Thin and lumpy: an experimental investigation of water quality trading

Jordan F. Suter [*] John Spraggon [†] Gregory L. Poe [‡]

May 4, 2011

Selected Paper prepared for presentation at the Agricultural & Applied Economics Associations 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24- 26, 2011.

[*]Department of Economics, Oberlin College, Oberlin, OH.
[†]Department of Resource Economics, University of Massachusetts- Amherst, Amherst, MA.
[‡]Department of Applied Economics and Management, Cornell University, Ithaca, NY.

Thin and lumpy: an experimental investigation of water quality trading

Jordan F. Suter [*] John Spraggon [†] Gregory L. Poe [‡]

May 4, 2011

Early Draft Please Do Not Cite

Abstract

Water quality trading schemes in the United States can predominantly be characterized by low trading volumes. In this paper we utilize laboratory economics experiments to explore the extent to which the technology through which pollution abatement is achieved influences market outcomes. Mirroring the majority of water quality trading markets, the sessions utilize small trading groups composed of six participants. To understand the extent to which abatement technology influences trading behavior, the experimental treatments vary the degree of heterogeneity in initial abatement costs and the potential for long-lived investments in cost-reducing abatement technology.

1 Introduction

Although the number of water quality trading schemes has increased substantially over the last two decades, nearly all active trading programs to date have been characterized by low trading volumes. Data from a comprehensive summary by Breetz et al. (2004) highlights over seventy proposed or active water quality trading programs in the United States, many of which are still under development. Of the nineteen active trading programs only four had experienced more than three trades (Morgan and Wolverton 2005). A more recent worldwide study of twenty six active trading programs by the World Resources Institute (Selman et al. 2009) found that only two programs, the Long Island Sound Nitrogen Credit Exchange Program in Connecticut and the Hunter River Salinity Trading Scheme in Australia, have exhibited continuous trading activity since their inception.

[*] Department of Economics, Oberlin College, Oberlin, OH.

[†] Department of Resource Economics, University of Massachusetts- Amherst, Amherst, MA.

[‡] Department of Applied Economics and Management, Cornell University, Ithaca, NY.

In this paper we utilize laboratory economics experiments to investigate the role of technological impediments to active emission credit trading. Mimicking the realities of most point-source water quality markets, our setup features a market composed of a small number of participants for which pollution reductions are realized through a combination of continuous, variable cost abatement and discrete capital investments in improved abatement technology. The consideration of both variable and fixed costs is a divergence from textbook examples pollution trading, which focus only on 'short-run' marginal cost trading in robust markets. Reflecting this focus, the majority of previous pollution trading experiments in the literature have ignored the role of fixed-cost, capital investments and long-run decision making.

Elsewhere, it has been argued that the limited number of observed trades in actual water quality markets can be attributed to a lack of a binding cap on emissions (Selman et al. 2009), lack of regulatory coverage (Faeth 2006), or imposed market structures (Woodward and co-authors 2002a, b; 2003). Individual water treatment plants may choose to over-comply and not trade due to demographic pressures (Earnhart 2004a, b), to allow for a margin of safety to account for variability in hydrological and other conditions (Bandyopadhyay and Horowitz 2006), or to protect opportunities for future growth (Hamstead and BenDor 2010). While there are undoubtedly a number of reasons for the observed lack of trading, our experiments focus on the existence of discrete capital costs and a limited number of traders as contributing factors. Both of these issues were recently raised in a study of municipal wastewater treatment plants in the Upper Passaic Watershed, N.J. (Sado, Boisvert and Poe 2010).

With relatively thin markets in specific watersheds, we conjecture that three interrelated technological features strongly impede trading. First, wastewater treatment technologies are relatively homogeneous, being limited to well-established chemical or biological treatment methods. Ceteris paribus, such homogeneity limits the gains that can be realized from trading. Second, most regulated firms cannot achieve the required pollution standard given their current abatement technology. In small markets, firms that do not upgrade may thus be subjected to monopoly powers by sellers or face a substantial penalty for noncompliance. Third, to achieve large scale emissions reductions, firms must make long-lived capital investments to upgrade their abatement technology. Hence, cost savings and trading opportunities are conditional on the allocation of investments across firms. These features are consistent with the observed overcapitalization in watersheds with trading programs (Hamstead and Ben-Dor, 2010), with too many firms investing in costly abatement technology and the absence of cost-saving trading activity. More generally this combination of features raises questions on the extent to which standard market trading can optimally allocate capital investment costs across traders.

An empirical analysis of active trading programs could provide insight into the challenges associated with thin markets and lumpy investments; however the overall lack of trades across nearly all trading programs makes this type of analysis difficult. We therefore turn to the experimental laboratory to provide feedback on the relative severity of the impediments described above. In the experiments, markets are composed of six participants that take on the role of point source emitters. Participants decide whether to pay the fixed cost necessary to upgrade to a technology that reduces their costs of abatement, which if adopted remains in place for the duration of the trial. Following the technology decision, participants choose a level of abatement and participate in a double auction where pollution credits can be purchased and sold.

While the experiments are primarily motivated by water quality trading, the results may also be relevant to on other trading scenarios characterized by thin markets and fixed costs associated with long-lived abatement technologies. A better understanding of the causes of inefficiencies in such markets will then allow for the development of improved trading regimes that allow buyers and sellers to better coordinate their capital decision making.

2 Relevant literature

In this section, we review several economic studies that utilize data from experiments that incorporate discrete "technology" choices in a pollution credit trading framework.

The most relevant studies in this literature come from Ben-David et al. (1999; 2000) who implement experiments that replicate features of the SO2 permit trading market to investigate the role of firm abatement cost heterogeneity. In the experiments, participants have an opportunity to choose between three irreversible abatement technologies after each trading period. Although including a technology decision makes these studies similar to our experiments, the technology investment itself is not a treatment variable. The authors instead focus their analysis on the influence of firm heterogeneity, uncertainty, and price dispersion on observed social efficiency. The authors find that increased cost-function heterogeneity reduces trade volume and efficiency. They also find that uncertainty does not influence trade volume or prices, potentially because participants use a "wait and see" strategy before making irreversible investments.

Two additional studies also investigate the effect of investment decisions on outcomes in emissions trading markets. A study by Camacho-Cuena, Requate and Waichman (2010) explores how the type of permit auction (ascending clock and double auction) influences investment and abatement decisions. The authors find that the "cleanest" firms tend to over-invest in abatement technology while the "dirtiest" firms tend to under-invest and that this is invariant to the type of auction mechanism. Experiments by Gangadharan, Farrell and Croson (*forthcoming*) explore permit banking behavior when firms have an opportunity to make investments in abatement technology. In one treatment, participants can make a one-time investment that allows them to produce goods more efficiently for the remainder of the treatment. The authors observe significant overinvestment in technology and overbanking relative to efficient levels. Importantly, the authors find that the combination of investment opportunities and banking lead to the most inefficient outcomes.

3 Theory

The theoretical model is motivated by the regulation of wastewater treatment facilities operating in a relatively small watershed. In this section we introduce both the social planner's problem and the problem faced by individual firms seeking to reduce emissions at least cost.

