*Title Page (With Author Details)

Title: Experimental Tests of Water Quality Trading Markets

Running Title: Tests of Water Quality Trading Markets

Authors: Luke R. Jones[1] and Christian A. Vossler[2*]

Acknowledgement: We thank Christopher Clark and Michael Price, and two anonymous referees, for their helpful comments and suggestions.

[1] Department of Economics and Finance, Valdosta State University, Valdosta, GA 31698. Email: lukjones@valdosta.edu.

[2] Department of Economics and Howard H. Baker Jr. Center for Public Policy, University of Tennessee, Knoxville, TN 37996. Email: cvossler@utk.edu.

[*]Corresponding Author. Complete mailing address is 916 Volunteer Blvd, 523 Stokely Management Center, Knoxville, TN 37996. Telephone: 865-974-1699. Fax: 865-974-4601.Email: cvossler@utk.edu.

Experimental Tests of Water Quality Trading Markets

Abstract

Many watersheds in the U.S. have established water quality trading programs in hopes of achieving cost-effective reductions in water pollution; however, the success of these programs has been limited. This study highlights some of the unique features of water-based credit trading markets that may explain the lack of success, and uses laboratory experiments to isolate their effects. In particular, we compare two forms of a baseline-and-credit institution, a Pigouvian tax/subsidy regulation, and – characteristic of air quality programs – a textbook cap-and-trade regulation. Across these institutions we examine the effects of abatement technology adoption. We find that a baseline-and-credit program, when it requires firms to make upfront investments to generate tradable credits, is less efficient than cap-and-trade and tax/subsidy institutions. Furthermore, we find that when efficient trading requires costly technology adoption, institutions that involve inter-firm trading, including cap-and-trade, are less efficient than the tax/subsidy.

Keywords: Water quality trading; baseline-and-credit; cap-and-trade; technology adoption; laboratory experiments

JEL classification: C91; C92; D80; L50; Q58

1. Introduction

Surface water pollution remains a prominent issue worldwide, including in many developed countries. Recent assessments indicate a substantial fraction of U.S. surface waters are still too impaired to support their designated uses (U.S. EPA, 2011). In order to meet water quality goals, policymakers have increasingly endorsed adoption, at the watershed level, of water quality trading (WQT) programs. This endorsement is no doubt related to the success of high-profile air quality trading programs such as the sulfur dioxide (SO_2) allowance trading component of the U.S. Acid Rain Program, and the EU Emissions Trading Scheme. As early as 2004, there were more than 70 WQT programs in some phase of development in the U.S., about twice as many as there were in 1999 (Breetz et al., 2004; Environomics, 1999).

Despite financial and political support, the numerous WQT initiatives established for U.S. watersheds have recognized little success. A recent report showed only 100 facilities had engaged in trade, with 80% coming from the Long Island Sound Trading Program (U.S. EPA, 2008).[1] In this study, we use laboratory experiments to investigate institutional and structural features of WQT markets that may help explain their lack of success. In particular, we look at the effects of using a "baseline-and-credit" system of allocating permits, requiring firms to have already undertaken and verified abatement beyond their baseline to generate tradable credits, and the performance of institutions when costly abatement technology adoption is necessary to achieve efficiency.

The WQT institutions in practice differ markedly from the textbook cap-and-trade institution that typifies air quality markets. The majority of WQT programs involve baseline-

[1] As we discuss below, the trading scheme associated with Long Island Sound differs substantially from the vast majority of credit trading markets in that "trades" are made with the regulator rather than through inter-firm trading.

and-credit trading institutions, wherein polluters have an emissions baseline and tradable credits are linked to emission reductions beyond this baseline (Breetz et al., 2004; Environomics, 1999).[2] That is to say, in contrast to cap-and-trade programs, there is no initial allocation of credits via free distribution, auction or otherwise. This approach is seemingly at odds with market fundamentals given inherent uncertainty over market prices and quantities, especially in the context of new markets. As highlighted by Cason and Plott (1996), the Clean Air Act Amendments governing the SO_2 market emphasize the importance of ensuring permit availability and providing clear price signals, and require the EPA to run annual auctions to help achieve this goal. This concern is exacerbated for WQT programs, which are often characterized by few (potential) buyers and sellers.

Regulators have discretion in defining how the generation of credits is determined. Some programs allow firms to receive credits based on proposed or estimated reductions. More common, however, are strict requirements that firms establish a reduction has occurred prior to receiving credits. Fundamental to this, firms face a risky financial decision as they must undertake costly activities (e.g. abatement, monitoring, and documentation) prior to the realization of credit demand. The theoretical analysis of Mailath et al. (2004) provides formal evidence that such investment activities are likely to be suboptimal whenever "people make current investment decisions whose returns depend upon future prices that are not currently contractible (pg.2)." Despite these potential concerns, the EPA guidelines (2007) strongly endorse this approach:

[2] A tradable right to emit an amount of a pollutant is typically referred to as a permit or allowance in the context of cap-and-trade, and is referred to as a credit when discussing baseline-and-credit institutions. For transparency, this paper uses the term "credit" regardless of the institution being discussed.

"A basic premise of water quality trading is that credits should not be used before the time frame in which they are generated. In general, a permitting authority should not allow for a pollutant reduction credit in a NPDES permit on the basis of the *proposed* treatment by another point source or an *unverified* commitment to install a BMP by a nonpoint source and their anticipated pollutant reduction (pg.34)."

In addition to baseline-and-credit institutional features, another potential impediment to efficient operation of WQT markets are the upfront costs associated with abatement technology adoption. It is typical in WQT markets that in order to reduce emissions below baseline and realize abatement cost advantages, firms must make investments in abatement technology (e.g. install a filtration device). As discussed by a number of sources, these investments often involve large fixed costs and result in substantial increases in abatement capabilities (Sado et al., 2010; Caplan, 2008; Boisvert et al., 2007; EPA 1996). Boisvert et al. (2007) and Sado et al. (2010) suggest that these fixed technology costs may lead to underinvestment in abatement technology and thus decrease WQT market performance.

Although baseline-and-credit WQT institutions are the most prevalent, some programs are fashioned instead after the Long Island Sound Trading Program – one of the few WQT success stories – which does not involve credit trading in a conventional sense.[3] In trading jargon, polluters that exceed their baseline "buy" credits and those with emissions below their baseline "sell" credits. However, the "trading" is with the regulator who automatically buys and sells credits at a pre-announced price at the end of a monitoring period. Further, there is no budget-balance condition: permits sold can exceed or fall short of permits purchased. Thus, this

[3] As stated in U.S. EPA (2008), "[s]ome [WQT] program interviewees noted that their program lacks the defining features of trading (e.g., buyers and sellers, credits) and felt that EPA and others may apply the term too freely (pg.3-3)."

mechanism is more accurately described as a Pigouvian tax/subsidy than a market-trading institution.[4] Because the tax/subsidy involves no financial risk due to market uncertainty, it provides an important yardstick from which to compare inter-firm trading mechanisms.

To explore the effects of different institutional features and costly technology adoption in WQT markets we use laboratory experiments, which have been a primary method of investigating the institutional features of cap-and-trade air pollution markets. Experiments are potentially informative for WQT program design, given (1) very few existing programs have recorded trades; and (2) identification issues arise from the many unobservable or difficult to measure factors in play (e.g. abatement costs; transaction costs; information; trading opportunities). In particular, our experiments allow us to compare tax/subsidy, baseline-and-credit (with and without binding abatement pre-commitment), and textbook cap-and-trade institutions, across treatments with and without fixed abatement technology costs.

It is worth noting that the primary inquiry in these experiments can be broadly characterized as an examination of how agents respond to risk that must be undertaken in order to trade efficiently in an emissions market, where the designed source of variation in this risk is the upfront investment required to generate (additional) credits. Upfront investment of this nature is introduced in our experiments when credit generation requires a binding, abatement pre-commitment and/or costly technology adoption, and is financially risky whenever uncertainty exists about credit demand. Hence, risky upfront costs for credit generation appear in the baseline-and-credit institution with abatement pre-commitment, and – when there are fixed abatement technology costs – in cap-and-trade and both baseline-and-credit institutions without

[4] Other examples of this include the Neuse River Basin Nutrient Sensitive Waters Management Strategy and Tar-Pamlico Nutrient Reduction Trading Program.

abatement pre-commitment.[5] The tax/subsidy institution incurs no financial risk regardless of whether there are fixed technology adoption costs, as this institution guarantees that an agent can buy or sell credits at an exogenous and known fixed price. By examining investment behavior and market performance across these treatment conditions we are able to analyze the ways that different sources of sunk investment costs endemic to WQT programs may be impacting their success.

2. Related Experiments

Though the experimental literature on emissions trading is vast (see Bohm, 2003 and Muller and Mestelman, 1999 for reviews), the majority of these involve cap-and-trade, and few have focused on issues related to WQT programs such as baseline-and-credit trading or abatement technology adoption. In a series of papers, Buckley et al. (2006, 2008, 2011) have compared cap-and-trade with baseline-and-credit institutions. The main distinction in this work is that they investigate a long-run setting where the baseline-and-credit "baseline" is an emissions intensity target tied to output, and agents can alter their emissions through adjusting output. Theoretically, this leads to higher aggregate emissions in the baseline-and-credit institution, and this result is confirmed experimentally. In concurrent experimental work specific to WQT, Ghosh, Kwasnica and Shortle (2011) study point-nonpoint pollution trading in a setting

[5] Note that in cap-and-trade, and baseline-and-credit without abatement pre-commitment, binding investments in technology are actually made during a trading period, as opposed to beforehand. However, given the parameters of our experiment, in general an agent is not expected to be able to recover the full fixed costs of technology adoption from the transaction that triggers the technology change. Therefore, technological investment still constitutes a risky investment whose profitability depends on conducting subsequent trades at sufficiently high prices.

where buyers have market power, and their experimental evidence favors the safety-first institution proposed by Ghosh and Shortle (2010) versus a trading-ratio-based market institution.[6]

Recent studies by Gangadharan et al. (2013) and Suter et al. (2013) investigate the impact of abatement technology adoption within the context of cap-and-trade markets.[7] In these studies, participants can make an irreversible abatement technology investment; specifically, for a one-time fixed cost, the marginal abatement cost curve is lowered for the remainder of a sequence of trading periods. Gangadharan et al. (2013) examine abatement technology investment when participants have the ability to bank permits. Suter et al. (2013) explore abatement technology investment along with the impact of initial abatement cost heterogeneity and limits on abatement capacity. Both studies find that participants tend to overinvest when given the opportunity to upgrade abatement technology. Suter et al. (2013) also find that risk aversion is important in explaining the decision heterogeneity across players. This effect is strongest when a participant has limited abatement capacity, in which case not investing exposes the participant to non-compliance fines if sufficient permits cannot be secured through trade.

