Simultaneous Versus Sequential All-Pay Auctions: An Experimental Study

Lian Jian, Zheng Li, and Tracy Xiao Liu*†

November 24, 2016

Abstract

While both simultaneous and sequential contests are mechanisms used in practice such as crowdsourcing, job interviews and sports contests, few studies have directly compared their performance. By modeling contests as incomplete information all-pay auctions with linear costs, we analytically and experimentally show that the expected maximum effort is higher in simultaneous contests, in which contestants choose their effort levels independently and simultaneously, than in sequential contests, in which late entrants make their effort choices after observing all prior participants' choices. Our experimental results also show that efficiency is higher in simultaneous contests than in sequential ones. Sequential contests' efficiency drops significantly as the number of contestants increases. We also discover that when participants' ability follows a power distribution, high ability players facing multiple opponents in simultaneous contests tend to under-exert effort, compared to theoretical predictions. We explain this observation using a simple model of overconfidence.

Keywords: sequential all-pay auctions, experiment
JEL Classification: C7, C91

*Jian: University of Southern California, Los Angeles, CA, U.S.A. Email: ljian@usc.edu. Li: Tsinghua University, Beijing, China. Email: lizh4.12@sem.tsinghua.edu.cn. Liu: Tsinghua University, Beijing, China. Email: liuxiao@sem.tsinghua.edu.cn.

†We thank Juan Carrillo and Isabelle Brocas for providing us with access to the Los Angeles Behavioral Economics Laboratory (LABEL). We would also like to thank Yan Chen, Jeffrey Mackie-Mason, Lijia Wei and seminar participants at the 2013 North American ESA meetings, the 2013 Annual Xiamen University International Workshop on Experimental Economics, and the 2014 Asian-Pacific ESA meetings for helpful discussions and comments, and Chao Tang for excellent research assistance. We thank the Editor, David Cooper, and two anonymous referees for their constructive comments and suggestions. The financial support from Tsinghua University, the National Natural Science Foundation of China (NSFC) under grant 71403140 and 71432004, is gratefully acknowledged. Lian Jian thanks the financial support from the APOC program at the Annenberg School of Communication, University of Southern California.

1 Introduction

Contests are effective mechanisms for eliciting individual effort in the workplace (Amegashie et al., 2007; Che and Gale, 2003; Dechenaux et al., 2015). Naturally, some contests, e.g., rentseeking (Tullock, 1980), can be modeled as simultaneous games, since the contestants choose their effort independently and simultaneously. Other contests are better modeled as sequential games, e.g., litigations (Baye et al., 2005) and elections (Morgan, 2003), since the participants enter the contests at different times and late contestants make their effort choices after observing prior contestants' efforts.

While a contest holder often has a limited choice in the format (simultaneous versus sequential) of the contest due to the nature of their problems, in many situations, the contest format is a possible decision variable. In principle, many sports contests (such as figure skating, diving and gymnastics) can be conducted either simultaneously or sequentially. Although in practice participants may perform one after another, as long as each participant's performance outcome is not revealed to other participants, it is a simultaneous contest. Conversely, in a sequential contest, such information is revealed and can potentially affect subsequent participants' performance, or even strategy (e.g., picking a different routine). Therefore, for an organizer whose objective is to break as many records as possible, it is useful to know if picking one of these contest mechanisms versus the other can make a difference. Further, online crowdsourcing contests for solving specific engineering or design problems can also be implemented as either simultaneous (Jeppesen and Lakhani, 2010) or sequential contests (Liu et al., 2014), and on online contest-hosting platforms switching between simultaneous and sequential could be as easy as a mouse click (e.g., on 99designs.com). These examples demonstrate the coexistence of simultaneous and sequential contests in practice, and the need to study these mechanisms simultaneously to guide practitioners.

To the best of our knowledge, little research exists to guide the choice between these two contest mechanisms, as most studies have examined them separately. A handful of studies have compared them (e.g., Fu, 2006; Konrad and Leininger, 2007; Leininger, 1993; Liu, 2015; Morgan, 2003), but they have either focused on contests with only two players (Fu, 2006; Leininger, 1993; Morgan, 2003) or have assumed a complete information environment (Dixit, 1987; Konrad and Leininger, 2007; Liu, 2015). This study is aimed at filling this gap by characterizing n-player n-stage sequential contests under incomplete information and comparing their performance with simultaneous contests.

By modeling contests as all-pay auctions under incomplete information with linear costs, we analytically compare simultaneous and sequential contests and experimentally test the theoretical predictions. Consistent with our theory predictions, simultaneous contests are superior to sequential contests in eliciting higher maximum effort and total effort. Simultaneous contests also achieve higher efficiency than sequential contests. Moreover, we discover that as the number of contestants increases, efficiency decreases in sequential contests, further deteriorating their performance relative to simultaneous contests.

Our experiment also reveals behavioral patterns in simultaneous contests that have not been highlighted in prior research. While high ability participants have been previously observed to over-exert effort compared to theoretical predictions, in our experimental environment with more than two players and a power distribution, high ability participants exhibit significant under-exertion. We explain this observation by overconfidence and use a simple model which utilizes belief data to provide supporting evidence.

2 Literature Review

Contests in which participants exert effort or spend resources to win an object, e.g., a prize, have been comprehensively studied in the economics literature (see Dechenaux et al., 2015, and Konrad, 2009, for comprehensive reviews). This literature suggests three canonical models: Tullock contests or lottery contests (Tullock, 1980), rank-order tournaments (Lazear and Rosen, 1981), and all-pay auctions (Nalebuff and Stiglitz, 1983; Dasgupta, 1986; Hillman and Riley, 1989). Terwiesch and Xu (2008) model various tasks in contests based on the relative importance of expertise and the degree of uncertainty. In this study, we focus on expertise-based contests (Terwiesch and Xu, 2008), in which tasks are well defined, output is mainly determined by effort, and the evaluation process has little noise. These features are well described by all-pay auctions (the focus of this study) in which there is no noise in individuals' production functions (unlike in rank-order tournaments) and the player with the highest effort wins with certainty (unlike in Tullock or lottery contests).[1]

Most extant theoretical work has focused on contests modeled as simultaneous games, with a few that have explored alternative time protocols that consider the dynamic nature of contests in practice. One line of studies involves contests with multiple stages. In these, at least some participants can make multiple moves. For example, in multi-stage elimination contests (Altmann et al., 2012; Fu and Lu, 2012; Gradstein and Konrad, 1999; Rosen, 1986; Zhang and Wang, 2009), each stage is a simultaneous contest among the eligible participants and only winners advance to the next stage. In races (Harris and Vickers, 1985), the same players take turns to make moves. In best-of-N contests (Harris and Vickers, 1987; Mago et al., 2013), participants engage in a series of simultaneous contests until one wins a certain number of times. Another line of studies focuses on sequential contests (Dixit, 1987; Fu, 2006; Konrad and Leininger, 2007; Leininger, 1993; Morgan, 2003; Segev and Sela, 2014), in which each participant only makes a single move. They follow a sequence, and all participants can view prior participants' choices before they make their own decisions. Some sequential models divide n participants into two stages (an early and a late) (Dixit, 1987; Konrad and Leininger, 2007), and other models allow only one participant to move in each stage (Fu, 2006; Leininger, 1993; Morgan, 2003; Segev and Sela, 2014). In our study, we focus on sequential contests of the latter type, i.e., contests with n players and n stages.

Prior analytical work in sequential contests varies in a number of aspects. First, the number of

[1]We assume that each participant's effort unambiguously determines the quality of the submission, and that the quality is objectively quantifiable. Therefore, we use the words "bid," "quality," and "effort" interchangeably throughout the paper. We also use "participant," "contestant," and "player" as synonyms.

players vary, i.e. two players (Morgan, 2003; Fu, 2006; Leininger, 1993) versus n players (Dixit, 1987; Konrad and Leininger, 2007; Segev and Sela, 2014). Second, the winning probability functions vary. Lottery contests (Morgan, 2003; Fu, 2006; Leininger, 1993), all-pay auctions (Konrad and Leininger, 2007; Segev and Sela, 2014), and contests with general winning probability functions (Dixit, 1987) have all been investigated in a sequential setting. Third, in determining the sequence of play, studies vary in assuming exogenous timing (Dixit, 1987; Segev and Sela, 2014), endogenous timing (Fu, 2006; Leininger, 1993; Konrad and Leininger, 2007), or allowing both (Morgan, 2003). And last, in terms of the information environment, most have assumed complete information (Morgan, 2003; Leininger, 1993; Konrad and Leininger, 2007; Dixit, 1987) while some have assumed incomplete information (Fu, 2006; Segev and Sela, 2014).

One important question when studying sequential contests is to compare them with simultaneous contests. Most such comparisons have focused on the total effort (bids) a contest receives. When players are asymmetric in values (or information or abilities) and can endogenously choose their move timing after learning their values, previous theoretical analysis predicts that sequential contests should generate lower total effort than simultaneous contests (see Leininger, 1993, and Fu, 2006, for lottery contests; and Konrad and Leininger, 2007, for all-pay auctions). This is because given one's own value, the weaker one will choose to move first and the stronger one will move last. To "calm-down" the stronger one, the weaker one will then reduce her effort compared to simultaneous games, which will likely lead to a reduction in the stronger player's effort, thereby reducing the overall total effort compared to simultaneous games. In contrast, participants in Morgan's (2003) model have to choose their timing of move before learning their values. In this context, sequential lottery contests generate higher total effort than simultaneous ones and achieve higher allocative efficiency.

Among experimental studies of all-pay auctions, the one most closely related to ours is by Liu (2015). Assuming a complete information environment, she focuses on two-player all-pay auctions in which participants can endogenously choose their bid timing. She finds that the revenue in simultaneous all-pay auctions is at least as high as that in sequential ones. In contrast, our study models contests as n-player all-pay auctions under incomplete information, which is arguably more relevant to many real-world scenarios in which there are multiple contestants, all of which have limited knowledge about one another's private abilities or types. Furthermore, our study investigates the size effect, i.e., whether auction outcomes vary based on the number of bidders. A few other studies have also tested equilibrium predictions of simultaneous all-pay auctions, without comparing them to sequential ones. Compared to theoretical predictions that assume risk-neutral contestants, experimental studies of incomplete information simultaneous all-pay auctions have reported that high ability individuals overbid and low ability ones underbid (Barut et al., 2002; Muller and Schotter, 2010; Noussair and Silver, 2006). An overbidding phenomenon is also observed in complete information all-pay auctions (Davis and Reilly, 1998; Gneezy and Smorodinsky, 2006). By assuming individual bounded rationality, Anderson et al. (1998) propose a logit equilibrium model to explain it. Typically, these experiments employ relatively large contest sizes, e.g., $n = 6$ (Noussair and Silver, 2006), and use uniform distribution of ability factors if the game is studied under incomplete information. In a few experiments with small contest sizes, e.g., $n = 2$ (Grosskopf et al., 2010; Potters et al., 1998), however, no significant over- or under-bidding behavior has been observed. Consequently, Dechenaux et al. (2015) conclude that whether overbidding can be observed is parameter-dependent and the findings from large all-pay auctions may not hold in smaller auctions. With respect to the effects of contest size on individual bids, a laboratory experiment conducted by Gneezy and Smorodinsky (2006) has found that increasing the number of bidders decreases the average bid in a complete-information all-pay auction with symmetric players, especially in later rounds.

This study fills a gap in the literature on contests by comparing simultaneous and sequential contests, modeled as all-pay auctions under incomplete information. Specifically, it differs from prior studies in a number of aspects. First, instead of modeling lottery contests (Morgan, 2003; Fu, 2006; Leininger, 1993), we follow Konrad and Leininger (2007) and Liu (2015) to focus on all-pay auctions. Second, while Konrad and Leininger (2007) and Liu (2015) examine complete information games, our framework assumes incomplete information. Third, as mentioned before, while in both ours and Konrad and Leininger's models there are n players, their model divides them into two stages (early and late) and our model allows n stages with only one unique player moving in each stage. Finally, while most prior comparisons of these two mechanisms focus on total effort as the main metric, our study focus on expected maximum effort as the main measure.

3 Theoretical Analysis

N risk-neutral contestants compete in a single task for a prize v which is normalized to 1. Each of them chooses an effort level: $q_i \geq 0$. $a_i \geq 0$ is i's ability factor, an *i.i.d.* draw from $[0, 1]$ with $F(a) = a^c$, where $c \in (0, +\infty)$.[2] A higher c indicates more high ability contestants.

Conditional on others' choices q_{-i}, contestant i's payoff is:[3]

$$\pi_i(q_i, q_{-i}) = \begin{cases} 1 - \frac{q_i}{a_i} & q_i > q_j, \forall j \neq i \\ -\frac{q_i}{a_i} & \text{otherwise} \end{cases}$$

In simultaneous all-pay auctions, each contestant chooses her effort q_i independently and all contestants choose simultaneously. The unique equilibrium effort function has been derived in prior studies (Chawla and Hartline, 2013; Hillman and Riley, 1989; Krishna and Morgan, 1997; Weber, 1985), and we reproduce it in Appendix A.

In sequential all-pay auctions, contestant i chooses q_i after observing $q_{j<i}$ and the position of each contestant is exogenously determined. The BNE effort function is derived by Segev and Sela (2014), and reproduced in Appendix A. In Lemma 1 (Appendix B), we show that there is a clear late-mover advantage in sequential contests.

[2]This assumption enables the existence of an explicit BNE in sequential contests (Segev and Sela, 2014).
[3]When there is a tie, the winner is randomly selected.

The expected maximum effort (EME) in simultaneous contests has been derived previously by Moldovanu and Sela (2006). Here, we reproduce their result under our model setup.

$$\text{EME}^{Sim} = \frac{n(n-1)c^2}{(1+(n-1)c)(2nc-c+1)} \quad (1)$$

It is straightforward to show that EME^{Sim} monotonically increases in both n and c. [4] [5]Next, we derive the general expression for the expected maximum effort in sequential all-pay auctions.[6]

Proposition 1. *The expected maximum effort in n-player sequential contests is:*

$$EME^{Seq} = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{c\left((1-d_{j-1})^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_{j-1}}} - (1-d_j)^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_j}}\right)}{(\prod_{m=1}^{j-1}(1-d_m)^c)(\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_i})}$$

Also, $\sup EME^{Seq} = 1/e$.

Proof. See Appendix C. □

EME^{Seq} monotonically increases in n, although such monotonicity does not hold for general ability distributions (Segev and Sela, 2014). Counterintuitively, EME^{Seq} is not monotonic in c (Segev and Sela, 2014).

Now we compare the two mechanisms and summarize the result below.

