Zero Intelligence Plus and Gjerstad-Dickhaut Agents for Sealed Bid Auctions

A. J. Bagnall and I. E. Toft
School of Computing Sciences
University of East Anglia
Norwich
England NR4 7TJ
{ajb,it}@cmp.uea.ac.uk

Abstract

The increasing prevalence of auctions as a method of conducting a variety of transactions has promoted interest in modelling bidding behaviours with simulated agent models. The majority of popular research has focused on double auctions, i.e. auctions with multiple buyers and sellers. In this paper we investigate agent models of sealed bid auctions, i.e. single seller auctions where each buyer submits a single bid. We propose an adaptation of two learning mechanisms used in double auctions, Zero Intelligence Plus (ZIP) and Gjerstad-Dickhaut (GD), for sealed bid auctions. The experimental results determine if a single agent adopting ZIP & GD bidding mechanisms is able to learn the known optimal strategy through experience. We experiment with two types of sealed bid auctions, first price sealed bid and second price sealed bid. Quantitive analysis shows that whilst ZIP agents learn a good strategy they do not learn the optimal strategy, whereas GD agents learn an optimal strategy in first price auctions.

1. Introduction

The increase in the level of Internet connectivity has allowed the WWW to become a hub for electronic trading places. Buyers and sellers are now able to trade in previously inaccessible markets. Some of the important questions facing market overseers and traders are: what are the optimal strategies for a given auction structure; how do agents learn the optimal strategy; and how does restriction of information prevent agents from learning a strategy? These questions have been addressed through auction theory [7, 12], field studies [8], experimental lab studies [9], and agent simulations [2, 4]. Recently there has been particular interest in the study of agents for continuous double auctions (CDA) [2, 3, 5, 6, 10].

We adapt learning mechanisms developed for CDA to investigate agent architectures for sealed bid auctions. We broadly classify agent architectures in the following way: *memory free agents, memory based agents* and *modelling agents*.

The simplest type of agent stores no explicit information about past auctions and simply reacts to the previous auction outcomes. These so called *memory free* reactive agents have been examined extensively in [1, 2]. The second type of agent we call *memory based* agents. These agents store some historical information about auctions and adjust their strategy based on some estimate of a global picture of auction outcomes. They are considered to be more sophisticated than *memory free* agents and have been used in [5, 10]. The third type of agent we consider to be a *modelling* agent, these agents also store information about past auctions. Rather than using the market information directly these agents form models of competitors behaviour to estimate the correct action or strategy (for example, see [6]).

It is our belief that prior to examining agent behaviour in complex, dynamic multi-agent systems, any agent architectures should be tested in learning environments where a known optimal strategy exists. In this paper we examine the success of a single adaptive agent in learning the optimal strategy when competing against a population of non-adaptive agents. We use the Private Values Model (PVM) for auctions because, under some constraints, there are provably optimal strategies. These strategies provide a metric with which we may assess the ability of an adaptive agent in learning a strategy. They also are an obvious choice of strategy for the non-adaptive agents.

The rest of this paper is structured as follows: Section 2 describes the auction model and the simulation structure. Section 3 details how the memory free ZIP [2] algorithm has been adjusted for sealed bid auctions and assesses how well it performs in simulations. Our experiments demonstrate that the complexity of the problem is such that memory free agents learn a good, but suboptimal strategy. Section 4 describes the memory based Gjerstad-Dickhaut (GD) [3] algorithm for sealed bid auctions. The GD agents perform better than ZIP agents, and on one class of sealed bid auctions learn an optimal strategy. In Section 5 we present our conclusions.

2. Simulated Auction Model

The PVM [12] is commonly assumed in auction research. For a PVM auction of N interested bidders, each bidder (or agent) i has a valuation x_i of the single object. Each x_i is an observation of an i.i.d. random variable X_i with range $[0, \phi]$ (ϕ is the universal maximum value) and distribution function F, assumed to be identical for all bidders.