[6] There are many features of the trading institutions studied by Ghosh, Kwasnica and Shortle (2011) that differ from our own. Although their setting can be characterized as baseline-and-credit in the sense that there is no initial permit allocation and an emissions cap, buyer and seller roles are clearly defined, only buyers face an emissions cap, and abatement and credit holdings are determined simultaneously through an iterated call market.

[7] We note that other studies, notably Ben-David et al. (1999, 2000), Cason et al. 2003, and Camacho-Cuena et al. (2012), incorporate abatement technology investment in their experimental designs, but their designs do not allow one to isolate the effects of this feature.

Relative to the above two studies, aside from exploring technology adoption for different market institutions, our experimental design differs in two important ways. First, technology adoption does not carry over into future periods (i.e. we implement a repeated but static game), allowing for greater learning opportunities. Second, technology adoption is characterized as a fixed cost that allows agents to achieve higher levels of abatement not otherwise obtainable, and this is portrayed as movement along the same marginal cost curve. Our characterization of abatement technology investment is similar to Cason et al. (2003), who tailored their experimental design based on firms in Port Phillip Watershed, and congruent with the theoretical analysis of Caplan (1998) and others. Although we do not model this interaction in our design per se, this type of investment decision could be viewed as a decision to pay a non-point firm to engage in abatement, leading to credits beyond what the point source could achieve with its existing abatement technology.

3. Experimental Design

The experimental design considers the tax/subsidy (TS, hereafter) mechanism and a prototype cap-and-trade (CAT) institution where all available credits are initially distributed free of charge. In addition, we consider two forms of a baseline-and-credit (BAC) mechanism. In both forms, agents pre-commit to an abatement level, and excess abatement generates tradable credits. To test the effect of having firms undertake costly actions prior to credit trading, in BAC1 the pre-commitment is nonbinding whereas it is binding in BAC2; in other words, the pre-commitment is "cheap talk" in BAC1 but has direct financial consequences in BAC2. These four trading mechanisms interacted with the presence ("Tech") or absence ("NoTech") of costly technology adoption cost yields eight treatments in our experimental design, which are summarized in Table 1. Each replication of the experiment is characterized by an eight-player group participating in a common market over a sequence of 10 trading periods.[8] There are four replications of each treatment, except for BAC2-Tech for which there are six.[9] The basic features of the design in terms of player types, abatement cost schedules, framing, and trading interface are loosely based on the cap-and-trade experiment of Cason and Gangadharan (2006). For clarity of exposition we describe the experiment in the context of *emissions* trading, although such framing was absent from the experiment instructions and software.

The basic decision setting for all treatments is as follows. In each round of the experiment, players receive an endowment ("initial earnings", in the instructions) of money and (possibly) emission credits ("coupons"). They face a regulation ("production rule"), which mandates that particular combinations of abatement ("production") and credit holdings be realized. This regulation is automatically enforced; there is no opportunity for noncompliance.[10] Players can alter their required abatement through credit trading; however, the specifics of how credits are initially distributed and traded differ across institutions. A player's earnings for a decision period are equal to her endowment minus abatement cost (including fixed abatement technology costs in the Tech treatments) and the cost of any credits purchased, plus earnings from any credits sold.

[8] For convenience, we use trading terminology when broadly characterizing the experiment, although we acknowledge that the TS treatments do not involve trading in a conventional sense.

[9] Consistent with List et al. (2010) the additional sessions of BAC2 were motivated by the higher variance in key outcomes realized for this treatment.

[10] Although compliance in emissions trading programs is an important issue examined in a number of studies (e.g., Cason and Gangadharan, 2006; Murphy and Stranlund, 2007; Stranlund et al., 2011), this feature is beyond the scope of the current paper.

3.1. Experiment parameters and theoretical predictions

Each player has (unregulated) emissions of 10 units, which means 80 units for the eight-player group. Regulation coincides with group emissions of 48 units, a 40% reduction. To provide incentives for trade, within each group there are four player "types" that differ in terms of initial credit allocation (or, equivalently, regulated emissions baseline), abatement cost schedule, and when applicable, fixed technology costs. Two players are randomly assigned to each type, and this role assignment is fixed for the duration of the experiment. In equilibrium, two of these types are buyers and two are sellers. However, buyer and seller roles are not explicitly assigned: each player is allowed to buy and sell credits. Initial credit allocations (emissions baselines), and abatement cost schedules for the four types are shown in Table 2.

The column labeled *MC* in Table 2 denotes the marginal abatement cost schedules, while the column labeled *FC* provides the fixed technology costs in play for the Tech treatments. For the Tech treatments, for simplicity, the fixed technology cost is automatically determined by the number of units abated. Lower levels of abatement are associated with a fixed technology cost of 100 lab dollars, while higher levels of abatement are associated with a cost of 300. The abatement level where this technology switch occurs varies between types. In the NoTech treatments, the fixed technology cost is zero for all levels of abatement, and there is no mention of different technologies in the instructions.

Given the parameters in Table 2, the equilibrium respectively has type 1s and 2s each buying three and four credits, and type 3s and 4s each selling four and three credits, for a total of 14 trades at an equilibrium price in the interval [220, 240], and potential gains from trade of 2400 at the group-level. In reaching the efficient allocation, type 1s and 4s remain with their initial technologies, whereas type 2s and 3s switch technologies. Our design is such that, under standard assumptions, the theoretical predictions are the same for all eight treatments: all four institutions are theoretically efficient both with and without fixed technology costs. However as we have previously conjectured, key differences between the institutions or/and the impacts of fixed technology costs may lead to empirical inefficiencies.

Expected gains from trade are 300 for all types under the Tech scenario. In the NoTech scenario type 1s and 4s have gains of 300, while type 2s have gains of 100, and type 3s have gains of 500. As a deliberate design choice, we adjust endowments and lab-to-US$ exchange rates such that – for a particular type – both expected earnings and marginal incentives are approximately equal across Tech and NoTech conditions.[11] This allows a more careful identification of the effects of risk that stem from institutional features; otherwise, any observed technology effect may simply be an artifact of reducing the available gains from trade.[12] For all player types under all treatments, for each decision period, if a player does not trade she would earn $0.40 whereas under equilibrium trading she would earn $2.40. Thus, there are large financial gains from trade. We now describe particulars on how the different market institutions are implemented.

3.2. Implementation of trading institutions

[11] In the Tech scenario the exchange rate is 150:1 for all types. In the NoTech setting, exchange rates are 150:1 for type 1s and 4s (who are not expected to switch technology), 50:1 for type 2s (who are expected to adopt the costly technology), and 250:1 for type 3s (who are expected to switch to the least costly technology).

[12] Of course, relative to the case where exchange rates were held constant, our estimated technology effect is a lower-bound measure (i.e. is conservative).

TS. With the TS mechanism, agents directly choose a level of abatement between one and 10 units. The relationship between the abatement choice and the emissions baseline directly determines the number of credits purchased (sold). Credits are automatically bought or sold (i.e. exchanged with the experimenter) at a fixed price of 230, which represents the midpoint of the equilibrium price interval. Thus, for example, since a type 1 player has baseline emissions of four (so six units of abatement are required in the absence of “trading”), if she chooses to abate three units she automatically would purchase three credits at the price of 230. If she instead chooses seven units of abatement, she would sell one credit at 230. There is no interaction between group members or feedback on the actions of others. Further, there is no aggregate constraint on the number of credits that can be bought or sold – there can be a net imbalance.

CAT. All available credits are initially freely distributed to players. The regulation is framed as a rule that abatement and credit holdings (at the end of the trading period) sum to 10. Trading takes place via a computerized double auction. Each player enters the market with her credit allocation, and then can trade credits with other group members. After the market closes, each player’s abatement is automatically determined according to the rule. For example, if a player with a credit allocation of four is a (net) buyer of one credit, then she would automatically abate five units. If she did not alter her initial credit holding, she would abate six units.

BAC1. Since there is no initial credit allocation, the regulation is instead framed as one where each player faces an emissions baseline (“initial production requirement”), and the final abatement level is adjusted up or down depending on the number of credits sold or bought. A decision period proceeds in two stages. In the first, players propose a *nonbinding* level of abatement between one and 10. For each unit of abatement that would lead to a reduction below a player’s emissions baseline, this generates a tradable credit. A player can instead opt out of the proposal and thus receive zero credits. In the second stage, players trade credits through a computerized double auction, as in CAT. As credits are only generated through the proposal, these are the only credits available to trade. At the end of the trading period, the player's abatement is automatically determined by her permit holdings. Overall, BAC1 mimics CAT with the exception of how the initial distribution of credits is determined.

BAC2. As in BAC1, each period in BAC2 consists of two stages. The fundamental difference between BAC1 and BAC2 is that the first stage abatement proposal is *binding*. That is, if a player commits to, for example, abating three units in excess of her baseline, she must do so regardless of whether she is able to sell all three credits in the (second stage) market. If a player does not propose excess abatement, she generates zero credits, and her actual abatement is determined – as in the other treatments – according to the regulation based on her emissions baseline and final credit holdings.

Double auction market. In the CAT, BAC1, and BAC2 treatments, players trade credits in a computerized double auction. Each trade is for one credit and all players can submit offers to sell and bids to buy, subject to standard improvement rules, as well as accept standing offers and bids. The improvement rule is that a new offer (bid) must be lower (higher) than a standing offer (bid). Only the most favorable offer and bid are displayed on all players' trading screens. When a player accepts an offer or a bid, the transaction occurs immediately, and the current offer and bid are cleared from the trading screen. While there is no explicit restriction on the number of transactions a player can make, the software is programmed to prevent players from reducing abatement below one unit. The trading screen, in addition to standing bids and offers, displays the history of (within-period) transaction prices and buyer and seller IDs, the player's current period earnings, current earnings from market trade, current level of required abatement, and current credit holdings, all of which are automatically updated as transactions occur. Also displayed is the time remaining before the market closes. The market is open for three minutes in the first two periods of each session and then is reduced to two minutes and 30 seconds for the remaining periods.