Theorem 1. $\forall$ n *and* c, $EME^{Sim} > EME^{Seq}$.

Proof. See Appendix D. □

The main intuition is as follows. In a two-player sequential game, a weak first mover would bid low, anticipating being beaten by a stronger participant later. This is called the "calm-down" effect (Fu, 2006; Leininger, 1993) which results in low maximum effort. In contrast, if the first mover is strong, one might think that she can engage in a deterrence strategy, i.e., bidding excessively high to deter high bids from her opponent (Morgan, 2003; Konrad and Leininger, 2007). However, with incomplete information, even a relatively strong first mover cannot know for sure that she is stronger than her opponent. Furthermore, unlike in lottery contests, in an all-pay auction the lower bidder loses for sure. Therefore, the deterrence strategy is not profitable but the calm-down effect is salient, resulting in an overall lower maximum effort in sequential contests.

[4]As a special case of ours ($c = 1$), Chawla et al. (2012) show that with a uniform distribution, the expected maximum effort monotonically increases in n.

[5]It is worthwhile to note that with more general distribution functions, the expected maximum effort in simultaneous all-pay auctions is not necessarily monotonic (Moldovanu and Sela, 2006).

[6]Segev and Sela (2014) derive the expected maximum effort for $n = 2$, and $n = 3$ under the condition that each participant has a different c, although they do not provide a generalized explicit solution for n players.

Table 1: Experimental Design

Contest Format	Contest Size	Treatment Name	Number of Subjects/Session	Number of Sessions	Total Number of Subjects
Sequential	3	SEQ3	12	4	48
	2	SEQ2	12	4	48
Simultaneous	3	SIM3	12	4	48
	2	SIM2	12,10,8	4	42

Pertaining to the expected total effort (ETE), when $c \geq 1$, it is higher in simultaneous contests since it is zero in sequential ones. When $c \in (0,1)$, the following Proposition, implied by Myerson (1981), holds for large contests (n sufficiently large).

Proposition 2. *$\forall\ c \in (0,1)$, $\exists$ an N, such that for any $n > N$, $ETE^{Sim} > ETE^{Seq}$.*

Proof. See Appendix E. □

4 Experimental Design

To test the theoretical predictions, especially to examine the effects of contest mechanism and contest size, we implemented a between-subject 2×2 factorial design, as shown in Table 1. We expected individual behaviors to vary with contest mechanism (simultaneous versus sequential) as well as contest size (two versus three participants).

Each treatment had four independent sessions, each with 12 subjects.[7] At the beginning of each round, subjects were randomly assigned into groups of two or three, depending on the treatment. As our theoretical model was based on one-shot games, this random re-matching protocol was used to minimize repeated-game effects in the experiment. In all treatments, the value of the reward was 100 tokens and every subject was given an ability factor, a_i ($a_i \in (0,1]$), which was randomly drawn from $F(a) = a_i^{0.25}$. They were informed that their ability factors were private knowledge. To prevent potential bankruptcies, we gave every subject 120 tokens as an endowment at the beginning of each round. Participants could choose an effort level from 0 to 120.[8] Finally, to capture any learning effect, all subjects played eighteen paying rounds after two practice rounds.

In simultaneous treatments, each participant made an effort independently and simultaneously. The sequential contests were implemented with games in multiple stages (2 stages for G2 and 3 stages for G3). In each stage a different participant had an exclusive opportunity to choose an effort level (including 0), after observing all previous participants' effort levels. Before the first

[7]Two simultaneous contest sessions with two players (SIM2) did not have enough subjects. Therefore, we had to run a session with 10 subjects and another with 8 subjects. However, the observed behaviors in these two sessions were not statistically significantly different from other sessions, so we treated them the same as other sessions of the same treatment.

[8]Four decimal points were kept for the display of ability factors, i.e., a_i, and for subjects' input of their effort levels, i.e., q_i.

stage started, the order of entry was randomly assigned by the computer and was announced to all participants in each group. Additionally, in both simultaneous and sequential treatments, before everyone chose an effort level, we elicited their beliefs of their winning probabilities. To incentivize the subjects to accurately report their beliefs, a quadratic scoring rule (Nyarko and Schotter, 2002) was implemented, which paid a maximum amount of two tokens for making accurate predictions.[9] A sample of the instructions is included in Appendix F. After the experiment, we gave each participant a post-experiment survey (see Appendix G) which collected demographic and personality trait information, including risk preferences.[10]

Based on our experimental parameter values, i.e., $c = 0.25$ and $n = 2, 3$, we computed all the theoretical values of interest and report them in Table 2. The table includes the equilibrium effort for each player, their expected winning probability, the expected maximum effort, the average effort, and the total effort. We also computed two efficiency measures: the expected proportion of efficient allocations and the value efficiency. Value efficiency is defined as the ratio of the winning contestant's ability to the highest ability among all contestants (Plott and Smith, 1978).[11]

In total, we conducted 16 independent computerized sessions at the Los Angeles Behavioral Economics Laboratory at the University of Southern California from January 2012 to February 2013, utilizing a total of 186 subjects. Our subjects were students recruited by email from a subject pool for economics experiments. We used z-Tree (Fischbacher, 2007) to program our experiments. Each session lasted about one and a half hours. The exchange rate was 110 tokens per \$1. The average amount earned by our participants was \$29, including a \$5 show-up fee.[12] Our experimental data are available from the authors upon request.

5 Experimental Results

We start with results aggregated at the session level and report the treatment effects with regard to both effort levels and efficiency. A detailed individual-level analysis follows to examine dynamics in participants' behaviors. In performing session-level analyses, we treated each session as one independent observation and computed the average within the session across multiple rematched groups and all periods. Therefore, we had four independent observations per treatment. In performing individual-level analyses, each bid was treated as a unit of analysis and standard errors were clustered at the session level. We also estimated an alternative model with random effects on each individual to account for their repeated measures across multiple rounds while clustering the standard errors at the session level. The results were qualitatively no different.

[9] We chose two tokens as the maximum earning since it was much below the average earning in our auctions in each round (146.67 tokens), so as to avoid hedging biases as suggested by Blanco et al. (2010).

[10] Due to time constraint and fear of experimental subjects' fatigue, we did not use an incentivized method for measuring risk preference (Charness et al., 2013; Eckel and Grossman, 2002; Holt and Laury, 2002).

[11] In some literature this measure is simply called "efficiency" (e.g., Noussair and Silver, 2006). Here, we call it *value efficiency* to differentiate it from the first measure.

[12] Seven subjects went bankrupt during the experiment and we decided to pay them a ten dollar flat fee to compensate them for the time, although such a payment was not pre-announced.

Table 2: Theory Predictions with $c = 0.25$

N	Game	Expected Effort			Winning Probability		
		Player 1	Player 2	Player 3	Player 1	Player 2	Player 3
2	Seq	2.487	1.193	NA	27%	73%	NA
3	Seq	2.836	2.835	2.069	13%	25%	63%
2	Sim	3.333	3.333	NA	50%	50%	NA
3	Sim	4.762	4.762	4.762	33%	33%	33%
N	Game	Expected Effort			Efficient	Value	
		Maximum	Average	Total	Allocations	Efficiency	
2	Seq	2.487	1.840	3.68	77%	86%	
3	Seq	4.684	2.580	7.74	65%	81%	
2	Sim	5.714	3.333	6.666	100%	100%	
3	Sim	11.111	4.762	14.286	100%	100%	

5.1 Effort Levels

Recall our theoretical prediction that simultaneous contests produce higher maximum effort than sequential contests (Theorem 1), which leads to Hypothesis 1:

Hypothesis 1 (Effort level comparison)**.** *For a given number of contestants and a given ability distribution, the expected maximum effort is higher in simultaneous than in sequential contests.*

Figure 1 plots the mean of the maximum effort under sequential (black line) and simultaneous (grey line) contests in each round. The left panel is for contests with two players and the right one is for contests with three players. The solid lines represent experimental data and the dashed lines represent theoretical predictions. Clearly, in contests with two players, simultaneous contests outperform sequential contests throughout all rounds, whereas the pattern in contests with three players is less clear at the beginning. However, toward the second half of the session (round 9 onwards), it becomes clear that simultaneous contests produce higher maximum effort than sequential contests. Formally, we conduct rank-sum tests to compare the two contests and report the results in Table 3, which shows the means of the maximum effort in different treatments for all rounds as well as for the first (Rounds 1-9) and second half (Rounds 10-18) of the session separately. For ease of comparison, we also include the theoretically predicted values at the bottom of Table 3. We summarize our first result below:

Result 1 (Effort level comparison)**.** *The maximum effort is higher in simultaneous contests than in sequential contests, and the effect is significant in later rounds.*

Support. *Across all rounds, the maximum effort is higher in simultaneous contests than in sequential contests, although neither of the comparisons are statistically significant (SIM2 vs. SEQ2: 11.01 vs. 4.35,* $p = 0.248$*; SIM3 vs. SEQ3: 9.01 vs. 6.93,* $p = 0.248$*; two-sided rank-sum tests). The comparison becomes weakly significant in the second half of the experiment (SIM2 vs. SEQ2: 10.61 vs. 4.55,* $p = 0.083$*; SIM3 vs. SEQ3: 9.45 vs. 5.3,* $p = 0.043$*; two-sided rank-sum tests).*

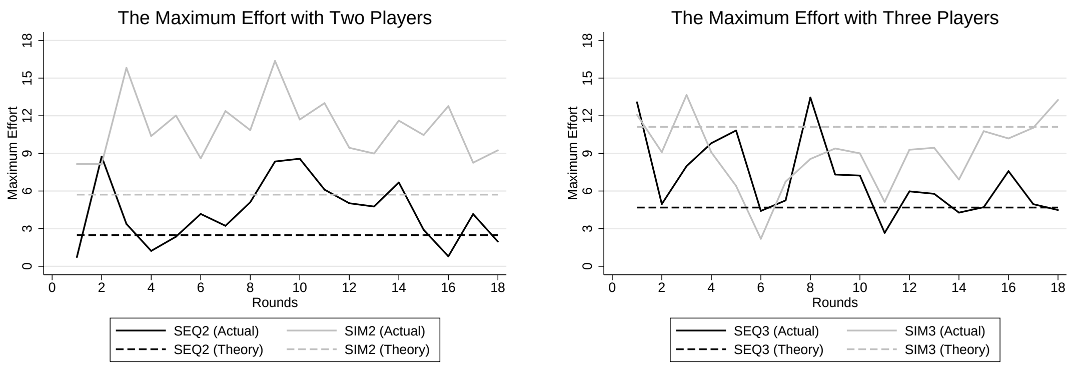


Figure 1: The Maximum Effort: Sequential vs. Simultaneous Contests

Table 3: Treatment Effects on the Maximum Effort

All Rounds	Seq	Sim	Seq vs. Sim
n=2	4.35	11.01	$p = 0.248$
n=3	6.93	9.01	$p = 0.248$
n=2 vs. n=3	$p = 0.248$	$p = 0.773$	
Rounds 1-9	Seq	Sim	Seq vs. Sim
n=2	4.14	11.41	$p = 0.149$
n=3	8.57	8.58	$p = 1.00$
n=2 vs. n=3	$p = 0.149$	$p = 0.773$	
Rounds 10-18	Seq	Sim	Seq vs. Sim
n=2	4.55	10.61	$p = 0.083$
n=3	5.3	9.45	$p = 0.043$
n=2 vs. n=3	$p = 0.564$	$p = 0.564$	
Theory	Seq	Sim	
n=2	2.487	5.714	
n=3	4.684	11.111	

Note: P-values from two-sided rank-sum tests are reported for all comparisons in this table. Each session is treated as an independent observation.

Due to Result 1, we reject the null in favor of Hypothesis 1. The difference between the two mechanisms' performance is large. For example, in two-player contests, the mean maximum effort in simultaneous contests is more than two times higher than that in sequential contests (Rounds 10-18: 10.61 vs. 4.55).[13]

In addition, consistent with the theoretical values computed for our experimental setting (see Table 2), we find that simultaneous contests also produce higher average effort than sequential contests. Although the comparison is not statistically significant across all rounds (SIM2 vs. SEQ2: 5.96 vs. 2.91, $p = 0.248$; SIM3 vs. SEQ3: 3.86 vs. 3.21, $p = 0.387$, two-sided rank-sum tests), it becomes significant when only the second half of the experiment is considered in contests with three players (SIM3 vs. SEQ3: 4.21 vs. 2.62, $p = 0.021$, two-sided rank-sum tests).[14] Taken together, these results suggest that simultaneous contests are more effective than sequential contests in inducing higher maximum effort and average effort from their contestants.

Next, we examine the size effect on effort level. Following the theoretical results in Section 3, we propose Hypothesis 2:

Hypothesis 2 (Size effect on effort level)**.** *In both sequential and simultaneous contests, the expected maximum effort increases with the number of contestants.*

Again, results are reported in Table 3. On average, the maximum effort is higher in SEQ3 than in SEQ2, but no statistically significant difference is observed either across all rounds or in the second half of the experiment (All Rounds: 6.93 vs. 4.35, $p = 0.248$; Rounds 10-18: 5.30 vs. 4.55, $p = 0.564$, two-sided rank-sum tests). Furthermore, no statistically significant difference is observed between SIM2 and SIM3 treatments ($p > 0.1$ for all comparisons) either. Therefore, we cannot reject the null hypothesis in favor of Hypothesis 2.

In addition, in contrast with the theoretical predictions in Table 2, the average effort is not statistically significantly higher in SEQ2 than in SEQ3 (All Rounds: 2.91 vs. 3.21, $p = 0.773$, Rounds 10-18: 3.12 vs. 2.62, $p = 0.773$, two-sided rank-sum tests). The comparison of the average effort between SIM2 and SIM3 is not statistically significant either ($p > 0.1$, two-sided rank-sum test).

5.2 Efficiency

In simultaneous contests, as all contestants are symmetric and the equilibrium effort function monotonically increases in ability, the player with the highest ability always wins and full efficiency is always achieved (Archak and Sundararajan, 2009). Therefore, simultaneous contests score one on both efficiency measures, i.e., the expected proportion of efficient allocations and value efficiency.

[13] To benchmark our experimental contests' performance against their BNE predictions, we report two-sided p-values from one-sample signed rank tests. In both simultaneous and sequential contests, the aggregate maximum effort does not appear different from BNE predictions ($p > 0.1$, two-sided tests).

[14] The comparison remains insignificant for the contests with two players (SIM2 vs. SEQ2: 5.80 vs. 3.12, $p = 0.149$, two-sided rank-sum tests).

