Collusion in Auctions for Emission Permits: An Experimental Analysis

Burtraw, Dallas, Jacob K. Goeree, Charles A. Holt, Erica Myers, Karen Palmer, and William Shobe *

February 2009

Abstract

Environmental markets have several institutional features that provide a new context for the use of auctions and which have not been studied previously. This paper reports on laboratory experiments testing three auction forms – uniform and discriminatory price sealed bid auctions and an ascending clock auction. We test the ability of subjects to tacitly or explicitly collude in order to maximize profits. Our main result is that the discriminatory and uniform price auctions produce greater revenues than the clock auction, both without and with explicit communication. The clock appears to facilitate successful collusion, both because of its sequential structure and because it allows bidders to focus on one dimension of cooperation (quantity) rather than two (price and quantity).

Introduction

Environmental policy makers are increasingly using market-based approaches to regulation of the environment and natural resources. These approaches, long advocated by economists, typically involve the creation of a limited property right for the use of some publicly regulated environmental resource. These "permits"[1] are transferred to individual agents who, in turn, are expected to manage the resources they own so as to maximize their economic value. The socially-determined environmental goals are met by setting the number of permits to be allocated and by enforcing the condition that any use of the environmental resource must be covered by the ownership of the requisite number of permits. These approaches, broadly referred to as 'cap and trade' programs are

* Burtraw, Myers, and Palmer: Resources for the Future, Holt: and Shobe: University of Virginia, and Goeree: Caltech. This research was funded in part, by NYSERDA and the National Science Foundation (SES 0098400). We wish to thank Andrew Barr, AJ Bostian, Ina Clark, Kendall Fox, Courtney Mallow, and Lindsay Osco for research assistance.

[1] Tradable emission rights are generally referred to as 'allowances'. The experiments reported on here used the more context-neutral word 'permits.' Consequently, we will refer to them as permits throughout.

expected to induce users to select low-cost measures for environmental improvement and to efficiently allocate the use of resources. As a result, these approaches have been shown to substantially reduce the overall costs of environmental regulation relative to traditional command and control approaches to regulation.

The use of environmental markets has expanded to include such diverse areas as the management of fisheries, allocation of hunting licenses, access to eco-tourism, management of water resources, and regulation of thermal and nutrient pollution in streams. The most economically important environmental markets have emerged in the management of air pollution. A crucial feature in the design of these programs is the mechanism by which property rights to environmental assets such as emissions permits are distributed initially. The approach used in most previous emissions trading programs has been free allocation of permits to shareholders of incumbent emitting facilities, a process known as "grandfathering." This approach has been subject to a broad critique based on results from the economics literature showing that there may be very substantial efficiency gains from selling permits initially. Recent policy proposals rely on markets to determine the initial distribution of environmental assets.

The largest environmental market in the world is the European Union (EU) market for carbon dioxide (CO_2) emission allowances, known as the Emissions Trading Scheme (ETS), which had an annual asset value of approximately $61 billion in 2008. In the first and second trading periods of the ETS (from 2005-2012) regulations required that the vast majority of allowances be distributed for free (Ellerman & Buchner, 2007). However, ensuing criticisms have led the European Commission to propose a major revision that would replace free allocation to the power sector with the use of an auction beginning in 2013, and would expand this approach to most other covered emission sources by 2020. Proposals for use of an auction have also emerged as an important feature of legislative proposals for US climate policy.

The second mandatory CO_2 cap-and-trade policy in the world, and the first to require the widespread use of auctions, began in 2009 and includes the 10 northeastern states that formed the Regional Greenhouse Gas Initiative (RGGI). This policy covers CO_2 emissions from electricity generators within the region. CO_2 emissions are capped initially at levels comparable to those at the beginning of this decade and then ramped down to 10 percent below initial cap levels by 2019. A number of other multi-state initiatives to limit carbon emissions are also underway. These programs have served as models for the development of programs at the national level to limit greenhouse gas emissions.

The RGGI proposal represents a substantial break with the past. Rather than give the permits (termed "emission allowances" in RGGI) away for free, the RGGI states have decided to auction close to 90 percent of their annual CO_2 permit budgets. As the first greenhouse gas cap-and-trade program to start with a substantial revenue-generating auction of permits, this initiative will have a global impact that will be felt beyond efforts to control climate change. Annual allowance auctions have been held since 1995 under the well-known SO_2 cap and trade program in the U.S., but no revenue is raised as all proceeds are refunded. The first known cases of allowance sale to raise revenues are the over-the-counter sale of NOx allowances by Kentucky starting in 2004 and the 2004 auction of NOx allowances by Virginia. (Holt, Shobe, & Smith, 2006; Porter et al., 2009)

A general measure of the efficiency of an auction is its ability to elicit bids that reflect actual valuations by bidders and thereby allocate resources to their highest-valued use. Bidding true willingness to pay is a sufficient, but not necessary condition for allocative efficiency. Collusive bidding drives a wedge between bidders' actual values for the goods being sold and the bids made at auction. Goods may still be allocated to their highest valued use, but this would depend on perfect collusion that reduced the bid below willingness to pay equally for all bidders. If successful, collusive bidding also will result in goods being sold for prices below the value that bidders place on those resources, thereby lowering revenues for the seller. The possibility of this outcome presents significant political risk to a government agency selling a good because revenues may be considerably below those expected, and the auction will not be viewed as a fair and transparent process for selling government assets, in this case, CO_2 permits. A collusive auction outcome can interfere with the efficiency of existing markets by generating false price signals and by increasing price volatility. These considerations are especially important when new classes of environmental assets are being created by governments and allocated to final users through market instruments.

The specific question we address in this paper is whether the type of auction used to sell environmental assets will affect the likelihood of collusive behavior by auction participants. In a series of experiments, we investigate the likelihood of collusive outcomes for a variety of auction types. We compare the performance of single round auction formats, both discriminatory ("pay as bid") and uniform price (all pay the highest rejected bid), with the multi-round ascending price English clock auction. In the single-round sealed-bid auctions, bidders submit bids that specify prices for blocks of permits, so the bids have both price and quantity dimensions. In the multi-round auctions, bidders submit quantities at a given price. If demand exceeds supply, the price is raised (analogous to advancing the time on a clock) until there is no excess demand.

It is difficult to compare the revenue generating and efficiency properties of sealed-bid, multi-unit discriminatory and uniform price auctions from a theoretical perspective because the strategy spaces are complex and multiple equilibria exist. There is currently no theoretical evidence to support using one auction type over the other as there is no clear ranking between the two in terms of revenue or efficiency (Ausubel & Cramton, 1998; Back & Zender, 1993; Bikhchandani & Huang, 1989; Binmore & Swierzbinkski, 2000).

Experimental studies comparing sealed-bid uniform and discriminatory auctions suggest that the preferred auction type for revenue generation may depend on market characteristics particular to the good being sold. The first experimental study of alternative auction formats is reported by Smith (1967). In his setup, bidders were rere restricted to bid for two units, which had a known value common to all bidders after the auction. The main finding is that when there is a small amount of excess demand (rejected bids), uniform price sealed-bid auctions generate more revenue than discriminatory (pay as bid) auctions. This difference disappears in competitive settings when there is large excess demand. Cox, Smith, & Walker (1985) compare both single unit and multi-unit demand for the uniform and discriminatory auction formats. They find that revenue drops for both auction types when the multi-unit demand is introduced, but find no difference in average auction revenues between the two formats. Miller and Plott (1985) consider another dimension, the elasticity of demand. They find that when demand is inelastic, discriminatory auctions generate more revenue than uniform auctions and when demand is elastic, uniform auctions are superior to discriminatory auctions. Results of Goswami, Noe, & Rebello (1996) suggest that uniform price auctions may raise more revenue when only tacit collusion is possible, but when explicit collusion is possible discriminatory raises more.