To begin, assume that n firms are operating in a watershed and that each firm accepts the waste of households and converts the waste into effluent, which enters a common waterway. It is assumed that the quantity of waste that a firm accepts is not a choice variable and does not influence the concentration of pollution in the effluent. The concentration of pollution in a firm's effluent, hereafter referred to as "emissions", is determined by the level of abatement effort that the firm engages in. The choice variable for firm i therefore is its level of abatement, a_i, which is related to its level of emissions, e_i, by the function $a_i = e_i^0 - e_i$, where e_i^0 is the status quo level of emissions for the firm.

The variable cost of abatement for firm i is determined by the function $C_{il}(a_i)$. The subscript l represents the particular abatement technology employed by firm i and we assume that $l = \{1, 2\}$. For each technology type, $C_{il}(0) = 0$ and it is further assumed that $C'_{il}(a_i) > 0$ and $C''_{il}(a_i) > 0$. To upgrade from technology $l = 1$ to technology $l = 2$ a firm must incur a fixed cost, K. The final assumption is that $C_{i1}(a_i) \geq C_{i2}(a_i)$ for all given values of a_i, which implies that the variable cost of abatement is higher with technology 1 compared to technology 2.

Assuming uniform mixing, the total quantity of ambient pollution in the water resource is given by $\sum_{i=1}^{n} e_i^0 - a_i$. Suppose that a social planner is interested in reducing pollution to some quantity G, which is less than $\sum_{i=1}^{n} e_i^0$, at least cost over T periods. The cost minimization problem can be formally stated as

$$Min \sum_{i=1}^{n} \left[\sum_{t=0}^{\tau-1} \rho^t C_{i1}(a_{it}) + I_i[\rho^\tau K_{i\tau} + \sum_{t=\tau}^{T} \rho^t C_{i2}(a_{it})] \right] \tag{1}$$
$$\text{s.t.} \quad \sum_{i=1}^{n} e_i^0 - a_{it} \leq G \qquad \forall \quad t = 0, ..., T,$$

where I_i is an indicator variable that is equal to one if firm i upgrades to $l = 2$ and zero otherwise and τ represents the time period in which firm i upgrades. Solving the social planner's problem is challenging because of the discrete nature of the technology decision.

Although a closed form solution to the mixed integer program above cannot be defined, the proposition below is based on the theoretical model.

Proposition 1 *If it is socially optimal for at least one firm to upgrade, it is always optimal for the upgrade to occur in period zero.*

Proof of Proposition 1 Without loss of generality, assume that $n = 1$ and $T = 1$. If the firm does not upgrade in either period, the present value of their costs across the two periods will be $D = C_{01} + \rho C_{11}$, where the first term in the subscript is the period (0 or 1) and the second term is the technology (1 or 2). If the firm upgrades in period zero, their discounted costs are $E = K + C_{02} + \rho C_{12}$. Finally, if the firm upgrades in period one, their discounted costs become $F = C_{01} + \rho K + \rho C_{12}$.

Now, suppose that it is optimal for the firm to upgrade in period one. Then it must be the case that $F < Min(D, E)$. From the condition that $F < D$ we get that $K < C_{11} - C_{12}$. From the condition that $F < E$ we get that $K > \frac{1}{1-\rho}(C_{01} - C_{02})$. However this brings us to a contradiction, since it is not possible for $C_{11} - C_{12} > \frac{1}{1-\rho}(C_{01} - C_{02})$ given that $C_{11} - C_{12} = C_{01} - C_{02}$.

In theory, a simple emissions permit trading system *could* induce optimal upgrade and abatement choices in which some firms upgrade their technology and sell permits to the remaining firms that do not undertake costly upgrades. With a small number of firms operating in a particular watershed, however, there is a great deal of coordination required to induce optimal investment. For example, if the penalty for noncompliance is high enough firms may be unwilling to take a chance that there will be sufficient permits for sale for them to achieve the emissions standard. The result would be overcapitalization, where too many firms engage in costly upgrades.

A firm operating in this market would seek to minimize costs according to

$$Min \sum_{t=0}^{\tau-1} \rho^t C_{i1}(a_{it}) + I_i[\rho^\tau K_{i\tau} + \sum_{t=\tau}^{T} \rho^t C_{i2}(a_{it})] + \sum_{t=0}^{T} \rho^t p_t(W_{it} - (e_i^0 - a_{it})) \qquad (2)$$
$$\text{s.t. } e_i^0 - a_{it} \leq W_{it} + m_{it} - y_{it} \qquad \forall \quad t = 0, ..., T.$$

where m_{it} represents the number of permits purchased by firm i and y_{it} is the number of permits sold by firm i. When equation (2) is summed across the n firms in the watershed, the result is equivalent to the social planner's problem in equation (1). This equivalence depends on the assumption that $\sum_{i=1}^{n} W = G$ and that $\sum_{i=1}^{n}[W_{it} - (e_i^0 - a_{it})] = 0$ in each time period.

Despite the functional equivalence, individual upgrade decisions may diverge from social efficiency under the permit trading policy for a number of reasons. The most obvious cause for a divergence stems from the fact that the price and availability of future permits is not known with certainty prior to firms making their upgrade decisions. Firms that are risk averse may upgrade rather than subject themselves to the chance that the price of permits will be excessively high or not available at all. Further, if there is a limitation on the quantity of abatement that firms with technology 1 are able to undertake, there is a risk that the supply of permits available at any price will be insufficient for them to meet the permit constraint unless they upgrade.

While uncertainty with respect to the price and supply of permits may lead to excessive upgrades, there is also an option value associated with waiting to upgrade that stems from the irreversible nature of the investment decision. In other words, there is a benefit to a firm associated with waiting to observe the number of firms that upgrade, since the number of firms that upgrade will influence the market price of permits. By not upgrading in period zero, a firm preserves its ability to upgrade in future periods when it can observe the number of other firms that upgrade and have a more accurate estimate of the price of permits. Finally, upgrade and abatement decisions may diverge from social efficiency if the discount rates of individual firms are different from the rate used by the social planner.

4 Experiment setup

To investigate the potential for a permit trading system to induce least cost upgrade and abatement behavior, we conduct a series of laboratory economics experiments to simulate the market environment described above. In this section we begin by outlining the baseline experimental setup and then describe the experimental treatments that are implemented.

Although we discuss the setup here with an environmental context, the instructions for the experiment (included in the appendix) are framed neutrally.

In the experiment, participants are randomly and anonymously placed into groups of six ($n = 6$). Each participant has one of two abatement cost functions, either "Type A" or "Type B". The marginal cost of abatement for Type A is defined as $MC_{i1}(a_i) = a_i^3$ while the marginal cost of abatement for Type B is $MC_{i2}(a_i) = a_i^2$. Therefore, abatement is more costly for Type A participants than it is for Type B participants. Abatement units are defined discretely on the interval $\{0,8\}$ and tokens are used as the currency within the experiment, which are exchanged for US Dollars at a rate of 200 tokens = 1 US Dollar.

Each experimental session consists of two distinct stages. Stage one implements a baseline scenario that is identical across all treatments and allows participants to become familiar with the trading framework. In each stage one round, participants are randomly assigned either Type A or Type B abatement costs. Participants undertake two practice rounds where there are 3 Type A and 3 Type B participants. There are then five actual rounds, and across the five rounds there are 1, 2, 3, 4, and 5 Type A participants (and the corresponding number of Type B participants), in random order.