In contrast, sequential contests are not always efficient, since late movers have positional advantages, as demonstrated by Lemma 1. As a result, the contestant with the highest ability does not always win. Such a difference suggests that simultaneous contests are more efficient than sequential contests, leading to Hypothesis 3:

Hypothesis 3 (Efficiency comparison)**.** *For a given number of contestants and a given ability distribution, simultaneous contests achieve higher efficiency than sequential contests.*

Table 4: Treatment Effects on the Proportion of Efficient Allocations

All Rounds	Seq	Sim	Seq vs. Sim
n=2	0.75	0.78	$p = 0.387$
n=3	0.56	0.74	$p = 0.021$
n=2 vs. n=3	$p = 0.021$	$p = 0.191$	
Rounds 1-9	Seq	Sim	Seq vs. Sim
n=2	0.74	0.78	$p = 0.234$
n=3	0.57	0.73	$p = 0.081$
n=2 vs. n=3	$p = 0.041$	$p = 0.561$	
Rounds 10-18	Seq	Sim	Seq vs. Sim
n=2	0.75	0.77	$p = 0.564$
n=3	0.55	0.75	$p = 0.029$
n=2 vs. n=3	$p = 0.043$	$p = 0.564$	
Theory	Seq	Sim	
n=2	0.77	1.00	
n=3	0.65	1.00	

Note: P-values from two-sided rank-sum tests are reported for all comparisons in this table. Each session is treated as an independent observation.

Tables 4 and 5 present the proportion of efficient allocations and value efficiency in different treatments for all rounds as well as for the two halves of the sessions separately. Comparing the efficiency measures across treatments, we find that, consistent with theoretical predictions, simultaneous contests produce more efficient allocations as well as higher value efficiency than sequential contests, and the comparison is significant for contests with three players. We summarize the results below:

Result 2 (Efficiency comparison)**.** *When $n = 3$, both the proportion of efficient allocations and value efficiency for simultaneous contests are superior to sequential contests.*

Support. *The proportion of efficient allocations is higher in simultaneous contests than in sequential contests. In particular, the comparison for the three-player case produces statistically significant results (All Rounds: 0.74 vs. 0.56, $p = 0.021$; Rounds 10-18: 0.75 vs. 0.55, $p = 0.029$, two-sided rank-sum tests). Similarly, in three-player conditions, value efficiency is also significantly higher in simultaneous contests (All Rounds: 0.86 vs. 0.70, $p = 0.021$; Rounds 10-18: 0.89 vs. 0.7, $p = 0.021$, two-sided rank-sum tests).*

Table 5: Treatment Effects on Value Efficiency

All Rounds	Seq	Sim	Seq vs. Sim
n=2	0.84	0.86	$p = 0.387$
n=3	0.70	0.86	$p = 0.021$
n=2 vs. n=3	$p = 0.021$	$p = 1.00$	
Rounds 1-9	Seq	Sim	Seq vs. Sim
n=2	0.84	0.85	$p = 0.564$
n=3	0.70	0.84	$p = 0.083$
n=2 vs. n=3	$p = 0.021$	$p = 0.564$	
Rounds 10-18	Seq	Sim	Seq vs. Sim
n=2	0.83	0.86	$p = 0.564$
n=3	0.70	0.89	$p = 0.021$
n=2 vs. n=3	$p = 0.043$	$p = 0.248$	
Theory	Seq	Sim	
n=2	0.86	1.00	
n=3	0.81	1.00	

Note: P-values from two-sided rank-sum tests are reported for all comparisons in this table. Each session is treated as an independent observation.

Due to Result 2, we reject the null hypothesis in favor of Hypothesis 3 for the three-player case and conclude that simultaneous contests are more efficient than sequential contests.[15] The difference is quite large in magnitude. For example, in three-player cases, sequential contests only achieve efficient allocations about half the time (55%) but simultaneous contests do much better (75%).

Last, we examine the size effect on efficiency. We do not expect any size effect on efficiency in simultaneous contests, since simultaneous contests are always efficient regardless of contest size. In sequential contests, as shown in Table 2, we should expect efficiency to decrease in n. Therefore we propose the next hypothesis:

Hypothesis 4 (Size effect on efficiency)**.** *In sequential contests, efficiency worsens as the number of contestants increases.*

Consistent with theoretical predictions, we find that in sequential contests both measures of efficiency decrease as the number of contestants increases, and in simultaneous contests no size effect is observed. Therefore, we have the following result:

Result 3 (Size effect on efficiency)**.** *In sequential contests, both the proportion of efficient allocations and value efficiency decrease with the number of contestants.*

[15]To benchmark the efficiency measures with their theoretical predictions (Table 2), we report results using one-sample signed rank tests. Overall, simultaneous contests' efficiency levels are lower than expected, i.e., achieving a proportion of efficient allocations lower than 100% (SIM2: 78%, $p = 0.068$; SIM3: 74%, $p = 0.068$, two-sided tests). SEQ2 contests' allocative efficiency levels are very close to their theoretical predictions, i.e., reaching a proportion of efficient allocations at 75% (vs. 77%, $p = 0.26$, two-sided test). The SEQ3 contests, however, are less efficient than predicted, i.e., achieving a proportion of efficient allocations of 56% (vs. 65%, $p = 0.07$, two-sided test).

Support. *The proportion of efficient allocations in SEQ2 is statistically significantly higher than SEQ3 (All Rounds: 0.75 vs. 0.56, $p = 0.021$; Rounds 1-9: 0.74 vs. 0.57, $p = 0.041$; Rounds 10-18: 0.75 vs. 0.55, $p = 0.043$, two-sided rank-sum tests). Consistently, the value efficiency in SEQ2 is statistically significantly higher than SEQ3 (All Rounds: 0.84 vs. 0.70, $p = 0.021$; Rounds 1-9: 0.84 vs. 0.70, $p = 0.021$; Rounds 10-18: 0.83 vs. 0.70, $p = 0.043$, two-sided rank-sum tests).*

Due to Result 3, we reject the null hypothesis in favor of Hypothesis 4. For sequential contests, we observe a significant negative size effect on the two allocative efficiencies. In particular, adding one more contestant can reduce efficiency significantly, e.g., the proportion of efficient allocations dropping from 0.75 to 0.55 in the second half of the sessions, although such differences are not observed for simultaneous contests. Such differential size effects in these two mechanisms highlight the impact of participant asymmetry in sequential contests, such that efficiency in sequential contests is highly sensitive to contest size.

5.3 Individual level analysis

We start by checking whether the last players in sequential contests best respond, since their strategies are the most straightforward. We find that the majority of last players (77% in SEQ2 and 65% in SEQ3) in sequential contests best respond. There is also a strong indication of learning: in the second half of the experiment, the percentage of best responses increases to 83% in SEQ2 and to 74% in SEQ3.[16]

Next, we examine all other players' strategies, except the middle players' in SEQ3 (which will be discussed later), since their choices depend on their preceding players'. Table 6 presents the OLS regression results, predicting the deviation of individuals' observed effort from their BNE predictions (i.e., $\delta =$ actual effort$-$predicted effort) for all players in simultaneous contests and the first players in sequential contests.[17] The independent variables include the ability factor and the period variable to control for learning effects. Since individual bidding behavior may be affected by their risk attitude and gender (Dechenaux et al., 2015), we also control for these two variables. Both risk attitude and gender were elicited in the post-experiment questionnaire (see Appendix G). Risk attitude was measured by a 7 point Likert scale with 7 being most risk seeking. Gender was coded as a dummy variable with 1 being female.

The results show that except in SIM3, no estimated coefficients are significantly different from zero.[18] In SIM3, the coefficient of the ability factor is negative (-10.37) and statistically signifi-

[16]Because z-Tree requires a maximum decimal point for any numerical input, subjects' efforts were restricted to four decimal points in our experiment. Therefore, we use 0.0001 as the increment of best responses, i.e., when a player tries to match the existing highest effort, she should simply make an effort that equals the existing highest effort plus 0.0001. To account for subjects' imprecision in entering their efforts, we checked our results using 0.001, 0.01 and 0.1, and found no significant changes to our results.

[17]We also ran probit regressions where the dependent variable was whether overbidding occurred, and the results were qualitatively consistent.

[18]Since the standard errors are clustered at the session level, they could be biased due to the small number of clusters in our data. As a robustness check, we ran regression analyses with a wild cluster bootstrap procedure to correct for the small number of clusters (Cameron et al., 2008) and found that the results remained the same.

Table 6: Deviation from BNE Predictions: OLS Regressions

	Simultaneous		Sequential	
	N=2	N=3	N=2	N=3
	(1)	(2)	(3)	(4)
Ability	-3.96	-10.37**	2.74	-0.45
	(2.16)	(2.76)	(3.23)	(3.15)
Period	0.02	-0.02	0.15	-0.10
	(0.06)	(0.06)	(0.09)	(0.14)
Gender	3.61	-1.70	-0.46	-1.90
	(3.02)	(1.06)	(1.25)	(1.09)
Risk	0.25	0.31	0.19	0.62
	(0.22)	(0.56)	(0.35)	(0.55)
Constant	0.77	1.15	0.47	0.33
	(1.19)	(1.99)	(1.29)	(1.18)
Obs.	756	864	432	288
R^2	0.01	0.20	0.01	0.06

Notes: Standard errors are clustered at the session level.

** $p < 0.05$, * $p < 0.1$

cantly different from zero ($p < 0.05$, two-sided t test). For low ability participants (e.g., 32% of the participants with $a < 0.01$), this coefficient indicates a negligible amount of under-exertion (0.1 on a scale of 0 to 100 points). For high ability participants (e.g., the 16% with $a > 0.5$), however, this coefficient indicates a fairly large amount of under-exertion (about 5 on a scale of 0 to 100 points).

These results are consistent with the amount of over- and under-exertion calculated directly from participants' effort levels across all treatments and all four quantiles of the ability distribution. That is, for each treatment and each quantile of the ability distribution, we compute the mean deviation, $\bar{\delta}$, and find that only high ability players in the SIM3 condition exert significantly lower effort than BNE predictions. Specifically, in the highest quantile of the ability distribution, i.e., $a > 0.32$, the average deviation is -5.05 and the difference is marginally significantly lower than zero ($p = 0.068$, two-sided one-sample signed rank tests).[19]

Taken together, these results reveal that with two contestants, the observed effort levels in simultaneous contests match their BNE predictions, in line with findings from previous experiments using similar experimental protocols (Grosskopf et al., 2010; Potters et al., 1998). In contrast, with three contestants, the effort levels of high ability contestants are significantly lower than BNE predictions. This contradicts previous findings (Muller and Schotter, 2010; Noussair and Silver, 2006) where high ability contestants (from a uniform distribution) over-exert effort. Muller and Schotter (2010) rationalize their data by assuming loss averse agents[20] and Noussair and Silver (2006)

[19]Following Muller and Schotter (2010), we used a switching regression model to examine whether the individuals' effort function was continuous. The switching regression model fit the data significantly better than BNE predictions based on the sum of squared deviations (SSD) measure (SIM2: 21.37 vs. 549.83, $p < 0.01$; SIM3: 25.04 vs. 695.77, $p < 0.01$, two-sided signed rank tests). Nevertheless, this switching regression model could not explain why the effort levels of high ability contestants in SIM3 were lower than BNE predictions.

[20]Under the assumption of loss-aversion, Mermer (2013) analytically shows that high ability contestants over-exert

explain the bidding pattern by Fibich et al.'s (2006) model of risk aversion. However, neither of these models can explain the seemingly counterintuitive result here. Below, we discuss why under-exertion is observed for high ability contestants in SIM3, but not in SIM2 and not for high ability first players in sequential contests.

We conjecture that under-exertion by high ability contestants in SIM3 is due to overplacement, a specific type of overconfidence (Moore and Healy, 2008) — people tend to overestimate their chances of success in comparison to their average peers. A recent study by Masiliunas et al. (2014) also reveals that bounded rationality, e.g., the difficulty of forming correct beliefs, explains the large behavioral variations in Tullock contest experiments. Similarly, Cooper and Fang (2008) find that a bidder's misperception of the opponent's strength, e.g., the value size of the object, significantly affects her behavior in second price auctions. In addition, compared to the uniform distribution function often used in previous experiments (e.g., Muller and Schotter, 2010; Noussair and Silver, 2006), the power distribution used in our study is relatively hard for participants to work with.[21] In particular, the power distribution function features a larger proportion of low ability contestants, which makes a draw of relatively high ability a rare event. We suspect that whenever a high ability factor is drawn for a contestant, it could easily lead to her overestimation of the winning probability.

Our first piece of evidence to support the overconfidence effect comes from the belief data (recall that in the experiment we elicited subjects' beliefs of their winning probabilities before letting them make their effort choices). Using this data, we can verify whether the subjects had misbeliefs about their winning probabilities. Specifically, in SIM3, subjects believed they were 7% more likely to win than predicted by BNE, though this difference is not significant ($p = 0.14$, two-sided signed rank tests). In contrast, in SIM2, no overestimate of winning probabilities was observed (-1%, $p = 0.71$, two-sided signed rank tests). This is understandable, since it is more difficult to compute one's winning probability in a three-player contest than in a two-player one. Indeed, overplacement is more likely to occur if one judges herself against the average of a group of others, compared to against another individual (Alicke et al., 2004).

To further quantify this overconfidence, we define individual players' misperceived winning probabilities as follows:

1. In SIM2, it is $\widetilde{P}(a_i) = a_i^{\widetilde{c}}$ for $i \in \{1, 2\}$;
2. In SIM3, it is $\widetilde{P}(a_i) = a_i^{2\widetilde{c}}$ for $i \in \{1, 2, 3\}$;
3. In SEQ2, it is $\widetilde{P}(a_1) = (a_1 * \widetilde{c})^{\frac{\widetilde{c}}{1-\widetilde{c}}}$;
4. In SEQ3, it is $\widetilde{P}(a_1) = (a_1 * (2\widetilde{c} - \widetilde{c}^2))^{\frac{2\widetilde{c}-\widetilde{c}^2}{1-\widetilde{c}^2}}$,

where $\widetilde{c}$ represents the corresponding parameter c in the misperceived winning probability. Apparently, the lower the $\widetilde{c}$, the stronger the overconfidence, i.e., the extent to which the probability of

effort while low ability contestants under-exert effort.

[21] The power distribution used in our experiment was described by its quantiles in the experimental instructions. See Appendix F for details.

winning is over-estimated. Using non-linear least square regression models, we obtain the point estimations of $\widetilde{c}$ based on the belief data (Table 7). We find that the best fitting $\tilde{c}$ in SIM3 is 0.18, which is marginally significantly smaller than 0.25 ($p = 0.06$, two-sided t test), while in the SIM2 treatment, the estimated $\tilde{c}$ is exactly 0.25 ($p = 0.99$, two-sided t test). This shows that there is some overconfidence among participants in SIM3 but not in SIM2.

