Noname manuscript No.
(will be inserted by the editor)

Strategic Behavior and Learning In All-Pay Auctions: An Empirical Study Using Crowdsourced Data

Yoram Bachrach · Ian A. Kash · Peter Key · Joel Oren[*]

the date of receipt and acceptance should be inserted later

Abstract We analyze human behavior in crowdsourcing contests using an all-pay auction model where all participants exert effort, but only the highest bidder receives the reward. We let workers sourced from Amazon Mechanical Turk participate in an all-pay auction, and contrast the game theoretic equilibrium with the choices of the humans participants. We examine how people competing in the contest learn and adapt their bids, comparing their behavior to well-established online learning algorithms in a novel approach to quantifying the performance of humans as learners.

For the crowdsourcing contest designer, our results show that a bimodal distribution of effort should be expected, with some very high effort and some very low effort, and that humans have a tendency to overbid. Our results suggest that humans are weak learners in this setting, so it may be important to educate participants about the strategic implications of crowdsourcing contests.

Keywords All-pay auctions · crowdsourcing contests · learning

Introduction

Crowdsourcing contests are becoming an increasingly popular mechanism for solving difficult problems. One prominent example is the Netflix challenge [6], where Netflix offered a $1 million prize to the team that managed to best improve their movie

[*] The project was completed while all authors were employed by Microsoft Research, Cambridge, UK.

DeepMind London, UK E-mail: yorambac@gmail.com

Chicago, IL, USA E-mail: iankash@uic.edu

Cambridge, UK E-mail: peter.key@outlook.com

Yahoo! Research Haifa, Israel E-mail: joren@oath.com

recommender system. This proved to be a great success in generating sophisticated algorithms through a large-scale competition among research groups. Other similar contests include TopCoder and CodeChef[1] which are examples of programming challenges, and the DARPA Network Challenge [37], where participants must harness the distributed nature of the Internet for performing computationally-demanding tasks.

Participants in such crowdsourcing contests perform tasks, and a reward is only given to the highest-performing participant. Regardless of who wins the contest, the expending of resources cannot be undone, so only one participant makes a positive net profit. Such contests raise a difficult strategic question: as all participants incur the cost of their bid but only one participant wins the prize, how much effort should you exert so as to maximize your expected gain? Reasoning about this requires factoring in beliefs about your opponents' decisions.

For example, consider a programming contest such as those described above. Suppose both you and your opponent's daily salary is $\$100$, whereas the single prize is also worth $\$100$. As both you and your opponent are equally competent programmers, the winner would be the person who invested the most time working on the solution. On the one hand, spending more than a day on coding is ill-advised, as this would yield a sure loss (even if you win, you would have made more money from your normal salary). On the other hand, spending very little time is also likely to result in a loss: the other participant is likely to invest more effort, so you'd lose the contest, get no prize and waste your effort in vain. How much time should you spend on the project, given what you know about your opponent?

The study of such strategic actors is the subject of game theory. This sort of contest has been extensively studied as an "all-pay auction," an auction where a single item is being sold to the highest bidder but *all bidders pay their bid* (and not just the winner). The happiness, or utility, of a participant is then value of the prize (if it is won) minus the amount bid. While the behavior of idealized agents in such auctions is well understood, the extent to which this reflects behavior of actual participants in a crowdsourcing setting is less clear.

Our contribution

We designed an online Facebook game that simulates an all-pay auction. We used Amazon's Mechanical Turk platform to source a pool of roughly $13{,}000$ instances of the game. We first perform a "static" analysis that assumes that the behavior of players is not changing over time. To do so, we examine the distributions of the bids and the utilities. We find that people tend to have a limited set of used bids, and that people are inclined to overbid. We cluster players by their bid distributions showing people can be classified into several types, such as bidders who are prone to extreme values or those who tend to choose high bids. These results are consistent with previous (laboratory) studies of human participants in various forms of all-pay auctions which find that few players play similarly to the mixed-strategy Nash equilibrium of the game [22, 16, 32]. In particular, researchers have noted that people in such com-

[1] www.topcoder.com, www.codechef.com

petitive settings tend to overbid, but some factors such as experience and reputation reduce overbidding [29].

However, our main goal is to study the *decision processes* by which the players set their bids in each game, and to determine how they *learn from past experience*. Our key conceptual contribution is thus evaluating the quality of human learning in an all-pay auction setting, contrasting human ability to learn from recent experience with the performance of simple learning algorithms that adjust their bids based on past experience. While we find that humans do modify their bids based on experience from past interactions in the game, our results indicate that their average gains are lower that the expected gains that can be achieved by employing relatively simple learning algorithms.

We find evidence that players choose their bids depending on the outcome of the previous game: they raise their bid after losses and decrease it after wins. This behavior is suboptimal in achieving high expected monetary gains. One interpertation of our results is that people try to *beat* their opponents by aggressively outbidding them, even at very a very high cost, ignoring the bad effect this may have on their *expected payoff*.

These results indicate a learning behavior of players. However, our results show that the average player achieves a *negative utility*, significantly lower than the utility obtained by players under the Nash equilibrium (under a Nash equilibrium, players have an expected utility of 0). Many learning algorithms for agents have been studied, but the quality of learning algorithms used by humans has received less attention.

We introduce a novel methodology to evaluate humans as learners. We take their history of opponents, and compare their utility to the utility achieved by a range of learning algorithms, including multiplicative weight updating, best-response dynamics, and fictitious play. Intuitively, the level of sophistication of learning algorithm needed to match human performance can be thought of as a rough metric for the performance of human learning in this setting. Unfortunately for humans, all these heuristics obtain higher average utilities than a typical player, and significantly outperform playing the Nash equilibrium, when playing against a random human adversary. This suggests that human players are relatively weak learners in our game. The only heuristic we tested that performs at approximately on par with the human players is an extremely simple policy that increases its bid by a fixed increment after a loss and decreases by the same increment after a win. These results also suggest that human players are more interested in winning each game rather than in maximizing their gain.

In summary:

- We collect a dataset containing strategies of participants playing an all-pay auction game, modeling a crowdsourcing contest. We find that, consistent with prior work, humans do not fit models of “rational” behavior.
- We introduce a novel methodology for assessing learning performance of humans, and show that our data suggests that humans are relatively weak learners in our game.

Related Work

A mathematical characterization of the behavior of rational agents in a single-item all-pay auction is given in [5], using the game theoretic solution of a mixed-strategy Nash equilibrium. They prove that there is a unique symmetric Nash equilibrium; In the specific case of two players, this equilibrium is simple: each player chooses a bid uniformly at random from the range $[0, m]$, where m is the value of the winner's prize. These results were extended to an incomplete information setting [1, 24] and to simultaneous and incomplete information settings [21, 13]. These papers point out the connection between all-pay auctions and crowdsourcing projects, where the efforts exerted by the agents can be thought of as bids. Further, they investigate equilibrium behavior as a function of the incentives (rewards). However, these models are quite different from the simple case examined in [5, 19]: these are multi-item auctions, where items correspond to the different tasks where each player has individual valuations for each task. Further, each player's action is a two-stage process, in which the player first selects a task to participate in, followed by submitting a bid (performing a task). An extensive literate exists on the design and analysis of contests more broadly [38].

Previous work studies sequential all-pay auctions on Taskcn.com [28, 40], a programming crowdsourcing site. However, the particular auctions studied there are different: the number of participants in such auctions is much larger, and the sequential auction formulation allows players to strategically defer their participation.

Various papers have examiend human behavior in all-pay auctions [22, 20] (see the recent survey [12] for a broad discussion of such work). Gneezy and Smorodinsky ([22]) empirically analyze all-pay auctions with agent *groups*, of sizes $4, 8,$ and 12, where each group participated in ten auction rounds. They study the bid distribution, finding evidence of players deviating from the theoretical equilibrium behavior, and showing that players bid more than expected under the game's symmetric Nash equilibrium. Roughly speaking, they find an overall tendency of overbidding on the players' part, resulting a consistent positive surplus, which diminishes with the number of steps. Overbidding in contests has also been reported in various other contest settings [33, 34]. [2] Towards the end of their paper, Gneezy and Smorodinsky suggest a possible explanation for players' behaviors in all-pay auctions, by proposing a simple two-stage process of first deciding on whether to participate in such an auction, and then deciding on a bid, without the use of any learning procedure. Ernst et al. [16] investigate the effect of maintaining fixed groups of two and three players on the bid distributions, over the course of ten rounds, relative to groups that were randomly set anew in each round. They find evidence of *collusion* in two-player games where the pairs of players remain unchanged, and find a bimodal distribution over the bids, where players tend to bid either very high or very low values, which again, stands in stark contrast to the bid distribution under the Nash equilibrium. Instead, they pro-

[2] Empirical results do not always exhibit overbidding. Some studies examining all-pay auctions between two players find no overbidding [32, 16], while studies examining auctions between more than two players find significant overbidding [22, 11] (recent work on this contains a more complete discussion [20]). Further, the degree of overbidding depends on the specifics of the contest and domain [7, 12, 19, 26, 25, 27].

pose an alternative explanation to their findings, by fitting their data to Kahneman and Tversky's prospect theory.

