Experimental Studies of Culture, Diversity and Crowdsourcing

by

Xiao Liu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Information)
in The University of Michigan
2012

Doctoral Committee:
Professor Yan Chen, Chair
Professor Jeffrey K. MacKie-Mason
Professor Scott E. Page
Associate Professor Lada A. Adamic
Assistant Professor Yusuf Can Masatlioglu

To my family and friends

Acknowledgments

First and foremost, I owe my deepest gratitude to my advisor, Professor Yan Chen, who has been the greatest mentor in both my academic and personal life. I still remember how, on the first day I came to the United States, she picked me up at the Detroit Airport and let me stay at her house for a week before I found an apartment. Since then, she has trained me as a rigorous researcher, trusted me as a young scholar with great potential, and has given me great support. Being her student is the most honorable thing in my past six years and all that I have learned from her will provide me valuable guidance in the future.

I am indebted to all my committee members. I want to thank Professor Jeffrey K. Mackie-Mason for his continuous guidance and support, especially during my job search. Professor Scott E. Page, who has been my mentor and coauthor throughout the past six years, taught me to think broadly and deeply. I appreciate all of the encouragement and help I received from Professor Lada A. Adamic in both teaching and research, and she is a role model for me as a successful researcher and supportive adviser. Professor Yusufcan Masatlioglu pushed me to work on the theory sections in my job market papers and he always believed that I was capable of meeting these challenges.

I am grateful to all of my collaborators who helped me to grow as a young scholar. In particular, I want to thank Professor Jenna Bednar for her confidence in my research ability. I really enjoyed working with Professor Margaret Shih and learned a lot from her about social psychology theory. Professor Sherry Li has been very supportive and shared her experiences as a young faculty member in economics with me.

I thank all faculty and student members in our Incentive-Centered Design group and the Behavioral and Experimental Economics Lab at the School of Information University of Michigan. I appreciate the support and advice given by Professors Erin L. Krupka, Paul Resnick, Rahul Sami and Neslihan Uler. I benefited greatly from my interactions with my senior fellows, including Drs. Lian Jian, Anna Osepayshvili, John Lin, Tapan Khopkar, Kan Takeuchi and Rick Wash. I also appreciate the help I received from other group members, such as Roy Chen, Jiang Jiang, Ming Jiang and Carrie Xu. Besides the support from my own research group, I am grateful for all the help that I received from other faculties, staffs and students at the School of Information, such as Professors Eytan Adar, Joan Durrance, Soo Young Rieh, Lionel Robert, Elizabeth Yakel, Tiffany Veinot and Ann Zimmerman. I also want to thank Drs. Beth St. Jean and Max Chen. I will remember all the joyful time that I spent with them during both school and leisure time. Additionally, I am grateful for the generous financial support I received, including NSF SBE-0524919, NSF SES-0720943, the Barbour Fellowship and the Rackham One-term Dissertation Fellowship.

I offer thanks as well to my friends, both the old ones in China and the new ones I have made in the last six years. I especially want to thank Drs. Chia-yin (Lyra) Chang and Wenjing Chen. Lyra, as my best friend, taught me to be more optimistic, caring and forgiving. She has made me a better person and I always appreciate being her friend. Wenjing, as my roommate for the past four years, has given me tremendous courage and unconditional trust. I also want to thank Dr. Li Yi, who opened her door and warmly welcomed me when I first came to Ann Arbor. I would like to thank Xin Liu and Lei Shi, whose apartment has been my warmest harbor in the past year. I am grateful to Dr. Jing Chen, Dr. Liuling Gong, Chenxi Shen, Chen Wang, Dr. Yipei Wang, Dr. Min Zhu and Uniko Chen for being so supportive and trustful. Last but not least, many thanks to my lifelong friends, Zhaoxiang Fu and Dr. Yue Xu. I know Zhaoxiang from childhood and our friendship will continue in the future. Yue, as the first friend that I met in high school, gave me advice at every turning point in my life and I trust her very much.

