

Casari, Marco; Ham, John C.; Kagel, John H.

Working Paper

Selection bias, demographic effects, and ability effects in common value auction experiments

Staff Report, No. 213

Provided in Cooperation with:

Federal Reserve Bank of New York

Suggested Citation: Casari, Marco; Ham, John C.; Kagel, John H. (2005) : Selection bias, demographic effects, and ability effects in common value auction experiments, Staff Report, No. 213, Federal Reserve Bank of New York, New York, NY

This Version is available at:

<http://hdl.handle.net/10419/60705>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Federal Reserve Bank of New York
Staff Reports

Selection Bias, Demographic Effects, and Ability Effects
in Common Value Auction Experiments

Marco Casari
John C. Ham
John H. Kagel

Staff Report no. 213
June 2005

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Selection Bias, Demographic Effects, and Ability Effects in Common Value Auction Experiments

Marco Casari, John C. Ham, and John H. Kagel

Federal Reserve Bank of New York Staff Reports, no. 213

June 2005

JEL classification: C9, D44, C24, J16

Abstract

We find clear demographic and ability effects on bidding in common value auctions: inexperienced women are much more susceptible to the winner's curse than men, controlling for SAT/ACT scores and college major, but they catch up quickly; economics and business majors substantially overbid relative to other majors; and those with superior SAT/ACT scores are much less susceptible to the winner's curse, with the primary effect coming from those with below median scores doing worse, as opposed to those with very high scores doing substantially better, and with composite SAT/ACT score being a more reliable predictor than either math or verbal scores by themselves. There are strong selection effects in bid estimates for both inexperienced and experienced subjects that are not identified using standard econometric techniques but rather through our experimental design effects. Ignoring these selection effects is most misleading for inexperienced bidders, as the unbiased estimates of the bid function indicate much faster learning and adjustment to the winner's curse for individual bidders than do the biased estimates.

Key words: common value auction experiments, selection effects, econometric methods, gender and ability effects

Casari: Purdue University (e-mail: casari@purdue.edu). Ham: Federal Reserve Bank of New York, Ohio State University, and the Institute for the Study of Labor (IZA) (e-mail: ham.25@osu.edu). Kagel: Ohio State University (e-mail: kagel.4@osu.edu). Earlier versions of this paper were presented at the Meetings of the Econometric Society in Washington, D.C.; the Simposio de Analisis Economico in Seville, Spain; the Conference on Experiments and Econometrics at New York University; the 11th Annual Panel Data Conference; the 2004 Society for Information Technology and Teacher Education conference; and seminars at Columbia University, the Ohio State University, Purdue University, the University of Torino, and the University of Siena. The authors thank Linda Babcock, Rachel Croson, John Duffy, Dan Levin, Geert Ridder, Jeffrey Smith, Lise Vesterlund, and participants at meetings and seminars for valuable comments. The authors thank Serkan Ozbeklik and Alex Brown for excellent research assistance; Robert Vandyne, Jr., from Student Enrollment and Research Services at the Ohio State University, for his help in obtaining detailed demographic and ability data on students; and Jo Ducey for editorial and administrative assistance. This research has been supported by National Science Foundation Grant 0136928. The views expressed in this paper are those of the authors, and do not necessarily reflect the position of the National Science Foundation, the Federal Reserve Bank of New York, or the Federal Reserve System.

1. Introduction

Experimental studies of common value auctions are characterized by negative average profits and large numbers of bankruptcies for inexperienced bidders (a “winner’s curse”), a result which has been replicated by a number of researchers with a variety of subject populations and with a variety of auction institutions. It is only with experience that bidders learn to avoid the worst effects of the winner’s curse and earn a respectable share of the profits predicted under the risk neutral Nash equilibrium (see Kagel and Levin, 2002, for a review of the literature). The transition from inexperienced bidders suffering persistent losses to experienced bidders earning respectable profits is characterized by large numbers of bidders going bankrupt, with these bankrupt bidders much less likely to return as experienced subjects.

These results raise a number of substantive questions regarding bidding behavior in common value auctions. First, are there some sorts of ability or demographic factors that characterize “better” bidders (bidders who are less susceptible to the winner’s curse, or who learn to recognize and overcome it faster)? There is a nagging suspicion that this must be the case. If so, what exactly are these characteristics and can individual learning serve as a partial or complete substitute for ability on this score?

Second, is the transition from out-of-equilibrium behavior (the winner’s curse) to respectable positive profits a result of a “market selection” effect (i.e. less able bidders going bankrupt and not returning for subsequent sessions), or learning on the part of individual bidders, or some mix of the two (and if so what is the mix)? There is essentially no information on this score from past experiments as they have almost exclusively relied on experienced subjects self-selecting whether to return or not for experienced subject sessions. In addition to being of inherent interest, distinguishing between these alternative “adjustment” processes affects the kinds of learning models one needs to develop to characterize the evolution of behavior over time in common value auctions. It also has potential public policy implications, as legislation on corporate bankruptcy and on procurement contracts is often directly related to these issues. For example, in all European Union countries competition for government procurement contracts has been regulated with the explicit goal of fostering the acquisition of expertise and to minimize the chances of contractors’ bankruptcies (CEE Directive n. 37, June 13, 1993). One rationale for these rules is the belief that “market selection”, if left unchecked, operates to the detriment of social welfare maximization in some contexts and does not leave time for “individual learning” to take place. We return to this point in the concluding section of the paper.

In answering these questions we obtain a number of methodological insights. First, using standard econometric estimators for dealing with selection effects in field data we are unable to identify any kind of selection effects in our data, in spite of having a relatively large sample by experimental standards and well identified econometric models.¹ However, the different experimental treatments designed to identify potential selection effects in the data serve to identify, measure, and verify such effects. Thus, standard econometric techniques are simply not powerful enough to identify selection effects even with relatively large samples by experimental standards. But good experimental design can substitute for technique. These results suggest that previous estimates of bid functions for both inexperienced and experienced bidders are seriously biased.

Second, we find clear demographic effects as women are much more susceptible to the winner's curse as inexperienced bidders than men. However, with experience the women catch up to the men and do as well as men by the end of the experiment. Note that the gender effect identified here is obtained while controlling for obvious confounding factors such as ability and college major, factors that are not typically controlled for in investigating gender effects in experimental economics. (In fact, ours is the first study we are aware of that includes SAT/ACT scores to control for ability effects in the experimental economics literature.) We discuss several possible explanations for the differences between men and women, and conjecture that they result from differential experience in dealing with problems involving strategic interactions. In addition, economics and business majors are much more susceptible to the winner's curse than other majors, and continue to do worse (and suffer from a winner's curse) even as experienced bidders. We discuss a possible explanation for this result as well.

Third, we find clear ability effects in the data as measured by SAT/ACT scores. Although this is not surprising, the nature of these ability effects are different from what one might expect as (i) composite SAT/ACT scores achieve a more consistent, statistically significant impact in the regressions controlling for ability than either verbal or math scores do alone and (ii) the biggest and most consistent impact of ability comes as a result of those with below median scores being more susceptible to the winner's curse, as opposed to those with very high scores doing exceptionally well. The latter continues to be observed for experienced bidders, so that bidders with below median composite SAT/ACT scores continue, on average, to suffer from a winner's curse even as experienced bidders.

¹ Our inexperienced subject sample includes 251 individuals for up to thirty auctions, which is very large by experimental standards.

With regards to whether learning or selection effects are driving bidders' ability to overcome the winner's curse its quite clear that both processes are at work. First, there is substantial learning by both men and women over time, regardless of major or SAT/ACT scores, which in most cases permit bidders to avoid losses and earn respectable profits. Nevertheless, business and economics majors, as well as those with below median SAT/ACT scores continue, on average, to suffer from a winner's curse (and bankruptcies) even at the end of the second, experienced subject session in which they participated.

In our experiment all subjects were recruited to participate in two experimental sessions to be conducted at the same time and day in two consecutive weeks (week 1 and week 2). We used a variety of techniques to address selection issues both within weeks and between weeks: To address selection bias resulting from bankruptcies of inexperienced (week 1) bidders we randomized initial cash balances of the subjects, and induced random shocks to these balances within the experiment. These manipulations also enable us to identify any potential cash balance effects on bidding (which are minimal at best). To address selection issues that arise because only a subset of (experienced) subjects returns for week 2, we provided differential incentives for returning. That is, we introduced into the experimental design instruments that could potentially help to identify selection effects using relatively sophisticated estimators borrowed from the applied econometrics literature (Heckman, 1979, Ryu 2001) that would allow us to obtain unbiased estimates of the bid equation by simply using an appropriate sub-sample of our data.

The paper proceeds as follows. Section 2 specifies the risk neutral Nash equilibrium (RNNE) bid function for the experimental design, along with some measures of when subjects have fallen prey to the winner's curse. Section 3 outlines the experimental procedures and provides some general descriptive statistics providing an overview of the changes in bidding between weeks 1 and 2, and the potential selection effects present in the data. Section 4 describes our ability and demographic measures. Section 5 looks at the effect of these demographic and ability measures on the conditional probability of bankruptcy in each period among inexperienced bidders. Section 6 discusses selection effects and demographic and ability effects for inexperienced (week 1) bidders. Section 7 addresses the question of selection effects, as well as ability and demographic effects, for experienced bidders. Section 8 discusses the gender effect identified and relates it to the existing literature on gender effects in economic experiments. The concluding section of the paper summarizes our main results.

2. Theoretical Considerations: First-Price Sealed-Bid Auctions

In each auction period the common value, x_0 , was chosen randomly from a uniform distribution with upper and lower bounds [\$50, \$950]. Each bidder i was provided with a private information signal, x_i , drawn from a uniform distribution on $[x_0 - \$15, x_0 + \$15]$.² Using a first-price sealed-bid auction procedure, bids were ranked from highest to lowest with the high bidder paying the amount bid and earning profits equal to $x_0 - b_1$, where b_1 is the high bid. (Note these profits are negative when the winning bid exceeds x_0 .) Losing bidders neither gain nor lose money.

Wilson (1977) was the first to develop the Nash equilibrium solution for first-price common-value auctions, with Milgrom and Weber (1982) providing significant extensions and generalizations of the Wilson model. In the analysis that follows, we restrict our attention to signals in the interval $65 \leq x \leq 935$ (called region 2), where the great bulk of the observations lie.³

For risk neutral bidders the symmetric risk neutral Nash equilibrium (RNNE) bid function $\gamma(x_i)$ is given by⁴

$$\gamma(x_i) = x_i - 15 + h(x_i) \text{ where} \quad (1)$$

$$h(x_i) = (30/n + 1) \exp\{-(n/30)(x_i - 65)\}, \quad (2)$$

and n is the number of active bidders in the auction. The term $h(x_i)$ quickly becomes negligible as x_i increases beyond 65. Thus for simplicity we ignore it in the discussion that follows, although we do include it in all relevant regressions.

In common-value auctions bidders usually win the item when they have the highest (or one of the highest) signals. Let $E[x_i|X=x_{1n}]$ be the expected value of the signal conditional on it being the highest among the n signal values drawn. For signals in region 2

$$E[x_i|X=x_{1n}] = x_0 + [(n-1)/(n+1)]15. \quad (3)$$

Equation (3) indicates that if individuals naively bid their signal, they will overbid and can expect to lose money. The failure of bidders to sufficiently discount their bids relative to their signals in order to avoid losing money is called the “winner’s curse”. To avoid losing money, i.e. break even in an expected value sense, bidders should use the bid function

² Given this informational structure, private signals are affiliated in the sense of Milgrom and Weber, 1982.

³ Within region 2, bidders have no end point information to help in calculating the expected value of the item, which simplifies the bid function.

⁴ Derivation of the RNNE bid function, as well as its characterization outside of interval 2 can be found in Kagel and Levin (1986) and Kagel and Richard (2001).

$$\tilde{\gamma}(x_i) = x_i - [(n-1)/(n+1)]15. \quad (4)$$

Since bids above (4) will incur negative expected profits in auctions in which the high signal holder always wins the item, the extent to which individuals bid above (4) provides a convenient measure of the extent to which bidders suffer from the winner's curse.

We define the bid factor as the amount an individual reduces his bid below his signal. Equation (4) gives the bid factor one would employ to just break even in an expected value sense, i.e. just to correct for the adverse selection effect from having the highest signal. On the other hand, equation (1) shows that the bid factor for the RNNE is approximately \$15. The bid factor needed to just break even is quite large relative to that for the RNNE: with $n = 6$, equation (4) implies that the bid factor required to generate zero expected profits is \$10.71, or approximately 71% of the bid factor of \$15 implied by equation (1). Of course bidders will want to do better than just break even, and the RNNE bid factor of \$15 implies positive expected profits for bidders.