Open auctions (auctions that allow bidders to observe other agents' bids) such as English auctions (ascending price) and Dutch auctions (descending price) allow for multiple bids by each bidder. Under the PVM, open auctions have strategically equivalent sealed bid auctions (auctions where each bidder can submit at most one hidden bid). Hence we restrict our attention to sealed bid auctions. An agent i forms a bid b_i with a bid function

$$\beta_i : [0, \phi] \rightarrow \Re+, \qquad \beta_i(x_i) = b_i.$$

The set of all bids for a particular auction is denoted $B = \{b_1, b_2, \ldots, b_N\}$. The winning agent, w, is the highest bidder,

$$w = \arg\max_{i \in N}(b_i \in B).$$

We consider two auction formats which differ in their method of price determination. In First Price Sealed Bid (FPSB) auctions, the winner pays the price they bid, i.e.

$$p = \max_{i \in N}(b_i \in B) = b_w.$$

Under the PVM, FPSB auctions are strategically equivalent to open Dutch auctions. In Second Price Sealed Bid (SPSB) auctions, the winner pays the amount bid by the *second* highest bidder

$$p = \max_{i \in N, i \neq w}(b_i \in B).$$

Under the PVM, SPSB auctions are strategically equivalent to open English auctions.

The benefits of the PVM are that for certain auction mechanisms and assumptions there is provably optimal behaviour. Hence we can measure the performance of intelligent adaptive agents and assess under what conditions learning is most effective. This is a necessary condition to studying more interesting (and realistic) scenarios where the assumptions under the PVM concerning the competitors behaviour do not necessarily hold true.

In any auction an agent's profit (or reward) is

$$r_i(x_i) = \begin{cases} x_i - p & \text{if } i = w \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

The problem facing an agent is to find the bid function that will maximize profit. If we assume that all the agents are using the same bid function, a symmetric equilibrium is a strategy, β^*, where no single bidder can do better by using any other function. A symmetric equilibria represents an optimal bidding strategy for any agent competing against other agents following the optimal strategy. For FPSB auctions the symmetric equilibrium is

$$\beta^*(x_i) = E[Y_{N-1} | Y_{N-1} < x]$$

where Y_{N-1} is the largest order statistic of the other $N-1$ bidders. When F is a uniform distribution on $[0, 1]$ the symmetric equilibrium strategy is

$$\beta(x_i) = \frac{N-1}{N} \cdot x_i$$

The symmetric equilibrium strategies in a SPSB auction are given by

$$\beta(x_i) = x_i.$$

The optimal strategy for a SPSB auction is independent of the form of the distribution function F and does not require that all bidders have the same value function. Proofs and a more complete description of auction formats are given in [7].

Our objective is to determine the most simple adaptive agent structure that, when competing with $N-1$ non-adaptive agents, is best able to learn the optimal strategy in simulated FPSB and SPSB auctions. The experimental structure we adopt is consistent with that used in simulations with human agents [8,9]. Agents compete in a series of k auctions indexed by $j = \{1, \ldots, k\}$. Each agent is aware of the distribution function common to all bidders, F, the universal maximum value, ϕ and the number of competitive bidders, N. For any auction j each bidder i is assigned a value $x_{i,j}$ by sampling F. Once the auction is complete, every agent is informed of the winning agent w, the price the winning agent must pay p and their own reward as defined by Equation 1. No other information relating to the other agents' bids is made available, agents are also unaware of the number of auctions in any experiment.

3. ZIP Agents for Sealed Bid Auctions

The ZIP algorithm has been primarily used for agents in auctions with multiple buyers and sellers [2]. We adapt the architecture for use in auctions with just a single seller.

The key feature of ZIP agents is that they learn a bidding strategy based only on information provided by the results of the previous auction (or shout) and their private value. It is our primary interest to assess whether simple agents that retain no explicit memory of previous auctions, such as ZIP agents, can learn an optimal strategy in sealed bid auctions.

A ZIP agent adopts a linear bid function given by

$$b_a = x_a(1.0 - \mu_a)$$

where μ_a represents the fraction above or below its value at which the agent bids. The optimal strategies, in terms of μ, are $\mu = \frac{1}{N}$ in FPSB auctions and $\mu = 0$ in SPSB auctions. The problem of learning an optimal strategy is reduced to learning these optimal margins. For a ZIP agent a in an auction j the margin $\mu_{a,j}$ is adjusted to $\mu_{a,j+1}$ using the rule

$$\begin{aligned} \mu_{a,j+1} &= \mu_{a,j} + \Delta_{a,j} \\ \Delta_{a,j} &= \beta(d_{a,j} - \mu_{a,j}) \end{aligned} \tag{2}$$

