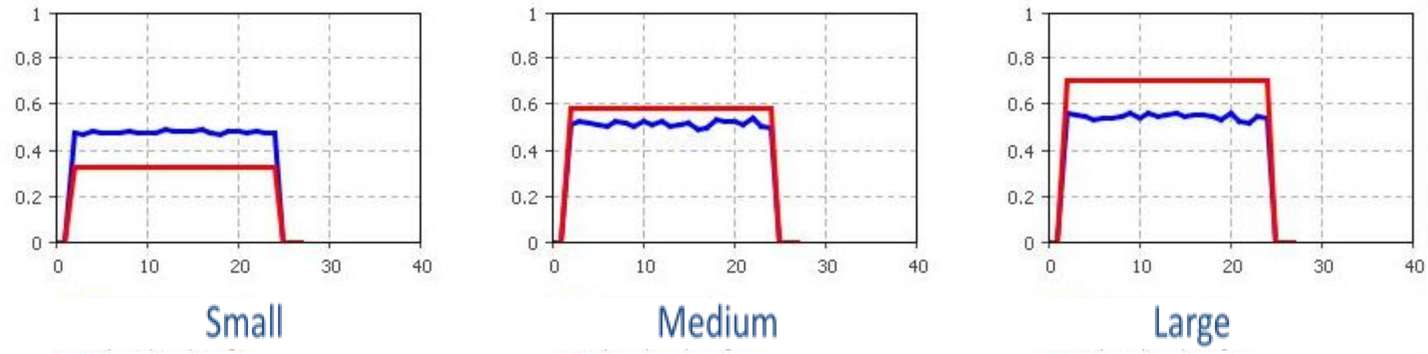


Figure 2.9 Adoption and Production Decisions by Size under No Information Treatment. The red lines indicate target levels and the blue lines indicate experiment simulated results.

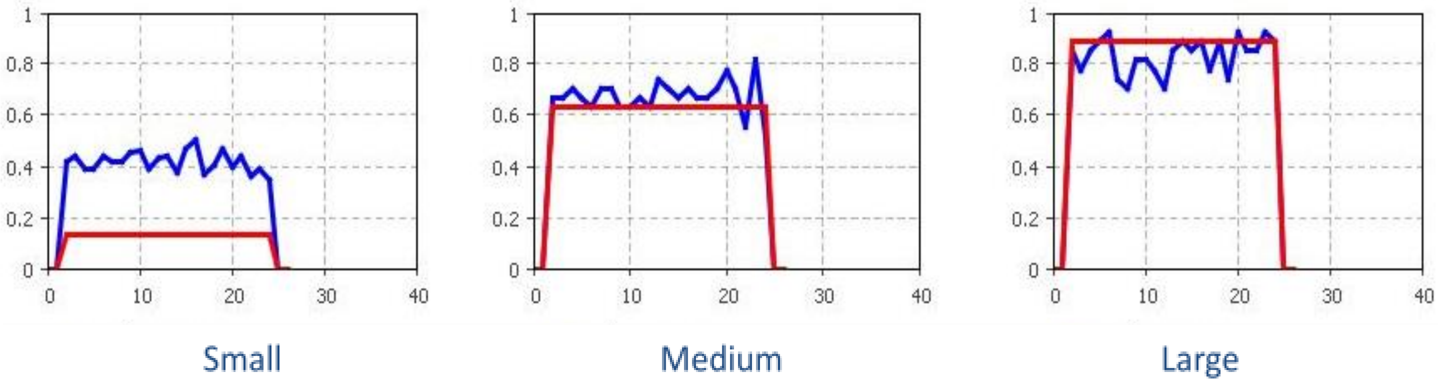


Panel a. Adoption Decision

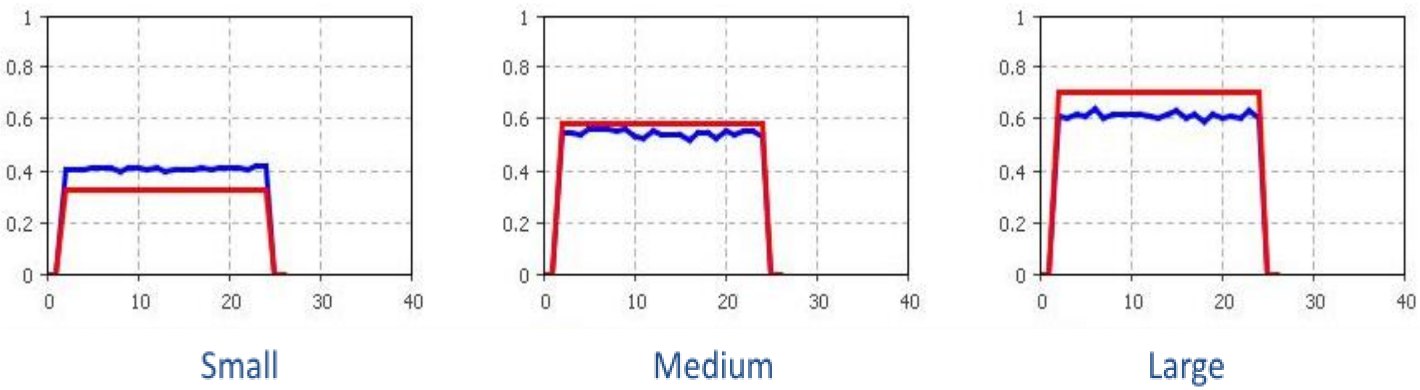


Panel b. Production Decisions

Figure 2.10 Adoption and Production Decisions by Size under Individual Level Information. The red lines indicate target levels and the blue lines indicate experiment simulated results.

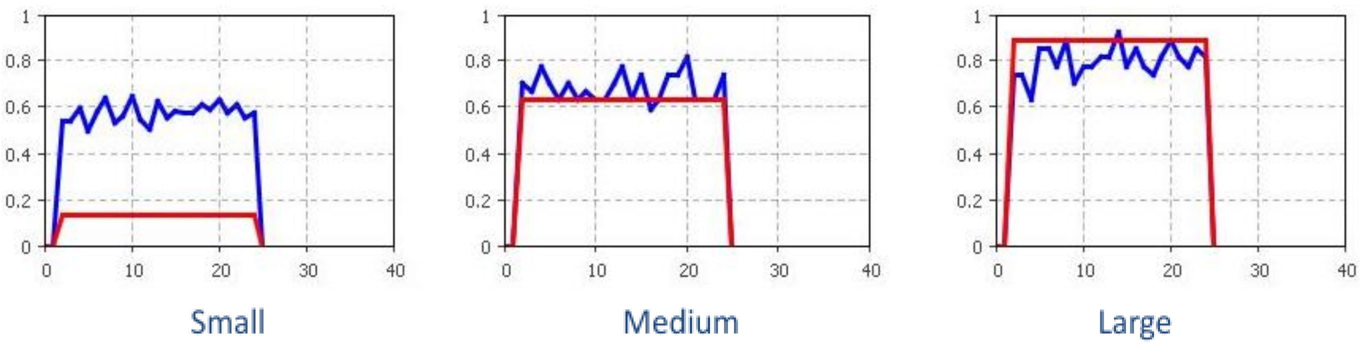


Panel a. Adoption Decisions

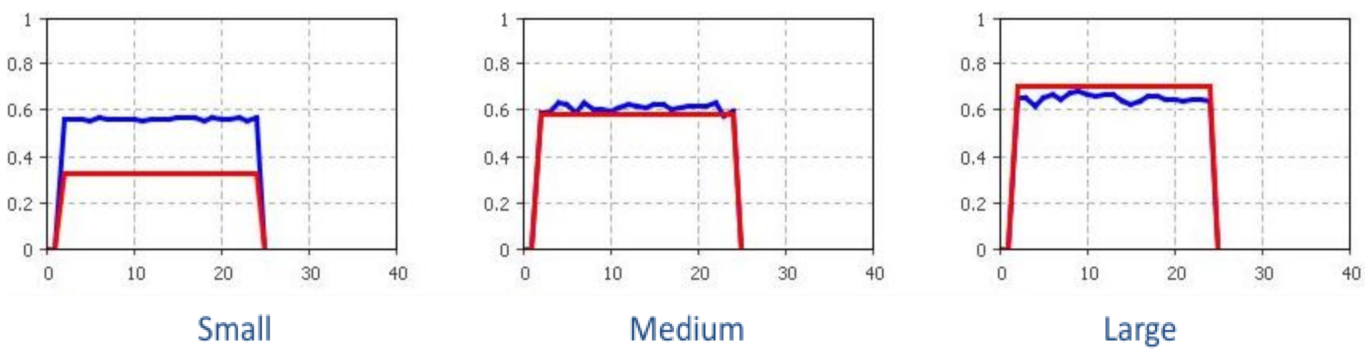


Panel b. Production Decisions

Figure 2.11 Adoption and Production Decisions by Size under Group Level Information. The red lines indicate target levels and the blue lines indicate experiment simulated results.



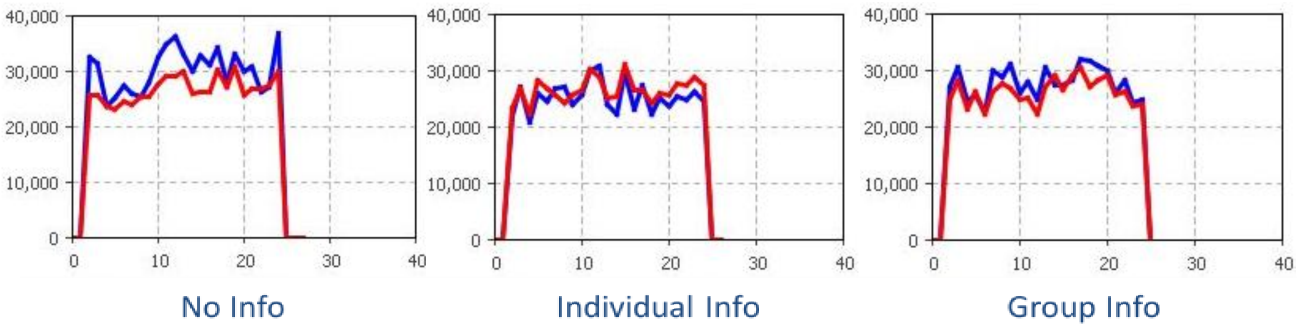
Panel a. Adoption Decisions



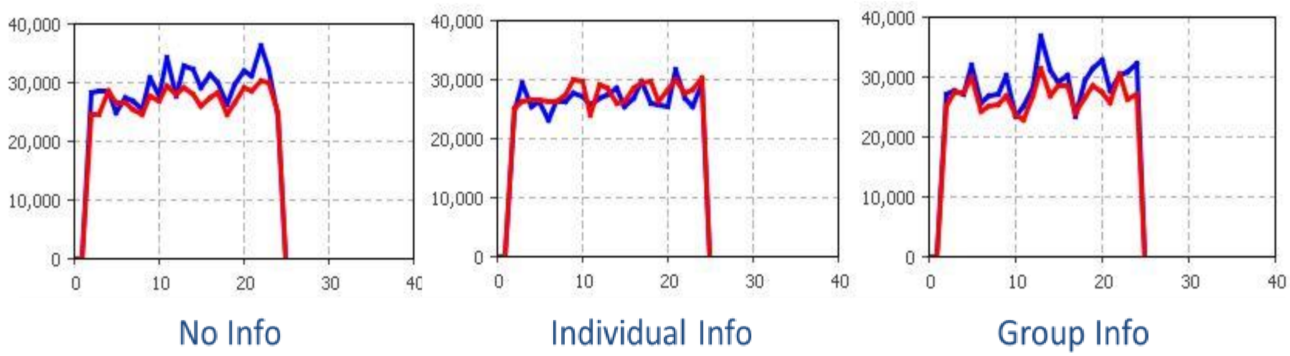
Panel b. Production Decisions

Figure 2.12 Sensitivity Test on Group Size. The red lines indicate target levels and the blue lines indicate experiment simulated results

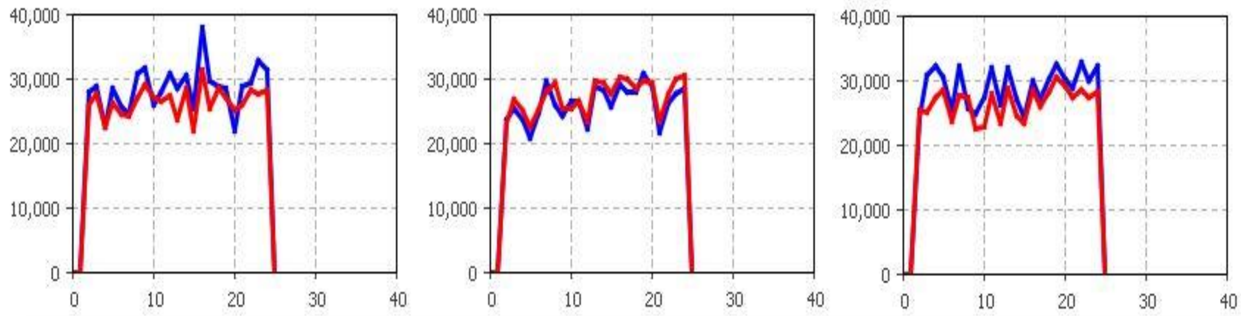
N = 5



N = 10

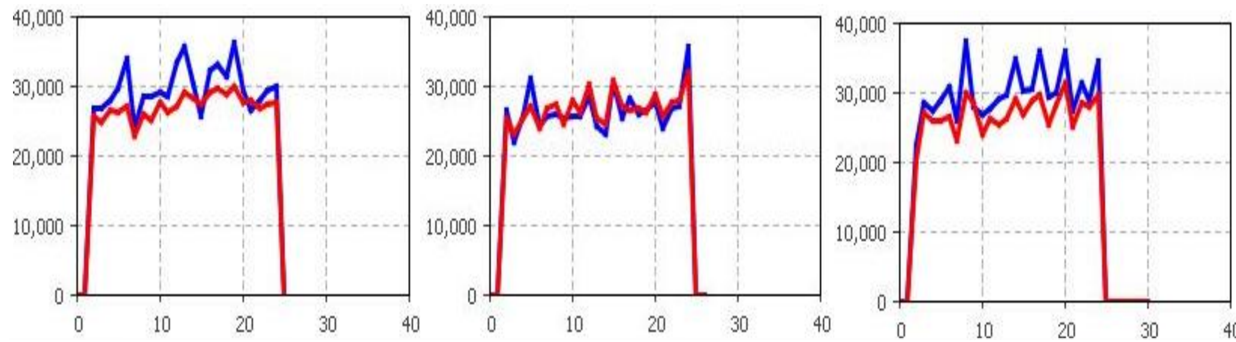


N = 15



No Info Individual Info Group Info

N = 20



No Info Individual Info Group Info

Chapter 3

AUCTIONS VERSUS POSTED PRICE IN EXPERIMENTS: COMPARISONS OF MEAN AND MARGINAL EFFECTS

3.1 Introduction

Economists frequently use auctions in experimental economics settings to measure consumers' preference for goods and services (List and Gallet, 2001; Lusk and Shogren 2007; Lusk et al., 2004). From a theoretical perspective, the bids in a well-designed and implemented auction are equivalent to consumers' true willingness to pay (WTP)—the maximum amount they would be willing to spend when faced with a purchase decision in a real market environment, since the auctions are incentive-compatible. Since a bid obtained from an auction is a point estimate of WTP, auctions are an attractive method as the data they generate is easier to work with econometrically and provide more efficient estimates than the information obtained from other methods such as yes/no decision in a posted prices format. Thus, it has become natural to emphasize auctions as a first-line valuation tool.

