Can auctions select people by their level-k types?

Choo, Lawrence and Kaplan, Todd R. and Zhou, Xiaoyu

University of Erlangen-Nuremberg and University of Haifa, University of Exeter and University of Haifa, University of Erlangen-Nuremberg

September 2019

CAN AUCTIONS SELECT PEOPLE BY THEIR LEVEL-K TYPES?*

Lawrence Choo[†]   Todd R. Kaplan[‡]   Xiaoyu Zhou[§]

September 9, 2019

Abstract

In this paper, we seek to determine if auctions can be used to select players according to their level-k types. To do so, we embed auctions into the p-beauty contest game. We find that by using different designs, we can get the auction winners to be either the lower level-k types or the higher level-k types. In particular, when the value of winning the auction is increasing in the level-k types of all the players, higher level-k players bid higher. When the value of winning the auction is decreasing in the level-k types of all the players, the lower level-k players bid higher. Taken together, our experiment confirms that we can use auctions to select players by their level-k types. This shows that auctions can allow an economic designer to affect the outcome of a game through the selection of level-k types entering to play the game.

JEL-classification: C72, C90, D44.
Keywords: p-beauty contest game, level-k, auctions, mechanism design.

1. INTRODUCTION

The *level-k* model (e.g., Costa-Gomes et al., 2001; Nagel, 1995; Stahl and Wilson, 1994, 1995) is often used to explain anomalous behaviour in strategic interactions. The model posits that people can be partitioned into L_k types, where $k \in \mathbb{N}$, and anchors upon a nonstrategic L_0 type who is often assumed to uniformly randomise over all strategies. The other L_k types ($k > 0$) believe that their opponents are L_{k-1} types and best respond to such beliefs through iterative thought experiments.[1] See Crawford et al. (2013) for an extensive survey of the literature.[2]

The level-k model attributes aggregated outcomes in a game to the distribution of level-k types. This implies that an economic designer (e.g., regulators, policy-makers) may be able to directly influence the outcomes through the selection of level-k types. Clearly, such selection is trivial when the designer observes everyone's type. What happens if he does not?[3]

In principle, the designer may infer players' level-k type from their decisions in previous games. However, Georganas et al. (2015) find that level-k types may not be stable across different games. There is a growing body of literature that finds people's cognitive abilities are correlated to level-k types (e.g., Brañas-Garza et al., 2012; Burnham et al., 2009; Carpenter et al., 2013; Gill and Prowse, 2016). In applications, cognitive abilities are indirectly measured through psychological tests such as the Raven's IQ test or the cognitive reflective test (Frederick, 2005). This however means that the reliability of the selection will also depend on whether the test can be 'gamed' or learned.[4] Finally, the designer may select people into a game by their previous decisions in that game. However, such a process could be costly (i.e, the designer has to first run the game with a larger

---

[1]A close variant of the level-k model is the *cognitive hierarchy* (Camerer et al., 2004) model where L_k ($k > 0$) types belief that others are mixture of lower types. There is also the *noisy introspection* (Goeree and Holt, 2004) model that uses the level-k structure but assumes that players best respond to the noisy best response of others.

[2]Variants of the level-k model have been used to explain behaviour in the *p*-beauty contest game (e.g., Nagel, 1995; Camerer et al., 2004), normal-form games (e.g., Stahl and Wilson, 1994, 1995; Costa-Gomes et al., 2001), auctions (e.g., Crawford and Iriberri, 2007; Georganas, 2011) and the centipede game (Kawagoe and Takizawa, 2012). More recently, the model is also used to explain betting behaviour (Östling et al., 2011) and the 'cold releases' of movies (Brown et al., 2012).

[3]In some circumstances, the designer might know the distribution of types in the population. This can inform the designer as to the prior probability of a randomly selected player being a specific type. This study examines whether a purposeful selection mechanism can give the designer a more accurate posterior of a selected player's type over the random selection mechanism.

[4]For example, Haigh (2016) find that performances in the cognitive reflective test strongly depend on whether subject have heard of the test.

group to find a subgroup of players). Also, since the game would need to be played twice, players' decisions could be influenced by learning from the initial game. In addition, the players' knowledge that the decisions in the game will affect their subsequent participation may lead to them strategically altering their behaviour in the initial game and their beliefs about others (potentially in an unknown way).

This paper proposes a more 'direct approach', using auctions to select people by their level-k types.[5] The idea is as follows. Suppose that the designer has some understanding as to how the different level-k types will behave in the game, he could design incentives such that the different types will have different willingness to pay to participate in the game. When auctions are used to allocate the participation rights to the game, the rights will be purchased by those with the highest willingness to pay. This willingness to pay will be 'calibrated' to be highest for the *desired* level-k type.

To test the idea that auctions can select people by their level-k types, we design an experiment where auctions are used to allocate the participation rights to the p-beauty contest game (Nagel, 1995).[6] The game involves $n \geq 2$ players who each choose a number between 0 and 100 (both limits included). The player whose chosen number is closest to $p = 2/3$ times the average of all numbers wins a fixed prize.[7] The equilibrium is for all players to choose 0. However, this is seldom observed in practice. Instead, people often choose numbers around $50p$ and $50p^2$ (see Nagel, 1995; Bosch-Domènech et al., 2002) which correspond to the predicted choices of the L_1 and L_2 types, respectively.[8]

Two treatments, the AVG and AVGINV treatments, are considered in our design. Each experimental session uses the 9th price auction to allocate 8 tickets – there are 24 bidders in each sessions.[9] A player can only purchase one ticket.

---

[5]We are aware that previous studies by Crawford and Iriberri (2007) and Georganas (2011) have attempted to use the level-k model to explain bidding behaviour in auctions. However, to the best of our knowledge, this study is the first to investigate whether auctions can select people by their level-k type.

[6]We do not claim that the p-beauty contest game may be the best game for the purposes of this study or that it is informative as to real world economic interactions. Instead, we use it because of the robust evidence that deviation from equilibrium in the game follow a coherent pattern that can be explained by the level-k model.

[7]In the case of a tie (i.e., more than one claimants) the prize is equally split amongst all claimants.

[8]When $n > 2$, a player may deviate from the equilibrium if he believes that others who do so. To address this, Grosskopf and Nagel (2008) consider the $n = 2$ version of the p-beauty contest game, where it is a weakly dominant strategy to choose 0. They find that chosen numbers in the $n = 2$ version are no different from those in the $n = 4$ version.

[9]We use the 9th price auction to allocate 8 tickets as bidding one's valuation is the weakly dominant strategy. This limits the strategical interplay that may arise with other auction formats.

2

Ownership of a ticket gives the player (a) the rights to play the $n = 8$ player p-beauty contest game and (b) an *uncertain* bonus payment – the bonus is separate from the p-beauty contest game prize. The two treatments only differ on how the bonus is determined. Applying the level-k model into this framework, we show that players' willingness to pay (for a ticket) in the AVG and AVGINV treatments will be decreasing and increasing, respectively, with their L_k types ($k > 0$). Given a fixed distribution of L_k types, auctions in the AVG and AVGINV treatments will hence select (allocate tickets to) the higher and lower L_k ($k > 0$) types, respectively, to participate in the p-beauty contest game. The model therefore predicts that:

- Auctions in the AVG treatment will select lower level-k types into the p-beauty contest game.

- Auctions in the AVGINV treatment will select higher level-k types into the p-beauty contest game.

Experimental data is often *noisy* and it is difficult to directly infer subjects' L_k type from their decisions in the p-beauty contest game. We hence use a novel approach to assign each experimental subject a L_k type.[10] To do so, we required subjects in both treatments play the p-contest game twice without feedback. In the first attempt, subjects played the standard $n = 8$ player p-beauty contest game. In the second attempt, subjects played the $n = 8$ player p-beauty contest game with the auction mechanism. We also use the *strategy method* (Selten, 1967) in the second attempt, subjects submitted their bids and also the p-beauty contest game number to be used if they were allocated a ticket. Based on their choices over both attempts, we devise a procedure to assign subjects a L_k type. Furthermore, by comparing subjects' choices in the second attempt between AVG and AVGINV, we can also shed light on the question as to whether the type of auction mechanism influences their behaviour in the p-beauty contest game.

Our results are summarised as follows. Firstly, we find that the type of auction mechanism has no significant influence on subjects' behaviour in the p-beauty contest game. Secondly, we find that auctions in the AVG treatment select subjects who are assigned lower L_k types. Thirdly, we find that auctions in the AVGINV treatment select subjects who are assigned higher L_k types. Taken together, these

---

[10]Prior studies (e.g., Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006) have used mixture models to estimate subjects' L_k type. This approach typically requires subjects to repeatedly play different variations of the same game in order to detect coherent patterns in their decision-making process. However, repeated play in the p-beauty contest game often results in behaviour converging towards the equilibrium (e.g., Nagel, 1995; Duffy and Nagel, 1997; Gill and Prowse, 2016) even when subjects receive no feedback (see Weber, 2003).

findings show that auctions can indeed be used to select people by their level-k types.

