NBER WORKING PAPER SERIES

GRADING ON A CURVE, AND OTHER EFFECTS OF GROUP SIZE ON ALL-PAY AUCTIONS

James Andreoni
Andy Brownback

Working Paper 20184
http://www.nber.org/papers/w20184

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2014

Andreoni gratefully acknowledges the support of the National Science Foundation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Grading on a Curve, and other Effects of Group Size on All-Pay Auctions
James Andreoni and Andy Brownback
NBER Working Paper No. 20184
May 2014
JEL No. C91,C92,D47,H52

ABSTRACT

We model contests with a fixed proportion of prizes, such as a grading curve, as all-pay auctions where higher effort weakly increases the likelihood of a prize. We find theoretical predictions for the effect of contest size on effort and test our predictions in a laboratory experiment that compares two-bidder auctions with one prize and 20-bidder auctions with ten prizes. Our results demonstrate that larger contests elicit lower effort by low-skilled students, but higher effort by high-skilled. Large contests also generate more accurate rankings of students and more accurate assignment of high grades to the high-skilled.

James Andreoni
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
andreoni@ucsd.edu

Andy Brownback
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
abrownback@ucsd.edu

An online appendix is available at:
http://www.nber.org/data-appendix/w20184

Consider a student deciding how much effort to devote to a course graded on the curve. Aside from her intrinsic motivation, she must consider her innate ability, the abilities of her classmates, and the percentage of students she believes will receive each grade. So, while some elements of her choice are unaffected by the choices of her classmates, many elements of her choice are strategic. This paper explores one specific element of that strategic choice: how the enrollment in the course affects the selection of effort. We will model this environment as an all-pay auction, with the enrollment of the course being the number of bidders in the auction and the number of "A" grades being the number of prizes in the auction.

Within the education community there is little discussion about the effect of enrollment on student strategic incentives. The education literature examines the impact of enrollment by proxy by exploring the effects of student-teacher ratio on test scores (Mosteller, 1995), attendance (Romer, 1993), and future earnings (Card and Krueger, 1996; Carniero and Heckman, 2003). The literature directly considers enrollment by estimating its impact on the grades assigned to students (Kokkelenberg, Dillon, and Christy, 2006), but frames the argument in terms of diseconomies of scale for teaching output, ignoring changing strategic incentives for student effort.

This gap in the literature can be attributed to the fact that much of the education research explicitly or implicitly assumes non-strategic interaction between students. Under this assumption, students will only react to changes in enrollment to the degree that enrollment affects the inputs—such as student teacher ratio—of their production function. Beginning with Becker and Rosen (1992), however, another branch of the literature arose and formalized students' incentives to select effort strategically when a course is graded according to an explicit curve. This framework generates testable predictions for courses that move from non-strategic to strategic settings. Paredes (2012) finds significant explanatory power in these predictions and discovers that students are indeed sensitive to a change between strategic and non-strategic environments. We hope to take this discussion of strategic sophistication in the classroom one step further and investigate students' reactions to subtle changes within the strategic classroom environment in a way that can be predicted by game theoretic models[1].

In everything from college admissions[2], to scholarship awards[3], to teacher salaries within schools[4] we see contests that use proportional awarding despite large differences in the number of participants. There appears to be an insensitivity among mechanism designers to the impact of contest size on behavior. We will show here, both theoretically and through an experimental test, that there are significant effects of a contest's size on individual effort even when the proportion of awards to participants is fixed. To fix ideas, our discussion will center on the application of this result to the practice of grading on a curve, but the results apply to a much more general class of contests. We will return to more general applications later in the paper.

1 Background

The branch of economics investigating competitive interaction with costly effort and uncertain payoffs began when Tullock (1967) and Krueger (1974) adapted the strategic framework of Nash (1951) to generate analytical predictions for wars of attrition and competitive rent-seeking, respectively. Within this field, there are three prominent models of competitive interaction, the rank-order tournament, the Tullock contest, and the all-pay auction.

[1]While our model and experiment exist within an environment with a strict grading curve, we believe that our results apply to all courses, even when instructors adhere to a less-strict grading curve. This is because the reality of modern education is such that relative position will always hold some influence over grades. Thus a rational student will certainly be aware of the significance of her relative position regardless of what her absolute position may be. Therefore, enrollment in a course affects not only the inputs to the student's production function, but also the strategic environment in which she operates.

[2]For example, in Texas, House Bill 588 grants high school students automatic admission to any state university if they graduate in the top 10% of their graduating class. In Kansas, the top 1/3 gain automatic admission. In California, high school students in the top 9% of their graduating class are guaranteed admission one of the University of California schools.

[3]As an example, the Bright Flight scholarship program in Missouri awards scholarships to high school seniors who score in the top 3% on their SAT or ACT tests.

[4]In North Carolina, Senate Bill 402 Section 9.6 (g) stipulates that each school administration must select the top performing 25% of their teachers to receive more favorable contracts.

In the decades since this field began, the breadth of applications has expanded well beyond the original motivating examples. Lazear and Rosen (1981) explore tournaments as optimal labor contracts. Becker and Rosen (1992) then consider relative grading to be a form of tournament and compare its results to those of an absolute grading scheme. Hillman and Riley (1989) and Baye, Kovenock, and de Vries (1993) model political rent-seeking as a form of all-pay auction. Amann and Leininger (1996) expand, and simplify to an extent, the model by introducing incomplete information as a form of Harsanyi (1973) purification, which allows for an equilibrium in pure strategies. It was not until Baye, Kovenock, and de Vries (1996), however, that the all-pay auction received a full characterization of its equilibrium, though the result offered no strong predictions, rather finding that there exists a continuum of equilibria in the model.

Moldovanu and Sela (2001) theoretically explore the optimal allocation of prizes in a contest given different objectives of the mechanism designer. Siegel (2009) makes great strides in generalizing the nature of contests and the equilibrium actions of contestants. More recently, Olszewski and Siegel (2013) theoretically address a very similar question to ours when they develop equilibrium predictions for large but finite contests.

1.1 Experimental results

A complete recounting of experimental evidence on all-pay auctions, rank-order tournaments, and related contests can be found in a review by Dechenaux, Kovenock, and Sheremeta (2012). For the sake of brevity, we will discuss only the most relevant experiments here.

Bull, Schotter, and Weigelt (1987) first brought rank-order tournaments into the laboratory, finding that bidders approached the equilibrium after learning, but demonstrated surprisingly high variance. Equilibrium predictions for all-pay auction experiments, on the other hand, have consistently missed the mark. Potters, de Vries, and van Winden (1998), Davis and Reilly (1998), Gneezy and Smorodinsky (2006), Barut, Kovenock, and Noussair (2002) all find substantial over-dissipation of rents, that is, aggregate bidding that exceeds the aggregate value of the prizes. Other papers introduce heterogeneity in the valuations of bidders, (Müller and Schotter, 2010; and Noussair and Silver, 2006), and find similar over-dissipation, but also find that bidders begin to separate themselves by type, with stark differences appearing in the bidding strategies of high- and low-valuation bidders.

With optimal contest design in mind, several papers have looked at the impact of small changes in the number of opponents on the effort selection of bidders (Gneezy and Smorodinsky, 2006; Müller and Schotter, 2010; Barut et al., 2002), but each of these allows the proportion of winners to change with the number of participants. Meaning that the ex ante equilibrium probability of receiving a prize changes with the size of the auction. Harbring and Irlenbusch (2005) make small variations in the size of tournaments while holding constant the proportion of winners, but they employ a convex cost function, common values, and a cap on the maximum effort choice. To our knowledge, no paper has ever focused on the policy implications of combining or separating multiple contests while maintaining the proportion of winners. We wish to address this gap in the literature by restricting our analysis to situations where the proportion of prizes to participants is fixed and testing the effects of a large-scale change in the number of participants in the contest on the effort selection of those participants. This specific focus allowed us to design the experiment in a way that more precisely captures systematic effort changes across contests of different sizes.