Potters et al. [32] study a rent-seeking problem, and define dissipation as the ratio of expected expenditure to the value of the rent (the prize). One of their two scenarios corresponds to all-pay auctions with two players, for 30 rounds, with anonymous players. Although the player's bids are not strictly uniform, they find no statistical evidence for overdisspiation (i.e. for overbidding). Lugovskyy et al. [29] conducted lab experiments with 144 participants, also focus on overdissipation, but arrive at different conclusions, and argue that overbidding, both individual and in aggregate, is a robust phenomenon. In contrast to these papers studying *static* behavior, we focus on how agents *learn and adapt*.

All of the above experimental studies mainly postulate underlying largely static model, through which the players make their decisions. Although they all acknowledge a significant deviation from the theoretical models, none of them consider a more dynamic, learning process, through which the agents "react" to the observed opponent bids. This is especially true in anonymous games, where the agents do not know the identity of their opponents.

We study the reasoning processes of agents. Similarly to our results, many other papers find that players choose their bids depending on the outcome of the previous game [35, 33, 34, 10, 31, 17, 30]. A similar approach to studying bidding reasoning was used to investigate Penny auctions [39, 8], which are very different from our setting. However, these papers also show a significant deviation from theoretical predictions and provide explanatory work on the observed bid sequences. Roth and Erev ([15]) studied the explanatory value of models of reinforcement learning using a game with only two strategies. As opposed to our findings, they show that strategies like fictitious play and regret minimization explain the agents gameplay reasonably well (possibly, humans could use advanced learning strategies due to the simpler game setup).

A Game Theoretic Analysis of All-Pay Auctions

We contrast our findings regarding human behavior in all-pay auctions with the mathematical solution of the game, so we first discuss this solution. In our setting, there is a set $N = \{1, \ldots, n\}$ of agents with a common value m for a single-item. The auction takes place in a single round, in which each agent $i \in N$ submits a single bid b_i. The highest bidder receives the auctioned item, while everyone pays their bid, regardless of who wins. Thus, the highest bidding player, i, wins the prize and thus obtains a utility of $m - b_i$, and each other player j gets a utility of $-b_j$.

A prominent mathematical "solution" of such games is the *mixed Nash equilibrium*, a profile of mixed strategies for each player which are mutual best-responses. In a mixed Nash equilibrium each player has a distribution over bids, such that no player can change their bid so as to get a higher expected utility (even knowing the distributions of the other players).

It is easy to show that the game doesn't admit a *deterministic* Nash equilibrium, where each player has a single bid (rather than a distribution over bids), and no player can unilaterally change their bid to improve their utility. For example, in a two player game, given her opponent's bid, a player can bid slightly higher than her opponent's, guaranteeing a win.

The analysis in [5] shows that in the simple all-pay auction, there is a unique symmetric mixed Nash equilibrium, in which each player samples a value from a common distribution over the range $[0, m]$; Specifically, in this equilibrium each player chooses a bid at random from the distribution with the following cumulative distribution function: $F(x) = (\frac{x}{m})^{\frac{1}{n-1}}$.

For two-player auctions, this reduces to choosing a bid from the uniform distribution over the range $[0, m]$, yielding an expected revenue of m to the auctioneer and zero utility to the players.

Methodology

We used the Amazon Mechanical Turk crowdsourcing marketplace to let human participants play against one another in all-pay auctions, which we interpret as a simple model of the decisions made in crowdsourcing contests. To this end, we constructed a two-player Facebook game called Doubloon Dash.

Although crowdsourcing contests may have more than two participants, we chose to focus on all-pay auctions with two players, for the sake of simplicity of the game and the analysis of learning. Each player was asked to play at least 30 games, in which they were matched against an opponent whose identity was unknown to them[3], and the outcome of each game was determined based on the simple all-pay auction as previously described, with bids between 0 and $m = 10{,}000$ doubloons. We recruited players in two sessions, each of which lasted a few hours from the time the tasks were posted until they were completed.

On top of their base payment b, players were paid based on performance: a player with average utility of u was awarded an amount of $\frac{u}{9{,}999} + 1$; so that bonuses ranged from \$0, for players who made the lowest possible utility in every game, to \$2 for players who got the maximal possible payoff in every game.

Empirical Analysis of Bids and Utilities

In this section we discuss our dataset and perform some basic analysis showing that, consistent with previous experiments with all-pay auctions, participants do not appear to follow the sort of strategy predicted by game theory.

Dataset

Our dataset consists of $12{,}899$ games by 518 players. The instructions for the task required players to play at least 30 games, and 178 **incomplete** players failed to do

[3] We also allowed players to play against a specific friend, but such games were not used for the analysis in this paper.

so. To discourage non-strategic play, players were asked not to make extensive use of “simple” bids such as 0, 1, 1000, 9999 or 10,000 (these “suspicious” bids were compiled after pilot rounds, where we observed a tendency to favor them). We labeled a player a **spammer** if they used the simple bids in more than a 25% of their games, and there were 85 such players. [4]

Furthermore, some participants attempted to *cheat* by using fictitious identities. Such **manipulators** used a fictitious Facebook profile designated a *“losing“* false identity (or a Sybil [14]), through which they consistently bid low values. The other player, controlled by the same individual, played games immediately after, hoping to be matched against the losing identity, and “win” by bidding slightly higher values, guaranteeing a very high total utility. For example, consider a player who creates a “loser Sybil”; If the loser bids 2 and the true identity bids 3, the true identity wins $10000 - 3 = 9777$ in every game against the Sybil.

This manipulation is effective in generating high utility for the manipulator, but it is *simple to identify* such players. The “losing” and “winning” identities enter their strategies at roughly the same time, are matched often against each other, and one always wins while the other always loses. We found 49 such manipulators, and they were often successful: nineteen of the top twenty players were manipulators.

The fact that people exploited the rules of our system is perhaps an unsurprising observation for crowdsourcing. We note that we did not see evidence of players engaged in *reciprocal* behavior where they could collude achieve high scores by taking turns as a low-bidding losing player. [5]

In order to focus on our interest in how people learn to participate in crowdsourcing contests, we exclude data from incomplete players, spammers and manipulators. A total of 6,383 out of the 11,327 original games had at least one player who was not in any of these categories. We call these 206 players **non-manipulators** and henceforth we restrict ourselves to studying only the performance of such players. The average number of games played by these players was 74.8. About a third of the players (67) played at least 70 games, and 31 of them played at least 100 games.

As they attempt to exploit the rules of the system, spammers and manipulators are arguably more sophisticated than non-manipulators. Thus, a limitation of our results is that we may underestimate the performance of the average participant because we have excluded those who would have been high performers. However, our data still reflects a substantial number of participants from a real crowdsourcing system, and thus we believe it is at least representative of a significant fraction of crowdsourcing participants. Finally, we are aware of the fact that our definition for ruling out spammers rules out some bids that have proven to perform well against the empirical distribution of bids, discussed in Section . We have thus repeated the key parts of

[4] We note that players who follow the mixed-strategy of the Nash equilibrium are extremely unlikely to be labeled as spammers. The symmetric mixed-strategy under the Nash equilibrium is choosing a bid uniformly at random over the range, making each possible strategy have a very small probability. In particular, for a player who selects bids uniformly at random, the probability of selecting the specific precluded bids in over 25% of the games is very low. This means that our spammer-detection rule is very unlikely to have a “false-positive”, and mistakenly labeling a player who is using the Nash mixed-strategy as a spammer.

[5] Notably, such collusive strategies are predicted by a different game theoretic analysis of this as a repeated game setting, showing another way human behavior differs from idealized mathematical models in crowdsourcing contests.

our analysis *with the spammers included*. This results of that analysis (which can be found in the appendix) are very similar to the results we report in Section and Section , so we believe our conclusions are somewhat robust to the choice of mechanism for dealing with spammers and manipulators.