This study also adds to the growing literature that investigates the influence of markets on strategic interactions. For example, prior experiments find that auctioning participation rights can improve efficiency in coordination games (e.g., Huyck et al., 1993; Crawford and Broseta, 1998), increase contributions in public goods games (Broseta et al., 2007), increase the prospects of collusion in oligopoly markets (Offerman and Potters, 2006) and affect behaviour in the ultimatum game (Shachat and Swarthout, 2013). On this issue, we find that auction markets in the p-beauty contest game (AVGINV treatment) can be used to drive behaviour closer to the equilibrium prediction.

The rest of this paper is structured as followed. Section 2 details our experiment design, Section 3 details our results and Section 4 concludes.

2. EXPERIMENT DESIGN

The AVG (3 sessions, 72 subjects) and AVGINV (3 sessions, 72 subjects) treatments were conducted in May 2018 with student subjects at the Shanghai Jiao Tong University (SUJT) Smith Experimental Economics Research Center. Subjects were recruited through ORSEE (Greiner, 2015) and the experiment was programmed with z-Tree (Fischbacher, 2007). Payoffs in the experiment were denoted in points and exchange into cash at the rate of one point to 0.45 RMB – subjects also received a 10 RMB show-up payment. Each session took about 60 minutes and the mean cash earning was approximately 60 RMB (8.72 USD).

In the following, we will first detail the AVG and AVGINV treatments. Thereafter, we will introduce the level-k model's predictions and our test hypotheses.

2.1. THE AVG AND AVGINV TREATMENTS

Each session in the AVG and AVGINV treatments involves 24 subjects who make decisions over two stages. The instructions for stage 2 are only distributed at the end of stage 1. Subjects also receive no feedback in between both stages.

- **Stage 1**: The 24 subjects are randomly matched into three groups. Each subject submits a number (up to two decimal places) to the $n = 8$ player p-beauty contest game for a winning prize of 15 points – there is one prize per group.

4

- **Stage 2:** The 24 subjects are each endowed with 100 points. We use the 9th price auction mechanism to allocate 8 tickets. Each ticket gives the owner:

  (a) the rights to again play the $n = 8$ player p-beauty contest game for a winning price of 15 points (i.e., ticket owners will play against other ticket owners),[11]

  and

  (b) an uncertain common bonus payment of $X \in [0, 100]$ – to be detailed in the later paragraph.[12]

  We use the strategy method in stage 2. Subjects submit their bids (between 0 and 100) and their numbers to p-beauty contest game if they purchased a ticket.

For the ease of exposition, we will refer to subjects' p-beauty contest game number in stages 1 and 2 as the BCG1 and BCG2 numbers, respectively. Also, we refer to auction *winners* and *losers* as subjects who are successful and unsuccessful, respectively, at purchasing a ticket.

In each session, define $\bar{X}$ as the average BCG1 number of 8 random subjects. The bonus in the AVG and AVGINV treatment sessions are computed as:

$$X_{AVG} = \bar{X},$$
$$X_{AVGINV} = 100 - \bar{X}.$$

As such, the treatments only differ on how the bonus is computed. We also elicited subjects' beliefs about the bonus payment: subjects submitted a number between 0 and 100, and received 10 points if the bonus was within 3 integers of their number. At the end of the experiment, subjects received full feedback as to their decisions in stages 1 and 2. Before receiving their cash payments, they also performed the three question cognitive reflect test (Frederick, 2005) and the Holt and Laury (2002) risk preference elicitation task.

Remarks: We choose to base the bonus on the BCG1 number as it seems the easiest for subjects to comprehend. It is also possible to base the bonus on the BCG2 numbers of auction winners. However, doing so for each auction winner

---

[11]Unlike stage 1, there is only one p-beauty contest game prize to be awarded amongst the 24 subjects.

[12]The auction here has a *common-value* flavour given that bonus payment is identical for all ticket holders. However, players' expectations about the bonus will depend on their L_k types. Since each L_k ($k > 0$) type does not allow for his beliefs to be incorrect, we believe that the auction is more closely aligned to the *independent private-value* setup.

would require the bonus to depend on the average number of the other seven auction winners – the bonus might be different for all auction winners.[13,14]

2.2. THEORETICAL ANALYSIS

The equilibrium is for the BCG1 and BCG2 numbers to be zero. Risk-neutral players in the AVG and AVGINV treatments will value the tickets at 15/8 and 100 + (15/8), respectively, and the tickets are randomly sold at the respective prices.[15]

In the following, we assume that the L_0 type uniformly randomises over all numbers in p-beauty contest game and bids randomly in the auction. Furthermore, assume that all players are risk neutral and the auction mechanism does not influence the L_k (k > 0) types' beliefs about their opponents' types. As such, a L_k (k > 0) number will submit the same BCG1 and BCG2 numbers – only the L_0 types can submit different numbers.

In the p-beauty contest game, a L_1 type believes that everyone else is a L_0 who uniformly randomises. He thus expects the average to be 50 (mean of the uniform distribution) and chooses 50p to minimise the distance between his number and p times his expected average – he neglects the influence of his own number on the average. A L_2 type believes that everyone else is a L_1 type who chooses 50p. From the above logic, the L_2 type best respond to such beliefs by choosing $50p^2$. More generally, a L_k type believes the average to be $50p^{k-1}$ and thus chooses $50p^k$.

Let v^k denote a L_k (k > 0) type's valuation of a ticket in stage 2,

$$
\begin{aligned}
v^k_{AVG} &= \overbrace{50p^{k-1}}^{(A)} + \overbrace{15m}^{(B)} \\
v^k_{AVGINV} &= \underbrace{(100 - 50p^{k-1})}_{(A)} + \underbrace{15m}_{(B)}
\end{aligned}
$$

where the component (A) is his expected bonus amount and the component (B)

---

[13]If the bonus only depended on the average BCG2 number of auction winners, then each subject in the AVG and AVGINV treatments has the strategic incentive to choose 100 and 0, respectively, as their BCG2 numbers in order to maximise their bonus payment.

[14]Such a more complicated design was used in Kaplan and Ruffle (2004) to test whether there was a self-serving bias where one may guess higher in the AVG if one's payment was higher when the other guessed higher.

[15]If the BCG1 and BCG2 numbers are zero for all players, then the expected bonus in the AVG and AVGINV treatments are 0 and 100, respectively. As such, a risk-neutral player will in the AVG and AVGINV treatments will value each ticket at 15/8 and 100 + (15/8) points, respectively – each ticket owner has an equal chance of winning the prize in stage 2. However, bids are bounded between 0 and 100.

is his expected cashflow from playing the p-beauty contest game if he purchased a ticket – here $m \in [0, 1]$ denotes the subjective probability of winning the prize and is assumed to be constant (does not vary with L_k type or treatment).

Given that each L_k ($k > 0$) type believes that all other player are type L_{k-1}, even though there is a common value to winning the auction, it is a weakly dominant strategy for him to bid his valuation.[16] As such, bids for L_k ($k > 0$) types in the AVG and AVGINV treatments are decreasing and increasing, respectively, with k. Finally, since choices in the p-beauty contest game are expected to be decreasing with k, the correlation of p-beauty contest game choices and bids will be positive and negative for L_k ($k > 0$) types in the AVG and AVGINV treatments, respectively. In contrast, bids and p-beauty contest choices are independent for L_0 types in both treatments.

The level-k model therefore predicts that auctions are able to *select* (i.e., allocate tickets to) people by their level-k types. For a fixed population of L_k types,

- Auctions in the AVG treatment will select the lower L_k ($k > 0$) types.

- Auctions in the AVGINV treatment will select the higher L_k ($k > 0$) types.

We remain nonchalant about the selection of L_0 types since they are assumed to bid randomly.

Comment. Given the above, each L_k ($k > 0$) always expects to win the auction in the AVGINV treatment. In contrast, the L_2 and higher types in the AVG treatment never expect to win the auction – the L_1 type may assign some chance of winning the auction since the L_0 type bids randomly. However, L_2 and above types in the AVGINV treatment may entertain the possibility that they overestimated the L_k types of the other players – this does not affect the L_1 type who believes that everyone else if a L_0 type. We assume that L_k types never underestimate the types of other players (i.e., they always think that they are the highest type in the population). Since in the AVGINV treatment, the players expect to win the auction and the bonus might be lower since it is inversely related to the BCG1 numbers of 8 random players, the L_2 and higher types may *shade* their bids, that is, bid less than their valuation. In contrast, in the AVG treatment, we expect entertaining the possibility of overestimating types to have limited

---

[16]A simplified version of our auction design in the 2nd price private-value auction. Crawford and Iriberri (2007) use the level-k model to explain bidding behaviour in auctions. Their model anchors upon two different L_0 types, one that bids randomly (*random* L_0) and one that bids truthfully (*truthful* L_0). This leads to the *truthful* and *random* L_1 types who best respond to a truthful and random L_0 types, respectively, as well as the *truthful* and *random* L_2 types who who best respond to a truthful and random L_1 type, respectively. The authors show that in a 2nd price auction, it is still a weakly dominant strategy for higher L_1 and L_2 types to bid their valuation.

influence on the bidding behaviours of L_2 and higher types in the AVG treatment since they do not expect to win the auction. Importantly, the above discussions suggest that even with two identical populations of players, it is possible for bids to be different from 100 minus the bids in the AVGINV treatment.[17] In particular, the latter can be expected to be higher.