In stage two of the experiment, participants undertake one five-round practice trial and then three actual five-round trials. The groups of six are reshuffled prior to the beginning of each of the trials. The trials in stage two of each session implement one of four treatment scenarios described below. The same treatment is played in each of the three actual trials within each session.

The social objective in each round is to reduce pollution from the status quo (zero abatement) of 48 to a target of 24. To achieve the desired reduction, a permit trading program is implemented wherein four permits are distributed to each of the six firms and each permit allows the firm to emit one unit. The permits are fully transferable, so that a firm can choose to emit fewer than four units and sell permits to firms that emit more than four units. Permits cannot be banked for use in future rounds.

The transfer of permits occurs in each round through a standard double auction mechanism. In the double auction, firms submit bids to purchase permits and offers to sell permits. Trades are consummated whenever a participant chooses to buy a permit for a price that has been offered or chooses to sell a permit for the highest posted bid price. Note that although it is possible for any given firm to be either a seller or a buyer of permits, it is socially optimal for Type A participants to be buyers of permits from Type B participants.

4.1 Treatments

The four treatments that are implemented vary the initial number of firms with upgraded technology, the maximum abatement possible for Type A firms, and the potential for firms Type A firms to upgrade. Treatment 1 serves as a baseline and implements a scenario with 4 Type A participants and 2 Type B participants. In this treatment, firms cannot upgrade and the objective of the treatment is to ensure that with cost function heterogeneity, firms are able to meet the pollution objective at least cost.

In treatment 2 there are again 4 Type A participants and 2 Type B participants at the beginning of each trial. In this treatment, however, Type A participants can upgrade to Type B at any point in the trial. To upgrade to Type B, a participant must pay a fixed cost of 400 tokens ($K = 400$). The abatement technologies are long-lived, so once the upgrade occurs the participant is Type B for the remaining rounds in the trial. This treatment provides a simple test of the extent to which the choice of technology influences permit prices and trading volumes.

Treatment 3 implements a scenario where, at the beginning of each trial, there are 6 Type A and 0 Type B participants and participants can upgrade at any time. This treatment is meant to test the extent to which inefficient upgrade behavior is a function of coordination challenges on the part of participants. Finally, in treatment 4 there are initially 6 Type A and 0 Type B participants, participants can upgrade at any time, and Type A participants can only abate three units of pollution without upgrading to Type B. This allows us to measure the extent to which technological constraints influence trading behavior. For example, firms may choose to undertake costly upgrades to eliminate the possibility that there will not be a sufficient supply of permits for them to purchase on the open market. Table 1 below provides a summary of the treatments described above.

Table 1: Treatment summary

	Initial # firms with tech. 1	**Initial # firms with tech. 2**	**Can firms upgrade?**	**Do firms face limited abatement?**
Treatment 1	4	2	No	No
Treatment 2	4	2	Yes	No
Treatment 3	6	0	Yes	No
Treatment 4	6	0	Yes	Yes

To implement the limited abatement treatment, we assume that if Type A firms do not acquire at least one permit (to bring them up to a total of five permits), then they will face a penalty of 300 tokens for every permit less than five that they hold. In other words, they pay the abatement costs associated with abating 3 units (36 tokens) and must pay an additional

300 tokens per permit below five that they hold.

4.2 Cost Scenarios

This section utilizes the parameterization employed in the experimental sessions to calculate the total costs of achieving the pollution standard and the individual incentives facing participants. To begin, we calculate the cost minimizing abatement choices necessary to achieve the pollution standard, contingent on the number of firms that upgrade. This is done by comparing the marginal costs of abatement for the firms that upgrade and those that do not. Variable costs are calculated by summing the marginal costs of abatement, starting from the lowest marginal cost units and proceeding until the standard is met. All upgrades are assumed to occur in the first round.

As an example, suppose three firms upgrade in round one. The firms that upgrade will optimally abate five units while the three firms that do not upgrade optimally abate two units. The sum of the abatement costs across the six firms and the five rounds of the experiment are equal to 1,365 while the fixed cost of the three firms that upgrade is 1,200. This implies that the minimum total cost for the six firms over five rounds is equal to 2,565 if three firms upgrade. Table 2 enumerates the minimum total cost of abatement contingent on the number of firms that upgrade.

Table 2: Comparison of compliance costs

Tech. 1 firms	**Units abated per firm**	**Tech. 2 firms**	**Units abated per firm**	**Variable cost**	**Fixed cost**	**Total cost**
6	4	0	-	3000	0	3000
5	3 (x3) 4 (x2)	1	7	2240	400	2640
4	3	2	6	1630	800	2430
3	3	3	5	1365	1200	2565
2	2	4	5	1190	1600	2790
1	2	5	4 (x3) 5 (x2)	1045	2000	3045
0	-	6	4	900	2400	3300

Based on the results in table 2, the cost minimizing means of abating 24 units in each of the five rounds is for two firms to upgrade to Type B in round one and abate six units, while the four firms that do not upgrade abate three units. The table provides a measure of the potential inefficiency associated with the upgrade decisions. For example, even if abatement decisions are made optimally, the total costs of achieving the pollution standard if all six firms upgrade will be more than 35% higher than if two firms upgrade.

The cost minimizing solution does not vary across treatments, which allows for relatively easy efficiency comparisons. Although table 2 focuses on the upgrade decision, inefficiency can enter in three ways; (1) More or fewer than two firms upgrade, (2) Firms that upgrade do not upgrade in round one, (3) Firms make inefficient abatement choices contingent on the number of firms that upgrade.

As an example of the costs of inefficient abatement choices, suppose that two firms upgrade in round one (optimal upgrade behavior). Instead of the firms that upgrade abating six units and the remaining firms abating three units, suppose that all firms abate four units in each round. The total fixed costs obviously remain the same, but the variable costs across the five rounds increase from 1,630 to 2,300. The total costs of abatement in this case therefore increase by nearly 30%.

An individual firm-decision maker will weigh the costs to the potential benefits of upgrading. If the firm ignores the potential for trades to occur and simply makes the upgrade decision by comparing the reduction in variable costs to the fixed cost, then it would never choose to upgrade. The total variable costs to a firm of not upgrading and abating four units in each of the five rounds is 500 tokens. The total variable cost to a firm that upgrades is 150 tokens. Given the 400 token fixed cost of upgrading the total cost of upgrading is 50 units higher than not upgrading.

The number of permits that are transacted and the expected price of the traded permits are a function of the number of firms that upgrade to Type B. The exact price of the permits that are transacted cannot be predicted with precision, but will instead depend on bargaining between firms. In the table below, we provide the theoretical minimum and maximum price of permits contingent on the number of firms that upgrade.

The predicted number of units abated by each firm is determined by ranking the marginal cost of abatement of all firms from lowest to highest. Units are abated starting from the lowest marginal cost and continuing until the standard is achieved. The minimum permit price is then equal to the highest marginal cost among the units abated (i.e., the last unit abated). If the price were to go below this point, then the firm would prefer to purchase an additional permit rather than abate that unit and the demand for permits would be greater than the supply. The maximum price is equal to the marginal cost of the first unit that is not abated. If the price were to exceed this level, then the firm would prefer to abate that unit and the supply of permits would be greater than the demand.