Average Revenue And Bid Distribution

Similarly to [22], we find that most players tend to overbid, yielding the auctioneer a higher revenue than what would be achieved in equilibrium. The average revenue (sum of bids) in games with at least one non-manipulator and no spammers was $13{,}730$. We ranked the bids by the number of their occurrences. Ignoring spammers and manipulators, the top-ten bids were (9000, 8000, 7000, 6000, 5000, 9900, 9999, 4000, 3000, 9990) with respective occurrence counts (1183, 932, 606, 589, 537, 510, 497, 382, 307, 286). The average number of distinct bids per player is 34.9.[6] It is thus apparent that most players have a preference for a relatively small set of bid values, most notably multiples of $1{,}000$, which we refer to as “focal points.” This is strengthened after considering the distribution of the unit digits of the bids, which shows that 75.7 percent of the bids had 0 as their unit digit, followed by 8.52 percent and 7.18 percent for the digits 1, and 2, respectively.

We analyzed the average distribution of bids of non-manipulators . For each player j, we compute a vector $\mathbf{f^{(j)}} = (f_1^{(j)}, \ldots, f_m^{(j)})$, such that $f_i^{(j)}$ denotes the fraction of occurrences of bids of value i. Normalizing each player’s frequency vector prevents the distortion of the analysis due to players who played significantly more games than others. We define the average cumulative distribution as follows: $\bar{\mathbf{f}} = \frac{1}{|j:j \text{ is not a manipulator}|} \sum_{j:j \text{ is not a manipulator}} \mathbf{f^{(j)}}$.

This bid distribution is given in Figure 1.

A sizeable portion is concentrated in the higher values, indicating an overall tendency to overbid. However, a substantial amount of bids are also concentrated in the low range $[0, 1000]$. This shows that people focus their attention on extreme points, perhaps viewing them as “safe bets”. Also, the bids tend to concentrate around multiples of 1000. This may suggest a reasoning process, in which some players initially narrow down the strategy space to only a few candidate bids.

We analyzed the average bid, over the course of 100 games. We grouped games into periods of length 4, so that period t’s average bid corresponds to bids submitted with games $4 \cdot t - 3, \cdots, 4 \cdot t$. The average bid over time is shown in Figure 2, where the bars for each step show the standard error in the corresponding set of bids. A striking aspect is the sharp initial increase in the average bid. Further, it seems that the average bids, despite fluctuations, tend to concentrate around values that are well above $6{,}000$ doubloons, reflecting players’ persistent tendency to overbid. From the perspective of the organizer of a crowdsourcing contest, this may actually be a good thing because it means that participants are exerting higher effort than theory predicts.

To get a refined view of focal points, we clustered players’ bid distributions (binned into ten equal size bins, $([0, \ldots, 1000], [1001, \ldots, 2000], \ldots, [9001, \ldots, 10000])$

[6] Note that the uniform distribution over the bids gives an expected number of distinct bids of 70. Shortly, we provide two more precise comparisons with the uniform distribution.

Fig. 1: Average player cumulative distribution of bids.

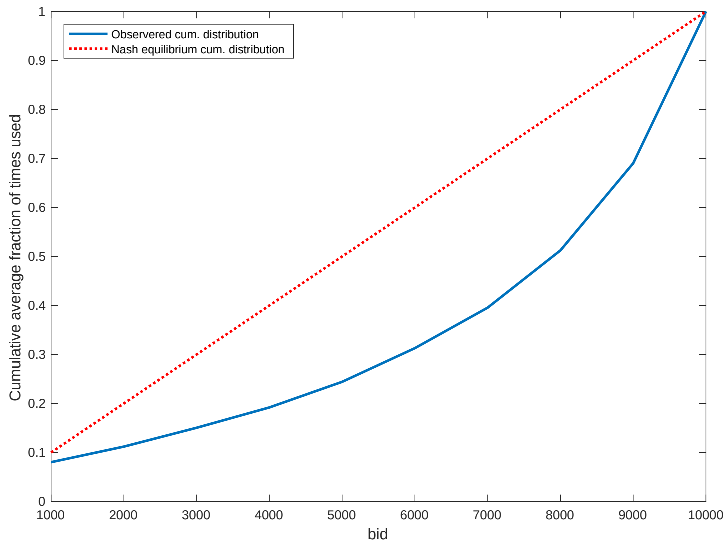
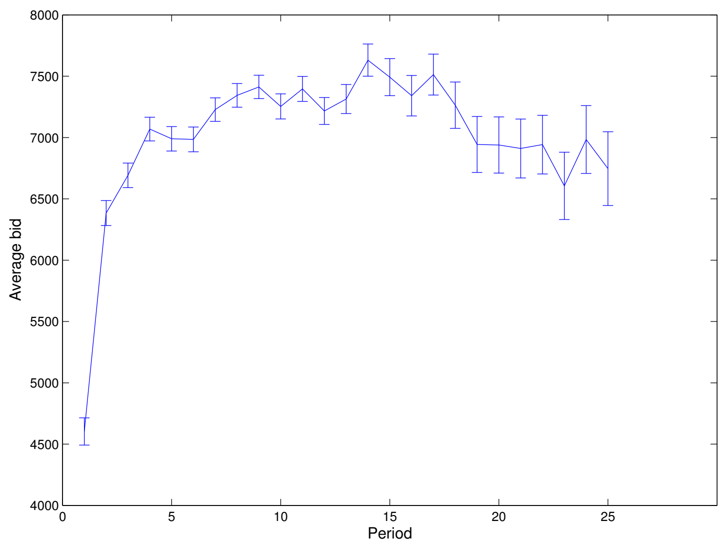


Fig. 2: Average bid per time period.



using the k-means algorithm. Intuitively, each cluster represents a group of players who bids in each bin with similar frequencies. Figures 3 and 4 show the bid histograms for two such clusters, selected to illustrate the variety of focal points. The x-axis is the bin number, and the y-axis is the number of times each bid range (bin) was played, and the histogram of each player is represented as a line. The figures show that despite the overall distribution observed in Figure 1, there was a significant

Fig. 3: Cluster 1 (47 players) player bid distributions

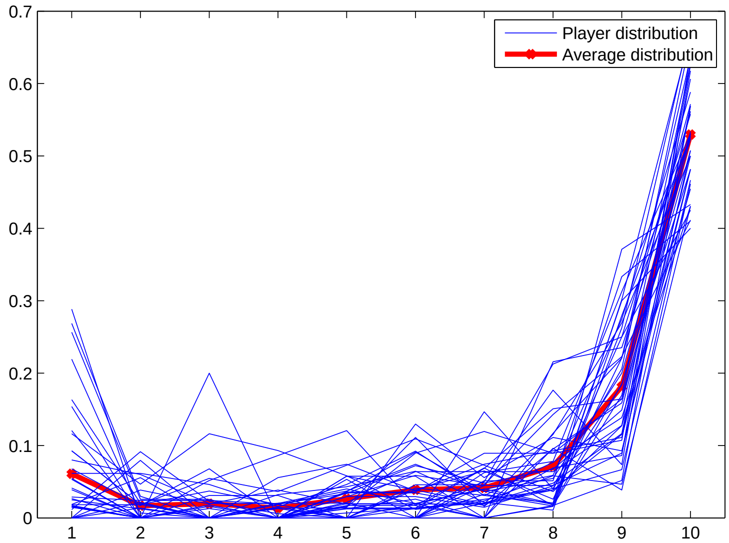
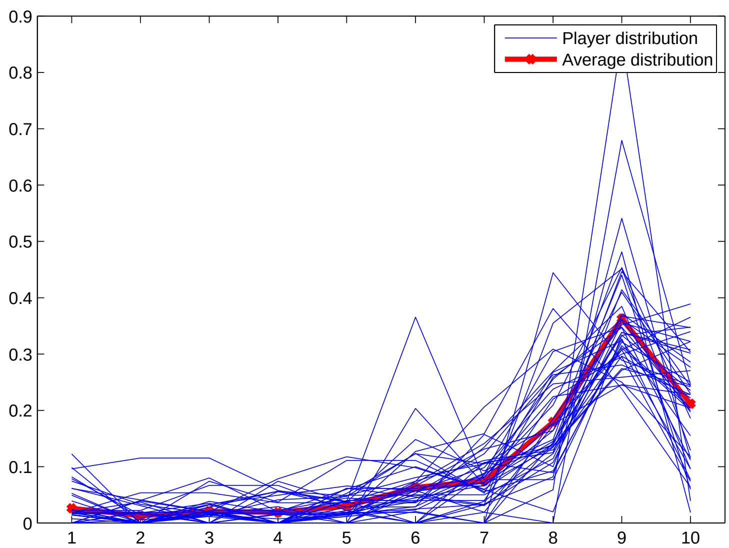