The early experimental papers were largely focused on comparisons of sealed bid auctions that were relevant for Treasury auctions, but multi-round auctions have received considerable attention as well. The theoretical equilibria of the sealed-bid uniform price auction and the English clock auction (also a uniform price auction) are indistinguishable (Porter & Vragov, 2006). However, because these auction types yield multiple equilibria in a multi-unit context, as bidders attempt to coordinate their actions (tacitly or explicitly), different auction designs may result in systematically different outcomes. For example, the dynamic nature of the English clock auction may aid in price discovery, especially in auction environments for new assets where a price has not already been established (Cramton, 1998). In addition, bidders in clock auctions interact in multiple rounds, and if attempting to collude they would coordinate their bids along only one dimension (quantity), which may make collusion easier than in sealed-bid, single round auctions where bids include two dimensions (quantity and price).

An experimental study by Alsemgeest, Noussair, & Olson (1998) suggests that the clock auction may be more susceptible to tacit collusion. They compare the performance of the sealed-bid uniform price auction with the English clock auction in environments with both single-unit and two-unit demand for each bidder. Communication between bidders is not allowed, but bids and bidder identification numbers are made public after each round of the clock auction and at the end of each sealed bid auction. The authors find that the sealed bid auction generates more revenue than the clock auction for both single and two unit demand.[2] Goeree, Offerman, and Sloof (2006) report a multi-unit auction experiment designed to assess the likelihood of demand withholding in a very simple environment with no opportunities for communication. In their experiment, groups of 3 participants were competing for 6 available units. Each bidder received a randomly determined private value for three units, so the total demand at a zero price could be as high as 9. The experiments were either run as discriminatory auctions or as clock auctions in which bidders were informed when anyone reduced their bid quantities. Tacitly coordinated "demand reduction" tended to stop the clock at low prices, and auction revenues were much higher in the discriminatory auction, i.e. 151 versus 40 for the clock, a difference that was statistically significant at the 1 percent level. In addition, there was less revenue variability across auctions in the discriminatory treatment. The revenue advantage of the discriminatory auction even persisted in other asymmetric treatments in which two of the bidders in each group were "incumbents" with strong incentives to bid high and exclude the third bidder ("entrant"). This paper did not consider a uniform price sealed bid auction, but the authors summarize the literature relevant to this comparison: "Pooling the results from these different studies suggests that demand reduction is more pronounced in ascending auctions than in uniform-price auctions." (p. 2)

[2] Computer simulations can also be used to study auction performance under pre-specified behavioral assumptions. For example, Sunnevåg (2003) simulates bidding behavior in ascending price auctions for emissions allowances when there is an oligopolistic product market. The simulations compare outcomes when firms bid sincerely and strategically, and find the profit-maximizing bidding strategy is sensitive to the relative emission intensity of production for the bidders as well as the rules of the auction.

In sum, the theoretical ranking of these different auction mechanisms remains ambiguous. Markets for environmental assets are relatively new and there is little experience with the use of auctions for distributing these assets. Our experimental analysis of the relative performance of these auctions with market characteristics similar to environmental markets and in particular the new CO_2 market provides valuable information to regulators.

We ran our first laboratory sessions with simple, generic, competitive market conditions. For example, cost varied randomly across auctions, emission permits were not bankable between auctions, there was no secondary market for trading outside of the auction, and there was no opportunity for explicit communication. In this setting the auction forms tested generated similar results with little evidence of differences in efficiency or revenues across auction types. There are, however, some particular features of emissions permit auctions focused on a single emitting sector, the electricity sector, that may be conducive to collusion: 1) These auctions are likely to be conducted on a regular basis. The RGGI auctions that begin in September 2008 will be held quarterly, for example. Collusion may be easier with a longer series of auctions. 2) Cost conditions are relatively stable at the firm level from one auction to the next, and such stability may facilitate collusion. 3) Permits are tradable, so auctions are preceded and followed by active spot markets, which let bidders acquire needed permits if collusion fails, e.g. if a bidder bids low and is overbid by others.[3] Also, to the extent that collusion involves equal price bids resulting in some randomness in quantities won at auction, there can be inefficiencies in the auction outcome, which can be remedied later in the spot market. 4)

[3] For example, looking at timber auctions, Haile (2001, 2003) finds that the opportunity for resale has important implications for behavior in an auction.

Permits are bankable, so the risk of getting no permits due to a failed attempt to collude is mitigated. 5) Although explicit collusion is illegal, it is possible to hide appeals among participants to cut back on bids by expressing these appeals in terms of the need to protect the environment and pollute less. (6) Compliance need only be demonstrated every third period. Generators may run a permit deficit in any period that is not a compliance period (or a surplus in any period). This provides additional protection against negative consequences of not winning permits in a given auction. Moreover, many of the industry representatives feel some antagonism to new caps on emissions, and these representatives meet regularly to discuss the management of the transmission grid and the functioning of the electricity spot markets.

We expect these institutional features could be important, and therefore, we ran two more series of laboratory sessions with conditions intended to capture some of these special features. These sessions had a longer series of auctions than our first series, the cost draws for bidders did not change between auctions, each auction was followed by a spot market, permits could be banked, and there were fewer bidders than in the baseline setting. In addition, the last series of laboratory sessions were conducted with "chat room" opportunities to communicate and collude explicitly. We compare the performance of single round auction formats, both discriminatory ("pay as bid") and uniform price (all pay the highest rejected bid) in theses settings, with the multi-round English clock auction.

Our main result for these sessions with a rich environment is that the discriminatory and uniform price auctions produce greater revenues than the clock action, both without and with explicit communication. The clock appears to be more subject to successful collusion because of its sequential structure and because it allows bidders to focus on one dimension of cooperation (quantity) rather than two (price and quantity). The experimental procedures and results are reported in the next several sections, followed by a more detailed discussion of the extent to which the observed data patterns are relevant for RGGI and other auction-based emissions control programs.

Procedures

Auctions for emissions permits are multi-unit auctions, since the blocks of permits being sold in a given auction are identical. Multi-unit auctions can be distinguished by whether or not there are multiple rounds of bidding, and by whether all winning bidders pay the same "uniform" price or whether they pay what they bid, which is termed a "discriminatory" price rule. In the experiment, we focused on the three auction formats that have received the most attention for emissions policy: discriminatory price sealed bid, uniform price sealed bid, and clock (multi-round, uniform price). In a discriminatory auction, bidders submit sealed bids on blocks of permits, and these are ranked from high to low. The highest bids are declared winners, and those bidders have to pay their own bid prices. In a uniform price auction, sealed bids on blocks are collected and ranked, with the cutoff for winning again being determined by the number of blocks being sold. The difference is that winning bidders need only pay a common, market-clearing price, which in the experiments was the highest rejected bid. In both formats, a bidder is free to bid differing amounts on blocks, subject to a restriction that bids exceed an announced reserve price. In contrast, the multi-unit clock auction begins at the reserve price, and bidders are asked to state the number of blocks desired at that price. If the total demand exceeds the amount being auctioned, the clock price is raised by a pre-announced bid increment. In the next round bidders can reduce their (quantity) bid or leave it unchanged, and this process of increasing prices and re-bidding continues until demand is at or below the auction quantity. In order to force bidders to participate actively in early rounds, there is an activity rule that prevents them from increasing their bid quantities from one round to the next. To maintain comparability, bids in the sealed-bid auctions were restricted to be at price levels determined by fixed bid increments above the same reserve price that was used in the clock auctions.

The auctions were evaluated in a stylized setting that was intended to capture key aspects of the market for permits, while keeping the setup simple enough to be relatively transparent for subjects. The experiments involved a total of 324 subjects who earned an average of about $30 for a session lasting about one and a half hours. Each experimental session involved either 6 or 12 participants, recruited from the undergraduate population at the University of Virginia. Participants were paid $6 for showing up, in addition to earnings from purchasing the auctioned "permits" at prices below their values in the experiment. Each participant was given the role of a firm with multiple "units" of capacity that could be used to produce a product that sold at a known price. The use of each capacity unit required that the person obtain permits.