2 Theoretical Model

In an all-pay auction, players simultaneously place irreversible bids for one of a limited number of prizes. Bidders must pay their bids regardless of the outcome of the auction. Since prizes are awarded to the highest bidders, an increase in the bid weakly increases the probability of receiving a prize.

2.1 Grading on a curve as an all-pay auction

While the economic study of all-pay auctions began as a model motivated by the examples of competitive rent-seeking behavior seen in lobbying efforts and research and development contests, the intuition behind the all-pay auction can naturally extend into other settings such as a course graded on a curve, job promotions, the allocation of bonuses among workers in a firm, or the method of awarding grants to applicants. Indeed, the framework of analysis can be applied to any environment with costly, deterministic effort and probabilistically awarded prizes. In a course graded on the curve, for example, each student in a classroom casts a "bid" by studying a given amount. These bids are measured against each other, and the students who cast the highest bids are awarded the "prizes" of higher grades.

There is an interesting question regarding changes in the number of participants of the contest. What are the independent effects of the number of participants on the effort chosen by each participant? Is the effect dependent on the ability of the student? How well do outcomes of the contest reflect the true relative ranking of participants' abilities when only effort is observable? Translated, does the course enrollment affect a student's effort in studying? Does it affect high-skilled students differently from low-skilled students? Is there an enrollment level where effort choices are more or less reflective of students' abilities? For each of these questions we will provide a theoretical prediction and a result from a controlled laboratory study. With these results established, we hope to encourage contest designers to take into account our insights about the independent effects of a contest's size.

2.2 Definitions and procedures

Let us suppose a contest environment with independent private values in the style of Vickrey (1961). Suppose the contest features N bidders competing for M prizes, with $M < N$. For simplicity, call the number of participants the size of the contest. Participants have varying levels of ability, and the cost of a given level of effort decreases as a participant's ability increases, making the surplus value of a prize higher for the higher ability participants. Since effort choices are invariant to affine transformations of the utility function, we can model heterogeneous costs of effort as heterogeneous values from winning prizes. Call this value of winning a player's valuation, v_i, and suppose that it is drawn privately from a commonly known distribution, $F(v_i)$. Under incomplete information no participant has access to the vector of other participants' true abilities, $\vec{v_{-i}}$. Effort is measured by a participant's bid value, b_i, and is costly regardless of the outcome of the auction. Auctions are one-shot, and bids are cast simultaneously, so bidders have no ability to condition their effort on other bids. The auctioneer determines the number of bidders and prizes in the same way that an administrator might choose the enrollment of a course. In order to optimally design the mechanism, the designer must take into account the effect that the number of participants has on the predicted effort levels.

With incomplete information about the valuations of other bidders, we can generate a symmetric Bayesian Nash Equilibrium in pure strategies. At equilibrium, there exists a continuous optimal bidding function, $B(v_i)$, that maps from the valuation space onto the bidding space: $B : v_i \to b_i$, where $v_i \in [0, 1]$ and $b_i \in [0, 1]$. Let $P_{N,M}(b_i)$ be the probability that bid b_i will win a prize in an auction with N participants and M prizes.

2.3 Bidder's Utility

For simplicity, we assume risk neutrality. We show in the appendix that the qualitative results generalize to risk aversion.[5] A bidder's utility is:

$$\begin{aligned} U(b_i; v_i, N, M) =& P_{N,M}(b_i)(v_i - b_i) + (1 - P_{N,M}(b_i))(-b_i), \\ =& v_i \times P_{N,M}(b_i) - b_i \end{aligned} \tag{1}$$

In this auction, the M bidders with the highest bids will each receive one of the prizes. Therefore the probability of receiving a prize is weakly increasing in the amount bid by construction.

[5]The generalization, however, does require common knowledge of risk aversion, common knowledge of rationality, and shared risk aversion parameters.

Proposition 1: ***Optimal Bidding is Weakly Monotonic in Valuation***

Proof: See Appendix[6].

Under monotonicity, there exists a probability function that maps valuations to probabilities of winning the auction. Denote this function $Z_{N,M} : v_i \to [0,1]$. This function will represent the probability that a bidder's valuation is higher than the valuations of at least $N - M$ of the opposing bidders. Mathematically, this is expressed in the form of an order-statistic:

$$Z_{N,M}(v_i) = \sum_{k=N-M}^{N-1} \left(\frac{(N-1)!}{(N-1-k)!k!} \right) F\left(v_i\right)^k \left(1 - F(v_i)\right)^{N-1-k},$$

where $F(v_i)$ is the cumulative distribution function of the valuations.

2.4 Optimal Bidding Function

So far, we have assumed that a symmetric bidding function, $B : v_i \to \mathbb{R}^+$, exists, is well-defined, and continuous. We then demonstrated its monotonicity. Continuous, monotone functions are invertible, implying that there exists a function, $B^{-1}(b_i)$, that maps bids cast into the valuations implied by those bids. Denote this inverse function $V(b_i)$.

To demonstrate the optimality of this bidding function, the bidding function must solve the first order condition of the bidder's utility,

$$U(b_i; v_i) = v_i \times Z_{N,M}(V(b_i)) - b_i, \tag{2}$$

where $Z_{N,M}(V(b_i))$ captures the bidder's incentive to misrepresent his valuation by casting a higher or lower bid than what his bidding function would prescribe. At equilibrium this term must be equal to $Z_{N,M}(v_i)$. Maximizing Equation (2) with respect to b_i returns

$$\frac{\partial U}{\partial b_i} = v_i \times \frac{\partial}{\partial b_i} \left[\sum_{k=N-M}^{N-1} \left(\frac{(N-1)!}{(N-1-k)!k!} \right) F\left(V(b_i)\right)^k \left(1 - F(V(b_i))\right)^{N-1-k} \right] - 1 \equiv 0.$$

[6] Also included in the appendix is a version of this proof under risk aversion.

Taking the derivative and rearranging, we find

$$\begin{aligned}\frac{1}{V'(b_i)} =& v_i \times \sum_{k=N-M}^{N-1} \left\{ \left(\frac{(N-1)!}{(N-1-k)!k!} \right) \times \right. \\ & \left[k \times f\left(V(b_i)\right) F\left(V(b_i)\right)^{k-1} \left(1 - F(V(b_i))\right)^{N-1-k} - \right. \\ & \left. \left. (N-1-k) \times f\left(V(b_i)\right) F\left(V(b_i)\right)^{k} \left(1 - F(V(b_i))\right)^{N-2-k} \right] \right\}.\end{aligned}$$

Since $V(b_i) \equiv B^{-1}(b_i)$, it follows that, $\frac{1}{V'(b_i)} = B'(v_i)$.

This derivation yields a first order differential equation from which we can solve the general form of the bidding function under risk neutrality

$$\begin{aligned}B'(v_i) =& v_i \times \sum_{k=N-M}^{N-1} \left\{ \left(\frac{(N-1)!}{(N-1-k)!k!} \right) \times \right. \\ & \left[k \times f\left(V(b_i)\right) F\left(V(b_i)\right)^{k-1} \left(1 - F(V(b_i))\right)^{N-1-k} - \right. \\ & \left. \left. (N-1-k) \times f\left(V(b_i)\right) F\left(V(b_i)\right)^{k} \left(1 - F(V(b_i))\right)^{N-2-k} \right] \right\}. \qquad (3)\end{aligned}$$

3 Experimental Model

Equation (3) provides general results and implications for any values of N and M and cumulative distribution $F(v_i)$. In our experiment, we will be testing the following two pairs: $(N, M) = (2, 1)$ and $(20, 10)$ with valuations uniformly distributed, $v_i \sim U[0, 1]$. The distributional assumption is without loss of generality, since the analysis is identical under any distribution, though the exact values of our predictions will vary with the distribution chosen. We will now proceed to derive closed-form solutions for the bidding functions of these two pairs.