As expected, table 3 reveals that both the minimum and maximum price decrease as the number of firms that upgrade increases. To illustrate how the benefits of upgrading depend on the number of firms that upgrade, we compare the costs to a firm when it is the only firm that upgrades to the costs to an upgraded firm when five firms upgrade. In the case where only one firm upgrades, the variable cost of abatement, plus the fixed cost, minus the revenue from the sale of permits is 140 tokens. For the firms that do not upgrade, the cost

Table 3: Predicted permit prices

# of firms that upgrade	# of Permits transacted per round	Minimum price	Maximum price
0	0	-	-
1	3-4	64	64
2	4	36	49
3	3	27	36
4	4	25	36
5	2	25	25
6	0	-	-

of abatement, plus the cost of permits is 500 tokens. When five firms upgrade, the cost of abatement, plus the fixed cost, minus the revenue from permit sales for a firm that upgrades is 675 tokens. This is compared to the total cost to the firm that does not upgrade of 295.

5 Results

The experimental sessions were completed at the Cleve E. Willis Experimental Economics Laboratory at the University of Massachusetts Amherst in the Spring of 2011. Each of the four treatments were completed in one session with 24 undergraduate participants, making the total number of participants equal to 96. In this section we summarize the results from the sessions by focusing on four linked outcomes of interest; upgrade behavior, transaction volume, transaction prices, and cost effectiveness.

5.1 Upgrade behavior

In each of the four treatments, the cost-minimizing outcome requires four type A and two type B participants. Further, to minimize costs, all upgrades must occur prior to the first of the five trading rounds. Treatment 1 implements the optimal upgrade scenario with two type B participants and does not allow further upgrades. Table 4 below provides the average number of observed type B participants across the four groups and three trials by treatment in each of the five rounds.

The results in Table 4 illustrate three clear outcomes. First, in each of the three treatments where participants are able to upgrade, the number of upgrades exceed the cost-minimizing level of two even in the first round (p ¡ 0.025 in each case). The second result from Table 4 is that the initial number of type B participants and whether or not the type A participants have limited abatement potential does not have a significant impact on upgrade behavior. There are no significant differences in the number of type B participants across the three

Table 4: Mean number of upgrades by treatment and round

	Treatment 2	**(S.E.)**	**Treatment 3**	**(S.E.)**	**Treatment 4**	**(S.E.)**
Round 1	3.33	(0.139)	2.92	(0.215)	2.92	(0.547)
Round 2	3.50	(0.220)	3.33	(0.241)	3.25	(0.449)
Round 3	3.58	(0.163)	3.50	(0.220)	3.50	(0.531)
Round 4	3.67	(0.241)	3.50	(0.220)	3.50	(0.531)
Round 5	3.67	(0.241)	3.50	(0.220)	3.58	(0.510)
N	300					
R^2	0.632					

treatments in any of the five rounds.

The third result related to upgrade behavior is that there is a small but significant increase in the number of type B participants between the first round and the fifth round. Across the three treatments there is an average increase of 0.5 in the number of type B participants. This suggests that there is some additional inefficiency that enters from firms choosing to upgrade after the first round.

In addition to exploring treatment effects, we also investigate the role that individual participant characteristics have on upgrade decisions. To accomplish this, we utilize two additional sources of information. First, we implement a post session questionnaire, which collects demographic and scholastic information from each participant. Second, we have each participant undertake a standard post-session risk preference test. We provided real financial incentives for participants by paying two subjects in each session based on the role of a die, which indicated the payoff scenario that they would be compensated for.

The results from the individual-level upgrade decision model is provided below in table 5. The dependent variable in the model is simply a binary outcome of whether or not the participant has chosen to upgrade by the end of the trial. Given that there are three trials in each session, the standard errors are clustered by subject. We utilize a simple linear probability model for ease of interpretation, as the coefficients represent how a particular characteristic influences a participant's probability of upgrading.

The participant-level results reveal some interesting characteristics that play a role in upgrade behavior. The starkest results are for year in school and gender. Freshmen are 57% more likely than seniors to upgrade from type A to type B, while female participants are 21% more likely to upgrade than their male counterparts. Interestingly, risk preference also influences upgrade decisions, while the number of economics classes and the number of prior experiments that a participant has undertaken do not have a significant influence. Based on the results, risk aversion is negatively correlated with decisions to upgrade. It appears that

Table 5: Upgrade decision- individual effects model

Variable	Coefficient	(Std. Err.)
Female	0.210**	(0.094)
Freshman	0.571***	(0.088)
Sophomore	0.175	(0.145)
Junior	0.271**	(0.115)
# of econ classes	-0.008	(0.009)
# of previous experiments	0.023	(0.023)
Level of risk aversion	-0.063**	(0.026)
Intercept	0.847***	(0.220)
N	216	
R^2	0.166	
$F_{(7,71)}$	11.202	

participants perceive the decision to upgrade decision as risky. This story could reasonably be told both ways. Upgrading to type B reduces the risk of high prices in the permit market. Upgrading is, however, risky in that participants must pay a high fixed amount and hope that there will be a sufficient number of buyers to allow them to recoup their investment in the lower variable cost technology.

5.2 Transaction volume

Based on the observed low numbers of actual transactions in most water quality trading markets, one motivation for the experimental design is to better understand the degree to which available abatement technologies influence transaction volumes. The predicted number of transactions in the experiment is a function of the number of participants that choose to upgrade, as shown in Table 3. If none of the participants choose to upgrade or if all of them upgrade, then there should not be any transactions, as abatement costs are homogeneous and there are no gains from trade. If fewer than five participants upgrade then there should be, at most, four permits transacted. Finally, if five of the six participants choose to upgrade, then there should be two permits that transact.

Table 5 provides the mean number of transactions per round in each of the treatments. The results appear to reveal a treatment effect on transactions, as the number of transactions in treatment 3 is significantly higher than in treatment 2. This apparent treatment effect, however, is likely caused by the idiosyncratic differences in participant behavior across the two treatments rather than a true treatment effect. Evidence for this assertion comes from the fact that the number of transactions in treatment 3 is also significantly higher than in treatment 2 in Part A of the experiment, where the scenarios are identical.

Table 6: Transaction volume

	Coefficient	**(Std. Err.)**
Treatment 1	6.48	(0.496)
Treatment 2	4.68	(0.637)
Treatment 3	7.20	(0.525)
Treatment 4	5.52	(0.740)
N	240	
R^2	0.883	

A second important observation from Table 5 is the total number of transactions that take place. Across the four treatments there are nearly six transactions per round on average, which is a 50% higher volume than is predicted by theory under even the most favorable trading conditions. This surprising result suggests that participants were attempting to gain extra revenue by engaging in arbitrage, even when few arbitrage opportunities exist. Indeed, participants both bought and sold permits in 28% of the rounds in Part B of the experiment.

Allowing participants to both buy and sell in the same round is a design component that we ultimately decided to allow because we assumed that it would help rationalize the market. Surprisingly, participants seeking arbitrage opportunities did not alter their strategies over time, as the number of transactions does not significantly diminish over rounds.