To keep the experiment from becoming too complicated, we used relatively small numbers of permits; with 60 permits being sold in each auction (or 30 permits in the sessions with only 6 bidders). Thus each permit in the experiment corresponds to a block of hundreds or thousands of "emissions permits" in practice. For example, RGGI allowances will be sold in lots of 1,000. All bidders were given 5 capacity units, but we introduced an asymmetric cost of compliance by requiring some subjects to obtain more permits to operate capacity than others. In particular, half of the subjects were "low users," who needed *one* permit for each capacity unit, and half were "high users" who were required to obtain *two* permits to operate each of their capacity units. Low users provide a simple representation of generators that use natural gas, and the high users represent generators that use coal, which produces approximately twice the CO_2 emissions as gas per unit of electricity produced. The equal numbers of low and high users was intended to roughly mimic the relative sizes of coal and gas generators in the region. This asymmetry is also important because coordinating collusive demand reductions may be more difficult in the presence of asymmetries that make it harder for bidders to agree on how to share the burden of demand reductions.[4]

Production costs for each unit were randomly generated in each round, in order to ensure that comparisons among auctions were not driven by particular configurations of production costs. The difference between the known price of the product and the randomly generated cost is the operating margin before permit costs, and permit values are determined by taking this margin and dividing by the required number of permits to operate a unit of capacity. For example, with a production cost of 5 and a price of 12, the margin is 7, and the value for a permit would be 7 for a low user who requires one permit to operate the capacity unit, whereas the value for each permit would be 3.5 = 7/2 for a high user who is required to have two permits to operate. The production costs for low users were set to be roughly twice as high as the costs for high users, to reflect the higher costs associated with natural gas generation. This cost difference also served to

[4] For example, Mason, Phillips, and Nowell (1992) report that subjects in the asymmetric duopoly games were less cooperative than was the case for symmetric Cournot markets.

approximately equalize earnings across subjects with different roles. The costs for low users were randomly drawn from the interval [5, 10], with all values in this interval being equally likely, and the costs for high users were drawn from the interval [2, 6]. With a fixed output price, the distribution of costs determines a range of permit values. Since costs are drawn from the range [5, 10] for low users, with all draws in this range being equally likely, then a product price of 12 will result in a range of permit values between 2 (= 12-10) and 7 (= 12-5). The values for high users are obtained by dividing operating margins by the required number of permits (2) per capacity unit, so a cost distribution from the range [2, 6] results in values between 3 (= (12 – 6)/2) and 5 (= (12 – 2)/2).

Using these basic value distributions and bidder types, we ran three series of experiments, with setup parameters shown in Table 1. The baseline series used a simple, more competitive setup with 12 bidders, with no explicit communication and costs were re-randomized for each round within a session, and with no banking or spot markets.

Table 1. Combinations of Settings and Auction Mechanisms Considered

Baseline Setting	Richer Setting	Richer Setting
No Communication	No Communication	Communication
12 bidders	6 bidders	6 bidders
60 permits	30 permits	30 permits
re-randomize costs	same costs	same costs
no banking or spot	banking and spot	banking and spot
8 auctions per session	12 auctions per session	12 auctions per session
Uniform (3 sessions)	Uniform (6 sessions)	Uniform (6 sessions)
Discriminatory (3 sessions)	Discriminatory (6 sessions)	Discriminatory (6 sessions)
Clock (3 sessions)	Clock (6 sessions)	Clock (6 sessions)

The two rich environment series were focused on the three auction formats that have been used (in some form or another) in previous emissions permit auctions (Burtraw & Palmer, 2006). In the second setting, shown in the center column of the table, we cut the numbers of bidders and permits in half, to increase the opportunities for tacit collusion, although explicit collusion was not permitted in this second setting. These sessions used the same uniform distributions of random cost draws as in the baseline, but there was only a single set of cost draws done at the start of each session, so a given bidder's costs for capacity units remained the same in all auctions. The possibility of tacit collusion was further enhanced by running more auctions, 12, in each session, and having a spot market that followed each auction, to let bidders acquire needed permits that were not obtained in the auction. We ran six sessions for each of the auctions formats: uniform sealed bid, discriminatory sealed bid, and English clock. The random draws were balanced in the sense that we used the same sequence of random number "seed" values for each auction format, so the random cost draws for the first clock session match the random draws for the first uniform price and discriminatory sessions. A new set of random draws was used for the second session in each format, etc.

The third setting, shown in the right column of Table 1, used the same 6 sets of random number draws and was the same as the second in other dimensions, except that explicit communications were permitted. This was done by letting bidders communicate in an electronic chat room for 2 minutes prior to each of the sealed bid auctions (discriminatory and uniform). Bidders were allowed to communicate for 1 minute prior to the first round of each clock auction and for an additional 15 seconds prior to each subsequent round of bidding in the clock auctions. The clock auctions generally only lasted for a few rounds, so the total time available for chat was roughly comparable. An alternative would have been to have no chat between rounds in a clock auction, but this would ignore the fact that if explicit collusion occurs, it may continue during a clock auction.

The spot markets, that followed the auctions in the second and third series, were structured so that participants could submit limit orders that specify a maximum quantity of permits and a maximum purchase price or a minimum sales price, e.g. sell up to 6 permits for at least $4. Buy orders were arrayed from high to low, sell orders were arrayed from low to high, and the price determined by the intersection of these arrays was the price at which transactions were executed. Then after the spot market cleared, subjects decided how many permits to use in production, and whether to bank permits or incur a deficit. It was announced that any deficit in permits was penalized at a rate of $9 (about three times the Walrasian auction price prediction) after the spot market that followed auctions 3, 6, 9, etc.[5]

Performance Measures

The economic problem posed by an emissions cap is to achieve the reduction with minimal cost, i.e. by maximizing the economic value associated with the limited supply of permits. For a firm, the value is represented by the difference between the product price and the production cost of the added output generated with an additional permit. The actual cost draws for each auction were used to calculate the maximum possible surplus that can be achieved with the 60 permits available. This maximum possible surplus can be represented by arraying the permit values from high to low as a demand function, drawing a vertical line at the fixed supply, and adding up the total area under the demand curve to the left of this line, indicated by the sum of the two shaded areas in the graph in Figure 1 for a hypothetical set of permit demands. In an actual auction, the permits may not all go to bidders with the highest values, and the actual efficiency is then calculated to be the actual value achieved as a percentage of the maximum possible value. In sessions with spot markets, the efficiency calculations are based on the permits actually used, after any spot market transactions following each auction.

[5] In only 5 out of a total of 144 compliance periods included in all of our experiments in the richer setting, did a participant end up having to pay the noncompliance penalty.