3.1 Bidding Functions

Proceeding from Equation (3) and substituting in $N = 2,\ M = 1,\ F(v_i) = v_i$ gives us

$$B'(v_i) = v_i.$$

Solving the differential equation yields[7]

$$B(v_i) = \frac{v_i^2}{2}.$$

Keeping the uniform distribution, but evaluating the bidding function for $N = 20$ and $M = 10$ we find

$$\begin{aligned} B'(v_i) =& v_i \times \sum_{k=10}^{19} \left[\left(\frac{(19)!}{(19-k)!k!} \right) \times \right. \\ & \left. \big((k)V(b_i)^{k-1}(1-V(b_i))^{19-k} - (19-k)V(b_i)^k(1-V(b_i))^{18-k} \big) \right]. \end{aligned}$$

The equilibrium assumption requires that bids reveal values truthfully, so $V(b_i) = v_i$. Substituting and rearranging,

$$B'(v_i) = \sum_{k=10}^{19} \left[\left(\frac{(19)!}{(19-k)!k!} \right) \left((k)v_i^k(1-v_i)^{19-k} - (19-k)v_i^{k+1}(1-v_i)^{18-k} \right) \right].$$

This implies

$$B'(v_i) = 923780 v_i^{10}(1-v_i)^9.$$

Solving this differential equation results in the optimal bidding function:

$$\begin{aligned} B(v_i) =& 83980v_i^{11} - 692835v_i^{12} + 2558160v_i^{13} - 5542680v_i^{14} \\ & + 7759752v_i^{15} - \frac{14549535}{2}v_i^{16} + 4564560v_i^{17} \\ & - 1847560v_i^{18} + 437580v_i^{19} - 46189v_i^{20}. \end{aligned}$$

Rather than using $f(v_i) = U[0,1]$, as above, our subjects will draw valuations uniformly from the set $\{\$0.01, \$0.02, \ldots, \$20.00\}$. We assert that subjects

[7] Technically, $B(v_i) = \frac{v_i^2}{2} + C$, however, we can use dominance to show that $C = 0$. Suppose that $C > 0$. Therefore, in equilibrium, a bidder with valuation $v_i = 0$ will make a bid, $b_i > 0$. That bid guarantees a negative payoff, so the same bidder would be made better off by deviating and choosing $b_i = 0$, which guarantees a zero payoff. Thus, $C \leq 0$. But, we restrict bids to positive values so $C = 0$.

view this setting as continuous, so we will use a rescaled version of the continuous bidding functions to generate predictions.

$$
\begin{aligned}
B_2(v_i) =& 10 \times \left(\frac{v_i}{20}\right)^2 \\
B_{20}(v_i) =& 20 \times \left(83980 \left(\frac{v_i}{20}\right)^{11} - 692835 \left(\frac{v_i}{20}\right)^{12} + 2558160 \left(\frac{v_i}{20}\right)^{13} \right. \\
& - 5542680 \left(\frac{v_i}{20}\right)^{14} + 7759752 \left(\frac{v_i}{20}\right)^{15} - \frac{14549535}{2} \left(\frac{v_i}{20}\right)^{16} \\
& + 4564560 \left(\frac{v_i}{20}\right)^{17} - 1847560 \left(\frac{v_i}{20}\right)^{18} + 437580 \left(\frac{v_i}{20}\right)^{18} \\
& \left. -46189 \left(\frac{v_i}{20}\right)^{20} \right).
\end{aligned}
$$

Graphically, the optimal bidding functions for our experiment can be seen in Figure 1.

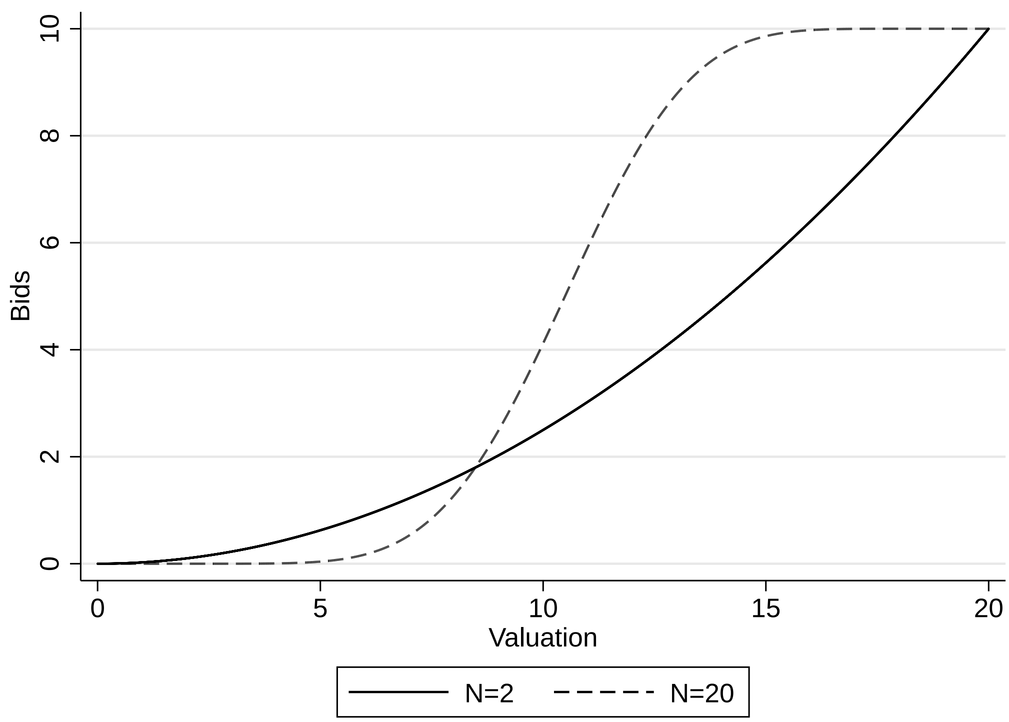

Figure 1: Optimal Bid Functions

To understand the intuition behind the shapes of the bidding functions, consider first the limit case where we maintain the proportion of prizes to participants ($M = \frac{N}{2}$), but let $N \to \infty$. In this case, the bidding function will look like a step function, where all bidders with $v_i < \$10$ bid near-zero and all bidders with $v_i > \$10$ bid near $b_i = \$10$. This is because only the top half of bidders receive prizes, so bidders below the median should best respond by bidding as little as possible, and bidders above the median should bid only just enough that they are guaranteed a prize.[8] The minimum winning bid in this case is equal to the valuation of the median participant, $v_i = \$10$.

Since our large auction is ten times the size of the small auction, the Law of Large Numbers will draw it closer to the limit case. That is, the distribution of valuations in the larger auction is expected to be more reflective of the underlying probability distribution than the distribution of valuations in the smaller auction. This convergence will cause the median of the realized distribution in the larger auction to be closer to the median of the probability distribution, giving less uncertainty to the minimum bid required to win a prize in the larger auction.

Bearing in mind this lower level of uncertainty in the larger auction, consider the net benefit of lowering a bid from the limit case of $b_i = \$10$ for a bidder with $v_i > \$10$ in each of our auctions. These costs depend both on the auction's size and the bidder's valuation. In either auction, the marginal *benefit* of lowering a bid is constant, since bids are paid with certainty, so foregone bids are recovered with certainty. On the other hand, the marginal *cost* of lowering a bid is paid stochastically by lowering the probability of winning a prize. For high-valuation bidders, this decrease in probability is greater in the large auction than in the small auction, so bids cast by high-valuation bidders are larger in the large auction.