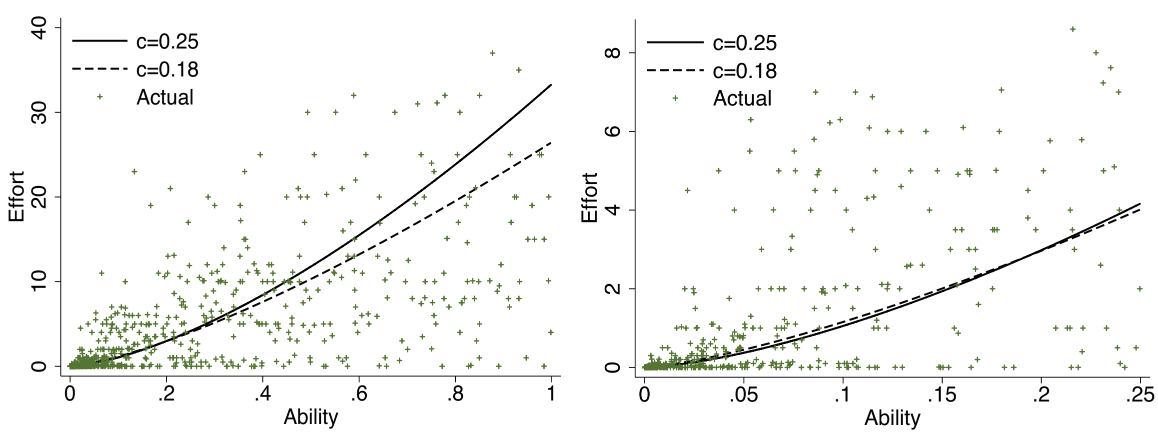
Table 7: The Estimated $\tilde{c}$ from Belief Data

	Simultaneous		Sequential	
	N=2	N=3	N=2	N=3
	(1)	(2)	(3)	(4)
$\tilde{c}$	0.25	0.18	0.21	0.15
Std.Err.	0.03	0.02	0.01	0.01
P-value	0.99	0.06	0.05	0.00
R^2	0.79	0.79	0.66	0.60

Note: P-values are reported from two-sided t-tests with the null hypothesis $\tilde{c} = 0.25$.

Furthermore, we use this best fitting $\tilde{c}$ to simulate participants' corresponding effort choices. For simplicity, we use the BNE effort function to generate the efforts, while replacing the actual c with the estimated $\tilde{c}$. Both the simulated efforts and BNE predicted efforts are plotted in Figure 2, with the actual effort choices made by the experiment subjects represented by scattered "+" signs. The solid line represents BNE efforts calculated with $c = 0.25$ and the dotted line represents efforts simulated with $\tilde{c} = 0.18$ (The left panel is for the full range of $a \in [0, 1]$ and the right panel shows the same curves but zooms in on the low ability range, i.e., $a \in [0, 0.25]$). We observe that the simulated efforts are significantly lower than BNE in the high ability range. This pattern is consistent with the previously described pattern of actual observed efforts, indicating that overconfidence can be a plausible explanation for high ability contestants' under-exertion behavior in SIM3.

Figure 2: BNE vs. Simulated Efforts in SIM3 Treatment



To assess model fit, we follow Muller and Schotter (2010) by computing the sum of squared deviations (SSD) of the BNE model (BNE) and the overconfidence model based on belief data (BLF). That is, for each individual i and across all t rounds, $SSD_i^{\text{BLF}} = \sum_{t=1}^{18}((q_i^t)^{BLF} - (q_i^t)^{obs})^2$ and $SSD_i^{\text{BNE}} = \sum_{t=1}^{18}((q_i^t)^{BNE} - (q_i^t)^{obs})^2$, where $(q_i^t)^{obs}$ denotes observed efforts. Table 8 reports the average SSD in each treatment. Column 2 reports the average SSD_i^{BNE} and Column 3 reports the average SSD_i^{BLF}. We find that in the SIM3 treatment, although only weakly significant ($p = 0.07$, two-sided signed rank tests), on average SSD_i^{BNE} (814.8) is 20% higher than SSD_i^{BLF} (678.9), indicating that our proposed overconfidence model fits the data much better than BNE predictions.

Table 8: Overview of the Sum of Squared Deviations (SSD)

	Average SSD based on		Wilcoxon Signed rank test
Treatment	BNE	BLF	Two-sided p values
SIM2	3912.5	3912.5	0.72
SIM3	814.8	678.9	0.07
SEQ2	806.0	815.2	0.27
SEQ3	340.4	337.8	0.72

Moving on to sequential contests, the individual behaviors there show a different pattern. Although the estimated $\widetilde{c}$ in the misperceived winning probabilities are significantly lower than $c = 0.25$ ($\tilde{c} = 0.21$, $p = 0.05$ for SEQ2 and $\tilde{c} = 0.15$, $p < 0.01$ for SEQ3 in Table 7), they do not significantly over- or under-exert effort, as shown in Table 6 (Column 3 and 4) earlier.[22] This suggests that overconfidence does not lead to high ability contestants' under-exertion in sequential contests. The reason, we speculate, is due to high ability contestants' raising their effort to compensate for their positional disadvantages (recall that here we only examine the first players in sequential contests, and according to Lemma 1, they face substantial positional disadvantages). Last, we examine the middle players' behaviors in SEQ3, and find that their behaviors are largely in line with the theoretically predicted best responses.

Summarizing our individual level analyses, we find some under-exertion for high ability contestants in SIM3, while other participants do not significantly deviate from BNE predictions. We construct a simple behavioral model of overconfidence to explain why these SIM3 players under-exert effort as opposed to over-exerting, as predicted by models of risk-aversion. The overconfidence model fits our data and rationalizes the observed behavior.

6 Conclusion and Discussion

In this study, by modeling contests as n-player n-stage all-pay auctions with incomplete information, under certain conditions, we analytically show that simultaneous contests generate higher maximum effort than sequential contests and produce more efficient outcomes. We conducted a

[22] Consistently, results from Table 8 show that the BLF model does not fit the data significantly better than BNE.

laboratory experiment which provides supporting evidence for these theoretical predictions. Our experimental data also show that the efficiency of sequential contests decreases in the number of contestants, which further reduces their performance relative to simultaneous contests. Under limited conditions, we analytically show that the total effort will be lower in sequential contests than in simultaneous ones, which is consistent with prior studies and has been supported by our experimental data. Additionally, we discover that in simultaneous contests with three players, when participant ability follows a power distribution, high-ability contestants exert significantly lower effort than their BNE predictions. We explain this phenomenon with a model of overconfidence in which high-ability players overestimate their winning probabilities and consequently choose to under-exert effort.

The results of our comparisons between simultaneous and sequential contests are largely consistent with prior research comparing the two in various settings, including in lottery contests with multiple stages (Klumpp and Polborn, 2006) and all-pay auctions with endogenous ordering of moves (Fu, 2006; Konrad and Leininger, 2007; Leininger, 1993). However, our results contradict findings by Irfanoglu et al. (2015), who conducted laboratory experiments to test the multi-stage lottery contest model analyzed by Klumpp and Polborn (2006), and found evidence refuting the theoretical predictions. That is, contrary to Klumpp and Polborn (2006) as well as our findings, they found that sequential contests generated higher total effort than simultaneous contests. They explained the empirical anomaly by the sunk-cost fallacy and participants' potential non-monetary utility of winning.

Though there is a significant size effect for efficiency comparison in sequential contests, our data provide no support for the size effect on the expected maximum effort. Specifically, for simultaneous contests, the comparison proves to be directionally opposite to theoretical predictions (n=2 vs. n=3: 11.01 vs. 9.01). This can be understood in light of our analysis of potential overconfidence by high-ability players in SIM3. The unexpected low effort from high-ability players in SIM3 may have driven this result. For sequential contests, the comparison is directionally consistent with theory prediction (n=2 vs. n=3: 4.35 vs. 6.93), but not statistically significant. This result can be potentially caused by a number of factors, including a relatively small sample size (the number of independent observation per treatment is 4), and insufficient differences in the theoretically predicted values. Behavior-wise, we suspect that when n increases from 2 to 3, the first player in the sequence becomes increasingly more likely to use the calm-down strategy (i.e., exert low effort to calm down subsequent players) than to use the deterrence strategy (i.e., exert high effort to drive subsequent players out of competition). This is evidenced in Table 6. Although not statistically significant, the coefficients on ability show that the high-ability first players in SEQ2 and SEQ3 behave differently. In SEQ2 they tend to over-exert effort (the coefficient on ability, 2.74, is positive) and in SEQ3 they tend to under-exert effort (the coefficient on ability, -0.45, is negative). We speculate that this difference is caused by the first player in SEQ3 being overly concerned by the late-mover advantage, since there are two players who have such advantages over her. Due to fear of losing, she would opt for under-exertion. Such a behavior would lead to the lower-than-predicted increase in the maximum effort, and offset the theoretically predicted size effect.

Overall, our contributions are two-fold. First, it fills a gap in the contest literature by comparing simultaneous and sequential contests under a new setting: n-player n-stage incomplete information all-pay auctions. Second, our experimental results reveal behavioral patterns that have never been highlighted before in the contest literature. Previously, experimental studies of behaviors in tournaments have observed that high ability participants over-exert effort and have identified a number of factors that could explain such deviation from BNE predictions, including the sunk cost fallacy, non-pecuniary utility of winning, and risk aversion (see a comprehensive literature survey by Dechenaux et al. (2015)). Our study reveals that high ability participants may also under-exert effort and we find that overconfidence can explain such behaviors. In particular, our study suggests that overconfidence occurs in contests when the following two conditions are satisfied: 1) the distribution of participants' private types are skewed such that high ability participants are rare, and 2) there is more than one opponent in the contest, which makes assessing one's own winning probability difficult.

Our findings also have practical implications for practitioners. In situations in which the contest mechanism can be a decision variable (e.g., in online contests, in conducting job interviews or sports contests), contest holders who care about eliciting higher maximum effort should consider implementing a simultaneous contest rather than a sequential one. For high ability participants in simultaneous contests with many players, our results on overconfidence suggest that they need to be mindful of the tendency to overestimate their winning probabilities, especially if high ability participants are outnumbered by low ability ones.

In future work, a number of assumptions made in our study need to be relaxed, especially our specific distributional assumption for contestants' ability factors, and the implicit assumption of the one-entry-per-participant rule and exogenous entry timing. Ultimately, theoretical predictions about the comparative performances of these two mechanisms could potentially be tested using field experiments (e.g., Carpenter et al., 2008; Harrison and List, 2004; Benz and Meier, 2008).

References

Alicke, M. D., M. L. Klotz, D. L. Breitenbecher, T. J. Yurak, and D. S. Vredenburg, "Personal Contact, Individuation, and the Better-than-Average Effect," *Journal of Personality and Social Psychology*, 2004, *68* (5), 804–825.

Altmann, Steffen, Armin Falk, and Matthias Wibral, "Promotions and Incentives: The Case of Multistage Elimination Tournaments," *Journal of Labor Economics*, 2012, *30* (1), 149–174.

Amegashie, J. A., C. B. Cadsby, and Y. Song, "Competitive Burnout: Theory and Experimental Evidence," *Games and Economic Behavior*, 2007, *59* (2), 213–239.

Anderson, S. P., J. K. Goeree, and C. A. Holt, "Rent Seeking with Bounded Rationality: An Analysis of the All-Pay Auction," *Journal of Political Economy*, 1998, *106* (4), 828–853.

Archak, N. and A. Sundararajan, "Attracting the best and the brightest: asymptotics and optimal prize structure of all-pay contests with heterogeneous participants," *Working paper*, 2009.

Barut, Y., D. Kovenock, and C. Noussair, "A Comparison of Multiple-Unit All-Pay and Winner-Pay Auctions Under Incomplete Information," *International Economic Review*, 2002, *43* (3), 675–708.

Baye, M. R., D. Kovenock, and C. G. de Vries, "Comparative analysis of litigation systems: an auction-theoretic approach," *The Economic Journal*, 2005, *115* (505), 583–601.

Benz, M. and S. Meier, "Do People Behave in Experiments as in the Field?-Evidence from Donations," *Experimental Economics*, 2008, *11* (3), 268–281.

Blanco, M., D. Engelmann, A. K. Koch, and H. T. Normann, "Belief elicitation in experiments: is there a hedging problem?," *Experimental Economics*, 2010, *13* (4), 412–438.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller, "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Carpenter, J., J. Holmes, and P. H. Matthews, "Charity auctions: a field experiment," *The Economic Journal*, 2008, *118* (525), 92–113.

Charness, Gary, Uri Gneezy, and Alex Imas, "Experimental methods: Eliciting risk preferences," *Journal of Economic Behavior & Organization*, 2013, *87*, 43–51.

Chawla, S. and J. D. Hartline, "Auctions with unique equilibria," *Proceedings of the fourteenth ACM conference on electronic commerce*, 2013, pp. 181–196.

_ , _ , and B. Sivan, "Optimal crowdsourcing contests," *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, 2012, pp. 856–868.

Che, Y.-K. and I. Gale, "Optimal Design of Research Contests," *American Economic Review*, 2003, *93* (3), 646–671.

Cooper, D. and H. Fang, "Understanding Overbidding in Second Price Auctions: An Experimental Study," *The Economic Journal*, 2008, *118* (532), 1572–1595.

Dasgupta, P., "The Theory of Technological Competition," in J.E. Stiglitz and F. Mathewson, eds., *New Developments in the Analysis of Market Structures*, London: Macmillan, 1986, pp. 519–548.

Davis, D. D. and R. J. Reilly, "Do Too Many Cooks Spoil the Stew? An Experimental Analysis of Rent-seeking and the Role of a Strategic Buyer," *Public Choice*, 1998, *95* (1), 89–115.

Dechenaux, E., D. Kovenock, and R. M. Sheremeta, "A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments," *Experimental Economics*, 2015, *18* (4), 609–669.

Dixit, A., "Strategic Behavior in Contests," *The American Economic Review*, 1987, *77* (5), 891–898.

Eckel, Catherine C and Philip J Grossman, "Sex differences and statistical stereotyping in attitudes toward financial risk," *Evolution and human behavior*, 2002, *23* (4), 281–295.

Fibich, G., A. Gavious, and A. Sela, "All-pay Auctions with Risk-averse Players," *International Journal of Game Theory*, 2006, *34* (4), 583–599.

Fischbacher, U., "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 2007, *10* (2), 171–178.