The RNNE is based on the assumption of risk neutral bidders, all of whom employ the same bid function and fully account for the adverse selection effect conditional on winning the item. However, as will be shown in the empirical analysis, the homogeneity assumption is not tenable as demographic characteristics and “ability” impact on bidding, and there is some residual, unexplained, heterogeneity in bidding, as evidenced by a statistically significant subject effect error term in the regressions. All of these deviations from the theory raise questions regarding the empirical relevance of the RNNE bid function (1), and of the breakeven bid function (4). They are, however, still relevant benchmarks for the following reasons. First, virtually all subjects are bidding above, rather than below, $x_i - 15$ in region 2, and the best response to such rivals is to bid $(x_i - 15)$ (Kagel and Richard, 2001). Second, as Kagel and Richard (2001) show, within region 2, regardless of what other explanatory variables are included in the empirically specified bid function, the coefficient value for own signal value, x_i , is indistinguishable from 1.0 (as it is here) and heterogeneity between bidders is almost exclusively confined to the intercept of the bid function. Third, although the impact of risk aversion on bids in first-price private-value auctions is unambiguous (it is to bid above the RNNE), it does not necessarily have the same impact in common-value auctions since bidding above $x_i - 15$ creates the possibility of losses.⁵ What we can

⁵ Thus, any impact of risk aversion should be less here than in private-value auctions. There is some controversy regarding the role of risk aversion in private-value auctions. For some of the latest results regarding the role of risk aversion in bidding above the RNNE in first-price, private-value auctions see Issac and James (2000) and Dorsey and Razzolini (2003).

say about risk aversion is that (i) a risk averse bidder clearly does not want to bid above (4) since to do so yields negative expected profits and (ii) from a strict empirical perspective any deviations from risk neutrality will impact on the size of the bid factor (the intercept term in the empirically estimated bid functions, 6a and b, reported on below).⁶ Finally, bidding above (4) yields negative expected profits both with strict homogeneity in bidding and in cases where all of one's rivals are bidding above (4), regardless of the heterogeneity in bid patterns. As such, for inexperienced bidders at least, (4) still provides a reasonable measure for whether individual bidders have fallen prey to the winner's curse.

3. Experimental Procedures and Basic Descriptive Statistics

Each experimental session consisted of a series of auctions in which a single unit of a commodity was awarded to the high bidder in a first-price sealed-bid auction. The value of the item, x_0 , was unknown at the time bids were submitted, with new values for x_0 and new signal values (x) drawn randomly in each auction. All of the information about the underlying distribution of x_0 and signal values was included in the instructions, which were read aloud to all subjects (each of which had a written copy to read).⁷ An admissible bid was any real number between zero and $x + \$17$. The latter is \$2 greater than any possible value of x_0 , with the restriction intended to prevent bankruptcies resulting from typing errors. A reservation price equal to $x_0 - \$30$ was in effect at all times, with the reservation price rule (but not its realizations) announced. Winning bids always exceeded the reservation price. At the end of each auction all bids were posted from highest to lowest along with the corresponding signal values (bidder identification numbers were suppressed) and the value of x_0 . Profits (or losses) were calculated for the high bidder and reported to all bidders as well.

Each experimental session began with two markets with six bidders each. Assignments to each market varied randomly between auction periods. To hold the number of bidders, n , constant in the face of potential bankruptcies, extra bidders were recruited for each session. Bidders were randomly rotated in and out of active bidding between auctions.⁸ In sessions where the total number of bidders fell below 12, the number of bidders in each market was reduced

⁶ This is because within region 2 the coefficient value on own signal is indistinguishable from 1. Of course, this is far from the only factor possibly influencing the size of the bid factor.

⁷ Instructions may be found at the following URL site: <http://www.krannert.purdue.edu/centers/vseel/research.asn>

⁸ Inactive bidders got to see the outcomes for one of the two markets. In the first seven auction sessions (all with inexperienced subjects) extra bidders only became active following a bankruptcy. Prior to this they were seated next to an active bidder, observing auction outcomes but with discussions between active and inactive bidders prohibited.

proportionately; e.g., with 11 bidders there would be 5 in one market and 6 in the other, with 10 bidders there would be 5 bidders in both markets, and so on.⁹ The number of active bidders in each market was always posted at the top of bidders' computer screens, with a bidder's status (active or inactive) clearly indicated as well.

We employed three different treatments designed to help identify selection effects:

Control treatment: This treatment was designed to match standard experimental practice for common-value auctions, with all subjects given starting capital balances of \$10 and a flat show-up fee of \$5. All subjects participating in week 1 were invited back for week 2 where all subjects were again given starting capital balances of \$10 and a flat show-up fee of \$5.

Bonus treatment: In this treatment starting cash balances were either \$10 or \$15, with half the subjects randomly assigned to each cash balance level. Further, following each auction, active bidders were automatically enrolled in a lottery with a 50% chance of \$0 and a 50% chance of \$0.50. Subjects were informed of the lottery structure, and were told that the lottery was designed as a source of extra earnings whose outcomes were completely unrelated to the auctions. Outcomes for the lottery were posted at the bottom of bidders' screens *after* reporting the auction results. In addition, a show-up fee of \$20 was paid only *after* completing participation in week 2's session (which was scheduled for the same time and day for the next week), and 50% of all earnings from week 1 were held in escrow, only to be paid after completion of week 2's session.¹⁰

Random treatment: The random treatment was the same as the bonus treatment with the exception that (i) bidders were given a \$5 show-up fee in week 1 along with all of week 1's earnings and (ii) when inviting bidders back for week 2, half the subjects (determined randomly) were assigned a show-up fee of \$5, with the other half assigned a show-up fee of \$15. Thus, the only difference between the random and bonus treatments was the incentive for returning for week 2.

In recruiting, all subjects were informed that they were needed for two sessions to be conducted at the same time in two consecutive weeks, and to only register if they could meet that commitment. Only after registering were subjects in the bonus treatment informed that they would receive a \$20 show-up fee conditional on attending both sessions and that they would lose half their earnings if they did not participate in a second session. They were then permitted to withdraw from

⁹ In the first seven auction sessions one market always had 6 bidders, with any required reduction in the number of bidders confined to the second market. Assignments to markets continued to vary randomly between auctions.

¹⁰ Armantier (2004) employs these procedures to induce subjects to return between sessions in a common value auction experiment aimed at determining the role of information feedback on learning.

participating, which no one did. There were also no noticeable differences in week 1 show-up rates between the bonus and the other two treatments.

Each inexperienced session (week 1) began with two dry runs, followed by thirty auctions played for cash. Earnings from the auctions were added to or subtracted from starting cash balances. Lottery earnings were added to cash balances as well. Once a bidder's cash balance was non-positive they were declared bankrupt and no longer permitted to bid. This was done since with a non-positive cash balance it is rational to bid *somewhat* above the risk neutral Nash equilibrium (RNNE) reference point (1), and there is no practical way of extracting money owed from bankrupt subjects.¹¹ Experienced subject sessions employed an abbreviated set of instructions, a single dry run, and thirty-six auctions played for cash (as the shorter instructions permitted more auctions).

Subjects were recruited by e-mail from the general student population at Ohio State University. Just under 93% were undergraduate students, with the remainder either graduate students or of unknown status. The consent form gave us permission to collect demographic information. Week 1 sessions lasted approximately 2 hours, with week 2 sessions being shorter as only a summary of the instructions were read and subjects were familiar with the procedures.

Table 1 shows the number of sessions and subjects (listed in parentheses) in each treatment in weeks 1 and 2. Note that because of large numbers of subjects choosing not to return in week 2 for the control and random treatments, in conjunction with the need to recruit enough subjects in week 2 to be assured of being able to conduct two auction markets simultaneously with $n = 6$, we had to cancel a total of three week 2 sessions scheduled for these two treatments. These subjects were invited back on different days or at different times in week 2.¹² Table 1 also reports some basic descriptive statistics: average profit per auction actually realized, average profit per auction predicted under the RNNE, percent of auctions won by the high signal holder (strict symmetry predicts that this will be 100%), the frequency with which there is a clear winner's curse for all bidders, and for the high bidder, (bidding above the breakeven bid $\tilde{y}(x_i)$ from (4)), the percentage of subjects going bankrupt, the percentage of week 1 subjects returning for week 2, and the percentage of subjects going bankrupt in week 1 returning for week 2.

¹¹ Because of limited-liability for losses, once cash balances drop below a certain level it is rational, in terms of a one-shot game, to bid modestly above the RNNE assuming that all other bidders employ the RNNE (see, Kagel and Richard, 2001, for details on the extent of bidding above the RNNE as a function of cash balances held). This overbidding increases monotonically the lower the cash balance. Thus, it is somewhat arbitrary to use a non-positive cash balance as the cut-off point. However, permitting subjects to accumulate large negative cash balances that could not be collected might create negative externalities for the remaining bidders, and would adversely effect the lab's reputation.

¹² The experiment actually took place over a four-week period with all inexperienced subjects invited back for the week following their initial experimental session.

Three key factors stand out from the descriptive statistics reported in Table 1. First, there is a strong improvement in performance going from week 1 to week 2, with average profits per auction going from -\$2.47 (large average losses in week 1) to \$0.75 (small average profits in week 2), while average predicted profits remain unchanged between weeks. This conclusion is confirmed and reinforced by the winner's curse measure ($b > \tilde{\gamma}(x_i)$), as the frequency of the winner's curse drops dramatically for all bidders, and for high bidders as well. Second, there is severe potential selection bias within week 1 across all treatments, given the large number of bankruptcies reported within week 1. Third, there is relatively strong selection bias between weeks 1 and 2 for the control group and the random group, with 40.0% and 25.9% of week 1 subjects *not* returning for week 2 for these two groups, respectively, versus 4% for the bonus group.

4. Demographic and Ability Measures

The consent form subjects signed gave us permission to collect demographic data from the University Enrollment Office. This requirement was noted in the recruiting process, and had no obvious adverse effect on recruiting. The consent form called for providing information regarding gender, SAT and ACT scores, major, and class standing (freshman, sophomore, etc.). This information was provided in machine-readable form by the University and was matched to the experimental data through social security numbers.¹³

The demographic and ability measures employed in the data analysis are as follows:

Gender: male/female.

College major: Three categories were established - business and economics, science and engineering, and a residual category that included mostly arts and humanities.

SAT/ACT scores: These scores result from standardized tests that most high school graduates take when seeking admission to a US college. A set of binary variables was constructed for three ability levels - high, median, and low - based on both SAT and ACT scores. The cut-off points chosen were below the median, between the median and the 95-percentile, and the 95-percentile or higher, where the scores are calibrated relative to the national distribution of scores for the years 2002-2003.¹⁴ Binary variables were generated for several reasons. First, ACT and SAT scores are not

¹³ Information was unavailable for eight subjects (3.2% of the sample) either because we could not read the social security numbers from the consent forms or there was no corresponding record at the University office.

¹⁴ The cut-off points employed are the following. Type of score (at or below median, at or above 95-centile): SAT Verbal (500,700), Mathematics (510, 700), Combination (505, 700); ACT English (20, 30), Reading (21, 32), Mathematics (19, 30), Science and Reasoning (20, 29), Combination (20, 30). For ACT the reference points are the

additive measures of ability but rather rank order measures so that binary scores are more appropriate to use than raw scores. Second, this specification is a simple way to capture possible non-linear impacts of different ability levels. Third, a number of students were missing SAT or ACT scores (39.4% SAT and 25.5% ACT), which would have increased the fraction of observations with missing values. Using these binary variables permits us to get around this fact by counting someone in the top 5% if she was in the top 5% of the national distribution for *either* SAT or ACT, below the median if they were below the median in *either* test, and in the middle group otherwise.¹⁵ Using both scores this way reduces the number of subjects with a missing value to 13.7% of the sample.¹⁶ Bidders were coded according to their verbal, mathematical, and combined skills. This is possible because the ACT test has separate sub-scores for Reading, English, Mathematics, and Science and Reasoning abilities, while the SAT test has separate sub-scores for Verbal and Mathematics abilities. The verbal ability measure is derived from ACT Reading (ACT English deals with grammar questions) and SAT Verbal, while the combined ability is based on ACT and SAT Composite test scores.¹⁷ Overall, there are twelve ability variables, corresponding to the three ability levels plus a missing score status times the three different ability measures (composite, math, verbal). Although the categories used for SAT/ACT scores are somewhat arbitrary, they do provide reasonable measures of high, medium and low ability that we believe are interesting, and the results are robust to small differences in the cut points.

We also explored a number of empirical specifications using grade point average (GPA) in place of, or in conjunction with, SAT/ACT scores to measure ability (here too using a binary specification). GPA proved to be a far inferior ability measure compared to SAT/ACT scores, rarely achieving statistical significance in any of our specifications. We suspect there are two primary reasons for this. First, we have a number of freshmen and sophomores in our sample for which GPA would be a very incomplete measure of academic performance. Second, there is likely to be a good deal of heterogeneity in grade scales within our specified college majors, no less between the different majors. The two together make GPA a much fuzzier measure of ability than SAT/ACT scores.

scores reported during years 2002-2003. For updates on the distribution see <http://www.act.org/aap/scores/>. For SAT I test the reference are the scores of 2002 distribution. See www.collegeboard.com for more information.

¹⁵ Many students take the SAT or ACT test more than once. In these cases we used the latest score reported.