Fig. 4: Cluster 2 (45 players) player bid distributions



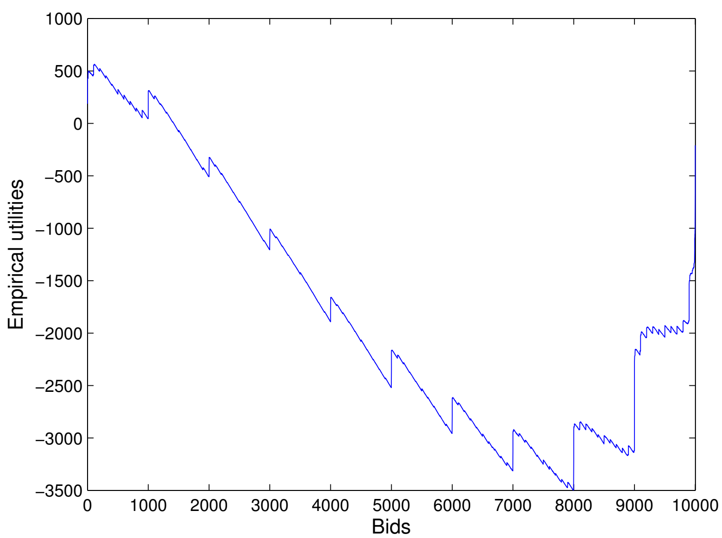
variance in the agents' focal points. For example, the players in Figure 3 follow a bimodal distribution that resembles the overall distribution of Figure 1, but players in Figure 4 show a monotonically increasing tendency to bid in higher values, save for the highest bin.

Our results show that very few players behaved consistently with the bids under the Nash equilibrium solution. To show this, we used the χ^2 Goodness-of-Fit test, examining the similarity of the bid distribution of a player with the bid distribution of the Nash equilibrium. To deal with the sparseness of the bid lists, we divided the $10,000$ bids into 10 equal-size bins. Using a statistical significance level of $p = 0.05$, there were only 9 players whose bid distribution was close enough that we could not reject the hypothesis it matched the Nash equilibrium distribution (i.e. uniform) with at least that significance level. Thus, the test shows that almost all players had a bid distribution that deviates from the Nash equilibrium behavior of bidding uniformly at random (with strong statistical significance). Since players may need to gain experience before adopting ideal play, we repeated this test using only the last 30 bids submitted by each player. This resulted in an increase to 29 players for whom we could not reject the uniform strategy behavior, which we believe is primarily due to the reduced amount of data rather than an actual change in behavior. Furthermore, this is still as substantial minority of players (less than 15%).

Empirical Analysis of the Utilities

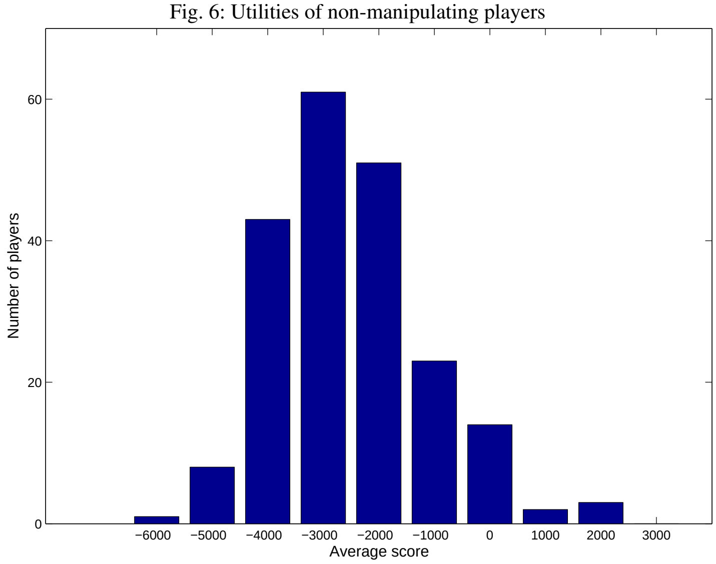
Given the empirical distribution of the bids submitted by non-manipulators, we computed the average utility of each bid. The optimal bid (i.e., the best-response to the empirical distribution) is 112, which yields an average utility of 564. A plot of the distribution of the average utility as a function of a fixed bid is given in Figure 5.

Fig. 5: Average utility against the empirical bid distribution



The figure shows that low bids, in the range $[0, 1000]$, tend to do well against the bids of our human participants, and yield an expected positive payoff. We note that most non-manipulators did not make positive average payoffs: the average utilities

Fig. 6: Utilities of non-manipulating players



ranged between $-5{,}122$ and $2{,}393$, with median $-2{,}274$ and average $-2{,}010$. This contrasts the expected payoff of 0 under Nash equilibrium bids. The distribution of the average utilities is given in Figure 6.

Note that because the maximum possible bonus payment was 2 dollars, the average payoff of $-2k$ doubloons translated into only a 20 cent loss, relative to the expected utility in the Nash equilibrium. This may be small enough that some players did not consider it significant, but is not trivial relative to common Mechanical Turk payments.

Learning From Experience

We now examine how players adapt their strategies and learn from experience. We show that although players do modify their behavior based on their experience in previous games, most players seem to follow a very simple learning heuristic: they increase their bids after losing and decrease them after winning. We contrast this with known simple machine learning approaches, and show that such algorithms can easily outperform humans in our all-pay auction game.

Effects of game outcomes

The results in the Section indicate that the Nash equilibrium mixed-strategy is a poor predictor of human behavior in all-pay auctions. Many settings exhibit a discreprency between equilibrium strategies and actual human behavior [7], with some researchers focusing on human behavior as a form of bounded-rationaliy (see [9] for a detailed discussion). For instance, in some settings, given a recent interaction with another player, it is easy to identify an alternative strategy that would have yielded a better ourcome. Clearly, in a future interaction an opponent may also adjust their behavior based on their own experience from the past interaction. Varoius models have been suggested regarding how humans reason in such settings. One of the simplest is some version of best or better reply dynamics, where players attempt to respond to the most recent strategy they have seen from their opponent. More generally, this logic is captured by models such as k-level iterative reasoning [36, 2, 3]

Consider the case of two players in an all-pay auction, who are about to bid b_1 and b_2 respectively, and assume without loss of generality that $b_1 > b_2$. Can player 1 do any better than her current bid, assuming that player 2 would bid b_2? Clearly she can lower her bid to b'_1 such that $b_2 < b'_1 < b_1$, and this would still result her still winning, but paying a lower amount. Similarly, player 2 is currently not winning anything, and can increase her bid to $b'_2 > b_1$ (but less than the maximal value $10,000$), and as a result win the auction and switch from a negative utility to a positive utility.

The above discussion illustrates how players might reason about changing their bids following a previous interaction. Assuming that they are going to face the same bid as in the past interaction (i.e. not taking into account that other players are also learning), a player who won the auction would lower their bid, and a player who lost the auction would raise their bid. [8] Thus, our hypothesis regarding players is that they react differently to wins and losses (in particular, in the above discussion players are more prone to increase their bid after losses and to decrease it after wins).

An alternative hypothesis is that players' bids are independently and identically distributed random variables (such as would be the case if they were simply choosing their bids from the Nash equiblibrium), resulting in bids following a win taking the same distribution as bids following a loss. To examine this issue, we performed an initial binning of the bids into ten equal size bins, and for each bin we examined the bids falling in that bin. We partitioned them into two groups: bids submitted by a player after winning the game, and bids submitted after losing a game. We observed that bids submitted after losing a game tend to be higher than those submitted after winning a game (i.e. they tend to fall into a higher bin). Let T_j be the total number of games played by player j, and $S^{(j)} = \{b_1^{(j)}, \dots, b_{T_j}^{(j)}\}$ denote the bins of the bids j has submitted ($b_i^{(j)} \in \{1, \dots, 10\}$, $\forall i = 1, \dots, T_j$). For each player j, let $W^{(j)} = \{i :$ player j wins in game $i\}$, and $L^{(j)} = \{1, \dots, T_i\} \setminus W^{(j)}$. We partition an agent's set of bids based on both the the outcome of the previous game, and the bins to which their previous bids belonged to: $A_\ell^{(j)} = \{b_t^{(j)} : (t-1) \in W^{(j)}$ and $b_{t-1}^{(j)} = \ell\}$, and similarly $B_\ell^{(j)} = \{b_t^{(j)} : (t-1) \in L^{(j)}$ and $b_{t-1}^{(j)} = \ell\}$, for $\ell = 1, \dots, 10$. By separating the subsequent bins indices based on their predecessors, we account for

[7] In particular, such settings include auctions, especially when the winner is simply the highest bidder, rather than ones where the probability of winning is higher for the highest bidder [11], and Blotto games [23, 10].