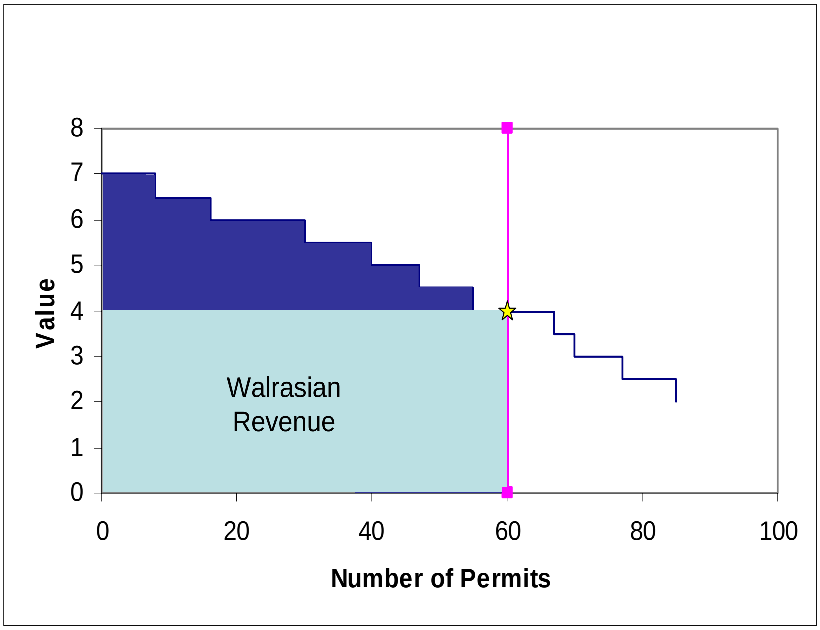

Figure 1. Walrasian Equilibrium and Revenue from the Auction

A second measure of auction performance is based on sales revenues, which is important since tacit or explicit collusion may result in low bids and revenues. If all bidders were to bid full value for each permit in a "pay as bid" auction, then the revenue would equal the area under the demand curve to the left of the vertical supply, again indicated by the sum of the two shaded areas in Figure 1. This area represents the maximum possible surplus that could be captured by the seller. A more reasonable benchmark is the revenue that would be generated if permits sell for the "Walrasian" or market-clearing price determined by the intersection of demand and supply. This Walrasian price is calculated numerically, based on the highest rejected bid value and the associated amount of revenue is indicated by the shaded rectangle in Figure 1. Note that the Walrasian revenue will be less than 100% of the maximum revenue and how much less depends on the steepness of the demand curve to the left of the market clearing price.

We analyze efficiency and revenue separately because they are distinct measures of auction performance, and an auction that yields high levels of efficiency may or may not yield revenues that would be expected in a competitive market.

Table 2: Summary Statistics

	Mean	St. Dev.	Min	Max
Baseline Setting				
Efficiency (proportion of maximum surplus)	.975	.018	.928	1.00
Revenue (experimental $)	204	14.9	165	246
Average Price (experimental $)	3.40	2.48	2.75	4.10
Richer Setting With No Chat				
Efficiency (proportion of maximum surplus)	.959	.039	.641	1.00
Revenue (experimental $)	90.4	15.4	42.0	128
Average Price (experimental $)	3.03	.493	2.00	4.25
Richer Setting With Chat				
Efficiency (proportion of maximum surplus)	.949	.038	.655	1.00
Revenue (experimental $)	77.5	15.2	40.0	128
Average Price (experimental $)	2.61	.475	2.00	4.27

Table 2 summarizes the performance measures for each of the three settings. These statistics indicate that the richer setting with no communication results in lower revenues than the baseline setting. To see this, note that the size of the market and thus the maximum surplus/revenue in the richer setting was half that of the baseline setting, but it can be seen that lower prices lead to lower relative revenue generation as well. In addition, the mean efficiencies are 2-3% lower in the richer settings than in the baseline settings. In the auctions with communication, the mean of the average price is lower than for auctions in the richer setting with no communication.

In the section that follows, we consider results the three treatment categories in sequence: baseline, rich environment with no communication, and rich environment with communication. In each case, we will present two sets of statistical tests: 1) nonparametric Wilcoxon tests using a single data point for each session, constructed as the average revenue for the last half of the auctions in that session, and 2) panel data regressions using a random effects specification to incorporate the fact that auction outcomes for the same session are not statistically independent. In order to construct Wald tests for separate treatment effects, we ran a single regression that incorporated baseline, rich no-chat, and rich chat auctions, which included dummy variables for all treatments and interactions, with the Clock dummy variable omitted. The structure and coefficient estimates for this regression is presented in Appendix A, and all Wald tests supporting empirical conclusions are derived from testing restrictions on those coefficients.

Results

Baseline Setting

The left panel of Figure 2 shows the average revenue percentage for the final 4 auctions, where each bar corresponds to a particular session. There are 3 sessions for each of the auction formats. Notice that revenue percentages are generally close to the Walrasian prediction of 79 percent, illustrated by the darker line. The revenues in the clock auction appear to be slightly lower than the revenues in the other two auction types, however we fail to reject the null hypotheses of equality between the uniform and clock coefficients (Wald χ^2 = 1.74, p=.1875) and the discriminatory and clock coefficients (Wald χ^2 = 1.12, p=.2898). The efficiency percentages for the final 4 auctions in each session are shown on the right side of figure 2, and we see no clear differences among auction formats, with all efficiencies being close to 100 percent.[6] These tests are the basis for our first conclusion:

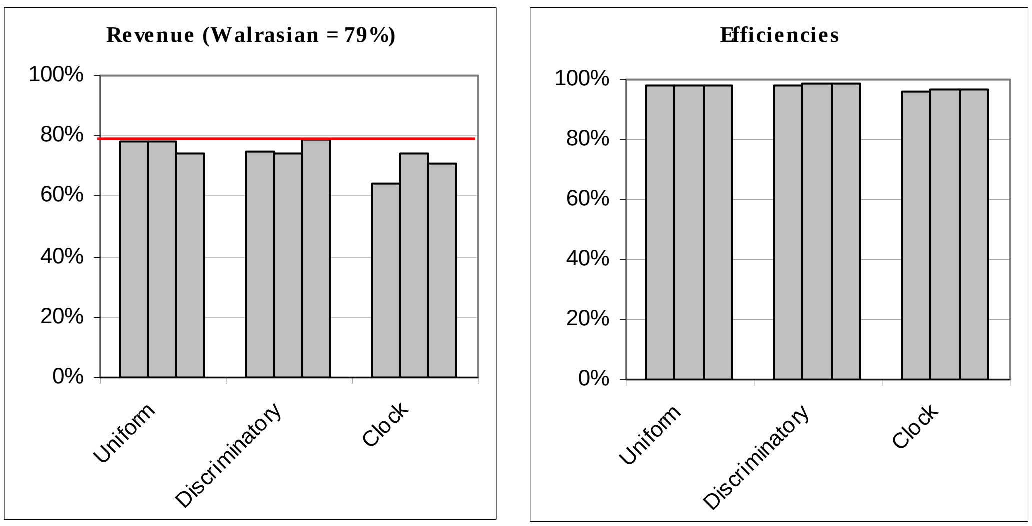

Figure 2: Average Revenues and Efficiencies by Auction Type for Baseline Setting (averages over 4 auctions in each session)

Baseline Performance: *There are no significant differences between auction formats in the competitive baseline environment, either in terms of revenue or efficiency. Revenues are near Walrasian predictions, and efficiencies are close to 100%.*

[6] In addition to the sets of 3 sessions shown reported here, we also ran a second set of sessions in which cost draws were made from a narrower range, which yielded a narrower range of permit values. The results of this "elastic demand" treatment are reported in Holt, Shobe, Burtraw, Palmer, & Goeree (2007). This treatment also failed to reveal any clear performance differences between auction formats in terms of revenue and efficiency.

These results differ slightly from those of Anderson and Holland (2006) who analyze auctions for fishing rights in New Zealand. They also find the discriminatory and uniform auction formats to be equally efficient, but they find the discriminatory auction to yield higher revenue than the uniform auction.