I thank my grandma, my parents and my brother. My grandma raised me for fourteen years with her great love. I appreciate all strict rules that my mother set for me when I was a little girl and they made me become an honest and responsible person. My father encouraged me to explore the world and provided me all the best things that he could get. I also thank my brother for things he taught me with great patience. Being a part of this family is the luckiest thing that could have happened to me and their love drives me to continue moving forward and excelling in both my career and life.

Table of Contents

List of Tables

Table

List of Figures

Abstract

Using experimental and behavioral economics methods, my dissertation focuses on culture, diversity and crowdsourcing. The first chapter investigates the role of culture and institutional contexts on individual behavior and institutional performance. In our experiments, subjects play two distinct games simultaneously with different opponents. Introducing entropy as an empirical measure of behavioral variation in a normal form game, we find that the strategies used in games with low entropy are more likely to be transferred to games with high entropy and are less subject to influence by other games.

The second chapter studies how Asians and Caucasians coordinate and cooperate with each other differently when their various dimensions of natural identities are salient. In an experiment conducted at two large public universities in the United States, we manipulate the salience of participants' multidimensional natural identities and find that Asians, especially those who are first-generation, exhibit significantly more in-group favoritism and out-group discrimination when their ethnic identity is salient. However, this significant group bias disappears when their school identity is primed.

The third and fourth chapters evaluate the performance of all-pay auctions in crowdsourcing labor markets, both in the field and the laboratory. In the field experiment conducted on Taskcn, we manipulate the size of the reward and the presence of a reserve in the form of the early entry of a high-quality submission. We find that a higher reward induces significantly more submissions and attracts higher quality users. However, unpredicted by theory, we find that high-quality users are significantly less likely to enter tasks where a high-quality solution has already been submitted, resulting in lower quality of subsequent submissions in such soft reserve treatments.

To directly compare the performance of simultaneous and sequential all-pay auctions both implemented in crowdsourcing sites, we conduct a laboratory experiment and find lower revenue in sequential than in simultaneous all-pay auctions. When the entry decisions are endogenous, individuals are more likely to enter an auction

early when it has a favor-early tie-breaking rule compared to a favor-late tie-breaking rule. In addition, in simultaneous all-pay auctions, our data is better rationalized by a risk-aversion model instead of the risk-neutral model.

Chapter 1
Introduction

A key assumption for traditional economic analysis is that individuals are fully rational and self-interested. Results from volumes of studies in experimental and behavioral economics challenge the theoretical predictions generated under this assumption. The first stream of research in my dissertation contributes to this literature by showing the effect of social context and natural identities on individual decisions and deepens our understanding of the modeling of individual behavior. In contrast, the second stream of research in my dissertation expands the scope of economic research by evaluating the performance of all-pay mechanisms implemented in crowdsourcing websites, which use online communities to outsource tasks and have recently become very popular online labor markets.

In Chapter 2, with my co-authors, Jenna Bednar, Yan Chen and Scott Page, I investigate the role of culture and institutional contexts on individual behavior and institutional performance. This work is motivated by the observation that policies geared toward the economic improvement of developing nations often fail. Furthermore, identical institutions implemented in different environments sometimes produce divergent outcomes. For example, identical institutional innovations implemented in northern and southern Italy in the 1970's performed quite differently. In contrast with the classical assumption in economics that an individual's decisions in one game are independent of their decisions in other games, we find evidence of behavioral spillovers from one game to another in laboratory experiments. In our experiments, subjects play two distinct games simultaneously with different opponents. We find that, consistent with our predictions, the strategies chosen and the efficiency of outcomes in one game depend on the nature of the other game. For example, subjects behave more selfishly in the prisoner's dilemma game when it is paired with a self-interest game than when it is played in isolation. Introducing entropy as an empirical measure of behavioral variation in a normal form game, we find that games with low entropy have a stronger influence on people's behavior than games with high entropy, and are less subject to influence by other games. Taken together, these findings suggest that people may not treat strategic situations in isolation, but may instead develop heuristics that they then apply across multiple games.