4. Experiment Participants and Procedures

We conducted 17 sessions between April and June of 2011, with either one or two groups (i.e. markets) in each session.[13] With the exception of the TS treatments, treatments were run between-subject. We ran two sessions for the TS treatments, with each session consisting of two groups and lasting 20 periods.[14] In one of these sessions, players faced the NoTech scenario in the first 10 periods, and then the Tech scenario in the second 10 periods. In the other session, this order was reversed. In these TS sessions, player types were randomly reassigned and the groups randomly re-matched after the first 10 periods. Although the number of periods is predetermined in all sessions, players were not informed of the number of periods until the experiment was completed.

A total of 240 players, recruited from the undergraduate student population at the University of Tennessee, Knoxville participated in the experiments. The experiments took place in a designated experimental laboratory. The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007). Upon entering the lab, players were randomly assigned

[13] Specifically, there were 13 two-group sessions, and four one-group sessions.

[14] This procedural variance is based on cost considerations, given that TS decision periods are much quicker (there is no market trading) and that there are no interactions between players (such that players rather than markets are the independent observational unit).

to individual computer stations separated by dividers, and were provided a calculator, scratch paper, and a paper copy of the experiment instructions.[15]

In each session, players first participated in a series of 11 risk preference elicitation tasks, similar to Holt and Laury (2002). In each task, participants were asked to choose between receiving a fixed payoff of $3 and a binary lottery. The binary lottery involved high and low payoffs of $5 and $0.50, with the probability of the high payoff increasing as the participant progressed through the tasks. The earnings from the risk elicitation, based on playing out a randomly chosen task, were not announced until the end of the session. Instructions for the lotteries were read aloud by an experiment moderator while players followed along on their copy. After the lottery, instructions for the WQT experiment were administered in the same manner, and questions were answered. Two quizzes were included as part of the instructions to help facilitate understanding of the experimental constructs and market institution. To encourage players to consider the quizzes carefully, they were incentivized with US$1 paid per quiz if all questions were answered correctly. Players participated in one or three unpaid practice periods, respectively, for the NoTech and Tech treatments. This design choice was motivated by the results of two pilot sessions and, if anything, works against our finding a technology effect.[16] After the practice period(s) players were given a final opportunity to ask questions before

[15] Representative instructions are available in an online appendix.

[16] Specifically, in the pilots (each with four markets), participants traded in double auctions facing the same parameters as in the reported experiment. The lone difference was that there was no notion of credits nor any regulation governing them (i.e. this was a more generic double auction experiment). In the pilot without technology cost, efficiency stabilized beginning with the second period, whereas stability was reached in the fourth period in the pilot with technology cost.

beginning the experiment. Following the market experiment, players completed a short demographic questionnaire. Players were then paid their earnings privately, in cash. Sessions lasted between 90 and 105 minutes, and players earned an average of about $25 (not including earnings from the quiz questions), with a range of $10 to $82.

5. Results

Differences in the trading programs we study can be described by the presence or absence of particular institutional features. Specifically, the TS institution differs from the other three in the sense that the latter institutions involve inter-firm trading through the double auction market and therefore uncertainty about market credit demand. The two BAC treatments differ from CAT in that credit allocations are requested rather than endowed. BAC2 differs from BAC1 in that credit requests bind players to specific levels of abatement prior to market trading. The analysis focuses on the marginal effects of changing these institutional features, and on the effect of fixed abatement technology costs.

In the analysis that follows, we first establish statistical differences in market efficiencies both in terms of means and variances. Then, we examine price volatility and trade volume across treatments involving inter-firm trading. Finally, we demonstrate that limited credit generation is a key factor in observed inefficiencies associated with BAC2. Throughout the analysis we rely on pooled OLS regressions, with standard errors clustered at the unit-level such that hypothesis tests are robust to unspecified heteroskedasticity and within-unit serial correlation. Conclusions drawn from hypothesis tests are based on a five percent significance level, unless otherwise noted.

5.1. Analysis of efficiency

Mean Efficiency. Figure 1 plots the mean time series of efficiency outcomes for each of the eight treatments. The efficiency of group i in period t is defined as $\frac{Actual\ Surplus_{it}}{Available\ Surplus_{it}} \times 100$, with the mean across groups plotted in Figure 1.[17] Figure 2 presents average deviations between actual and efficient levels of abatement, by player type, averaged across all periods. The figures reveal several notable patterns. First, Figure 1 shows noticeable efficiency decreases for all treatments when technology costs are in effect, although this effect appears small for the TS. Figure 2 reveals that efficiency losses stem in particular from the two seller types under-abating (i.e. selling too few credits) and the two buyer types over-abating. Without technology costs, the TS, CAT and BAC1 time-series track one another closely, with near-efficient abatement; in contrast, there are noticeable efficiency losses associated with BAC2. Averaging across all periods, in the absence of technology costs, the rank-order of institutions in terms of efficiency is TS (96.6% efficiency), CAT (94.8%), BAC1 (91.0%) and BAC2 (72.8%). With technology costs, TS (92.5% efficiency) outperforms the other institutions, while CAT (74.4%) and BAC1 (72.5%) are similar, and BAC2 (54.6%) again has the lowest efficiency. Last, efficiencies increase as the experiment progresses and further appear to level off by the final period.

To formally test for efficiency differences attributable to differences in institutional features and fixed technology costs, we present as Model 1 in Table 3 estimates from a regression of group-level efficiencies. The indicator $Market$ equals one for CAT and BAC treatments. The interaction $Market \times Proposal$ equals one for the two BAC treatments. The

[17] Actual (available) surplus is calculated by taking the difference between actual (available) earnings and what earnings would be in the absence of trade. Note that it is possible for the efficiency measure to be negative if the actual earnings are lower than they would have been in the absence of trading.

indicator $Market \times Proposal \times Binding$ equals one for BAC2. The variable $1/Period^2$ captures possible (nonlinear) trends in the data. Note that a negative coefficient on the trend variable indicates that efficiencies are increasing over time. Further, since the time effect goes to zero as the period number gets large, estimates are interpreted as long-run effects.[18] The indicator $Technology\ Cost$ equals one for treatments where the technology cost feature is present, and interactions between $Technology\ Cost$ and the previously defined variables are also included.

First focusing on NoTech treatments, the model intercept indicates that estimated efficiency is near 100% for the TS. Neither introducing market trading nor incrementally requiring a proposal to generate credits has significant effects on efficiency (the joint effect is also insignificant). The former result is consistent with previous experiments, which have shown the double auction to be highly efficient empirically. The incremental effect of having a binding credit proposal is statistically significant, and large, suggesting a decrease of 18 percentage points. Linking this to our institutions, the evidence suggests that the TS, CAT and BAC1 efficiency outcomes are close to theoretical expectations in the absence of fixed technology costs, nearly achieving full efficiency (in the long-run). It is the binding aspect of the credit proposal present in BAC2 that leads to a large efficiency loss. The estimated time trend suggests that efficiency is increasing over time, consistent with what can be gleaned from Figure 1.

Turning to technology effects, the coefficient on the variable $Technology\ Cost$ suggests a three percentage point decrease in efficiency when technology costs are introduced in the TS setting, but this effect is not significant. Thus, in our baseline treatment of sorts, subjects are able to individually optimize in this slightly more complicated setting involving technology costs. In

[18] We note that, since the time variable is completely orthogonal to the other control variables, its inclusion/exclusion only affects the estimated intercept.

contrast, the effects of technology appear prominent when one moves to mechanisms involving inter-firm trading, under which the estimated efficiency loss is greater than 16 percentage points. There are no additional significant interaction effects with respect to BAC-specific institutional features, and the time effect is statistically equal to that for the NoTech settings. Thus, overall, the binding pre-commitment required in BAC2 as well as the presence of fixed technology costs in inter-firm trading settings are the main sources of inefficiency.

Variance of Efficiency. Using the same set of control variables, we also examine the variance of efficiency and present this as Model 2 in Table 3. Similar to Gilpatric et al. (2011), we calculate the variance of efficiency for group i in period t as $(Efficiency_{it} - \overline{Efficiency_t})^2$, where $\overline{Efficiency_t}$ is the treatment-specific mean efficiency in period t. This measure thus captures the within-period variation across groups (i.e. replications). The estimated intercept is not statistically different from zero, indicating that there is little variation in efficiency outcomes across groups for the TS treatment. This result is not surprising given that efficiency is virtually 100% in these treatments. Other than this, the results suggest that the two institutional features that reduce efficiency – requiring proposed (excess) abatement levels be verified prior to trade and, in the presence of fixed technology cost, inter-firm trading – also lead to an increased variance in efficiency. Thus, one central finding is that not only does the BAC2 institution lead to the lowest efficiencies it also leads to the greatest variation in efficiency outcomes, suggesting the possibility of fairly extreme poor outcomes.

5.2. Price Volatility

In this portion of the analysis we consider alternative measures of price volatility and compare them across treatments in order to examine the effects of institutional features and abatement technology adoption on market uncertainty. Although upfront investment is the controlled source of risk in our experiments, market uncertainty could play a secondary role in explaining differences across treatments. If for example, binding abatement pre-commitment and/or fixed technology costs increase price volatility in addition to adding sunk costs, then there is a secondary channel through which these features could impact risk and efficiency. We find limited evidence that binding abatement pre-commitment increases within-group price volatility, while none of the other experimental factors play a significant role.[19]

We present two models related to within-group price volatility in Table 4. Model 3 in Table 4 examines a measure of volatility at the transaction level, defined as $(Price_{jit} - \overline{Price}_{it})^2$, where $Price_{jit}$ is the observed price of the j^{th} transaction for group i in period t and $\overline{Price}_{it}$ is the mean price for group i in period t. Thus, this a measure of within-group, within-period price variation that would reflect the variation in what a player observes in terms of transaction prices over the course of a typical decision period. Using the same covariates as in the previous models, only the time trend variable, $1/Period^2$, has a statistically significant coefficient. The sign of the coefficient suggests that price volatility decreases over time. None of the institutional or technology cost variables are significant in either model, indicating that these factors have little influence on within-period price dispersion. If we instead estimate this model

[19] In addition to examining price volatility, we also looked for possible patterns in mean prices. Similar to price volatility results, mean prices are fairly consistent across treatments. Averaging over all periods, CAT, BAC1, and BAC2 have mean prices of \$237, \$235, and \$236, respectively, all within the equilibrium price interval of \$220 – \$240, and there are no statistically significant differences when stratifying by Tech and NoTech. A price analysis also suggests consistency in mean prices over time. Test statistics that support these conclusions are available upon request.

using only Period 1 data, there are likewise no significant effects.