Implicit in the decision to use auction is that, while some error in the value elicitation process may be inevitable, the WTP estimates from these auctions have applicability to decisions made in the more common post-price markets, such as those in grocery stores or on Amazon.com, where consumers make yes/no decisions on whether or not to purchase an item at a given ‘posted’ price. However, typical consumers rarely use auctions as their primary shopping method. Even with training and practice, their decision-making in an auction setting may diverge from the daily purchasing formats they generally use. Thus, an open question is whether consumers behave in an auction the way that is consistent with how they behave in a posted price market.

Some researchers studied the estimated mean WTP disparities in experiments that use auctions versus posted prices and found that auctions in general provide lower mean WTP than posted prices for the same goods (Xie and Gao 2013). Our study contributes to the literature by verifying the existence of and further offering explanations for such WTP estimate discrepancy. Moreover, economists do not only use experiments to elicit consumer WTP for a product or service. Other important outcomes of these experiments include estimations on how WTP changes with certain product attributes, how individuals respond to different treatments, and how demographic variables contribute to WTP differences. For example, many researchers and policymakers are interested in the WTP premium for specific environmental attributes in a product, such as the location (Wu et al., 2015) and growing methods (Loureiro et al., 2003). Surprisingly, little attention has been paid to comparing the marginal effect estimates between these two elicitation mechanisms. In this research, we test the mean WTP and marginal effects of product characteristics using an artefactual field experiment. The experiment provided adult participants the opportunity to purchase different jars of honey using both a sealed-bid, second-price auction and a posted-price, dichotomous-choice mechanism. We avoid drawbacks in the existing literature with careful controls and detect the difference using both within-subject and between-subject tests.

The results suggest that estimated mean WTP in auction is smaller than the posted price mechanism WTP in the range of 32%-39%. We then seek to explain this result by testing different possible explanations. We found no evidence of anchoring on the posted price in the auction or of an asymmetric inconsistent preference effect in the dichotomous-choice setting. We did find evidence suggesting that the cause of low WTP estimates from auctions is due to some characteristic inherent to the auction setting and perhaps associated with consumers' lack of familiarity with auctions. In terms of marginal effects of different product attributes, we find that the auction and posted price mechanisms provide consistent signs, which indicates that consumer preferences for different product attributes do not vary with the elicitation methods. While the signs of coefficients are consistent, the significance level is much higher in auctions. Therefore, a posted price mechanism requires a larger sample size to detect the same preference change.

3.2 Background

Researchers and policy-makers are often interested in consumer evaluation for products or services in order to estimate values for welfare, demand elasticity, and other market information. Such information is used to set prices for new products and services and to inform policy decisions and legal proceedings. However, accurately measuring consumer preferences is not an easy task. Many techniques have been adopted to measure WTP for goods that lack an existing well-defined or easily observable market. The many variations on auctions that have been used in laboratory economic experiments are particularly appealing for this purpose since they give the researcher a great deal of control over the data being observed and allow observations of actual decisions involving real financial incentives. In essence, researchers can directly ask an individual "How much are you willing to pay for this item?" Auction methods have been generally eschewed in research on stated preferences associated with environmental valuation as poor indicators of actual WTP (Diamond and Hausman, 1994) since an auction differs from the normal price-taking setting in which consumers react to posted prices (Loomis et al. 1997). In response to such criticisms, a panel convened by the National Oceanic and Atmospheric Administration (NOAA) recommended using a dichotomous-choice format in contingent-valuation surveys (Arrow et al., 1993).[2] However, auctions have been more widely accepted in experiments for valuing private goods, because these choices are non-hypothetical.

Using posted prices in a laboratory environment should more closely mimic a market setting, such as a grocery store, since participants are price-takers. In this design, participants are asked a yes/no question: "Are you willing to purchase this item at $A?" Participants will spend $A to purchase the item if they choose "Yes," while they will not get the item nor pay anything if they choose "No." Since this framing of the

---

[2] The dichotomous choice also is referred to as a posted-price, take-it-or-leave-it, and a discrete-referendum design.

purchase question resembles decisions consumers make every day about purchasing items at different posted prices, the design is easy for participants to understand. However, a disadvantage is that the experiment does not elicit the exact WTP for each participant – instead it only indicates if WTP is above or below a certain value. Consequently, the mechanism is less statistically efficient and requires large sample sizes to produce the same level of precision as other methods (Loomis et al. 1997) such as auctions.

3.2.1 Comparisons of Posted Prices and Auctions

Approaches involving incentive-compatible auction mechanisms (e.g., Vickrey, English, Becker-DeGroot-Marschak (BDM), and random n^{th} price) are widely used in experimental economics research to elicit values for consumer WTP as they provide a point estimate of WTP for each participant (Vickrey 1961; Becker, DeGroot, and Marschak 1964; Shogren et al. 2001). An auction is considered to be theoretically incentive compatible if the dominant strategy for participants is to bid their true values. Two common auction formats are the Vickrey auction (a second-price sealed-bid auction) and the English auction.

In the context of private-value auctions, where each participant knows what the item is worth to her but is uncertain of its value to other participants, both Vickrey and English auctions are theoretically incentive compatible (Vickrey 1961). This study implements a variation of the second-price Vickrey auction that combines the ascending price feature of the English auction with the sealed bids of the Vickrey auction (Bernard

2006; Dillaway et al. 2011). According to Bernard (2006), participants appear to learn the dominant strategy faster than in a Vickrey auction, and retain it better than those in an English auction.

Economists have also adopted various posted price mechanisms in evaluation studies. For example, double-bounded dichotomous choice models are widely used to elicit consumer WTP for new technologies (Li and McCluskey, 2017), and researchers have studied how to best implement such a mechanism (e.g., Yoo and Yang, 2001). In experimental settings with real monetary incentives, a single-bounded posted price format becomes popular recently (Li et al., *forthcoming*; Venkatachalam, 2004). This is mainly because posted price choice activities easy to implement, especially in field experiments that usually take place in real market places surrounded by many distractions.

One increasing stream of literature involves the comparison of Real Choice Experiments (RCE) and auctions. In RCE, participants are presented with combinations of products at different price levels and are asked to choose the one they prefer most. Most studies on this topic found empirical WTP from RCE are significantly higher than WTP from auctions (Lusk and Schroeder 2006; Gracia et al., 2011). RCE and posted prices are similar in the way that consumers make decisions rather than submitting bids, and no point estimates of WTP can be directly estimated. The price levels presented in RCE are usually chosen from a set of price vectors that were pre-determined based on sales prices in local supermarkets or national retail prices of similar products (Lusk and Schroeder 2006; Gracia et al., 2011). While it might not be an issue for common products in standard size, it would be difficult to determine appropriate price levels for novel products with additional attributes, such as labeling and packaging. However, introducing more flexible prices in RCE comes with the risk of lower power because a portion of the results would solely be driven by the presence of a very low price. Although in the literature, posted price designs also generally pick price offers from a set of pre-determined prices (Hanemann 1984; Frykblom and Shogren 2000), it is easily extendable, and as discussed later, beneficial, to allow for more flexible price offers. As a result, posted price is especially useful in situations where credible price levels for new products or attributes are hard to obtain.

Despite the extensive literature related to RCE and auctions, fewer studies have compared relative WTP from posted price offers and auctions. Frykblom and Shogren (2000) compared a non-hypothetical dichotomous-choice question to a Vickrey auction using a market good and claimed to have eliminated two potential explanations (strategic behavior and hypothetical bias), leaving anchoring, asymmetric inconsistent preferences, and lack of familiarity with open-ended questions untested. However, the study did not actually find significant difference in resulting WTP estimates of the two methods. Besides, the experiment lacked appropriate training and practice rounds for the participants, which helps the participants to understand the dominant strategy is to bid their true value (Lusk et al., 2004). Moreover, the student participants had to enclose the entire bid in an envelope, which might lead the students to neglect the fact that they would only have to pay the second highest bid and result in underbidding in the auction. Kaas and Ruprecht (2006) proposed a model to explain that with valuation uncertainty, subject bids were lowest in a Vickrey, followed by BDM and stated preference methods. But it was actually not empirically tested since their posted price section was hypothetical. Roosen et al. (2010) explored how BDM compare with a discrete choice mechanism (BMS) that evaluates WTP by measuring the propensity of substitution between two goods and found that differences in WTP disappear when considering only engaged bidders with non-zero bids. BMS is more similar to RCE than posted price since participants are making a series of choices between two goods with different price vectors.

It is worth noting that a similar question has been discussed in the literature on operations management, especially in the context of “Buy it now” versus auction bids used on eBay. With different specifications on the cost of the auction, the reserve price, the cost to participants, and agent information, “Buy it now” and the auction yield different WTP estimates (Boyer et al., 2014; Wang et al., 2008; Wang 1993). Hammond (2010) empirically tests both auction and posted price online markets for compact discs on two internet selling platforms. A conclusion is that while auctions sell at a higher probability, posting a fixed price sells at a higher price.

3.2.2 Potential Explanations of the WTP Difference

In the existing literature, several possible reasons on what might have caused the difference in WTP estimates of auction and posted price have been investigated. These candidate explanations include the anchoring effect, the asymmetric inconsistent preferences effect, and the lack of familiarity with auction formats.

The anchoring effect (also known as starting-point bias) occurs when respondents' valuations are influenced by and biased toward the posted offer in dichotomous choice questions (Tversky and Kahneman 1974; Herriges and Shogren 1996). This anchoring effect could influence both the decisions in the posted price setting and the subsequent auction bids (Ariely et al., 2003). While Frykblom and Shogren (2000) did not observe anchoring effect in posted price decisions and Kriström (1993) observed no anchoring effect in the auction bids, Green et al. (1998) found strong evidence of anchoring on both tasks.

The asymmetric inconsistent preferences effect originates from the “yea-saying” effect in the contingent valuation literature that describes a tendency for some respondents in hypothetical choice settings to choose affirmatively in a dichotomous setting regardless of their true preferences (Couch and Keniston 1960; Ready, Buzby, and Hu 1996). Therefore, it leads to an overestimation of overall WTP in the posted price setting. For instance, Kanninen (1995) concluded that 20% of respondents in the sample were yea-sayers. Ready et al. (1996) found similar evidence with 20–22% of the respondents being yea-sayers in a split sample contingent valuation study for food safety improvements. However, as Frykblom and Shogren (2000) noted “nay-saying” has received little attention and seems to have been generally neglected in the contingent-valuation literature, while this effect would lower WTP from dichotomous choice settings. In the posted price setting with real economic incentives, it is possible that similar effects might still be present. If these effects resulted in difference in WTP estimates between posted price and auction, we could treat the auction bids as the “undisturbed preferences” and test whether the participants deviated significantly to one side from the bids. For example, one inconsistency resulting from “yea-saying” would be when the auction bid is lower than the posted price offer, but the participant accepted the posted price; the inconsistency resulting from “nay-saying” would be when the auction bid is higher than posted price offer, but the participant rejected the price. These two inconsistent preferences would cause WTP discrepancies between posted price and auction if their effects were asymmetric.

Plott (1996) in the Discovered Preference Hypothesis (DPH) casts economic decision-making as a process of discovery that assumes that participants have stable underlying preferences that are consistent with expected utility maximization. If there is appropriate feedback, decision-making converges to expected utility behavior in a series of three steps, starting with myopic “impulsive” behavior and gradually advancing to behavior that is more systematic as the decision-maker gains additional information through familiarization and feedback. As the NOAA panel pointed out, open-ended questions typically lack realism and is sensitive to trivial characteristics of the scenario presented. In contrast, dichotomous-choice questions better approximate an actual purchasing environment and are easier for respondents to answer accurately Arrow et al. (1993). Although one cannot claim that either posted price or auction reveals the “correct” WTP, posted price is obviously the format that is more familiar, easier to understand and similar to a real-world purchasing decision. Familiarity with auctions is a form of institutional information and choice framing, and many consumers may not be familiar with auction formats because they do not routinely participate in any form of auction. In that case, we would expect to see an experience effect as an auction's rounds progress.

3.2.3 Contribution to the Literature

Our study contributes to the literature in several ways. We carefully design an experiment to compare the homegrown-value WTP estimates between an auction and a posted price elicitation format. In addition, we examine possible explanations for such WTP estimate discrepancy. Few studies have compared important findings generated by auctions versus posted price mechanisms other than the mean WTP. However, auctions are not mainly used to measure average WTP for products. Rather, they are often adopted to measure relative WTP for product attributes, information and policy treatment effects, and heterogeneous demographic responses. Therefore, we further extend the research question to comparing the sign and statistical significance of coefficient estimates.