In chapter 3, along with my co-authors Yan Chen, Sherry Li, and Margaret Shih, I study the effect of individuals' natural identities and stereotypes on their coordination and cooperation. This study is inspired by the observations that today's workplaces are increasingly diverse across multiple demographic categories and that many collaborative projects are carried out globally. Organizations frequently face the challenge of integrating a diverse workforce. In practice, common identities which are shared by everyone have often been used to create common goals and values; however, this practice has not been evaluated. Combining theories and methods from economics and social psychology, we prime students' school identity as their common identity and their different ethnic identities as the fragmenting identities at both the University of Michigan (UM) and UCLA.[1] We find that Asians, especially first-generation Asians, are more likely to coordinate and cooperate with their Asian matches than with their Caucasian counterparts when their ethnic identity is primed. However, this significant in-group favoritism disappears when their school identity, rather than their ethnic identity, is primed. Interestingly, the school priming alleviates the negative effects of the competitiveness stereotype on cooperation among UCLA Asians but not among UM Asians, indicating different understanding of school identity between these two student populations. This is the first experiment in economics which evaluates the effects of common identity as a non-pecuniary source of worker motivation among an ethnically diverse group of participants. Furthermore, the first-generation Asians' stronger inter-ethnic group discrimination, as well as their positive response to the priming of a common organizational identity, has policy implications for the integration of immigrants into the workforce.

In contrast with Chapters 2 and 3, which reveal the missing factors in traditional economic models, Chapters 4 and 5 apply the game theory framework to study the new economic phenomenon. Specifically, we study the performance of all-pay auction mechanisms on crowdsourcing markets. With the increasing popularity of the Internet, Web 2.0 technology facilitates knowledge exchange and crowdsourcing on a scale never before experienced. Many people use tools such as wikis and Yahoo! Answers to both obtain and contribute information. Crowdsourcing sites, such as Amazon's Mechanical Turk and Topcoder.com, have changed the way people work by

[1] Priming is an experimental technique in psychology that entails introducing certain stimuli, such as a questionnaire, to activate individuals' social knowledge structures(Bargh and Chartrand, 1999).

enabling collaboration between geographically dispersed workers and by outsourcing tasks to individuals across the globe. In particular, some crowdsourcing sites implement all-pay auctions as their exchange mechanism. On these websites, a requester posts a task (such as designing a company logo) offering a certain amount of money as a reward, and then anyone can submit a solution. After receiving responses, the requester selects the best solution and rewards the person who provides it. In this process, every contestant expends effort and time, but only the winner gets the reward; hence, the term "all-pay". For example, Taskcn.com uses sequential all-pay auctions where contestants submit their solutions sequentially and can observe the content of prior solutions. In contrast, Topcoder.com, the world's largest software development website, implements simultaneous all-pay auctions where one cannot read others' solutions before submitting their own.

To study the impact of different design features on all-pay auction crowdsourcing sites, along with my co-authors Yan Chen, Jiang Yang, and Lada A. Adamic, I conduct a randomized field experiment on Taskcn.com. We systematically vary the size of the reward and the presence of a soft reserve in the form of the early entry of a high-quality submission. We find that a higher reward induces significantly more submissions and attracts higher-quality users. However, unpredicted by theory, we find that high-quality users are significantly less likely to enter tasks where a high-quality solution has already been submitted, resulting in lower quality in subsequent submissions in such soft reserve treatments.

Continuing this line of research, in Chapter 5, I compare the performance of sequential all-pay auctions with simultaneous all-pay auctions in a laboratory experiment. I find that the amount of bids, which approximates submission quality, is significantly higher in simultaneous all-pay auctions than in sequential all-pay auctions. In addition, regarding individuals' entry decisions, which approximate users' submission timing on crowdsourcing sites, I find that a favor-early tie-breaking rule, which is similar to "first come, first serve" among multiple best solutions, attracts more early entries. However, this effect is attenuated as players acquire more experience. In addition, in simultaneous all-pay auctions, individuals do not play a mixed strategy as predicted by the risk-neutral model. Instead, the data is better rationalized by a risk-aversion model.