$d_{a,j}$ is the desired margin, i.e. the margin a bidder a should have adopted in an auction j to maximize its reward $r_{a,j}$. We calculate $d_{a,j}$ by first calculating, or estimating, the optimal bid $o_{a,j}$. The desired margin is then defined as

$$d_{a,j} = 1 - \frac{o_{a,j}}{x_{a,j}} \ . \tag{3}$$

In selecting an optimal bid $o_{a,j}$ the ZIP agent considers any bid which could have increased their reward. When a ZIP agent loses an auction ($a \neq w$) and the price was greater than its value ($p_j > x_{a,j}$) there is no bid which could have increased the agents reward, since the best reward achievable is 0. Hence we restrict adaptive agents to updating in two situations.

1. *The agent wins* ($a = w$). We characterise this situation as the agent being *greedy*, hence it increases its margin by estimating the optimal bid to be lower than the current bid.
2. *The agent loses* ($a \neq w$) but could have made a profit ($p \leq x_{a,j}$). In this situation the agent is *fearful* and becomes more cautious. The agent reduces its margin by estimating an optimal bid to be greater than the current bid $b_{a,j}$.

Large variations in margin can result from the wide range of observable optimal bids, hence ZIP agents employ a *momentum* co-efficient γ_a to smooth the update variable. $\Delta_{a,j}$ is replaced by $\Gamma_{a,j}$ in the first line of Equation 2. $\Gamma_{a,j}$ is determined by the formula

$$\Gamma_{a,j+1} = \gamma_a \Gamma_{a,j} + (1 - \gamma_a)\Delta_{a,j}.$$

where $\gamma_a \in [0, 1]$. Larger values of γ result in greater smoothing (i.e. reduce the effect on the margin of the current update). The update on the margin then becomes

$$\mu_{a,j+1} = \mu_{a,j} + \Gamma_{a,j}.$$

ZIP agents select their optimal bid by randomly sampling a range of values given by

$$o_{a,j} = b_{a,j} \cdot R_a + A_a$$

where R and A are observations of independent random variables with a uniform distribution. When a ZIP agent wins $A \in [A_{min}, 0.0]$ and $R \in [R_{min}, 1.0]$. If the agent loses $A \in [0.0, A_{max}]$ and $R \in [1.0, R_{max}]$.

The only difference between the ZIP agent for CDA described in [2] and the ZIP agent for sealed bid auctions is that the situations in which it is allowed to update are different.

3.1. ZIP Results

The ZIP agent competes in first and second price sealed bid auctions versus fifteen non-adaptive agents. Our simulations have the following experimental parameters: agents compete in a series of 10,000 auctions; private values are drawn from a uniform distribution function F with a universal maximum $\phi = 1.0$; the ZIP agent adopts a learning rate $\beta = 0.1$; an initial margin $u_1 = 0.5$; a momentum co-efficient $\gamma_a = 0.7$; $R_{max} = 1.05$, $R_{min} = 0.95$, $A_{max} = 0.05$ and $A_{min} = -0.05$; the number of agents $N = 16$ is fixed throughout a run.

First Price Sealed Bid

In FPSB auctions the objective of the ZIP agent is to learn the optimal margin $\mu = \frac{1}{16}$. Figure 1 shows the ZIP traders margin for a single run where the straight line represents the optimal strategy. Table 1 shows a quantitive summary of accumulated profits received by all agents. It can

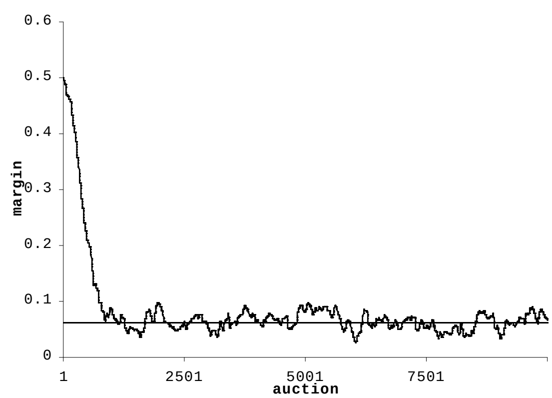

Figure 1. Plot of μ for ZIP type agent in FPSB auctions.

be clearly seen from Figure 1 that the margin converges towards to a value close to the optimal in ~1200 auctions.