First, we carefully design an experiment that avoids many drawbacks of existing ones in the literature. We only use experienced shoppers as experiment participants since it has been shown that experience with the good can reduce market anomalies List (2003). Compared to the literature, our experiment includes more extensive training, including written instructions, oral presentations and two training rounds to give participants better understanding on their tasks. In a setting with unfamiliar tasks, extensive training is necessary because even if subjects are told it is in their best interest to bid their "true value," subtle misconceptions about how the elicitation mechanism works might trigger subjects to default to the strategies associated with familiar auctions (Plott and Zeiler, 2005). Moreover, we argue that for our purposes, running an experiment in a more controlled environment in terms of information and feedback introduces less noise into participants' decision making process compared to a field setting (Plott 1996). Second, we test if discrepancies exist using both a within-subject and a between-subject design. Compared to the literature where only one kind of comparison is used, combining both within- and between-subject design adds robustness to our results. Third, we introduce more flexible price vectors into the posted price section, since prices vary randomly for each posted price question, we control for the possibility that consumers treat the price offers as quality signal and therefore alleviates valuation being anchored to the price offers. Using flexible price vectors also avoids picking inappropriate price offers in the situation where it is difficult to form fixed price points or appropriate widths between each price point. Fourth, we explicitly test for several possible explanations for the discrepancy and provide our own explanation. Lastly, our participants made choices on otherwise homogeneous honey with different shapes of jars. Therefore, it allows us to easily compare how individuals respond to each jar under both mechanisms. We run regressions based on models commonly used in the literature to examine the signs and statistical significance of the coefficients.

3.3 Experimental Design

We design a homegrown-value artefactual field experiment in which we offered adult subjects the opportunity to purchase honey presented in a variety of jars. This research was conducted in an experimental economics laboratory at a large university in the Northeastern United States. We recruited 115 adult participants through various sources that included the university's online newspaper, local community meetings, emails to staff members, and the laboratory's website. We endeavored to recruit adult consumers rather than students so that the sample would better represent the community as a whole and to ensure that participants were experienced buyers (Gracia et al., 2011; Chang et al., 2009; List 2003).

Table 3.1 describes the socio-demographic characteristics of the participants. The average participant age was about 42 years. Most of the participants were female and most of the participants were primary shoppers in their households. Average household income was between \$70,000 and \$80,000 and the average number of years of education was 16. The relatively high education level and income among participants likely reflects the population of a university town.

Fifteen one-hour sessions were held with participants receiving \$20 in cash and/or products for the session (\$5 show up fee and \$15 to be spent during the experiment). Participants were informed that they could keep any portion of the money that they did not spend and that they would be given the opportunity to purchase a jar of honey during the session. Participants received the money and products purchased at the end of the session.

At the beginning of the experiment, the administrator randomly assigned the participants to computer terminals equipped with privacy screens to ensure confidentiality. Participants were asked to read information about the experiments once they were seated (see Review Appendix). A presentation then was given to explain the steps involved and how to use the program. No communication among participants was permitted, but participants were welcomed to ask questions to the administrator at any time. Data was collected through the use of Excel files that were programmed with Visual Basic with Applications and stored in an Access database.

The experiment involved investigating the effects of labeling and packaging on consumers' WTP for honey products. Specifically, we tested WTP for honey of three origins (local, domestic, and international) that were each distributed to five types of jars that had different shapes but the same volume (12 ounces), making fifteen jar/origin combinations. In the auction, participants bid on all fifteen honey products. In the posted-price rounds, they made purchase decisions for the five jars of U.S. honey only. Therefore, each participant made twenty honey-purchasing decisions in total. In this paper, we limit our comparison of WTP estimates to purchases of U.S. honey because it is most commonly sold in grocery stores and is most familiar to the general public. A set of labeled jars (Jar 1, Jar 2,…, Jar 5) of honey was placed on the administrator's desk and on the desk of each participant throughout the experiment, and participants were encouraged to examine the appearance of the jars, but not open the jars. Since the three types of honey (U.S., international, local) were indistinguishable in terms of appearance, we just displayed the U.S. honey due to desk space constraints. The sequence of the posted-price experiment and the auction experiment was randomly determined before the session, and the order in which the products were presented was also randomized.

To address the concern of demand reduction, at the end of each session, only one of the twenty decisions made by participants (fifteen in the auction and five in the posted prices) was selected at random to determine which product would be binding and used to calculate cash earnings (Lusk et al., 2004; List and Lucking-Reiley 2000; Messer et al. 2010). This binding decision was selected by having a volunteer draw a labeled ball from a cage containing twenty balls, each representing one decision. In order to reinforce the understanding of this concept, demonstrations of how the ball would be drawn to determine the binding round were shown to participants prior to them making any decisions. It was also emphasized that no decision was affected by prior or subsequent decisions. As explaining the dominant strategy to participants in homegrown-value experiments is regarded as "best practice" and is widely used, we also informed the participants that it was in their best interest to bid as close to the worth of the item to them as possible (Rutström 1998; Lusk et al., 2004).

In the posted-price experiment, the question to participants was "Are you willing to purchase Jar Y of U.S. honey at $A?" The price of each product varied randomly for each decision and was distributed uniformly between $0 and $15. The effects of pre-set range have been discussed in the payment card literature and it has been shown that using different range and center would not lead to a bias as long as the upper limit is sufficiently high (Rowe, Schulze and Breffle, 1996). Participants were informed that clicking "yes" was a decision to purchase the jar of honey at the posted price; clicking "no" meant they would not receive Jar Y nor pay the price.

In the second-price auction, a number representing the participant's bid for the item was shown on the screen in front of each participant. Once the auction started, this bid increased incrementally at a speed about $0.10 per second from $0 to $15.[3] Participants were asked to click the "withdraw from auction" button when they saw the bid representing the maximum amount they were willing to pay for the product displayed on the screen. When they indicated a desire to withdraw from the auction, a second box appeared that asked them to confirm the number on their screen as their bid. Participants could choose to restart the auction round (incremental ascending increases in the number) from $0 and bid again or could confirm the bid and submit it. The auction stopped either when all participants' bids were confirmed or when the bid reached the pre-set upper limit of $15. The bids by each participant were stored in a database and the auction then proceeded to a new bidding decision.

To help participants better understand the bidding procedure, two practice rounds were held first. Participants were given $3 in the practice rounds and were asked to submit bids on a pencil and a ballpoint pen. In the practice auction, the winner and the second highest bidder were announced after each round. It was emphasized to participants that the winner pays only the amount of the second highest bid so it was in

[3] Since participants started the program by themselves, the participants' bids were not synchronized making it impossible for other participants to know whether they stopped the program on a low or high bid.

their best interest to focus on determining their own value for the item and to bid as closely to that as possible.

After the practice rounds, participants were asked to submit bids on different jars of honey following the same procedure with an initial balance of $15. This research followed the proposed “best practice” in Harrison et al. (2004) to clearly train and inform the subjects that their dominant strategy is to bid their true value. At the beginning of each new purchasing decision, participants were provided with the list of items already auctioned and bids they submitted for each. After each decision, no feedback was given to participants with regard to the winner or the winning price as a means of reducing market feedback (Corrigan et al. 2011). At the end of the session, participants were asked to fill out a survey about their demographics background and consumer behavior.

The only announcement was the winner of the binding round at the end of the experiment. This was done by having a volunteer draw a ball to determine which of the twenty purchase decisions was binding. Each participant’s screen then displayed a chart showing their decisions and products. Based on this binding decision, the computer program calculated each participant’s earnings and products purchased (if any) and displayed them on that person’s screen to assist them in filling out receipts.

3.4 Model and Testable Hypotheses

In this section, we describe the model and the hypotheses that we will be testing in the experiment. We proceed by first verifying if WTP estimate difference exists between the two value elicitation methods. We then test if the observed difference (if any) is a result of the inter influence between the posted price and the auction parts. Next, we examine two other behavioral factors that may result in WTP estimate differences. We conclude by comparing the marginal effects offered by the two elicitation methods.

3.4.1 Comparison of Posted Price and Auction

The series of hypotheses tested in this research are summarized in Table 3.2. The first hypothesis is that the WTP estimates from the posted-price mechanism equal to those from the second-price auction.

H_0: $WTP_{Posted\_Price} = WTP_{Auction}$

Where WTP_{Posted_Price} denotes willingness-to-pay estimates obtained from the posted price questions and $WTP_{Auction}$ denotes willingness-to-pay estimates derived from the experimental auctions.

The posted price generates binary responses while the auction generates continuous bids. To make the two types of data comparable, the auction data may be transformed to simulated binary responses, or average WTP point estimates can be inferred from posted price responses. For consistency with the literature, we follow the procedure documented in Green et al. (1998) and Frykblom and Shogren (2000) where the auction data is transformed into synthetic binary responses and compared to the actual responses. Let b_{ij} denote the bid that participant i submit for good j, and p_{ij} denote the posted price offer of participant i for good j, and δ_{ij} denote whether participant i responded yes in the posted price section for good j. Since each participant responded in both the auction and posted price formats, we can compare their auction responses, b_{ij} to the binary response that would be consistent with the prices they see in posted price section, p_{ij}, for the same good. We generate a synthetic dichotomous choice response variable δ'_{ij}, where $\delta'_{ij} = 1$ if $b_{ij} \geq p_{ij}$; $\delta'_{ij} = 0$ if $b_{ij} < p_{ij}$. Here, δ'_{ij} can be interpreted as when facing the price offers, what participants' response would be based on the bids they indicated. Theoretically, if the null hypothesis holds, we would not observe a significant difference in the WTP inferred from δ and δ'.

To test if δ and δ' significantly differ from each other, we perform both parametric and non-parametric tests. The advantage of a non-parametric test is that no distributional assumption is placed on the variables. Since δ and δ' are binary variables and since these are considered as paired observations, we use McNemar's non-parametric test (McNemar 1947).

Since non-parametric tests generally have lower power than parametric tests, we also do a parametric test assuming a normal/logistic distribution on the underlying WTP (Frykblom and Shogren 2000). Formally, we assume a consumer's WTP, w, follows some probability distribution with μ as the location parameter and σ as the scale parameter. We denote F as the cumulative distribution function and S as the survival (or duration) function. Therefore, for a given posted price offer p, $F(p) = Prob(w \leq p)$, $S(p) = 1 - F(p) = Prob(w > p)$, $f(p) = dF(p)/dp$. So the survival function, in this case, represents the probability that a "yes" response in the posted price format will continue above a given price. We estimate μ and σ by maximizing the log-likelihood function L, which is written as:

$$L = \sum (\delta_{ij} \log \left(S(p_{ij})\right) + \left(1 - \delta_{ij}\right) \log \left(F(p_{ij})\right))$$

where δ_{ij} is equal to 1 if participant *i* accepted posted price offer for the *j*-th object (p_{ij}), and equal to 0 otherwise. The survival function for normal distribution is:

$$S(p_{ij}) = 1 - F(p_{ij}) = 1 - \Phi\left(\frac{p_{ij} - \mu}{\sigma}\right).$$

For logistic distribution, the corresponding function is:

$$S(p_{ij}) = 1 - F(p_{ij}) = \frac{1}{1 + \exp\left(\frac{p_{ij} - \mu}{\sigma}\right)}.$$

The estimated mean for both distributions are μ, the estimated variance for normal distribution is σ^2, while for the logistic distribution it is $\sigma^2\pi^2/3$.

From this maximum likelihood estimation, we would be able to infer the distribution of WTP that generated the posted price responses. To test if the estimated mean from the two samples are different, we follow the same test as Frykblom and Shogren (2000), which is recommended by Kmenta (1986):

$$Z_{\overline{w}_1 - \overline{w}_2} = (\overline{w}_1 - \overline{w}_2) / \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}$$

where Z is an approximately standard normal variable, w_k is the estimated mean in offer format k, s_i^2 is the estimated sample variance, and n_i is the sample size.

3.4.2 Possible Task Inter Influences

Since our experiment consists of different tasks within a subject, we address the most common problem for a within-subject design—the tasks potentially influencing each other. We do this in two ways: first we test if the difference still exists if we only utilize the first task that each participant completed; second we test for anchoring effect to see if the bids in the auction are influenced by posted price offers that were presented to the participant.

3.4.2.1 Testing for a Difference Using First Task Only

The intuition is to make comparisons only from data of the first task that a participant did. Specifically, we estimate WTP only from participants who went through the auction first and compare to posted price WTP estimates from participants who did the posted price first. In this way, we are actually making a between-subject comparison. The procedure used for this test is similar to the one described earlier where we transfer auction data into yes/no responses and compare it to the posted price data. One issue in generating the synthetic yes/no responses is that there does not exist a corresponding relationship between the auction-first group's bids and the posted price-first group's price offers. Therefore, we use a complete combinatorial approach similar to the one suggested in Poe, Giraud and Loomis (2005). For every auction bid (suppose n_1 total observations), we generate a yes/no response according to every posted price offering (suppose n_2 total observations), resulting in n_1*n_2 pairs of observations on bids (b), synthetic yes/no (δ'), price offer (p) and real yes/no response (δ). Next, we compare δ to δ' following the procedure discussed before.

3.4.2.2 Testing for Anchoring Effect

We randomized the posted price offered for each decision in the experiment to control for possible anchoring of participants' valuation of each item to the posted price. However, posted price offers might still be affecting consumers' value formation process in two ways. First, the WTP estimates from posted price could increase if the participant saw a higher posted price offer for the item (Frykblom and Shogren 2000). Second, the WTP estimates from the auction could be affected by the posted price offers if the subject participated in posted price first (Kriström 1993). We assume that if the underlying valuation of the product is changed by the posted price offer, it is likely reflected in both the posted price part and the auction part, meaning that the presence of the two presentations of anchoring effects are positively related. The design allows us to test for the second type of anchoring effect by a Tobit model that includes posted price offers as an independent variable. Since bids were limited to a range of \$0 to \$15, a two-limit random-effects Tobit model was appropriate to analyze WTP.[4] The dependent variable is defined based on a latent variable y_{ijk}^{*} that cannot always be observed and is specified as

$$y_{ij} = \begin{cases} y_{ij}^{*} & if \quad 0 < y_{ij}^{*} < 15 \\ 0 & if \quad y_{ij}^{*} \leq 0 \\ 15 & if \quad y_{ij}^{*} \geq 15 \end{cases} .$$

4 An OLS model without censoring gives very similar estimates.

For subject *i* and item *j*, y^*_{ij} is limited to a value between 0 and 15 and depends linearly on X_{ij} via a parameter (vector), β. The following random-effects Tobit model was used:

$$y^*_{ij} = \alpha + \beta X_{ij} + U_i + u_{ij} = \alpha + \beta_1 Jar\ type\ 2_{ij} +$$

$$\beta_2 Jar\ type\ 3_{ij} + \beta_3 Jar\ type\ 4_{ij} + \beta_4 Jar\ type\ 5_{ij} + U_i + u_{ij}$$

where α is the average bid for the entire population, U_i represents the individual random effects, and u_{ij} is the error term for individual *i* for product *j*. We also include a specification with bootstrap standard errors. The variables *Jar type 2* through *Jar type 5* are dummy variables indicating which item was auctioned. The variable *Jar order* is a dummy that included to control for order effects. We define a variable *auction_first* equals one when the posted price treatment follows the auction and equals zero otherwise.