To get a better idea for what may be driving these "arbitragers" we regress the number of transactions made by round on the same set of participant information used in Table 4 above. The results presented in Table 6 suggest that experience may play a role in participant behavior. Both the number of previous experiments that a participant has undertaken in the past and to a lesser extent the number of economics classes that a participant has taken reduce the number of observed transactions.

5.3 Prices

The observed transaction prices from the experimental sessions reveal important information about the effect of upgrade decisions on market outcomes. The greater the number of participants that choose to upgrade, the greater the supply of permits on the market and the lower the predicted price at which permits transact. In addition, observed market prices provide a basis for participants in making future upgrade decisions. If permit prices are relatively high, this suggests higher gains associated with upgrading and becoming seller of permits.

The mean prices presented in Table 8 reveal three outcomes of interest. First, in each of the treatments there is an inverse relationship between the number of type B participants and the mean price at which permits trade. This result is in line with predictions and reinforces

Table 7: Transaction volume model, individual effects

Variable	Coefficient	(Std. Err.)
Female	0.410**	(0.199)
Freshman	0.288	(0.383)
Sophomore	0.198	(0.240)
Junior	0.157	(0.263)
# of econ classes	-0.032*	(0.018)
# of previous experiments	-0.122**	(0.041)
Level of risk aversion	-0.047	(0.075)
Intercept	2.309	(0.625)
N	1440	
R^2	0.059	
$F_{(7,95)}$	3.218	

Table 8: Mean prices by treatment and number of type B participants

# Type B	Tr 1	(S.E.)	Tr 2	(S.E.)	Tr 3	(S.E.)	Tr 4	(S.E.)
2	48.79	(3.21)	48.00	(24.87)	48.54	(6.90)	70.09	(7.18)
3	-	-	39.01	(4.79)	44.24	(5.86)	54.49	(5.30)
4	-	-	32.96	(4.54)	34.57	(5.08)	42.56	(6.65)
5	-	-	24.08	(17.58)	29.55	(11.12)	29.36	(9.40)
N	299							
R^2	0.734							
$F_{(15,284)}$	52.26							

the idea that the returns to upgrading diminish with the number of participants that have already upgraded.

The second outcome of interest from the transaction prices is that in treatments 1 - 3 mean prices, conditional on the number of Type B participants, are not statistically different. The prices fall in the upper range or slightly above the theoretical predictions made in Table 3. Despite this outcome, prices in treatments 2 and 3 are not high enough to provide an explanation for the excessive number of upgrades that are observed. It can be shown that when there are two type B participants, a participant with type A can only recoup the fixed cost of upgrading if the price of permits is at least 50 tokens. Mean prices never exceed 50 in treatments 2 and 3, although the difference is not large.

The final outcome of interest with respect to prices is that prices in treatment four are significantly higher than prices in the first three treatments when there are only 2 or 3 type B participants. Recall that in treatment 4, type A participants have limited abatement choices and are penalized if they do not purchase a permit from a type B participant. The threat of the penalty gives type B participants extra market power and results in higher observed transaction prices. Prices in Treatment 4 are high enough to explain the excessive number of upgrades that occur in that treatment.

5.4 Cost effectiveness

Markets for water quality will only be judged as a success if water quality objectives are achieved and there are realized cost savings associated with trading. In the experiments that we conduct, water quality objectives are assumed to be achieved and therefore we focus here on the cost-effectiveness of the market outcomes. Specifically, we measure the degree of inefficiency of the observed market outcomes by calculating the percent by which observed costs exceed the minimum possible cost. Formally inefficiency is measured as,

$$Inefficiency = \frac{(ObsCost - MinCost)}{MinCost} = \frac{(UpCost - MinCost)}{MinCost} + \frac{(ObsCost - UpCost)}{MinCost}. \quad (3)$$

In equation (3), there are two sources of inefficiency which additively determine the total inefficiency measure. The variable *UpCost* in equation (3) represents the cost to the group if trading is perfectly efficient, conditional on the number of type B participants. The first component in equation (3) therefore represents inefficiency stemming from the upgrade decision, while the second component represents inefficiency that results from trading. A separate component of inefficiency for upgrades that occur after the first round could be included separately, but late upgrades account for only about a one percent increase in costs and so are lumped into the upgrade inefficiency measure. As a point of comparison, if there are exactly two type B participants in the group and no trading occurs, then upgrade inefficiency is zero and trading inefficiency is 27.6%.

Table 9: Observed cost inefficiency

	Inefficiency	**(S.E.)**	**Upgrade Ineff.**	**(S.E.)**	**Trade Ineff.**	**(S.E.)**
Treatment 1	0.237	(0.027)	-	-	0.237	(0.027)
Treatment 2	0.174	(0.027)	0.118	(0.023)	0.055	(0.011)
Treatment 3	0.300	(0.043)	0.109	(0.017)	0.191	(0.060)
Treatment 4	0.263	(0.038)	0.119	(0.044)	0.144	(0.056)
N	60		60		60	
R^2	0.514		0.490		0.429	

The inefficiency results reveal that on average, total costs are nearly 25 percent higher than the minimum possible cost across the four treatments. Slightly less than half of this inefficiency, in treatments 2 - 4, is a result of the excessive number of upgrades that occur. The fact that the upgrade inefficiency does not vary significantly across treatments 2 - 4 is not surprising given that upgrade behavior is nearly identical in these treatments. It is surprising, however, that there is so much variation in trading inefficiency. In the first treatment, where there is no potential for upgrades, costs are 23.7% higher than the cost minimizing outcome, strictly because of inefficient trading. In fact, the trade inefficiency is not significantly different from the case where no trading occurs at all. The inefficiency associated with trading is significantly lower in treatment 2 than in any of the other treatments, although it is not entirely clear what drives this result. There is an inverse relationship between the volume of transactions reported in Table 6 and the trading efficiency reported above. It appears that the participants that were bent on engaging in arbitrage served to drive up the cost of the market outcomes.

6 Discussion

While water quality markets have been advanced as a cost-effective way to achieve water quality objectives, the market environment is often vary different that the textbook examples on which with trading programs are advocated. With only a few firms operating in the market (thin) and cost savings largely driven by large capital investments (lumpy), the efficiency of market outcomes is likely to be case specific. To date, only a handful of experiments have investigated trading in small markets where heterogeneity is largely driven by investment decisions. In this paper, we present the results of market experiments with groups made up of six traders. Importantly, we vary the ability of participants to invest in lower variable cost technology, the initial number of firms with the lower cost technology, and the nature of the higher cost technology.

The findings from our experiments should be seen as a first step towards understanding behavior in thin and lumpy markets. Overall we find a degree of over-confidence in upgrade and trading decisions that permeates all of our treatments. While this was not intended to be a "bubble" experiment (e.g. Smith et al. 1988), our results appear to exhibit some of this bubble-type behavior. Investment and trading volume appear to be correlated, at least in part, to what falls under the category of experience, as the number of upgrades and the number of trades that a participant undertook was negatively correlated with their class standing and the number of past economics experiments that they had participated in.