Bidders with low valuations face the constant cost of increasing their bid from the limit case of $b_i = \$0$ and the probabilistic benefit of increasing the likelihood of receiving a prize. At low enough valuations, bidders see greater

[8]In the limit case equilibrium, the marginal player will be mixing to ensure that bidders above the median maintain their bids of \$10 instead of cheating down towards \$0.

increases in the probability of winning a prize in the small auction, so bids in the small auction rise above those of the large auction.[9][10]

4 Experimental Hypotheses

Principally, a test of our model is a specific test that bidders behave according to the risk-neutral Nash Equilibrium displayed in Figure 1, but we can also consider several more general predictions of the model. We will restrict our focus to predictions that could serve as components of a mechanism designer's objective function. We have identified three such objectives for which our model generates predictions that our experimental data can put to the test. Objective 1 is the total effort of the subjects, Objective 2 is the distribution of effort across subject types, and Objective 3 is the ability of the auction to create an accurate ranking of subjects based on their effort alone.

In the event that the equilibrium prediction holds exactly, these objectives will also match their predictions exactly. But, even if the precise equilibrium prediction is rejected by statistical tests, the relative predictions about the effects of contest size may still be informative for the policy debate about the merits of different class sizes. Thus, we will test the objectives of the mechanism designer independently of the equilibrium test.

Hypothesis: ***Equilibrium predictions hold***

Given the complexity of the equilibrium bidding functions, strict hypotheses about equilibrium bidding seem overly restrictive. In addition to the computa-

[9]Surprisingly, the equilibrium is not strongly affected by adding in risk-aversion to the bidders' utilities, as shown in the appendix. This results mainly from the fact that bidding is driven mostly by changes in probability, which is unaffected by risk aversion, not changes in surplus, which is affected by risk aversion. See Figure **??** for a graph of the equilibrium bidding functions under common risk aversion parameters.

[10]In the appendix, we also include a graph of the equilibrium under a common joy-of-winning value. Trivially, this improves the fit of our model to the data. This is mechanical, since the joy-of-winning specification is simply the Nash Equilibrium specification with an added degree of freedom. We do not include this specification of the model in the discussion, because it provides no alternative qualitative predictions for us to compare with the Nash Equilibrium model.

tional difficulty, there appears to be no obvious heuristic that subjects might adopt. We might believe that our subjects could instinctively discover the equilibrium through experimentation, but even with experience, the complexity of the bidding functions makes a tight fit between the data and the theory unlikely in either auction size. Nonetheless, we will consider equilibrium as a starting point and move from there to the more general objectives identified below.

4.1 Qualitative Predictions: The Designer's Objectives

Objective 1: ***Maximize aggregate bidding.***

In our education example, this means maximizing the total effort of students. Our equilibrium bidding function predicts that aggregate bidding will be higher in the larger auction. Increasing the size of the auction gives bidders less uncertainty about the valuation of the minimum winning bidder. With tighter predictions about the minimum winning bidder's valuation, bidders face higher probabilistic costs when decreasing their bids, which raises bids for the majority of bidders and increases the rent dissipation.

Objective 2: ***Generate the desired distribution of bids.***

Translated into the education context, this implies that the instructor may value effort exerted by one type of student over another. For example, the instructor may prefer a mechanism that promotes effort by low-ability students. As the auction size increases and the uncertainty decreases, we expect bids in the larger auction to dominate for high valuations. At the same time, among bidders with low valuations, the decrease in uncertainty drives the bids towards zero in the large auction. Therefore, we expect to see higher bids in the small auction across low valuations. Our equilibrium predictions place the crossover point at $v_i = \$8.48$. This point corresponds to the point of intersection between the optimal bidding functions in Figure 1.

Objective 3: ***Accurately order the valuations of bidders based only on their bids.***

An instructor may wish to employ the mechanism that generates the most accurate ordering of his students' unobservable abilities based only on their observable effort. Indeed, any mechanism designer may have a desire to rank of bidders, inferring their true abilities by their bids. Thus, regardless of the outcome of any one auction, the designer wants to be able to infer the true ranking of abilities across auctions from only the bids cast. Our theory asserts that, in equilibrium, there should be a one-to-one, monotonic matching of bids to valuations for both auctions, making the bids a perfect proxy for relative ability. In practice, deviations from the equilibrium bid will almost surely make the inferred ranking imprecise. Therefore, one objective of an instructor might be to employ the mechanism that minimizes the impact of these deviations on the accuracy of the ordering of bidders.

5 Experimental Procedures

To test the sensitivity of effort to a change in the size of a contest, we employ an independent private value all-pay auction with multiple prizes. We use a paired bidding design similar to Kagel and Levin (1993) and Andreoni, Che, and Kim (2007) in which we elicit bids for both the large auction and the small auction from all bidders every round. Using this design, we can pair data perfectly and analyze the difference in bidding at each valuation rather than relying on subject fixed effects for statistical power. These pairs also give us more power to test our hypotheses about the relative levels of bids at any point in the distribution of valuations.

5.1 Recruitment and Participation

60 undergraduate students from the University of California, San Diego participated in our experiment. We recruited all of our subjects by means of online advertisements on the EconLab website. The experiment was conducted in 3

sessions in February of 2013 in the EconLab at UCSD. Each session required exactly 20 subjects. All subjects received a \$20 participation payment in addition to the money gained or lost in the experiment. Our sessions lasted approximately 90 minutes, and subjects earned between \$15 and \$45.

5.2 Instructions

We first read the instructions aloud to all subjects and then administered a quiz to test their comprehension of the auction formats and the payoffs from different outcomes. The instructions and quiz can be found in the appendix.

We designed instructions to clarify the means by which we awarded prizes and the payoffs conditional on receiving or not receiving a prize. We carried out the entire experiment on computers using the Z-Tree Economics Software (Fischbacher, 1999). This allowed us to keep key instructions posted at the top of every decision screen. Specifically, we reminded subjects that any bid would be deducted from their winnings regardless of the outcome of the auction, and that their final payment for the experiment would be based on one round selected at random. We reinforced the number of opponents they faced and the number of prizes available in each round. We also reminded subjects that they would see two auctions for each valuation they drew, first the small auction, then the large auction.

5.3 Auction Design

Every round, subjects received valuations drawn uniformly in 1 cent increments between \$.01 and \$20.00. That is, $v_i \sim U[\$.01\,,\,\$20.00]$. With this valuation, subjects participated in both a "2-Person Auction," for which they were randomly and anonymously paired, and a "20-Person Auction," which consisted of all participants.

Subjects cast bids for both auctions in each round. Subjects were repeatedly reminded of the two auctions they would see and the order. Upon submission of both bids, we revealed the auction outcomes to the subjects. This included the subject's own payoffs from both auctions as well as all bids cast by all bidders in both auctions that round. All bids were cast without knowledge of opponents' bids or valuations, so we consider them simultaneous bids under incomplete information. No communication was allowed, and we never revealed the valuations associated with any bids.

The auction stage was repeated 20 times in session 1 and 15 times in sessions 2 and 3. In every repetition, subjects drew new valuations and new pairs were randomly assigned. For the sake of uniformity, we exclusively analyze the first 15 rounds of each session, but our results only strengthen with the inclusion of the final 5 rounds of the first session.[11]

After completing the auction rounds, subjects made 3 incentivized choices between gambles in order to elicit their preference for risk. We employed a discretization of the method detailed in Andreoni & Harbaugh (2009). These measures of risk aversion will be used as a control in the analysis.

6 Results

With 15 rounds in which subjects drew valuations and cast bids in two different auctions, we collected a total of 900 different valuations and 1800 different bids. For the analysis, however, we will drop any bid that exceeds the valuation of the bidder. These bids guaranteed a negative profit for the bidders. There are 14 such bids across both auction sizes. The dropped bids were evenly distributed across periods and auction sizes. It is not clear if they resulted from typing errors, confusion, apathy, or a desire to win at any cost. We will indicate when our analysis is substantively affected by removing these bids.

Figure 2 graphs the bidding behavior in the 2-Person and 20-Person Auctions. We split valuations split into 10 different, equally sized bins. Adjacent dots represent bidding in different auction sizes for bidders with the same valuations. They are placed alongside each other to facilitate comparison. Every dot represents the mean bid for that bin with the 95% confidence intervals

[11] We planned each session to fit 20 rounds into 90 minutes. During the first session we discovered that 20 rounds would not work due to a slow server. The second and third session were shortened to accommodate this.

overlaid. The dashed lines represent the equilibrium predictions for the two auctions.