Fu, Q. and J. Lu, "The optimal multi-stage contest," *Economic Theory*, 2012, *51* (2), 351–382.

Fu, Qiang, "Endogenous timing of contest with asymmetric information," *Public Choice*, 2006, *129* (1), 1–23.

Gneezy, U. and R. Smorodinsky, "All-pay Auction: An Experimental Study," *Journal of Economic Behavior and Organization*, 2006, *61* (2), 255–275.

Gradstein, M. and K. A. Konrad, "Orchestrating rent seeking contests," *The Economic Journal*, 1999, *109* (458), 536–545.

Grosskopf, B., L. Rentschler, and R. Sarin, "Asymmetric Information in Contests: Theory and Experiments," *Working paper*, 2010.

Harris, Christopher and John Vickers, "Perfect Equilibrium in a Model of a Race," *Review of Economic Studies*, 1985, *52* (2), 193–209.

_ **and** _ , "Racing with uncertainty," *Review of Economic Studies*, 1987, *54* (1), 1–21.

Harrison, G. W. and J. A. List, "Field Experiments," *Journal of Economic Literature*, December 2004, *42* (4), 1009–1055.

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

Holt, C. A. and S. K. Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 16441655.

Irfanoglu, Zeynep, Shakun Mago, and Roman M Sheremeta, "New Hampshire Effect: Behavior in Sequential and Simultaneous Election Contests," *MPRA paper*, October 2015, pp. 1–36.

Jeppesen, L. B. and K. R. Lakhani, "Marginality and Problem-Solving Effectiveness in Broadcast Search," *Organization Science*, 2010, *21* (5), 1016–1033.

Klumpp, Tilman and Mattias K Polborn, "Primaries and the New Hampshire Effect," *Journal of Public Economics*, May 2006, *90* (6-7), 1073–1114.

Konrad, K. A., *Strategy and Dynamics in Contests*, New York, NY: Oxford University Press, 2009.

_ **and W. Leininger**, "The Generalized Stackelberg Equilibrium of the All-pay Auction with Complete Information," *Review of Economic Design*, 2007, *11* (2), 165–174.

Krishna, V. and J. Morgan, "An Analysis of the War of Attrition and the All-Pay Auction," *Journal of Economic Theory*, 1997, *72* (2), 343–362.

Lazear, E. P. and S. Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," *The Journal of Political Economy*, 1981, *89* (5), 841–864.

Leininger, Wolfgang, "More efficient rent-seeking - A Munchhausen solution," *Public Choice*, 1993, *75* (1), 43–62.

Liu, T. X., "All-Pay Auctions with Endogenous Bid Timing: An Experimental Study," 2015. Working paper.

_ **, J. Yang, L. A. Adamic, and Y. Chen**, "Crowdsourcing with All-pay Auctions: a Field Experiment on Taskcn," *Management Science*, 2014, *60* (8), 2020–2037.

Mago, Shakun D., Roman M. Sheremeta, and Andrew Yates, "Best-of-three contest experiments; Strategic versus psychological momentum," *International Journal of Industrial Organization*, 2013, *31* (3), 287–296.

Masiliunas, A., F. Mengel, and J.P. Reiss, "Behavioral Variation in Tullock Contests," *Working paper*, 2014.

Mermer, A. G., "Contests with expectation-based loss-averse players," 2013. Working paper.

Moldovanu, B. and A. Sela, "Contest Archiecture," *Journal of Economic Theory*, 2006, *126* (1), 70–96.

Moore, D. A. and P. J. Healy, "The trouble with overconfidence," *Psychological Review*, 2008, *115* (2), 502–517.

Morgan, J., "Sequential Contests," *Public Choice*, 2003, *116* (1), 1–18.

Muller, W. and A. Schotter, "Workaholics and dropouts in organizations," *Journal of the European Economics Association*, 2010, *8* (4), 717–743.

Myerson, R. B., "Optimal Auction Design," *Mathematics of operations research*, 1981, *6* (1), 58–73.

Nalebuff, B. J. and J. E. Stiglitz, "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *The Bell Journal of Economics*, 1983, *14* (1), 21–43.

Noussair, C. and J. Silver, "Behavior in all-pay auctions with incomplete information," *Games and Economic Behavior*, 2006, *55* (1), 189–206.

Nyarko, Y. and A. Schotter, "An Experimental Study of Belief Learning Using Elicited Beliefs," *Econometrica*, 2002, *70* (3), 971–1005.

Plott, C. and V. Smith, "An Experimental Examination of Two Exchange Institutions," *Review of Economic Studies*, 1978, *45* (1), 133–153.

Potters, J., C. G. de Vries, and F. van Winden, "An Experimental Examination of Rational Rent-Seeking," *European Journal of Political Economy*, 1998, *14* (4), 783–800.

Rosen, Sherwin, "Prizes and incentives in elimination tournaments," *American Economic Review*, 1986, *76* (4), 701–715.

Segev, E. and A. Sela, "Multi-Stage Sequential All-Pay Auctions," *European Economic Review*, 2014, *70*, 371–382.

Terwiesch, C. and Y. Xu, "Innovation Contests, Open Innovation, and Multiagent Problem Solving," *Management Science*, 2008, *54* (9), 1529–1543.

Tullock, G., *Efficient Rent Seeking*, College Station, Tx: Texas A&M University Press,

Weber, R. J., "Auctions and competitive bidding," in H. Peyton Young, ed., *Fair Allocation, American Mathematical Society Proceedings of Symposia in Applied Mathematics*, Vol. 33, Providence, Rhode Island: American Mathematical Society, 1985, pp. 143–170.

Zhang, J. and R. Q. Wang, "The Role of Information Revelation in Elimination Contests," *Economic Journal*, 2009, *119* (536), 613–641.

Online Appendices

A Equilibrium Characterizations

The unique BNE effort function in simultaneous contests has been derived in prior research (Chawla and Hartline, 2013; Hillman and Riley, 1989; Krishna and Morgan, 1997; Weber, 1985). For completeness, we reproduce the equilibrium effort function according to our model setup in Equation (2):

$$q^{Sim}(a_i) = \frac{(n-1)c}{1+(n-1)c} a_i^{1+(n-1)c} \tag{2}$$

The fact that $q^{Sim}(a_i)$ monotonically increases in a_i implies that in a simultaneous contest the contestant with the highest ability always wins (Archak and Sundararajan, 2009). $q^{Sim}(a_i)$, however, is not monotonic in n or in c, as can be revealed by taking derivatives with respect to these two variables.

For sequential contests, the BNE effort function depends on the value of c. The case of $c \geq 1$ is trivial, because in this case $F(x)$ is convex and each individual's maximization reaches a corner solution, resulting in all participants bidding zero in equilibrium. Therefore, for the rest of the paper, unless otherwise specified, we focus on the case of $c \in (0, 1)$, which implies that $F(x)$ is concave. Under this condition, Segev and Sela (2014) have derived the equilibrium effort function (Equation (3)) as follows.

$$q^{Seq}(a_i) = \begin{cases} 0 & \text{if } 0 \leq a_i < \overleftarrow{a}_i \\ \max\{q_j(a_j)\}_{j<i} & \text{if } \overleftarrow{a}_i \leq a_i < \overrightarrow{a}_i, \\ (a_i(1-d_i))^{\frac{1}{d_i}} & \text{if } \overrightarrow{a}_i \leq a_i \leq 1. \end{cases} \tag{3}$$

where $\overleftarrow{a}_i = [\max\{q_j(a_j)\}_{j<i}]^{d_i}$, $\overrightarrow{a}_i = \frac{1}{1-d_i}[\max\{q_j(a_j)\}_{j<i}]^{d_i}$, and $d_i \equiv (1-c)^{n-i}$.

This function is broken into three parts based on the range in which contestant i's ability factor falls, relative to the maximum effort by all preceding contestants, i.e., $\max\{q_j(a_j)\}_{j<i}$. In the first part, $\max\{q_j(a_j)\}_{j<i}$ is too large compared to a_i, and thus contestant i's best option is to "fold", i.e., to make zero effort and forgo the opportunity of winning. In the second part, $\max\{q_j(a_j)\}_{j<i}$ is not too large, so contestant i's best option is to play the "call" strategy, i.e., to make an effort equal to $\max\{q_j(a_j)\}_{j<i}$. In the third part, $\max\{q_j(a_j)\}_{j<i}$ is relatively small, so contestant i "raises" by choosing an effort, $(a_i(1-d_i))^{\frac{1}{d_i}}$, that is higher than $\max\{q_j(a_j)\}_{j<i}$. This effort level is a global maximizer of her expected payoff. Here, we use q_i^* to denote this global maximizer for player i, and further analyze its properties.

B Probability of Winning

Lemma 1. *The expected winning probability of contestant i, $P(i)$, increases in i ($i \in \{1, \cdots, n\}$).*[23]

Proof. Consider two contestants i and j ($j > i$) with ability factors a_i and a_j respectively. We only consider the cases in which either i or j wins the contest, and see if j has a higher chance of winning than i. As the distribution of a_i and a_j are *i.i.d.*, both have an equal chance to be larger.

First consider the case when $a_j > a_i$. If the winner is between i and j, it must be that $q_j \geq q_i$. Note that here $q_j = q_i$ is a shorthand for $q_j = q_i + \epsilon$, which would imply that the later player outperforms the earlier one. If player i folds, i.e., $q_i = 0$, then by definition $q_j \geq 0$, so $q_j \geq q_i$ holds. If player i raises, i.e., $q_i = q_i^*$, player j will either call or raise, since $a_j > (q_i^*)^{d_j}$ (due to $a_j > a_i$). Last, if player i calls, i.e., $q_i = \max\{q_k(a_k)\}_{k<i}$, then j will also either call or raise, again due to $a_j > a_i$.

Next, we consider the opposite case in which $a_i > a_j$. If $q_i = q_i^*$, then as long as $a_j > (q_i^*)^{d_j}$, player j will either call or raise. This is possible as we can easily verify that $(q_i^*)^{d_j} < a_i$. If $q_i = \max\{q_k(a_k)\}_{k<i}$, then $[\max\{q_k(a_k)\}_{k<i}]^{d_i} < a_i$. Since $d_j > d_i$, we arrive at $[\max\{q_k(a_k)\}_{k<i}]^{d_j} < a_i$, which shows that $a_j > [\max\{q_k(a_k)\}_{k<i}]^{d_j}$ is possible, in which case j will either call or raise. $\square$

Lemma 2 provides the closed-form solution for the winning probability of player i.

Lemma 2. *The expected probability of winning of player i in an n-player sequential contest is the sum of the following three items, A, B, and C:*

$$A = \sum_{m=1}^{i-1} \sum_{l=m+1}^{i} \frac{d_m}{\prod_{k=1}^{l-1}(1-d_k)^c} \Bigg[\frac{(1-d_{l-1})^{\frac{c}{d_{l-1}} \sum_{k=1,k\neq l\sim i}^{n} d_k} - (1-d_l)^{\frac{c}{d_l} \sum_{k=1,k\neq l\sim i}^{n} d_k}}{\sum_{k=1,k\neq l\sim i}^{n} d_k} - \frac{(1-d_{l-1})^{\frac{c}{d_{l-1}} \sum_{k=1,k\neq l\sim i-1}^{n} d_k} - (1-d_l)^{\frac{c}{d_l} \sum_{k=1,k\neq l\sim i-1}^{n} d_k}}{\sum_{k=1,k\neq l\sim i-1}^{n} d_k} \Bigg]$$

$$B = \sum_{m=1}^{i-1} \frac{d_m (1-d_i)^{\frac{c}{d_i} \sum_{k=1}^{n} d_k} (1-(1-d_i)^c)}{\left(\prod_{k=1}^{i}(1-d_k)^c\right) \sum_{k=1}^{n} d_k}$$

$$C = \frac{d_i (1-d_i)^{\frac{c}{d_i} \sum_{k=1}^{n} d_k}}{\left(\prod_{k=1}^{i}(1-d_k)^c\right) \sum_{k=1}^{n} d_k}$$

Note, the notation $\sum_{k=1,k\neq l\sim i}^{n} d_k$ means $\sum_{k=1}^{n} d_k - \sum_{k=l}^{i} d_k$.

Proof. For a given player i to win, there are only three scenarios:

[23] The explicit expression of $P(i)$ is derived in Lemma 2 in Appendix B.

A. There is an existing highest effort initially made by player m ($m < i$). $q_m = q_m^*$ and $\frac{(q_m^*)^{d_i}}{1-d_i} > 1$. Player i wins by calling, thus $q_i = q_m^* + \epsilon$. For any player j ($j > i$), $a_j < (q_m^*)^{d_j}$.

B. There is an existing highest effort initially made by player m ($m < i$). $q_m = q_m^*$, $\frac{(q_m^*)^{d_i}}{1-d_i} < 1$, and $(q_m^*)^{d_i} < a_i < \frac{(q_m^*)^{d_i}}{1-d_i}$. In this case, player i wins by calling, thus $q_i = q_m^* + \epsilon$. For any player j ($j > i$), $a_j < (q_m^*)^{d_j}$.

C. Player i's q_i^* is higher than any previous player's effort, such that player i wins by raising, thus $q_i = q_i^*$. For any player j ($j > i$), $a_j < (q_i^*)^{d_j}$.

We use A (B, C) to denote the probability of player i winning under the scenario A (B, C). For ease of exposition, let us define $\widehat{d_a d_b} = \frac{(1-d_a)^{\frac{d_b}{d_a}}}{1-d_b}$, which implies $\widehat{d_n d_b} = 0$. So $A = \sum_{m=1}^{i-1} \sum_{l=m+1}^{i} \int_{\widehat{d_l d_m}}^{\widehat{d_{l-1} d_m}} (\prod_{k=1,k\neq m}^{l-1} P(q_k^* < q_m^*)) P(a_i > (q_m^*)^{d_i}) \prod_{j=i+1}^{n} P(a_j < (q_m^*)^{d_j}) dF(a_m)$. $B = \int_0^{d_i d_m} \prod_{k=1,k\neq m}^{i-1} P(q_k^* < q_m^*) P((q_m^*)^{d_i} < a_i < \frac{(q_m^*)^{d_i}}{1-d_i}) \prod_{j=i+1}^{n} P(a_j < (q_m^*)^{d_j}) dF(a_m)$. $C = \int_0^1 \prod_{k=1}^{i-1} P(q_k^* < q_i^*) \prod_{j=i+1}^{n} P(a_j < (q_i^*)^{d_j}) dF(a_i)$.

□

C Proof for Proposition 1

Proof. We start by deriving the expected maximum effort if contestant i's effort is the highest, denoted by EME_i^{Seq}. For contestant i's effort to be the highest, first, it has to be that any preceding participant k's ($k \in \{1, \ldots, i-1\}$) effort, q_k, satisfies $\frac{q_k^{d_i}}{1-d_i} \leq a_i$.