In addition to pervasive over-capitalization and excessive trade volume relative to theoretical predictions, we also find more predictable outcomes with respect to the price at which the permits transact. Specifically, the greater the number of firms with the lower-cost technology, the lower the mean permit price. Further, mean trading prices are very consistent across treatments, with the exception of the treatment in which participants with the high variable cost technology are penalized if they did not purchase at least one permit. In this treatment, prices are significantly higher when only two or three participants have chosen to upgrade to the lower-cost technology, as these participants gain market power.

There are a number of future directions for this research to head. A first consideration is the fact that it is unlikely that experienced wastewater treatment plant operators would engage in excessive attempts at arbitrage. One way to limit arbitrage in the experimental sessions would be to forbid both buying and selling permits in the same round. Eliminating speculation might then cause traders to focus more on making only beneficial trades. A second potential direction to explore would be to make the upgrade decision sequential rather than simultaneous. In the current design the number of participants with the lower cost technology is a treatment condition, but we could also implement a treatment where participants can observe the upgrade decisions of other market participants in real time.

The ultimate objective of this research is to improve the cost effectiveness of water quality trading programs. Therefore using these baseline data, we hope to investigate the role of institutions in altering upgrade and trading behavior in thin and lumpy markets such as those that categorize most water quality trading programs.

References

Bandyhyopadhyay, S. and J. Horowitz, 2006. "Do Plants Overcomply with Water Pollution Regulations? The Role of Discharge Variability." *Topics in Economic Analysis and Policy*, 6(1):1-30

Ben-David, S., D. S. Brookshire, S. Burness, M. McKee, and C. Schmidt. 1999. Heterogeneity, Irreversible Production Choices, and Efficiency in Emission Permit Markets. *Journal of Environmental Economics and Management.* 38, 176-194.

Breetz, H.L., K. Fisher-Vanden, L. Garzon, H. Jacobs, K. Kroetz, R. Terry, 2004. *Water Quality Trading and Offset Initiatives in the U.S.: A Comprehensive Survey.* Hanover: Dartmouth College. (http://www.dartmouth.edu/~kfv/waterqualitytradingdatabase.pdf)

Camacho-Cuena, E. , T. Requate, I. Waichman, 2010. Investment Incentives under Emission Trading: An Experimental Study. Working Paper. January 2010.

Earnhart, D., 2004a. "Regulatory Factors Shaping Environmental Performance at Publicly-Owned Treatment Plants." *Journal of Environmental Economics and Management*, 48:6555-681.

Earnhart, D., 2004b. "The Effects of Community Characteristics on Polluter Compliance Levels." *Land Economics*, 80(3): 408-432.

Faeth, P., 2006. "Point-Nonpointillism: The Challenges that Water Quality Trading Faces and What We Might Do About It." Paper presented at the Second Nation Water Quality Trading Conference, May 23-25 Pittsburgh, PA.

Gangadharan L., A. Farrell and R. Croson. Investment Decisions and Emissions Reductions: Experimental Results in Emissions Permit Trading, in The Handbook of Experimental Economics and the Environment, M. Price and J. List, editors, *forthcoming*.

Hamstead, Z.A., and T.K. BenDor, 2010. "Over-compliance in Water Quality Trading Programs: Findings from a Qualitative Case Study in North Carolina." *Environment and Planning C: Government and Policy* 28(1): 1-17.

Morgan, C., and A. Woverton, 2005. "Water Quality Trading in the United States." *NCEE Working Paper Series* 200507.

Sado, Y., R.N. Boisvert and G.L. Poe, 2010. "Potential Cost Savings from Discharge Permit Trading: A Case Study and Implications for Water Quality Trading." *Water Resources Research* 46, W02501.

Selman, M., S. Greenhalgh, E. Branosky, C. Jones, and J. Guiling, 2009. "Water Quality Trading Programs: an International Overview." *WRI Issue Brief* 1: 1-15.

Woodward, R.T., 2003. "Lessons about Effluent Trading from a Single Trade." *Review of Agricultural Economics* 25(1):235-245.

Woodward, R.T. and R. A. Kaiser, 2002. "Market Structures for U.S. Water Quality Trading." *Review of Agricultural Economics* 24(2):366-383.

Woodward, R.T., R.A. Kaiser, and A-M B. Wicks, 2002. "The Structure and Practice of Water Quality Trading Markets." *Journal of the American Water Resources Association* 38(4):967-979.

Appendix (Instructions)

Introduction

This experiment is a study of how people make economic decisions. If you follow these instructions and make careful decisions, you will earn money that will be paid to you privately in cash at the end of the session. The amount you earn will depend both on the decisions that you make and the decisions of others. You will also earn a show-up fee of $5. Your earnings will be measured in tokens which will be added up over the course of the experiment. At the end of the experiment you will be paid in cash at the rate of **200 tokens = $1**. Please do not communicate with other participants in the experiment. If you have questions regarding these instructions, raise your hand and a monitor will come by to answer your questions.

In the experiment you will buy and sell coupons over a series of rounds. In each round, you will be part of a group made up of six participants. You will not know the identity of the other participants in your group. Your earnings in each round are made up of three components, which will be described in detail below. In particular, round earnings will be determined as follows:

Round Earnings = Coupon Revenue + Proceeds from Coupon Sales – Cost of Coupon Purchases

Your **Coupon Revenue** will be determined by the number of coupons that you hold at the end of the round. Each participant will **start each round with four coupons** and can adjust the number of coupons they hold by buying and selling them in a market that will operate over the computer network. The details of how to buy and sell coupons will be described shortly. In general, the more coupons you hold, the higher your Coupon Revenue.

In each round of the experiment you will be either **Type A** or **Type B**. The Coupon Revenue associated with different quantities of coupon holdings is different for Type A and Type B. In general, coupon revenue is lower for Type A compared to Type B. The coupon revenue for Type A and Type B participants associated with each coupon quantity can be found in the "Coupon Revenue Sheet", which has been provided to you. The sheet also indicates the **change** in coupon revenue that comes from increasing or decreasing your coupon holdings by one coupon.

Prior to the beginning of each trading round, you will be told whether you are Type A or Type B. You will also be told the total number of Type A and Type B participants in your group of six. An example of the screen indicating your type is provided below in Figure 1.

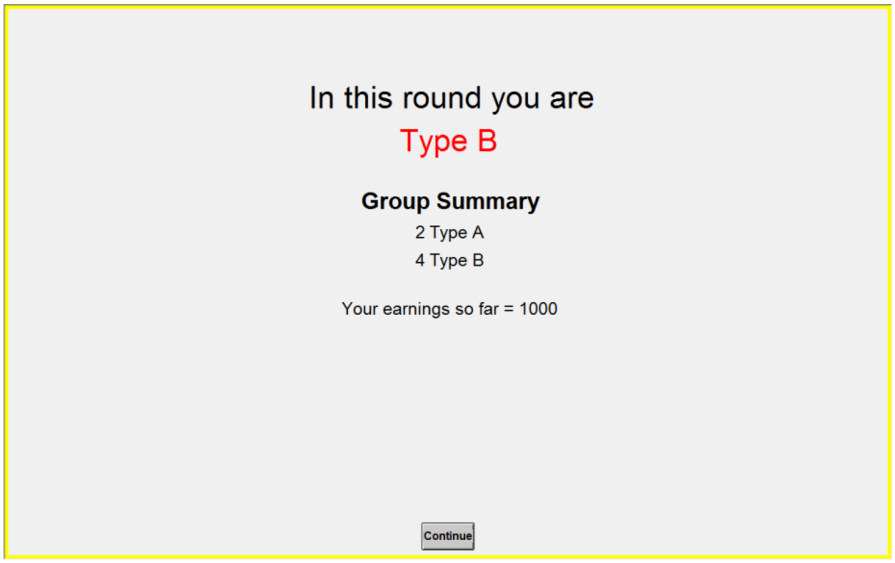

Figure 1

Buying and Selling Coupons

Each participant will start each round with four coupons and can adjust their coupon holdings by buying and selling them in a market. If you sell coupons your earnings will increase by the sale price and if you buy coupons your earnings will decrease by the sale price.