Model 4 in Table 4 examines a measure of within-group, between-period price volatility, defined at the group-level as $(\overline{Price}_{it} - \overline{Price}_{i,t-1})^2$. Thus, this measure picks up on dispersion in average price between consecutive periods for a particular group. The time trend is again statistically significant and indicates decreasing volatility over time. In addition, binding abatement pre-commitment significantly increases price volatility. The magnitude of the effect is fairly large, suggesting an average difference in price between consecutive periods of almost 30 lab dollars. Hence, there is some evidence that upfront costs introduced through binding-abatement pre-commitment do increase the volatility of prices from one period to the next, which market participants may perceive as an additional source of risk, and which could therefore contribute to some of the observed efficiency loss associated with the BAC2 institution. A second possible measure of between-period price volatility, $(\overline{Price}_{it} - \overline{Price}_{i})^2$, is based on deviations between mean price for a period and the mean price across all periods. Our analysis of this measure suggests that price volatility decreases over time, and that there are no statistically significant effects linked to model covariates.

5.3. Trade Volume

As one of the main issues cited with existing water quality trading programs is limited or no trading, we analyze the effects of institutional features and costly technology adoption on trading volume (i.e. number of realized trades). We begin with simple comparisons of mean market-level trade volumes, and follow with regression analyses. The observed average volume over all treatments is roughly 14 trades per market, which is equal to the number predicted in all treatments according to competitive equilibrium. However, the average volume in treatments without technology costs or abatement pre-commitment is substantially higher than this number, at 18.5 trades, while the volume in treatments that have these features is slightly lower, 12.6. The mean volumes by treatment are: CAT-NoTech = 18.9; CAT-Tech = 15.7; BAC1-NoTech = 18.1; BAC1-Tech = 11.6; BAC2-NoTech = 12.5; and BAC2-Tech = 11.2. Note that trading volume does not necessarily equate to realized changes in abatement (e.g., if a participant buys then sells, then this has zero net effect on abatement), and that the higher-than-equilibrium volumes in some treatments suggest the occurrence of either reselling or trades for losses. Thus, we do not necessarily expect to find results that parallel our efficiency analysis. Nonetheless, our analysis of trade volume does reveal a similar pattern to that which emerges in the efficiency results, in the sense that sunk costs associated with purchasing a technology upgrade and pre-committing to (excess) abatement to generate credits appear to be thinning the market.

Table 5 presents a regression of trade volumes on the same set of covariates used in previous models. In this model, the interaction $Proposal \times Binding$ is again significant, and the coefficient suggests a 5.6 unit decrease in trade volume between BAC1 and BAC2. As in previous analysis, there is no discernable effect of simply moving from an exogenous initial allocation of credits to one where credit allocation is determined by a non-binding abatement proposal. There is also a marginally significant decrease in volume due to technology cost. The interaction terms together suggest that the technology effect is small for BAC2 at -1.2, and is not statistically significant (p=0.35), such that the technology effect is limited to CAT and BAC1 institutions. Somewhat peculiar is that the interaction $Proposal \times Binding \times Technology\ Cost$ is weakly significant and positive. One possible explanation is that, when the proposal is binding, some players may be willing to “dump” their credits in order to recover – at least partially – some of their initial investment costs.

5.4. Credit Generation

One potential source of inefficiency is limited credit availability in the market, resulting from sellers generating too few credits when faced with market risk posed by upfront abatement requirements or fixed technology costs. Credit availability is only a potential issue with the baseline-and-credit institutions, where there is no initial allocation and instead availability depends on proposed or binding levels of abatement in excess of baselines. If the group of sellers in BAC1 or BAC2 begins with fewer than 14 credits, there will necessarily be efficiency losses. We first discuss the treatment-specific means and then present an individual-level regression analysis that controls for risk aversion.

The mean number of credits generated at the group level by sellers is as follows: BAC1-NoTech = 17.5; BAC1-Tech = 10.2; BAC2-NoTech = 9.4; and BAC2-Tech = 8.1. In the latter three treatments, the mean number of credits generated is significantly different, and less than, 14 credits.[20] The data suggest an effect of the binding credit generation feature as well as a fixed technology cost effect.

If binding abatement requirements and/or fixed costs reduce credit generation through increased market risk, then risk aversion is likely to exacerbate the effect. To examine formally the possible role of risk aversion, we use choices from the risk elicitation tasks faced by participants prior to the market experiment. In particular we use the number of choices where the participant chose the fixed payoff instead of the binary lottery. Noting that a risk neutral person would select the fixed payoff for six of the 11 lotteries, we define the indicator variable $Risk\ Averse$ to equal one for players that chose the fixed payoff in at least seven lotteries.

[20] The results reported here are based on *t*-tests from a pooled linear regression of market-level credit generation on a set of treatment indicator variables, with cluster-robust standard errors.

Table 6 presents regression results using individual-level data from the BAC treatments, where the dependent variable is the difference between the number of credits actually generated by a seller in a given period and the number of credits that would be sold under the efficient equilibrium. Explanatory variables include a full set of treatment-specific indicator variables (thus, no intercept is included), treatment interactions with the covariate $Risk\ Averse$, and controls for time trends. Thus, the coefficients on the treatment-specific indicator variables are interpretable as the estimated (long-run) mean deviations for a risk-neutral seller, and the interactions with $Risk\ Averse$ capture the mean effect of risk aversion.

The results suggest that, for a risk-neutral seller, there is no significant deviation between the number of credits generated and the efficient level in either BAC1 treatment. Of course, confirming the raw data, there is *excess* credit generation in BAC1-NoTech. The story is much different for BAC2, where there is a statistically significant shortfall in both treatments, with the degree of shortfall (roughly 0.8 units per seller) similar across treatments. Thus, assuming that our risk aversion measure is accurate, risk aversion alone does not explain the shortfall in credit generation in BAC2. However, the BAC2 interactions with $Risk\ Averse$ are both negative and statistically significant suggesting that risk aversion does nevertheless appear to play a role, and in the expected direction. The risk aversion effect is three times larger with technology cost in play, which is intuitive given the larger degree of financial risk. On the other hand, risk aversion does not explain deviations from efficient credit generation for BAC1. Overall, since proposals are “cheap talk” in BAC1, it would make sense that neither risk aversion nor technology cost would play a role. The risk aversion results for BAC1 can thus be viewed as a falsification test of sorts, providing some credence that risk aversion is a partial driver of behavior in BAC2.

6. Conclusions

Motivated by the puzzling observation that most water quality trading (WQT) programs have experienced little success, this study uses laboratory experiments to isolate the effects of some of the features that distinguish WQT programs from the widely-studied and more successful extant air quality trading programs. Fundamentally, the allocation of tradable credits, which are typically freely allocated or auctioned in the cap-and-trade air quality programs, are based on reducing emissions below a baseline in WQT. Tied to the latter "baseline-and-credit" institution is that most programs require firms to undertake (excess) abatement, and have this quantified through a stringent verification process, prior to receiving tradable credits. This alternative method of credit allocation exposes firms to increased financial risk given uncertainty over credit demand and the related possibility of not recovering investment costs through trading. A second characteristic endemic to WQT is that opportunities for trade in many watersheds are linked to a firm's willingness to adopt costly technology, which amplifies concerns over the effects of credit market uncertainty.

There are several findings. First, requiring firms to make a binding abatement choice prior to market trading – which reflects suggested EPA guidelines – results in a large decrease in efficiency, as well as a large increase in the variation in efficiency outcomes across markets. The empirical result is consistent with theoretical work on sunken investments (Mailath et al. 2004). Sellers tend to pre-commit to inefficiently low abatement levels, which leads to a lower initial allocation of tradable credits. This effect is found to be exacerbated by risk aversion. Second, there is no effect of simply moving from a setting where firms are endowed with credits to one where credits are generated through a non-binding abatement choice. Third, the presence of fixed abatement technology costs significantly reduces efficiency in all of the inter-firm trading settings examined. Fourth, the tax/subsidy mechanism we explore is nearly efficient in settings with and without fixed technology costs. As under this mechanism there is no uncertainty over credit demand and supply – firms simply buy/sell credits at a fixed price – this result is largely expected. However, this emphasizes that our findings are due to risk tied to inter-firm trading rather than confusion tied to experiment complexity.

Our findings provide some evidence that institutional rules (requiring a binding abatement pre-commitment to generate tradable credits) as well as inherent features of WQT settings (fixed costs associated with abatement technology) are impediments to trade in existing WQT markets. From a policy perspective, consideration should be given to determining the appropriateness of requiring binding abatement pre-commitment in order to generate credits in a WQT program. Furthermore, where high fixed technology costs are a concern, a tax/subsidy scheme may be more efficient than baseline-and-credit or cap-and-trade, as the absence of market uncertainty appears to make the tax/subsidy more robust in such circumstances.

However, a potential caveat with a tax/subsidy program is that the regulator's uncertainty over the optimal tax/subsidy rate (i.e. credit price) has the ability to invoke efficiency losses and distort short and long-run incentives for investment in abatement technology. Indeed, the Long Island Sound Trading Program has required the State of Connecticut to deal with large imbalances between the number credits purchased and sold, implying that the program has more than met its water quality goals in some years, and has failed to meet them in others. For example, over the first three years of the program, the State spent $2.6 million to purchase excess credits; in the fourth, $1.2 million worth of credits were sold (Connecticut State Treasury 2003, 2004, 2005, 2006).