¹⁶ Most of the subjects with missing values were students who transferred to OSU after the first year from satellite campuses as these scores are not required for these students.

¹⁷ Correlations between SAT and ACT sub-scores are rather high, 0.85 for Mathematics, 0.76 for Verbal, and 0.86 for Composite.

Table 2 gives the sample composition according to these demographic and ability variables for inexperienced and experienced bidders. Column 1 shows the sample composition for inexperienced subjects, with column 2 showing the sample composition for returning, experienced bidders. Men comprise 58-60% of the sample, with the breakdown by major being approximately one-third economics/business majors, one-third engineering and science majors, and the remainder primarily humanities majors. Some 20-25% of the sample are in the top 5% (of the national average) with respect to composite SAT/ACT scores, with less than 10% scoring below the median, and around 12% not having any SAT/ACT scores. About 50% of the sample is freshmen and sophomores, with the remaining 50% being upper-division undergraduates. When demographic variables are included in the regressions, the reference bidder is a male subject with a humanities major that has a median ability level.

Finally, a number of observations were dropped from the dataset. There were three possible reasons for this. First, the signal was not in region 2, [65, 935] (3.0% of total observations). Second, problems with the software, such as crashes, that made some data points unreliable or unavailable (1.9% of total observations). Third, the bid was an outlier, defined as a bid of more than \$60 below the signal or more than \$30 above it (1.0%).

5. A Duration Model of Bankruptcy

As Table 1 shows there are a large number of bankruptcies for inexperienced bidders (week 1). The question to be addressed here is what are the factors behind these bankruptcies, since avoiding bankruptcy is one measure of bidding skill. In particular, are there any demographic or ability factors that we would otherwise be unaware of behind the bankruptcies? A natural format for modeling bankruptcies is to employ a hazard, or duration, model, since once a subject goes bankrupt they are no longer permitted to bid in the week 1 session.

We assume that the probability that a person goes bankrupt in period t , conditional on not having gone bankrupt in the previous $t-1$ periods, is given by the discrete-time (logit) hazard function

$$\lambda_i(t|\theta_i, Z_{it}; \delta, \gamma) = 1/(1 + \exp(-y_{it})) \quad (5a)$$

where

$$y_{it} = Z_{it}\delta + \sum_{r=k}^K \gamma_k \ln(t)^k + \theta_i \quad (5b)$$

In (5b) Z_{it} is a vector of explanatory variables and t measures the number of auctions the individual has actively participated in (including the current one). We are particularly interested in the coefficient vector δ since Z_{it} consists in part of demographic and ability variables. Our ability measure employs the composite SAT/ACT score, with dummy variables to distinguish between subjects who score in the top 95th percentile, below the median, and with no SAT/ACT scores.¹⁸ There are also dummy variables for women, engineering and science majors, and economics and business majors. In addition to the demographic and ability variables, Z_{it} consists of a number of control variables. These include the variable *lagcumcash*, defined as subjects' starting cash balance plus any lottery earnings they may have received, but excluding auction earnings. (The exclusion of auction earnings is necessary to avoid endogeneity problems.) We anticipate a negative coefficient value for this variable as bidders with larger exogenous sources of cash are less likely to go bankrupt. Additional control variables include whether a bidder had the highest, or second highest, signal in auction t , since the 'winner' almost always has the highest or second highest signal. Given the large negative average profits reported in Table 1, we anticipate that the coefficient values for these variables will be positive, since when subjects 'win' they lose money on average (the winner's curse) and therefore increase their probability of going bankrupt. We also include the fraction of past periods that a bidder has received the highest signal and the fraction of time they have received the second highest signal. These variables can have a positive or negative effect on the probability of going bankrupt as there are opposing forces at work. On the one hand, having the high signal in the past increases the likelihood that the subject has lost money from falling prey to the winner's curse. On the other hand, having fallen prey and surviving has probably taught this subject to bid more conservatively; or it may simply be that subjects who survived high signals in the past weren't prey to the problem in the first place. The polynomial in $\ln(t)$ reflects duration dependence, i.e. how the probability of going bankrupt changes with the number of periods the subject has participated in. There may be positive or negative duration dependence indicating that the probability of going bankrupt increases or decreases with t . We then use the Schwartz criterion to choose the order of the polynomial, and in all specifications this procedure yielded a first order polynomial in $\ln(t)$ for the duration dependence.

¹⁸ Results employing verbal and mathematics SAT/ACT scores in place of the composite score are reported in Tables B3a and b in the online Appendix B available at http://www.krannert.purdue.edu/centers/vseel/papers/CHK_new_append_June_2.pdf

Finally, the term θ_i is a random variable designed to account for unobserved heterogeneity. (Recall that Z_{it} controls for observed heterogeneity.) It takes on the value θ_1 with probability P_1 and the value θ_2 with probability $1 - P_1$. The terms θ_1 , θ_2 , and P_1 are parameters to be estimated. We include unobserved heterogeneity for two reasons. First, if we ignore it, we run the risk of biasing the duration dependence in a negative direction and biasing the absolute value of the δ coefficients towards zero. Second, we can potentially use the unobserved heterogeneity distribution in Ryu's (2001) correction for selection bias (resulting from bankruptcies) when analyzing bidding behavior in week 1. We estimate the model by maximum likelihood; see, for example, Ham and Rea (1987) for an explicit expression for the likelihood function.¹⁹

The results for no demographics and the case where we use demographics are reported in Table 3, with separate estimates reported with and without accounting for unobserved heterogeneity in both cases (columns labeled yes and no, respectively). The analogous results for the case of math scores or verbal scores are in Tables B3a and B3b respectively of the online Appendix. Focusing first on the demographics, women have a higher probability of going bankrupt in a given period, conditional on having survived to that period. Students who are in the 95th percentile and above have a significantly lower probability of going bankrupt – the reference group consists of those between the median and the 95 percentile. Also, students below the median have a significantly higher probability of going bankrupt than the reference group, while those with no aptitude score are not significantly different from the reference group. Further, choice of major does not significantly affect the probability of bankruptcy after having controlled for gender and aptitude scores. Thus ability, as measured by these standard aptitude tests, plays a significant role (in the anticipated direction) in avoiding, or limiting the impact of, the winner's curse. In contrast, the gender effect was totally *unanticipated* but, as will be shown, is one of the strongest results reported. We relate this gender result to the limited literature on gender effects in experimental economics in Section 8 below.

The amount of cash from the initial allocation and the lottery that a subject has on hand (*lagcumcash*) has a very significant negative effect on the probability of going bankrupt. While this is largely an accounting as opposed to a behavioral effect, it suggests that to limit selection problems in the bidding equation due to bankruptcies we will need to focus on the bonus and

¹⁹ A more detailed discussion of estimating duration dependence and unobserved heterogeneity parameters are contained in the online Appendix, section A.1, available at http://www.krannert.purdue.edu/centers/vseel/papers/CHK_new_append_June_2.pdf

random treatment subjects who got the larger (\$15) starting cash balance. Also, as expected, receiving the high signal or the second highest signal in the auction has a very significant positive effect on the probability of bankruptcy. On the other hand, the fraction of past auctions that a subject received the high signal has a negative effect on bankruptcy, and the fraction of previous periods that the subject received the second highest signal has a negative sign but is not statistically significant. This is consistent with substantial learning on the part of subjects: Those that get hit with losses in a given auction but survive to later periods learn to bid less aggressively. The log duration variable is negative in sign and statistically significant indicating that the longer a subject is able to bid, the less likely he or she is to go bankrupt. This, in conjunction with the negative coefficient values for the fraction of past auctions with the highest or second highest signal value, indicates that subjects have learned to bid more conservatively from their own past losses as well as from others having fallen prey to the winner's curse.²⁰

When we move from the model with no unobserved heterogeneity ($\theta_1 = \theta_2 = \text{the constant}$) to the case where we have unobserved heterogeneity ($\theta_1 \neq \theta_2$), the value of θ_1 does differ somewhat from θ_2 . However, the value of the log-likelihood only increases by 0.4, which leads to accepting the null hypothesis of no unobserved heterogeneity. As always, care must be exercised when not rejecting the null hypothesis, since these results may also reflect our inability to identify unobserved heterogeneity in a sample of 251 individuals.²¹ This issue is of practical importance, since it can be argued that one can ignore selection bias if one finds no evidence of unobserved heterogeneity – for example Ryu's correction cannot be implemented. However, Bijwaard and Ridder (2005) demonstrate, in the context of a duration model, that there can still be selection bias even after finding no evidence of unobserved heterogeneity. Thus, when analyzing the week 1 bid function we consider an alternative approach to selection bias due to bankruptcy that exploits our sample design, and indeed find that such bias is important.

²⁰ The conclusions from this section are robust to employing only verbal or only math SAT/ACT scores as ability measures, with the exception that the variables for being above the 95% percentile and for log duration lose their significance at conventional levels, although they maintain the same signs.

²¹ Monte-Carlo evidence from Heckman and Singer (1984) suggests that it is much easier to avoid bias in the parameters of the hazard function by including unobserved heterogeneity than it is to actually recover the distribution function of the unobserved heterogeneity, so that it may be difficult to implement Ryu's correction.

6. Analyzing the Bidding Behavior of Inexperienced Subjects

6.1 Basic Specification of the Bidding Equation

Our goal in this section is to estimate a bidding equation for inexperienced bidders. The dependent variable is signal minus bid.²² In our basic specification the explanatory variables include the nonlinear bidding term $h(x)$ in equation (2) and a learning term equal to $1/\ln(1+t)$, where t measures the number of auctions in which a bidder was active. This specification for learning has the attractive feature that the term becomes smaller as t gets large. Our specification differs from previous work in that it also includes dummy variables for the various ability and demographic factors, all of which are coded exactly as in the duration model. Thus, the constant represents the bid factor for a male whose ACT or SAT score was between the median and the 95th percentile, and whose major was not in engineering, science, business or economics.

Our second departure from the standard bidding equation involves including cash balances as an endogenous explanatory variable. We do this for two reasons. First, Ham, Kagel and Lehrer (2005; hereafter HKL) found that it affected bidding behavior in private value auctions, and it is interesting to ask whether it also affects common value auctions. There are a number of avenues through which cash balances might affect bidding, including a strictly behavioral one in which subjects enter the auction with some target earnings in mind so that their cash balance impacts on the perceived value of trying to win the auction. Secondly, and more importantly, our failure to include it may give misleading results in terms of the selection bias analysis – this issue is best discussed in the next section.

Cash balances must be treated as endogenous. For the controls, cash balances only change if the subject ‘wins’ a previous auction – thus cash balances are entirely determined by lagged bidding behavior for this group. For the random and bonus treatments, cash balances will depend on past bidding, as well as on the (randomized) initial balances and the lottery earnings. The last two variables are obvious instrumental variables for bidders not in the control group. Further, following HKL we use the fraction of previous periods that the subject received the high signal and the fraction of previous periods that the subject received the second highest signal as instrumental variables. All of these instrumental variables are statistically significant in the first-stage equation.

Thus, for the case of no gender by learning interaction, we write the regression equation as

²² Alternatively we could use signal as a regressor with bid as the dependent variable. If we do this, we obtain a coefficient on signal of approximately 1.0 with a t-statistic over 1000. The other coefficients do not change.

$$x_{it} - bid_{it} = \beta_0 + \beta_1 F_i + \beta_2 App_i + \beta_3 Maj_i + \beta_4 C_{it} + \beta_5 (1/\ln(1+t)) + \beta_6 h(x_{it}) + \varepsilon_{it} \quad (6a)$$

where x_{it} denotes signal, F_i is a dummy variable equal to 1 if the subject is a woman, App_i denotes a vector of dummy variables based on the subjects' aptitude score, Maj_i is a vector of dummy variables indicating the subjects major, C_{it} denotes cash balances, the term $1/\ln(1+t)$ captures learning, $h(x_{it})$ is the nonlinear term in equation (2) and ε_{it} is an error term. The bid equation for the case of a gender by learning interaction is given by

$$x_{it} - bid_{it} = \beta_0 + \beta_1 F_i + \beta_2 App_i + \beta_3 Maj_i + \beta_4 C_{it} + \beta_5 (1 - F_i)(1/\ln(1+t)) + \beta_6 F_i(1/\ln(1+t)) + \beta_7 h(x_{it}) + \varepsilon_{it}. \quad (6b)$$

6.2 Addressing Selection Bias Due to Bankruptcy

In considering the behavior of inexperienced bidders, there is an obvious potential selection bias as bidding is only observed as long as a subject does not go bankrupt, and 37.4% of all bidders went bankrupt in week 1. Further, as the hazard function estimates of the previous section indicate, the bankruptcy rate decreases when the experimenter gives higher cash endowments to subjects. Our experimental treatments manipulated cash endowments through variations in the initial cash balance and in lottery earnings, producing three distinct groups:

- (i) Low cash endowment, with \$10 initial balance and no lottery; these are fairly standard conditions in common-value auction experiments and corresponds to our Control treatment which had a bankruptcy rate of 46.3%
- (ii) Medium cash endowment, with \$10 initial balance and lottery earnings in the Random and Bonus treatments, which had a bankruptcy rate of 42.3%.
- (iii) High cash endowment, with \$15 initial balance and lottery earnings in the Random and Bonus treatments, with a bankruptcy rate of 20.3%.