2.3. HYPOTHESES

We do not perfectly observe subjects' L_k types. However, the level-k model predicts that choices in the p-beauty contest games are decreasing with subjects' L_k type. This leads us to the following within-treatment and between-treatment test hypotheses:

H1a : *(within-treatment). BCG1 and BCG2 numbers in the AVG treatment are higher for auction winners relative to auction losers.*

H2a : *(within-treatment). BCG1 and BCG2 numbers in the AVGINV treatment are lower for auction winners relative to auction losers.*

H3a : *(between-treatment). Auction winners' BCG1 and BCG2 numbers are higher in AVG treatment relative to the AVGINV treatment.*

H4a : *(between-treatment). Auction losers' BCG1 and BCG2 numbers are lower in AVG treatment relative to AVGINV treatment.*

Building on the hypotheses H1a-H4a, we introduce a procedure to assign each subject a L_k type based on their BCG1 and BCG2 numbers – we discuss the procedure later. Based on this assignment procedure the corresponding within-treatment and between-treatment test hypotheses are:

H1b : *(within-treatment). The L_k (k > 0) types in the AVG treatment are lower for auction winners relative to auction losers.*

H2b : *(within-treatment). The L_k (k > 0) types in the AVGINV treatment are higher for auction winners relative to auction losers.*

H3b : *(between-treatment). The L_k (k > 0) types for auction winners are lower in the AVG treatment relative to the AVGINV treatment.*

H4b : *(between-treatment). The L_k (k > 0) types for auction losers are higher in the AVG treatment relative to the AVGINV treatment.*

---

[17]If the population of players are identical and players always bid their valuation, a L_k (k > 0) type's bid in the AVG treatment will be identical to 100 minus the same type's bid in the AVGINV treatment.

3. RESULTS

3.1. PRELIMINARIES

We do not find any significant between-treatment differences in BCG1 numbers (Mann-Whitney (MW), P = 0.61), cognitive reflective test performances (MW, P = 0.67) and elicited risk preferences (MW, P = 0.48).[18,19]

As an empirical warmup, we first investigate whether the auction mechanism affects subjects' behaviour in the p-beauty contest game. One possibility is that the players foresee that the auction mechanism selects different types of players and adjust their behaviour accordingly. If so, we can expect BCG2 numbers to differ in the AVG and AVGINV treatments. However, we find no significant between treatment differences in BCG2 numbers (MW, P = 0.36).

An alternative possibility is that participating in auctions lead to subjects submitting different numbers in the first and second p-beauty contest games through affecting their thought processes. Subjects may also 'learn' about the equilibrium through repeated play even when they receive no feedback (see Weber, 2003). Indeed, we often find subjects' BCG2 numbers to be slightly lower than their BCG1 numbers. However, the differences are not significant in the AVG (Wilcoxon Signed-rank (WSR), P = 0.09) and the AVGINV (WSR, P = 0.22) treatments. Furthermore, the nominal differences in BCG1 and BCG2 numbers are no more than five for 58% and 61% of subjects in the AVG and AVGINV treatments, respectively.

We test whether there is any coherent differences in how AVG and AVGINV treatment subjects behave over the first and second p-beauty contest games – this builds on the logic that both treatments offer contrasting predictions as to the selection of L_k types.[20] We define Δ as subjects' BCG1 number minus their BCG2 number.[21] Between-treatment comparison finds no significant differences in Δ (MW, P = 0.67).

---

[18]Three sessions were conducted for each treatment. We also find no significant between-sessions differences in BCG1 and BCG2 numbers for the AVG (Kruskal-Wallis, P $\geq$ 0.27) and AVGINV (Kruskal-Wallis, P $\geq$ 0.09 treatments).

[19]We do not find any significant correlations between BCG1 or BCG2 numbers with subjects' performances in the cognitive reflective test nor their risk preferences.

[20]Kaplan and Ruffle (2004) provide one possibility as to how the bonus computation can affect behaviour. Subjects in the AVG treatment might expect the tickets to be purchased by those other subjects who believe the average BCG1 number to be high and consequently submit a high BCG2 number. As such, subjects in the AVG treatment would submit a higher BCG2 number than their BCG1 number. By the same logic, subjects in the AVGINV treatment would submit a lower BCG2 number than their BCG1 number.

[21]A negative (resp. positive) Δ implies that the subject's BCG2 (resp. BCG1) number is relatively higher.

The above observations suggest that the auction selection mechanism has no significant influence on subjects' behaviour in the p-beauty contest game.[22] The discussions henceforth will mainly focus on the BCG2 numbers. The summary statistics of the above observations are reported in the Appendix.

3.2. AUCTIONS AND *P*-BEAUTY CONTEST NUMBERS

Figure 1 details the scatter plot of bids and BCG2 numbers for subjects in the AVG and AVGINV treatments.[23] We observe bids and BCG2 numbers to be positively (Spearman $\rho = 0.47$, P < 0.01) and negatively (Spearman $\rho = -0.42$, P < 0.01) correlated for subjects in the AVG and AVGINV treatments.
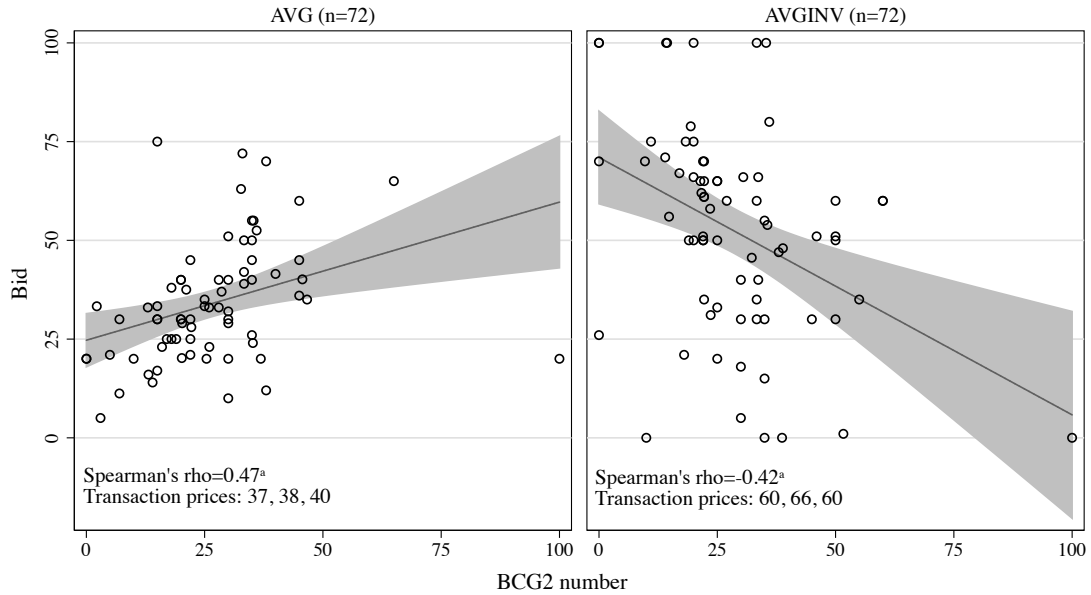
Figure 2 details the Boxplot distribution of BCG2 numbers for auction winners and losers in the AVG and AVGINV treatments. Within-treatment comparisons find BCG2 numbers to be significantly higher (MW, P < 0.01) for auction winners relative to losers in the AVG treatment. In contrast, BCG2 numbers are significantly lower (MW, P < 0.01) for auction winners relative to losers in the AVGINV treatment. Turning our attention to the between-treatment differences, we find auction winners' BCG2 numbers to be significantly lower (MW, P < 0.01) in the AVG treatment relative to AVGINV treatment. Finally, we find auction losers' BCG2 numbers to be significantly higher (MW, P < 0.01) in the AVG treatment relative to AVGINV treatment. The above within-treatment and between-treatment comparisons also hold (at the 1% significance level) when the BCG1 numbers are considered.[24] This leads us to the following result.

---

[22]Perhaps higher types are more capable in anticipating that the auction will affect the selection of types entering the game – they adjust their BCG2 numbers to reflect this. In particular, the BCG2 numbers of higher types will be lower in the AVGINV relative to the AVG treatment. To study this, we only focus on subjects whose BCG1 numbers were at most $50p^2$, that is, the level-k model's prediction for the L_2 types and above. We find no significant between-treatment differences for the BCG2 numbers of these subjects (MW, P $= 0.08$) – mean AVG and AVGINV BCG2 numbers for these subjects are 17.54 (n $= 21$) and 22.49 (n $= 20$), respectively. If anything, BCG2 numbers are instead higher in the AVGINV relative to AVG treatment, which is contradictory to the idea that higher types anticipate the selection outcome.

