Rational and actual behavior in lowest unique bid auctions.

Simone Pigolotti\textsuperscript{1}, Sebastian Bernhardsson\textsuperscript{1}, Jeppe Juul\textsuperscript{1}, Gorm Galster\textsuperscript{1}, and Pierpaolo Vivo\textsuperscript{2}

\textsuperscript{1}Niels Bohr Institute, Blegdamsvej 17, DK-2100, Copenhagen, Denmark
\textsuperscript{2}Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy
(Dated: May 5, 2011)

In lowest unique bid auctions, \( N \) players bid for an item. The winner is whoever places the lowest bid, provided that it is also unique. We derive an analytical expression for the equilibrium distribution of the game as a function of \( N \) and study its properties, which are then compared with a large dataset of internet auctions. The empirical collective strategy reproduces the theoretical equilibrium with striking accuracy for small \( N \), while for larger \( N \) the quality of the fit deteriorates.

A crucial issue in evolutionary game theory is the understanding of how agents choose and adapt their strategies. In games where the number of players is large, the analogy with statistical physics has provided powerful tools to study both equilibria \cite{1,2} and dynamical properties of games \cite{3,4}. So far, most of the efforts in this field have been devoted to study games in which interactions among players are pairwise, the most notable and studied example being the prisoner’s dilemma \cite{5}. However, there are many examples of real collective games, where a unique winner is singled out from a large population. In such cases, comparing theory with empirical data is particularly challenging: as the number of players increases, a statistical description becomes more appropriate. However, the complexity of the rational strategy may also increase, thus making it more difficult for players to infer which behavior is rational \cite{6}.

Online auctions provide a fertile, yet scarcely investigated, ground for exploring this problem \cite{7,8,9,10}: the availability of large datasets provides a unique opportunity to study whether the behavior of real players is actually rational, that is whether it maximizes the player’s winning chances. Here, we focus on lowest unique bid auctions. This game is interesting for two reasons. First, it is sufficiently simple to allow for a comprehensive game theoretical analysis \cite{11,12} under the assumption of a fluctuating number of players. Second, detailed data of many auctions are freely available online. The game is formulated as follows: \( N \) players can bid for an item of value \( V \). The bid must be a multiple of a minimum amount, so that one can effectively consider bid amounts as natural numbers. The winner (if any) is the player who places the lowest bid, which is unique, i.e. no other player bid on that amount. All players must pay a fee \( c \) to take part in the auction; additionally, the winner has to pay the amount bid.

In this Letter, we derive an explicit analytic expression for the symmetric Nash equilibrium of the lowest unique bid auction and explore its dependence on the total number of players \( N \). We then compare the expression with a large dataset taken from the website auctionair.co.uk where bidders are notified in advance of the total number of allowed bids \( N \). We find a strong dependence of the actual behavior of players on the total number \( N \): the theory describes the behavior of small auctions (fewer than \( \approx 200 \) players) very well, while the bid distribution for large auctions deviates from the theoretical one and resembles an exponential distribution. This unexpected feature has the crucial consequence that the game is effectively a lottery for small \( N \), in the sense that each player has roughly the same chance of winning on each number. For larger \( N \), these chances are highly non-homogeneous as a consequence of the non-optimality of the strategy.

In a symmetric Nash equilibrium, all players adopt the same strategy, and no player can benefit from changing strategy unilaterally, i.e. should any of the players change strategy, his expected payoff would be equal or worse. Consider \( N \) individuals playing according to the same strategy \( p \), where \( p \) is the probability of bidding the number \( k \). Then, the distribution of bids is multinomial:

\[
P(n) = \frac{N!}{n_1!n_2!\cdots}p_1^{n_1}p_2^{n_2}\cdots
\]

(1)

where \( n = (n_1, n_2, n_3, \ldots) \) are the number of bids placed on each number \( k \). For convenience, we introduce the generating function:

\[
G_N(x) = \sum_{\{n\}} P(n) x_1^{n_1}x_2^{n_2}\cdots = (x \cdot p)^N.
\]

(2)

We now assume that the number of players \( N \) is not fixed, but fluctuates according to a Poissonian distribution of mean \( \lambda \), leading to a grand canonical generating function:

\[
\tilde{G}_\lambda(x) = \sum_{N} \frac{\lambda^N e^{-\lambda}}{N!} \tilde{G}_N(x) = \exp[\lambda(p \cdot x - 1)],
\]

(3)

i.e. the number of bids on each number is also Poisson distributed with mean \( f_k = \lambda p_k \). By differentiating the
generating functions with respect to $x$, it is straightforward to compute the probabilities $w_k$ that $k$ will be the winning number, i.e. there is a unique bid on $k$ and no unique bids on lower numbers:

$$w_k = f_k e^{-f_k} \prod_{j=1}^{k-1} (1 - f_j e^{-f_j}), \quad (4)$$

It is also useful to compute the probability $c_k$ that the number $k$ is a potential winning number, i.e. that there are no bids on $k$ and no winners on lower numbers:

$$c_k = e^{-f_k} \prod_{j=1}^{k-1} (1 - f_j e^{-f_j}). \quad (5)$$

We note that the average probability of each bid to be a winner on a given number, $w_k/f_k$, is equal to the probability, $c_k$, of numbers to be potential winning numbers, which is a peculiar property of Poisson games.