The time pattern of the bankruptcy rates is given in Figure 1. From this figure we see that we lose observations on a substantial number of subjects in the overall sample and in groups (i) and (ii) in particular. This attrition can lead to inconsistent estimates of the bidding function, so that we test for selection bias and adjust for it, if necessary.

6.2.1 Testing for Selection Bias

There are several ways of proceeding. One is to consider panel data maximum likelihood. Another is to estimate a two-step Heckman (1979) type model. Both alternatives tend to be

computationally and data intensive; e.g., Ridder (1990), and do not seem like sensible procedures given that we have only 251 individuals.²³

An alternative due to Ryu (2001) is based on the duration model of attrition estimated above in Section 5. Ryu provides a feasible and natural means of dealing with selection bias based on the estimates of the parameters of the duration model.²⁴ Unfortunately, his approach relies on obtaining relatively precise estimates of the parameters of the density for θ_i in the duration model (equations 5a and 5b). As already noted, this is a very hard estimation problem given our sample size, and we did not find any evidence of unobserved heterogeneity in the duration model in Section 5. One response would be to argue that since we cannot estimate the unobserved heterogeneity distribution, there really is no heterogeneity and thus no selection problem. However, we believe this finding simply reflects the difficulty of identifying unobserved heterogeneity using standard econometric techniques with our relatively small sample size. Thus, we turn to a simple alternative based on our experimental design. This alternative approach is based on the following idea: Suppose one's sample consists of two sub-samples (created randomly at the beginning of the auction). There is no attrition in the first sub-sample but there is substantial attrition in the second sub-sample. Under the null hypothesis of no selection bias, the estimates of equations (6a) (or 6b) from the two sub-samples should not differ. However, if there is selection bias, the estimates should differ, since the estimates from the first sample are consistent, but the estimates from the second sub-sample are inconsistent (because of selection bias). Alternatively, assume that both samples experience attrition, but that it is much more serious in the second sample. Then using the arguments of White (1982), we would expect the estimates not to differ if there is no selection bias, but to differ if there is selection bias, since the bias will be much greater in the second sample.²⁵

In what follows one of our sub-samples, which we will refer to as the *high bankruptcy group*, consists of the control treatment (group (i) above) and the medium cash endowment group (group (ii) above). We combine these two groups as (a) they have very similar (high) bankruptcy rates and (b) they yield essentially the same results as the control group alone, but with substantially more statistical power in testing for selection bias (given that combining the two groups essentially

²³ We also considered a simpler Heckman type model where we looked only at those who never went bankrupt and estimated a simple probit equation for whether a subject went bankrupt at all during week 1. We could find no evidence of selection bias using this approach. (See section 7 and the online Appendix A.3, for a description of the Heckman approach.)

²⁴ See Appendix A.2 where we present his procedure for a discrete time hazard, rather than the continuous time hazard he considered.

²⁵ Our approach is similar to that taken by Verbeek and Nijman (1992) in panel data.

doubles the sample size).²⁶ We will compare the coefficients from this sub-sample with what we will refer to as the *low bankruptcy group*, which consists of group (iii) above (the high cash endowment group), since they have a bankruptcy rate of 20.3%. The latter is just at the margin of where empirical researchers would worry about selection bias.²⁷

We test for selection bias by testing whether the low bankruptcy and high bankruptcy groups produce the same coefficients. We use random effects instrumental variable estimation to account for (i) the correlations across periods for a given subject and (ii) the potential endogeneity of cash balances.²⁸ The last row of columns 1 and 2 of Table 4 contain the p-value for the null hypothesis that the high bankruptcy group and the low bankruptcy group produce the same coefficients *when we do not control for ability and demographics*. We reject this null hypothesis at the .01 level. The last row of columns 3 and 4 show the p-value for the null hypothesis that the two samples produce the same results when we add ability and demographic measures. Again, we reject the null hypothesis of no selection bias at the .01 level. Finally, the last row of columns 5 and 6 provides the p-value for the case where we interact gender with learning (this will be an important specification in what follows.) Again we reject the null hypothesis that the coefficients are equal across the two sub-samples at any reasonable confidence level. The results are the same if we substitute the verbal or mathematical score for the composite score.²⁹

These results provide strong evidence of selection bias arising from attrition due to bankruptcy, whether or not we condition on the demographics. The implication is that we failed to find any evidence of selection bias using Ryu's (2001) approach due to a lack of power, not because this bias did not exist. Further, our results suggest that experimenters are more likely to avoid selection problems by changing the design of the experiment (e.g. giving high initial balances) than to be able to deal with selection problems using econometric techniques that require large samples. Of course, one could argue that the results from the high bankruptcy sub-sample differ from the low bankruptcy group results because the groups have different cash balances in each period, and that these cash balances affect bidding. However, because we have conditioned on cash balances, this problem is taken into account in the estimation. Since we find selection bias both with and without the inclusion of demographic variables, and since selection bias has not been dealt with prior to this,

²⁶ See our working paper (Casari, Ham and Kagel, 2004) for results for the control group by itself.

²⁷ Note that the bankruptcy numbers somewhat overstate the actual level of attrition, since those who go bankrupt contribute observations up until the point that they go bankrupt – see Figure 1.

²⁸ The instruments are those discussed in Section 6.1.

²⁹ See online Appendix Tables B4a and B4b for the results for verbal and math scores respectively.

our results suggest biases in estimating learning effects and bid factors in virtually all earlier common-value auction experiments. Finally, note that while selection bias due to bankruptcies is a particularly acute problem in common-value auctions, it has the potential to impact in a number of other experimental environments as well, whenever subjects are exposed to potential losses.³⁰

6.2.2 Bidding Behavior of Inexperienced Subjects Absent Selection Bias

Given the above evidence of selection bias, we focus on the results from the low bankruptcy (high initial balance) sub-sample to represent the true bidding equation. Implicitly we are assuming that a bankruptcy rate of 20% will not create selection bias. However, one may interpret these results as being biased, but less biased than one would obtain using the standard experimental design and practice that ignores selection bias entirely.

Considering the results in column 3 of Table 4 for the high initial balance sub-sample, there is substantial and statistically significant learning going on.³¹ Further, women are bidding significantly higher than men by a little over \$3, as are those with no aptitude score (by over \$5) and those who are an economics or business major (by almost \$3). Although all individuals are overbidding relative to (4), individuals in these groups are doing a significantly poorer job of bidding. Transfers to the main Ohio State campus from satellite campuses dominate the no aptitude score group, and these transfers would, in general, not be as strong academically as students who start at the main campus.³² Thus, this provides some evidence (in addition to that from the duration model in Section 5) that those with lower ability as measured by SAT/ACT score do a poorer job of bidding. There are several possible explanations for the result that those with an economics or business major do a poorer job of bidding. One possibility is that these students are by nature aggressive in business-type transactions so that they fall into an even worse winner's curse trap. Although this aggressiveness might help in some situations, it does not help here.

To shed further light on the differences between men and women, consider the results reported in column 5 of Table 4 where we interact the learning term with gender. The no-aptitude score and economics/business-major coefficients are still significant and essentially the same as when we do not interact gender and learning. However, interacting learning with gender produces

³⁰ For example, in asset market studies subjects can typically lose money and are typically barred from further trading once they have negative cumulative earnings (Bossaerts and Plott, 2004).

³¹ Given that the learning variable is $1/\ln(1+t)$, a negative coefficient value indicates that the bid factor is increasing over time.

³² Transfers from the satellite campuses are not required to have SAT or ACT scores, and only require a C average or better to transfer to the main campus.

the very interesting result that women are learning at a much faster rate than men. These differences in the bid factors over time are reported in Table 5 for this specification. We see that the bid factors for women are much below those of men in period 1. For business/economics majors, the female bid factor is only 11.3 per cent of the bid factor for male business/economics majors. However, by period 30 the female bid factors are 73.9 per cent of the male bid factor for this group. For those with other majors, the female bid factor is only 34.4 per cent of the male one in period 1, but by period 30 rises to 79.6 per cent of the male bid factor. Thus, women start out bidding much more poorly than men but close much of the gap by period 30. This phenomenon also occurs when we do not control for other demographics. See Figure 2, which presents the actual average bid factors for men and women by period, without controlling for major or ability score. There are a number of possible explanations for this gender effect which we will discuss in Section 8 below, where we relate our results to other reports of gender effects in the experimental literature.

It is interesting to compare the bid factors in Table 5 with the RNNE bid factor of \$15 from (1) and the breakeven bid factor from equation (4) of \$10.71 (for $n=6$). Quite clearly the bid factor for women and economics/business majors is well below the minimum required to avoid the winner's curse in period 1. And even the largest (mean) bid factor in period 1 is not significantly larger than the breakeven bid factor. Further, the mean bid factor for women and economics/business majors is still below the breakeven bid factor in period 30, but the bid factor for best group (male, other majors) is now large enough to avoid the winner's curse.

The coefficients on cash balances are positive and statistically significant in all of our specifications, meaning that those with larger cash balances have a larger bid factor (bid less aggressively).³³ Calculating the effect of this, other things equal, we find that those with \$15 starting cash balances bid \$0.53 less than those with \$10 starting balances in period 1. Further, someone having earned \$30 (near the maximum of cash balances across subjects) would be bidding \$2.13 less than when they have a \$10 cash balance.³⁴ We surmise from this that although statistically significant, the cash balance effect is by itself not very important economically in terms of its direct impact on the bid factor. Where cash balances seem more important is in keeping subjects in the auction, giving them the opportunity to learn from their mistakes (see the discussion immediately below).

³³ The cash balance coefficient is significant at only the 10% level when we interact gender with learning.

³⁴ These calculations are based on the estimates in column 5 of Table 4.

It is interesting to note the differences between the coefficients for the low bankruptcy sample (columns 3 and 5 of Table 4) and those for the high bankruptcy sample (columns 4 and 6 of Table 4), since the bankruptcy rates for the latter are typical of past experimental results. First, considering the results in columns 3 and 4, learning is much slower in the high bankruptcy sample. This, at first blush, seems counterintuitive, since one would expect more of the less able subjects to go bankrupt in the low initial balance sample as they have less cash reserves to keep them in the game, and that this would show up in terms of a larger learning coefficient as the less able bidders went bankrupt. However, given heterogeneity in initial bid factors, it is precisely those subjects who have the most to learn that are being eliminated due to bankruptcy in the high bankruptcy group. That is, those subjects who would learn the most are being eliminated in the high bankruptcy group due to their lower starting cash balances, so that they do not stick around long enough to learn. This suggests that learning can serve as a substitute for initial abilities in terms of successful bidding in common value auctions, provided subjects have sufficient opportunity to learn. Second, we are able to identify significant demographic and cash balance effects for the high bankruptcy group in column (4) of Table 4. Specifically, the female dummy variable is significantly negative and its coefficient is two-thirds the value of that for the low-bankruptcy group, the below the median SAT/ACT variable is negative and statistically significant, and the cash balance variable is statistically significant and the same sign (and value) as for the low-bankruptcy group. Further, when we interact gender with learning in column 6, we see that there is a significant learning effect for women only, although this coefficient is less than a third as large as the term for the low-bankruptcy group in column 5.³⁵

Overall, the inclusion of the demographic and ability variables in the bid function does little to reduce the standard error of the bid function. This can be seen directly in Table 4 comparing the standard errors of the coefficients for the variables included in the no demographics bid function with the same variables in the bid functions with demographics. Thus, we conclude that including demographic and ability measures in the bid function, while providing a number of interesting insights, does not serve as a substitute for larger sample size in terms of the precision with which the bid function can be estimated.