Richer Setting with No Communication

Recall that averages of revenues percentages were sharply lower in the rich environment with no chat than in the baseline setting. This effect is captured in panel regression (AR(1) GLS random effects) shown in (1), which uses revenues for all auctions in the baseline and rich (no-chat) sessions. The dependent variable, *Rev*, has been normalized by dividing the auction revenue by 2 for the 12-person baseline sessions to make them comparable to the 6-person session in the rich environment with longer series of auctions, spot markets, banking, and compliance penalties. The dummy variables, *Uniform*, *Disc*, and *Rich* take values of 1 for uniform price, discriminatory and rich settings, and Walrasian_Rev is the normalized Walrasian revenue prediction for the random draws for that auction. The omitted treatment dummy is for the Clock auction.

$$(1)\quad Rev = \underset{(24.01)}{48.93^{***}} + \underset{(4.65)}{17.73^{***}}\,Uniform + \underset{(4.65)}{11.10}\,Disc - \underset{(4.06)}{13.54^{***}}\,Rich + \underset{(220)}{400^{*}}\,Walrasian\_Rev$$

The significant negative coefficient on *Rich* indicates that revenues are lower in this environment, which is indicative of tacit collusion in the rich environment. The magnitude of this coefficient, \$13.54, is economically significant relation to levels of \$70-\$80 that were typically observed in 6-person auctions. The significant and positive coefficient on *Uniform* in (1) indicates that the uniform price auctions raise more revenue than Clock auctions (the omitted dummy), but this result masks the fact that auction format results vary by treatment. More precise results can be obtained from Wald tests on coefficient restrictions for the panel data regression in Appendix A that has all interaction terms included. The results can be summarized:

No-Chat Rich Environment: *Revenue generation is significantly lower than in the baseline series. In the no-chat environment, the revenues from Uniform and Discriminatory auction formats cannot be ranked, but both of these raise significantly more revenue than the Clock auction.*

Support: We find that overall the richer setting with no communication generates less revenue than the baseline setting (Wald χ^2 = 12.69, p=.0054). A within auction format comparison shows that both the discriminatory and clock auctions raise more revenue in the baseline setting than in the richer setting with no communication (Wald χ^2 =4.38, p=.0364; χ^2 =7.07, p=.0078), and we find no difference in revenue generation between the two settings in uniform auctions (Wald χ^2 =1.27, p=.2603). Table 3 below displays the results of between auction format comparisons in the richer setting with no communication. We also report two-tailed Wilcoxon rank sum tests, using revenue percentages averaged over the last six auctions (with each session producing a single data point). The Wilcoxon and Wald tests yield similar results with one exception: the Wald test indicates higher revenues for the discriminatory auction than for the clock auction, although this result is not apparent from the less sensitive Wilcoxon test based on a single average revenue for each session.

Table 3: Pair-Wise Revenue Comparisons for Last Six Auctions: Richer Setting With No Communication (p-values are reported in parentheses, where * and ** indicate significance at the 0.01 and .05 levels respectively)**

Findings	Wald chi^2	Wilcoxon Rank Sum
Uniform no chat > Clock no chat	14.50*** (0.000)	25** (0.026)
Discriminatory no chat> Clock no chat	4.97** (0.026)	32 (0.310)
Uniform no chat ~ Discriminatory chat	2.49 (0.115)	33 (0.3939)

Richer Setting with Communication

Figure 3 shows a comparison of revenues and efficiencies for the rich environment sessions with no chat (light gray) and with chat (dark gray). In each case, the order of the bars from left to right corresponds to the random number seed. Revenues are generally way below the Walrasian predictions, both with and without chat, which is a clear difference from the baseline results shown in Figure 1. In contrast, the efficiency remains close to that in the baseline.[7]

[7] The efficiency results reported in Figure 3 are with respect to the distribution of permits after the secondary market clears. We have also looked at the efficiency results prior to trading on the secondary market and we find that all auctions are close to 100% efficient with no meaningful differences across the different auction types.

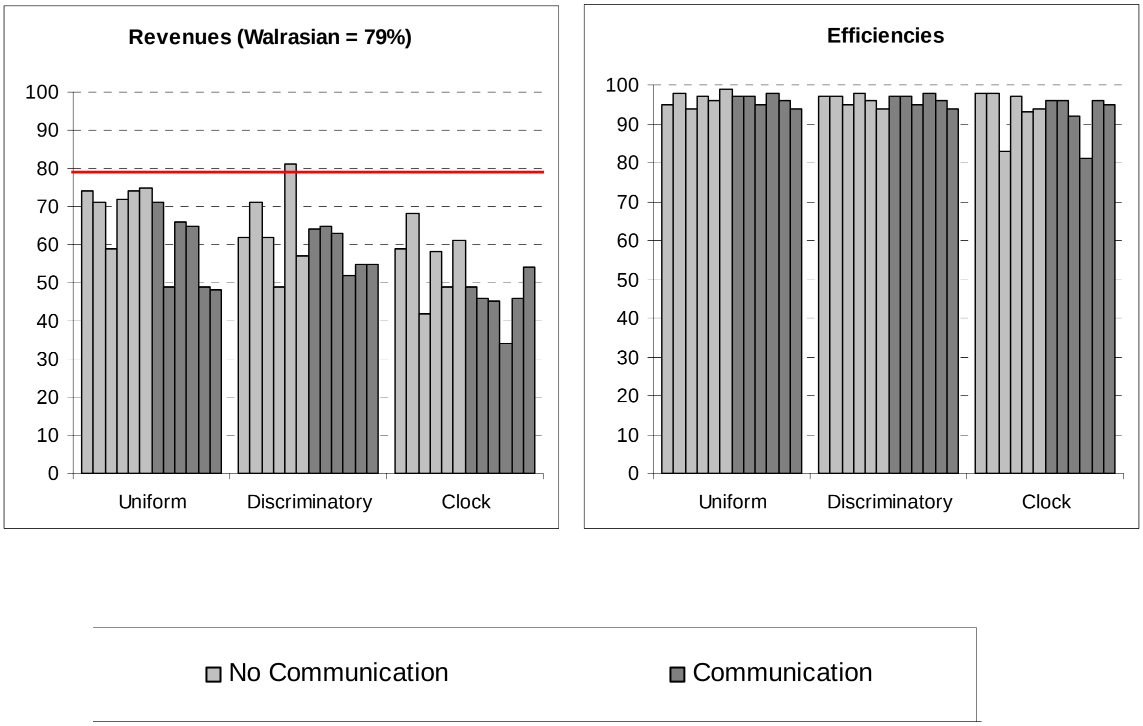

Figure 3: Revenues and Efficiencies for Last Six Auctions by Auction Type:

No Communication Versus Communication

Figure 4 shows average purchase prices for permits by setting with no communication in the left panel and with communication in the right panel. The effect of communication is to lower auction prices for all auction mechanisms. In addition the prices for the clock auction are lower both without and with communication. Notice that the average price for the discriminatory price auction converged to the average price for the uniform price auction in a few rounds. For reference, the average Walrasian prediction (where supply equals demand), averaged over all 6 sessions, is shown as a thick dashed horizontal line just above $3.50.