ZIP Agent	Fixed Strategy Agents			
Profit	Min	Max	Mean	Std Dev
17.75	16.94	19.34	18.36	0.75

Table 1. Profits for ZIP Agent in FPSB auctions.

The average margin over the final 5000 auctions is 0.063, suggesting that the agent may be learning too greedy a strategy. Table 1 show that although the profit of the adaptive agent is within the range of the optimal agents, it is nevertheless below the average figure. To test further whether the ZIP agent was reaching the optimal strategy, we repeated the experiment 100 times and measured the average margin and the profit achieved over the last 5000 auctions. The sample mean margin was 0.0648 and the median 0.0649. We can reject the null hypothesis that the average level is 0.0625 using a t-test for the mean and a Wilcoxon sign rank test for the median at the 1% level.

To demonstrate that the reduced margin does actually result in a suboptimal strategy and hence less profit, the profit obtained by the ZIP agent was compared to the average profit made by the optimal agents. Over 100 runs the mean and median ZIP profit was 17.95 and 17.87 respectively, compared to the mean and median average optimal agent profit of 18.40 and 18.41. We can reject the null hypothesis that the mean difference between the profit figures is zero using a t-test for paired samples and the null hypothesis that the median difference is zero using the Wilcoxon sign rank test for paired samples at the 1% level. Hence we can conclude that for FPSB auctions, a ZIP agent competing against non-adaptive agents following the optimal strategy learns on average a suboptimal strategy and hence receives on average a lower profit than the average obtained by the optimal agents.

Second Price Sealed Bid

In SPSB auctions the optimal margin is $\mu = 0$. Figure 2 shows the ZIP agent learns a margin close to optimal after $\sim$1200 auctions. The average margin over the final 5000 auctions is 0.0084, which suggests that the ZIP agent is being marginally too greedy. Table 2 shows the profit of the ZIP agent is within the range of profits achieved by the optimal agents, although as with FPSB it is below the average figure.

As with FPSB, a sample of 100 independent experiments was conducted. The mean and median margin were found to be significantly different from the optimal margin of zero and the mean and median profit for the adaptive agent (17.95 and 17.92) to be significantly less than the mean and median profit of the non-adaptive agents (18.58 and 18.56).

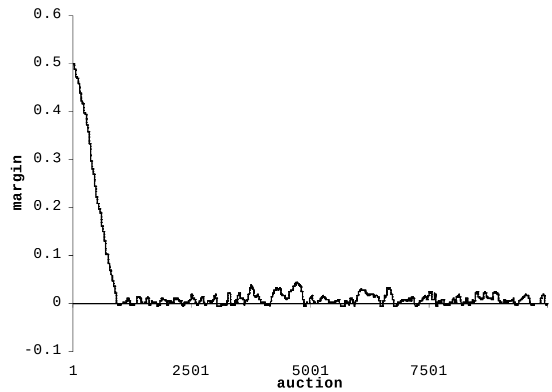

Figure 2. Plot of μ for ZIP type agent in SPSB auctions.

ZIP Agent	Fixed Strategy Agents			
Profit	Min	Max	Mean	Std Dev
17.01	15.48	21.24	18.59	1.62

Table 2. Profits for ZIP Agent in SPSB auctions.

From these experiments we conclude that the ZIP agent is able to learn an average margin close to optimal for both FPSB and SPSB auction. However, there is evidence that the average margin is significantly different to the optimal and the profit achieved is significantly less than that achieved by the non-adaptive agents. ZIP agents perform at least as well as the similar memory free agents used in [1]. In [11] we fully evaluate the effect on margin and profit of the use of momentum and alternative margin updates and find that there is no simple implementation that achieves optimal profit. We also demonstrate the effect of alternative values of N.