Under the null hypothesis that the anchoring effect is present, we would expect that when the posted price section is before auction (*Auction_first=0*), the effect of parameter of *b* on *p* would be significantly different from 0, we denote this as: $\beta_{p,\ Auction\_first=0} \neq 0$. Meanwhile, it is expected that when the posted-price section is after auction (*Auction_first=1*), such an effect should not be observed ($\beta_{p,\ Auction\_first=1} = 0$). This hypothesis is listed as Hypothesis 2 in Table 3.1. In summary, we test:

H_0: $\beta_{p,\ Auction\_first=0} \neq 0$

H_0: $\beta_{p,\ Auction\_first=1} = 0$.

3.4.3 Testing for Behavioral Factors

After testing for inter influences between the tasks, we investigate behavioral factors that may result in WTP estimate differences between the two methods. As explained previously, asymmetric inconsistent preferences and the fact that participants are more unfamiliar with auctions may both lead to discrepancies in WTP estimates.

3.4.3.1 Asymmetric Inconsistent Preferences Hypothesis

If asymmetric inconsistent preferences were the cause of the WTP discrepancy, we would expect to observe a difference in the following two inconsistencies: 1) when the bid is higher than the price and 2) when the bid is lower than the price. When a participant answers "yes" to a dichotomous choice question even though the price is higher than their bid, we define it as "affirmative inconsistent preference". In contrast, when a participant answers "no" to a dichotomous choice question even when the price is lower than their bid, we define it as "negative inconsistent preference". Affirmative inconsistent preference can be denoted as: WTP in posted price offer (p) > bid in auction (b), meaning when $\delta' = 0$, $\delta = 1$. Negative inconsistent preference can be denoted as: WTP in posted price offer (p) < bid in auction (b), meaning when $\delta' = 1$, $\delta = 0$. If the inconsistent preferences cause the WTP discrepancies, we would expect one inconsistency would be more prevalent than the other. We test whether the probability of a affirmative inconsistent preference is larger than the probability of a negative inconsistent preference—specifically, whether $\Pr(\delta = 1 \mid \delta' = 0) > \Pr(\delta = 0 \mid \delta' = 1)$. If this hypothesis is rejected, it means participants are not more likely to have affirmative inconsistent behavior than negative inconsistent behavior, and asymmetric inconsistent preference does not explain any discrepancy.

3.4.3.2 Familiarity Hypothesis

As compared to answering a posted price question, auction is a mechanism that is relatively unfamiliar to most participants. Even if participants do not receive direct feedback after each round, all information available to a participant may evolve due to additional opportunities for introspection, belief reinforcement, learning, and similar mechanisms. In that case, we would expect to see an experience effect as an auction's rounds progress. We test if *roundnumber* (the number of bidding decisions a participant has made) has an effect on the bids. Under the null, $\beta_{\text{Auction, RoundNumber}}$ would be significantly different from 0. Specifically:

H_0: $\beta_{\text{Auction, RoundNumber}} = 0$

H_1: $\beta_{\text{Auction, RoundNumber}} \neq 0$.

3.4.4 Comparing Marginal Effects in the Two Methods

Despite any WTP estimate differences that may exist and the reasons that may cause the differences, in practical research, we are often not only interested in the absolute WTP estimates of a homegrown good, but also care about the marginal effects, or the ability that the estimation method is able to provide relative comparison conclusions on the effects of some particular attributes. When the research question is not about estimating absolute WTP values but instead about testing marginal effects of attributes, it is important to learn if the two mechanisms provide similar results. We compare the marginal effect estimates on jar attributes from the two parts in terms of the signs and significance levels of the coefficients.

3.5 Results

3.5.1 Descriptive Statistics on Bids and Yes/No choices

A histogram on the frequency distribution of the bids is displayed in Figure 3.1. As expected, the number of bids into each price category decreases as the price increases. The mean of the bids is $2.91 and the standard deviation is 1.97. Figure 3.2 shows the percentage distributions of posted prices conditioning on whether the posted price was accepted or declined. As expected, number of acceptance decreases as prices go up, and the number of declines increase as price increase. In general, we do not observe fat tails in the distributions

In total there were 45 zero bids in the auction. Out of the 115 participants, 4 people (3.5% of the total participants) bid zero for all five auctions of honey. This seems to be a reasonable proportion of people who would not be interested in purchasing honey at any price. Of these four participants, three also declined the honey in all the posted price questions. So their behavior appears to be generally consistent.

3.5.2 Hypothesis 1: Test for WTP Difference, H0: WTPPosted_Price = WTPAuction

As shown in Panel A of Table 3.3, for a within-subject comparison, the average of the actual binary response (δ) in posted price is 0.2904; the average of the generated synthetic binary response (δ') based on bids in the auction and posted price offer is

0.1652. Since McNemar's chi-squared test statistic equals 150.45 and the corresponding p value is less than 0.0001, we reject the null hypothesis.

As discussed earlier, we also do parametric tests assuming the underlying WTP distribution is either normal or logistic (Panel B of Table 3.3). With a normal distribution, the estimated WTP from auction bids has a mean of 2.4889, while the estimated mean of WTP from posted price is 4.0587. A Z test rejects the null hypothesis that the two WTP means are equal. With a logistic distribution assumption, results are similar. The estimated mean of WTP is 2.4579 for auction bids and 4.0570 for posted price. The Z test also rejects the null at 1% level. The results suggest that WTP estimate from the auction is approximately 39% lower than that from posted price.

3.5.3 Test for Task Inter Influences

To address potential concerns that a within-subject design involving two tasks might influence each other, we test for the discrepancy using first task only and then test for anchoring effects between the tasks.

Hypothesis 2.1: Test for Discrepancy using First Tasks Only, H_0: $WTP_{Posted\_Price|Posted\_Price\_First} = WTP_{Auction|Auction\_First}$

We conducted a between-subject comparison using only data from the first task each participant completes. In other words, we generate WTP estimates for auction from participants who did auction first, and generate WTP estimates for posted price from those participants who did posted price first. Since there is no one-to-one corresponding relationship between the bids and posted price offers, we do a complete combinatorial procedure (Poe, Giraud and Loomis, 2005) on bids and price offers to generate a synthetic binary response (δ^') and compare it to the corresponding actual binary response (δ). Again a McNemar's test rejects the null that the probability of accepting is equal.

In a similar fashion, we conducted parametric tests assuming either normal or logistic distribution on the underlying WTP. Under normal distribution assumption, the estimated mean WTP is 3.171 for auction and 4.647 for posted price. Under the assumption of logistic distribution, the estimated mean WTP is 3.137 for auction and 4.603 for posted price. In both cases, Z test rejects the null that the two estimated means are equal. This indicates that estimated mean WTP from auction is approximately 32% lower than that from posted price.

Hypothesis 2.2: Anchoring Effect

Anchoring effect might happen if participants perceive the posted price as a quality signal of the product and therefore anchor their value of the product to the price offer. We perform a test similar to Kriström (1993) where we examine if the respondents' auction bids are anchored to the posted-price offers when they participated in the posted-price setting first. Meanwhile, their bids should not be affected if the auction was held first. To test H_0: $\beta_{p,\text{ Auction\_first=0}} \neq 0$, we regressed the bids from sessions in which the posted-price mechanism was conducted first (Table 3.5). The left panel of Table 3.5 reports a random effects Tobit model, while the right panel reports the same model with bootstrapped standard errors included. As shown in both panels of Table 3.5, the posted price in the posted-price mechanism did not affect subsequent bids in the auction. Therefore, H_0: $\beta_{p,\ Auction\_first=0} \neq 0$ is rejected and the anchoring effect appears not to be responsible for differences in WTP.

Similarly, to test $\beta_{p,\ Auction\_first=1} = 0$, we regressed the bids from sessions in which the posted-price mechanism was conducted first. Again, the left panel of Table 3.6 reports a random effects Tobit model, while the right panel includes bootstrap standard errors. As both panels demonstrate, posted-price offers do not have an effect on bids when auction was conducted first. Therefore, no anchoring effect is observed.

Therefore, we conclude that WTP estimates from auction significantly differ from WTP estimates from posted price and it is not likely a result of the two tasks influencing each other but rather due to behavioral reasons.

3.5.4 Tests for Behavioral Factors

3.5.4.1 Hypothesis 3.1: Asymmetric Inconsistent Preference Effect

As mentioned in previous sections, one argument against the accuracy of WTP estimates from posted price markets is that some consumers might respond affirmatively to a posted-price question without actually forming a solid WTP, as opposed to being forced to form a value by open-ended questions such as auctions. This tendency of providing affirmative answers (if exists) would boost WTP estimates in posted price. However, we show that the percentage of affirmative inconsistent preference behavior is not significantly greater than negative inconsistent preference behavior.

We test the hypothesis that the proportion of affirmative inconsistency is greater than the proportion of negative inconsistency. Of the 480 times when WTP estimated from the posted price setting was higher than WTP estimated from the auctions, affirmative inconsistency occurred 89 times. Of the 95 times when WTP under posted prices was lower than under auctions, negative inconsistency happened 17 times. A proportion test of equality does not reject the null hypothesis that the two proportions are equal (p-value of 0.88). Thus, the proportion of affirmative inconsistency is not significantly greater than the proportion of negative inconsistency. Therefore, this hypothesis is rejected and asymmetric inconsistent preferences should not be driving the differences in WTP.

3.5.4.2 Hypothesis 3.2: Lack of Familiarity with Auction Settings

Participants' institutional information might be affected by their lack of familiarity with auction formats. We test if *roundnumber* (the number of bidding decisions a participant has made) has an effect on the bids by specifically testing whether $\beta_{\text{Auction, RoundNumber}} = 0$ holds. In order to gain more insights from the data, the regression in this part involves auction bids for all of the honey products (local, US and international). We test this hypothesis with a Tobit model, adding a set of experimental controls. The experiment controls include three information treatments, origin-information interactions, survey variables on consumer attitude towards honey, and other socio-demographic variables.[5] As shown in Table 3.7 column 1, $\beta_{\text{Auction, RoundNumber}}$, is significantly different from zero with a coefficient estimate of –0.039. Besides, in a logit model examining the probability that a participant submits a zero bid, we find that a zero bid is more likely to appear as the auction progresses (as shown in column 4, Table 3.7). Thus, the null hypothesis that $\beta_{\text{Auction, RoundNumber}} = 0$ is rejected. As the auction rounds progress, participants tend to adjust their behavior based on information gathered through the process.

The underlying reason for the change of WTP in the auctions in successive rounds is not obvious, especially since there was no feedback regarding the price and winners. Meanwhile, it is possible that some participants lost interest and stopped bidding. Thus, we considered if off-margin and on-margin bidders behaved differently. Given the size of the bids, it is reasonable to define "on-margin" bidders as those whose bids are less than \$1 below the second highest bid and the rest as "off-margin" bidders:

On margin: Bid > Second Highest Bid – \$1;

Off margin: Bid ≤ Second Highest Bid – \$1.

Column 2 and 3 in Table 3.7 show Tobit regression results for on-margin and off-margin bidders respectively with experimental controls. The results are significant and robust to inclusions of demographic and attitude controls. In sum, bids by on-margin bidders increase \$0.024 each round and bids by off-margin bidders decrease \$0.078 each round. Therefore, on-margin bidders seem to show a gradually increasing

---

5 For a detailed list of control variables used, reference Wu et al. (2015)

pattern in their revealed WTP. Note that even though we did 15 rounds of auction, in the end auction is still a relatively unfamiliar task and participants might still be relatively unfamiliar and inexperienced with it. Therefore even the increase in bids of on-margin bidders may not explain the entire discrepancy, it could be a plausible explanation for this discrepancy.

3.5.5 Hypothesis 4: Marginal Effects between Two Methods

In this part, we compare the marginal effect estimations in the two elicitation mechanisms. Since the auction bids implement a random effects Tobit model and the posted price binary responses use a random effects Logit model, the magnitudes of the variables are not directly comparable. However, the signs and significance levels of the attribute coefficients should be comparable. As demonstrated in Table 3.8, we examine the sign estimates and significance levels for the jar attributes in the posted price and auction parts. A positive sign in the coefficient would indicate a WTP premium for that attribute and a negative sign indicates the opposite. Significance levels indicate the ability to detect a preference for attributes.

The first three columns show results for auction bids. Compared to the baseline Jar1, participants are willing to pay more for honey packaged in Jar2, Jar4 and Jar5. As shown in the last three columns for the posted price part, with the same number of observations we can only demonstrate that participants are significantly willing to pay more for Jar2, while the rest of the jar attributes are highly insignificant. However, if we focus on the sign estimates in the posted price part, the result suggest that participants are willing to pay more for Jar4 and Jar5, compared to Jar1, which is consistent with the auction. The coefficient is negative but highly insignificant for Jar2, which is also highly insignificant in the auction part. Therefore, even though we obtain less significance in the posted price part, the sign estimates mostly agree with the auction part. The above analysis suggests that posted price and auction generate similar qualitative marginal effects for attributes, but auctions are more efficient in revealing these underlying preferences.

3.6 Conclusion

Experimental auctions are a popular instrument for measuring consumer WTP for various attributes of a product or service. A key attractive feature of auction mechanisms is that they provide point estimates of WTP. However, posted price formats are how most consumer choices are made. Inferring consumers' WTP for a posted price market from auction bids can be problematic since consumers generally may have relatively limited experience with auctions and may not behave in a consistent manner in both mechanisms. Therefore, some attention has been paid to comparing the estimated mean WTPs using these two mechanisms. On the other hand, the comparison of other important aspects, such as the estimation of how WTP varies with certain product attributes has not been thoroughly examined in the literature. In this research, we test the mean WTP differs in the two elicitation methods and further offer explanations of such a discrepancy using an artefactual field experiment.

Moreover, we compare the signs and significance levels of marginal effects for different product characteristics.