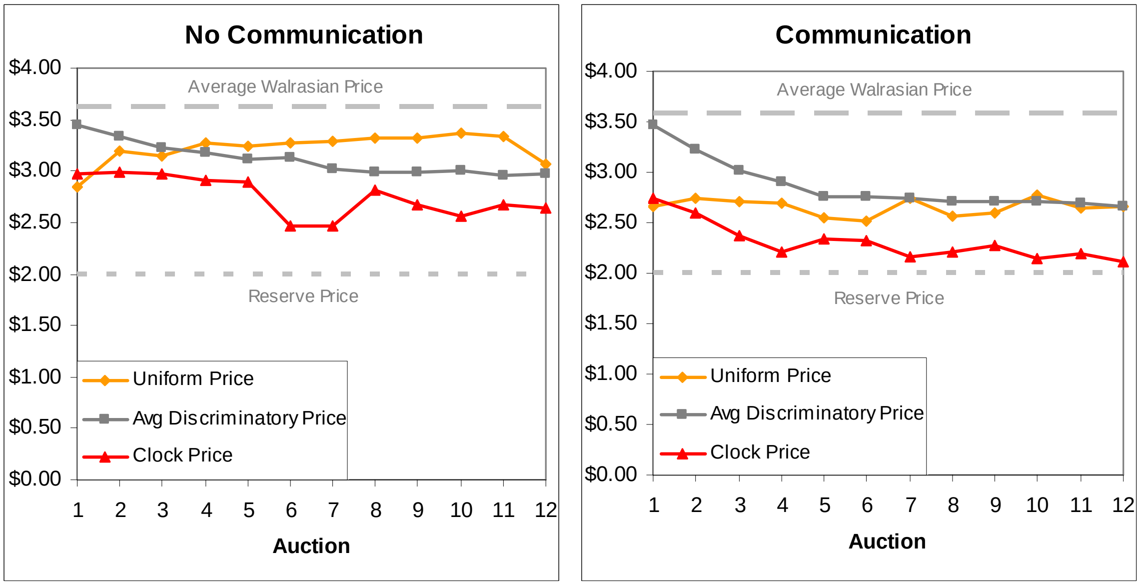

Figure 4: Average Price by Auction

As before, we begin with a simple panel data regression (GLS AR(1) random effects) that only includes data from auctions in the two rich environment treatments, with the Clock dummy omitted. The significant positive coefficients on *Uniform* and *Disc* indicate that these auction formats yield higher revenues that the Clock auction in this rich environment. The negative coefficient estimate for *Chat* is indicative of the effectiveness, on average, of explicit collusion. All of these coefficients are economically significant relative to the typical revenue levels of \$70-\$80 with no chat and \$60-\$80 with chat.

$$(2)\quad Rev = \underset{(26.66)}{44.85^{*}} + \underset{(4.28)}{19.29^{***}}\ Uniform + \underset{(4.28)}{15.52^{***}}\ Disc - \underset{(3.49)}{13.85^{***}}\ Chat + \underset{(243)}{294}\ Walrasian\_Rev$$

Specific comparisons among auction formats are derived from two-tailed Wilcoxon rank-sum tests using average revenues by session (last half of the auctions) and from Wald tests (for the panel data regression in the appendix with dummies and interactions). The results of these tests can be summarized:

Explicit Collusion: *Revenues were significantly lower in the clock auction with chat than in either the uniform or discriminatory auctions with chat, and 2) The effect of chat was to reduce revenues in both the uniform and clock auctions.*

Support: The relevant tests are shown in Table 4.

Table 4: Pair-Wise Revenue Comparisons for Last Six Auctions: (p-values are reported in parenthesis)

Findings	Wald chi^2	Wilcoxon Rank Sum
Chat		
Uniform chat > Clock chat	9.72 (0.002)	27 (0.056)
Discriminatory chat > Clock chat	11.15 (0.001)	23 (0.009)
Uniform chat ~ Discriminatory chat	.05 (.8251)	38 (0.900)
Chat vs. No Chat		
Clock chat > Clock no chat	6.92 (0.009)	28 (0.093)
Uniform chat > Uniform no chat	11.02 (0.001)	25 (0.024)
Discriminatory chat ~ Discriminatory no chat	2.32 (.1281)	31 (0.243)

The averages shown in Figure 4 to some extent mask the sharply increased tendency for clock auctions with collusion to stop at the reserve price or at one or two bid increments above the reserve, as shown as a histogram in Figure 5. Thirty-eight percent of the 72 clock auctions stop at the reserve price of $2.00 and the average price is $2.29, whereas the average prices for the uniform and discriminatory auctions are $2.77 and $2.83 respectively.[8, 9]

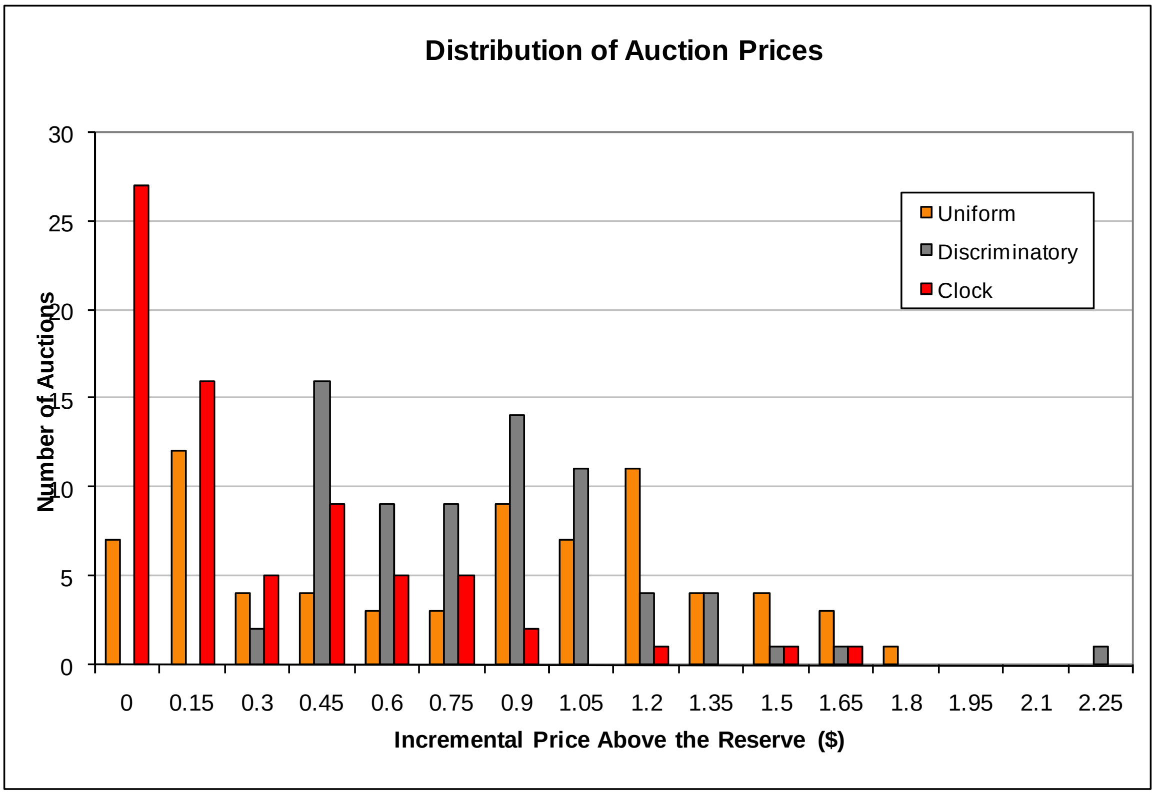

Figure 5: Price Distributions with Communication

The recorded chat between subjects provides some insight into why collusion is more successful in clock auctions. Most of the initial proposals made by participants were based on suggesting quantity reductions for low and high users. The focus on quantity reduction in the clock auction sessions is revealed in some of the participants'

[8] Average prices for each discriminatory auction were rounded to the nearest bid increment.

[9] These results are consistent with Haile (2000), who finds in a theoretical context a partial pooling of bids around a reserve price in auctions with a secondary market.

comments: "again, bid for fewer permits earlier on so we can get permits cheaper" and "this will go 5X faster and will all make LOTS more money if everyone just cooperates the first time." One person suggested "so why doesn't everyone bid exactly the same amount that we ended last round [auction] with, since we keep getting the same clearing price." This plan permitted participants to obtain the same final allocation without the run up in prices that occurred previously. Of course, quantity discussions occurred with the other auction formats too, but the effect of the clock is to take out the price dimension so that bidders only have to reach an agreement in a single dimension, quantity.