Our treatment of providing relatively large starting capital balances and lottery earnings provides one device for controlling/minimizing selection effects in the initial inexperienced subject

³⁵ Limiting the high-bankruptcy group to the control group we are unable to identify any significant demographic or cash balance effects (Casari, Ham and Kagel, 2004). No doubt the expanded sample size from combining the control group with the low cash balance random group bidders enables us to identify these effects.

session. Two alternative devices that have been offered in the literature for minimizing (or even eliminating) bankruptcies have been to (i) employ sellers' markets (Lind and Plott, 1991) and (ii) create "deep pockets" for bidders (Cox, Dinkin, and Swarthout, 2001). In the sellers' market subjects are endowed with a single unit of unknown value with the option to keep it and collect its value, or to sell it. In this setup, all sellers earn positive profits, including the winner of the auction, but the winner's curse can still express itself as the opportunity cost of selling the item for less than its true value. While this procedure clearly eliminates bankruptcies, to keep costs down valuations and bids are all in terms of experimental dollars, with a relatively low conversion rate into US dollars. This in turn is likely to reduce the marginal incentives for equilibrium behavior; i.e., reduces the sting when bidders succumb to the winner's curse. In the deep pockets treatment of Cox et al. subjects were given sufficiently large starting cash balances so that they could not go bankrupt in any given auction even when bidding well above their signal value, and these cash balances were replenished following each auction. However, in order not to make the experiment prohibitively expensive, subjects were paid off on only 3 out of the 30 auctions conducted, selected at random at the end of each session. (Bidders were given feedback regarding whether or not they won the auction and their potential profits/losses following each auction.) In regressions comparing this treatment with otherwise identical treatments in which subjects were paid off following each auction, the deep pockets treatment produces a statistically, and economically, significant *increase* in the magnitude of the winner's curse. Thus, this treatment appears to limit learning/adjusting to the winner's curse, perhaps because the pain of potential losses does not arouse as much attention as that of immediate actual losses.³⁶

7. Analyzing the Bidding Behavior of Experienced Subjects

7.1 Addressing Selection Bias from Not Returning

7.1.1 Using Heckman's (1979) Approach

There is a clear potential for selection effects impacting on estimates of the experienced subject bid function since only 75.3% of all subjects return for week 2, with only 60.6% of bankrupt bidders returning. The percentage of returning subjects is substantially lower if we exclude the bonus group with its 96% return rate (recall Table 1). To obtain further insight on the potential for the attrition between week 1 and week 2 to cause selection bias in the estimates of the week 2 bid

³⁶ This result is consistent with Garvin and Kagel's (1994) study of feedback effects promoting larger bid factors for inexperienced bidders, which showed that the strongest effects came from their own losses and seeing others lose money as well; i.e., it takes real losses to capture bidders' full attention.

function, consider Figure 3. This figure provides a histogram of the bid factors from the first five auctions in week 1 for those bidders who returned in week 1 versus those who did not return. If there is no selection bias resulting from the attrition between week 1 and week 2, we would expect the distributions to be the same. However, this figure shows clear cut evidence of selection bias since those who return in week 2 clearly have much larger week 1 bid factors than those who do not return.³⁷

As is known from, e.g. Heckman (1979), ignoring this selection bias will bias the estimates of the intercept as well as the coefficients of the independent variables in the week 2 bid function if there is overlap (or correlation) between the variables in the bid function and the variables that affect the probability of returning. Estimates of a probit function for returning (see Table 6) show that both major and composite SAT/ACT scores significantly affect the probability of returning. Therefore, we need to account for selection bias in week 2 if we are to have unbiased estimates of the coefficients on the independent variables in the bid function, as well as the constant (the week 2 bid factor for the reference group). See Appendix A.3 for a discussion of these issues.

Dealing with selection bias is a difficult and controversial area of applied economics. Heckman (1979) provides one possible solution: put a bias correction term in the regression functions (6a) and (6b).³⁸ There are two major concerns with using this correction that are directly relevant to our case. First, the estimator is only unbiased asymptotically and was proposed for research problems with large data sets. Thus, our sample of 251 individuals may not be large enough for the asymptotics to come into play. Second, we need variables in the probit equation for returning bidders that do not affect bidding behavior. Variables included in the probit but not in the regression are dummy variables for: (i) the bonus group (coded 1 for bidders in the bonus group, 0 otherwise), (ii) the random group receiving the high, \$15, return show-up fee (coded a 1; 0 otherwise), (iii) the random group receiving the low, \$5, return show-up fee (coded as 1; 0 otherwise) and (iv) those subjects scheduled to return at the same time and day in week 2 as in week 1 (coded 1; 0 otherwise).

³⁷ For those readers familiar with the estimates of treatment effects in the training literature, this comparison of the initial week 1 bid factors in Figure 3 is analogous to comparing pre-training earnings of training participants to those of non-participants to see if there is non-random selection in those who undertake training. Note the bonus group is included in week 2 with its extraordinarily high return rate, so that selection bias is likely to be even higher in the standard experimental design.

³⁸ We discuss the Heckman approach to selection bias, along with Lee's (1982) generalization to the case of nonnormal error terms, in Appendix A.3.

As noted above, Table 6 reports the probit estimates for the probability of returning in week 2 that constitute the first step in the Heckman-Lee bias correction procedure. The probit estimates for the case where we do not use demographics are in column 1 of Table 6, while the probit estimates for the case where we include demographics are in column 2 of this table. Note that the bonus, high-return fee and same-day dummy variables are all positive and statistically significant in both columns. Hence, the selection model is well identified in the sense of having variables in the probit equation but not the bidding equation.³⁹ In terms of the demographic and aptitude variables included in the probit equation in Table 6, being an engineering or science major, and having a comprehensive SAT/ACT score in the 95th percentile or higher, leads to a higher probability of returning.⁴⁰

When we implemented the Heckman-Lee procedure, in no case are the selection terms close to statistical significance at standard confidence levels, independently of whether we control for demographics – see Table A1 in Appendix A. As with our test for selection effects in the week 1 bid function, there are two interpretations for this result. One interpretation is that there is no selection bias among the subjects returning as experienced bidders. The second, of course, is that our sample is too small for the Heckman estimator to be effective.⁴¹

7.1.2. An Alternative Test of Selection Bias

Our experimental design permits an alternative approach to correcting for selection bias in estimating the experienced subject bid function. Basically we break the sample into high return and low return sub-samples, and test whether the parameter estimates are equal across the sub-samples. Our *high return sub-sample* consists of the bonus group where 96% of all subjects returned in week 2. Our *low return group* consists of the control group, where only 60% of the subjects returned, combined with the low return fee subjects in the random group, who had a return rate of 69.1%. If there is no selection bias, we would expect that low return and high return sub-samples to produce the same coefficients; while if there is selection bias we would expect the two sub-samples to produce different coefficients.

³⁹ The positive coefficient for the low fee return group reflects the fact that they are more likely to return than the control group as earnings for these subjects were higher, on average, in week 1 than the controls due to higher starting cash balances and/or the lottery payments.

⁴⁰ This result remains if we substitute the verbal or math ACT/SAT scores for the composite scores (see Tables B6a and B6b in Appendix B).

⁴¹ Here too using the verbal or math aptitude score by itself gives the same results – again see Tables A2 and A3 in the online Appendix A.

Table 7 contains the results comparing the low return sub-sample and the high return sub-sample. Columns 1 and 2 report the results for the case where the bid function does not contain ability and demographic measures, while columns 3 and 4 present the results when the bid function contains ability and demographics and the learning term is not interacted with gender. Columns 5 and 6 extend the results to the case where learning is interacted with gender. Again we use the comprehensive aptitude score to measure ability. The bottom line of the table contains the p-value of the test statistic for the null hypothesis of no selection bias; i.e., the p-value for the null hypothesis that the estimates of the bid function are the same from the two sub-samples. We reject the null hypothesis of no selection bias at the 5% level, independently of whether we include demographics. Thus, as with the results for inexperienced bidders, we conclude that our sample is simply not large enough to identify selection effects using econometric selection models, but we find such effects when taking account of the diversity in return rates built into the experimental design.⁴² We also conclude that previous studies of common value auctions are likely to suffer from selection bias. Although this is a particularly acute problem in common value auctions, there is likely to be a potential problem in other experimental environments as well, as basic economic theory would suggest that more successful players (those earning higher average profits) are more likely to return for experienced subject sessions.

7.2 Estimates of the Experienced Subjects' Bidding Equation

Given this evidence of selection bias, we focus on the results for the high return sub-sample (the bonus group) for unbiased estimates of the bid function for experienced bidders in Table 7.⁴³ Consider first the estimates in column 1 for the case where we do not control for demographics. There is a learning effect for this sub-sample with the expected sign and a small cash balance effect. Further, the term $h(x)$ is statistically significant and has the correct sign, although it is still well below the predicted value of -1.0 . These effects are also present when we include demographic

⁴² This is true if we use the verbal or math ability score – see Tables B7a and B7b, although we can only reject the null at the 10% level when we use verbal scores. See our working paper (Casari, Ham and Kagel, 2004) for results in which we confine the low return sub-sample to the control group. For this case we find no selection bias, as the standard errors of the low return sub-sample are quite large relative to the coefficients, because of high autocorrelation in the random effects estimation for this sub-sample.

⁴³ The strategy of choosing a sample on the basis of an exogenous set of variables, for which the probability of participation is approximately one, to estimate the unconditional regression equation and avoid selection bias is known in the sample selection literature as ‘identification at infinity’ (see Chamberlain 1986). Note that we are not concerned with attrition within week 2 since bankruptcy is a much smaller problem in week 2 than in week 1 especially for those outside the control group, which constitute the source of our unbiased week 2 bid function estimates (recall Table 1).

variables in column 3. In terms of the demographic variables, the results from column 3 indicate that individuals with an aptitude score below the median do a significantly worse job of bidding in week 2⁴⁴. This is also true for economics and business majors, as was the case in week 1. Note that there is no longer any significant difference between men and women, although now the female coefficient is positive, in contrast to the week 1 results. However, when we interact learning with gender in column 5 we see that the learning term is only significant for women, and it has the expected sign; i.e., women are bidding closer to equilibrium/best responding over time.⁴⁵

In Table 8 we present the bid factors for men and women for different majors at different time periods based on the results from column 5. Women start out with a slightly lower bid factor than men but by period 30 their bid factor is approximately \$0.50 higher. From the week 1 results we saw that the gap between men and women was closing, with experience, as the experiment progressed. These results for experienced bidders suggest that the rest of the gap disappeared between week 1 and week 2, presumably as subjects fully absorbed the lessons from week 1's experience. Indeed, women may even be doing a slightly better job of bidding in week 2. From Table 8 we see that female business/economics majors have a bid factor that is below the breakeven bid factor of \$10.71 in periods 1 through 20 (and only slightly above it in period 30), while male business/economics majors are below this level for all periods. The non-economics/business majors have bid factors that are substantially above the breakeven bid factor, although these bid factors are considerably below the RNNE bid factor of \$15.00. Bidders with below median SAT/ACT scores have bid factors that are below the breakeven bid factor of \$10.71 for all periods regardless of major.⁴⁶

Table 7 also suggests that the results for the high return sub-sample are quite different than from the large low return sub-sample. Looking at the no demographics regressions first, the cash balance effect is much larger for the low return sub-sample than for the high return sub-sample. Further, the learning coefficient has a positive sign in the large low return sub-sample, indicating more aggressive bidding over time, which is opposite to best responding and opposite in sign to the high return sub-sample. These learning and cash balance effects for the low return sub-sample will tend to offset each other as cash balances are growing over time, while in the high return sub-

⁴⁴ Interestingly, this variable remains significant when we use the math score, but not when we use the verbal score (see Tables B7a and b in the online appendix).

⁴⁵ The gender coefficient is positive, and the coefficients in column 5 imply that the female bid factor is larger than the male bid factor for $\tau > 1$, assuming that their other characteristics are the same.

⁴⁶ These results are not shown in Table 8 to save space.

sample these effects reinforce each other. On the other hand, the intercepts of the bid functions are not significantly different between the two sub-samples.

Comparing specifications with the demographic and ability measures, the only demographic variable that is statistically significant in the low return sub-sample is the engineering/science major, with this group doing a better job of bidding. The insignificance of the other demographic variables in the low return sub-sample is not due to the standard errors increasing, since the standard errors in column 4 are comparable or smaller than those in column 3 for the high return sub-sample. Note that cash balances are still statistically significant for the low return sub-sample when we add demographics, and we continue to be unable to reject a null hypothesis that the intercept of the bid function is the same as for the high return sub-sample.

8. Discussion of Gender Effects

Perhaps the most surprising result reported here is that women start out bidding substantially more than men, suffering from a strong and severe winner's curse, but close the gap with men relatively quickly. In this section we briefly relate this result to what is known about gender differences that might explain this result. Two known factors that immediately come to mind as possibly responsible for the more severe winner's curse on the part of women are (i) women are generally identified as being more risk averse than men (see Eckel and Grossman, 2002 and 2003 for a survey)⁴⁷ and (ii) men tend to be over represented in the upper tail of mathematical reasoning (Geary, 1996, Benbow and Stanley, 1980, 1983). However, neither of these factors can account for the gender effect reported here. First, risk aversion cannot explain succumbing to the winner's curse as the latter is defined as a bidding strategy that insures *negative* expected profits conditional on winning the item. Second, our regression analysis explicitly controls for ability as measured by SAT/ACT composite scores. While the composite score summarizes in one index both mathematical and verbal abilities, these results are robust to including only SAT/ACT mathematical scores. In addition, our regressions include a variable for college major with the category science and engineering picking up subjects who would be most likely to have had more extensive courses in mathematics and deductive reasoning. Thus, even after controlling for these two factors, we identify a significant pure gender effect in our experiment.

⁴⁷ However, some studies show limited or no differences (Holt and Laury, 2002, Schubert, Gysler, Brown and Brachinger, 1999). See Croson and Gneezy (2004) for a review of the literature on gender differences as it relates to economics.