Second, for q_i to be the highest, all participants after contestant i, thus contestant j ($j \in \{i+1, \ldots, n\}$), have to make zero effort or q_i ($q_i + \epsilon$, to be exact). Recall that contestant j's equilibrium effort function has three parts. The third part does not exist if $\frac{q_i^{d_j}}{1-d_j} \geq 1$, in which case, contestant j only has the choices of making zero effort or q_i, regardless of the value of a_j. For $\frac{q_i^{d_j}}{1-d_j} \geq 1$ to hold for participant j, a_i has to be in the range of $[\frac{(1-d_j)^{\frac{d_i}{d_j}}}{1-d_i}, 1]$.

Further, since $(1-d_j)^{\frac{1}{d_j}}$ is a decreasing function of j, if $\frac{q_i^{d_j}}{1-d_j} \geq 1$ is satisfied for contestant j, it is satisfied for all contestants after j. So to account for all possibilities, we start by integrating a_i over $[\frac{(1-d_{i+1})^{\frac{d_i}{d_{i+1}}}}{1-d_i}, 1]$, the range in which the value of q_i will make all subsequent participants choose either zero or q_i, followed by integrating a_i over the next lower range, $[\frac{(1-d_{i+2})^{\frac{d_i}{d_{i+2}}}}{1-d_i}, \frac{(1-d_{i+1})^{\frac{d_i}{d_{i+1}}}}{1-d_i}]$, and so on. Following the notations used in Lemma 2, we again define $\widehat{d_a d_b} = \frac{(1-d_a)^{\frac{d_b}{d_a}}}{1-d_b}$, which implies $\widehat{d_n d_b} = 0$.

Taken together,

$$
\begin{aligned}
\text{EME}_i^{Seq} &= \sum_{j=i+1}^{n} \int_{\widehat{d_j d_i}}^{\widehat{d_{j-1} d_i}} q_i \prod_{k=1}^{i-1} P(\frac{q_k^{d_i}}{1-d_i} < a_i) \prod_{m=i+1}^{j-1} P(a_m < \frac{q_i^{d_m}}{1-d_m}) f(a_i) da_i \\
&= \sum_{j=i+1}^{n} \int_{\widehat{d_j d_i}}^{\widehat{d_{j-1} d_i}} q_i \prod_{k=1,k\neq i}^{j-1} \left(\frac{(a_i(1-d_i))^{\frac{d_k}{d_i}}}{1-d_k} \right)^c f(a_i) da_i \\
&= \sum_{j=i+1}^{n} \int_{\widehat{d_j d_i}}^{\widehat{d_{j-1} d_i}} q_i a_i^{-c} \prod_{k=1}^{j-1} \left(\frac{(a_i(1-d_i))^{\frac{d_k}{d_i}}}{1-d_k} \right)^c c a_i^{c-1} da_i \\
&= \sum_{j=i+1}^{n} \frac{c(1-d_i)^{\frac{1}{d_i}}}{\prod_{m=1}^{j-1}(1-d_m)^c} \int_{\widehat{d_j d_i}}^{\widehat{d_{j-1} d_i}} a_i^{\frac{1}{d_i}-1} (a_i(1-d_i))^{\frac{c}{d_i}\sum_{k=1}^{j-1} d_k} da_i \\
&= \sum_{j=i+1}^{n} \frac{c(1-d_i)^{\frac{1}{d_i}}}{\prod_{m=1}^{j-1}(1-d_m)^c} \int_{\widehat{d_j d_i}}^{\widehat{d_{j-1} d_i}} a_i^{\frac{1}{d_i}-1} (a_i(1-d_i))^{\frac{c}{d_i}\frac{(1-c)^n}{c}(\frac{1}{(1-c)^{j-1}}-1)} da_i \\
&= \sum_{j=i+1}^{n} \frac{c\left((1-d_{j-1})^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_{j-1}}} - (1-d_j)^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_j}} \right)}{(\prod_{m=1}^{j-1}(1-d_m)^c)(\frac{1+c\sum_{m=1}^{j-1} d_m}{d_i})}
\end{aligned} \tag{4}
$$

Note that the above derivation uses the formula for the sum of a geometric series, i.e., $\sum_{k=1}^{j-1} d_k = (1-c)^n \sum_{k=1}^{j-1} \frac{1}{(1-c)^k} = \frac{(1-c)^n}{c}(\frac{1}{(1-c)^{j-1}} - 1)$.

Summing up EME_i^{Seq} over all $i \in \{1, \cdots, n\}$, we obtain

$$
\begin{aligned}
\text{EME}^{Seq} &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{c\left((1-d_{j-1})^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_{j-1}}} - (1-d_j)^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_j}} \right)}{(\prod_{m=1}^{j-1}(1-d_m)^c)(\frac{1+c\sum_{m=1}^{j-1} d_m}{d_i})} \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{c\left((1-d_{j-1})^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_{j-1}}} - (1-d_j)^{\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_j}} \right)}{(\prod_{m=1}^{j-1}(1-d_m)^c)(\frac{1+(1-c)^{n-j+1}-(1-c)^n}{d_i})}
\end{aligned} \tag{5}
$$

As n increases to infinity, EME^{Seq} has an upper bound of $\frac{1}{e}$. This can be easily seen since $EME^{Seq} < (1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}}$ and $\lim_{n\to\infty}(1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} = \frac{1}{e}$.

We have shown $EME^{Seq}(n,c) \leq \frac{1}{e}$, so $\sup EME^{Seq} \leq \frac{1}{e}$. On the other hand,

$$
ME^{Seq}(n,c) = max\{(a_i(1-d_i))^{\frac{1}{d_i}} | i = 1, 2, \cdots, n-1\}
$$

which implies,

$$ME^{Seq}(n,c) \quad \geq (1-d_{n-1})^{\frac{1}{d_{n-1}}}(max\{a_i|i=1,2,\cdots,n-1\})^{\frac{1}{d_1}}$$

Given that $F(a)=a^c$, we have

$$\begin{aligned} EME^{Seq}(n,c) &\geq c^{\frac{1}{1-c}} \int_0^1 a^{\frac{1}{(1-c)^{n-1}}}(n-1)ca^{(n-1)c-1}da \\ &= c^{\frac{1}{1-c}} \frac{(n-1)c}{(n-1)c+\frac{1}{(1-c)^{n-1}}} \end{aligned}$$

Since

$$\lim_{\substack{n\to\infty \\ c\to 1}} c^{\frac{1}{1-c}} \frac{(n-1)c}{(n-1)c+\frac{1}{(1-c)^{n-1}}} = \frac{1}{e}$$

i.e. $\sup EME^{Seq}(n,c) \geq \frac{1}{e}$. Taken together, $\sup EME^{Seq} = \frac{1}{e}$.

□

D Proof for Theorem 1

Proof. First of all, the case of $c \geq 1$ is trivial. In this case, $EME^{Seq}(n,c)=0$ whereas $EME^{Sim}(n,c)>0$, therefore $EME^{Sim}(n,c) > EME^{Seq}(n,c)$. For the rest of the proof, we focus on the case of $c \in (0,1)$.

Due to the complexity of the expression for EME^{Seq}, standard methods involving directly taking derivatives do not work. We start by simplifying the inequality we aim to prove. First, since the last bidder never raises, we have

$$ME^{Seq}(n,c) = max\{(a_i(1-d_i))^{\frac{1}{d_i}}|i=1,2,\cdots,n-1\}$$

which implies,

$$ME^{Seq}(n,c) \quad \leq (1-d_1)^{\frac{1}{d_1}}(max\{a_i|i=1,2,\cdots,n-1\})^{\frac{1}{d_{n-1}}} \tag{6}$$

$$ME^{Seq}(n,c) \quad \leq (1-d_1)^{\frac{1}{d_1}}(max\{a_i|i=1,2,\cdots,n-1,n\})^{\frac{1}{d_{n-1}}} \tag{7}$$

Given that $F(a)=a^c$, the distribution of the highest a_i out of $\{a_1,a_2,\cdots,a_n\}$ is $f_{a(n)}(a) = na^{nc-c}ca^{c-1} = nca^{nc-1}$, where $f_{a(n)}(a)$ denotes the n^{th} order statistic of a. Therefore, Equation (7) implies

$$EME^{Seq}(n,c) \leq (1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} \int_0^1 a^{\frac{1}{1-c}} nca^{nc-1} da$$
$$= (1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} \frac{nc}{nc+\frac{1}{1-c}} \tag{8}$$

Similarly, Equation (6) implies

$$EME^{Seq}(n,c) \leq (1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} \frac{(n-1)c}{(n-1)c+\frac{1}{1-c}} \tag{9}$$

The rest of the proof proceeds as follows. Starting with Equation (8), Steps $1 \sim 6$ prove that $EME^{Sim}(n,c) > EME^{Seq}(n,c)$ holds for $n \geq 5$. Starting with Equation (9), Step 7 proves that $EME^{Sim}(n,c) > EME^{Seq}(n,c)$ holds for $n < 5$. The reason why the proof has to be broken into two parts is because while Equation (8) is relatively easier to work with in proving the cases of $n \geq 5$, it cannot be used to prove the cases of $n < 5$ because it enlarges $EME^{Seq}(n,c)$ too much such that $(1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} \frac{nc}{nc+\frac{1}{1-c}} < EME^{Sim}(n,c)$ does not hold for $n < 5$.

Step 1. Show that inequality $\frac{Nc}{2Nc+c+1} > (1-(1-c)^N)^{\frac{1}{(1-c)^N}}$ ($N = n-1$) implies $EME^{Sim}(n,c) > EME^{Seq}(n,c)$.

Now, we re-write the expression for $EME^{Sim}(n,c)$ as $\frac{(n-1)c}{1+(n-1)c} \frac{nc+\frac{1}{1-c}}{2nc-c+1} \frac{nc}{nc+\frac{1}{1-c}}$. Comparing this expression and Equation (8), we know that if we could show $(1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} < \frac{(n-1)c}{1+(n-1)c} \frac{nc+\frac{1}{1-c}}{2nc-c+1}$, we would have shown $EME^{Sim}(n,c) > EME^{Seq}(n,c)$. Replacing $n-1$ with N, we obtain that if we could show that the following inequality holds for $N \geq 4$, we would have completed the proof (note that $\frac{nc+\frac{1}{1-c}}{1+(n-1)c} > 1$):

$$(1-(1-c)^N)^{\frac{1}{(1-c)^N}} < \frac{Nc}{2Nc+c+1} \tag{10}$$

Step 2. Show that if we could prove that Equation (11) has a unique solution of $c = 0$, we would have shown that inequality (10) holds.

$$(1-(1-c)^N)^{\frac{1}{(1-c)^N}} - \frac{Nc}{2Nc+c+1} = 0 \tag{11}$$

First, given that $N \geq 4$, when $c = 1$ we must have

$$\frac{Nc}{2Nc+c+1} - (1-(1-c)^N)^{\frac{1}{(1-c)^N}} = \frac{N}{2N+2} - \frac{1}{e} > 0.$$

If $\exists\, c \in (0,1)$, s.t.

$$\frac{Nc}{2Nc+c+1} - (1-(1-c)^N)^{\frac{1}{(1-c)^N}} < 0.$$

By continuity, $\exists\, \widetilde{c} \in (c, 1)$ s.t.

$$\frac{N\widetilde{c}}{2N\widetilde{c}+\widetilde{c}+1} - (1-(1-\widetilde{c})^N)^{\frac{1}{(1-\widetilde{c})^N}} = 0,$$

which contradicts the assumption that Equation (11) has a unique solution.

Step 3. For any $N > 0$ and $t \in \{1, 2, \cdots\}$, we construct a sequence c_t^N that satisfies the difference equation (Equation (12)), and show that, for any given N, proving that Equation (11) has a unique solution of $c = 0$ is equivalent to showing that the sequence c_t^N converges to zero, i.e., its infima is zero.

To construct the sequence c_t^N for any given N, we define a difference equation:

$$\frac{Nc_{t+1}}{2Nc_{t+1}+c_{t+1}+1} - (1-(1-c_t)^N)^{\frac{1}{(1-c_t)^N}} = 0. \tag{12}$$

Solving c_{t+1} from Equation (12), we get

$$c_{t+1}(N, c_t) = \frac{(1-(1-c_t)^N)^{\frac{1}{(1-c_t)^N}}}{N-(2N+1)(1-(1-c_t)^N)^{\frac{1}{(1-c_t)^N}}} \tag{13}$$

Define function $h(N, x) = \frac{x}{N-(2N+1)x}$, which is an increasing function of x. Define function $g(N, c) = (1-(1-c)^N)^{\frac{1}{(1-c)^N}}$, and we know $g(N, c)$ increases in c. Taken together, the properties of both function $h(N, x)$ and $g(N, c)$ imply that $\frac{\partial c_{t+1}(N, c_t)}{\partial c_t} > 0$.

Equation (12) induces a sequence $\{c_t^N\} =: \{c_t^N \text{ satisfy } (12) \mid c_0^N = 1\}$, which has two properties: 1) the sequence monotonically decreases; and 2) $c_t^N > 0$ for any $N > 0$ and any $t \geq 0$. We first establish property 1) by induction. Due to Equation (12) and $c_0^N = 1$, we know $c_1^N = \frac{1}{(e-2)N-1} < 1$, i.e. $c_1^N < c_0^N$. Given that $c_t^N < c_{t-1}^N$ and $\frac{\partial c_{t+1}(N, c_t)}{\partial c_t} > 0$, we know $c_{t+1}(N, c_t^N) < c_t(N, c_{t-1}^N)$, i.e. $c_{t+1}^N < c_t^N$, $\forall t \geq 0$. To establish property 2), since $\{c_t^N\}$ monotonically decreases, we know $c_t^N \leq 1$, which implies $0 \leq (1-(1-c_t^N)^N)^{\frac{1}{(1-c_t^N)^N}} \leq \frac{1}{e}, \forall t \geq 0$. It also should be noted that $h(N, x) \geq 0$, $\forall N \geq 4$, $\forall x \in (0, \frac{1}{e}]$. Therefore we have shown that $c_t^N > 0$ for any $N \geq 1$ and any $t \geq 0$.