This study represents a first step toward investigating fundamental features that distinguish WQT from textbook cap-and-trade, and there are a number of features remaining in this avenue to explore. One characteristic of WQT programs that has garnered attention from economists is trading ratios, which define the amount of additional emissions one firm would be allowed for each credit purchased from another firm. As trading ratios are in part based on topography, this could serve to increase or decrease opportunities for trade. A second characteristic is the potential for an unregulated non-point source to generate credits for regulated point sources. Third, various aspects of uncertainty such as those associated with decentralized trading (e.g. bilateral trade), the number of viable market participants, abatement costs, and future regulatory intensity are absent from the experimental design in this study but are likely important in the field. Given the ability of experiments to parse the effects of particular market features and the lack of sufficient real world data on WQT markets, experiments can play a central role in designing future WQT programs and (ideally) rescuing existing ones.

References

Ben-David, S., Brookshire, D.S., Burness, S., McKee, M., & Schmidt, C. (1999). Heterogeneity, Irreversible Production Choices, and Efficiency in Emission Permit Markets. *Journal of Environmental Economics and Management,* 38(2), 176-194.

Ben-David, S., Brookshire, D.S., Burness, S., McKee, M., & Schmidt, C. (2000). Attitudes Toward Risk And Compliance in Emission Permit Markets. *Land Economics*, 76(4), 590-600.

Bohm, P. (2003). Experimental evaluations of policy instruments. In K.-G. Maler, & J. R. Vincent (Eds.), *Handbook of environmental economics* (Vol. 1, pp. 437-460). Amsterdam: Elsevier Science.

Boisvert, R. N., Poe, G. L., & Sado, Y. (2007). Selected economic aspects of water quality trading: A primer and interpretive literature review. Cornell University.

Breetz, H. L., Fisher-Vanden, K., Garzon, L., Jacobs, H., Kroetz, K., & Terry, R. (2004). *Water quality trading and offset initiatives in the U.S.: A comprehensive survey.* Hanover, NH: Dartmouth College.

Buckley, N. J., Mestelman, S., & Muller, R. A. (2006). Implications of alternative emission trading plans: Experimental evidence. *Pacific Economic Review, 11*(2), 149-166.

Buckley, N. J., Mestelman, S., & Muller, R. A. (2008). Baseline-and-credit emission permit trading. In T. L. Cherry, S. Kroll, & J. F. Shogren (Eds.), *Environmental economics, experimental methods* (pp. 9-28). Routledge Explorations in Environmental Economics. London and New York: Taylor and Francis, Routledge.

Buckley, N. J., Mestelman, S., & Muller, R. A. (2011). *Production capacity and abatement technology strategies in alternative emissions markets.* Paper presented at the 18th Annual Conference of the European Association of Environmental and Resource Economists.

Camacho-Cuena, E., Requate, T., Waichman, I. (2012). Investment incentives under emission trading: an experimental study. *Environment and Resource Economics*, 53(2), 229–249.

Caplan, A. J. (2008). Incremental and average control costs in a model of water quality trading with discrete abatement units. *Environmental and Resource Economics, 41*(3), 419-435.

Cason, T. N., & Gangadharan, L. (2006). Emissions variability in tradable permit markets with imperfect enforcement and banking. *Journal of Economic Behavior & Organization, 61*(2), 199-216.

Cason, T.N., Gangadharan, L., & Duke, C. (2003). Market power in tradable emission markets: a laboratory testbed for emission trading in Port Phillip Bay, Victoria. *Ecological Economics*, 46(3), 469-491.

Cason, T.N., & Plott, C.R. (1996). EPA's New Emissions Trading Mechanism: A Laboratory Evaluation. *Journal of Environmental Economics and Management*, *30*, 133-160.

Connecticut State Treasury (2003). 2003 clean water fund report.

Connecticut State Treasury (2004). 2004 clean water fund report.

Connecticut State Treasury (2005). 2005 clean water fund report.

Connecticut State Treasury (2006). 2006 clean water fund report.

Environomics (1999). *A summary of U.S. Effluent trading and offset projects*. Report to U.S. EPA, Office of Water.

Farrow, S. R., Schultz, M. T., Celikkol, P., & Van Houtven, G. L. (2005). Pollution trading in water quality limited areas: Use of benefits assessment and cost-effective trading ratios. *Land Economics*, *81*(2), 191-205.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Gangadharan, L., Farrell, A., & Croson, R. (2013). Investment decisions and emissions reductions: experimental results in emissions permit trading. In Price, M. & List, J. (Eds.), *The Handbook of Experimental Economics and the Environment*. Edward Elgar.

Ghosh, G., Kwasnica, A., & Shortle, J. (2011). A Comparison of Two Oligopsonistic Permit Trading Markets when Emissions are Variable. Paper presented at the 18th Annual Conference of the European Association of Environmental and Resource Economists.

Ghosh, G. and Shortle, J. (2010). Emissions trading when pollution loads are stochastic and unmeasurable. Paper presented at the Fourth World Congress of Environmental and Resource Economists.

Gilpatric, S. M., Vossler, C. A., & McKee, M. (2011). Regulatory enforcement with competitive endogenous audit mechanisms. *The RAND Journal of Economics*, *42*(2), 292-312.

Hoag, D. L., & Hughes-Popp, J. S. (1997). Theory and practice of pollution credit trading in water quality management. *Review of Agricultural Economics*, *19*(2), 252-262.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, *92*(5), 1644-1655.

King, D. M. (2005). Crunch time for water quality trading. *Choices, 1st Quarter, 20*(1), 71-75.

King, D. M., & Kuch, P. J. (2003). Will nutrient credit trading ever work? An assessment of supply and demand problems and institutional obstacles. *The Environmental Law Reporter, 33*, 10352-10368.

List, J. A., Sadoff, S., & Wagner, M. (2010). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. National Bureau of Economic Research, Inc, NBER Working Papers: 15701.

Mailath, G. J., Postlewaite, A., & Samuelson, L. (2004). Sunk investments lead to unpredictable prices. *The American Economic Review, 94*(4), 896-918.

Mestelman, S., Moir, R., & Muller, R. A. (1999). A laboratory test of a canadian proposal for an emissions trading program. In R. M. Isaac, & C. Holt (Eds.), *Emissions permit experiments* (pp. 45-91). Research in Experimental Economics, vol. 7. Stamford, Conn.: JAI Press.

Murphy, J. J., & Stranlund, J. K. (2007). A laboratory investigation of compliance behavior under tradable emissions rights: Implications for targeted enforcement. *Journal of Environmental Economics and Management, 53*(2), 196-212.

Sado, Y., Boisvert, R. N., & Poe, G. L. (2010). Potential cost savings from discharge allowance trading: A case study and implications for water quality trading. *Water Resources Research, 46*(2), 1-12.

Stranlund, J. K., Murphy, J. J., & Spraggon, J. M. (2011). An experimental analysis of compliance in dynamic emissions markets. *Journal of Environmental Economics and Management, 62*(3), 414-429.

Suter, J.F., Spraggon, J.M., & Poe, G.L. (2013). Thin and lumpy: An experimental investigation of water quality trading. *Water Resources and Economics*, 1, 36-60.

U.S. EPA (1996). *Draft framework for watershed-based trading*. Office of Water, No. EPA 800-R-96-001.

U.S. EPA (2002). *2000 national water quality inventory*. No. EPA-841-R-02-001.

U.S. EPA (2007). *Water quality trading toolkit for permit writers (updated 2009)*. No. EPA 833-R-07-004.

U.S. EPA (2008). *Epa water quality trading evaluation: Final report*.

U.S. EPA (2011). Watershed assessment, tracking & environmental results: National summary of state information. http://iaspub.epa.gov/waters10/attains_index.control. Accessed August 6, 2011.

Table 1. Experiment Design Summary

Treatment	Fixed Technology cost?	Inter-firm trading?	Initial Credit Allocation	Is Proposal Binding?	No. of Replications (Groups)
TS-NoTech	No	No	Exogenous, at no cost	N/A	4
CAT-NoTech	No	Yes	Exogenous, at no cost	N/A	4
BAC1-NoTech	No	Yes	Based on proposed (excess) abatement	No	4
BAC2-NoTech	No	Yes	Based on proposed (excess) abatement	Yes	4
TS-Tech	Yes	No	Exogenous, at no cost	N/A	4
CAT-Tech	Yes	Yes	Exogenous, at no cost	N/A	4
BAC1-Tech	Yes	Yes	Based on proposed (excess) abatement	No	4
BAC2-Tech	Yes	Yes	Based on proposed (excess) abatement	Yes	6

Notes:
(1) TS: Tax/Subsidy policy; CAT: Cap-and-Trade policy; BAC: Baseline-and-Credit policy
(2) A "replication" is a unique eight-player group (two of each type as defined in Table 2) that participated in the same treatment over 10 distinct decision periods.

Table 2. Abatement Costs and Credit Allocations (Emissions Baselines), by Type

	Buyers				Sellers			
Abatement	Type 1		Type 2		Type 3		Type 4	
	MC	FC	MC	FC	MC	FC	MC	FC
1	100	100	155	100	17	100	25	100
2	150	100	170	100	18	100	27	100
3	**200**	**100**	185	100	*19*	*100*	30	100
4	260	100	200	100	20	100	35	100
5	330	100	**215**	**100**	50	300	40	300
6	*400*	*100*	240	100	130	300	*50*	*300*
7	475	100	250	100	**220**	**300**	60	300
8	550	300	260	300	310	300	130	300
9	625	300	*270*	*300*	450	300	**200**	**300**
10	700	300	425	300	575	300	300	300
Credit Allocation or Emissions Baseline	4		1		7		4	

Notes:
(1) MC is the marginal abatement cost for the indicated unit of abatement. FC is the fixed technology cost associated with the indicated level of abatement, which applies only to the Tech scenario.
(2) Italicized numbers reflect the abatement level under the status quo (no trading). Bold numbers reflect the expected outcome after efficient trading.