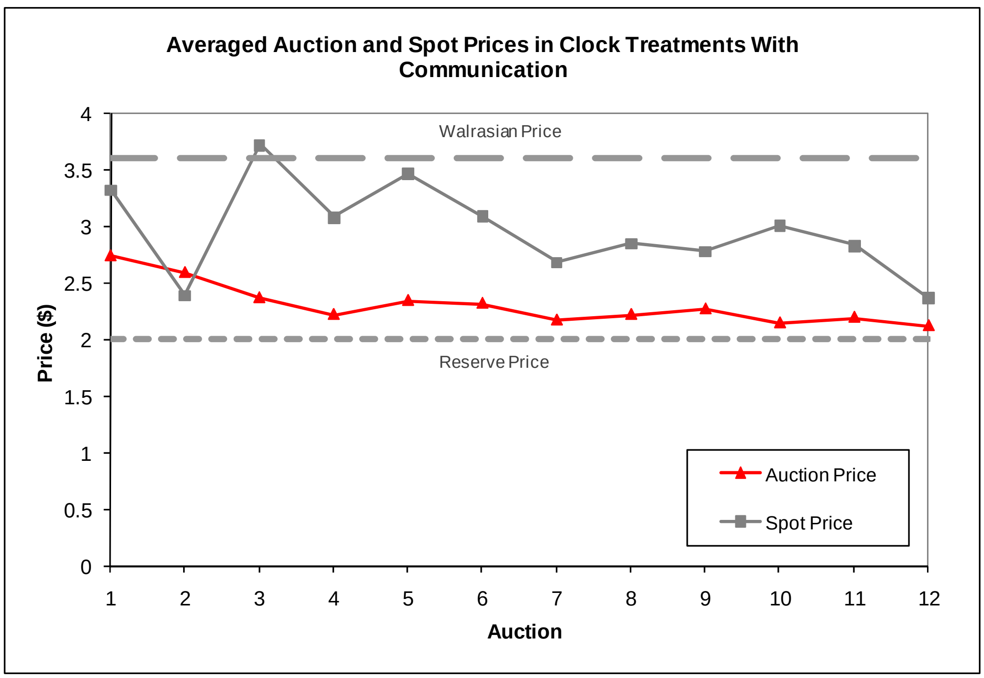

Figure 6. Collusion: Low Clock Auction Prices and High Spot Prices

One interesting feature of the data for the clock sessions with chat is that, while auction prices were typically near the reserve price, the subsequent trading in the spot market tended to be at much higher price levels that were closer to the Walrasian price, as indicated in Figure 6. The auction settings where the spot price is higher than the auction price are the clock with both chat and no chat and the uniform with chat, as shown in Table 5. Interestingly, these are the three auction settings that we found to be statistically significantly lower than other auction types, i.e. these are the three auction settings that appear on the right sides of the inequalities for the bullet point tests below each of the key results listed above. The bidders in these settings were defeating the auctions with successful collusion, which typically results in inefficient allocations that were, to some extent, corrected by trading in the spot markets.

Table 5: Spot vs. Auction Prices Averaged Across All Auctions

(Average Walrasian Price = $3.60)

Auction Setting	Auction Price ($)	Spot Market Price ($)
Uniform No Communication	3.21	3.20
Discriminatory No Communication	3.11	3.07
Clock No Communication	2.75	3.01
Uniform Communication	2.66	2.86
Discriminatory Communication	2.86	2.74
Clock Communication	2.31	2.96

There are several widely cited cases in which coordinated demand reductions are used to defeat ascending price auctions for broadcast spectrum. Just prior to the 2001 Austrian third generation mobile telecommunication spectrum auction, Telekom Austria announced that it "… would be satisfied with just two out of the 12 blocks for offer and if the [five] other bidders behaved similarly, it should be possible to get the frequencies on sensible terms … but that it would bid on a third block if one of its rivals did …" The other bidders clearly understood and the bidding stopped after a couple of rounds, with each bidder obtaining 2 blocks (Klemperer, 2004, p. 136). The extreme symmetry of this situation differs from the bidding environment faced by bidders in RGGI auctions, but the success of collusion triggered by a public announcement is disturbing.

One issue in the implementation of an English clock auction is whether to announce the excess demand, that is, the amount by which the bid quantities exceed the number of units for sale at each round. It has been suggested that releasing this information can enhance the price discovery function of the English clock auction. The clock auction experiments reported here did not provide excess demand information to auction participants. On average, the clock auction prices observed with collusion were quite close to the reserve levels, and therefore we did not think it worthwhile to redo these sessions with the additional information that would likely facilitate collusive behavior. With ex post quantity information after each round, bidders would be able to "signal" an intent to cooperate by reducing their demands to stop the clock early in one auction, in the hopes of inducing reciprocal cooperation from others in subsequent auctions.

In fact, experience from the Virginia NOx auction conducted in 2004 suggests that end-of-round quantity information may even induce bidders acting individually to reduce quantities bid in order to stop a clock auction. In the NOx auction, bidders were not provided with end-of-round quantity information. Late in the Virginia NOx auction, there was one bidder with a large share of the allowance requests; this bidder could have ended the auction by unilaterally reducing demand by a relatively small increment. As the clock price increased, this bidder apparently began shading requests by relatively small amounts, in an apparent attempt to "feel around" for the edge needed to stop the auction. We believe that this bidder would have ended the auction sooner, with a quite modest "demand reduction," if the needed information about demand "overhang" had been available (Porter, Rassenti, Shobe, Smith, & Winn, 2009).

One possible justification for the release of excess demand information in a clock auction could be that it may help "price discovery." We have recently finished conducting a series of sessions, after several auctions with stationary conditions, there is an unanticipated demand shift that raised permit values for some bidders and not for others. We had thought releasing excess demand information in a clock auction would help bidders discover and react to the change during the first auction following the demand shift, but the two flavors of clock, with and without excess demand information, tracked the shift up in the Walrasian price equally well, and neither did noticeably better then a simple sealed-bid uniform-price auction (Burtraw, et al. 2009).

Conclusion

The increasing use of market instruments in managing environmental assets has lead, in turn, to increased attention to how these new environmental assets should be allocated. A large body of economic literature argues for charging for environmental assets rather than granting them for free, the usual practice until quite recently. The recent sales of NOx allowances and of EU CO_2 allowances represented a significant break from past practice, but proposals to auction allowances in the Regional Greenhouse Gas Initiative, in the EU ETS, and in many of the cap and trade proposals before

Congress point to a strong trend in favor of the auction of environmental assets. The experiments reported in this paper investigate how opportunities for collusion affect the choice of auction type in the allocation of environmental assets. One of the key innovations in this study is the institutional setting, which includes key features of markets for CO_2 allowances including permit banking, the presence of secondary markets, and compliance periods spanning multiple market periods. The addition of these institutional elements results in significant changes in the likelihood of collusive outcomes across auction types.

In laboratory auctions with communication among participants, collusion is more effective in clock auctions than in discriminatory and uniform price auctions. An analysis of the 'chat' (instant message communications suggests that clock auctions may facilitate collusion by allowing bidders to focus on a single dimension (quantity reductions). The effects of this collusion are reflected in clock prices at or near reserve price levels, with subsequent trading at much higher prices in the spot markets.

This research also supports two key points make in the literature on auctions of public resources. First, auction design must be responsive to the institutional context. Because each context will imply different information and different strategies available to participants, results for different auction forms may differ dramatically as between different institutional settings (Binmore & Klemperer, 2002). Second, in general, increasing the competitiveness of an auction will be associated with better auction outcomes (Whitford, 2007). The design of the institutional setting for the auction of environmental assets should emphasize features that increase competition among bidders.