The expected payoff for a player bidding $k$ will be $(V - k)w_k/f_k - c$. At equilibrium, the expected payoff should be independent of $k$. Otherwise, players could benefit from bidding on numbers with high payoffs more frequently. Imposing this condition leads to a recurrence relation for the average bidding frequencies

$$f_{k+1} = \ln(e^{f_k} - f_k) + \ln\left(1 - \frac{1}{V - k}\right) \quad (6)$$

To avoid a negative number of bids, the support of the distribution must be limited to the region where all $f_k$'s are positive. In this region $k < V$ holds, so that the $f_k$'s are strictly decreasing. The initial condition $f_1$ has to be determined iteratively from the condition $\sum_j f_j = \lambda$.

If the frequencies $f_k$ tend to zero for values of $k$ much smaller than the value of the item $V$ (a condition that is always fulfilled in our dataset), the last term in (6) can be disregarded. In this limit (formally corresponding to $V \to \infty$) the recurrence relation has been derived in [12] and it is well defined for all values of $k \in N$. We extend this solution by noting that the normalization condition $\sum_k f_k = \lambda$ implies an explicit expression of $f_1$: eq. 6 can be written as $f_k = e^{f_k} - e^{f_{k+1}}$, which summed over $k$ yields the initial condition $f_1 = \ln(\lambda + 1)$. This allows for an explicit expression for any specific $f_k$'s by iteration. Substituting the solution into (4) also shows that the average chance of winning of each bid is equal to $(\lambda + 1)^{-1}$, which is also equal to the chance of having no winner in the auction.

Before discussing the comparison with the data, we briefly study some properties of the $V \to \infty$ equilibrium distribution. When $f_k \gg 1$, a continuous approximation of (6) shows that $f_k$ decreases logarithmically with $k$. For small $f_k \ll 1$, one can approximate $f_{k+1} \approx f_k^2/2$, showing that the distribution has a super-exponential cutoff. The scaling of the value $k^\ast(\lambda)$, around which the cutoff occurs, can be estimated in the following way. By removing all the players that bid on $k = 1$ we change the average number of players by the amount $f_1$ giving $\lambda_{new} = \lambda_{old} - \ln(\lambda_{old} + 1)$. This is equivalent to shifting the whole distribution by one along the $k$ axis, resulting in $k_{new}^\ast = k_{old}^\ast - 1$. This scaling transformation in the continuum limit becomes $d\lambda/dk^\ast = \ln(\lambda + 1)$, with the solution $k^\ast(\lambda) = \ln(1 + \lambda) + C$, where $\ln(\lambda) = \int_0^\lambda dt/\ln t$ is the logarithmic integral and $C$ is a constant of $O(1)$.

We now move to the comparison of the equilibrium solution with empirical data from the website auction-air.co.uk. The dataset includes 724 online auctions from April 2007 to January 2011 with a variable number of bids ranging from $N = 26$ to $N = 4748$. The number of allowed bids in a particular auction is announced before bidding starts, and the auction closes when this number is reached. Each player is allowed a limited number of bids. The average number of bids per player in the dataset was only 1.88, so in the subsequent analysis we will neglect the effect of multiple bids by the same player and treat the bids as statistically independent.

In Fig. 4 we compare the theoretical and empirical
bidding frequencies in different auctions having different numbers of players and different values of the item. In order to make the histograms smoother, we averaged each panel over \( L \) different auctions having similar numbers of players and same value of the item (shown in the figure).

On the fine scale of single numbers, the data show a structure dictated by known psychological effects. Players tend to favor odd numbers over even numbers. Some specific numbers (like 17 and other primes) are perceived to be ‘original’ by some players, and are chosen with significantly larger probability than neighboring ones.

On a coarser scale, the agreement between theory and data is striking for smaller auctions (i.e. fewer than 200 players, panels A and B of Fig. 1). It is particularly remarkable that the empirical histograms reproduce the sharp cutoff, considering the non-trivial dependence of \( k^* \) on the number of players.