Two other possible explanations for this gender effect - overconfidence and aversion to competition - are not supported in the data either. Experimental evidence shows that although men and women both tend to be overconfident, men are generally significantly more overconfident than women (Deaux and Farris, 1977, Lundeberg, Fox, and Puncchohar, 1994). When facing a difficult task such as bidding in a common value auction, bidders with lower confidence might be expected to adopt the safe strategy of bidding their signal minus \$15, or bidding even lower than that. In contrast, more confident bidders might be expected to place more “competitive” bids. But our data clearly does not show such a differential pattern between men and women, but rather just the opposite of this.

Recent research indicates that women tend to shy away from, or under-perform, in competitive situations compared to men (Gneezy, Niederle, and Rustichini, 2003; Niederle and Vesterlund, 2005). One immediate implication of this is to suggest that women should bid more conservatively (lower) than men. But again we do not observe this for inexperienced women, but rather more aggressive bidding resulting in a more severe initial winner’s curse than men.

Absent a solid explanation for these gender differences in the literature, we conjecture that the overbidding may reflect a relative lack of experience with strategic interactions on the part of women compared to men, perhaps as a result of the above mentioned phenomenon of women shying away from competition more than men. It is in competitive situations that strategic interactions would come most into play, and relative lack of familiarity with such interactions might be sufficient to induce more aggressive bidding as a consequence of the failure to fully think through the implications of more aggressive bidding in this setting. We find this conjecture supported by the fact that women stop bidding more aggressively over time and do not differ from men as experienced bidders. By comparison, the business and economics students bid overly aggressively as both inexperienced and experienced bidders. Identifying similar gender effects in other experimental settings would provide support for this conjecture.

There have been a handful of other studies looking at gender and ability effects in auctions. Rutstrom (1998) looks for gender and racial effects in second-price and English private-value auctions with subjects bidding for a box of high quality chocolates. She finds no gender effect, but that non-whites bid significantly higher than whites. Since her auction is best modeled as an independent private value (IPV) auction in which bidders have a dominant strategy to bid their value, the differences could simply represent taste differences. Chen, Katuscak, and Ozdenoren (2005) check for gender effects in first- and second-price sealed-bid IPV auctions. They find that

women bid significantly more than men in the first-price auctions, but find no differences in the frequency of bidding their valuations (as the dominant bidding strategy requires) in the second-price auctions. They argue that the latter indicates no differences in intelligence between the men and women in their sample, while the fact that women are known to be more risk averse than men explains the overbidding in the first-price auctions. Ham and Kagel (2005) report gender differences in a two-stage (indicative) auction where first-stage bids are not binding but determine a short-list from which binding second-stage bids will be solicited in a sealed-bid private value auction.⁴⁸ They find that women are significantly more likely than men to get into stage-two with stage-one values that are so low that they result in earning negative average profits, and that women were significantly more likely to go bankrupt than men. Once again these results cannot be explained by risk aversion since they involve systematic overbidding that generates negative average earnings and bankruptcies.

In the papers cited in the previous paragraph there are no explicit controls for ability as measured by SAT/ACT scores, or grade point average, as in the present paper. Chen et al. (2005) collect information on numbers of courses taken by fields (e.g., science and engineering, economics and business, etc.).⁴⁹ Rutstrom (1998) uses self-reported information on income, age, and marital status but finds no significant effects from any of these variables. Ham and Kagel (2005) have no additional demographic or ability data on their subjects.

A closely related environment in which gender effects similar to those reported here have been identified is in the *Acquiring a Company* game (Charness and Levin, 2005). This game also involves a potential winner's curse should subjects ignore the adverse selection effect associated with bidding high enough to meet the owner's reservation price when attempting to acquire a company. Charness and Levin employ a series of questions testing for statistical sophistication in Bayesian updating in an effort to directly control for the relevant ability characteristics in checking for a gender effect in their task. They find a gender effect similar to the one reported here, with men less likely to commit a winner's curse than women, and that this gender effect is robust and relatively large in regressions that include the number of correct answers in their quiz for statistical sophistication. Thus, it appears the gender effect identified here with respect to the winner's curse is robust to the details of the game and the types of controls one could employ to account for possible omitted variable bias. Note that none of the above studies, except for Ham and Kagel (2005),

⁴⁸ See Kagel, Pevnitskaya and Ye (2004) for a thorough description of the two-stage (indicative) bidding procedure.

⁴⁹ Chen et al. (2005) incorporate other extrinsic factors that they believe might affect bidding in their analysis, discussion of which goes well beyond the scope of this section.

investigates differential learning on the part of women and men. They did not find evidence of the type of learning by gender interaction effects observed here, perhaps because the feedback promoting learning was not as clean cut as in the common value auctions.

9. Summary and Conclusions

We have investigated selection bias, demographic effects, and ability effects in common value auction experiments. We employed three different treatment effects and collected demographic and ability data on our subjects with the aid of the university enrollment office. As is typical of common value auction experiments, inexperienced bidders suffered from a strong winner's curse, earning negative average profits and suffering numerous bankruptcies. More experienced bidders did substantially better, earning positive average profits with relatively few bankruptcies. We obtained a number of substantive insights into this learning/adjustment process by which bidders start to overcome the winner's curse, as well as a number of methodological insights.

First, we find that some bidders are better than others: There are clear ability effects in the data as measured by SAT/ACT scores, as one might suspect. However, the pattern is different than one might have anticipated. The major impact of ability comes from bidders with below median composite test scores performing worse than those with above median test scores in terms of bidding more aggressively and going bankrupt, with the highest ability subjects (those scoring in the 95% or better on composite SAT/ACT scores) doing a bit better than the rest. Further, although our ability results are reasonably robust to whether we employ composite SAT/ACT scores or math or verbal scores alone, the composite scores achieve a more consistent, statistically significant impact in the regressions that control for ability effects than do either math or verbal scores alone. This suggests a general comprehension issue underlying better or worse performance in the auctions.

There are two interesting demographic effects. Most surprising, and least anticipated, is that women start out bidding substantially higher than men. However, they learn a lot faster so that by the end of the inexperienced subject sessions the differences between men and women have narrowed substantially. Further, women do as well as men as experienced bidders. This gender effect among inexperienced bidders cannot be attributed to risk aversion (as with a winner's curse bidders are earning negative average profits), or to ability (as we control for this with our SAT/ACT scores as well as data on undergraduate major). We have reviewed other potential causes for the gender effect identified and found them lacking as well. Our conjecture is that the overbidding may reflect a relative lack of experience with strategic interactions on the part of women compared to

men.

The second demographic effect identified is that economics and business majors tend to bid more aggressively than do other majors. The explanation for this effect that comes most immediately to mind is that economics and business students have a relatively aggressive mind set in commercial environments that gets in the way of maximizing profits in this environment. Note, we are not arguing that these students get extra utility from winning. Just that they are more aggressive in “business type” settings which may help in a number of situations but not here.⁵⁰ One possible implication of this effect is that the winner’s curse might be particularly hard to eliminate in a number of field settings where bidders with business and/or economics majors might be expected to be responsible for formulating bidding strategies.⁵¹

Aggregate adjustment toward the risk neutral Nash equilibrium is achieved in the different treatments through a variable mix of market selection and individual learning. When bidders have a low or medium cash endowment, they have less opportunity to learn and go bankrupt much faster. What the larger cash balances do is to provide bidders with protection from bankruptcy while permitting them to learn to avoid the winner’s curse. As such learning can serve as a partial substitute for ability in terms of overcoming the winner’s curse and as a substitute for market selection. This result provides some justification for the common practice of “qualified bid lists” (limiting who can bid on a project). The European Union adopts qualified bid lists to actively limit entry into government procurement auctions in an attempt to foster individual learning and minimize market selection (CEE Directive n.37, June 13, 1993). The common pattern here is to require bidders to have proven themselves in smaller projects before they can move on to large ones of the same type. Although this does limit competition in the short run, it has the benefit of insuring that bidders have the experience necessary to successfully complete the project. In the commercial construction industry, and a host of other environments, letting a contract only to have the high bidder default or delay completion because of financial distress causes enormous problems for the seller, ranging from badly needed but stalled completion of construction projects to tying up

⁵⁰ See Holt and Sherman (1994) for an experiment designed to distinguish between the joy of winning as opposed to the failure to appreciate the adverse selection effect conditional on winning in the closely related Acquiring a Company game. They clearly reject the joy of winning argument.

⁵¹ We are reminded of an anecdote told to us from the early spectrum auctions where an advisor to a major company urged his client not to bid on so many licenses as the net result would be to bid up the prices of all licenses, while greater profits could be earned by winning fewer licenses but at lower prices. In response to repeated advice along these lines, the CEO posted a bumper sticker on the advisors computer that read: “Winning isn’t everything, it’s the only thing” (a quote from the late, great football coach Vince Lombardi).

valuable assets in litigation, as bankrupt bidders are loath to give up what they have won in the auction.⁵²

We have also identified clear selection effects in the data, since without special inducements (i.e., the bonus treatment and the high return fee random treatment), less successful bidders are much less likely to return for experienced subject sessions. Although this is a particularly acute problem in common value auctions with less successful bidders going bankrupt, it might well be a problem in other experimental environments, as basic economic theory would suggest that more successful players (those earning higher average profits) are more likely to return for experienced subject sessions. The interesting methodological point here is that standard econometric techniques are of little value in identifying, no less measuring, these selection effects in our experiment in spite of the fact that by experimental standards we had a very large sample population (some 251 subjects). This probably has to do with the limited power of these techniques for such relatively small sample sizes by the standards of microeconomic work with field data. Rather, we are able to identify and measure selection effects through introducing treatment effects that yield the relevant contrasts directly. That is, the power to control the environment yields superior outcomes relative to more sophisticated statistical techniques. This is, in some sense, good news for experimenters as it means that they do not have to tool up on very advanced econometric techniques to do their analysis, or obtain prohibitively expensive large samples. They do, however, have to be aware of *potential* sources of estimation bias and take steps in designing their experiments (and in their data analysis) to account for these biases.

Finally, our experimental manipulations enable us to clearly identify the impact of cash balances on bidding. Inexperienced bidders with larger cash balances bid somewhat less aggressively than those with smaller balances, but the differences in cash balance are not nearly enough to account for bidding above the expected value conditional on winning, which is so prominent among inexperienced bidders. Cash balances have an even smaller role to play in experienced subjects' bid function. This too is good news for experimenters since it appears that real payoffs are more important than hypothetical payoffs (at least in this environment), but that the

⁵² The latter perils are vividly illustrated by the 1996 FCC auction of a block of spectrum for personal communications services reserved to small business ("C-block"). The bulk of that spectrum was either tied up in bankruptcy courts or returned to the FCC for re-auctioning, hence leaving a valuable asset unused for years because of the inability of auction winners to pay their bids.

changes in the subjects overall earnings during the course of the experiment have minimal impact on behavior.

References

- Armantier, O. 2004. "Does Observation Influence Learning," Games and Economic Behavior 46: 221-239.
- Benbow, C.P. and J.C. Stanley 1980. "Sex-Differences in Mathematical Ability Fact or Artifact," Science 210: 1263-1264.
- Benbow, C.P. and J.C. Stanley 1983. "Sex-Differences in Mathematical Reasoning Ability – More Facts," Science 222: 1029-1031.
- Bijwaard, Govert and Geert Ridder 2005. "Correcting for Selective Compliance in a Re-employment Bonus Experiment," Journal of Econometrics, 125: 77-111.
- Bossaerts, Peter and Charles Plott. 2004. "Basic Principles of Asset Pricing: Evidence from Large-Scale Experimental Financial Markets," Review of Finance 8: 125-169.
- Casari, Marco, John C. Ham, and John H. Kagel 2004. "Selection Bias, Demographic Effects and Ability Effects in Common Value Auction Experiments," mimeo, Ohio State University.
- Chamberlain, Gary 1986. "Asymptotic Efficiency in Semiparametric Models with Censoring," Journal of Econometrics 32:189-218.
- Charness, Gary and Dan Levin 2005. "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity and Affect," mimeo, Ohio State University.
- Chen, Yan, Peter Katuscak, and Emre Ozdenoren 2005. "Why Can't a Woman Be More Like a Man?" mimeo, University of Michigan.
- Cox, J. C., S. Dinkin, and J. T. Swarthout 2001. "Endogenous Entry and Exit in Common Value Auctions," Journal of Experimental Economics 4: 163-182.
- Croson, Rachel. and Uri Gneezy 2004. "Gender Differences in Preferences," mimeo, University of Chicago.
- Deaux, Kay and Eliazbeth Farris 1977. "Attributing Causes for One's Own Performance: The Effects of Sex, Norms, and Outcomes," Journal of Research in Personality 11: 59-72.
- Dorsey, R. and L. Razzolini 2003. "Explaining Overbidding in First Price Auctions Using Controlled Lotteries," Journal of Experimental Economics 6: 123-140.
- Eckel, Catherine C. and Philip J. Grossman 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk," Evolution and Human Behavior 23: 281-295.
- Eckel, Catherine C. and Philip J. Grossman 2003. "Men, Women and Risk Aversion: Experimental Evidence," forthcoming The Handbook of Experimental Economics Results.
- Garvin, Susan. and John H. Kagel 1994. "Learning in Common Value Auctions: Some Initial Observations," Journal of Economic Behavior and Organization 25: 351-372.
- Geary, D.C. 1996. "Sexual Selection and Sex Differences in Mathematical Abilities," Behavioral and Brain Sciences 19: 229.
- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini 2003. "Performance in Competitive Environments: Gender Differences," Quarterly Journal of Economics CXVIII: 1049-1074.