[8] A player who lost the auction can also lower their bid in an attempt to improve the negative utility they incur, but achieving a *positive* utility requires increasing the bids so as to win the auction.

cases where the preceding bids are likely to lead to similar responses, regardless of the outcome of the previous games. For instance, a bin 1 bid is likely to be followed by a higher bin bid, regardless of whether or not the game was won or lost. Lastly, we let $A_\ell = \bigcup_j A_\ell^{(j)}$, $B_\ell = \bigcup_j B_\ell^{(j)}$, for $\ell = 1, \ldots, 10$.

Our focus in the paper is on the learning process of people in the all-pay environment, which violate the premises of a static, non-learning, behaviour assumption taken by many statistical tests, such as the Mann-Whitney U-test. However, we do note that for the vast majority of bins in which a player's bid occurs in round i, following a loss the bid for round $i+1$ tends to fall in a higher bin, and following a win the next bid tends to fall in a lower bin. In other words, no matter what a player has bid in a previous round, they are likely to increase it in the next round if they lose, and decrease it if they win.

The result above indicates that the data is not consistent with players behaving the same after a win or a loss. However, we do not claim that players only take into account the result of the one last interaction they've experienced. To the contrary, we devote a section below to investigating various models of how players learn in all-pay auctions (not just from the single last interaction). Further, we note that the analysis above checked for the effects on the entire set of player bids, but does not characterize the behavior of *individual* players. To test individual player effects, we investigated the change in bid value following either a loss or a win. First, following loss (resp. won) games, we found that, on average, players submit an identical bid following a win (resp. loss), 13.6% (resp. 16.6%) of the time. This shows that players mostly tend to change their bids from turn to turn. Next, conditional on a change in the bid relative to the previous one, we measure the number of players who lower their bids in more than $2/3$ of their games played after winning. We found 93 players matching this criterion (only a single player *raised* his bids in more than two-thirds of these won games). Similarly, 135 players raised their bids in at least a $2/3$ of their games played after losing (conditioned on a change in bid) and only two players lower their bids in more than two thirds of such lost games). A natural interpretation is straightforward: a player may infer that a losing bid has a higher probability of losing, which causes him to increase his bid. Experiencing a win may encourage a player to try to lower bids, in the hope of still winning while keeping more money.

On the Additional Value of Knowing the Bids in Previous Rounds

Our analysis above indicates that knowing whether a player won or lost in the previous round is predictive of whether they choose a higher bid for this round. Consider a player x who had participated in a game with player y at round $t-1$, and denote by W_{t-1} the variable indicating whether x won or lost (i.e. $W_{t-1} = 1$ if x won, and $W_{t-1} = 0$ otherwise). Denote by x^{bid}_{t-1} the bid that x had used, and by y^{bid}_{t-1} the bid that y had used. Further, denote by I_t the variable indicating whether x increases their bid between round $t-1$ and round t, i.e. $I_t = 1$ if $x^{bid}_t > x^{bid}_{t-1}$. Our analysis shows that a loss is predictive of a bid increase (i.e. W_{t-1} is correlated with I_t). However, do players base their decision regarding raising or lowering the bid on whether they won the previous round (W_{t-1}), or do they also take the bids used in that round ($x^{bid}_{t-1}, y^{bid}_{t-1}$) into consideration?

One could in principle answer the above question by conducting a controlled experiment or by asking players to explain the reasoning behind their chosen bids across rounds, but these are very expensive to run. As an alternative, we may ask whether the actual bids in the previous round are more predictive of a bid increase than only knowing whether the player won or lost the previous round. Clearly, knowing the exact bids used in a round ($x^{bid}_{t-1}, y^{bid}_{t-1}$) is more informative than knowing who won the round (W_{t-1}) — the winner variable W_{t-1} is a deterministic function of the bid variables $x^{bid}_{t-1}, y^{bid}_{t-1}$, since $W_{t-1} = 1$ if $x^{bid}_{t-1} > y^{bid}_{t-1}$ and $W_{t-1} = 0$ otherwise. However, does knowing the previous bids ($x^{bid}_{t-1} > y^{bid}_{t-1}$) allow us to predict I_t (whether player x would raise their bid) better than we could when we only knew whether the player won or lost the previous round (W_{t-1})?

We examine the above question by training logistic regression models to predict bid increases (the variable I_t). The baseline model uses W_{t-1} as the sole feature, whereas the second model uses the feature W_{t-1} *as well as* the features $x^{bid}_{t-1}, y^{bid}_{t-1}$. We use the R^2 as our measure for the quality of the model's predictions. We train the models on the data from all the players across all rounds, and apply a 10-fold cross validation. The baseline model, which only uses W_{t-1} as a predictor, achieves $R^2 = 0.35$, and the model which also uses the bids $x^{bid}_{t-1}, y^{bid}_{t-1}$ achieve $R^2 = 0.33$. In other words, using the bids as additional features does not increase the predictive performance. [9]

These results indicate that knowing the actual previous bids does not increase predictive power beyond knowing only whether the player previously won or lost, at least for a logistic regression model. We believe that this provides some evidence that our ability to predict whether a player is likely to increase their bid in the next rounds stems mostly from knowing whether they won or lost the previous auction (at least for simple linear models).

We note that while we grounded the intuition behind our analysis in players lowering bids after *winning* and raising them after *losing*, one can also state them in a way that relates to imitating the opponent or choosing bids that are closer to those chosen by the opponent. In other words, a different conjecture is that players tend to increase bids after observing higher bid by the opponent and decrease bids after observing lower bid by the opponent. As the winner in the auction is the player with the higher bid and the loser is the player with the lower bid, Choosing a bid closer to the bid chosen by the opponent would result in the winner lowering their next bid, and the loser choosing a higher bid. Thus the data is consistent with such "imitation based learning", and further research is needed to determine the reasoning behind the participants' choice of bids. For instance, future research could present participants with questionnaires regarding why they chose the bids they chosen in various auction rounds.

[9] We have also tried using only the bid difference $d_{t-1} = x^{bid}_{t-1} - y^{bid}_{t-1}$ as a feature (i.e. using the two features W_{t-1}, d_{t-1}, which achieves $R^2 = 0.34$, still lower than the baseline model.

Limited and perfect recall response models

Our results show that players tend to increase bids after losing and decrease after winning. This is a plausible decision process, but it is also possible that players use a more complex learning dynamic that results in this behavior as a side effect. A natural more complex model of adaptive behavior is one where players make bids based on a (possibly limited-recall) view of previous opponent bids. Such models have a long history in the study of learning in games [41]. However, we show that the players' behavior is consistent only with the simple adaptive behavior model, not with the more complex one.

For a player j, let $S^{(j)} = \{b_1^{(j)}, \ldots, b_{T_j}^{(j)}\}$ be the set of T_j submitted bids. Let $O^{(j)} = \{o_1^{(j)}, \ldots, o_{T_j}^{(j)}\}$ be the set of corresponding opponent bids in each of player j's games. For a window length d, let α_{ji}^d be the fraction of the previous d opponent bids defeated by player j's i'th bid: $\alpha_{ji}^d = \frac{1}{T_i}\{o_{i'}^{(j)} : o_{i'}^{(j)} < b_i^{(j)} \text{ and } i - d \leq i' < i\}$. Let $\overline{\alpha}_j^d$ be the mean value of the α_{ji}^d values. Testing for $d = 1, 2, 3, 4$ and $d = \infty$, i.e., perfect recall, player j is considered a d-level responder, if $\overline{\alpha}_j^d > \theta$, for a given threshold value θ.

Checking the above criterion for all non-manipulators using a threshold of $\theta = 0.6$, accounts for 115 of the players (over half the non-manipulators). Increasing θ to 0.7 results in a decrease in the number of the players to sixty (about a third). Note however, that this does not mean that even these players are following such a strategy. Random play would satisfy $\theta = 0.5$, so even a slight bias would be enough to explain these results. Indeed, the adaptive behavior discussed in the previous section provides exactly such a bias.