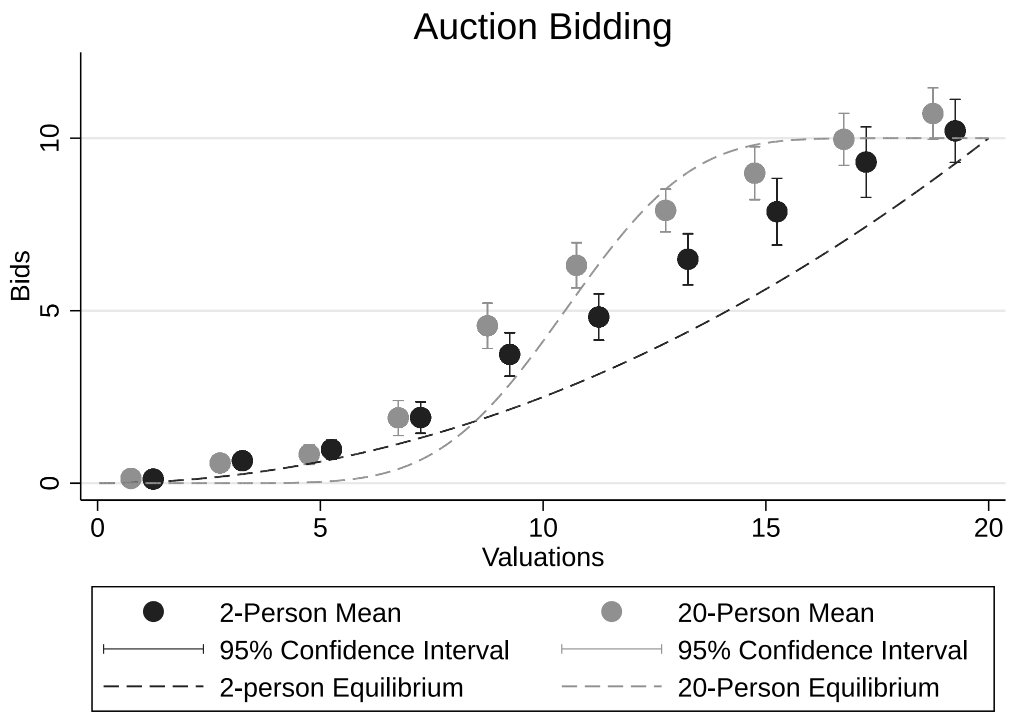

Figure 2: Auction Bidding

6.1 Hypotheses

It is clear from Figure 2 that, on average, our subjects performed rather closely to the risk neutral Nash Equilibrium predictions. Despite this, when we test the difference between observed bids and the bids predicted by the Nash Equilibrium we can reject equilibrium bidding in both the small ($t = 7.03$, $p < .001$, $n = 893$) and large ($t = 4.73$, $p < .001$, $n = 893$) auctions.[12] If we employ a

[12]We calculate the difference between the observed bid and the predicted bid and perform an uncontrolled linear regression with standard errors clustered at the individual level. The statistics of the constant capture the mean deviation. It is that set of statistics that are reported.

seemingly more conservative notion of over-bidding and consider the revenue of a given round as the unit of observation, we can reject equilibrium bidding at an even higher significance level in both the small ($t = 10.46$, $p < .001$, $n = 45$) and large ($t = 5.82$, $p < .001$, $n = 45$) auctions.

Of course, one should not be surprised that in this unique environment with complex and precise predictions, our model fails these statistical tests. As the literature on auction experiments reveals, the theory often provides a good benchmark but seldom predicts precise outcomes. Indeed, as we noted in regards to Figure 2, the qualitative features of the theory are captured by the plotted bids. We frame our objectives in terms of the relative performance of the two auction sizes, so, while noisy behavior may be enough to cause our statistical tests to reject the equilibrium prediction, it may not affect the relative performance with respect to our stated objectives. Therefore, the question becomes whether, despite the lack of a precise fit, the qualitative predictions of our model and the intuitions gleaned from our theory extend to the observed data. We will explore this question within the context of the following three objectives.

Objective 1: ***Maximize aggregate bidding.***

In 33 of the 45 rounds, relative revenues were consistent with our theory. That is, the larger auction generated greater bidding than the smaller auction. We perform a paired t-test on the difference in total revenue between auctions and find that the larger auctions generate significantly more revenue ($t = 4.74$, $p < .001$, $n = 45$). Our paired design allows us to test this hypothesis on the individual level and perform a paired t-test that matches each bid in the large auction with its counterpart in the small auction. We can reject the null hypothesis of equal bidding in both auctions, in favor of the alternate hypothesis that bidding in the large auction is significantly larger than bidding in the small auction ($t = 5.02$, $p < .001$, $n = 888$).[13]

[13] These results are unchanged by including all outlying bids except one particularly extreme outlier, where the bid was 9,900 times the valuation of the bidder.

Objective 2: ***Generate the desired distribution of bids.***

Our model predicts that bids in the small auction will be greater than bids in the large auction for all $v_i < \$8.48$. Figure 3 graphs the predicted difference along with the observed mean differences across 10 equally spaced bins. While the magnitudes are clearly attenuated, the qualitative predictions appear to be captured quite well.

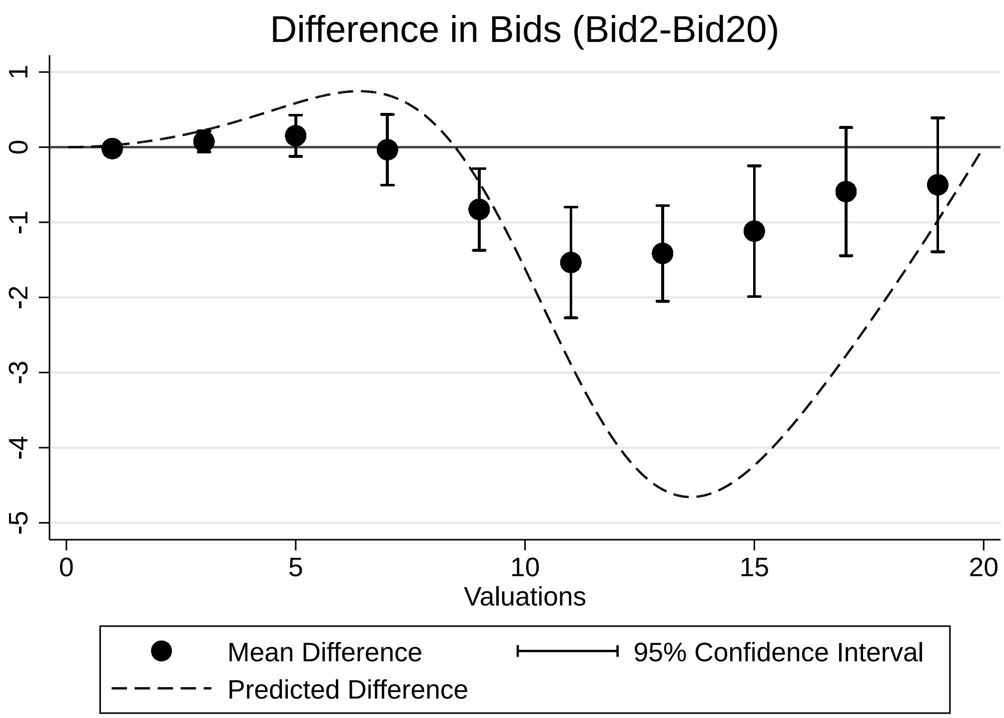

Figure 3: Bidding Differences

We test this prediction more formally in three ways. First by estimating the point at which bids in the large auction cross over bids in the small auction, second with a direct test of the sign of the bidding difference predicted by our model at the theoretical and fitted crossover points, and finally with a test of the relative revenues on either side of the theoretical and fitted crossover points.