Table 3. Analysis of Efficiency

Variable	Model 1: Efficiency	Model 2: Efficiency Variance
Market	−1.76	25.68
	(2.40)	(22.78)
Market × Proposal	−3.81	77.25
	(3.00)	(61.70)
Market × Proposal × Binding	-18.20^{***}	178.12^{***}
	(5.66)	(62.28)
1/Period2	-20.55^{***}	361.55^{**}
	(6.24)	(182.05)
Technology Cost	−2.99	61.53
	(1.93)	(46.66)
Market × Technology Cost	-16.34^{***}	253.55^{***}
	(5.98)	(82.12)
Market × Proposal × Technology Cost	1.94	-222.06^{*}
	(8.27)	(113.71)
Market × Proposal × Binding × Technology Cost	0.28	419.42
	(11.37)	(343.34)
(1/Period2) × Technology Cost	−7.03	−290.33
	(8.86)	(291.20)
Intercept	99.75^{***}	−38.86
	(1.64)	(29.43)
Observations	340	340
F	11.31^{***}	19.08^{***}
R^2	0.43	0.09

Notes: Cluster-robust standard errors are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4. Within-Group Price Volatility

Variable	Model 3: Within-Period Volatility	Model 4: Across-Period Volatility
Proposal	1003.27	25.38
	(1013.83)	(73.66)
Proposal × Binding	−961.47	835.32**
	(976.84)	(313.65)
$1/\text{Period}^2$	4537.26***	742.26**
	(1029.87)	(325.40)
Technology Cost	1534.89	105.04
	(1450.09)	(275.88)
Proposal × Technology Cost	−2588.27	170.18
	(2147.71)	(439.14)
Proposal × Binding × Technology Cost	715.04	−442.05
	(1670.53)	(567.91)
$(1/\text{Period}^2)$ × Technology Cost	2076.62	1502.74
	(1945.46)	(991.56)
Intercept	1594.11***	92.57
	(445.91)	(78.08)
Observations	3740	234
F	7.39***	3.98***
R^2	0.04	0.17

Notes: Cluster-robust standard errors are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 5. Trade Volume

Variable	Model 5: Trade Volume
Proposal	−0.80
	(1.74)
Proposal × Binding	-5.60^{***}
	(1.31)
1/Period^2	−1.52
	(1.07)
Technology Cost	-3.09^{*}
	(1.62)
Proposal × Technology Cost	−3.25
	(2.74)
Proposal × Binding × Technology Cost	5.16^{*}
	(2.51)
(1/Period^2) × Technology Cost	−0.70
	(2.14)
Intercept	19.11^{***}
	(1.29)
Observations	260
F	6.67^{***}
R^2	0.36

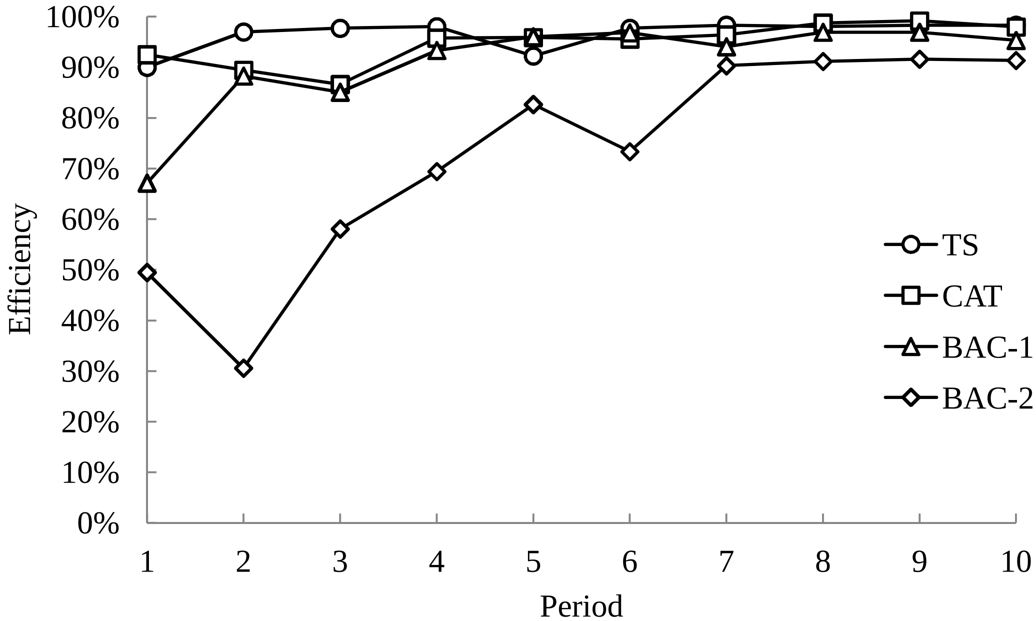
Notes: Cluster-robust standard errors are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 6. Deviation from Efficient Credit Generation (Sellers, BAC1 and BAC2 data)

Variable	Model 6: Efficient Generation
BAC1-NoTech	1.02^{***}
	(0.33)
BAC1-Tech	−0.63
	(0.63)
BAC2-NoTech	-0.77^{***}
	(0.20)
BAC2-Tech	-0.81^{***}
	(0.20)
BAC1-NoTech × Risk Averse	0.07
	(0.60)
BAC1-Tech × Risk Averse	−0.33
	(0.64)
BAC2-NoTech × Risk Averse	-0.27^{**}
	(0.45)
BAC2-Tech × Risk Averse	-0.86^{***}
	(0.21)
$1/\text{Period}^2$	-1.23^{***}
	(0.28)
(1/Period2) × Technology Cost	0.19
	(0.31)
Observations	720
F	37.47^{***}
R^2	0.46

Notes:
(1) The dependent variable is the difference between the number of credits generated by a seller and the number of credits required by the seller in order to conduct the efficient number of trades.
(2) Sellers are defined as those players in the experiment who are net sellers of credits in the efficient outcome, type 3 and type 4 players.
(3) Cluster-robust standard errors are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

No Fixed Technology Cost (NoTech)



Fixed Technology Cost (Tech)

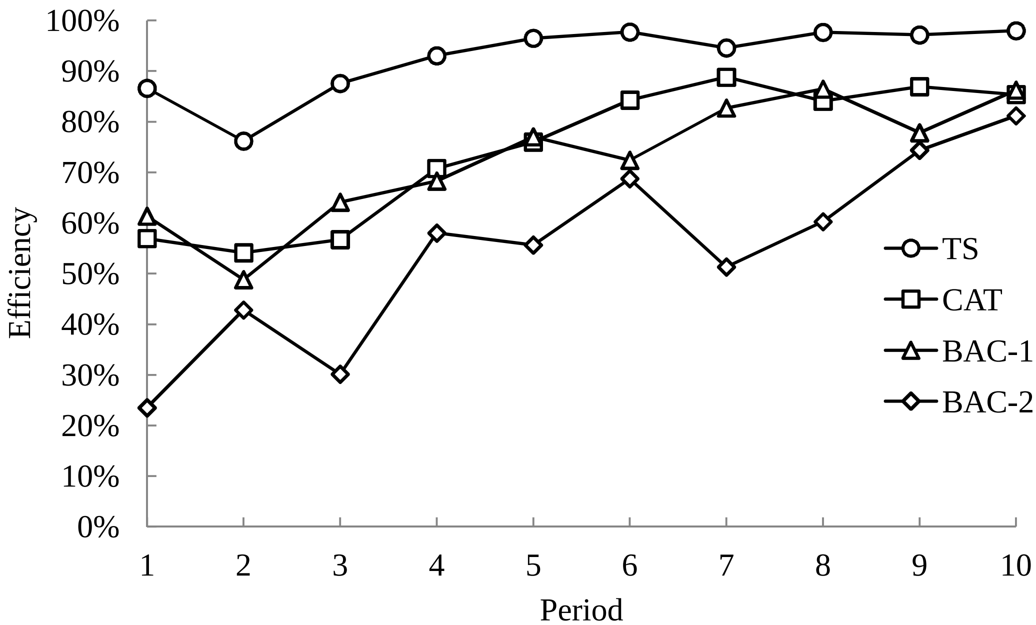


Figure 1. Market Efficiency

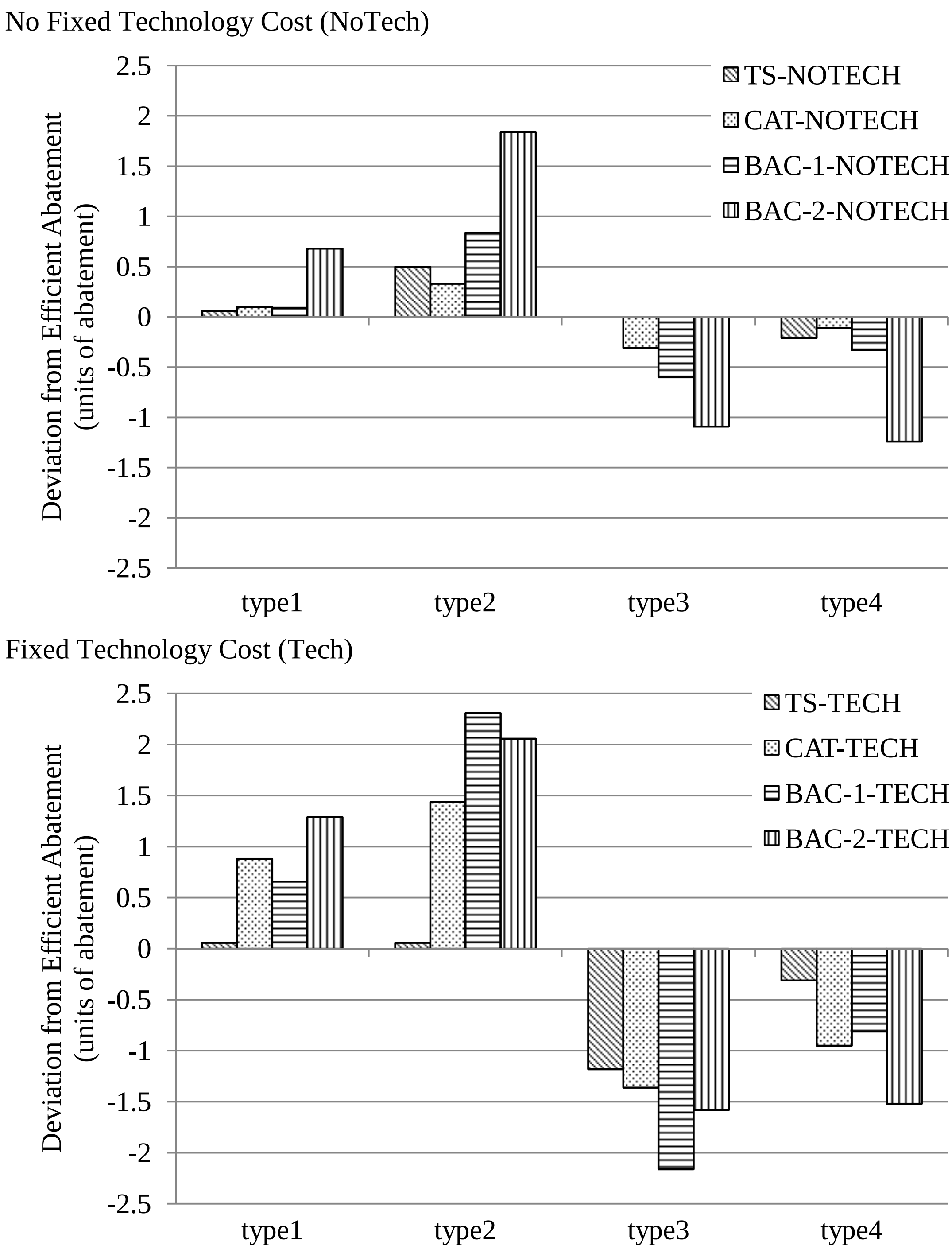


Figure 2. Individual-Level Deviations from Efficient Abatement

Online Appendix for "Experimental Tests of Water Quality Trading Markets"