First, in our second price auction, estimates of WTP from bids are significantly less than estimates of WTP for the same product via the posted-price mechanism. We conducted both within-subjects and between-subjects tests and the results are robust. We test several potential explanations related to information and framing effects. The differences in WTP do not appear to be due to either an anchoring effect or asymmetric inconsistent preferences. The results do suggest that the reason for the difference in auctions is that research participants' lack of familiarity with auctions. Second, we run regressions to test the marginal effects of different product attribute on WTP. The signs of coefficients are consistent in the auction and posted price mechanisms. Third, we find that the significance level is much higher using auctions for each confident.

Our research sheds light on which economic evaluation elicitation format, namely auctions and posted price mechanisms, is more suitable under different circumstances. We show that a WTP estimate difference does exist between the two mechanisms. Participants do demonstrate an adaption process in the auction format. In the meanwhile, the posted price mechanism is more familiar to the general public and participants may focus more on the task itself. This is particularly true in a field setting where the researchers usually recruit participants from busy market places, where attention and time allocated to experiments are generally limited. However, we show that both methods elicit similar signs for the marginal effects of specific product attributes. Thus, either using auctions or posted price mechanisms can provide credible prediction on the marginal effects of important product characteristics. But auctions have clear advantage over posted price in terms of statistical power, which indicates that a larger sample size is required for a posted price mechanism to reveal the preference for specific product attributes. Therefore, it is reasonable to consider using posted price when one's goal is to understand absolute WTP values and to use auctions when one is interested in relative WTP comparisons associated with different attributes.

Table 3.1 Sample Demographic Characteristics

Variable Definition	Mean	Std. Dev.
Gender		
1 = female 0 = male	0.77	0.42
Age (years)	41.93	14.27
Years of Education	16.39	2.85
Household Yearly Income	$76,086	48,373
Primary Shoppers	0.77	0.42

Table 3. 2 Hypotheses

Question	Hypothesis	Result
1. Is there a difference in WTP between the posted-price mechanism and second-price auction?	H_0: $WTP_{Posted\_Price} = WTP_{Auction}$	Reject - There is a difference between measured WTP.
2. Is the difference due to the two tasks influencing each other?		
2.1 Is there a difference only comparing the first task completed?	H_0: $WTP_{Posted\_Price\mid Posted\_Price\_First} = WTP_{Auction\mid Auction\_First}$	Reject – There is a difference even for the first tasks completed
2.2 Is this difference due to anchoring effect?	H_0: $\beta_{p,\ Auction\_first=0} \neq 0$ H_0: $\beta_{p,\ Auction\_first=1} = 0$	Fail to Reject - No evidence of anchoring
3. Is the difference due to behavioral factors?		
3.1 Is this difference due to asymmetric inconsistent preferences?	H_0: $\Pr(Accept=1 \mid ShouldAccept=0) = \Pr(Accept=0 \mid ShouldAccept=1)$	Fail to Reject - no evidence of asymmetric inconsistent preferences
3.2 Is this difference due to a lack of familiarity with an auction setting?	H_0: $\beta_{Auction,\ RoundNumber} = 0$ H_1: $\beta_{Auction,\ RoundNumber} \neq 0$	Reject - There is evidence that the difference decreases with learning
4. Are the marginal effects comparable?	H_0: The signs and significance levels are similar	The signs are similar, significance levels are higher in auction.

Table 3.3 Within-subject comparison of estimated WTP from Posted Price and Auction

Panel A: Non-parametric	Mean(δ) (std. dev.)	Mean(δ') (std. dev.)	McNemar (p-value)
	0.2904 (0.4544)	0.1652 (0.3717)	150.45 (<0.0001)
Panel B: Parametric Assumption	$WTP_{Auction}$ (std. dev.)	WTP_{PP} (std. dev.)	Z (p-value)
Normal	2.4889 (2.0898)	4.0587 (4.5021)	7.5838 (<0.0001)
Logistic	2.4579 (1.1698)	4.0570 (2.5562)	7.5199 (<0.0001)

Table 3.4 Between-subject comparison of estimated WTP from Posted Price and Auction

Panel A: Non-parametric	Mean(δ) (std. dev.)	Mean(δ') (std. dev.)	McNemar (p-value)
	0.3396 (0.4736)	0.2192 (0.4137)	14272 (<0.0001)
Panel B: Parametric Assumption	$WTP_{Auction}$ (std. dev.)	WTP_{PP} (std. dev.)	Z (p-value)
Normal	3.1710 (2.1021)	4.6466 (4.9929)	78.0702 (<0.0001)
Logistic	3.1369 (1.1694)	4.6026 (2.8807)	74.4968 (<0.0001)

Table 3.5 Test for Anchoring When Posted Price Is before Auction

	Random Effects Tobit			Random Effect Tobit with Bootstrap Std. Err.		
	Marginal Effect	Std. Err	P>|z|	Marginal Effect	Std. Err	P>|z|
Price	–.0004	0.001	0.971	–.0004	0.009	0.966
Jar type 2	0.459	0.117	0.000	0.458	0.087	0.000
Jar type 3	-0.044	0.117	0.711	-0.044	0.116	0.707
Jar type 4	0.110	0.117	0.347	0.110	0.109	0.314
Jar type 5	0.342	0.117	0.003	0.342	0.160	0.033
_cons	2.660	0.360	0.000	2.280	0.318	0.000
Wald chi^2	28.33			43.33		
Prob> chi^2	0.000			0.000		
Log likelihood	-322.606			-322.606		
Number of Obs	265			265		
Left-censored observations	31			31		
Uncensored observations	234			234		
Right-censored observations	0			0		

Table 3.6 Test for Anchoring When Posted Price Is after Auction

	Random Effects Tobit			Random Effect Tobit with Bootstrap Std. Err.		
	Marginal Effect	Std. Err	P>|z|	Marginal Effect	Std. Err	P>|z|
Price	0.004	0.012	0.714	0.004	0.009	0.633
Jar type 2	0.303	0.136	0.026	0.303	0.105	0.004
Jar type 3	0.115	0.136	0.397	0.115	0.095	0.226
Jar type 4	0.246	0.137	0.073	0.246	0.151	0.104
Jar type 5	0.361	0.136	0.008	0.361	0.143	0.011
_cons	2.926	0.300	0.000	2.926	0.300	0.000
Wald chi^2	9.39			18.08		
Prob> chi^2	0.094			0.003		
Log likelihood	-452.642			-452.642		
Number of Obs	310			310		
Left-censored observations	14			14		
Uncensored observations	296			296		
Right-censored observations	0			0		

Table 3.7 The Effect of Round Number

	WTP Bid Amount – Random Effects Tobit			Likelihood of Zero WTP – Random Effects Logit
	All Bidders	On Margin Bidders	Off Margin Bidders	All Bidders
RoundNumber	-0.0393***	0.0237**	-0.0776***	0.0984
Experimental Controls	X	X	X	X
On Margin Bidder	X	X		X
Off-Margin Bidder	X		X	X
Jar type 2	0.421***	0.411***	0.441***	-0.470*
Jar type 3	0.180*	-0.0424	0.130	-0.244
Jar type 4	0.328***	0.202*	0.299***	-0.308
Jar type 5	0.510***	0.537***	0.406***	-0.479
_cons	2.004***	1.658***	2.257***	
Wald chi^2	942.25	772.78	650.34	190.66
Prob> chi^2	0.000	0.000	0.000	0.000
Log likelihood	-2812.153	-1165.610	-1254.617	-543.289
Number of obs	1725	773	952	1725

Notes: ***, **, * represent significance at the 1%, 5%, and 10% levels, respectively. Estimates include subject random effects. Experimental Controls include several order effects and information treatments, details can be found in Wu et al. (2015)

Table 3.8 Marginal Effect Estimation Comparison in Posted Price and Auction

	Auction			Posted Price		
	Coefficient	Std. Err	P>|z|	Coefficient	Std. Err	P>|z|
Jar type 2	0.374	0.092	0.000	0.710	0.313	0.023
Jar type 3	0.043	0.092	0.639	-0.054	0.329	0.869
Jar type 4	0.183	0.092	0.048	0.351	0.318	0.270
Jar type 5	0.355	0.092	0.000	0.303	0.319	0.342
_cons	2.643	0.198	0.000	-1.311	0.253	0.000
Wald chi^2	28.06			7.79		
Prob> chi^2	0.000			0.099		
Log likelihood	-786.041			-335.720		
Number of Obs	575			575		
Left-censored observations	45			0		
Uncensored observations	530			575		
Right-censored observations	0			0		

Figure 3.1 Frequency Distribution of Bid Amounts in the Auction

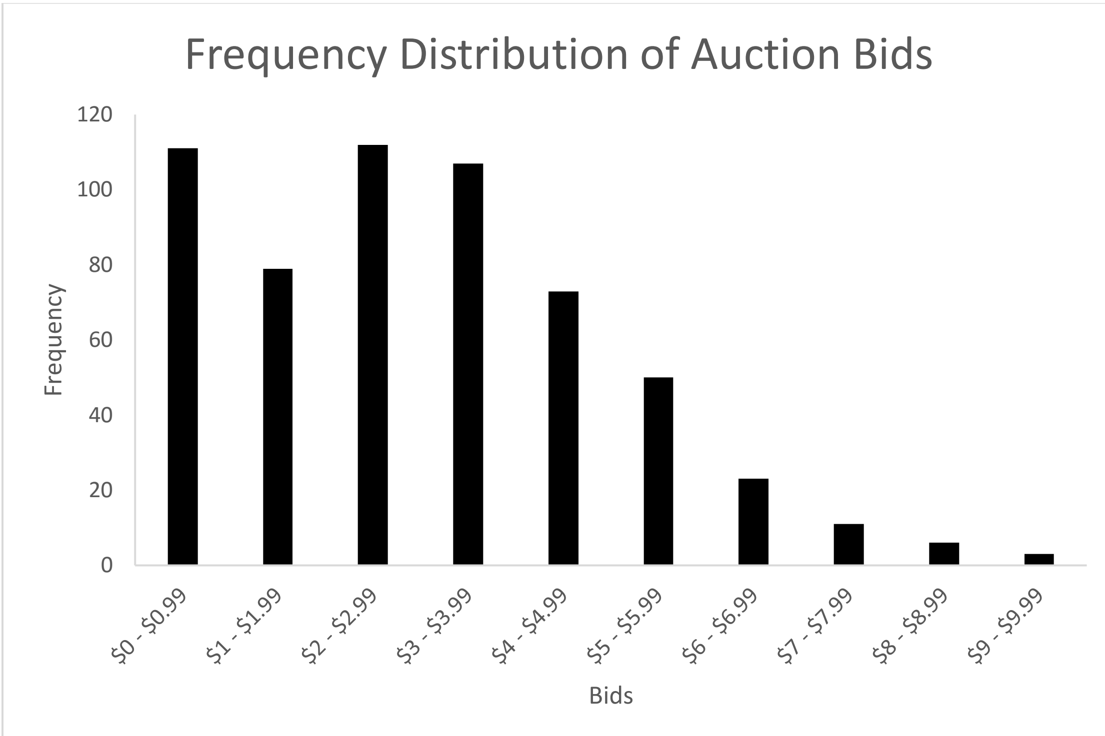
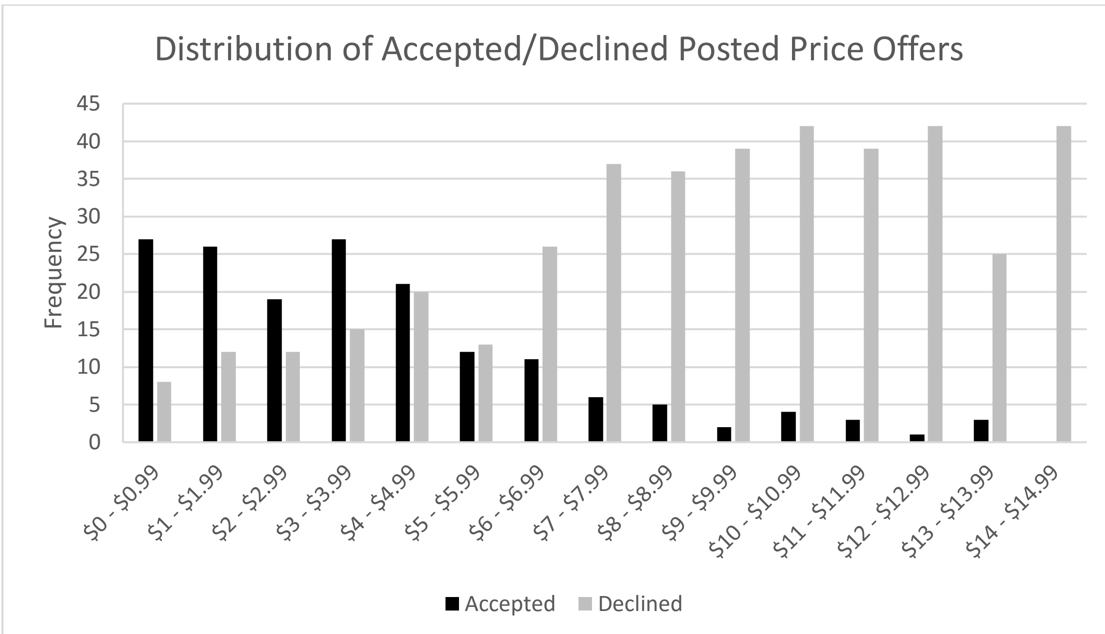


Figure 3.2 Frequency Distribution of Accepted and Declined Posted Price Offers



REFERENCES

Alfnes, Frode. 2009. “Valuing Product Attributes in Vickrey Auctions when Market Substitutes are Available.” *European Review of Agricultural Economics*: jbp013.

Allcott, H., 2011. “Social norms and energy conservation.” *Journal of Public Economics*, 95(9), pp.1082-1095.

Alpízar, Francisco, Till Requate, and Albert Schram. 2004. “Collective Versus Random Fining: An Experimental Study on Controlling Ambient Pollution.” *Environmental and Resource Economics* 29 (2): 231-252.

An, L. 2012. “Modeling human decisions in coupled human and natural systems: review of agent-based models.” *Ecological Modelling*, *229*, 25-36.

Ariely, Dan, George Loewenstein, and Drazen Prelec. 2003. ““Coherent arbitrariness”: Stable demand curves without stable preferences.” The Quarterly Journal of Economics 118 (1): 73-106.