Why might you want to buy a coupon? Remember that your Coupon Revenue increases when you hold more coupons. If, for example, you are Type A and currently hold 4 coupons, then your current coupon revenue is 150 [all quantities below are in terms of tokens]. If instead you hold 5 coupons, your coupon revenue will increase to 214. Therefore if you can buy a coupon for less than 64 [214 – 150 = 64], it will allow you to increase your earnings for that round. For example, if you bought a coupon for 40, your round earnings will increase from 150 to 214 – 40 = 174.

Why might you want to sell a coupon? Suppose that you are Type A and currently hold 6 coupons. If you sell a coupon then your coupon revenue will decrease from 241 to 214. Therefore, if you can sell a coupon for more than 27 [241 – 214 = 27], the sales revenue that you gain will exceed the coupon revenue that you lose going from 6 to 5 coupons. For example, if you sell a coupon for 50, your round earnings will increase from 241 to 214 + 50 = 264.

In each round you will have **1 minute** to make transactions. In a given round it is possible for you to both buy coupons and/or sell coupons. Note that coupons **cannot** be carried over between rounds. Figure 2 shows an example of the trading screen.

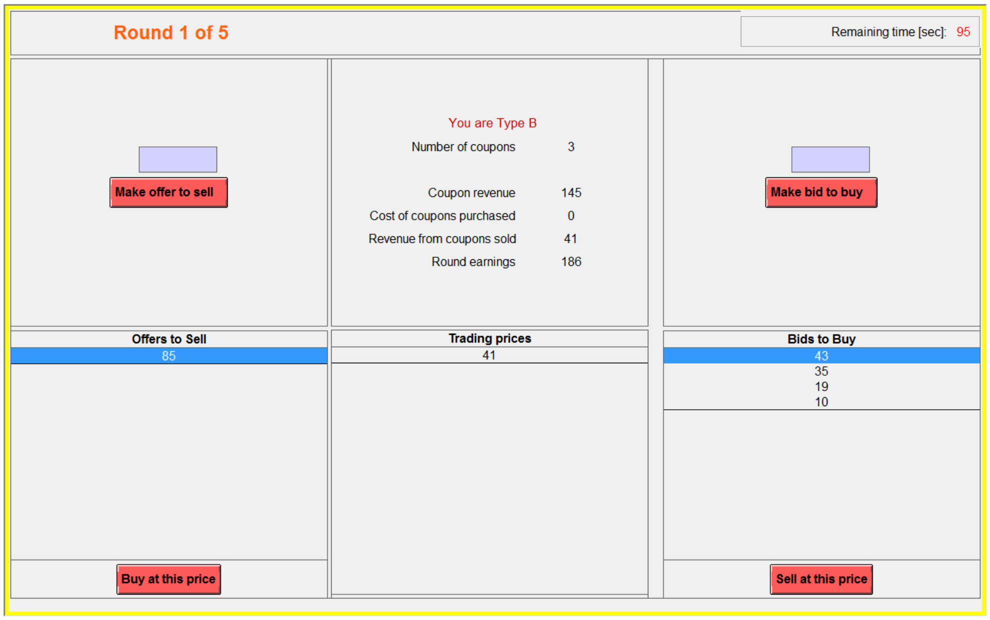

Figure 2

There are two ways that you can **buy coupons**.

1. Participants interested in buying a coupon can submit bid prices by typing their bid into the box on the right side of the screen, and then clicking the "**Make bid to buy**" button. The bid is immediately displayed to all six members of your group in the lower right part of the screen labeled "**Bids to Buy.**" Once a bid is submitted, anyone wishing to sell a coupon at this price can accept the bid. Such an acceptance results in an immediate trade at that price. Sellers prefer higher prices, so any new bids must be higher than the current highest bid.

2. The offer prices from participants interested in selling coupons will be listed in ascending order in the "**Offers to Sell**" section on the left. In addition to posting bids to buy, as described above, anyone wishing to buy a coupon can accept the lowest offer price by simply clicking the "**Buy at this price**" button on the bottom left of the screen.

There are also two ways that you can **sell coupons**.

1. Participants interested in selling a coupon can submit offer prices by typing their offer into the box on the left side of the screen and then clicking on the "**Make offer to sell**" button. This offer price is immediately displayed to all six members of your group in the lower left part of the screen labeled "**Offers to Sell**." Once an offer is submitted, anyone wishing to buy a coupon at this price can accept the offer. Such an acceptance results in an immediate trade at

that price. Buyers prefer lower prices, so any new offers to sell must be lower than the current lowest offer.

2. Bid prices from participant interested in buying coupons will be listed in descending order in the “Bids to Buy” column on the far right. In addition to posting Offers to Sell, as described above, anyone wishing to sell a coupon can accept the highest bid price by simply clicking the “**Sell at this price**” button on the bottom of their screen.

Once a transaction is made, the accepted bid or offer will be displayed in the “**Trading prices**” list in the center of your computer screen. Note that when you buy or sell a coupon, all of your previous bids and offers will disappear from the screen.

Round Results

At the end of the round a summary of the results are provided, as shown in Figure 3. The results screen provides information on your earnings for the round as well as the total amount that you have earned up to that point in the experiment. Once you have reviewed this information, please click “continue” to begin the next round.