[23]We do not find any significant between-session differences in bids for the AVG (Kruskal-Wallis, P $= 0.50$) and AVGINV (Kruskal-Wallis, P $= 0.36$) treatments. The auction transaction prices (i.e., the 9th highest bid price) in the three AVG treatment sessions are 37, 38 and 40. The corresponding prices in the three AVGINV treatment sessions are 60, 60 and 66.

[24]The correlation of bids and BCG1 numbers in the AVG and AVGINV treatments to be positive (Spearman $\rho = 0.53$, P < 0.01) and negative (Spearman $\rho = -0.23$, P $= 0.04$), respectively. Similarly, between-treatment comparisons find BCG1 numbers to be significantly (MW, P < 0.01) higher in the AVG treatment for auction winners and significantly lower in the AVG treatment for auction losers.

FIGURE 1.
Scatterplot of bids and BCG2 numbers.

AVG (n=72)

AVGINV (n=72)

Bid

BCG2 number

Spearman's rho=0.47[a]
Transaction prices: 37, 38, 40

Spearman's rho=-0.42[a]
Transaction prices: 60, 66, 60

Note. The hollow circles denote an observation per subject and the shaded lines denote the confidence interval of the best-fit line. The Spearman correlation ρ and the observation auction transaction prices (3 sessions per treatment) are reported on the bottom left of each panel.
[a] and [b] denote significance at $P < 0.01$ and $P < 0.05$, respectively.

FIGURE 2.
Boxplot distribution of BCG2 numbers for auction winners and losers.

BCG2 numbers

Auction Losers   Auction Winners

AVG

Auction Losers   Auction Winners
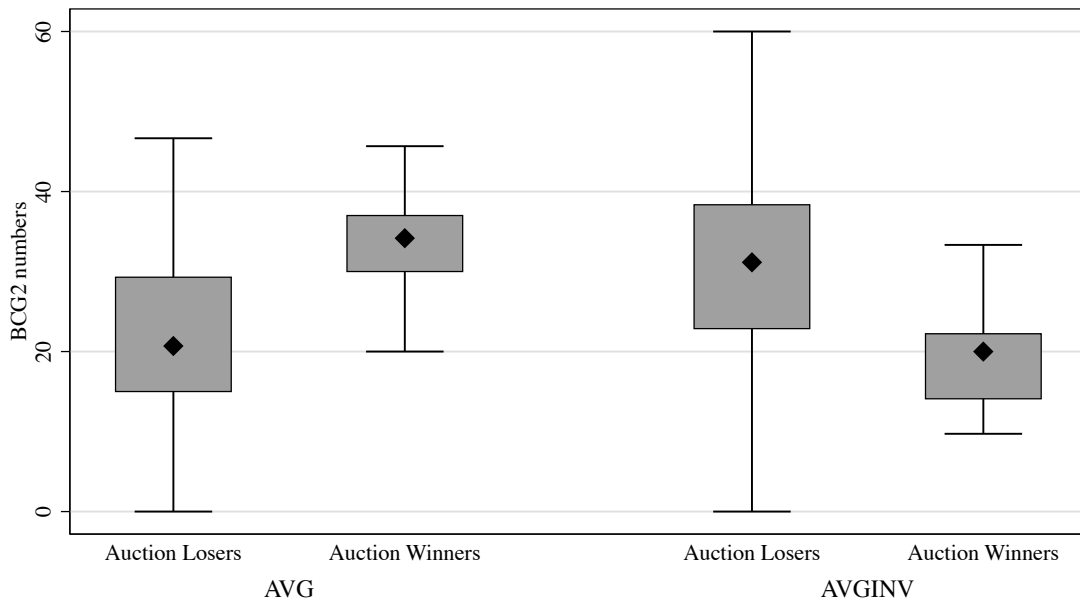
AVGINV

Note. Outliers ($n = 6$ observations) are omitted from the graph.

Result 1 *(H1a) BCG1 and BCG2 numbers in the AVG treatment are significantly higher for auction winners relative to auction losers. (H2a) BCG1 and BCG2 numbers in the AVGINV treatment are significantly lower for auction winners relative to auction losers. (H3a) BCG1 and BCG2 numbers for auction winners are significantly higher in the AVG treatment relative to the AVGINV treatment. (H4a) BCG1 and BCG2 numbers for auction losers are significantly lower in the AVG treatment relative to the AVGINV treatment.*

3.3. AUCTION AND LEVEL-K TYPES

Panel A of Figure 3 details the kernel density estimates of BCG2 numbers in the AVG and AVGINV treatments – the distributions appear to be bimodal with the modes around the L_1 and L_2 types' predicted choices. When only auction winners are considered (Panel B of Figure 3), the distribution of BCG2 numbers in the AVG and AVGINV treatments now seem to be unimodal with the mode at L_1 and L_2 types', respectively, predicted choice. Finally, when only auction losers are considered (Panel C of Figure 3), the AVG treatment's distribution seems unimodal with the mode at the L_2 types' predicted choice. In contrast, the AVGINV treatment's distribution is bimodal with the modes at L_1 and L_2 types' predicted choices.

The following details our strategy to infer subjects' L_k types from their BCG1 and BCG2 numbers.[25] The approach is based on the idea that whilst each L_k ($k > 0$) type might pick slightly different BCG1 and BCG2 numbers, both numbers will be close to the same predicted L_k type number – no such coherent behaviour is expected for the L_0 type.

3.3.1. How we assign level-k types

For each subject, let c^j be his p-beauty contest game number in stage $j = 1, 2$ and compute $\alpha_k^j = |c^j - 50p^k|$ for $k = 1, 2, ..., K$.[26] Here, α_k^1 (resp. α_k^2) is the absolute distance of the subject's BCG1 (resp. BCG2) number from the L_k ($k > 0$) type's predicted choice. Given the above, we derive the set $S^j = \{\alpha_1^j, \alpha_2^j, .., \alpha_K^j\}$ for each subject. Denote $s^j = \min S^j$ as the smallest element in set S^j.[27] Let θ^j be a subject's

---

[25]We have too few observations per subject to reliably estimate the type of each subject. We are hesitant to increase the number of p-beauty contest game played by each subject as this would complicate the auction design. Furthermore, increasing the number of games played may induce large learning effects – we would have need to specify a learning model.

[26]The level-k model does not impose a limit on the highest possible type. In such cases, it is also possible for behaviour to converge to the equilibrium as L_k types goes to infinity. For practical purposes, researchers often impose an arbitrary 'highest' type K.

[27]In all our discussions, the variable s^j is always observed to be unique.

FIGURE 3.
Kernel density estimates of BCG2 numbers.
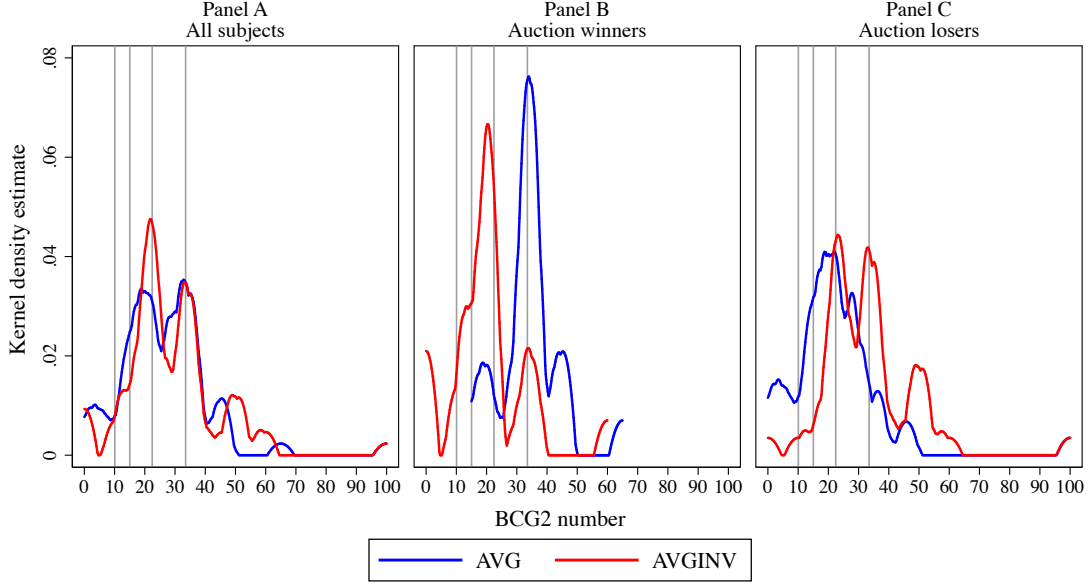
Note. Panels A, B and C detail the kernel density estimates for all subjects, auction winners and auction losers, respectively – we use a kernel bandwidth of 2. The vertical lines denote the 50p, $50p^2$, $50p^3$ and $50p^4$ points, the level-k model predicted choices for the L1, L2, L3 and L4 types, respectively.

pseudo-type in stage j:

$$
\theta^j = \begin{cases}
1 & \text{if } s^j = \alpha_1^j \text{ and } s^j \le e, \\
2 & \text{if } s^j = \alpha_2^j \text{ and } s^j \le e, \\
\vdots & \vdots \\
K & \text{if } s^j = \alpha_K^j \text{ and } s^j \le e, \\
K+1 & \text{if } c^j < 50p^K - e, \\
0 & \text{if otherwise,}
\end{cases}
$$

where $e > 0$ is some *tolerance* level. A subject is therefore classified as pseudo-type $\theta^j = k$ type in stage j if c^j is closest to $50p^k$ and $|c^j - 50p^k| \le e$ or otherwise $\theta^j = 0$. For a finite K, we simply cluster subjects' whose c^j is lower than $50p^K - e$ as pseudo-type $K + 1$. Given the above, a subject is assigned to be type $\hat{L}_k$ ($k > 0$) if $\theta^1 = \theta^2 = k$ or otherwise $\hat{L}_0$. In this case, a type $\hat{L}_0$ subject is one whose $\theta^1 = \theta^2 = 0$ or $\theta^1 \neq \theta^2$.[28]

---

[28]Suppose that $c^1 = 30$ and $c^2 = 28$. This subject will be assigned to be type $\hat{L}_1$ if $e = 6$ and $\hat{L}_0$ if $e = 4$.