Arrow, Kenneth, Robert Solow, Paul R. Portney, Edward E. Leamer, Roy Radner, and Howard Schuman. 1993. Report of the NOAA panel on contingent valuation. *Federal register*, 58(10), pp.4601-4614.

Axelrod, R. M. 1997. “The complexity of cooperation: Agent-based models of competition and collaboration.” *Princeton University Press.*

Axtell, R. L., Epstein, J. M., Dean, J. S., Gumerman, G. J., Swedlund, A. C., Harburger, J., ... & Parker, M. 2002. “Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley.” *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7275-7279.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak. 1964. “Measuring Utility by a Single−response Sequential Method.” *Behavioral Science* 9 (3): 226-232.

Bernard, John C. 2006. “Finding and Retaining the Dominant Strategy: The Second-Price, English and ‘sealed offer’ English Auctions.” *Journal of Economic Behavior & Organization* 60 (4): 553-561.

Bernedo, M., Ferraro, P.J. and Price, M., 2014. "The persistent impacts of norm-based messaging and their implications for water conservation." *Journal of Consumer Policy*, 37(3), pp.437-452.

Boyer, Christopher N., B. Wade Brorsen, and Tong Zhang. 2014. "Common-Value Auction Versus Posted-Price Selling: An Agent-Based Model Approach." *Journal of Economic Interaction and Coordination* 9 (1): 129-149.

Cason, Timothy N. and Lata Gangadharan. 2013. "Empowering Neighbors Versus Imposing Regulations: An Experimental Analysis of Pollution Reduction Schemes." *Journal of Environmental Economics and Management* 65 (3): 469-484.

Chang, Jae Bong, Jayson L. Lusk, and F. Bailey Norwood. 2009. "How Closely do Hypothetical Surveys and Laboratory Experiments Predict Field Behavior?" *American Journal of Agricultural Economics* 91 (2): 518-534.

Cherry, Todd, Peter Frykblom, Jason Shogren, John List, and Melonie Sullivan. 2004. "Laboratory Testbeds and Non-Market Valuation: The Case of Bidding Behavior in a Second-Price Auction with an Outside Option." *Environmental and Resource Economics* 29 (3): 285-294.

Colson, Gregory, Jay R. Corrigan, and Matthew C. Rousu. 2010. "The Impact of Perceived Prices on Willingness to Pay in Experimental Auctions." *Journal of Agricultural & Food Industrial Organization* 8 (1).

Conservation Technology Information Center, Purdue University. "Conservation Buffers.", accessed 09/08, 2016, http://www.ctic.purdue.edu/Core4/Conservation%20Buffers/.

Corrigan, Jay R., Andreas C. Drichoutis, Jayson L. Lusk, Rodolfo M. Nayga, and Matthew C. Rousu. 2011. "Repeated Rounds with Price Feedback in Experimental Auction Valuation: An Adversarial Collaboration." *American Journal of Agricultural Economics*: aar066.

Couch, Arthur and Kenneth Keniston. 1960. "Yeasayers and Naysayers: Agreeing Response Set as a Personality Variable." *The Journal of Abnormal and Social Psychology* 60 (2): 151.

Cummings, R. G., Holt, C. A., & Laury, S. K. 2004. "Using laboratory experiments for policymaking: an example from the Georgia irrigation reduction auction." *Journal of Policy Analysis and Management*, *23*(2), 341-363.

Dia, H. 2002. “An agent-based approach to modelling driver route choice behaviour under the influence of real-time information.” *Transportation Research Part C: Emerging Technologies*, *10*(5), 331-349.

Diamond, Peter A. and Jerry A. Hausman. 1994. “Contingent Valuation: Is some Number Better than no Number?” *The Journal of Economic Perspectives* 8 (4): 45-64.

Dillaway, Robin, Kent D. Messer, John C. Bernard, and Harry M. Kaiser. 2011. “Do Consumer Responses to Media Food Safety Information Last?” *Applied Economic Perspectives and Policy* 33 (3): 363-383.

Duffy, J. 2006. “Agent-based models and human subject experiments.” *Handbook of computational economics*, *2*, 949-1011.

Evans, T. P., Sun, W., & Kelley, H. 2006. “Spatially explicit experiments for the exploration of land-use decision-making dynamics.” *International Journal of Geographical Information Science*, *20*(9), 1013-1037.

Farmer, J. D., & Foley, D. 2009. “The economy needs agent-based modelling.” *Nature, 460*(7256), 685-686.

Ferraro, P. J., & Price, M. K. 2013. “Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment.” *Review of Economics and Statistics*, *95*(1), 64-73.

Festinger, L. 1954. “A theory of social comparison processes.” *Human relations*, *7*(2), 117-140.

Fooks, J. R., Higgins, N., Messer, K. D., Duke, J. M., Hellerstein, D., & Lynch, L. 2016. “Conserving spatially explicit benefits in ecosystem service markets: Experimental tests of network bonuses and spatial targeting.” *American Journal of Agricultural Economics*, aav061.

Frykblom, Peter and Jason F. Shogren. 2000. “An Experimental Testing of Anchoring Effects in Discrete Choice Questions.” *Environmental and Resource Economics* 16 (3): 329-341.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. 2008. “A room with a viewpoint: Using social norms to motivate environmental conservation in hotels.” *Journal of Consumer Research*, *35*(3), 472-482.

Gracia, Azucena, Maria L. Loureiro, and Rodolfo M. Nayga. 2011. “Are Valuations from Nonhypothetical Choice Experiments Different from those of Experimental Auctions?” *American Journal of Agricultural Economics* 93 (5): 1358-1373.

Green, Donald, Karen E. Jacowitz, Daniel Kahneman, and Daniel McFadden. 1998. “Referendum Contingent Valuation, Anchoring, and Willingness to Pay for Public Goods.” *Resource and Energy Economics* 20 (2): 85-116.

Grimm, V., & Railsback, S. F. 2005. “*Individual-based modeling and ecology”* (Vol. 2005). Princeton: Princeton university press.

Hammond, Robert G. 2010. “Comparing Revenue from Auctions and Posted Prices.” *International Journal of Industrial Organization* 28 (1): 1-9.

Hanemann, W. Michael. 1984. “Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses.” *American Journal of Agricultural Economics* 66 (3): 332-341.

Harrison, Glenn W., Ronald M. Harstad, and E. Elisabet Rutström. 2004. “Experimental Methods and Elicitation of Values.” *Experimental Economics* 7 (2): 123-140.

Heckbert, S. 2009. Experimental economics and agent-based models. In *18th World IMACS/MODSIM Congress* (pp. 2997-3003).

Heckbert, S., Baynes, T., & Reeson, A. 2010. “Agent-based modeling in ecological economics.” *Annals of the New York Academy of Sciences*, *1185*(1), 39-53.

Herriges, Joseph A. and Jason F. Shogren. 1996. “Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning.” *Journal of Environmental Economics and Management* 30 (1): 112-131.

Johnson, Liz, Linda Lemasters, and Prasun Bhattacharyya. 2017. “A Complexity Context to Classroom Interactions and Climate: ABM Forecasting Achievment.” *Journal on Policy and Complexity Systems*, 3(1), 28-47.

Kaas, Klaus and Heidrun Ruprecht. 2006. “Are the Vickrey Auction and the BDM-Mechanism really Incentive Compatible? Empirical Results and Optimal Bidding Strategies in the Case of Uncertain Willingness-to-Pay.” *Empirical Results and Optimal Bidding Strategies in the Case of Uncertain Willingness-to-Pay*.

Kanninen, Barbara J. 1995. “Bias in Discrete Response Contingent Valuation.” *Journal of Environmental Economics and Management* 28 (1): 114-125.

Khouja, M., Hadzikadic, M., & Zaffar, M. A. 2008. “An agent based modeling approach for determining optimal price-rebate schemes.” *Simulation Modelling Practice and Theory*, *16*(1), 111-126.

Kimenta, Jan. 1986. Elements of Econometrics. *Ann Arbor: MI: University of Michigan Press.*

Kriström, Bengt. 1993. "Comparing Continuous and Discrete Contingent Valuation Questions." *Environmental and Resource Economics* 3 (1): 63-71.

Li, Jingyuan, Holly A. Michael, Joshua M. Duke, Kent D. Messer, and Jordan F. Suter. 2014. "Behavioral Response to Contamination Risk Information in a Spatially Explicit Groundwater Environment: Experimental Evidence." *Water Resources Research* 50 (8): 6390-6405.

Li, Tongzhe, Maik Kecinski, and Kent D. Messer. 2017. "Heterogeneous Responses and Demographic Differences for Oyster Consumption: Evidence from Field Experiments." Agricultural and Resource Economics Review, *Forthcoming.*

Li, Tongzhe, and Jill J. McCluskey. 2017. "Consumer preferences for second-generation bioethanol." Energy Economics 61: 1-7.

List, John A., and Craig A. Gallet. 2001. "What experimental protocol influence disparities between actual and hypothetical stated values? " Environmental and Resource Economics 20 (3): 241-254.

List, John A. 2003. "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics-Cambridge Massachusetts-* 118 (1): 41-72.

List, John A. and David Lucking-Reiley. 2000. "Demand Reduction in Multiunit Auctions: Evidence from a Sportscard Field Experiment." *The American Economic Review* 90 (4): 961-972.

Liu, Zhuo, Jordan F. Suter, Kent D. Messer, Joshua M. Duke, and Holly A. Michael. 2014. "Strategic Entry and Externalities in Groundwater Resources: Evidence from the Lab." *Resource and Energy Economics* 38: 181-197.

Loomis, John, Thomas Brown, Beatrice Lucero, and George Peterson. 1997. "Evaluating the Validity of the Dichotomous Choice Question Format in Contingent Valuation." *Environmental and Resource Economics* 10 (2): 109-123.

Loureiro, Maria L., Jill J. McCluskey, and Ron C. Mittelhammer. 2002. "Will consumers pay a premium for eco- labeled apples?." Journal of Consumer Affairs 36 (2): 203-219.

Lusk, Jayson L., Ty Feldkamp, and Ted C. Schroeder. 2004. "Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods." *American Journal of Agricultural Economics* 86 (2): 389-405.

Lusk, Jayson L. and Ted C. Schroeder. 2006. "Auction Bids and Shopping Choices." *Advances in Economic Analysis & Policy* 6 (1).

Lusk, Jayson L. and Jason F. Shogren. 2007. *Experimental Auctions: Methods and Applications in Economic and Marketing Research* Cambridge University Press.

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. 2007. "Agent-based land-use models: a review of applications." *Landscape Ecology*, *22*(10), 1447-1459.

McNemar, Quinn. 1947. "Note on the Sampling Error of the Difference between Correlated Proportions Or Percentages." *Psychometrika* 12 (2): 153-157.

Messer, K. D., Kaiser, H. M., & Schulze, W. D. 2008. "The problem of free riding in voluntary generic advertising: parallelism and possible solutions from the lab." *American Journal of Agricultural Economics*, *90*(2), 540-552.

Messer, Kent D., Gregory L. Poe, Daniel Rondeau, William D. Schulze, and Christian A. Vossler. 2010. "Social Preferences and Voting: An Exploration using a Novel Preference Revealing Mechanism." *Journal of Public Economics* 94 (3): 308-317.

Miao, H., Fooks, J. R., Guilfoos, T., Messer, K. D., Pradhanang, S. M., Suter, J. F., … & Uchida, E. 2016. "The impact of information on behavior under an ambient-based policy for regulating nonpoint source pollution." *Water Resources Research*, *52*(5), 3294-3308.

Ng, T. L., Eheart, J. W., Cai, X., & Braden, J. B. 2011. "An agent-based model of farmer decision-making and water quality impacts at the watershed scale under markets for carbon allowances and a second-generation biofuel crop." *Water Resources Research*, *47*(9).

Palm-Forster, L.H., J.F. Suter, and K.D. Messer. 2017. "Linking Eligibility for Agricultural Subsidies to Water Quality." Applied Economics and Statistics Working Paper, APEC RR17-01. Available at: https://dspace.udel.edu/bitstream/handle/19716/21176/APEC%20RR17-01%20FINAL.pdf?sequence=2&isAllowed=y [Accessed May 24, 2017].

Parker, D. C. 2014. "An Economic Perspective on Agent-Based Models of Land Use and Land Cover Change." *The Oxford Handbook of Land Economics*, 402.

Plott, Charles R. 1996. *Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis* McMillian.

Plott, Charles R. and Kathryn Zeiler. 2005. "The Willingness to Pay-Willingness to Accept Gap, the'Endowment Effect,'Subject Misconceptions, and Experimental Procedures for Eliciting Valuations." *American Economic Review, June*.

Poe, Gregory L., Kelly L. Giraud, and John B. Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87 (2): 353-365.

Poe, G. L., Schulze, W. D., Segerson, K., Suter, J. F., & Vossler, C. A. 2004. "Exploring the performance of ambient-based policy instruments when nonpoint source polluters can cooperate." *American Journal of Agricultural Economics*, *86*(5), 1203-1210.

Ready, Richard C., Jean C. Buzby, and Dayuan Hu. 1996. "Differences between Continuous and Discrete Contingent Value Estimates." *Land Economics*: 397-411.

Roosen, Jutta, Stephan Marette, and Sandrine Blanchemanche. 2010. "Value Elicitation using BDM and a Discrete Choice Mechanism." *Economics Bulletin* 30 (2): 1554-1563.

Rowe, R.D., Schulze, W.D. and Breffle, W.S., 1996. A test for payment card biases. *Journal of Environmental Economics and Management*, 31(2), pp.178-185.

Rutström, E. Elisabet. 1998. "Home-Grown Values and Incentive Compatible Auction Design." *International Journal of Game Theory* 27 (3): 427-441.

Schnier, Kurt Erik. 2009. "Spatial Externalities and the Common-Pool Resource Mechanism." *Journal of Economic Behavior & Organization* 70 (1): 402-415.

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

Shogren, Jason F., Michael Margolis, Cannon Koo, and John A. List. 2001. "A Random Nth-Price Auction." *Journal of Economic Behavior & Organization* 46 (4): 409-421.

Spraggon, John. 2002. "Exogenous Targeting Instruments as a Solution to Group Moral Hazards." *Journal of Public Economics* 84 (3): 427-456.

Spraggon, John. 2004. "Testing Ambient Pollution Instruments with Heterogeneous Agents." *Journal of Environmental Economics and Management* 48 (2): 837-856.