4. GD Agents for Sealed Bid Auctions

The Gjerstad-Dickahaut (GD) trader algorithm for CDA is a *memory based* agent architecture described in [3] and refined in [10]. GD traders have a strategy for shout price selection based on maximising expected profit. The maximisation of expected profit relies on the GD trader forming a belief and payoff function. A GD trader maintains a history H of length m storing messages pertaining to the last L transactions. For double auctions, GD traders employ their memory of previous shout and transaction prices to form a belief function $\hat{q}(b)$. For a particular bid history a buyer forms a belief that a bid b will win is

$$\hat{q}(b) = \frac{TBL(b) + AL(b)}{TBL(b) + AL(b) + RBG(b)}$$

where $TBL(b)$ is the number of transacted bids less than or equal to b, $AL(b)$ is the number of asks less than or equal to b and $RBG(b)$ is the number of rejected bids greater than or equal to b. A trader forms a profit function on the space of possible bids, $r^*(b)$, then finds the expected profit function. It selects the bid that maximizes expected profit.

The GD algorithm requires some alteration for sealed bid traders because of the different information available. The agent maintains an auction history H of length m. Suppose an agent were able to record the price paid by the winner, p, and the winning bid, b_w in the history H. For a FPSB auction $p = b_w$, but in a SPSB $p \leq b_w$. With this historcial information we can adopt a similar assertion to GD in forming a belief function: If a bid b_w was the winning bid then a bid $b \geq b_w$ would also have won. An agents' belief that a bid b will be accepted is

$$\hat{q}(b) = \frac{T(b)}{m} \tag{4}$$

Where the $T(b)$ is the number of auctions in which b would have been a winning bid in previous m auctions. $\hat{q}(b)$ is an estimate of the the probability of winning the auction with a bid of b or less. The GD trader evaluates $\hat{q}(b)$ for the values of b given by the winning bids in H. For example, if the bid history contained the 5 winning bids and prices for a SPSB auction

$$H = \{(0.85, 0.9), (0.78, 0.8), (0.64, 0.65), (0.86, 0.95), (0.83, 0.88)\}.$$

$\hat{q}(b)$ is estimated at points $b = \{0.65, 0.8, 0.88, 0.9, 0.95\}$ to be

$$\hat{q}(0.65) = 0,\ \hat{q}(0.8) = 0.2,\ \hat{q}(0.88) = 0.4,\ \hat{q}(0.9) = 0.6,\ \hat{q}(0.95) = 0.8\ ,\ \hat{q}(b) = 1\ ;\ b > 0.95$$

The agent then estimates the true payoff function were it to win,

$$r^*(b) = x - p$$

with the empirical function $\hat{r^*}$ at the same values of b. Based on the prices in the bid history and the current value, x, $r^*(b)$ is estimated as

$$\hat{r^*}(b_w) = x - p\ \ \forall (p, b_w) \in H.$$

Using the same example history as before, an agent with value 0.8 would estimate the payoff function at

$$\hat{r^*}(0.65) = 0.16,\ \hat{r^*}(0.8) = 0.02,\ \hat{r^*}(0.88) = -0.03,\ \hat{r^*}(0.9) = -0.05,\ \hat{r^*}(0.95) = -0.06$$

The expected profit is then the product of the probability of winning and the estimate of the profit assuming the agent won,

$$\hat{E}(b) = \hat{q}(b) \cdot \hat{r^*}(b).$$

The agent selects the bid that maximizes the estimated expected profit. Thus the expected profit for each bid in the history given above would be

$$\hat{E}(0.65) = 0,\ \hat{E}(0.80) = 0.004,\ \hat{E}(0.88) = -0.012,\ \hat{E}(0.90) = -0.030,\ \hat{E}(0.95) = -0.048$$

The agent would choose to bid 0.8. In cases when the maximum expected reward is less than or equal to zero, the agent chooses a bidding margin based on a summary of the margin it had adopted in previous auctions.

The problem with this method is that in sealed bid auctions the agent is only informed of the price the winner paid. In FPSB auctions the price and the winning bid are the same, hence the procedure can be followed as described. However, in SPSB auctions the agent is only aware of both price and winning bid if it won the auction itself. This means that a further adaptation of the bidding mechanism is required. The profit function is estimated based on a separate record of the auctions the agent itself won and hence has accurate pairs of winning bids and prices. The estimate of the probability of a bid winning is based on the history of auctions H, but the points at which it is estimated, b_w are now also estimated based on the experience of the agent. A more detailed implementation description can be found in [11].

4.1. GD Results

The GD agent was assessed with the same experimental parameters as the ZIP agent. The history length, m, was set to 1000.