Theoretically, players should adjust their bidding strategies according to the number of players rather than the value of the item, which can be assumed to be infinite. In all auctions in the dataset, the cutoffs of the corresponding theoretical distributions occur at bid values much smaller than the item values (shown in panels). To test whether the empirical bidding distributions are independent of \( V \), panel D compares two sets of auctions with the same number of players, but with item values that differ by a factor of three. The bidding distribution for the pricey item have a slightly heavier tail, but overall the distributions are very similar.

The agreement between theory and data progressively deteriorates as the number of players increases (see panel C). For a large number of players (more than 2000) an exponential distribution fits the data better, as shown in the lower panels of Fig. 1. This can be understood by arguing (see e.g. [14–16]) that players having incomplete information about the game play according to a strategy defined as \( p_k = \exp(-\beta E_k)/\sum \exp(-\beta E_j) \) where \( E_k \) is the expected payoff when placing a bid \( k \) and \( \beta \) quantifies the degree of uncertainty about the game. If players have poor knowledge about the optimal strategy, they could assume a uniform prior probability to win on each \( k \), resulting in \( E_k \) decreasing linearly with \( k \) due to the cost of the bid when winning. Substituting this prior into the logit strategy yields a strategy exponentially decreasing with bid size.

A quantitative measure of how the data deviates from theory is shown in Fig. 2 where, for each auction, we plot the \( l_2 \) distance \( d \) between the theoretical probability distribution and the empirical one, \( d = N^{-2} \sum_k (f_k - \phi_k)^2 \), where \( \phi_k \) is the number of bids on \( k \) in a given auction. If bids were randomly drawn from the theoretical distribution, the expected distance would decrease with the number of players as \( \langle d \rangle = N^{-1} \). Empirically, the distances have a large spread around the expected value for small auctions and are consistently larger than expected for \( N > 200 \).

Another interesting quantity is the distribution of actual winning numbers. At equilibrium, the distribution of winning numbers should be equal to the distribution of bids. As shown in Fig. 3, the empirical distribution of winning numbers supports this feature. The vast majority of winning numbers fall within the theoretical cutoff (pink shaded area of Fig. 3). In the figure, we also show the analytical estimate \( k^* \approx N/\ln(N)(1+1/\ln(N)) \) based on the asymptotic expansion of the logarithmic integral. The black line is the average winning number, which in the relevant range, scales approximately in the same way as the cutoff.

Binning the data yields average winning numbers in excellent agreement with the theory, even for large \( N \) where the empirical distribution of bids departs from the theory. However, the variance of winning numbers becomes much smaller than the theoretical prediction for \( N > 10^3 \). To further explore this phenomenon, we can calculate the actual probabilities to win on a certain number, given the empirical distributions of bids for auctions of various sizes. This probability is given by \( w_k/f_k \) which, since \( c_k = w_k/f_k \), is the same as the probability to win on a certain number for an additional player entering the game. The results are shown in Fig. 4 for the same examples considered in Fig. 1. For simplicity we have rescaled the probability with \( N+1 \), such that 1 corresponds to the winning probability of each bid in the Nash equilibrium. For auctions with few players, the relative probabilities are close to one with small scale fluctuations. The largest chance of winning in these smaller auctions is not more than 4 – 5 times higher than the average. In contrast, for large auctions, the normalized \( w_k/f_k \)’s are...
sharp peaked around the theoretical average winning number. As a consequence, the winning chances on low and high numbers are very small, while intermediate bids can have a chance of winning more than 60 times higher than the average bid in the Nash equilibrium, providing opportunities for exploitation and thus potential room for adaptation.

To conclude, in this paper we have studied the equilibrium bidding distribution for the lowest unique bid auction and compared it with data from real auctions. The emerging picture is that players are able to infer the optimal strategy when the number of players is not too large. In auctions with a large number of players, the strategy is non-optimal and the behavior seems to be controlled by the simple prior of not assuming a particular preference for any number while avoiding the cost of a large bid. Consequently, the game can be considered as a lottery (each number has equal probability to win) only for small auctions, when players adopt the equilibrium solution. When $N$ is large and players depart from the equilibrium solution, the chance of winning depends much more on the chosen strategy.

This result raises non-trivial questions about the effectiveness of adaptive dynamics in collective games. To verify whether the outcome was not a simple consequence of the larger number of auctions for smaller $N$, we checked whether there was any sign of adaptation in the dataset from older to more recent auctions at a fixed number of players, without finding any significant correlation. A probable reason is that empirical populations are made up of a mixture of players of different experience and the turnover rate is such that the distribution does not change with time in a significant way. This suggests that, in this class of games, adapting to the optimal strategy is inherently more difficult when the number of players increases, and might be virtually impossible for very large populations.

We acknowledge useful discussions with M. Marsili, K. Sneppen, C. Strandkvist and P. Kempson.

---