Spraggon, John. 2013. “The Impact of Information and Cost Heterogeneity on Firm Behaviour under an Ambient Tax/subsidy Instrument.” *Journal of Environmental Management*, 122, pp.137-143.

Spraggon, John, and Robert J. Oxoby. 2010. “Ambient-based policy instruments: the role of recommendations and presentation.” *Agricultural and Resource Economics Review*, 39(2), pp.262-274.

Suter, Jordan F., Joshua M. Duke, Kent D. Messer, and Holly A. Michael. 2012. “Behavior in a Spatially Explicit Groundwater Resource: Evidence from the Lab.” *American Journal of Agricultural Economics*: aas058.

Suter, J. F., & Vossler, C. A. 2013. “Towards an understanding of the performance of ambient tax mechanisms in the field: evidence from upstate New York dairy farmers.” *American Journal of Agricultural Economics*, aat066.

Suter, Jordan F., Christian A. Vossler, and Gregory L. Poe. 2009. “Ambient-Based Pollution Mechanisms: A Comparison of Homogeneous and Heterogeneous Groups of Emitters.” *Ecological Economics* 68 (6): 1883-1892.

Suter, Jordan F., Christian A. Vossler, Gregory L. Poe, and Kathleen Segerson. 2008. “Experiments on Damage-Based Ambient Taxes for Nonpoint Source Polluters.” *American Journal of Agricultural Economics* 90 (1): 86-102.

Tesfatsion, L., Rehmann, C. R., Cardoso, D. S., Jie, Y., & Gutowski, W. J. 2017. “An agent-based platform for the study of watersheds as coupled natural and human systems.” *Environmental Modelling & Software*, *89*, 40-60.

Tsai, Y., Zia, A., Koliba, C., Bucini, G., Guilbert, J., & Beckage, B. 2015. “An interactive land use transition agent-based model (ILUTABM): Endogenizing human-environment interactions in the Western Missisquoi Watershed.” *Land Use Policy*, 49, 161-176.

Tversky, A. and D. Kahneman. 1974. “Judgment Under Uncertainty: Heuristics and Biases.” *Science (New York, N.Y.)* 185 (4157): 1124-1131.

USDA NASS. 2012 Census of Agriculture.

Veldkamp, A., & Verburg, P. H. 2004. “Modelling land use change and environmental impact.” *Journal of Environmental Management*, *72*(1), 1-3.

Venkatachalam, L. 2004. “The contingent valuation method: a review,” Environmental Impact Assessment Review 24(1): 89–124.

Vickrey, William. 1961. “Counterspeculation, Auctions, and Competitive Sealed Tenders.” *The Journal of Finance* 16 (1): 8-37.

Vossler, C. A., Mount, T. D., Thomas, R. J., & Zimmerman, R. D. 2009. “An experimental investigation of soft price caps in uniform price auction markets for wholesale electricity.” *Journal of Regulatory Economics*, *36*(1), 44-59.

Wang, Ruqu. 1993. “Auctions Versus Posted-Price Selling.” *The American Economic Review*: 838-851.

Wang, Xin, Alan Montgomery, and Kannan Srinivasan. 2008. “When Auction Meets Fixed Price: A Theoretical and Empirical Examination of Buy-it-Now Auctions.” *Quantitative Marketing and Economics* 6 (4): 339-370.

Weersink, Alfons, John Livernois, Jason F. Shogren, and James S. Shortle. 1998. “Economic Instruments and Environmental Policy in Agriculture.” *Canadian Public Policy/Analyse De Politiques*: 309-327.

Wu, S., J. Fooks, K.D. Messer, and D. Delaney. 2015. “Consumer Demand for Local Honey.” *Applied Economics.* 47(14): 4377-4394.

Wu, S., Palm-Forster, L., & Messer, K. D. 2017. “Heterogeneous agents and information nudges in non-point source water pollution management.” *APEC working paper RR17-04,* University of Delaware.

Xepapadeas, Anastasios. 2011. “The Economics of Non-Point-Source Pollution.” *Annu.Rev.Resour.Econ.* 3(1): 355-373.

Xepapadeas, Anastasios P. 1992. “Environmental Policy Design and Dynamic Nonpoint-Source Pollution.” *Journal of Environmental Economics and Management* 23 (1): 22-39.

Xie, Jing and Zhifeng Gao. 2013. “The Comparison of Three Non-Hypothetical Valuation Methods: Choice Experiments, Contingent Valuation, and Experimental Auction.” Working Paper.

Yoo, Seung-Hoon, and Hee-Jong Yang. 2001 “Application of sample selection model to double-bounded dichotomous choice contingent valuation studies.” *Environmental and resource economics* 20 (2): 147-163.

Zenobia, B., Weber, C., & Daim, T. 2009. “Artificial markets: A review and assessment of a new venue for innovation research.” Technovation, 29(5), 338-350.

Zia, A., Messer, K., Ding, S., Miao, H., Suter, J., Fooks, J., … & Koliba. C., 2016a. "Spatial effects of sensor information in inducing cooperative behaviors for managing non-point source pollution: evidence from a decision game in an idealized watershed." Ecological Economics (in review).

Zia, A, Tsai, Y., Turnbull, S., Ding, S., Miao, H., Koliba, C., … & Uchida, E. 2016b. "Simulating the effects of alternate control strategies on heterogeneous farmer behaviors and water quality outcomes: an agent-based modeling application in the Mississquoi Watershed of Lake Champlain Basin." SwarmFest2016: 20th Annual Meeting on Agent-based Modeling and Simulation. Burlington VT, July 31 – August 2, 2016.

Zia, A. and Koliba, C., 2015. "The emergence of attractors under multi-level institutional designs: agent-based modeling of intergovernmental decision making for funding transportation projects." AI & society, 30(3), pp.315-331.

Appendix A

EXPERIMENT INSTRUCTIONS FOR CHAPTER 1 AND 2

Thank you for participating!

Please return the signed consent form to the administrator.

Please read and follow the instructions carefully and do not communicate with others during the experiment.

INTRODUCTION

This is an experiment about the economics of decision making. You will earn money during this experiment if you follow these instructions carefully and make informed decisions; otherwise, you may end up losing money. Any money earned during this experiment will initially be recorded as experimental dollars. At the end of this experiment, we will convert your experimental dollars into actual US dollars that will be handed to you as you leave. The more experimental dollars you earn the more actual US dollars you will receive. At the end of the experiment, your earnings will be converted at a rate of $1 US dollar for 50 experimental dollars. Please read these instructions carefully and do not communicate with any other participants during the experiment.

General Instructions: Today's experiment has several parts. Each part will have five rounds. Each round is independent, meaning that decisions during a round do not affect future rounds in any way. The only value that gets carried over across rounds is the cumulative amount of money you earn, which will be used to calculate your cash earnings at the end of the experiment.

Your role: You own and operate a firm. You will make decisions that affect the amount of money your firm earns. This money will be called your **Firm Profit**.

Groups: Throughout the experiment, you will be in a group of eight people, each will play the role of a firm. Think of your firm and the seven other firms as being located near a river. Groups are randomly reassigned after each part of the experiment and you will not know who is assigned to each group.

Production and Production Income: Each business owner produces output that creates **Production Income**. Production income only depends on how much is produced. The more a firm produces, the more production income the firm will get.

Pollution: Production also generates pollution that goes into the river. In general, the higher the output being produced, the more pollution is being generated. Some concentration of this pollution is harmless. However, if the concentration is too large, the pollution has negative effects to the environment.

Total Pollution: This is measured by a sensor downstream and is the sum of pollution for everyone in the same group.Capacity: The firms may have a different production capacity, which is the maximum amount your firm can produce. Each firm's capacity will be shown on the calculator in the corresponding part for that firm. There are three types of capacities: Large firms with a capacity of 125; medium firms with a capacity of 100; small firms with a capacity of 75.

Technology: At the beginning of each round, the firms may choose to adopt a technology at a cost proportional to your firm capacity. When adopted, the technology will reduce the firm's pollution to a certain percentage of the original level for that round.

Location: The firms may either be located in the same location or at different locations along a river. As shown in Figure 1, when the region is separated by lines, it means the region is being divided into Region 1 to Region 4. In this case, Region 1 is the most upstream and Region 4 is the most downstream. The further downstream your firm is the more pollution per unit of production will be recorded by the sensor. As shown in Figure 2, when there are no lines separating the region, it means all of the firms are placed in the same region. The actual capacity and location of the firm that you operate will be shown on your computer screen.

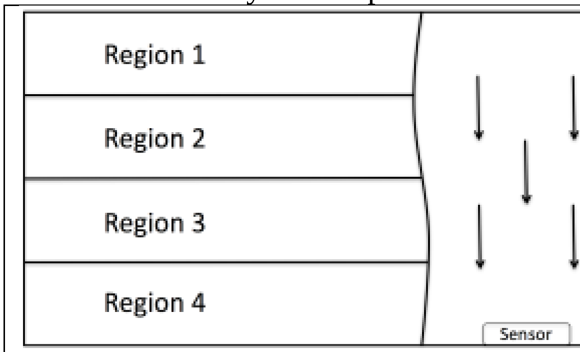


Figure 1. Different Locations

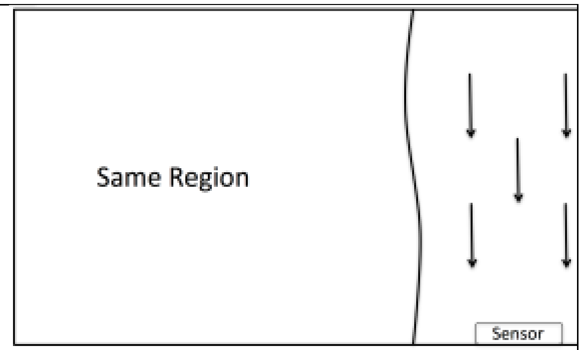


Figure 2. Same Location

Decisions: In each round, you will make two decisions:

(1) **Production Decision** – You will decide your firm's production level, between 0 and your firm's capacity.

(2) **Technology Decision** – You will choose whether to adopt a technology at a certain cost, labeled "Not Adopt" or "Adopt".

Pollution Table: To help you better understand the relationship of production, technology, location and pollution, you are given a **Pollution Table** that has pollution levels of a firm corresponding to different production decisions, technology decisions and location. Use this table to understand how your production would affect pollution based on your location and technology decision.

Firm Profit: Your **firm profit** is calculated based on your production decision and technology decision and will be explained to you in further details in each part of the experiment.

Decision Calculator: A **Decision Calculator** is provided to test different scenarios to see how the decisions of other firms in your group could affect Total Pollution and your Firm Profit. Follow the instructions on how to use this calculator provided on the next page.

In summary:

- In each part of the experiment, you will be given additional instructions and all calculations will be described.
- Your earnings from the experiment depend on your cumulative firm profit.
- Use the decision calculator to test out different scenarios and determine your own production and technology decision.
- Choose your own production and technology decision and click “Confirm”.
- Your production income is affected by your production decision, technology decision, and firm capacity.
- Your pollution depends on your production decision, technology decision and firm location.
- A round of the experiment is complete when all eight players have made their production and technology decisions.
- After each part, participants will be randomly reassigned to a new group.

HOW TO USE THE DECISION CALCULATOR AND MAKE DECISIONS

In each round, you will be provided with a decision calculator like the one in the attached handout.

The layout of all firms and their corresponding capacity in your group is shown in the calculator.

Your firm is labeled “Your Firm” and marked with a black box.

Step 1. On the left part of the page, assume what everyone in your group will be doing by choosing a production and technology decision for every firm. To choose a production decision, move the slider or type in the amount that you think other firms will be producing; to choose a technology decision, simply choose between the “Not Adopt” and “Adopt” options. Note that your firm is labeled in the black box and you do not have to choose technology decision for your firm.

Step 2. On the top right part of the page, click “Calculate” and your pollution, total pollution and your profit of “Not adopt” and “Adopt” will be shown to you in the table right under the “Calculate” button.

Keep in mind that the decisions you make in the decision calculator are for <u>informational purposes only</u> and other firms can make their own decisions regardless of what you choose for them.

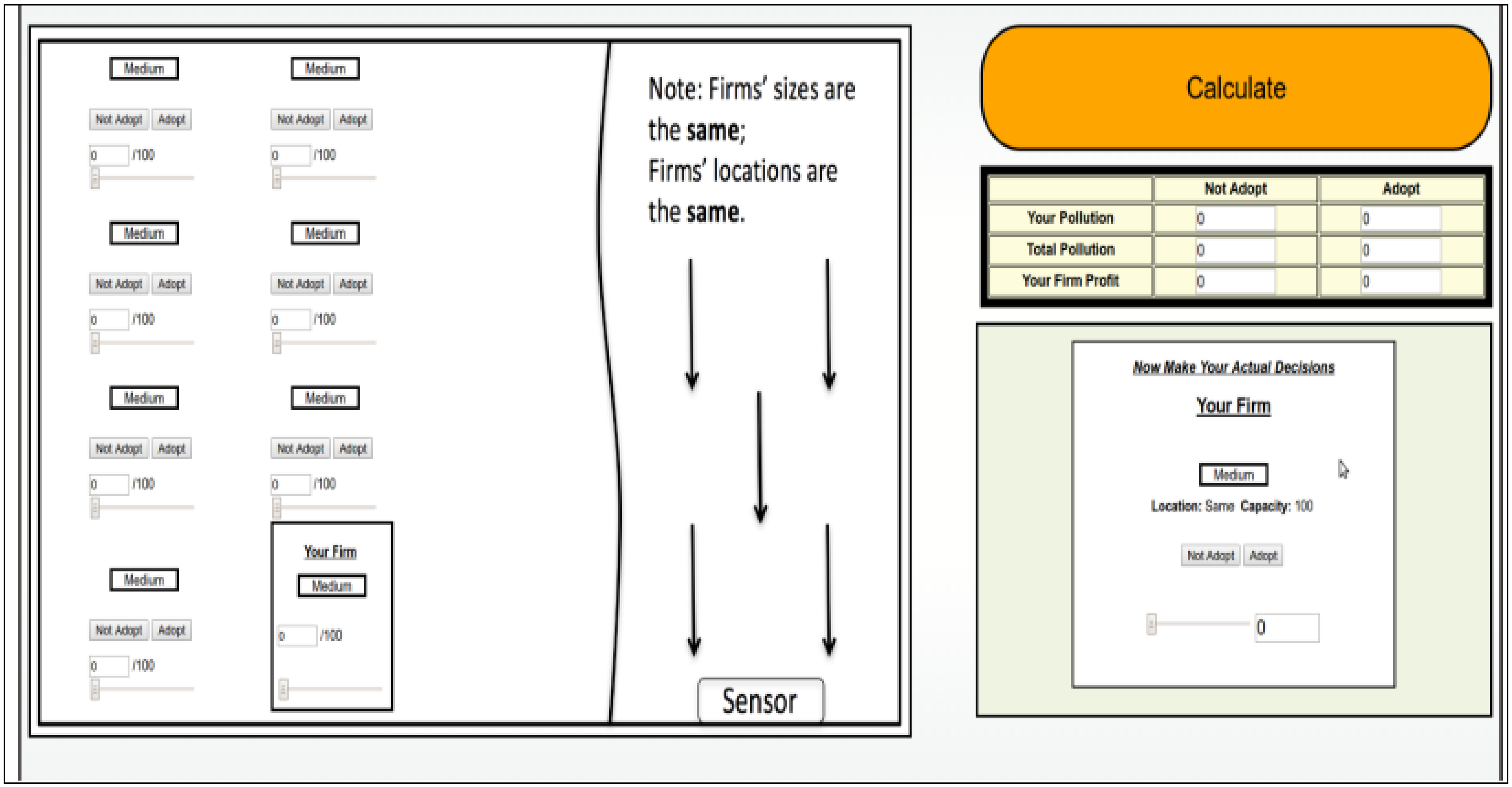
After you decide what your decision will be, make your actual decision in Step 3.