Furthermore, this heuristic does not take into account the average utilities, with respect to the d opponent bids: it assumes players simply play in order to win, not necessarily optimizing their bids in order to increase their resulting revenue. To strengthen this point, we measure each agent's average utility against the previous d opponent bids. An agent's bid is a p-approximate best-response if his average utility against the previous d opponent bids is at least a p fraction of the optimal average such utility. Agent i is considered a p-approximate best-responder if at least a fraction θ of his bids are p-approximate best-responses. With a mild threshold $\theta = 0.6$, and checking over $d = \infty, 1, 2, 3, 4$, we found that only twelve players were 0.3-approximate best-responders (ten of which previously classified as d-level responders). Note that $d = 1$ corresponds to approximate best-response dynamics and $d = \infty$ corresponds to approximate fictitious play, both well-known simple learning heuristics. Thus, while player behavior appears consistent with simple updates with the goal of winning at reasonable cost, it does not appear to be consistent with simple learning dynamics that assume players optimize their utility.

Learning Algorithms

We showed that players adapt their play based on their histories, but their behavior does not appear to be consistent with standard simple learning dynamics from the literature such as best-response dynamics or fictitious play. We now examine how players *perform* as learners. We introduce a novel methodology to do so. We showed that the average player score was roughly -2000. This gives us *benchmark* to compare the performance of humans as learner, against *standard learning algorithms* from the literature. We show that humans perform significantly worse than all but the weakest learning algorithms. Thus, not only do players not play equilibrium strategies or use standard learning heuristics, they also do not effectively improve their performance over time.

Our tests are performed by running the learning algorithms on the sequence of opponent bids observed by a non-manipulating player (we report averages based on a sample of 100 players). We note that most of these players played against many different opponents, so we believe introducing an agent based on a learning algorithm would not have significantly affected the opponents' bids (as each would only encounter the agent very few times). We further note that simulating how various approaches perform on small size sample of 100 random player bids avoids the risk of overfitting: each algorithm may only update its bids by observing a tiny fraction of the entire bid population, similarly what the human participants do as they adjust their bids in our experiment. In other words, agents are never exposed or trained on the entire bid distribution (or entire human dataset), but rather the data available to them is the same small sample data available to a single human participant.

Multiplicative Weights

We begin with the celebrated Multiplicative Weights update (MWU) algorithm for on-line learning [18], which works as follows (see [4]). Starting from uniform weights $w_i(1) = 1$ for every bin i, on each step t the algorithm selects a bid from the set of candidate bids $C = \{\frac{i \cdot 10{,}000}{B}\}_{i=1}^{B}$, with probability proportional to its current weight $(w(t))$. Given the observed opponent bid, compute the normalized utilities of the candidate bids: $\{u_i\}_{i=1}^{B}$ (note the possible payoffs are in the range $[-(m-1), m-1]$), and update the weights, so that the weights of high-payoff and low-payoff bids increases and decreases accordingly. Specifically, if $u_i \leq 0$, set $w_i(t+1) = w_i(t) \cdot (1-\lambda)^{-u_i}$, and otherwise $w_i(t+1) = w_i(t) \cdot (1+\lambda)^{u_i}$. The amount by which weights are updated, λ, is called the "learning rate".

For each sampled bid sequence, we tested the algorithm with varying values of the learning rate λ and number of candidate bids. We reran the algorithm for each sequence 100 times, to minimize sampling errors.

We ran the algorithm for $B = 10$ and $B = 100$, while increasing λ in increments of 0.05. The mean utility for each combination of B and λ is displayed in Table 1 (the largest standard error was 12.21). As the results demonstrate, even with a very limited set of bids, by using a high learning rate the algorithm obtains a slightly positive average payoff, considerably higher than the mean utility of human players. Moreover, by increasing B to 100, and setting $\lambda = 0.9$, the algorithm is able to make an even higher payoff, on average. Thus, the performance of MWU is much better than that of a typical player. We expect that this would be also true of other established "good" learning algorithms. Note that for $\lambda = 0$, the algorithm plays

Table 1: Mean payoffs for varying B and λ

λ	0	0.05	0.1	0.15	0.2
$B = 10$	-1577	-1194	-871	-647	-472
$B = 100$	-1758	-1374	-1048	-800	-591
λ	0.25	0.3	0.35	0.4	0.45
$B = 10$	-342	-243	-163	-95	-48
$B = 100$	-425	-300	-183	-101	-21
λ	0.5	0.55	0.6	0.65	0.7
$B = 10$	3	33	61	80	104
$B = 100$	44	98	149	194	222
λ	0.75	0.8	0.85	0.9	0.95
$B = 10$	121	133	141	146	142
$B = 100$	262	292	316	340	350

the (discretized) Nash equilibrium strategy, which is poor overall given the human tendency to overbid, but still does better than the average human.

Limited-recall based responses

Previously, we saw that players tended to play in such a way that they beat their previous opponent(s), but were not best responding. We now test how well they would have done had they executed either of these strategies consistently. This represents a simpler form of learning than MWU. In both cases, we let the set of candidate bids be $C = \{1{,}000 \cdot i\}_{i=1}^{10}$. For a window length d define the following two strategies:

- *Limited-recall based minimal bid (LRMB):* For a fraction $\theta = 0.75$, select the *minimal* bid that wins against at least a θ-fraction of the last d opponent bids.
- *Limited-recall best response (LRBR):* Submit the optimal bid (on average) against the previous d opponent bids. Testing this heuristic on each non-manipulator's sequence of opponent bids.

The pseudo-code for the LRMB and LRBR is given as Algorithm 1 and Algorithm 2.

Input: m – reward, B – number of bins, T – number of games, θ – fraction, d – window length
Candidate bids: $C = \{\frac{m}{B} \cdot i\}_{i=1}^{B}$.
Step $t = 1$: select a random bid $b_1 \in_R C$.
foreach *step* $t = 2$ *to* T **do**
  $\ell_1 = \max\{1, t-d\}$
  $\ell_2 = t - 1$
  Let $O_t = \{o_{\ell_1}, \ldots, \ldots, o_{\ell_2}\}$ be the previously observed opponent bids.
  Set: $b_t = \arg\min_{b \in C}\{\frac{|\{o \in O_t : b > o\}|}{|O_t|} \geq \theta\}$
end

Algorithm 1: Limited Recall, Minimal Bid (LRMB)

Table 2 contains the average payoffs obtained by the two heuristics for the players sequences of observed opponent bids. The table shows that both of these relatively

Input: m – reward, B – number of bins, T – number of games, θ – fraction, d – window length
Candidate bids: $C = \{\frac{m}{B} \cdot i\}_{i=1}^{B}$.
Let $u(a, b)$ denote the utility of a player bidding a against an opponent bid of b.
For a given set of bids O, let $u(a, O) = \frac{1}{|O|} \sum_{o \in O} u(a, o)$ denote bid a's average payoff against the bids in O.
Step $t = 1$: select a random bid $b_1 \in_R C$.
foreach *step* $t = 2$ *to* T **do**
  $\ell_1 = \max\{1, t - d\}$
  $\ell_2 = t - 1$
  Let $O_t = \{o_{\ell_1}, \ldots, \ldots, o_{\ell_2}\}$ be the previously observed opponent bids.
  Set: $b_t = \arg\max_{b \in C}\{u(b, O)\}$
end

Algorithm 2: Limited Recall, Best Response (LRBR)

simple limited recall heuristics do worse than MWU, but perform better than the typical player. Although there was a considerable variance in results (with the largest standard error being 76.17), this illustrates that for this particular domain, the learning behavior of people under-performs even relatively simple algorithms.

d	∞	1	2	3	4
LRMB	-230.1	-545.8	-103.9	79.5	-331.1
LRBR	210.5	-544.8	36.9	172.8	270.9

Table 2: Average payoffs of limited-recall heuristics

A simple incremental response

Our final approach mimics the simplistic outcome-driven change in the bids that we observed in the previous section. This heuristic increments (decrements) the previous bid by $1{,}000$ (within the limits) following a loss (win). Testing this on the players' sequences of observed opponent bids yielded a mean payoff of $-2{,}308$, with a standard error of 89.2. This performance is comparable with the actual mean payoff obtained by our players, and as this is particularly a simple adaptive dynamic, this suggests that whatever approach players are taking, they are in general quite weak learners.

Of course, as discussed this conclusion does exclude arguably the most rational players, those who attempted to exploit the rules of the task, but still is representative of a substantial number of participants.

Conclusion

We empirically studied the behavior of crowdsourcing workers in an all-pay auction, which is a standard model for crowdsourcing contests. Our results show that human bids substantially deviate from the mixed strategy Nash-equilibrium bids. For the crowdsourcing contest designer, our analysis shows that a bimodal distribution of effort should be expected, with some very high effort and some very low effort. Our analysis also suggests that that contests tend to generate more effort than what would be exerted under Nash-equilibrium behavior. Given the weak performance of participants as learners in our game, it may be important to educate participants about the strategic implications of using contests for crowdsourcing.