13

The above approach is sufficient to assign the types for most subjects. However, it is plausible for subjects to 'learn' in the p-beauty contest game even when there is no feedback.[29] To allow for such learning, we apply a *correction* to our assignment of types. We reassign subjects type from $\hat{L}_0$ to $\hat{L}_g$, where $g = \theta^2$, if $\theta^2 - \theta^1 = 1$ (i.e., we assume that subjects can only perform a one step jump in L_k type).

How meaningful is our assignment of types? To shed light on this question, it is useful to return the model in Section 2.2. The model predicts that bids and BCG2 numbers are independent for the L_0 type. In contrast, bids and BGC2 numbers are positively and negatively correlated for L_k ($k > 0$) types in the AVG and AVGINV treatments, respectively. Also the model predicts that L_k ($k > 0$) types in AVG and AVGINV treatments expect the bonus amount to be $\mathbb{E}(X^k) = 50p^{k-1}$ and $\mathbb{E}(X^k) = 100 - 50p^{k-1}$, respectively. We hence compare the $\mathbb{E}(X^k)$ against $\mathbb{E}(X^B)$, subjects' elicited beliefs about the bonus amount, for L_k ($k > 0$) types.[30]

3.3.2. Assignment of level-k types

Panel A of Table 1 details the proportion of $\hat{L}_k$ types for K = 3 and *e* = 0, 2.5, 5, 7.5 and 10 – we also report the number of corrections performed.[31] Whilst we do not observe any significant between-treatment differences in the distribution of assigned types (Fisher exact, $P \geq 0.35$), the table demonstrates the tradeoffs with the choice of *e*. More specifically, decreasing *e* increases the *precision* of the assignment procedures (i.e., subjects have to choose closer to predicted L_k ($k > 0$) types' choice to be assigned as $\hat{L}_k$ type). However, decreasing *e* also increases the biases towards the $\hat{L}_0$ type (i.e., a greater proportion of subjects may be wrongly assigned as $\hat{L}_0$ type).

Panel B of Table 1 reports the Spearman correlation of bids and BCG2 numbers for $\hat{L}_0$ and $\hat{L}_1$-$\hat{L}_4$ types.[32] When $e \geq 5/2$, we observe bids and BGC2 numbers to be positively and negatively correlated (significant at the 1% level) for higher types in the AVG and AVGINV treatments. For $\hat{L}_0$ types, we only fail to reject (10% level) the null hypothesis of independence between bids and BCG2 numbers

---

[29]For example, a subject might behave like a L_1 type in the first stage and a L_2 type in the second stage.

[30]Recall that subjects submit their beliefs about the amount and receive a payoff of 10 points if their beliefs are within 3 integers of the realised bonus amount.

[31]We fix K = 3 as the distance between the L_k predicted choices above this threshold becomes incrementally small. In this case, all subjects whose c^j is less than $50p^3 - e$ will be assigned the pseudo-type $\theta^j = 4$.

[32]The correlation of bids and BCG1 numbers yield similar conclusions.

TABLE 1.
Summary statistics of assigned $\hat{L}_k$ types for all subjects.

	$e = 0$		$e = 2.5$		$e = 5$		$e = 7.5$		$e = 10$	
Treatment	AVG	AVGINV	AVG	AVGINV	AVG	AVGINV	AVG	AVGINV	AVG	AVGINV
♯ of correction	0	0	10	5	10	4	9	5	9	6
PANEL A: Proportion of $\hat{L}_k$ types										
$\hat{L}_0$	0.90	0.94	0.61	0.60	0.32	0.42	0.26	0.38	0.22	0.36
$\hat{L}_1$	0.00	0.00	0.17	0.21	0.28	0.22	0.31	0.26	0.33	0.28
$\hat{L}_2$	0.00	0.00	0.07	0.10	0.22	0.25	0.25	0.25	0.25	0.25
$\hat{L}_3$	0.00	0.00	0.07	0.06	0.13	0.07	0.13	0.08	0.15	0.08
$\hat{L}_4$	0.10	0.06	0.08	0.04	0.06	0.04	0.06	0.03	0.04	0.03
Fisher Exact P:	$P = 0.532$		$P = 0.787$		$P = 0.602$		$P = 0.591$		$P = 0.350$	
Panel B: Correlation of Bids and BCG2 number (Spearman ρ)										
$\hat{L}_0$ type:	0.34^a	-0.35^a	0.14	-0.30^b	0.10	-0.25	0.16	-0.28	0.04	-0.28
$\hat{L}_1 - \hat{L}_4$ types:	-0.35	0.23	0.73^a	-0.64^a	0.68^a	-0.59^a	0.60^a	-0.52^a	0.61^a	-0.50^a
Panel C: Proportion of $\hat{L}_1$-$\hat{L}_4$ types whose										
$\|\mathbb{E}(X^E) - \mathbb{E}(X^A)\| = 0$	0.00	0.00	0.04	0.07	0.02	0.05	0.02	0.04	0.02	0.04
$\|\mathbb{E}(X^B) - \mathbb{E}(X^k)\| \leq e$	0.71	0.25	0.50	0.35	0.41	0.48	0.43	0.49	0.40	0.50
♯ of subjects	7	4	28	29	49	42	53	45	56	46

Note. [a] and [b] denote $P < 0.01$ and $P < 0.05$, respectively.

TABLE 2.
Proportion of types $\hat{L}_1$-$\hat{L}_4$ for auction winners and losers.

	$e = 0$		$e = 5/2$		$e = 5$		$e = 15/2$		$e = 10$	
	(A)	(B)	(A)	(B)	(A)	(B)	(A)	(B)	(A)	(B)
PANEL A: Auction winners										
$\hat{L}_1$	0.00	0.00	0.82	0.22	0.80	0.14	0.76	0.20	0.79	0.20
$\hat{L}_2$	0.00	0.00	0.18	0.11	0.20	0.36	0.24	0.33	0.21	0.33
$\hat{L}_3$	0.00	0.00	0.00	0.33	0.00	0.29	0.00	0.33	0.00	0.33
$\hat{L}_4$	0.00	1.00	0.00	0.33	0.00	0.21	0.00	0.13	0.00	0.14
n	0	4	11	9	15	14	17	15	19	15
FE	NA		$P = 0.001^a$		$P = 0.001^a$		$P = 0.001^a$		$P = 0.001^a$	
PANEL B: Auction losers										
$\hat{L}_1$	0.00	0.00	0.18	0.65	0.24	0.50	0.25	0.53	0.24	0.55
$\hat{L}_2$	0.00	0.00	0.18	0.30	0.38	0.46	0.39	0.43	0.38	0.42
$\hat{L}_3$	0.00	0.00	0.29	0.05	0.26	0.04	0.25	0.03	0.30	0.03
$\hat{L}_4$	1.00	0.00	0.35	0.00	0.12	0.00	0.11	0.00	0.08	0.00
n	7	0	17	20	34	28	36	30	37	31
FE:	NA		$P = 0.005^a$		$P = 0.007^a$		$P = 0.006^a$		$P = 0.002^a$	

Note. (A) = AVG and (B) = AVGINV. We also report the Fisher exact (FE) p-value for between-treatment comparison of types.
[a] and [b] denote $P < 0.01$ and $P < 0.05$, respectively.

when $e \geq 5$.[33] Panel C of Table 1 details the proportion of $\hat{L}_1$-$\hat{L}_4$ types whose $|\mathbb{E}(X^B) - \mathbb{E}(X^k)| = 0$ and $|\mathbb{E}(X^B) - \mathbb{E}(X^k)| \leq e$. We note that the absolute difference are no more than the tolerance level e for around 40% of $\hat{L}_k$ ($k > 0$) types when $e \geq 5$. The above observations suggest that our assignment procedures may capture meaningful properties that are consistent with the level-k model in Section 2.2.