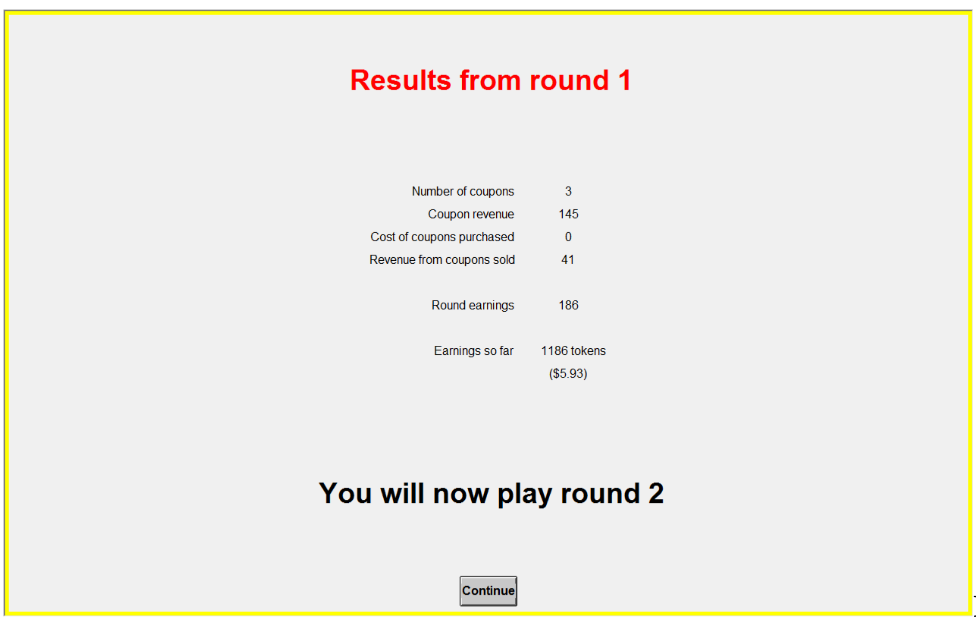

Figure 3

To begin, you will participate in **two practice rounds** that will allow you to get comfortable with the trading software. **The practice rounds do not affect your earnings in any way**. In each of the practice rounds there will be three Type A and three Type B participants. Whether you are type A or Type B will be determined randomly by the computer software.

After the practice rounds, you will play **five actual rounds for real money**. The number of Type A and Type B participants in your group will be different in each round. You will play rounds with 1, 2, 3, 4 and 5 Type A participants. The order of these rounds is randomly determined. Whether you are Type A or Type B in a given round will also be determined randomly. In each round, every participant has an **equal probability** of being either Type A or Type B.

You will begin the first actual round with earnings equal to the show up fee of 1000 tokens = $5. Groups will be randomly shuffled prior to the first actual round and will stay the same for each of the five actual rounds. After the fifth round, you will receive additional instructions.

Summary

- Each participant begins each round with 4 coupons.
- Your Coupon Revenue increases the more coupons you hold.
- In each round you will be Type A or Type B with equal probability.
- Your Round Earnings = Coupon Revenue + Proceeds from Sales – Cost of Purchases.
- You will play 2 practice rounds and then 5 actual rounds for real money.
- Prior to the first actual round you will be paired with a new group of six participants.

Quiz: Please circle the correct answer.
A monitor will come around shortly to check your answers and answer any questions you have.

1. Suppose that you are **Type A**. If you hold 3 coupons, what is your **coupon revenue** for the round?

   (a) -191 (b) 25 (c) 150 (d) 195

2. Suppose that you are **Type B** and that you hold 4 coupons. By how much will your **coupon revenue increase** if you acquire a fifth coupon?

   (a) 16 (b) 64 (c) 214 (d) 236

3. Suppose you are **Type A**. During the trading period you sell one coupon at a price of 50 tokens and do not purchase any coupons. What are your **Round Earnings**?

   (a) 25 (b) 75 (c) 200 (d) 264

4. Suppose you are **Type B**. During the trading period you sell one coupon at a price of 50 tokens and a second coupon at a price of 20 tokens. What are your **Round Earnings**?

   (a) 70 (b) 189 (c) 159 (d) 229

5. Suppose you are **Type A**. During the trading period you sell one coupon at a price of 50 tokens and then are able to buy a coupon at a price of 20 tokens. What are your **Round Earnings**?

   (a) 180 (b) 250 (c) 30 (d) 55

Part 2 Instructions

In Part 2 you will play **three trials** and each trial will be made up of **five rounds**. In each round you will make trades in the same way as you did in Part 1. There are, however, two differences. The first difference is that, at the **beginning of each trial** all six participants in your group will be **Type A**.

The second difference is that each Type A participant will have the opportunity in each round to **upgrade** to Type B. The cost to upgrade to Type B is **400 tokens**. If you choose to upgrade you will be Type B for the remaining rounds of the trial and will only pay the upgrade cost once. For example, if you choose to upgrade in round 2 you will pay 400 tokens in round 2 only and will be Type B for rounds 2 – 5. The upgrade decision screen is shown in Figure 4 below.

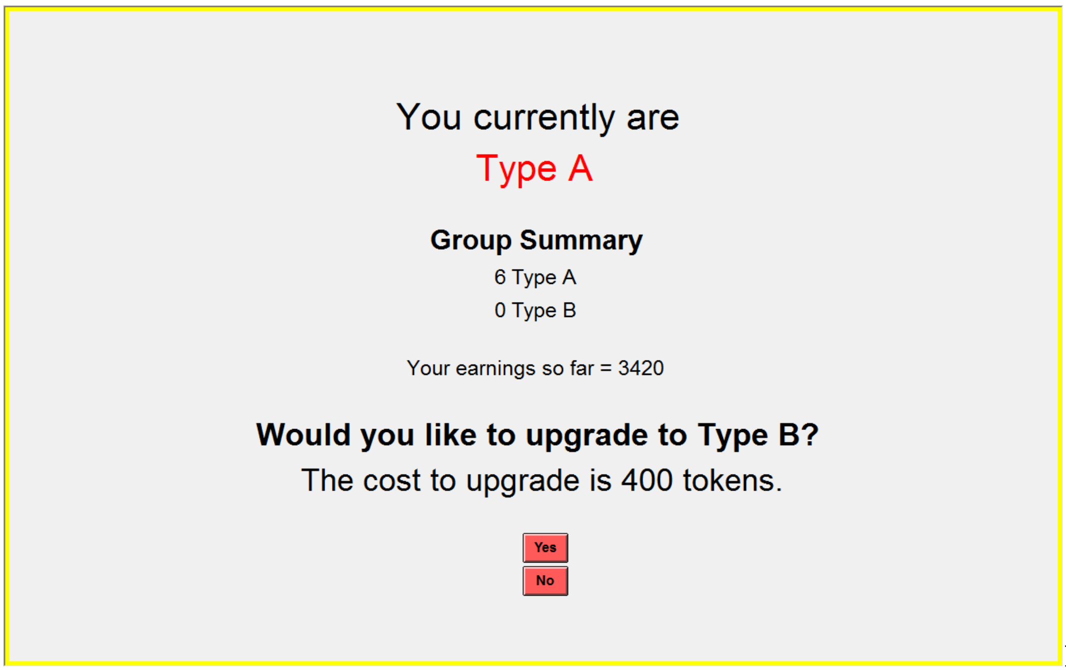

Figure 4

You will have up to 30 seconds in which to make your upgrade decision prior to each trading round. After making your upgrade decision you will see a screen that indicates the total number of Type A and Type B participants in your group.

You will begin Part 2 by participating in one five-round practice trial that will not affect your earnings. After completing the practice trial, you will then play **three five-round trials for real money**. Groups will be randomly reshuffled prior to the beginning of each trial.

If your cumulative earnings fall below 0 tokens you will be excused from the remainder of the trial that you are currently participating in. You will then begin the next trial with 200 tokens in earnings.

Summary

- At the beginning of each trial, all participants will be Type A.
- In each round, Type A participants will have the opportunity to pay 400 tokens to upgrade to Type B.
- If you upgrade to Type B, you will be Type B for the remaining rounds of the trial.
- You will play 1 practice trial and then 3 actual 5-round trials for real money.