Appendix A. Panel Data Regression for all Treatments

In order to assess the effect of the different auction formats on revenue generation in these settings, we estimated a panel regression model, restricting the data to auctions in the last half of each session (4 auctions for the baseline setting and 6 auctions in the richer environment). Because there were twice as many participants in the baseline environment than in the richer settings, we normalized the revenues and Walrasian revenues in the baseline auctions by dividing the values by two. We explain revenue generation as a function of normalized Walrasian revenue and treatment indicator variables:

$$(3)\quad Rev_{it} = \beta_0 + \beta_1 Uniform_i + \beta_2\, Disc_i + \beta_3\, Rich_i + \beta_4\, Uniform_i\, Rich_i + \beta_5\, Disc_i\, Rich_i + \beta_6\, Chat_i + \beta_7\, Uniform_i\, Chat_i + \beta_8\, Disc_i\, Chat_i + \beta_9\, Walrasian\_Rev_i + \varepsilon_{it}$$

In equation (3) above, Rev_{it} is the normalized revenue captured by the auctioneer in session *i* and auction *t*, and $Walrasian\_rev_{it}$ is the normalized revenue that would be captured if all units sold for the competitive Walrasian price. The constant term, β_0, represents the average revenue for the omitted auction (clock phase 1). $Uniform_i$ and $Disc_i$ take on the value of 1 when the session used the sealed-bid uniform price or discriminatory auction format respectively. The $Rich_i$ variable indicates auctions in the richer setting and the $Chat_i$ variable indicates sessions with explicit communication opportunities. We estimate the model using a GLS panel regression where the error term, ε_{it}, is assumed to have an AR1 correlated structure with an unobserved, random effects component. The results of the estimation of equation (3) are displayed in Table 6 below. The coefficients on the treatment dummies are interpreted as the incremental amount of revenue raised, on average, relative to the omitted auction (in this case clock in the baseline environment).[10]

Table 6. Revenue Capture for All Treatments:

AR(1) panel GLS with Random Effects

Constant [phase 1 clock no chat]	56.08*** (19.81)
Uniform	10.58 (8.030)
Discriminatory	8.491 (8.022)
Rich	-18.38*** (6.913)
Uniform*Rich	10.61 (9.770)
Discriminatory*Rich	3.921 (9.764)
Chat	-14.64*** (5.66)
Uniform*Chat	-3.837 (7.871)
Discriminatory*Chat	6.174 (7.871)
Normalized Walrasian Revenue	3.63** (.178)

[10] For example, a clock auction with chat in the rich environment raises on average $\beta_3 + \beta_6$ more/fewer experimental dollars than the baseline clock auction. Likewise, a uniform price auction with chat raises on average $\beta_1 + \beta_4 + \beta_7$ more/fewer experimental dollars than a clock auction with chat.

A more complete assessment of the effects of different combinations of settings and auction types requires an assessment of the effects of restrictions on combinations of the estimated coefficients. We use a Wald statistic to test hypotheses whether the unrestricted estimates of the coefficients violate restrictions on the coefficients by a significant amount.[11]

[11] We tested linear hypotheses in the form $Rb = r$ where b is the estimated coefficient vector. If V is the estimated variance-covariance matrix, the Wald statistic is $W = (Rb-r)'(RVR')^{-1}(Rb-r)$ which has a chi squared distribution and degrees of freedom equal to the number of restrictions. For example, to test if Uniform Chat revenues are different than Clock Chat revenues we use the Wald test to see if the sum of the coefficients on the uniform and uniform chat dummies is different from zero.

References

Alsemgeest, P., Noussair, C., & Olson, M. (1998). Experimental comparisons of auctions under single- and multi-unit demand. Economic Inquiry, 36(1), 87-97.

Anderson, Christopher M. and Daniel S. Holland, (2006). ”Auctions for Initial Sale of Annual Catch Entitlement,” Land Economics, 82(3),333 – 352.

Ausubel, L., & Cramton, P. (1998). Demand Reduction and Inefficiency in Multi-unit Auctions, Working Papers. College Park: University of Maryland Department of Economics.

Back, K., & Zender, J. (1993). Auctions of Divisible Goods: On the Rationale for the Treasury Experiment. The Review of Financial Studies, 6(4), 733-764.

Bikhchandani, S., & Huang, C.-f. (1989). Auctions with Resale Markets: An Exploratory Model of Treasury Bill Markets. The Review of Financial Studies, 2(3), 311-339.

Binmore, K., & Klemperer, P. (2002). The Biggest Auction Ever: The Sale of the British 3G Telecom Licences. The Economic Journal, 112, C74-C76.

Binmore, K., & Swierzbinkski, J. (2000). Treasury Auctions: Uniform or Discriminatory. Review of Economic Design, 5(4), 387-410.

Burtraw, D., & Palmer, K. (2006). Summary of the Workshop to Support Implementing the Minimum 25 Percent Public Benefit Allocation in the Regional Greenhouse Gas Initiative Discussion Papers. New York: Resources for the Future.

Burtraw, D., Goeree, J., Holt, C., Myers, E., & Shobe, W. (2009). Price Discovery in Emissions Permit Auctions, draft report, University of Virginia.

Cox, J. C., Smith, V. L., & Walker, J. M. (1985). Expected Revenue in Discriminative and Uniform Price Sealed-Bid Auctions. In V. L. Smith (Ed.), Research in Experimental Economics (Vol. 3, pp. 183-208). Greenwich, CT: JAI Press Inc.

Cramton, P. (1998). Ascending Auctions. European Economic Review, 42(3-5), 745-756.

Ellerman, A. D., & Buchner, B. K. (2007). The European Union Emissions Trading Scheme: Origins, Allocation, and Early Results. Review of Environmental Economics and Policy, 1(1), 66-87.

Goeree, J. K., Offerman, T., & Sloof, R. (2006). Demand Reduction and Preemptive Bidding in Multi-Unit License Auctions, Discussion Paper: CalTech.

Goswami, G., Noe, T., & Rebello, M. (1996). Collusion in uniform-price auctions: Experimental evidence and implications for treasury auctions. Review of Financial Studies, 9, 757-785.

Haile, Philip A. (2001). "Auctions with Resale Markets: An Application to U.S. Forest Service Timber Sales," American Economic Review*,* 91(3), 399-427.

Haile, Philip A. (2000). "Partial Pooling at the Reserve Price in Auctions with Resale Opportunities," Games and Economic Behavior, 33, 231-248.

Haile, Philip A. (2003). "Auctions with Private Uncertainty and Resale Opportunities," Journal of Economic Theory*,* 108, 72-110.

Holt, C., Shobe, W., Burtraw, D., Palmer, K., & Goeree, J. K. (2007). Auction Design for Selling CO2 Emission Allowances under the Regional Greenhouse Gas Initiative. Albany: New York State Energy Research and Development Authority.

Holt, C., Shobe, W., & Smith, A. M. (2006). An Experimental Basis for Public Policy Initiatives. In E. Patashnik & A. Gerber (Eds.), Promoting the General Welfare:

New Perspectives on Government Performance (pp. 174-199). Washington, DC: Brookings Institution Press.

Klemperer, P. (2004). Auctions: Theory and Practice. Princeton: Princeton University Press.

Mason, Charles F., Owen R. Phillips, and Clifford Nowell (1992). “Duopoly Behavior in Asymmetric Markets: An Experimental Evaluation,” Review of Economics and Statistics, 1992, vol. 74 (4), pages 662-670.

Porter, D., & Vragov, R. (2006). An experimental examination of demand reduction in multi-unit versions of the Uniform-price, Vickrey, and English auctions. Managerial and Decision Economics, 27(6), 445-458.

Porter, D., Rassenti, S., Shobe, W., Smith, V., Winn, A. (2009). The Design, Testing, and Implementation of Virginia's NOx Allowance Auction, Journal of Economic Behavior and Organization, 69 (2), 190 - 200.

Smith, V. (1967). Experimental studies of discrimination versus competition in sealed-bid auction markets. Journal of Business Research, 40, 56–84.

Sunnevåg, Kjell J. (2003). “Auction Design for the Allocation of Emissions Permits in the Presence of Market Power,” Environmental and Resource Economics*,* 26, 385-400.

Whitford, A. B. (2007). Designing Markets: Why Competitive Bidding and Auctions in Government Often Fail To Deliver. Policy Studies Journal, 35(1), 61-85.