Panels A and B of Table 2 detail the observed proportions of $\hat{L}_1$-$\hat{L}_4$ types who are auction winners and losers, respectively, in the AVG and AVGINV treatments.[34]

Within-treatment comparisons ($e \geq 5/2$) find the proportion of $\hat{L}_k$ ($k > 0$) types amongst auction winners and losers to be significantly different in the AVG (Fisher exact, $P < 0.01$) and AVGINV (Fisher exact, $P < 0.01$) treatments. Between-treatment comparisons ($e \geq 5/2$) finds the proportion of $\hat{L}_k$ ($k > 0$)

---

[33]For all e considered, we do not find $\hat{L}_k$ types to be significantly correlated with cognitive reflective test performances or risk preferences.

[34]We omit the $\hat{L}_0$ to focus our comparisons on the L_k ($k > 0$) types. For $e \geq 5/2$, we nevertheless find no significant between-treatment differences in the proportion of $\hat{L}_0$ types amongst auction winners (Fisher Exact, $P \geq 0.34$) and losers (Fisher Exact $P \geq 0.28$).

types in the AVG and AVGINV treatments to be significantly different for auction winners (Fisher exact, P < 0.01) and losers (Fisher exact, P < 0.01). This leads us to our third result.

Result 2 *(H1b) $\hat{L}_k$ (k > 0) types in the AVG treatment are significantly lower for auction winners relative to auction losers. (H2b) $\hat{L}_k$ (k > 0) types in the AVGINV treatment are significantly higher for auction winners relative to auction losers. (H3b) $\hat{L}_k$ (k > 0) types for auction winners are significantly lower in the AVG treatment relative to the AVGINV treatment. (H4b) $\hat{L}_k$ (k > 0) types for auction losers are significantly higher in the AVG treatment relative to the AVGINV treatment.*

3.4. THE WINNER'S CURSE?

In common value auctions, the 'winner's curse' is the phenomenon where the winner of the auction overpays for the object – in equilibrium players avoid the winner's curse by shading their bids. In their analysis of second price common value auctions, Crawford and Iriberri (2007) show that it can be possible for higher L_k types to be less susceptible to the winner's curse. This would also be the case if higher L_k types are more reflective about the causes of a winner's curse.

Given the above, it is natural to ask if higher types are indeed less susceptible to the winner's curse. To study this, we use the fact that auction winner's L_k types are higher in the AVGINV treatment relative to the AVG treatment. For each auction winner, denote π as the bonus amount received minus the auction ticket price paid plus the expected value of the p-beauty contest game, which is 15/8. We find that $\pi < 0$ (Signrank P = 0.001) and $\pi > 0$ (Signrank P < 0.001) for auction winners in the AVG and AVGINV treatments, respectively – mean π for auction winners in the AVG and AVGINV treatments are -4.54 and 10.02, respectively. These observations imply that only auction winners in the AVG treatment suffer from the winner's curse.

The results insofar indicate that the proportion of L_k types in the AVG and AVGINV treatments are similar. Do they exhibit similar bidding behaviours? The comment discussions on section 2.2 suggest that L_2 types and above in the AVGINV treatment may shade their bids more aggressively than those in the AVG treatment to account for the possibility of the winner's curse.[35] To see this we compute γ as the subjects' bids in the AVG treatment. For subjects in the AVGINV treatment, we compute γ as 100 minus their bids. If bidding behaviours

---

[35]In the AVGINV treatment, the winner's curse occurs when the L_k type overestimates the types of other players and hence overpays for the ticket.

FIGURE 4.
Cumulative density function of γ.
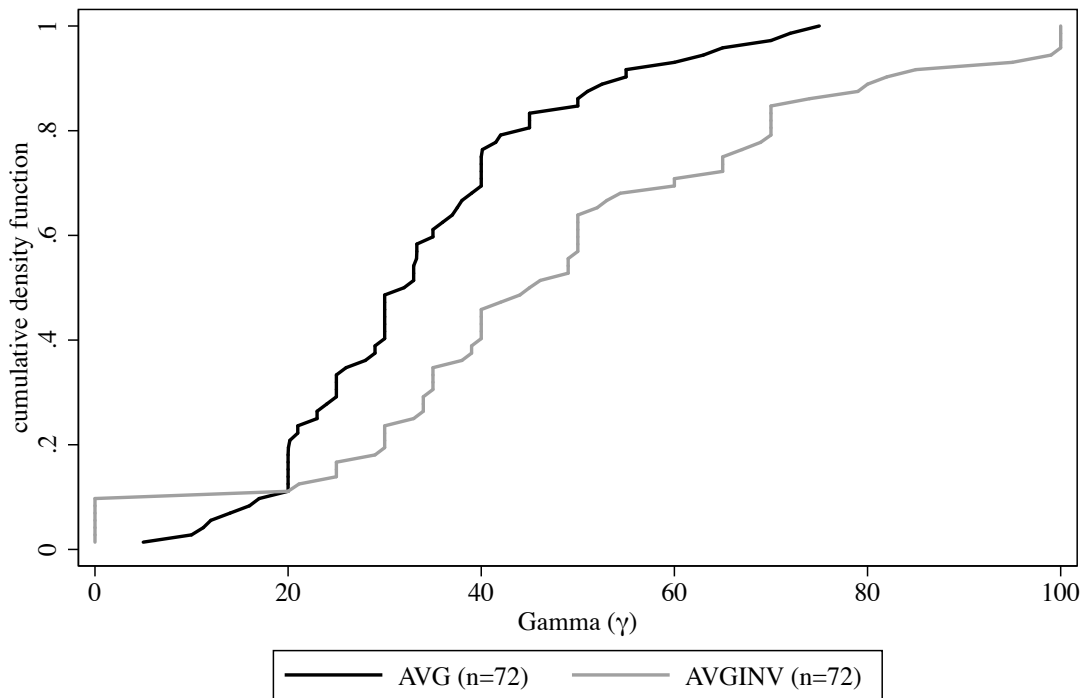
Note. γ is computed as subjects' bid in the AVG treatment and 100 minus bids in the AVGINV treatment.

are similar across both treatments, we also expect the distribution of γ to be similar across both treatments. Figure 4 details the cumulative density function of γ in the respective treatments. We find γ to be significantly higher (MW, P < 0.001) in the AVGINV treatment relative to the AVG treatment, suggesting that subjects in the AVGINV treatment may be shading their bids more aggressively than those in the AVG treatment.

3.5. ROBUSTNESS

Central to Results 1 and 2 is the idea that a 'purposeful' auction mechanism is able to separate subjects by their level-k types. This implies that 'non purposeful' auction mechanism should therefore fail to separate subjects by their level-k types. To test this, we conducted the RAND treatment (2 sessions, 48 subjects) which differs from the AVG and AVGINV treatments only in that the bonus is now a random (uniform) variable between 0 and 100.[36] As a consequence, the

---

[36]Subjects were informed that the bonus X is a random number between 0 and 100.

auction mechanism will no longer be able to select people by their level-k types.[37]

Between-treatment comparisons does not find BCG1 (Kruskal-Wallis, P = 0.71) and BCG2 (Kruskal-Wallis, P = 0.49) numbers in the RAND treatment to be significantly different to those reported in the AVG and AVGINV treatments. However, we find no significant correlation in bids and BCG2 numbers (Spearman ρ = −0.05, significance P = 0.70) for subjects in the RAND treatment.[38] As a consequence, we do not find BCG1 (MW, P = 0.81) and BCG2 (MW, P = 0.45) numbers to be significantly different for auction winners and losers in the RAND treatment. These observations clearly show that a 'non purposeful' auction mechanism is unable to select subjects by their level-k types.

4. CONCLUSION

This study investigates whether auctions can select people by their level-k types. To do so, we embed auctions into the p-beauty contest game. Two treatments are considered, the AVG and AVGINV treatments. Using the level-k model we show that auctions should select lower and higher L_k (k > 0) types to play the p-beauty contest game in the AVG and AVGINV treatments, respectively. Our findings provide robust evidence inline with this prediction. For robustness, we also consider a treatment where the auctions not expected to select people by their level-k types. Again, our experiment provides evidence inline with this prediction.

Taken together, we find that auctions can indeed select people by their level-k type. As mentioned in the Introduction, this has important implications for economic designers since it implies that it might be possible to influence outcomes in any game by using auctions to select level-k types. We hope to build on this findings and extend the research on market selection to our games commonly associated with the level-k model.

Many economies use markets to allocate decision making tasks (e.g., licenses, permits, firms). In most case, the elimination of behavioural anomalies in such task requires higher level-k types. However, our findings suggest that market competition does not necessary eliminate behavioural anomalies in such task. If the lower level-k types value the decision making rights more than the higher level-k types, it is possible that market competition might actually exacerbate

---

[37]Whilst the L_0 type will still bid randomly, L_k (k > 0) types will now bid 50 + 15m, where 50 is the expected bonus payment (mean of the random variable) and m is the expected probability of winning the second p-beauty contest game prize.

[38]The ticket transaction price in the two RAND treatment sessions are 50 and 51.

behavioural anomalies.

REFERENCES

Bosch-Domènech, Antoni, José G. Montalvo, Rosemarie Nagel, and Albert Satorra, "One, two, (three), infinity, ... : Newspaper and lab beauty-contest experiments," *The American Economic Review*, 2002, *92* (5), 1687–1701.