Journal of Environmental Economics and Management

This appendix contains (unabridged) experiment instructions for treatment BAC2, baseline-and-credit regulation with binding abatement pre-commitment, with fixed technology costs. Instructions for other treatments are available from the authors upon request.

INTRODUCTION

This experiment is a study of group and individual decision making. The amount of money you earn depends on the decisions that you make, so please read the instructions carefully. The money you earn will be paid privately to you, in cash, at the end of the experiment. A research foundation has provided the funds for this study.

You will make decisions privately, that is, without consulting other participants. Please do not attempt to communicate with other participants in the room during the experiment. Further, please refrain from text messaging, emailing, or other activities not related to the experiment.

If you have a question as we read through the instructions or at any time during the experiment, please raise your hand and an experiment moderator will answer it. We are happy to clarify anything about the rules of the experiment and what determines your earnings. However, please refrain from discussing strategy (e.g. "I think X is a good choice…") as this may affect the choices and earnings of others.

There are two parts to this experiment. You earn money in each part, and your earnings in the second part are denominated in experimental dollars. The experimental dollars will be exchanged for U.S. dollars at the end of the experiment at a known exchange rate. We are now ready to begin Part A of the experiment.

PART A

At this time, please read the information on your computer screen and then click the "Begin Part A" button after you are finished. Know that the software will not continue the experiment until everyone has indicated they are ready by clicking on the button. Now, follow along as the instructions below (which you also appear on your computer) are read aloud. *Please* refrain from making any choices until the instructions are finished and any questions are addressed.

> For each scenario below we ask you to choose between option "A" and option "B". Only ONE of the scenarios will be used to determine your earnings. What this means is that you should treat each choice as independent from the others, and thus should consider each choice carefully.
>
> After we complete the other part of the experiment, the computer will randomly select ONE of these scenarios to be played out. Those who chose "A" for the selected scenario will receive \$3.00. Those who chose "B" will receive either \$5.00 or \$0.50 according to the chances identified in the scenario. As an example, *suppose* that Scenario 6 were randomly selected to be played out. This means that, if you selected "A" you will earn \$3 for this part of the experiment. If you selected "B" you have a 50% chance of earning \$5 and a 50% chance of earning \$0.50. No exchange rate will be used in this part (i.e., values on the screen are in U.S. dollars).

Before you make your choices, are there any questions?

After you are comfortable with the choices you have made, please click the "submit" button on your screen.

PART B

In this part we are going to conduct a number of periods. You will not learn the number of periods until the end of the experiment. Each period is completely independent in the sense that your decisions in one period do not directly affect your earnings in other periods. Each period has two stages. In the first stage you propose a level of production, which determines the number of "coupons" available to be traded. In the second stage, you will have the opportunity to buy and sell coupons in a computerized auction. Based on the number of coupons you buy or sell, you will pay production costs.

Your earnings for a particular period are determined as follows:

Period Earnings = Initial Earnings – Total Production Costs – Technology Cost + Amount Earned from Selling Coupons – Amount Spent from Buying Coupons

Your Initial Earnings do not depend on any actions you take and do not change throughout the experiment. We will now go over the other factors important for your Earnings.

PRODUCTION COSTS

You must pay production costs when you produce units. The cost of each unit produced is typically different from the cost of other units produced, and your costs may or may not be different from the costs of other participants. Your production costs are always shown on the left side of your computer screen, as illustrated in **Figure 1** (the numbers on this example screen are for illustration only). You can produce up to 10 units, and the cost of each unit is shown separately.

For instance, based on the numbers shown in the example, your first unit produced would cost $200, your second unit produced would cost $400, etc. If, for example, these were your unit production costs and you produced 3 units, your total production costs would be $200 + $400 + $600 = $1200. So you must recognize that the costs shown on your screen are the <u>extra</u> costs associated with each <u>additional</u> unit produced.

TECHNOLOGY COSTS

In order to produce anything, you need to have <u>one</u> of two possible technologies. This technology is automatically determined by the number of units you choose to produce. In particular, for this example, if you produce 1 to 7 units, you pay for Technology A. If you produce 8 to 10 units you pay for Technology B. Based on the example, if you produced 5 units you would need Technology A, which costs you $250 (not $250 per unit!). Note that your technology costs, as well as the amount of production that is possible with a given technology, may be different from those in the example.

COUPONS

We've already explained that your Initial Earnings never change, but each additional unit you produce costs you additional money. So why should you ever produce any units? The reason comes from today's production rule:

Required Production = Initial Production Requirement – Coupons Purchased + Coupons Sold

Everyone starts each period with an Initial Production Requirement (it is 6 units in the example). As suggested by this rule, you decrease your production by buying coupons and increase your production if you sell coupons.

Why might you want to buy a coupon? Buying a coupon will cost you money. However, for each coupon you buy you produce one unit less and so do not have to pay the production cost of that unit. Suppose you had the production costs illustrated in the example and your initial production requirement was 7 units. If you bought a coupon through the auction, this means you would only have to produce 6 rather than 7 units. Therefore, you would save the production cost of the 7^{th} unit, which is $1400 in the example. Depending on how many coupons you buy, you may also save money by being able to use a less costly technology.

Why might you want to sell a coupon? You will receive money by selling a coupon. However, for each coupon you sell you produce one more unit and so have to pay the production cost of that additional unit. Using the production costs in the example, suppose that your initial production requirement was 4. If you sold a coupon, this means you would need to produce 5 instead of 4 units to meet the rule. Thus, you would have to pay the production cost of the 5th unit, which is $1000 in the example. Depending on how many coupons you sell, you may incur additional cost by needing to use a more costly technology.

PROPOSED PRODUCTION DECISION

No one is automatically given free coupons that they can then trade in the market. The number of coupons available for trade instead depends upon proposed production decisions. This proposed production decision represents a production pre-commitment and works as follows:

> If you want to receive coupons, which you can then sell in the auction market, you propose a level of production that is *higher* than your Initial Production Requirement. You will receive one coupon for each unit of production that exceeds your Initial Production Requirement. Also, regardless of whether you sell the coupon(s), you will *automatically* produce the number of units you propose. The production rule no longer applies (in fact, you already exceeded the requirement), and you will have no opportunity to lower your production by buying coupons.
>
> If you do not want to receive coupons prior to entering the market, then you simply indicate this by selecting the "I do not wish to generate any coupons" option.

For those who decide not to have coupons prior to entering the market, the computer will make sure that you satisfy the Production Rule. This means that, after you buy/sell coupons in the auction the computer will automatically determine what production is required. For example, suppose you have an Initial Production Requirement of 6 units and you buy 2 coupons in the auction. Since buying the 2 coupons reduces your required production by 2 units, the computer will automatically require you to produce 4 units.

We realize that we have given you a lot of information for you to absorb. Before we continue, are there any questions?

We now ask you to work through a couple of examples. The purpose of the examples is to help you better understand how the experiment works before real money is on the line. For these examples, assume you face the initial earnings, productions costs, etc., as shown in **Figure 1**. To provide you an incentive to work through these examples carefully, we will give $1 U.S. if you answer them correctly.

EXAMPLE 1. Assume you did not generate any coupons prior to entering the market. Further, suppose you neither purchased nor sold any coupons in the auction. Then,

(a) You would be required to produce ______________ units.

(b) Your earnings for the period would be (please show your calculations):

(+) Initial Earnings:

(−) Total Production Costs:

(−) Technology Costs:

(+) Amount Earned from selling coupons:

(−) Amount Spent from buying coupons:

Period Earnings: ______________________

EXAMPLE 2. Assume you did not generate any coupons prior to entering the market. Consider what would happen if you purchased 1 coupon.

(a) You would then have to produce __________ units.

(b) What price would you need to buy the coupon for to *increase* your earnings by $1?

EXAMPLE 3. Now consider the tradeoffs associated with pre-committing to produce extra units.

(a) If you wanted to generate one coupon prior to entering the auction, you would then have to propose to produce __________ units.

(b) What price would you need to sell the coupon for in order to *increase* your earning by $1, relative to the situation where you did not generate any coupon?

Please raise your hand when you are ready to have your answers checked.

THE MARKET

We now will discuss how you buy and sell coupons in the auction market. Please pay close attention to the rules for both buying and selling, as this will help you determine the best strategy for you – you will not automatically know if it is in your best interest to buy or sell coupons (or both).

Figure 2 is an example of the market trading screen for a player that chose not to generate coupons prior to entering the market (these numbers are for illustration only). Notice that production costs are automatically displayed on the left side of the screen. Based on the number of coupons you have bought or sold, units that you are required to produce in order to meet the production rule have "required" next to them. In the example, the Initial Production Requirement is 6 units and this player has not made any trades. Therefore, notice that the first 6 units are labeled as "required". This information is updated when transactions are made. For instance, in the example, if the player bought a coupon the number of coupons would go from 0 to 1 and only the first 5 units would be labeled as "required".