Last, we show that the sequence $\{c_t^N\}$ converges to zero if and only if Equation (11) has a unique solution $c = 0$. To prove sufficiency, since $\{c_t^N\}$ is monotonic and bounded, it must converge. We denote its limit as $\widehat{c} = \lim_{t\to\infty} c_t^N$. Substituting $\{c_t^N\}$ into Equation (12) and taking limits on both sides, we arrive at

$$\frac{N\widehat{c}}{2N\widehat{c}+\widehat{c}+1} - (1-(1-\widehat{c})^N)^{\frac{1}{(1-\widehat{c})^N}} = 0,$$

Given that this equation has a unique solution of zero, we know $\widehat{c} = 0$, i.e. the sequence $\{c_t^N\}$ converge to zero.

To prove necessity, suppose there exists another solution to Equation (11), i.e., a $\bar{c} \in (0,1)$, s.t.

$$\frac{N\bar{c}}{2N\bar{c}+\bar{c}+1} - (1-(1-\bar{c})^N)^{\frac{1}{(1-\bar{c})^N}} = 0,$$

it means $c_{t+1}(\bar{c}, N) = \bar{c}$ by (13). Because $\bar{c} \in (0,1)$ and $\{c_t^N\}$ converges to zero, $\exists\, t,$ s.t. $c_t^N \geq \bar{c} > c_{t+1}^N$. But $c_t^N \geq \bar{c}$ and $\frac{\partial c_{t+1}(N,c_t)}{\partial c_t} > 0$ together imply $c_{t+1}(N, c_t^N) \geq c_{t+1}(N, \bar{c})$, i.e. $c_{t+1}^N \geq \bar{c}$, leading to a contradiction.

Taken together, given that $\{c_t^N\}$ monotonically decreases, if it converges to zero for any N, inequality (10) must hold. More precisely, if we can show that the infima of $\{c_t^N\}$ is zero for all $N \geq 4$, inequality (10) must hold. The next two steps (Step 4 and 5) establish this by induction.

Step 4. Show that for $N = 4, \forall c \in (0,1)$, the infima of the sequence $\{c_t^N\}$ is zero.

In preparation for the proof, we first establish a few useful limits and an inequality:

$$\lim_{c\to 0}(1-(1-c)^N)\ln(1-(1-c)^N) = 0 \tag{14}$$

$$\lim_{c\to 0}(1-(1-c)^N)^{\frac{1}{(1-c)^N}-1} = 1 \tag{15}$$

$$x < -\ln(1-x) < \frac{x}{1-x}, \forall x \in (0,1) \tag{16}$$

To prove Equation (14),

$$\begin{aligned}
\lim_{c\to 0}(1-(1-c)^N)\ln(1-(1-c)^N) &= \lim_{c\to 0}\frac{\ln(1-(1-c)^N)}{(1-(1-c)^N)^{-1}} \\
&= \lim_{c\to 0}\frac{(1-(1-c)^N)^{-1}N(1-c)^{N-1}}{-N(1-c)^{N-1}(1-(1-c)^N)^{-2}} \\
&= \lim_{c\to 0} -(1-(1-c)^N) \\
&= 0
\end{aligned}$$

To prove Equation (15),

$$\lim_{c\to 0}(1-(1-c)^N)^{\frac{1}{(1-c)^N}-1} = e^{\lim_{c\to 0}\frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N}} = 1 \text{ (by (14))}$$

To prove the inequalities in (16), we observe that when $x = 0$, $0 = -\ln(1-0) = \frac{0}{1-0}$. Taking derivatives with respect to x and note that $1 < \frac{1}{1-x} < \frac{1}{(1-x)^2}$, $\forall x \in (0,1)$, it is straightforward to show that inequalities in (16) hold.

Next, we define function $f(N,c) = \frac{Nc}{2Nc+c+1}$, and recall that we have defined $g(N,c) = (1-(1-c)^N)^{\frac{1}{(1-c)^N}}$. Then we derive their first, second, and third order partial derivatives with respect to c as follows,

$$
\begin{aligned}
\frac{\partial f(N,c)}{\partial c} &= \frac{N}{(2Nc+C+1)^2} \\
\frac{\partial^2 f(N,c)}{\partial c^2} &= \frac{-2N(2N+1)}{(2Nc+c+1)^3} \\
\frac{\partial^3 f(N,c)}{\partial c^3} &= \frac{6N(2N+1)^2}{(2Nc+c+1)^4} \\
\frac{\partial g(N,c)}{\partial c} &= (1-(1-c)^N)^{\frac{1}{(1-c)^N}}\left(\frac{N}{(1-c)(1-(1-c)^N)} + \frac{N\ln(1-(1-c)^N)}{(1-c)^{N+1}}\right) \\
\frac{\partial^2 g(N,c)}{\partial c^2} &= (1-(1-c)^N)^{\frac{1}{(1-c)^N}-1}\left(\frac{2N^2+N}{(1-c)^2} + \frac{N^2(1-(1-c)^N)(\ln(1-(1-c)^N))^2}{(1-c)^{2N+2}}\right. \\
&\quad \left. + \frac{2N^2\ln(1-(1-c)^N)}{(1-c)^{N+2}} + \frac{N(N+1)(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^{N+2}}\right).
\end{aligned}
$$

Observing that for any given N, we have $f(N,0) = g(N,0) = 0$, $\frac{\partial f(N,c)}{\partial c}|_{c=0} = \frac{\partial g(N,c)}{\partial c}|_{c=0} = N$ (by Equations (14), (15)), if we can find an open interval $(0, c(N))$, s.t. $\frac{\partial^2 f(N,c)}{\partial c^2} > \frac{\partial^2 g(N,c)}{\partial c^2}$, $\forall c \in (0, c(N))$, we must have $f(N,c) > g(N,c)$, $\forall c \in (0, c(N))$. In the following steps (a) $\sim$ (e), we will show that $c(N)$ exists for all N.

Step (a). By $\frac{\partial^3 f(N,c)}{\partial c^3} > 0$, we know $\frac{\partial^2 f(N,c)}{\partial c^2} > \frac{\partial^2 f(N,c)}{\partial c^2}|_{c=0} = -2N(2N+1)$, so $\frac{\partial^2 g(N,c)}{\partial c^2} < -2N(2N+1)$ would imply $\frac{\partial^2 f(N,c)}{\partial c^2} > \frac{\partial^2 g(N,c)}{\partial c^2}$.

Step (b). $\frac{\partial^2 g(N,c)}{\partial c^2} < -2N(2N+1) \Leftrightarrow (1-c)^2\frac{\partial^2 g(N,c)}{\partial c^2} < -(1-c)^2 2N(2N+1) \Leftarrow (1-c)^2\frac{\partial^2 g(N,c)}{\partial c^2} < -2N(2N+1)$. Thus, we just need to show $(1-c)^2\frac{\partial^2 g(N,c)}{\partial c^2} < -2N(2N+1)$.

Step (c). We observe, first, the last term of $\frac{\partial^2 g(N,c)}{\partial c^2}$ is less than zero. Second, by the inequalities in (16), we have $-1 < \frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N} < 0$, so $(1-(1-c)^N)^{\frac{1}{(1-c)^N}-1} = e^{\frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N}} \in (e^{-1}, 1)$, which implies $2N^2+N > (1-(1-c)^N)^{\frac{1}{(1-c)^N}-1}(2N^2+N)$. Utilizing these two observations, we have

$$
\begin{aligned}
&(1-c)^2\frac{\partial^2 g(N,c)}{\partial c^2} < 2N^2+N+(1-(1-c)^N)^{\frac{1}{(1-c)^N}-1} \\
&\left(\frac{N^2(1-(1-c)^N)(\ln(1-(1-c)^N))^2}{(1-c)^{2N}} + \frac{2N^2\ln(1-(1-c)^N)}{(1-c)^N}\right).
\end{aligned}
$$

If we can show $-2N(2N+1)$ is greater than the RHS of the above inequality, we would have shown $(1-c)^2\frac{\partial^2 g(N,c)}{\partial c^2} < -2N(2N+1)$. Collecting items and simplifying, we need to show the following,

$$-\frac{3(2N+1)}{N} > (1-(1-c)^N)^{\frac{1}{(1-c)^N}-1}\frac{\ln(1-(1-c)^N)}{(1-c)^N} \left(\frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N}+2\right) \tag{17}$$

Step (d). Substituting $x=(1-c)^N$ into inequalities (16), we have $-1 < \frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N} < 0$, implying $1 < \left(\frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N}+2\right) < 2$. Therefore, the following inequality holds,

$$\frac{\ln(1-(1-c)^N)}{(1-c)^N} > \frac{\ln(1-(1-c)^N)}{(1-c)^N}\left(\frac{(1-(1-c)^N)\ln(1-(1-c)^N)}{(1-c)^N}+2\right).$$

That is, the following inequality would imply that inequality (17) holds,
$-\frac{3(2N+1)}{N} > (1-(1-c)^N)^{\frac{1}{(1-c)^N}-1}\frac{\ln(1-(1-c)^N)}{(1-c)^N}$
By $(1-(1-c)^N)^{\frac{1}{(1-c)^N}-1} \in (e^{-1},1)$ and $(1-c)^N < 1$, we have

$$\begin{aligned}(1-(1-c)^N)^{\frac{1}{(1-c)^N}-1}\frac{\ln(1-(1-c)^N)}{(1-c)^N} &< e^{-1}\frac{\ln(1-(1-c)^N)}{(1-c)^N}\\ &< e^{-1}\ln(1-(1-c)^N).\end{aligned}$$

Therefore, to prove inequality (17) we need to establish the following inequality,

$$-\frac{3(2N+1)}{N} > e^{-1}\ln(1-(1-c)^N) \tag{18}$$

Step (e). Solving inequality (18), we arrive at

$$c(N) =: 1-(1-e^{\frac{-3e(2N+1)}{N}})^{\frac{1}{N}}$$

It is easy to check that $0 < c(N) < 1$ for any given N. By our definition of $c(N)$, whenever $c \in (0, c(N))$, we have $\frac{\partial^2 g(N,c)}{\partial c^2} < \frac{\partial^2 f(N,c)}{\partial c^2}$, which implies $g(N,c) < f(N,c)$. Going back to the case of $N=4$, we have $c(4) = 1-(1-e^{\frac{-27e}{4}})^{\frac{1}{4}} > 2.6\times10^{-9}$. That is, for $c \in (0, 2.6\times10^{-9})$, we have $f(N,c) > g(N,c)$ for $N=4$. Now, as long as we can show that the iteration process indicated by the difference equation (Equation (12)) can reach a t such that $c_t^4 < 2.6\times10^{-9}$, we would have shown that the infima of the sequence $\{c_t^4\}$ is zero. Indeed, with the help of MATLAB, we can calculate that such a t exists ($t = 6,246,758$) such that $c^4_{(6246758)} < 2.6\times10^{-9}$.

Step 5. Compare the two sequences $\{c_t^N\}$, and $\{c_t^{N+1}\}$. Show that $c_t^N > c_t^{N+1}$, $\forall t \geq 1$.

First, we show that if $g(N, c_t^N) \geq g(N+1, c_t^{N+1})$, then $g(N, c_{t+1}^N) > g(N+1, c_{t+1}^{N+1})$. By Equation (13) and our definition for $g(N, c)$, $h(N, x)$, we get

$$\begin{aligned} g(N, c_{t+1}(c_t, N)) &= \left(1-(1-c_{t+1}(N, c_t))^N\right)^{\frac{1}{(1-c_{t+1}(N,c_t))^N}} \\ &= \left(1-(1-h(N, g(N, c_t)))^N\right)^{\frac{1}{(1-h(N,g(N,c_t))^N}} \\ &= g(N, h(N, g(N, c_t))) \end{aligned}$$

Treating $g(N, c_t) = x \in (0, \frac{1}{e}]$ as a constant and consider the function

$$g(N, h(N, x)) = (1-(1-h(N, x))^N)^{\frac{1}{(1-h(N,x))^N}}$$

where $x \in (0, \frac{1}{e}]$. We claim that $\frac{\mathrm{d}g(N,h(N,x))}{\mathrm{d}N} < 0$.

$$\begin{aligned} \frac{\mathrm{d}g(N, h(N, x))}{\mathrm{d}N} &= \frac{\partial g}{\partial c}\frac{\mathrm{d}c}{\mathrm{d}N} + \frac{\partial g}{\partial N}\bigg|_{c=\frac{x}{N-(2N+1)x}} \\ &= (1-(1-c)^N)^{\frac{1}{(1-c)^N}}\left(\frac{1}{1-(1-c)^N} + \frac{\ln(1-(1-c)^N)}{(1-c)^N}\right) \\ &\quad \left(-\frac{Nx(1-2x)}{(1-c)(N-(2N+1)x)^2} - \ln(1-c)\right)\bigg|_{c=\frac{x}{N-(2N+1)x}} \end{aligned}$$

Using inequality (16), we obtain $\frac{1}{1-(1-c)^N} + \frac{\ln(1-(1-c)^N)}{(1-c)^N} > 0$, so the sign of $\frac{\mathrm{d}g(N,h(N,x))}{\mathrm{d}N}$ is the same as that of $-\frac{Nx(1-2x)}{(N-(2N+1)x)^2} - (1-c)\ln(1-c)$. Using inequality (16) again, we have $-(1-c)\ln(1-c) < c$. So if $-\frac{Nx(1-2x)}{(N-(2N+1)x)^2} + c < 0$, then $\frac{\mathrm{d}g(N,h(N,x))}{\mathrm{d}N} < 0$. Since $c = \frac{x}{N-(2N+1)x}$, inequality $-\frac{Nx(1-2x)}{(N-(2N+1)x)^2} + c < 0$ is equivalent to inequality $-\frac{N(1-2x)}{(N-(2N+1)x)} + 1 < 0$. Since $x \in (0, \frac{1}{e}]$, we have $-\frac{N(1-2x)}{(N-(2N+1)x)} + 1 = -\frac{x}{(N-(2N+1)x)} < 0$. So we get $\frac{\mathrm{d}g(N,h(N,x))}{\mathrm{d}N} < 0$. It means that if we have two points c_t^N and c_t^{N+1} in sequences $\{c_t^N\}$ and $\{c_t^{N+1}\}$ respectively, s.t. $g(N, c_t^N) = g(N+1, c_t^{N+1}) = x \in (0, \frac{1}{e}]$ (e.g., $g(N, c_0^N) = g(N+1, c_0^{N+1}) = \frac{1}{e}$), we must have $g(N, c_{t+1}^N) > g(N+1, c_{t+1}^{N+1})$. Further more, if we have $g(N, c_t^N) > g(N+1, c_t^{N+1})$,we also have $g(N, c_{t+1}^N) > g(N+1, c_{t+1}^{N+1})$. To see this, since $\frac{\partial g(N,c)}{\partial N} > 0$, we have $g(N+1, c_t^N) > g(N, c_t^N)$. By continuity of function $g(N, c)$, we know there must exist a $\widetilde{c_t}^{N+1} \in (c_t^{N+1}, c_t^N)$ s.t. $g(N, c_t^N) = g(N+1, \widetilde{c_t}^{N+1})$, which implies $g(N, c_{t+1}^N) > g(N+1, \widetilde{c_{t+1}}^{N+1})$. By $\frac{\partial c_{t+1}(N,c_t)}{\partial c_t} > 0$ we know $c_{t+1}(N+1, \widetilde{c_t}^{N+1}) > c_{t+1}(N+1, c_t^{N+1})$, i.e. $\widetilde{c_{t+1}}^{N+1} > c_{t+1}^{N+1}$, which implies $g(N+1, \widetilde{c_{t+1}}^{N+1}) > g(N+1, c_{t+1}^{N+1})$ (as function $g(N, c)$ monotonically increases in c). Taken together, we have $g(N, c_{t+1}^N) > g(N+1, c_{t+1}^{N+1})$.