Brañas-Garza, Pablo, Teresa García-Muñoz, and Roberto Hernín González, "Cognitive effort in the beauty contest game," *Journal of Economic Behavior & Organization*, 2012, *83* (2), 254–260.

Broseta, Bruno, Enrique Fatas, and Tibor Neugebauer, "Asset markets and equilibrium selection in public goods games with provision points: An experimental study," *Economic Inquiry*, 2007, *41* (4), 574–591.

Brown, Alexander L., Colin F. Camerer, and Dan Lovallo, "To review or not to review? Limited strategic thinking at the movie box office," *American Economic Journal: Microeconomics*, May 2012, *4* (2), 1–26.

Burnham, Terence C., David Cesarini, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace, "Higher cognitive ability is associated with lower entries in a p-beauty contest," *Journal of Economic Behavior & Organization*, 2009, *72* (1), 171–175.

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, 2004, *119* (3), 861–898.

Carpenter, Jeffrey, Michael Graham, and Jesse Wolf, "Cognitive ability and strategic sophistication," *Games and Economic Behavior*, 2013, *80*, 115–130.

Costa-Gomes, Miguel A. and Vincent P. Crawford, "Cognition and behavior in two-person guessing Games: An experimental study," *American Economic Review*, December 2006, *96* (5), 1737–1768.

Costa-Gomes, Miguel, Vincent P. Crawford, and Bruno Broseta, "Cognition and behavior in normal-form games: An experimental study," *Econometrica*, 2001, *69* (5), 1193–1235.

Crawford, Vincent and Bruno Broseta, "What price coordination? The efficiency-enhancing effect of auctioning the right to play," *The American Economic Review*, 1998, *88* (1), 198–225.

Crawford, Vincent P. and Nagore Iriberri, "Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?," *Econometrica*, 2007, *75* (6), 1721–1770.

__ , **Miguel A. Costa-Gomes, and Nagore Iriberri**, "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications," *Journal of Economic Literature*, 2013, *51* (1), 5–62.

Duffy, John and Rosemarie Nagel, "On the robustness of behaviour in experimental 'beauty contest' games," *The Economic Journal*, 1997, *107* (445), 1684–1700.

Fischbacher, Urs, "z-Tree: Zurich toolbox For ready-made economic experiments," *Experimental Economics*, 2007, *10*, 171–178.

Frederick, Shane, "Cognitive reflection and decision making," *Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

Georganas, Sotiris, "English auctions with resale: An experimental study," *Games and Economic Behavior*, 2011, *73* (1), 147–166.

__ , **Paul J. Healy, and Roberto A. Weber**, "On the persistence of strategic sophistication," *Journal of Economic Theory*, 2015, *159*, 369–400.

Gill, David and Victoria Prowse, "Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis," *Journal of Political Economy*, 2016, *124* (6), 1619–1676.

Goeree, Jacob K. and Charles A. Holt, "A model of noisy introspection," *Games and Economic Behavior*, 2004, *46* (2), 365–382.

Greiner, Ben, "Subject pool recruitment procedures: Organizing experiments with ORSEE," *Journal of the Economic Science Association*, 2015, *1* (1), 114–125.

Grosskopf, Brit and Rosemarie Nagel, "The two-person beauty contest," *Games and Economic Behavior*, 2008, *62* (1), 93 – 99.

Haigh, Matthew, "Has the Standard Cognitive Reflection Test Become a Victim of Its Own Success?," *Advances in Cognitive Psychology*, 2016, *12* (3), 145–149.

Holt, Charles A. and Susan K. Laury, "Risk aversion and incentive effects," *American Economic Review*, December 2002, *92* (5), 1644–1655.

Huyck, John B. Van, Raymond C. Battalio, and Richard O. Beil, "Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication," *Games and Economic Behavior*, 1993, *5* (3), 485–504.

Kaplan, Todd R. and Bradley J. Ruffle, "The self-serving bias and beliefs about rationality," *Economic Inquiry*, 2004, *42* (2), 237–246.

Kawagoe, Toshiji and Hirokazu Takizawa, "Level-k analysis of experimental centipede games," *Journal of Economic Behavior & Organization*, 2012, *82* (2), 548–566.

Nagel, Rosemarie, "Unraveling in guessing games: An experimental study," *The American Economic Review*, 1995, *85* (5), 1313–1326.

Offerman, Theo and Jan Potters, "Does auctioning of entry licences induce collusion? An experimental study," *The Review of Economic Studies*, 2006, *73* (3), 769–791.

Östling, Robert, Joseph Tao yi Wang, Eileen Y. Chou, and Colin F. Camerer, "Testing game theory in the field: Swedish LUPI lottery games," *American Economic Journal: Microeconomics*, August 2011, *3* (3), 1–33.

Selten, Reinhard, "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," *Beiträge Zur Experimentellen Wirtschaftsforschung*, 1967, pp. 136–168.

Shachat, Jason and J. Todd Swarthout, "Auctioning the right to play ultimatum games and the impact on equilibrium selection," *Games*, 2013, *4* (4), 738–753.

Stahl, Dale O. and Paul W. Wilson, "Experimental evidence on players' models of other players," *Journal of Economic Behavior & Organization*, 1994, *25* (3), 309–327.

_ **and** _ , "On players' models of other players: Theory and experimental evidence," *Games and Economic Behavior*, 1995, *10* (1), 218–254.

Weber, Roberto A., "'Learning' with no feedback in a competitive guessing game," *Games and Economic Behavior*, 2003, *44* (1), 134–144.

APPENDIX

A. INSTRUCTIONS

Remark: Below is the translated version of the instructions (the instructions were written in Mandarin). Where necessary, we use "*text*" and "<mark>text</mark>" to distinguish between the instructions for the AVG and AVGINV, respectively. To avoid confusing the subjects, the instructions for Part II (stage 2) do not mention the strategy method. In the experiment, subjects first submitted their bids. Thereafter, they were asked to submit their p-beauty contest number if they purchased a ticket. Finally, subjects were informed about the beliefs elicitation payment (e.g., their beliefs about the bonus) at the end of part II.

A.1. GENERAL INFORMATION

Thank you for participating in this study. Please note that you are not allowed to talk with other participants during the study. If you have a question, please raise your hand and we will answer your question in private. In order to minimize distractions, please turn off your mobile phone and put away anything else that could distract you from the study, e.g. books, study notes or electronic devices. You are only allowed to use the functions of the computer pertained to the current study. Note that violation of these rules may lead to an immediate exclusion from the study and from all payments. At no time during this study will you learn the identity of the other participants and no other participants will learn anything about your identity. Also, no other participant will learn what you earn during the study: At the end of the study, the amount of money you have earned will be paid out to you in private. Hence, no other participant will know your choices and how much money you earn in this study. The experiment will consist of two parts (Part I and Part II). Your earnings in this experiment will depend on your decisions in Parts I and II. During the study, we do not speak of RMB, but in points. Your entire earnings will be calculated in points. At the end of the study the total amount of points you have earned will be converted to RMB at the following rate:

$$1 \text{ point} = 0.45 \text{ RMB}$$

Please read the instructions carefully, because your earnings in the study depends on how well you understand the instructions. If you follow the instructions and apply them carefully, you can earn some money in additional to the 10 RMB

show-up fee which we will give you in any case. Therefore, your total earning in today's study is:

Total Earning = Show-up Fee + Money Earned in Part I + Money Earned in Part II

In the following, we present the instructions for Part I of the experiment. The Part II instructions will be available at the end of Part I.

A.2. INSTRUCTION PART I

There are 24 participants in this room. You will be randomly matched with 7 other participants to form a group and play the Number-Game. There are 8 participants per group.

How the Number-Game works? Each participant chooses a number between 0 and 100 (including both 0 and 100; up to two decimal places). After everyone has chosen a number, we will compute a "Target-Number". The Target Number is calculated as follows:

1. We first calculate the group average, which is the average of all 8 numbers chosen by the participants in your group.

2. Next, we multiply the group average by 2/3 to get the target number

$$\text{Target Number} = 2/3 * \text{Group Average}$$

The winner in your group is the participant whose chosen number is closest to the target number.

The winner will receive a prize of 15 points. In the case of a tie (i.e., where there are more than one winners), the prize money will be equally split between all winners.

Some important points: It is important to note that you will never learn the identity of other group members and others will never know the identity of you. Also, all participants choose their numbers simultaneously. This means when you are making your decision, you do not know the number chosen by other group members. Similarly, when other group members are making their decisions, they also do not know the number chosen by you.

24

A.2.1. Control questions to part I

1. How many participants are there in each group?

2. Do you and your group members choose the number simultaneously? In other words, do you know each other's chosen number when you making decisions? (True/False)

3. The target number is computed by the group average multiplied by 2/3. (True/false)

4. The participant whose number is closest to the target number will win the prize. (True/False).

5. Suppose the numbers chosen by one group are: 10, 20, 30, 30, 50, 60, 80, 100:

   - What is the group average?
   - What is the Target Number?
   - Which number wins the game?
   - How many points the winners get?