To estimate the valuation at which bids in the large auction begin to dominate those of the small auction, we perform a random effects regression of the difference in bids on the valuation of the bidder to the first, second, and third power.[14] The regression results can be seen in Table 1.

Table 1: Predicting the Crossover Point

	Difference in Bids (Small-Large)
Valuation	.211
	(0.169)
$Valuation^2$	-0.047**
	(0.023)
$Valuation^3$	0.002**
	(0.001)
Constant	-0.073
	(0.258)
N	888

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Standard errors clustered by subject.

Using these results, we can predict the difference in bids at each valuation, and determine at what valuation the predicted difference is equal to zero. Call this point the fitted crossover point. The predicted difference in bids is plotted in Figure 4. Notice that the shape of the predicted difference in bids captures many of the characteristics of the equilibrium predictions, but with a smaller magnitude. From Figure 4 you can see that our fitted crossover valuation is $v_i \approx \$5.32$, and our theoretical crossover valuation is $v_i \approx \$8.48$.

To test if our model accurately predicts the relative levels of bidding on either side of the theoretical crossover, $v_i = \$8.48$, we will split our data into two bins, one on each side of the crossover point. We generate a dummy variable, *LowerBin*, to indicate valuations below the crossover. To best exploit our paired data, our dependent variable will be the difference between the bid in the small auction and the bid in the large auction. Our theory predicts that the difference will be positive until $v_i = \$8.48$ and negative afterwards.

[14] We can use random-effects instead of fixed-effects because our valuations were randomly assigned and so are mechanically uncorrelated with other explanatory variables.

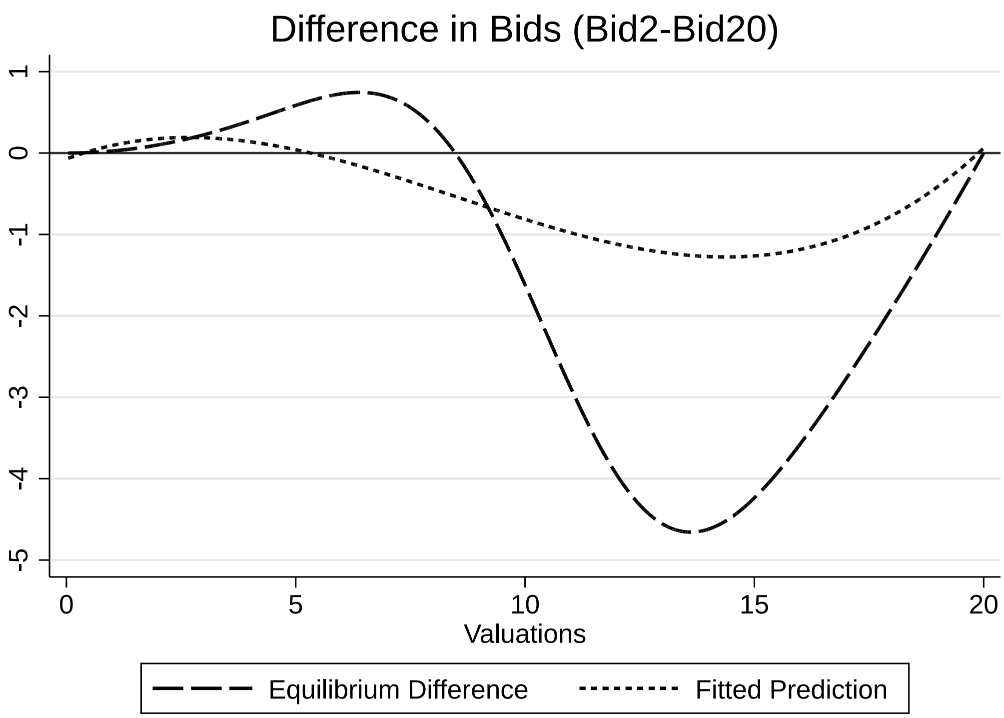

Figure 4: Predicted Difference in Bids

We regress this difference onto $LowerBin$ and risk aversion controls using a random effects regression with subject-specific random effects terms. Mathematically,

$$D_{i,t} = \beta_0 + \beta_1 LowerBin_i + \rho RA_{i,t} + \epsilon_{i,t}, \quad (4)$$

where

$$D_{i,t} = Bid2_{i,t} - Bid20_{i,t},$$

and[15]

$$RA_{i,t} = \{\alpha_{1,i}, \alpha_{2,i}, \alpha_{3,i}, \alpha_{1,i} \times LowerBin_{i,t}, \alpha_{2,i} \times LowerBin_{i,t}, \alpha_{3,i} \times LowerBin_{i,t}\}.$$

The results can be seen in the left-hand side of Table 2.

[15]Here the set, $RA_{I,t}$, is a set of risk aversion parameters and interactions between those parameters and the lower bin. We generate the interactions because there is no overall directional prediction for risk averse bidders as some bids are predicted to increase and some predicted to decrease. The appendix includes a longer discussion of risk aversion.

Repeating this analysis with the fitted crossover point is straightforward. We generate a new dummy variable, $LowerFittedBin$, for valuations below the fitted crossover point, $v_i = \$5.32$, and replicate Equation (4) replacing $LowerBin$ with $LowerFittedBin$, and interacting our risk aversion measures with the new bin. The results are found on the right-hand side of Table 2.

Table 2: The difference in bids across different valuations

	Difference in Bids (Small-Large)					
	Predicted Cutoff: \$8.48			Fitted Cutoff: \$5.32		
Periods:	1-15	1-15	9-15	1-15	1-15	9-15
LowerBin=1 if						
Valuation $\leq$ \$8.48	1.070***	1.079***	1.151***			
	(0.33)	(0.30)	(0.43)			
LowerFittedBin=1 if						
Valuation $\leq$ \$5.32				0.827***	0.862***	1.154***
				(0.26)	(0.24)	(0.35)
Constant	-0.994***	-0.973***	-1.192***	-0.770***	-0.765***	-1.038***
	(0.32)	(0.29)	(0.41)	(0.25)	(0.24)	(0.32)
R.A. Controls	No	Yes	Yes	No	Yes	Yes
N	888	888	417	888	888	417

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$
Standard errors clustered by subject

The difference in bids in the upper bin is captured by the constant, while the difference in bids in the lower bin is captured by the sum of the constant and the coefficient of the dummy variable for $LowerBin$. Clearly, the difference in bids in the upper bin is large in magnitude and highly statistically significant. This difference in bids is also statistically distinguishable from the same difference in the lower bin. However, the difference in bids in the lower bin is not statistically distinguishable from zero. Recall Figure 3, which graphs the predicted differences in the bidding functions. We can see that there are not large differences in the equilibrium predictions for bidders with low valuations, so it is not surprising that we failed to find a significant difference from zero for these bidders. For valuations just above the median, however, we predict large differences between bids cast in the different auction sizes, and this separation shows up strongly in both statistical tests, and graphical analysis.[16]

Finally, we consider the distribution of revenues for the two auction sizes. We split revenues according to the same theoretical and fitted crossover points as before, and report revenues from the large auction as a percentage of revenues from the small auction. Our results can be seen in Table 3. The first two columns represent either side of the theoretical crossover point, and the second two columns represent either side of the fitted crossover point.

Table 3: Mean revenue in the large auction relative to the small auction

	$v_i \leq \$8.48$	$v_i > \$8.48$	$v_i \leq \$5.32$	$v_i > \$5.32$
Large/Small	94%	114%	89%	113%

To calculate the percentage, we divided bidders using the cutoff valuation and summed all bids within a period in that interval in each auction. We then divided the aggregate bidding in the large auction by the aggregate bidding in the small auction for each interval. We report the mean of this percentage across all 45 pairs of auctions here.