Figure 3 is an example of the market trading screen for a player that chose to generate coupons prior to entering the market (these numbers are for illustration only). In particular, the player had an Initial Production Requirement of 6 units and proposed to produce 7 units. Therefore, the player has 1 coupon available to sell. Notice that the word "produced" is written next to the first 7 units. And, remember that in this case the level of production stays the same regardless of how many coupons the player buys or sells.

Technology costs are also displayed on the bottom left side of the screen. The technology labeled as "required" is determined based on your current level of required production. This will change if you need a different technology due to a change in production.

HOW YOU BUY COUPONS…

There are two ways for you to buy a coupon: by submitting a "Buy Bid" (and having a seller accept it) or accepting an offer made by a participant wishing to sell.

You submit a Buy Bid by typing a number in the "Buy Bid" box on the right side of the screen, and then clicking on the "Make Bid" button that appears below the box. This bid is immediately displayed on all traders' computers on the left part of the screen, underneath where it says "Best Buy Bid". Once a bid has been posted, any seller can accept it. Such an acceptance results in an immediate trade at that price.

If you or another participant has already placed a Buy Bid, it will remain until a trade is made, or until anyone submits a new bid that represents more favorable terms to potential sellers. Sellers prefer higher prices, so any new bid submitted must be higher than the posted Best Buy Bid. Your computer will give you an error message if you try to offer a lower price than the current Best Buy Bid. From the example in Figure 2, notice that there is displayed a Best Buy Bid of $1230. In order to have your own bid displayed you would have to submit a bid higher than $1230.

Alternatively, at any time, you can accept the current Best Sell Offer by clicking the “Buy” button directly below the offer. This results in an immediate trade at that price. From the example, notice that – since the “Best Sell Offer” is $1075 – this would be the price paid if you clicked the “Buy” button.

How you Sell Coupons…

There are two ways for you to sell a coupon: by submitting a “Sell Offer” (and having another participant accept it) or accepting a bid made by a participant wishing to buy.

You submit an offer by typing in a number in the “Sell Offer” box on the left side of the screen, and then clicking on the “Make Offer” button below this box. This offer price is immediately displayed on all traders’ computers on the right part of the screen, underneath where it says “Best Sell Offer.” Once this offer price has been submitted, any buyer can accept it. Such an acceptance results in an immediate trade at that price.

If you or another participant has already posted a Sell Offer, it will remain until a trade is made, or until a seller submits a new offer that represents more favorable terms to the buyers. Buyers prefer lower prices, so any new sell offer submitted must be lower than the posted Best Sell Offer. Your computer will give you an error message if you try to offer a higher price than the current Best Sell Offer. From the example in Figure 2, notice that there is displayed a Best Sell Offer of $1075. In order to have your own offer displayed you would have to submit an offer lower than $1075.

Alternatively, at any time, you can accept the Best Buy Bid by clicking the “Sell” button that appears directly below the bid. This results in an immediate trade at that price. From the example, notice that – since the “Best Buy Bid” is $1230 – this would be the price received if you clicked the “Sell” button.

Market Organization and Transactions…

The first two trading periods will last three minutes. The remaining periods will last two minutes and thirty seconds. The time left in the period is displayed in the upper right corner of the computer screen. Displayed in the upper left part of the screen is some useful information. Your “Period Earnings” are calculated using the formula we previously discussed. The particular amount displayed is based on your current level of required production. “Market Earnings” reflect your earnings associated with trades only. In other words, it is the difference between the Period Earnings prior to trading and your current Period Earnings. “Required Production” is the number of units you are required to produce. “Number of Coupon(s)” is the number of coupons you currently have.

When a trade occurs, any offer/bid associated with the trade will be cleared. Both the buyer and seller involved will be notified of the transaction, and will see their Period Earnings, Market Earnings, and Number of Coupon(s) change accordingly. For those who did not pre-commit to a level of production, their Required Production will be updated (e.g. will decrease by one if they bought a coupon).

In the middle of the screen the ID number of the buyer and seller along with the transaction price are recorded. This information can be seen on all traders' computers and will remain during the trading period. On the example screens (Figure 2 or Figure 3), notice that two trades have already occurred.

Transactions can occur quickly and so it is in your best interest to keep track of your coupons, production costs and technology cost prior to submitting a bid/offer or accepting one.

At the end of each trading period, you will see a results screen that summarizes the outcomes of the trading period. You will see how many coupons you have bought or sold and number of units produced. You will see your Period Earnings calculated for you, based on your initial earnings, production costs, and what you paid for additional coupons (or received from coupon sales)

Before we begin with the paid decision periods, we will first go through **2** unpaid training periods. In the training periods you will face the same scenario (unit production costs, initial production requirement, technology cost, etc.) as you will in the paid decision periods that follow.

Are there any questions before we proceed?

Before the training period, we ask that you work through a couple more examples. The purpose of the examples is to help you better understand how the auction works before you start trading and real money is on the line. To provide you an incentive to work through these examples carefully, we will give $1 U.S. if you answer all of them correctly. Please show your calculations.

EXAMPLE 4. For this example, assume you face the initial earnings, productions costs, etc., as shown in **Figure 2**. Consider buying a coupon in this market. Look carefully at **Figure 2** and notice that there is already both a Best Sell Offer and Best Buy Offer posted.

(a) If you clicked the "Buy" button, how much would you pay?
(b) Relative to the situation where you did not buy a coupon, by how much did you increase (or decrease) your earnings? [Hint: consider what you saved in production costs by buying]

EXAMPLE 5. For this example, assume you face the initial earnings, productions costs, etc., as shown in **Figure 3**. Consider selling a coupon in this market. Look carefully at **Figure 3** and notice that there is already both a Best Sell Offer and Best Buy Offer posted.

(a) If you clicked the "Sell" button, how much would you receive?
(b) Relative to a situation where you did not pre-commit to produce the extra unit, by how much are you better or worse off? [Hint: consider what the additional amount you paid in production cost to generate the coupon]

Please raise your hand when you are ready to have your answers checked. Please click the "Begin Part B" button on your screen <u>only</u> after being asked to do so.

Figure 1. Example Market Entry Screen (Part B)

Figure 2. Example Market Trading Screen for a coupon non-generator (Part B)

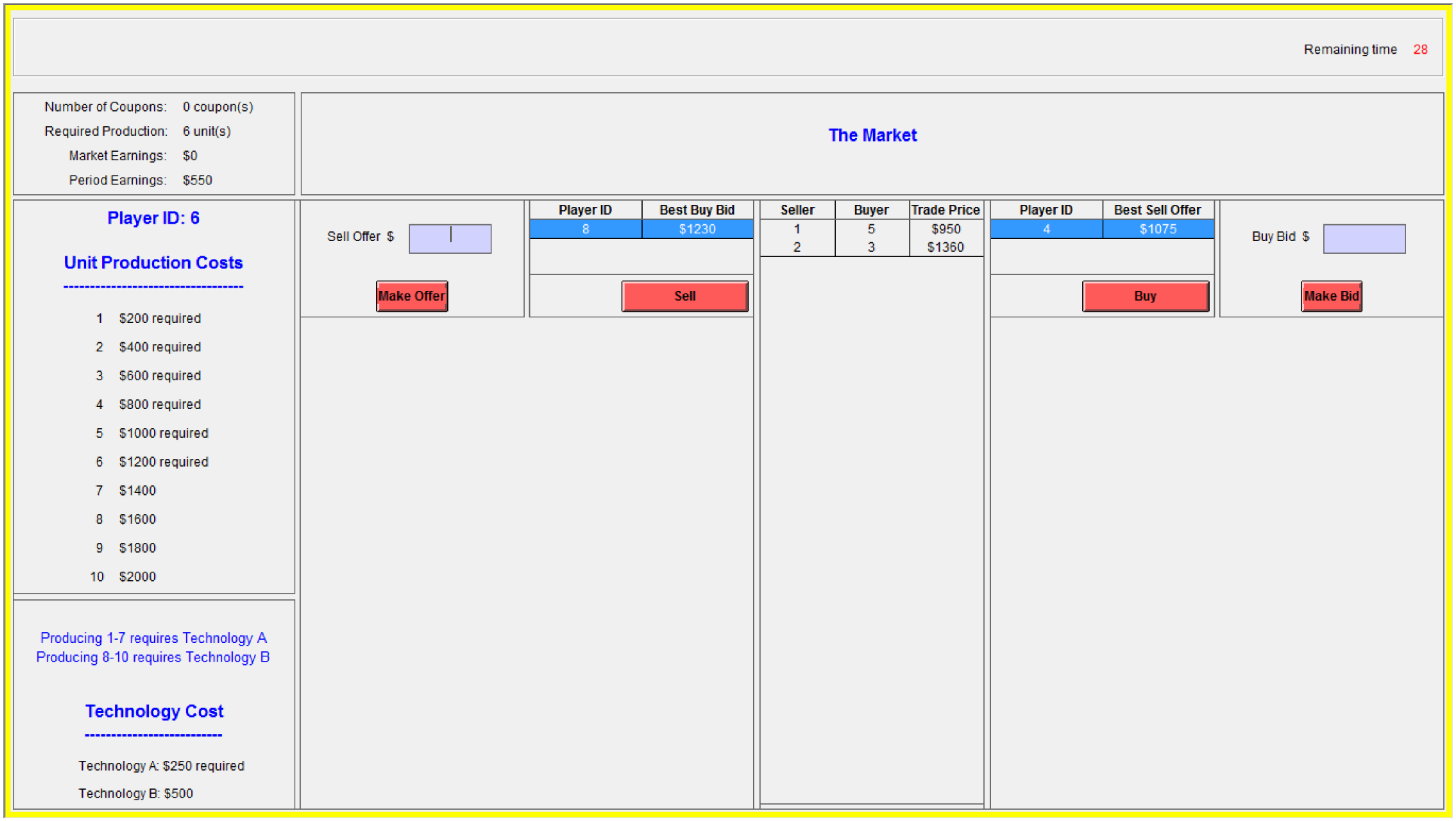


Figure 3. Example Market Trading Screen for a coupon generator (Part B)