A.3. INSTRUCTION PART II

In part II of the experiment, there are 8 tickets to be sold. Participants who bought a ticket will receive a payment of X points for sure (we call this the Redemption-Amount). In addition, each ticket holder is allowed to participate in another round of the Number-Game and hence have a chance win a bonus of 15 points. In the following, we will first explain how the tickets are sold. Thereafter, we will explain the Redemption-Amount. Finally, we again briefly explain the number game.

A.3.1. How are tickets sold

There are 24 participants in this experiment and only 8 tickets to be sold. To determine who gets a ticket, each participant will be endowed with 100 points and submits a "bid".

- Bid = the maximum amount you are willing to pay for a ticket

After all participants have submitted their bid, we will rank all 24 bids and sell the eight tickets to the participants with the highest 8 bids. However, the ticket holders only need to pay the ninth highest price. This means that if your bid is amongst the eight highest bids, you will never pay more than your bid amount and can possibly pay less than your bid amount.

Here is a simple example to help you better understand how the tickets are sold. Here, there are 24 participants (A-X). Each participant submits a bid and we rank the bids.

In Example 1, participants Q, R, S, T, U, V, W and X will get to buy a ticket. However, they each will only pay the 9th highest price, which is 12 points.

EXAMPLE 1.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Bid	1	2	3	4	5	6	7	8	9	9	9	10	10	11	11	12	13	14	15	15	17	18	18	19
Rank																9	8	7	5	5	4	2	2	1
Buy ticket																	X	X	X	X	X	X	X	X

There can also be situations where participants submit the same bid price. In example 2, participants S, T, U, V, W and X will get to buy a ticket since they are the 6 highest bids. However, there is a tie between participants P, Q and R for the 7th and 8th highest bids. In this case, we randomly pick two participants to buy the ticket – participants P and R in the example. All participants who get to buy a ticket will only pay 12 points.

EXAMPLE 2.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Bid	1	2	3	4	5	6	7	8	9	9	9	10	10	11	11	12	12	12	15	15	17	18	18	19
Rank																7	7	7	5	5	4	2	2	1
Buy ticket																X		X	X	X	X	X	X	X

A.3.2. How the Redemption-Amount is computed

If you purchased a ticket, your ticket will be redeemed at an amount which we call the Redemption-Amount. Again, you can think about the redemption-amount as the price which we buy back the ticket from you at. Recall that there were 24 participants who submitted their chosen number to the Number-Game in Part I of the Experiment. This results in a total of 24 chosen numbers. To determine Redemption-Amount, we will first randomly pick eight numbers from the 24 chosen numbers in part I. Thereafter, we will compute the average of the 8 randomly chosen numbers. The Redemption-Amount is computed as follows:

Redemption-Amount = (Average of the 8 randomly chosen number)

Redemption-Amount = 100- (Average of the 8 randomly chosen number)

Note that you will only learn about the Redemption-Amount at the end of the experiment. This means that when you submit your bid, you do not know the redemption-Amount.

Remark: For the RAND treatment, we omit the text in red. Subjects were informed that the Redemption-Amount is a random number between 0 and 100.

A.3.3. Playing the Number-Game

If you purchased a ticket, you can participate in the Number-Game. The Number-Game is exactly like the description in Part I. Briefly, 8 participants each choose a number between 0 and 100. Given the chosen numbers in the Number-Game, we calculate a Target-Number which is 2/3 of the average of all 8 chosen numbers in the Number-Game. The participants whose chosen number is closest to the Target-Number receives a prize of 15 points. In the case of a tie, the prize money will be equally split between all winners.

A.3.4. Payoff in Part II

Your payoff in Part II will depend on (a) Whether you purchased a ticket, (b) The Redemption-Amount if you purchased a ticket and (c) The Number-Game winner if you purchased a ticket. If you did not purchase a ticket, your payoff is simply your endowment

$$\text{Payoff} = 100 \text{ points}$$

If you purchased a ticket and is a winner in the Number-Game, your payoff is:

$$\text{Payoff} = 100 - (\text{Ticket Price}) + (\text{Redemption-Amount}) + (15) \text{ points}$$

If you purchased a ticket and is NOT a winner in the Number-Game, your payoff is:

$$\text{Payoff} = 100 - (\text{Ticket Price}) + (\text{Redemption-Amount})$$

A.3.5. Control questions to part II

1. Will everyone be able to get a ticket? (Yes/No)

2. How many tickets will be sold?

3. What is the amount of the lowest bid that you can make?

4. What is the amount of the highest bid that you can make?

5. If you purchase a ticket, will the ticket price ever be higher than your bid? (True/False)

6. If you purchase a ticket, can it be possible for the ticket price to be lower than your bid? (True/False)

7. Will you be informed about the Redemption Amount before you purchase your ticket?

8. Suppose the eight randomly selected numbers from the 24 numbers in part I are 10, 20, 30, 40, 50, 60, 70, 80.

   - What is the Redemption Amount?

   - Suppose you bought the ticket with a cost of *40* 50 points and win the second Number game, how many points do you earn in part II?

   - Suppose you bought the ticket with a cost of *40* 50 points and do not win the second Number game, how many points do you earn in part II?

B. SUMMARY STATISTICS

Panels A and B of Table B1 detail the summary statistics of BCG1 and BGC2 numbers, respectively, in the AVG, AVGINV and RAND treatments. Each cell reports the summary statistics for all subjects, followed by auction winners (circle brackets) and auction losers (square brackets). As expected, we do not find any significant between-treatment differences in BCG1 (Kruskal-wallis (KW), P = 0.71) and BCG2 (KW, P = 0.49). We see that choices in the p-beauty contest game are clearly different from the equilibrium – only a small proportion of subjects choose 0 or numbers below 10.

TABLE B1.
Summary statistics.

Treatment	AVG	AVGINV	RAND
♯ of subjects	72 (24) [48]	72 (24) [48]	48 (16) [32]
Panel A: BCG1 numbers.			
Median BCG1 number	26 (35) [25]	28 (25) [32]	25 (25) [27]
Frac. BCG1 = 0	1% (0%) [2%]	6% (12%) [2%]	2% (0%) [3%]
Frac. BCG1 ≤ 5	6% (0%) [6%]	7% (12%) [4%]	4% (0%) [6%]
Frac. BCG1 ≤ 10	6% (0%) [10%]	8% (16%) [4%]	6% (0%) [9%]
Panel B: BCG2 numbers.			
Median BCG2 number	26 (34) [21]	25 (20) [31]	25 (25) [25]
Frac. BCG2 = 0	2% (0%) [2%]	6% (13%) [2%]	4% (0%) [6%]
Frac. BCG2 ≤ 5	7% (0%) [10%]	6% (13%) [2%]	4% (0%) [6%]
Frac. BCG2 ≤ 10	11% (0%) [17%]	8% (16%) [4%]	14% (6%) [18%]
Panel C: Cognitive Reflective test score[i].			
Mean	2.68 (2.42) [2.81]	2.61 (2.62) [2.60]	2.83 (2.75) [2.88]
Panel D: Risk preference[ii].			
Mean	5.57 (5.62) [5.54]	5.62 (5.48) [2.60]	5.79 (6.25) [5.56]
Panel E: Subjects Δ = BCG1 − BCG2.			
Median Δ	0.05 (0.00) [0.12]	0.00 (2.89) [0.00]	0.00 (0.00) [0.25]
Frac. |Δ| = 0	18% (16%) [18%]	28% (13%) [35%]	27% (31%) [25%]
Frac. |Δ| ≤ 5	58% (33%) [70%]	61% (58%) [64%]	83% (31%) [81%]
Frac. |Δ| ≤ 10	80% (70%) [85%]	80% (80%) [80%]	93% (100%) [90%]

Note. Each cell details the summary statistics for all subjects, followed by the auction winners (circle brackets) and auction losers (square brackets).
[i] Maximum score is 3.
[ii] Risk preference is a scale from 1-12. The risk-neutral scale is XX.

Panels C and D of Table B1 detail the mean cognitive reflective test (KS, P = 0.28) performances and elicited risk preferences (KW, P = 0.55) – we again do not find any significant between-treatment differences. There does not seem to be

any notable within-treatment differences in cognitive reflective test performances and risk preferences amongst auction winners and losers.[39] We note that subjects' cognitive reflective test performances in all treatments seem to be somewhat high compared to the literature.[40] We cannot exclude the possibility that the subject pool might have prior exposure to the cognitive reflective test in previous experiments or classes.

Finally, Panel E of Table B1 details the median $\Delta = BCG1 - BCG2$ in the all treatments as well as the proportion of subjects who absolute value of Δ is 0, at most 5 and at most 10. We do not find any significant between-treatment differences in Δ (KW, P = 0.89). Furthermore, within-treatment comparison does not find Δ to differ significantly amongst auction winners and losers (Mann-whitney, P $\geq$ 0.18 in all treatments).

---

[39]Only in the AVG treatment do we find the cognitive reflective test performances to be significantly lower for auction winners relative to losers (Mann-Whitney, P = 0.02).

[40]In general, the literature finds that mean performances between 1.5-2.