On average, the general predictions of the model with respect to relative revenue hold quite well. The interval of valuations over which we predict greater revenues from the large auction do indeed have greater revenues from the large auctions, do indeed dominate, and the same is true for the small auction. These results demonstrate that on aggregate and individually the predictions of the model capture many distributional characteristics of the data.

Objective 3: *Accurately order the valuations of bidders based only on their bids.*

As indicated by Figure 1, we expect bidders in the larger auction with low valuations to pool near zero, and bidders with high valuations to pool near \$10. In the smaller auction we expect a more gradual increase of bids, which means that the predicted bids of any two bidders are less likely to be very close to each other in the small auction than in the large auction. This provides

[16] These results are also unchanged with the inclusion of all outlying bids except for the particularly extreme outlier mentioned in objective 1.

an advantage to the small auction with respect to maintaining a well ordered sorting of bidders in the presence of errors. On the other hand, bids in the large auction pool at opposite ends of the bidding space. This provides an advantage to the larger auction as it creates a greater distinction between bids by high- and low-valuation bidders.

In venturing down the path of off-equilibrium dynamic responses to bidding, it becomes clear that our model can no longer provide strong predictions, and the intuition gleaned above represents the fullest extent to which we are comfortable speculating.

We test which auction best accomplishes Objective 3 with Kendall's (1938) tau rank correlation coefficient, K-T. The K-T coefficient measures the relationship between the number of concordant pairs, C, and discordant pairs, D, in a sample with N observations. Mathematically, the K-T coefficient is

$$K = \frac{C - D}{\left(\frac{N \times (N-1)}{2}\right)}.$$

A pair of bids in a given auction is concordant if the ranking of the two bids is consistent with the ranking of the valuations of the bidders ($b_i < b_j \iff v_i < v_j$), while a pair of bids is discordant if it is inconsistent with the rankings of the bidders' valuations ($b_i > b_j \iff v_i < v_j$) [17] [18]

The K-T coefficients and their significance tests for the pooled data are reported in the first and second row of Table 4. We then split the data by period and generate the K-T coefficient for each auction in that period before running a pairwise analysis of the two auctions.[19]

[17] We can also make a correction found in Kendall (1970) that accommodates ties in bids or in valuations. The signs and significance of our results are maintained or increased under this specification.

[18] A possibly more recognizable test of rank-correlation is the Spearman (1904) coefficient. We have chosen Kendall's tau for two reasons: First, Spearman's coefficient is more heavily influenced by outliers. This will bias our results in the direction of the large auction, since there are many more outlying bids in the small auction. Second, Kendall's tau has a more straightforward interpretation.

Repeating the analysis using Spearman's coefficient returns the same signs, but with larger magnitudes.

[19] Given that we have 45 pairs of K-T statistics, we invoke the Central Limit Theorem and perform a standard t-test on the difference between the pairs.

Table 4: Measuring the Well-Ordering of Bids

Kendall-Tau Coefficient of Well-Ordering				
	Obs	Small Auction	Large Auction	Difference
Pooled†	900	0.546	0.586	-0.040**
		(0.014)	(0.012)	
Pooled⋆	888	0.549	0.591	-0.043***
		(0.014)	(0.012)	
Per Period	45	0.565	0.608	-0.042**
		(0.014)	(0.014)	

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

† *Standard errors clustered at the period level.*

⋆ *Outliers omitted, and standard errors clustered at the period level.*

Our results show that the ability to generate a perfect ordering of valuations based only on bids is compromised in both the large and small auctions, but unevenly so. Observing a properly ordered pair of bids is roughly 55% more likely than observing an improperly ordered pair of bids in the small auction, while in the large auction the percentage is closer to 59%. The difference in likelihood is highly statistically significant. Changing the unit of analysis from individual bids to means across auctions has no meaningful effect on the magnitude or significance of the coefficients.

Digging deeper into the K-T coefficients reveals some of the mechanisms that drive bids out of their well-ordered ranking. While the small auction equilibrium bidding function showed a more gradual increase in Figure 1, the larger auction showed starker contrasts between bids of high- and low-valuation bidders. These competing mechanisms are clearly visible in the data.

Consider two groups of bidders, those above the median of the distribution and those below it, call them expected winners and expected losers. Rows 1 and 3 of Table 5 report the K-T coefficients for expected winners and losers, respectively. Clearly, the coefficients within a group are much different than the pooled coefficients reported in Table 4, and indeed are relatively more favorable towards the small auction. In the second row, we maintain the interval length, but now compare sorting between expected winners and losers. This sorting strongly favors the larger auction, indicating that bidders identified the probabilistic distinction of expected outcomes more in the larger auction, causing them to separate based on the side of the median on which their valuation lies. This separation is consistent with the predicted influence of the law of large numbers. These two sorting mechanisms closely follow the intuition of our model, which demonstrates that the larger auction should see pooling among expected winners and losers, but strong separation of the two groups, while the smaller auction should see more gradual separation of all types along the distribution.

Table 5: The Well-Ordering of Bids for Expected Winners and Losers

Kendall-Tau Coefficient of Well-Ordering				
	Obs	Small Auction	Large Auction	Difference
$Valuation_i < 10$	414	0.375	0.295	0.079***
		(0.028)	(0.029)	
$5 < Valuation_i < 15$	474	0.404	0.482	-0.078***
		(0.029)	(0.021)	
$10 < Valuation_i$	474	0.334	0.345	-0.011
		(0.031)	(0.028)	

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Outliers omitted, and standard errors clustered at the period level.

Interpreting Table 4 in light of the results from Table 5 gives us insight into the relative strengths of each sorting mechanism. It appears that as the auction size grows, the superior sorting into expected winners and losers dominates the diminished sorting within expected winners and losers.

7 Discussion

Our results show that subjects are sensitive to changes in the size of their competitive environment in a way that closely follows many predictions of our model. While choices are statistically different than the Nash equilibrium, the comparative static predictions generated by our model served as useful benchmarks in our analysis.

Taking our results from the abstract environment and returning them to the environment of an undergraduate course will provide context and useful policy recommendations. We would be remiss if we failed to offer a caveat to our interpretation: our analysis takes all other inputs as held equal. That is, we abstract away from systematic correlations between the quality of inputs to the students' production function and the size of the classroom. Of course, in any discussion of policy these correlations will need to be weighed against the effects we discovered. We do not feel that the effects we uncovered trump other effects of classroom size, but by bringing them to light we will allow policy-makers to factor them into the discussion of optimal classroom size.

Recall that we supposed that administrators were restricted to fixed proportions of students receiving each grade, but were given flexibility with respect to the enrollment of the course. Since the larger auctions generated more revenue, we predict that, conditional on a relative grading scheme, larger classes will likewise generate greater effort from students. In smaller classes, the greater uncertainty surrounding the minimum effort required for a given grade may cause students to strategically lower their effort. This could be in an attempt to exploit the randomness of the environment, an effect of loss aversion, or a discouragement effect from previously receiving a low grade despite substantial investment of effort. In larger classes, however, there is less uncertainty about the minimum effort required for a given grade, and students will likely respond by increasing their effort to match this minimum required effort for their desired grade.

Aggregate effort, however, may not be the primary goal of the instructor. We demonstrated that the distribution of effort was strongly affected by the size of the contest in the direction predicted by our model. From our results, we can infer that high ability students will likely offer higher effort in larger classes, while students with lower-middle abilities may exert slightly more effort in smaller classes. Unfortunately, our data do not provide strong predictions for how the lowest ability students will respond to changes in enrollment, since they did not demonstrate any sensitivity to the size of an auction.

Another advantage of the larger auction was its ability to sort bidders by their valuation when only effort was observable. In the classroom, student abilities are unobservable and so assignments, test, and other evaluations in the course must serve as a sorting mechanism, attempting to distinguish students' abilities by students' outcomes. What we discovered suggests that the ability of a relative grading scheme to sort students could be systematically undermined in courses with low enrollment. That is, a student's ability seems to be less correlated to her relative performance in courses with fewer students.