- Ham, John C. and Samuel Rea Jr. 1987. "Unemployment Insurance and Male Unemployment Duration in Canada," Journal of Labor Economics, 5: 325-53.
- Ham, John C., John H. Kagel, and Steven Lehrer 2005. "Randomization, Endogeneity, and Laboratory Experiments," Journal of Econometrics 125: 175-205.
- Ham, John C and Kagel, John H. 2005. "Gender Effects in Private Value Auctions," mimeo, Ohio State University
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error," Econometrica 47: 157-61.
- Heckman, James J. and Burton Singer 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Duration Data," Econometrica 52: 271-320.
- Holt, Charles A. and Susan K. Laury 2002. "Risk Aversion and Incentive Effects," American Economic Review 92: 1644-1655.
- Holt, Charles A. and Roger Sherman 1994. "The Loser's Curse and Bidder's Bias," American Economic Review 84: 642-52.
- Isaac, R. M. and D. James 2000. "Just Who Are You Calling Risk Averse?" Journal of Risk and Uncertainty 20: 177-87.
- Kagel, John H. and Dan Levin 1986. "The Winner's Curse and Public Information in Common Value Auctions," American Economic Review 76: 894-920.
- Kagel, John H. and Dan Levin 2002. "Bidding in Common Value Auctions: A Survey of Experimental Research," in John H. Kagel and Dan Levin, Common Value Auctions and the Winner's Curse, Princeton, Princeton University Press.
- Kagel, John H., Svetlana Pevnitskaya, and Lixen Ye 2004. "Indicative Bidding: An Experimental Analysis," mimeo, Ohio State University.
- Kagel, John H. and Jean-Francois Richard 2001. "Super-Experienced Bidders in First-Price Common Value Auctions: Rules of Thumb, Nash Equilibrium Bidding and the Winner's Curse," Review of Economics and Statistics 83: 408-419.
- Lee, Lung-Fei 1982. "Some Approaches to the Correction of Selectivity Bias," Review of Economic Studies 49: 355-72.
- Lind, B. and C.R. Plott 1991. "The Winner's Curse: Experiments with Buyers and with Sellers," American Economic Review 81: 335-46.
- Lundeberg, Mary A., Paul W. Fox, and Judith Puncchohar 1994. "Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgements," Journal of Educational Psychology 86: 114-121.
- Milgrom, Paul and R.J. Weber 1982. "A Theory of Auctions and Competitive Bidding," Econometrica 50: 1485-527.
- Niederle, Muriel and Lise Vesterlund 2005. "Do Women Shy Away from Competition," mimeo, Stanford University
- Ridder, Geert 1990. "Attrition in Multi-Wave Panel Data," in J. Hartog, G. Ridder and J. Theeuwes, Panel Data and Labor Market Studies, Amsterdam, North Holland.

- Rutstrom, E. E. 1998. "Home-Grown Values and Incentive Compatible Auction Design," International Journal of Game Theory 27: 427-441.
- Ryu, K. 2001. "A New Approach to the Attrition Problem in Longitudinal Studies," in C. Hsiao et al Nonlinear Statistical Modeling: Essays in Honor of Takeshi Amemiya, Cambridge, Cambridge University Press.
- Schubert, Renate, Matthias Gysler, Martin Brown and Hans-Wolfgang Brachinger 1999. "Financial Decision-Making: Are Women Really More Risk Averse?" American Economic Review Papers and Proceedings 89: 381-385.
- Verbeek, Marno and Theo Nijman 1992. "Testing for Selectivity in Panel Data Models," International Economic Review 33: 681-703.
- White, Halbert 1982. "Maximum Likelihood Estimation of Misspecified Models," Econometrica 50: 1-26.
- Wilson, Robert 1977. "A Bidding Model of Perfect Competition," Review of Economic Studies 44: 511-18.

Figure 1: Bankruptcy rate for inexperienced bidders

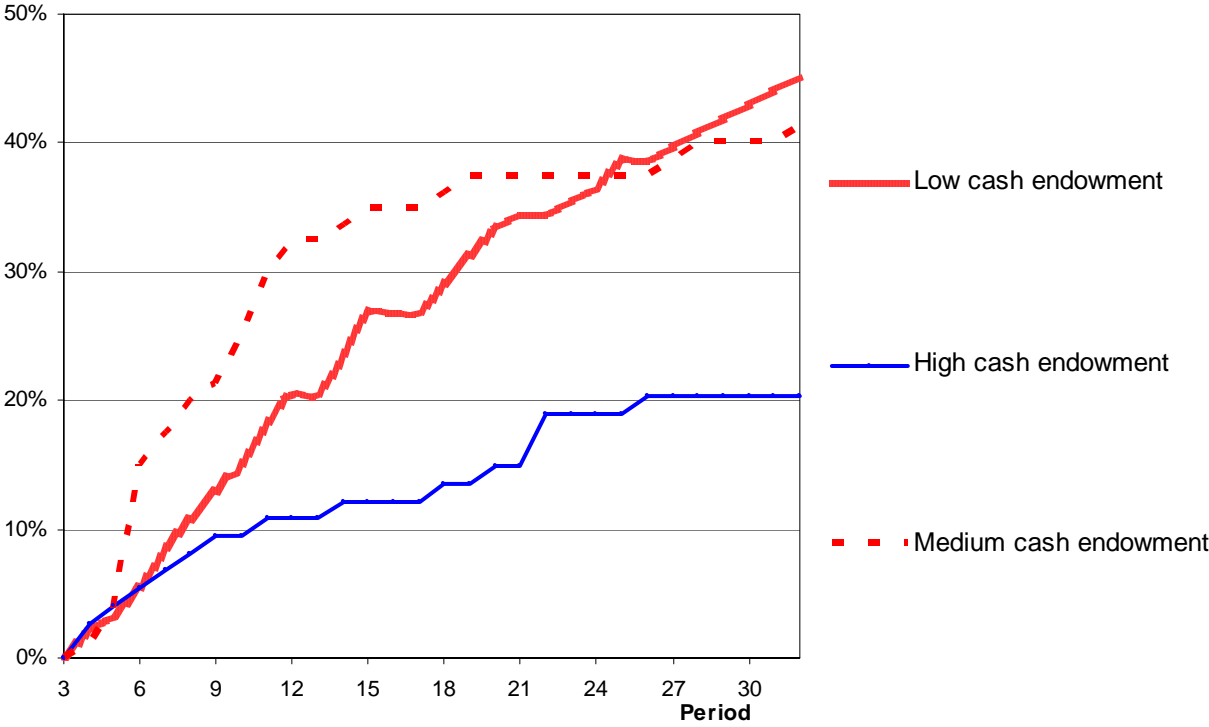
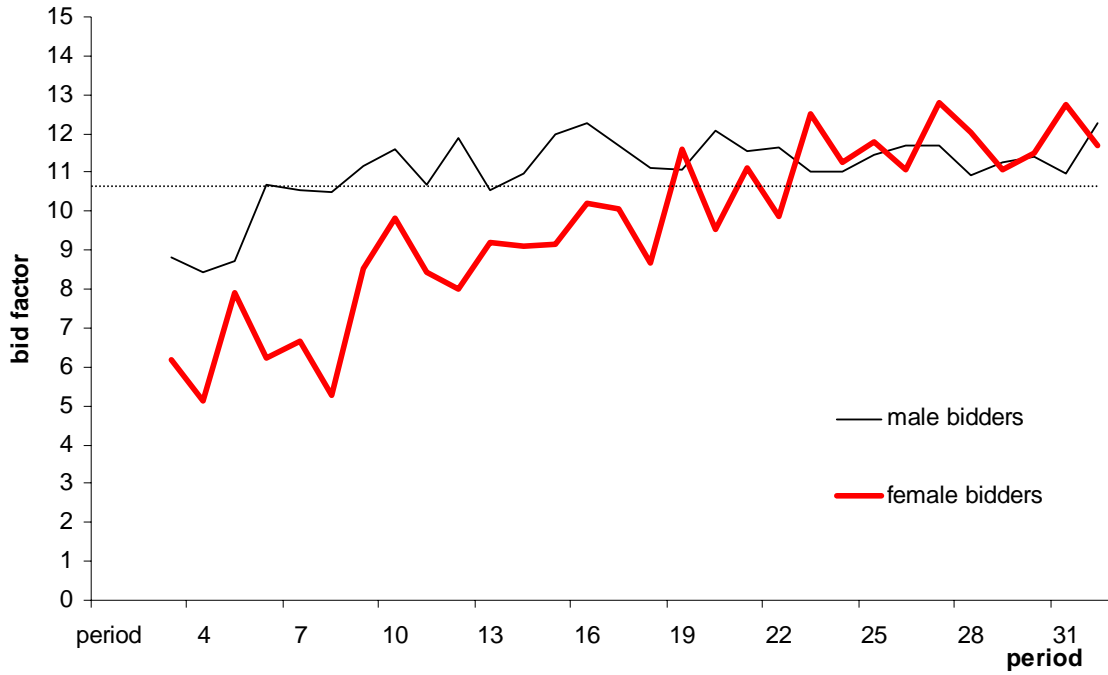
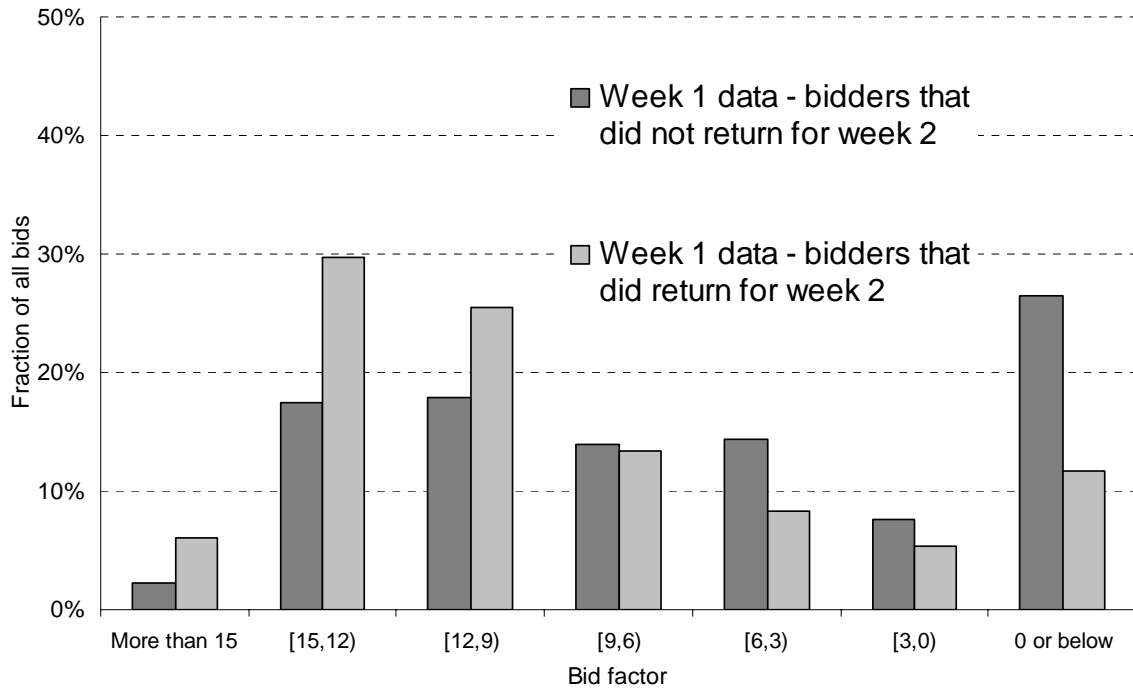


Figure 2: Bid factor by gender (week 1, Low Bankruptcy Group)



Notes: A bid equal to the signal (private estimate) is represented as zero; The risk neutral Nash equilibrium bid (RNNE) is at about 15; The cut-off point for the definition of the winner's curse (expected zero profits) is at 10.71 and it is marked as a dashed line. There is no correction for selection effects due to bankruptcies or not returning for week 2. Data are for region 2 and markets with 6 bidders only.

Figure 3: Distribution of the bid factor of inexperienced bidders
(Signal minus bid)



Notes: First five bidding periods only (periods 3-7). Data from all treatments are included (week 1, region 2, markets with 6 bidders only). A bid factor of zero means a bid equal to the signal. The risk neutral Nash equilibrium bid (RNNE) is about 15; The cut-off point for the definition of the winner's curse (expected zero profits) is at 10.71.