Second, we show that $g(N, c_t^N) \geq g(N+1, c_t^{N+1})$ implies $c_{t+1}^N > c_{t+1}^{N+1}$. Re-write Equation (13) as $c_{t+1}(N, c_t) = h(N, g(N, c_t))$. Then, since $\frac{\partial h(N,c)}{\partial N} < 0$, $g(N, c_t^N) =$

$g(N+1, c_t^{N+1})$ implies $c_{t+1}^N > c_{t+1}^{N+1}$. Similarly, $g(N, c_t^N) > g(N+1, c_t^{N+1})$ also implies $c_{t+1}^N > c_{t+1}^{N+1}$. This is because, since $g(N+1, c_t^N) > g(N, c_t^N)$ (due to g monotonically increasing in N) and the fact that g is a continuous function, there must exist a $\widetilde{c_t}^{N+1} \in (c_t^{N+1}, c_t^N)$ s.t. $g(N, c_t^N) = g(N+1, \widetilde{c_t}^{N+1})$, which implies $c_{t+1}^N > \widetilde{c_{t+1}}^{N+1}$ as we just proved. Further, as $\frac{\partial c_{t+1}(N, c_t)}{\partial c_t} > 0$, we know $\widetilde{c_{t+1}}^{N+1} > c_{t+1}^{N+1}$. That is, $c_{t+1}^N > c_{t+1}^{N+1}$.

The previous two steps have established that $g(N, c_t^N) \geq g(N+1, c_t^{N+1})$ implies both $g(N, c_{t+1}^N) > g(N+1, c_{t+1}^{N+1})$ and $c_{t+1}^N > c_{t+1}^{N+1}$. Now consider the sequences $\{c_t^N\}$ and $\{c_t^{N+1}\}$. Since $g(N, c_0^N) = g(N+1, c_0^{N+1}) = \frac{1}{e}$, we know $g(N, c_1^N) > g(N+1, c_1^{N+1})$ and $c_1^N > c_1^{N+1}$. Then $g(N, c_1^N) > g(N+1, c_1^{N+1})$ implies $g(N, c_2^N) > g(N+1, c_2^{N+1})$ and $c_2^N > c_2^{N+1}$. This iterative process proves $c_t^N > c_t^{N+1}, \forall t$.

Step 6. Show (10) hold for $N \geq 4, \forall c \in (0,1)$.

Step 5 established that $c_t^N > c_t^{N+1}, \forall t > 0$ and $\forall N > 0$, which implies that the infima of the sequences $\{c_t^N\}$ are non-increasing in N. Together with the fact that the infima of the sequence $\{c_t^4\}$ is zero (step 4) and $c_t^N > 0$ (step 3), the infima of the sequence $\{c_t^N\}$ (for any N) must be zero, which implies inequality (10) holds for $N \geq 4, \forall c \in (0,1)$.

Step 7. Show that $EME^{Sim}(n,c) > EME^{Seq}(n,c)$ holds for $n < 5$.

Recall that we only need to show that

$$(1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} \frac{(n-1)c}{(n-1)c + \frac{1}{1-c}} \leq EME^{Sim}(n,c)$$

(Equation (9)) holds for $n = 2, 3$ and 4. Again, we rewrite $EME^{Sim}(n,c)$ as $\frac{\frac{1}{1-c}+(n-1)c}{1+(n-1)c} \frac{nc}{2nc-c+1} \frac{(n-1)c}{(n-1)c+\frac{1}{1-c}}$. Since $\frac{\frac{1}{1-c}+(n-1)c}{1+(n-1)c} \geq 1$, we only need to show

$$(1-(1-c)^{n-1})^{\frac{1}{(1-c)^{n-1}}} < \frac{nc}{2nc-c+1}$$

Replacing $n-1$ with N, the above inequality becomes

$$(1-(1-c)^{N})^{\frac{1}{(1-c)^{N}}} < \frac{(N+1)c}{2Nc+c+1}$$

Now, in Step 4. (steps (a) $\sim$ (e)), we have proved that for any N, there exists a $c(N)$, such that $\forall c \in (0, c(N))$, $(1-(1-c)^N)^{\frac{1}{(1-c)^N}} < \frac{Nc}{2Nc+c+1}$. This would imply that

$$(1-(1-c)^{N})^{\frac{1}{(1-c)^{N}}} < \frac{(N+1)c}{2Nc+c+1}$$

Therefore, using the same logic as in Step 4., for the rest of the proof, we only need to check that the sequence induced by

$$\frac{(N+1)c_{t+1}}{2Nc_{t+1}+c_{t+1}+1} - (1-(1-c_t)^N)^{\frac{1}{(1-c_t)^N}} = 0$$

can iterate from $c_0 = 1$ to $c(N)$, for $N = 1, 2$ and 3 (thus $n = 2, 3,$ and 4). It is easy to compute that $c(3) > 1.8 \times 10^{-9}$, $c(2) > 6 \times 10^{-10}$ and $c(1) > 2 \times 10^{-11}$. With the help of MATLAB, we can show $c^3_{(61)} < 1.8 \times 10^{-9}$, $c^2_{(42)} < 6 \times 10^{-10}$ and $c^1_{(31)} < 2 \times 10^{-11}$.

□

E Proof for Proposition 2

Proof. Let $\psi(a) = a - \frac{1-a^c}{ca^{c-1}}$. $\psi'(a) = 1 - \frac{1-a^c-c}{ca^c}$. $\psi'(a) > 0$ if and only if $a > \hat{a} = (\frac{1-c}{1+c})^{\frac{1}{c}}$. And $\psi(a) > 0$ if and only if $a > \tilde{a} = (\frac{1}{1+c})^{\frac{1}{c}}$. It is easy to see that $\tilde{a} > \hat{a}$.

As Myerson (1981) shows, the expected total effort can be expressed as

$$ETE = \int_A \sum_{i \in N} \psi(a_i) p_i(a) \, \mathrm{d}F^N(a)$$

where $A = (0, 1)^N$.

Then whenever there exists at least one player with ability higher than $\tilde{a}$, the part $\sum_{i \in N} \psi(a_i)p_i(a)$ in a simultaneous contest will be higher than that in a sequential contest, since the prize will be allocated to the one with the highest ability. The probability that everyone's ability is lower than $\tilde{a}$ is $F^N(\tilde{a}) = (\frac{1}{1+c})^N \xrightarrow{N \to \infty} 0$, which implies our Proposition 2.

□

F Experimental Instructions: SEQ2

Name: ID Number: Total Payoff:

This is an experiment in decision-making. You will make a series of decisions in the experiment, followed by a post-experiment questionnaire. **Please note that you are not being deceived and everything you are told in the experiment is true.**

Each of you has been assigned an experiment ID, i.e. the number on your index card. The experimenter will use this ID to pay you at the end of the experiment.

Rounds: The experiment consists of 20 rounds of three-person games. The first 2 rounds are practice rounds, i.e., you will not receive payments from these two rounds. The payment you earned in each of the remaining 18 rounds will be cumulated toward your final payment.

Grouping: At the beginning of each round, you will be randomly grouped with two other people in the room. You are equally likely to be grouped with anyone in the room.

Endowment and Prize: At the beginning of each round, each of you will be given 120 tokens. You will use these tokens to compete for a prize which is worth 100 tokens.

Winning: In each round of the game, you will choose an effort level and the person with the highest effort level among the three will receive the prize of 100 tokens.

Procedure: The three participants in each group are called: Participant 1, Participant2 and Participant 3 based on the order in which they enter the game. The order is randomly assigned by the computer at the beginning of each round.

1. If you are Participant 1, you will be the first one to choose your effort. After observing your effort, Participant 2 and 3will choose their efforts subsequently.

2. If you are Participant 2, you will first observe Participant 1' effort and then choose yours. Both yours and the Participant 1's efforts will be observed by Participant 3 before he/she makes a decision.

3. If you are Participant 3, you will observe Participant 1's and 2's efforts first and then choose yours.

Effort Range: Your effort can be any number with four decimal points between [0, 120]. Tie-Breaking: If two or three of you make the same highest effort, the computer will randomly choose one as the winner.

Ability Factor: The ability factor captures the idea that it sometimes costs more or less to make the same effort: A person with a higher ability has a lower cost. At the beginning of each round, the computer will randomly draw a different ability factor for everyone, following the distribution described below. Your ability factor is private information: only you know it and do not inform any other participant of your private information.

Your ability factor is a number randomly drawn between $[0, 1]$, according to the following distribution function: $F(x) = x^{0.25}$. Do not worry if you do not understand this equation the

Table 9: Ability Distribution

Ability Factor	Percentile
0	0
0.004	25%
0.06	50%
0.32	75%
1	100%

distribution will be illustrated in a table and a graph below. The table below shows the percentiles of the distribution of ability factors.

For instance, the 25th percentile is 0.004. It indicates that with a 25% chance, your ability factor is below 0.004. Thus with a 75% chance, it is above 0.004.

The 50th percentile is 0.06. It indicates that with a 50% chance, your ability factor is below 0.06 whereas with a 50% chance, it is above 0.06.

For your reference, we also provide you with a graph of the distribution of the ability factor below. As you can observe in the graph, you are more likely to receive a low ability factor than a high ability factor.

[Insert Figure]

Your net earnings in a game: In each round, your earnings in the game will be determined by (1) your effort; (2) other participants' efforts; (3) and your ability factor. Specifically,

Your net earnings in a game= The amount of the prize if you win-Your Cost = 100 or 0-your effort/ability factor For example, if in a given round, the computer draws an ability factor of 0.5 for you, and you choose an effort of 10, then

(1) If you win, then your earning will be 100-10/0.5 = 80 tokens

(2) If you lose, then your earning will be 0-10/0.5 = -20 tokens

Calculator: To help you calculate your earning, we provide a calculator on the top-right corner of your screen. If you enter an effort level that you are considering in the box and then click the "compute" button, your cost, net earnings if you win, net earnings if you lose will be shown in the result box. Remember, you will always pay your cost, no matter whether you win or lose. Moreover, if the calculator shows that your net earnings are negative even when you win, it suggests that the effort level you considered is too high. Predicting Your Winning Probability:

Before you choose an effort, you will be given an opportunity to earn extra money by predicting how likely you will win the game. You will be asked the following questions:

Given your ability factor: x, and your are Participant y, estimate the probability that you will win ()?

Estimate the probability that any other participant will win ()?

If you think there is a 90% chance that you will win and a 10% that someone else will win, answer 90 to the first question and 10 to the second. If you think there is a 67% chance that you will win and hence a 33% that someone else will, answer 67 to the first question and 33 to the second. Each number you answer should be an integer between 0 and 100 and the two numbers have to add up to 100.

You are paid based on the accuracy of your prediction. The more accurate your belief is, the more you will earn. Since your prediction is made before you know others' efforts, the best thing you can do is to simply state your true belief.

If you believe that you will win with a 100% chance and you actually win, you will be rewarded 2 tokens. If you believe that someone else will win with a 100% chance (in other words, you will lose for sure) and someone else actually wins, you will be rewarded 2 tokens as well.

Here's another example. Suppose you believe that you will win with a 90% chance and someone else will win with a 10% chance.

1. If you win, your prediction payoff is:

$$2 - (1 - 90\%)^2 - (0 - 10\%)^2 = 1.98$$

2. If someone else wins, your prediction payoff is:

$$2 - (0 - 90\%)^2 - (1 - 10\%)^2 = 0.38$$

Your Total Payoff in Each Round:

Your Payoff in each round= Your net earnings in the game +Your endowment of 120 tokens + Your payoff in predicting your winning probability

Cumulative Payoff: Your cumulative payoff will be the sum of your payoff in all paying rounds.

Feedback: At the end of each round, you will receive the feedback on your screen about the round.

History: Your ability factor, all three participants' efforts, your Participant ID in the group (1/2/3), and your net earnings in the game in each previous round will be displayed in a history box.

Exchange Rate: At the end of the experiment, the tokens you earned will be converted to U.S. dollars at the rate of $1 = 110$ tokens.

Please do not communicate with one another during the experiment or use your cell phone. No food is allowed in the lab either. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

G Post-Experiment Questionnaire

We are interested in whether there is a correlation between participants' decision behavior and some socio-psychological factors. The following information will be very helpful for our research. This information will be strictly confidential.

1. Gender
    (a) Male
    (b) Female
2. Ethnic Background
    (a) White
    (b) Asian / Asian American
    (c) African American
    (d) Hispanic
    (e) Native American
    (f) Other
3. Age
4. How many siblings do you have?
5. Grad/Year
    (a) Freshman
    (b) Sophomore
    (c) Junior
    (d) Senior
    (e) > 4 years
    (f) Graduate student
6. Major
7. Would you describe yourself as (Please choose one)
    (a) Optimistic
    (b) Pessimistic
    (c) Neither
8. Which of the following emotions did you experience during the experiment? (You may choose any number of them.)

(a) Anger
(b) Anxiety
(c) Confusion
(d) Contentment
(e) Fatigue
(f) Happiness
(g) Irritation
(h) Mood swings
(i) Withdrawal

9. In general, do you see yourself as someone who is willing, even eager, to take risks, or as someone who avoids risks whenever possible? [7 point likert]

10. Concerning just personal finance decisions, do you see yourself as someone who is willing, even eager, to take risks, or as someone who avoids risks whenever possible? [7 point likert]

11. In general, do you see yourself as someone who, when faced with an uncertain situation, worries a lot about possible losses, or someone who seldom worries about them? [7 point likert]

12. Concerning just personal finance decisions, are you someone who, when faced with an uncertain situation, worries a lot about possible losses, or someone who seldom worries about them? [7 point likert]

13. In general, how competitive do you think you are? [7 point likert]

14. Concerning just sports and leisure activities, how competitive do you think you are? [7 point likert]