With multiple grade levels (A,B,C,D,F, for example), each student will face greater uncertainty with regards to their relative position in low-enrollment courses. Our results indicate that this uncertainty may cause effort to diminish and weaken the correlation between ability and performance. On the other hand, with large enough enrollment, we still expect students to match the minimum required effort, but at a finer scale, now matching the minimum score required for the highest grade they can receive.[20] With a finer grading scheme, the ordering of bids will be more dependent on between-group sorting, likely giving an even stronger advantage to courses with higher enrollment. This speculation, however, needs to be verified by further studies.

8 Conclusions

We designed and executed an experiment to examine how the size of a contest will affect the effort choices by players of different types. We leveraged the paired-auction design of our experiment to focus on the difference in bids between auctions of different sizes and analyzed how that difference evolved across the type space. With this approach we were able to abstract away from the exacting question of on- or off-equilibrium bidding and focus on more general, comparative static differences in bidding across auction sizes. We discovered economically and statistically significant results that hold implications for a broad category of contest design. These results will allow mechanism designers to select the size of their contests in order to optimize with respect to

[20]For a deeper discussion of optimally selecting the fineness of a grading scheme, see Dubey and Geanakoplos (2010).

aggregate effort, the distribution of effort, and the well-ordering of players. In addition to selecting the size of a contest, there is room for further research to illuminate how adjusting other contest characteristics affect large and small contests differently.

Our equilibrium predictions correctly identified that a larger contest size would lead to higher aggregate bidding, larger bids from high-valuation bidders, and a more accurate between-group ordering of relative ability. These predictions drew attention to some undesirable effects of smaller contests on participant behavior that our results then verified. Specifically, as the size of a contest decreased, the uncertainty faced by each bidder increased, causing them to shade down their bids. Shrinking the size of a contest also limited the ability of observers to determine an accurate ordering of contestants' underlying quality, demonstrating that the ordering of bidders between groups of expected winners and losers was relatively more important than the ordering within groups, which was actually superior in the small contest.

With respect to bidding levels and revenues, holding constant the proportion of winners, a policy maker should seek to implement the largest contest if maximizing total effort, or maximizing effort among high types is the goal. The lowest types seem unaffected by the size of a contest, and offer similarly low levels of aggregate effort regardless of their environment, while bidders with lower-middle valuations exert more effort in smaller contests. An interesting topic for study could address endogenous selection of a contest's size, because it is not clear from the data the degree to which players can predict the responses of other players to a change in the contest's size.

As policy makers, administrators, and other mechanism designers continue to employ contests, this research sheds light on the effect of contest's size on the effort exerted by its participants. While this study considered only two sizes of a stylized contest, we believe it serves as a strong foundation for further research on the independent effect of a contest's size on the performance of its participants. With specific focus on the classroom application of this model, a field experiment testing grading schemes under different enrollment levels could confirm the external validity of our results.

References

Amann, E., & Leininger, W. (1996). Asymmetric All-Pay Auctions with Incomplete Information: The Two-Player Case, *Games and Economic Behavior*, 14, 1-18.

Andreoni, J., Che, Y. K., & Kim, J. (2007). Asymmetric information about rivals' types in standard auctions: An experiment. *Games and Economic Behavior*, 59(2), 240-259.

Andreoni, J. & Harbaugh, W. (2010). Unexpected Utility: Experimental Tests of Five Key Questions about Preferences over Risk. *Mimeo.*

Barut, Y., Kovenock, D., & Noussair, C., (2002). A comparison of multiple-unit all-pay and winner-pay auctions under incomplete information. *International Economic Review* 43, 675707.

Baye, M., Kovenock, D. & de-Vries, C.G. (1996). The All-Pay Auction with Complete Information. *Economic Theory*, 8, 291-305.

Baye, M.R., Kovenock, D. & de Vries, C.G. (1993). Rigging the lobbying process: an application of the all-pay auction. *American Economic Review*, 83, 289-294.

Becker, W.E., & Rosen, S. (1992). The Learning Effect of Assessment and Evaluation in High School. *Economics of Education Review* 11, 2, 107-18. EJ448 452.

Bull, C., Schotter, A. & Weigelt, K. (1987). Tournaments and Piece Rates: an Experimental Study. *Journal of Political Economy*, 95, 1-33.

Card, D. & Krueger, A. (1996). School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina, *Journal of Economic Perspectives, American Economic Association*, vol. 10(4), pages 31-50, Fall.

Carneiro, P. & Heckman, J. (2003). Human Capital Policy. NBER Working Paper No. w9495.

Davis, D. & Reilly, R. (1998). Do Many Cooks Always Spoil the Stew? An Experimental Analysis of Rent Seeking and The Role of A Strategic Buyer. *Public Choice*, 95, 89-115.

Dechenaux, Kovenock, and Sheremeta (2012). A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments. mimeo.

Dubey, P. & Geanakoplos, J. (2010). Grading Exams: 100,99,98,... or A,B,C? *Games and Economic Behavior*, 69, 72-94.

Fischbacher, U., (1999). Z-TreeZurich toolbox for ready-made economic experimentsexperimenters manual. *Mimeo.*

Gneezy, U. & Smorodinsky, R. (2006). All-Pay Auctions An Experimental Study. *Journal of Economic Behavior and Organization*, 61, 255-275.

Harbring, C. & Irlenbusch, B. (2005). Incentives in Tournaments with Endogenous Prize Selection. *Journal of Institutional and Theoretical Economics*, 127, 636-663.

Harsanyi, J. C. (1973). Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1), 1-23.

Hillman, A. & Riley, J.G. (1989). Politically contestable rents and transfers. *Economics and Politics*, 1, 17-40.

Kagel, J. H., & Levin, D. (1993). Independent private value auctions: Bidder behaviour in first-, second-and third-price auctions with varying numbers of bidders. *The Economic Journal*, 868-879.

Kendall, M.G., (1938). A new measure of rank correlation. *Biometrika*, Vol 30, 1938, 81-93.

Kendall M. G. (1970). Rank correlation methods, fourth ed., *Charles Grifn & Co.*, London.

Kokkelenberg, E.C., Dillon,M, & Christy, S.M. (2006). The Effects of Class Size on Student Grades at a Public University. Cornell University ILR School. Working Paper.

Krueger, A.O. (1974). The Political Economy of the Rent-Seeking Society. *American Economic Review*. 64, 291-303.

Lazear, E.P. & Rosen, S. (1981). Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89, 841-864.

Moldovanu, B. & Sela, A. (2001) The optimal allocation of prizes in contests. *American Economic Review*, 91, 542558.

Mosteller, F. (1995). The Tennessee Study of Class Size in the Early School Grades. *The Future of Children*, 5, 113-127.

Müller, W. & Schotter, A. (2010). Workaholics and Dropouts in Organizations. *Journal of the European Economic Association*, 8, 717-743.

Nash, J. (1951). Non-cooperative games. *The Annals of Mathematics*, 54(2), 286-295.

Noussair, C. & Silver, J. (2006). Behavior in All Pay Auctions with Incomplete Information. *Games and Economic Behavior*, 55, 189-206.

Olszewski, W. & Siegel, R. (2013). Large Contests. *Mimeo.*

Paredes, Valentina. (2012). Grading System and Student Effort. *Mimeo.*

Potters, J.C., de Vries, C.G. & van Winden, F. (1998). An Experimental Examination of Rational Rent Seeking. *European Journal of Political Economy*, 14, 783-800.

Romer, D. (1993). Do Students Go to Class? Should They?. *Journal of Economic Perspectives*, 7, 167-174.

Siegel, R. (2009). All-Pay Contests. *Econometrica*, 77(1): 71-92.

Spearman, C. (1904). The proof and measurement of correlation between two things. *American Journal of Psychology,* 15, 72101.

Tullock, G. (1967). The Welfare Costs of Tariffs, Monopolies, and Theft. *Western Economic Journal*, 5, 224-232.

Vickrey, W. (1961). Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance*, 16: 8-37.