First Price Sealed Bid

Table 3 shows the profit achieved by a GD agent over the last 5000 auctions for a typical run. Over 100 runs the mean and median GD profit was 18.24 and 18.27 respectively, compared to the mean and median average optimal agent profit of 18.34 and 18.33. We cannot now reject the

GD Agent	Fixed Strategy Agents			
Profit	Min	Max	Mean	Std Dev
18.20	17.26	19.45	18.34	0.73

Table 3. Profits for GD Agent in FPSB auctions.

null hypothesis that the mean difference between the profit figures is zero.

We can conclude that for FPSB auctions, a GD agent competing against non-adaptive agents following the optimal strategy learns a strategy that is not significantly different to the optimal and hence receives on average a profit as high as the average obtained by the optimal agents.

Second Price Sealed Bid

The results for a single run with a GD agent in SPSB auctions is given in Table 4. Over 100 runs the mean and median GD profit was 17.77 and 17.64 respectively, compared to the mean and median average optimal agent profit of 18.05 and 18.03.

The mean and median profit over 100 runs is significantly less than the optimal. It is interesting to note that the GD agent in SPSB frequently learns to bid above it's value. This means that the non-adaptive agents make significantly less profit when competing against GD as opposed to ZIP agents. This is likely to be an artifact of the method used to estimate the belief function, and is worthy of further investigation.

GD Agent	Fixed Strategy Agents			
Profit	Min	Max	Mean	Std Dev
17.89	15.11	21.14	18.51	1.50

Table 4. Profit for GD Agent in a single SPSB auctions.

Preliminary investigations indicate that this suboptimal performance is caused by the sensitivity of the agent to variation in the estimate it has to make of the winning bid to form a belief function.

5. Conclusion

This paper discusses the problem of autonomous adaptive agents learning the optimal strategies for sealed bid auctions under the constrained PVM model.

Single seller auctions provide an excellent test bed for assessing the ability of alternative agent architectures to learn complex behaviour. There is a well established theory about optimal behaviour in constrained forms of auction, extensive studies on how well human agents learn have been conducted and real world observations have been made from online auctions.

The inability of an adaptive agent to find an optimal strategy reflects real world problems with respect to learning behaviour in auctions. Field studies have shown that real world strategies are often suboptimal [8].

Our work in developing adaptive intelligent agents for simulated auctions will continue by studying behaviour in multi adaptive agent environments, considering more complex auction scenarios with for example, agents leaving and entering market and agents with changing reward functions.

References

[1] A. J. Bagnall and I. Toft. An agent model for first price and second price private value auctions. In *Proceedings of the 6th International Conference on Artificial Evolution*, pages 145–156, 2003.

[2] Dave Cliff. Minimal-Intelligence Agents for Bargaining Behaviours in Market-Based Environments. Technical report, June 1997.

[3] Steven Gjerstad and John Dickhaut. Price Formation in Double Auctions. *Games and Economic Behaviour*, 22(1):1 – 29, 1998.

[4] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119–137, 1993.

[5] Minghua He, Ho fung Leung, and Nicholas R. Jennings. A Fuzzy Logic Based Bidding Strategy for Autonomous Agents in Continuous Double Auctions. *IEEE Transactions on Knowledge and Data Engineering*, 15(6), 2002.

[6] Junling Hu and Michael P. Wellman. Multiagent Reinforcement Learning: Theoretical Framework and an Algorithm. In *Proc. 15th International Conf. on Machine Learning*, pages 242–250. Morgan Kaufmann, San Francisco, CA, 1998.

[7] Vijay Krishna. *Auction Theory*. Academic Press, San Diego, California, 2002.

[8] David Lucking-Reily. Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet. *American Economic Review*, 89(5):1063–1080, 1999.

[9] Vernon Smith. An Experimental Study of Market Behavior. *Journal of Political Economy*, 70(2):111–137, 1962.

[10] Gerald Tesauro and Rajarshi Das. High-Performance Bidding Agents for the Continuous Double Auction. In *Third acm Conference on Electronic Commerce*, pages 206–209, 2001.

[11] I. Toft and A. J. Bagnall. Adaptive agents for sealed bid auctions. Technical Report CMP-C04-03, 2004.

[12] William Vickery. Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance*, 16:8–37, 1961.