Step 3. On the bottom right part of the page, choose your <u>actual production decision</u> with the slider, and pick your <u>actual technology decision</u>. When you are done, click “Confirm”. Once you have clicked this button, the button will turn gray and it is no longer possible to change your decisions for that round.

<u>Results</u> – While you are waiting for the other players to make their decisions, you can review the results of past rounds, which will be shown on your screen. After all eight players have clicked the Confirm button, the results of the current round will appear, including Your Pollution, the Total Pollution from all members of your group, your Production Income, and Your Firm Profit.

DECISION CALCULATOR

The image below are examples of the interactive Decision Calculator that you will use on your computer.



Pollution Table

This Pollution Table helps you to better understand how your firm's production decision, technology decision and location affect your pollution. Use this table along with the Decision Calculator to help you make more informed decisions.

How to read this table?

1. The first column (Production) indicates how much is being produced.
2. Find where your firm is located from the Decision Calculator. If every firm is in the same region, use the last two columns (marked as "Same Region").
3. Your firm's pollution for each level of production under "Not Adopt" and "Adopt" are listed in the columns corresponding to your region.

	Your Firm Pollution									
Production	**Region 1**		**Region 2**		**Region 3**		**Region 4**		**Same Region**	
	Not Adopt	Adopt	Not Adopt	Adopt	Not Adopt	Adopt	Not Adopt	Adopt	Not Adopt	Adopt
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	1.20	0.60	1.40	0.70	1.60	0.80	1.80	0.90	1.50	0.75
10	2.40	1.20	2.80	1.40	3.20	1.60	3.60	1.80	3.00	1.50
15	3.60	1.80	4.20	2.10	4.80	2.40	5.40	2.70	4.50	2.25
20	4.80	2.40	5.60	2.80	6.40	3.20	7.20	3.60	6.00	3.00
25	6.00	3.00	7.00	3.50	8.00	4.00	9.00	4.50	7.50	3.75
30	7.20	3.60	8.40	4.20	9.60	4.80	10.80	5.40	9.00	4.50
35	8.40	4.20	9.80	4.90	11.20	5.60	12.60	6.30	10.50	5.25
40	9.60	4.80	11.20	5.60	12.80	6.40	14.40	7.20	12.00	6.00
45	10.80	5.40	12.60	6.30	14.40	7.20	16.20	8.10	13.50	6.75
50	12.00	6.00	14.00	7.00	16.00	8.00	18.00	9.00	15.00	7.50
55	13.20	6.60	15.40	7.70	17.60	8.80	19.80	9.90	16.50	8.25
60	14.40	7.20	16.80	8.40	19.20	9.60	21.60	10.80	18.00	9.00
65	15.60	7.80	18.20	9.10	20.80	10.40	23.40	11.70	19.50	9.75
70	16.80	8.40	19.60	9.80	22.40	11.20	25.20	12.60	21.00	10.50
75	18.00	9.00	21.00	10.50	24.00	12.00	27.00	13.50	22.50	11.25
80	19.20	9.60	22.40	11.20	25.60	12.80	28.80	14.40	24.00	12.00
85	20.40	10.20	23.80	11.90	27.20	13.60	30.60	15.30	25.50	12.75
90	21.60	10.80	25.20	12.60	28.80	14.40	32.40	16.20	27.00	13.50
95	22.80	11.40	26.60	13.30	30.40	15.20	34.20	17.10	28.50	14.25
100	24.00	12.00	28.00	14.00	32.00	16.00	36.00	18.00	30.00	15.00
105	25.20	12.60	29.40	14.70	33.60	16.80	37.80	18.90	31.50	15.75
110	26.40	13.20	30.80	15.40	35.20	17.60	39.60	19.80	33.00	16.50
115	27.60	13.80	32.20	16.10	36.80	18.40	41.40	20.70	34.50	17.25
120	28.80	14.40	33.60	16.80	38.40	19.20	43.20	21.60	36.00	18.00

125	30.00	15.00	35.00	17.50	40.00	20.00	45.00	22.50	37.50	18.75

For Example:

1. A firm in Region 1, producing 75 units. Firm Pollution for not adopt: 18; adopt: 9.
2. A firm in Region 4, producing 75 units. Firm Pollution for not adopt: 27, adopt: 13.5.
3. A firm in Same Region, producing 100 units. Firm Pollution for not adopt: 30; adopt: 15.

ID# _________

UNDERSTANDING THE EXPERIMENT

This short exercise is designed to help you understand how the experiment works. The profit you earn in this section does not affect your real earnings.

Please use the decision calculator on the computer in front of you to figure out what your firm profit will be under the following scenarios:

You will be guided through Scenario A, and you will complete scenario B by yourself.

Scenario A:
Please fill in your profit for the following hypothetical decisions. The steps listed below will guide you through scenario A.

Everyone else		**You**		
Technology	**Production**	**Your Production**	**Your Technology**	**Your Profit**
Not Adopt	80	50	Not Adopt	
Not Adopt	80	50	Adopt	

Step 1: On the left part of the page, select “Not Adopt” for everyone else except your firm.
Step 2: Use the slider or type in the boxes to change everyone else’s production to 80 units.
Step 3: Still on the left part of the page, find the box that lists “Your Firm”, change the production decision to 50 units.
Step 4: Click “Calculate”. Your pollution, total pollution and your firm profit should be shown to you.
Step 5: Find “Your Firm Profit” for “Not Adopt”, which should be “33.75” in this case. Type in “33.75” in the first row under profit for scenario A.
Step 6: Find “Your Firm Profit” for “Adopt”, which should be “25.55” in this case. Type in “25.55” in the second row under profit for scenario A.
Step 7: Click “Check answer for scenario A” when you are done. If the program asks you to try again, please check answers for the highlighted parts.

Now please complete scenario B on your own, please raise your hand if you have any questions.

Scenario B:
Please fill in your profit for the following hypothetical decisions on the computer screen.

Everyone else Technology	**Every else Production**	**Your Production**	**Your Technology**	**Your Profit**
Not Adopt	80	50	Not Adopt	
Not Adopt	80	50	Adopt	
Not Adopt	80	80	Not Adopt	
Not Adopt	80	80	Adopt	
Everyone else		**You**		
Technology	**Production**	**Your Production**	**Your Technology**	**Your Profit**
Adopt	100	100	Not Adopt	
Adopt	100	100	Adopt	

You may refer to instructions for Scenario A to help you complete Scenario B.

Input your firm profit for Scenario B on the computer program and check if it is correct by clicking "check answers". When the program asks you to "try again", it means your answer is not correct and will be highlighted. In that case, please use the calculator to recalculate the answer.

When you get both scenarios correct, you may click the continue button to move on to the next part.

INSTRUCTIONS FOR PRACTICE

You will now play five practice rounds to learn how the experiment works. The outcomes of these rounds will not affect your cash earnings.

In each round of this part, you will make your Production Decision and your Technology Decision. Use the Decision Calculator to see how your decision and others' decisions affect your earnings.

In this practice part, pollution does not affect firm profits. The more you produce, the more your firm profit will be.

After everyone makes their decisions, you will see the results screen that will display your
Firm Profit and Pollution. In this part, your Firm Profit will be calculated as follows:

Firm Profit = Production Income.

MOVING on to PART 1 through PART 8

After you have finished the practice rounds, you will participate in Part 1 through Part 8 of the experiment. In these parts, the experimental dollars you earn from your firm's profits in each round will affect your cash earnings.

In each round of Part 1 through Part 8, you will make a Production Decision and a Technology Decision. Groups will be randomly reassigned after each part.

INSTRUCTIONS FOR PART 1-4

1. In these parts, your Firm Profit only depends on your production and technology decisions; the production and pollution generated by other firms do not affect your Firm Profit.

2. Note that the location and capacity of firms may or may not be different. The capacity of each firm is shown on the calculator. When firms have different locations, the region will be divided in 4 sub-regions by solid lines; when firms have the same location, the region will not be divided. Refer to the **Pollution Table** to see how location influences pollution. We will indicate each scenario at the beginning of each part.

3. Use the **Decision Calculator** to make more informed decisions. Although the results are for informational purposes only, the location and capacity of each firm is the same as the real decisions.

4. To make your actual decision for this round, choose a Production Decision and a Technology Decision. Once done, click “Confirm”.

5. In these parts, pollution does not affect firm profits. The more you produce, the more your firm profit will be.

In these parts: **Firm Profit = Production Income**

<u>**INSTRUCTIONS FOR PART 5-8**</u>

In these parts, an **environmental regulator** has set a **target total pollution level**. There will be a tax or subsidy based on the total pollution of your firm compared with the target level. The target will change between parts and the specific value will be shown to you.

Your profit will be adjusted by a tax or subsidy (from here on referred to as **tax/subsidy**). This tax/subsidy can be either negative (a tax) or positive (a subsidy) and is determined based on how much pollution is in the river relative to the **Target** determined by the regulator. The pollution level in the river is the aggregation of pollution from all firms. There will be a subsidy for zero concentration, but the amount of subsidy gets smaller as concentration increases. If the measured concentration level is exactly the same as the target, there will be neither a tax nor a subsidy. As concentration increases beyond the target, the tax gets larger.

Pollution in one round does not affect pollution in other rounds. However, at the end of the experiment, your earnings will be the sum of the profits you earned from all of the rounds.

In each round, you will make a Production Decision and a Technology Decision. **Total Pollution** in your group affects the profits of firms in your group.

The **Tax Payment for each firm** in your group is calculated as follows:

Total Pollution ≤ Target	Subsidy Received = Target – Total Pollution
Total Pollution > Target	Tax Payment = Total Pollution – Target

For example, if the target is set at 60, then

- If the Total Pollution in your group is less than or equal to 60, each firm in your group receives 1 experimental dollar in subsidy for every unit of total pollution under 60 units.
- If the Total Pollution in your group is greater than 60, each firm pays 1 experimental dollar in taxes for every unit of total pollution above 60 units.

The amount of the Tax/Subsidy Payment is determined by decisions of everyone in your group. Your Firm Profit in these parts will be calculated as:

If Total Pollution ≤ Target,
Firm Profit = Production Income + Subsidy Payment

If Total Pollution > Target,
Firm Profit = Production Income – Tax Payment

<u>Use the Decision Calculator to help you make more informed decisions, otherwise, you may lose money. Note that in these parts, it is not true that the more you produce, the more profit you will get.</u>

Appendix B

IRB APPROVAL LETTERS FOR EXPERIMENT IN CHAPTER 1 AND 2

UNIVERSITY OF DELAWARE

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE:	September 7, 2016
TO:	Kent Messer
FROM:	University of Delaware IRB
STUDY TITLE:	[573740-9] NEWRNet Water Quality Sensing Resolution
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVED
APPROVAL DATE:	September 7, 2016
EXPIRATION DATE:	December 3, 2016
REVIEW TYPE:	Expedited Review
REVIEW CATEGORY:	Expedited review category # *(7)*

Thank you for your submission of Amendment/Modification materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE:	September 7, 2016
TO:	Kent Messer
FROM:	University of Delaware IRB
STUDY TITLE:	[573740-9] NEWRNet Water Quality Sensing Resolution
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVED
APPROVAL DATE:	September 7, 2016
EXPIRATION DATE:	December 3, 2016
REVIEW TYPE:	Expedited Review
REVIEW CATEGORY:	Expedited review category # *(7)*

Thank you for your submission of Amendment/Modification materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.

Appendix C

IRB APPROVAL LETTER FOR EXPERIMENT IN CHAPTER 3

UNIVERSITY OF DELAWARE

RESEARCH OFFICE

210 Hullihen Hall
University of Delaware
Newark, Delaware 19716-1551
Ph: 302/831-2136
Fax: 302/831-2828

DATE:	August 27, 2012
TO:	Kent Messer, PhD
FROM:	University of Delaware IRB
STUDY TITLE:	[371015-1] Consumer's Perception on Honey Attributes
SUBMISSION TYPE:	New Project
ACTION:	APPROVED
APPROVAL DATE:	August 27, 2012
EXPIRATION DATE:	August 26, 2013
REVIEW TYPE:	Expedited Review
REVIEW CATEGORY:	Expedited review category # 7

Thank you for your submission of New Project materials for this research study. The University of Delaware IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

Please remember that informed consent is a process beginning with a description of the study and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the study via a dialogue between the researcher and research participant. Federal regulations require each participant receive a copy of the signed consent document.

Please note that any revision to previously approved materials must be approved by this office prior to initiation. Please use the appropriate revision forms for this procedure.

All SERIOUS and UNEXPECTED adverse events must be reported to this office. Please use the appropriate adverse event forms for this procedure. All sponsor reporting requirements should also be followed.

Please report all NON-COMPLIANCE issues or COMPLAINTS regarding this study to this office.

Please note that all research records must be retained for a minimum of three years.

Based on the risks, this project requires Continuing Review by this office on an annual basis. Please use the appropriate renewal forms for this procedure.