We point out some limitations of our experiments. Despite the popularity of the Mechanical Turk platform, the behavior of its users may not extend to general crowdsourcing settings. In particular, our population may have contained users who did not make an effort to understand the game, but rather played to quickly win the base payment only. As a result, one could try alternative methods of payment, used in previous studies such as an initial endowment (instead of a base payment), or letting the players perform a tedious task, using the time they dedicated to it as their 'bid'.

Another limitation of this work is our perpective of viewing players as rational expected-utility maximizers. Our results indicate that if humans are attempting to maximize their expected monetary gain in the all-pay auction, there exist alternative learning algorithms that could achieve a higher performance. However, this only holds true for our monetary payoff function; if participants are risk-averse, they would be optimizing for a good trade-off between the expected payoff and the variance in payoffs. Similarly, if participants care only about winning or losing the auction (rather than the monetary gain), their true utility is different from the one we focus on in our analysis. Alterrnatively, if participants care about their payoff *relative* to that of other players (rather than their absolute gain), they would be optimizing for a different function (for a discussion of such alternative functions humans might be optimizing for see recent work on contests [35, 30]). Thus, an interesting line of future research is examining whether human behavior is more consistent with strong learning under such different utility functions.

Some other questions are also open for future research. A learning-theoretic model where the agents have limited learning capacities may give some theoretical traction to our results. Also, one can test the performance of learning algorithms against human opponents, and study their responses, as has been done in similar games [23]. One could also study the effect of social ties on the strategies of the users, or the effect of varying the rewards on players' behaviors. Finally, our approach of evaluating the ability of humans as learners by comparing their score to a range of learning algorithms would be interesting to apply to other settings.

References

1. E. Amann and W. Leininger. Asymmetric all-pay auctions with incomplete information: The two-player case. *Games and Economic Behavior*, 14(1):1–18, 1996.
2. A. Arad and A. Rubinstein. Colonel blottos top secret files. Levine's Working Paper Archive 814577000000000432, David K. Levine, Jan. 2010.
3. A. Arad and A. Rubinstein. The 11-20 money request game: A level-k reasoning study. *American Economic Review*, 102(7):3561–73, September 2012.

4. S. Arora, E. Hazan, and S. Kale. The Multiplicative Weights Update Method: a Meta-Algorithm and Applications. *Theory of Computing*, 8(1), 2012.
5. M. Baye, D. Kovenock, and C. d. Vries. The all-pay-auction with complete information. *Economic Theory*, pages 291–305, January 1996.
6. J. Bennett and S. Lanning. The netflix prize. In *In KDD Cup and Workshop in conjunction with KDD*, 2007.
7. M. Burkart. Initial shareholdings and overbidding in takeover contests. *The Journal of Finance*, 50(5):1491–1515, 1995.
8. J. W. Byers, M. Mitzenmacher, and G. Zervas. Information asymmetries in pay-per-bid auctions: How swoopo makes bank. *CoRR*, abs/1001.0592, 2010.
9. C. F. Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2011.
10. S. M. Chowdhury, D. Kovenock, and R. M. Sheremeta. An experimental investigation of colonel blotto games. *Economic Theory*, 52(3):833–861, 2013.
11. D. D. Davis and R. J. Reilly. Do too many cooks always spoil the stew? an experimental analysis of rent-seeking and the role of a strategic buyer. *Public Choice*, 95(1-2):89–115, 1998.
12. E. Dechenaux, D. Kovenock, and R. M. Sheremeta. A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669, 2015.
13. D. DiPalantino and M. Vojnovic. Crowdsourcing and all-pay auctions. In *Proc. 10th ACM conf. on Electronic commerce*, EC '09, pages 119–128, New York, NY, USA, 2009. ACM.
14. J. R. Douceur. The sybil attack. In *Revised Papers from the First Int. Workshop on Peer-to-Peer Systems*, IPTPS '01, pages 251–260, London, UK, UK, 2002. Springer-Verlag.
15. I. Erev and A. E. Roth. Predicting how people play games: Reinforcement learning . . . *American Economic Review*, 88:848–881, 1998.
16. C. Ernst and C. Thni. Bimodal bidding in experimental all-pay auctions. Technical report, 2013.
17. F. Fallucchi, E. Renner, and M. Sefton. Information feedback and contest structure in rent-seeking games. *European Economic Review*, 64:223–240, 2013.
18. Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, 55(1):119–139, 1997.
19. X. A. Gao, Y. Bachrach, P. Key, and T. Graepel. Quality expectation-variance tradeoffs in crowdsourcing contests. In *AAAI*, pages 38–44, 2012.
20. A. Gelder, D. Kovenock, and R. M. Sheremeta. Behavior in all-pay auctions with ties. 2015.
21. A. Ghosh and P. McAfee. Crowdsourcing with endogenous entry. In *Proceedings of the 21st international conference on World Wide Web*, WWW '12, pages 999–1008, New York, NY, USA, 2012. ACM.
22. U. Gneezy and R. Smorodinsky. All-pay auctions–an experimental study. *Journal of Economic Behavior & Organization*, 61(2):255–275, 2006.
23. P. Kohli, M. Kearns, Y. Bachrach, R. Herbrich, D. Stillwell, and T. Graepel. Colonel blotto on facebook: the effect of social relations on strategic interaction.

In *WebSci*, pages 141–150, 2012.
24. V. Krishna and J. Morgan. An analysis of the war of attrition and the all-pay auction. *Journal of Economic Theory*, 72(2):343–362, 1997.
25. O. Lev, M. Polukarov, Y. Bachrach, and J. S. Rosenschein. Mergers and collusion in all-pay auctions and crowdsourcing contests. In *The Twelfth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Saint Paul, Minnesota, May 2013. To appear.
26. Y. Lewenberg, O. Lev, Y. Bachrach, and J. S. Rosenschein. Agent failures in all-pay auctions. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pages 241–247. AAAI Press, 2013.
27. Y. Lewenberg, O. Lev, Y. Bachrach, and J. S. Rosenschein. Agent failures in all-pay auctions. *IEEE Intelligent Systems*, 32(1):8–16, 2017.
28. T. X. Liu, J. Yang, L. A. Adamic, and Y. Chen. Crowdsourcing with all-pay auctions: A field experiment on taskcn. *Proceedings of the American Society for Information Science and Technology*, 48(1):1–4, 2011.
29. V. Lugovskyy, D. Puzzello, and S. Tucker. An experimental investigation of overdissipation in the all pay auction. *European Economic Review*, 54(8):974 – 997, 2010.
30. S. D. Mago, A. C. Samak, and R. M. Sheremeta. Facing your opponents: Social identification and information feedback in contests. *Journal of Conflict Resolution*, 60(3):459–481, 2016.
31. M. Mash, Y. Bachrach, and Y. Zick. How to form winning coalitions in mixed human-computer settings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 465–471, 2017.
32. J. Potters, C. G. de Vries, and F. van Winden. An experimental examination of rational rent-seeking. *European Journal of Political Economy*, 14(4):783 – 800, 1998.
33. C. R. Price and R. M. Sheremeta. Endowment effects in contests. *Economics Letters*, 111(3):217–219, 2011.
34. C. R. Price and R. M. Sheremeta. Endowment origin, demographic effects, and individual preferences in contests. *Journal of Economics & Management Strategy*, 24(3):597–619, 2015.
35. R. M. Sheremeta and J. Zhang. Can groups solve the problem of over-bidding in contests? *Social Choice and Welfare*, 35(2):175–197, 2010.
36. D. O. Stahl and P. W. Wilson. On players models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1995.
37. J. C. Tang, M. Cebrian, N. A. Giacobe, H.-W. Kim, T. Kim, and D. B. Wickert. Reflecting on the darpa red balloon challenge. *Communications of the ACM*, 54(4):78–85, Apr. 2011.
38. M. Vojnovic. *Contest theory*. Cambridge University Press, 2015.
39. Z. Wang and M. Xu. Learning and strategic sophistication in games: The case of penny auctions on the internet. Working paper., 2012.
40. J. Yang, L. A. Adamic, and M. S. Ackerman. Crowdsourcing and knowledge sharing: strategic user behavior on taskcn. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, EC ’08, pages 246–255, New York, NY, USA, 2008. ACM.