Table 1**Experimental Treatments and Descriptive Statistics**

	Number of Sessions (Number of subjects)	Average Auction Profits ^a (S_m)	Average Profits RNNE ^a (S_m)	Percent Auctions x_{1n} Wins ^a	Percent of Bids $b > \tilde{y}(x_i)^a$ (Above Breakeven Bid)		Percentage Bankrupt	Percent Returning from Week 1	Percent Bankrupt in Week 1 Returning
					All Bidders	High Bidders			
Week 1									
Control	6 (95)	-2.84 (0.478)	4.34 (0.281)	55.8	50.1	73.4	46.3	NA	NA
Random	5 (81)	-1.89 (0.448)	4.73 (0.277)	56.7	43.0	67.1	30.9	NA	NA
Bonus	5 (75)	-2.91 (0.566)	3.49 (0.271)	68.9	38.5	57.1	33.3	NA	NA
Combined	16 (251)	-2.47 (0.285)	4.26 (0.163)	59.9	43.9	66.2	37.5	NA	NA
Week 2									
Control	4 (57)	0.79 (0.374)	4.45 (0.277)	79.2	19.8	34.8	19.3	60.0	47.7
Random	4 (60)	1.05 (0.298)	4.27 (0.233)	81.9	14.2	29.3	6.7	74.1	56.0
Bonus	5 (72)	0.48 (0.288)	4.09 (0.204)	79.4	18.7	33.4	13.9	96.0	88.0
Combined	13 (189)	0.75 (0.183)	4.25 (0.135)	80.1	17.6	32.5	13.2	75.3	60.6

^a Auctions with 6 active bidders only NA = not applicable.

Table 2

Sample Composition in Terms of Ability and
Demographic Characteristics

	Inexperience Subjects	Experienced Subjects
SEX		
Female	42.2%	39.7%
Male	57.8%	60.3%
ACADEMIC MAJOR		
Economics and Business	29.9%	29.6%
Engineering and Science	24.7%	29.1%
Humanities	45.4%	41.3%
VERBAL STANDARDIZED TESTS – ACT AND SAT		
Top 5%	17.9%	21.2%
Above the median but not top 5%	51.0%	49.2%
Below the median	17.5%	17.5%
No ACT and No SAT	13.5%	12.2%
MATH STANDARDIZED TESTS – ACT AND SAT		
Top 5%	24.7%	30.2%
Above the median but not top 5%	52.2%	48.1%
Below the median	9.6%	9.5%
No ACT and No SAT	13.5%	12.2%
COMPOSITE (VERBAL AND MATH) STANDARDIZED TESTS		
Top 5%	19.9%	23.8%
Above the median but not top 5%	57.8%	56.1%
Below the median	8.8%	7.9%
No ACT and No SAT	13.5%	12.2%
ACADEMIC GRADES – GPA AVERAGE		
A+, A, A-	8.4%	10.1%
B+, B	32.3%	33.9%
B- or below	8.0%	7.9%
Freshmen, Sophomore, or no GPA	51.4%	48.1%
TOTAL NUMBER OF SUBJECTS	251	189

TABLE 3- ESTIMATES OF THE CONDITIONAL PROBABILITY OF GOING
BANKRUPT IN WEEK 1

Unobserved heterogeneity?	NO DEMOGRAPHICS		COMPREHENSIVE SCORES USED	
	No	Yes	No	Yes
Lagcumcash	-0.141 (0.038)***	-0.150 (0.041)***	-0.154 (0.038)***	-0.162 (0.041)***
Received high signal this period	2.569 (0.313)***	2.589 (0.316)***	2.695 (0.319)***	2.722 (0.324)***
Received 2nd highest signal this period	2.175 (0.327)***	2.175 (0.329)***	2.303 (0.334)***	2.307 (0.337)***
Female	-	-	0.571 (0.242)**	0.590 (0.254)**
Above 95th percentile	-	-	-0.716 (0.375)*	-0.739 (0.388)*
Below median	-	-	1.252 (0.331)***	1.272 (0.347)***
No aptitude score	-	-	0.064 (0.337)	0.183 (3.606)
Fraction prev. periods received high signal	-3.140 (0.988)***	-3.133 (1.002)***	-2.916 (0.974)***	-2.863 (0.991)***
Fraction prev. periods received 2nd high	-1.033 (0.693)	-1.057 (0.714)	-0.718 (0.693)	-0.734 (0.708)
Engineering/science major	-	-	-0.225 (0.338)	-0.236 (0.353)
Economics/business major	-	-	0.138 (0.270)	0.121 (0.282)
Log duration	-0.316 (0.126)**	-0.281 (0.134)**	-0.292 (0.128)**	-0.261 (0.137)*
Constant	2.481 (0.516)***	-	-2.806 (0.577)***	-
Theta1	-	-2.668 (0.582)***	-	-2.926 (0.618)***
Theta2	-	-1.690 (0.944)*	-	-1.848 (1.029)*
P	-	0.679 (0.393)*	-	0.793 (0.317)**
Log likelihood	-366.3	-365.6	-349.0	-348.6

Note: Standard errors in parentheses.

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

TABLE 4- BIDDING EQUATION FOR INEXPERIENCED SUBJECTS(COMPREHENSIVE APTITUDE SCORES USED)

	No Demographics		Learning without Gender Interaction		Learning with Gender Interaction	
	Low Bankruptcy Group	High Bankruptcy Group	Low Bankruptcy Group	High Bankruptcy Group	Low Bankruptcy Group	High Bankruptcy Group
Cash balances	0.1285 (0.0566)**	0.0962 (0.0240)***	0.1209 (0.0561)**	0.0966 (0.0239)***	0.1064 (0.0549)*	0.0953 (0.0238)***
h(x)	0.2386 (0.3196)	0.0783 (0.2401)	0.2337 (0.3181)	0.0706 (0.2391)	0.2174 (0.3133)	0.0835 (0.2387)
1/ln(t+1)	-3.0306 (0.4269)***	-0.8690 (0.2749)***	-2.9907 (0.4246)***	-0.7830 (0.2740)***	-	-
Male*(1/ln(t+1))	-	-	-	-	-1.3624 (0.5560)**	-0.1715 (0.3408)
Female*(1/ln(t+1))	-	-	-	-	-5.6477 (0.6356)***	-1.8044 (0.4214)***
Female	-	-	-3.2186 (0.9688)***	-2.1020 (0.7973)***	-1.0932 (1.0940)	-1.1040 (0.8577)
Above 95th percentile	-	-	-0.5589 (1.0286)	1.1837 (0.9669)	-0.4903 (1.0392)	1.1443 (0.9651)
Below median	-	-	-0.6999 (1.4911)	-4.4050 (1.3753)***	-0.8991 (1.5032)	-4.3570 (1.3727)***
No aptitude score	-	-	-5.1348 (1.5101)***	-0.6704 (1.0850)	-4.9508 (1.5227)***	-0.6500 (1.0829)
Engineering/science major	-	-	-0.1976 (1.1277)	0.0167 (1.0031)	-0.3472 (1.1393)	-0.0043 (1.0012)
Economics/business major	-	-	-2.7001 (1.0765)**	-0.6048 (0.9192)	-2.8859 (1.0861)***	-0.6370 (0.9175)
Constant	8.9864 (1.0562)***	7.6498 (0.5089)***	11.8951 (1.4903)***	8.8545 (0.8889)***	11.4600 (1.5180)***	8.5493 (0.8950)***
Number of observations	1702	3279	1702	3279	1702	3279
P-value(coefficients same in both sub-samples)	0.00		0.00		0.00	

Note: Standard errors in parentheses.

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

TABLE 5- ESTIMATED BID FACTORS FOR MALES AND FEMALES-INEXPERIENCED SUBJECTS (a)

"BIDFACTORS-MEAN CASH BALANCES USED"(b)

	FEMALE		MALE	
	Bus/Econ	Other majors	Bus/Econ	Other majors
Period1	0.9290 (1.2212)	3.8149 (0.8236)***	8.2046 (0.9493)***	11.0905 (1.2595)***
Period10	6.6303 (1.1444)***	9.5162 (0.5115)***	9.6915 (0.6716)***	12.5774 (1.0279)***
Period20	7.1408 (1.1570)***	10.0267 (0.5241)***	10.0423 (0.6891)***	12.9282 (1.0116)***
Period30	7.7238 (1.2582)***	10.6097 (0.5629)***	10.4456 (0.7615)***	13.3315 (1.0266)***

Note: Standard errors in parentheses.

*** significant at the 1% level.

(a) Estimates used are taken from column 5 of Table 4. Results are for subjects with a comprehensive ability score between the median and 95%.

(b) In each period, cash balances were assumed to take on the mean value in that period for the respective subgroup.

TABLE 6-PROBIT ESTIMATES OF PROBABILITY OF RETURNING TO WEEK 2 SESSION
(COMPREHENSIVE APTITUDE SCORES USED)

	(1) No Demographics	(2) With Demographics
Bonus	1.4362 (0.2763)***	1.6153 (0.3048)***
High return fee	0.6794 (0.2741)**	0.8564 (0.2930)***
Low return fee	0.3837 (0.2494)	0.5472 (0.2638)**
Female	-	-0.1096 (0.2071)
Engineering/science major	-	0.7123 (0.2939)**
Economics/business major	-	0.2147 (0.2312)
Above 95th percentile	-	0.5432 (0.2992)*
Below median	-	-0.2239 (0.3194)
No aptitude score	-	-0.2528 (0.2781)
Sameday	0.4174 (0.2130)*	0.4874 (0.2253)**
Constant	-0.1186 (0.2224)	-0.4301 (0.2916)
Observations	255	255
Log likelihood	-125.97	-115.863

Note: Standard errors in parentheses.

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

TABLE 7-BIDDING EQUATION FOR EXPERIENCED SUBJECTS (COMPREHENSIVE APTITUDE SCORES USED)

	No Demographics		Learning without Gender Interaction		Learning with Gender Interaction	
	High Return Group	Low Return Group	High Return Group	Low Return Group	High Return Group	Low Return Group
Cash balance	0.0409 (0.0148)***	0.1062 (0.0198)***	0.0395 (0.0148)***	0.1064 (0.0198)***	0.0390 (0.0147)***	0.1061 (0.0198)***
h(x)	-0.4490 (0.2238)**	-0.1835 (0.1958)	-0.4321 (0.2208)**	-0.1854 (0.1954)	-0.4366 (0.2209)**	-0.1853 (0.1954)
1/ln(t+1)	-0.6116 (0.2979)**	0.1941 (0.2550)	-0.5399 (0.2950)*	0.2164 (0.2547)	-	-
Male*(1/ln(t+1))	-	-	-	-	-0.2883 (0.3682)	0.2070 (0.2931)
Female*(1/ln(t+1))	-	-	-	-	-0.8781 (0.3951)**	0.2277 (0.3874)
Female	-	-	0.4275 (0.9577)	0.2614 (0.5308)	0.7033 (0.9766)	0.2514 (0.5700)
Above 95th percentile	-	-	0.0585 (1.0204)	-0.6837 (0.5971)	0.0616 (1.0105)	-0.6837 (0.5965)
Below median	-	-	-3.9275 (1.4102)***	-0.2475 (0.9935)	-3.9280 (1.3969)***	-0.2480 (0.9926)
No aptitude score	-	-	-1.5014 (1.3571)	-0.2850 (0.7475)	-1.4865 (1.3442)	-0.2866 (0.7469)
Engineering/science major	-	-	1.0021 (1.1090)	1.7739 (0.6574)***	1.0082 (1.0983)	1.7730 (0.6567)***
Economics/business major	-	-	-2.2637 (1.1214)**	1.0682 (0.6216)	-2.2661 (1.1106)**	1.0675 (0.6210)*
Constant	11.2935 (0.4737)***	10.2982 (0.4334)***	11.7665 (1.1547)***	9.5110 (0.6796)***	11.6536 (1.1517)***	9.5204 (0.6820)***
Number of observations	1996	2360	1996	2360	1996	2360
P-value(coefficients same in both sub-samples)	0.05		0.02		0.02	

Note: Standard errors in parentheses.

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

TABLE 8- ESTIMATED BID FACTORS FOR MALES AND FEMALES-EXPERIENCED SUBJECTS (a)

"BIDFACTORS-MEAN CASH BALANCES USED" (b)

	FEMALE		MALE	
	Bus/Econ	Other majors	Bus/Econ	Other majors
Period1	9.2873 (1.0628)***	11.5534 (0.8588)***	9.4741 (0.9338)***	11.7402 (1.1036)***
Period10	10.1881 (1.0011)***	12.4543 (0.7870)***	9.8824 (0.8878)***	12.1485 (1.0656)***
Period20	10.4169 (1.0007)***	12.6830 (0.7825)***	10.0857 (0.8878)***	12.3518 (1.0619)***
Period30	10.7777 (1.0066)***	13.0438 (0.7805)***	10.2985 (0.8911)***	12.5646 (1.0603)***

Note: Standard errors in parentheses.

*** significant at the 1% level.

(a) Estimates used are taken from column 5 of Table 7. Results are for subjects with a comprehensive ability score between the median and 95%.

(b) In each period, cash balances were assumed to take on the mean value in that period for the respective subgroup.