41. H. P. Young. *Strategic learning and its limits*. Oxford University Press, 2004.

Including Spammers: Revised Results

In this appendix section, we revisit our findings without the exclusion of the so-called Spammers (the 85 players whose at least 25% of their bids were taken from the set $\{0, 1, 1000, 9{,}999\}$). Recall that in total, there are 340 players who played a total of $11{,}327$ games. The average revenue over all games, was $11,832$ (previously $13,730$).

Figure 7 shows the average player bid distributions, which is quite close to that in the original analysis (though of course including more "spam" bids).

Fig. 7: Average player distribution of bids (with spamming users).

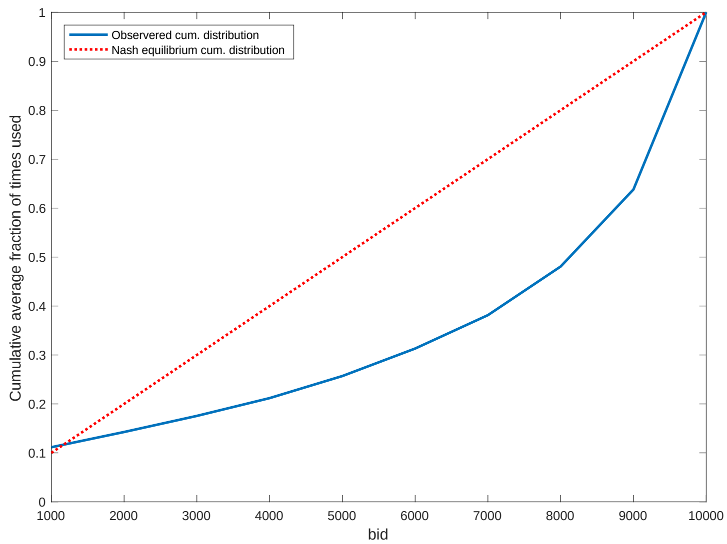


Figure 8 shows the average bid per time period (and again, the results are similar to those occuring with spammers, although unsuprisingly there are more players who do not adjust their bid when including spammers, as these are by definition players who use the same bids frequently).

Figure 9 and Figure 10 show key clusters when including spammers in the clustering analysis. Again, the results are qualitatively similar to the original analysis.

Figure 11 shows the average utility of bids against the empirical bid distribution. The figure is very similar to that in the original analysis. In other words, although best-responses against "spam" bids do well against the bids taken only from spammers, the *overall* best responses (when reacting to the general population of all players, including both spammers and non-spammers) are very similar to what we found in the original analysis. Figure 12 shows the utility distribution of players, which is again very similar to the distribution found in our original analysis.

To conclude, repeating the analysis carried in the main paper which *not* filtering our players who frequently use the same "focal-bids" (but do filtering our uses who used fake profiles) does not yield significantly different results. This indicates that our results are relatively robust to our choice of mechanism for eliminating spammers.

Fig. 8: Average bid per time period.

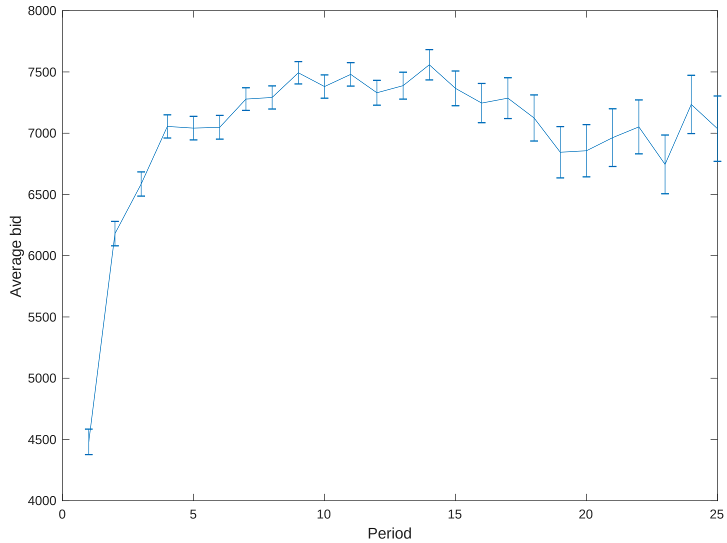


Fig. 9: Cluster 1 (64 players) player bid distributions

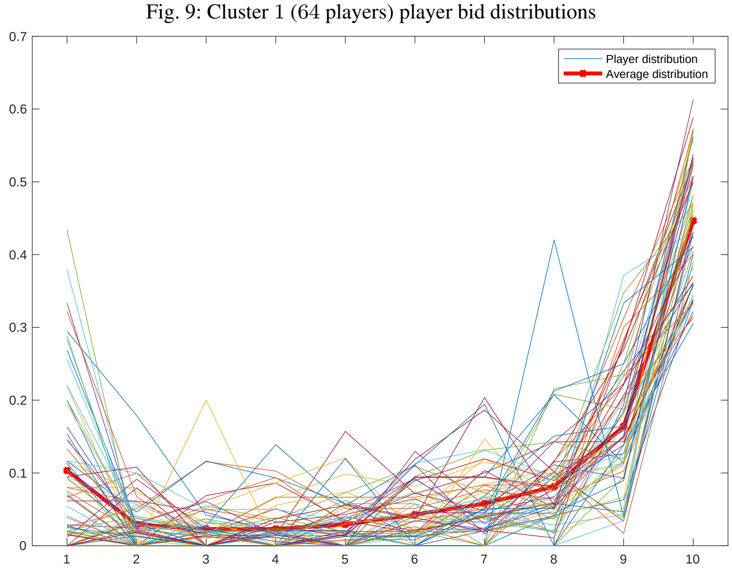


Fig. 10: Cluster 2 (43 players) player bid distributions

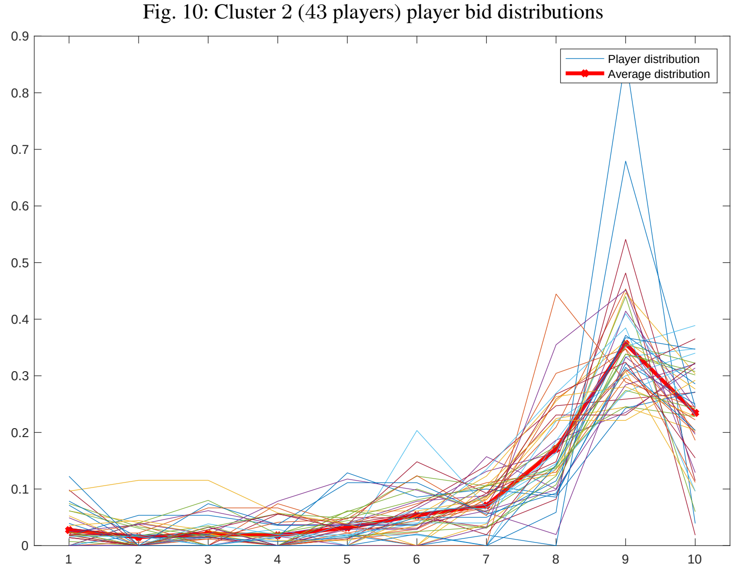


Fig. 11: Average utility against the empirical bid distribution

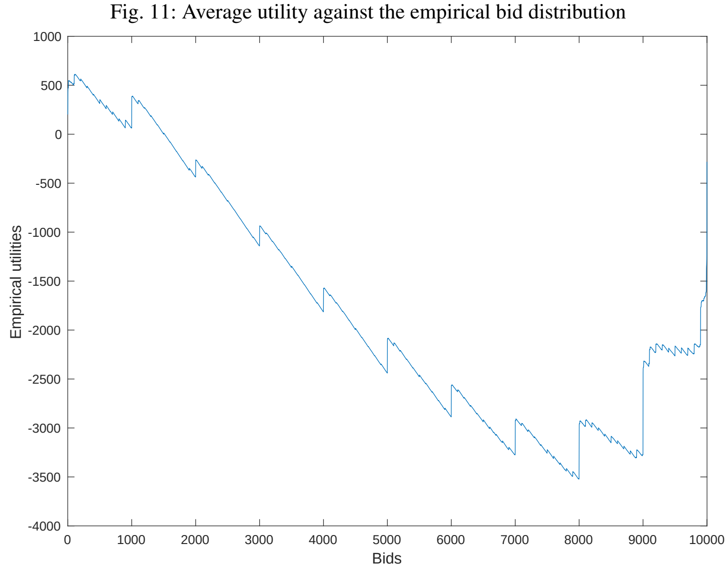


Fig. 12: Utility distribution of players