Altogether, the first two studies help us understand the effect of social context and natural identities on individual behavior separately. The second two studies systematically investigate the performance of different all-pay auction mechanisms on crowdsourcing labor markets. In Chapter 6, I conclude and discuss future directions of my dissertation work.

Chapter 2

Behavioral Spillovers and Cognitive Load in Multiple Games: an Experimental Study

2.1 Introduction

In this paper, we describe laboratory experiments in which individuals simultaneously and repeatedly play two games with different opponents. We test whether an individual's play in one game is influenced by the other strategic interaction she faces. Multiple games can increase cognitive load preventing individuals from choosing efficient or even equilibrium behaviors. They can also induce behavioral spillovers in which individuals choose similar strategies in the two games. We find evidence of both psychological processes.

First, we find that although individuals are free to apply distinct strategies in each game, they instead develop and apply common behaviors across the two games. For example, when playing the prisoner's dilemma paired with a game of chicken, players alternate on the off-diagonals more often than they cooperate, as compared to when the prisoner's dilemma is played alone. Thus, behavior in one game appears to spill over into the second game.

The extent and direction of that spillover appears to depend on the complexity of the outcome dynamics. We introduce a measure game complexity: *entropy* of outcomes. Entropy captures behavioral variation, which should be correlated with the cognitive load induced by the game. For example, in a game with a dominant strategy, the efficient equilibrium is not complex. In the control sessions, when subjects play a single game of this type, the play of this game produces the same outcome in almost every period, and therefore has low entropy. Thus, using entropy as a measure, this game is less cognitively taxing than a prisoner's dilemma game, which produces substantial behavioral variation.

In our experiments, we compare outcomes in single-game controls to outcomes in two-game ensembles, as well as outcomes between ensembles. Using entropy of outcomes measured in the single-game control sessions, we can then make an *out-of-sample* forecast of game play in treatments with two-game ensembles. We find that in these two-game ensembles, prevalent strategies in games with low entropy are more likely to be used in the games with high entropy, but not vice versa.[1] In other words, subjects develop strategies for easier games and apply them to more complex games.

The subjects' performance in the two-game ensembles supports hypotheses of both cognitive load and behavioral spillovers. By cognitive load, we mean that subjects' cognitive constraints prevent them from playing both games optimally. By behavioral spillovers, we mean that strategies in one game bleed over into the other game. Cognitive load need not imply behavioral spillovers. A subject suffering from overload could opt to play a simple strategy unrelated to her strategy in the other game. Nor is cognitive load necessary for behavioral spillovers. For example, cooperation in the prisoner's dilemma could be enforced through either Tit for Tat or a grim trigger strategy. Either of these strategies could behaviorally align with an efficient strategy in some other game. Thus, the spillover would exist even though the subjects were not overloaded. That said, spillovers do imply a cognitive cost reduction. Also, given the games that we consider in this paper, the spillovers that we identify all include loss of efficiency, which is consistent with cognitive load.

To test for behavioral spillovers we posit hypotheses that are distinct from those that would be created by cognitive load alone. As just mentioned, cognitive load would imply that behavior varies significantly from control sessions when multiple games are played simultaneously but that the form of that variation would be independent of the other game in the ensemble (provided that game demands equal cognitive attention). That is, cognitive load would not be sufficient to enable the direction of behavioral deviance. Yet, we find that behavior in one game depends significantly and predictably upon what other game is included in the two-game ensemble. This finding suggests that variance in actions cannot be attributed exclusively to cognitive load but instead indicates the presence of behavioral spillovers.

Our findings have an important implication for the study of games and for social science research more generally. If behavior in one game depends on other games an individual plays, then social scientists, whether doing experimental, theoretical, or

[1]Our results are distinct from but complement earlier research on sequential behavioral spillovers, which have been interpreted as a form of priming or framing (Tversky and Kahneman, 1986).

empirical research, may need to consider the full *ensemble* of games that an individual faces (Samuelson 2001, Bednar and Page 2007). To date, almost all game-theoretic research focuses on individual games, as do most experiments. That norm is changing. A recent flurry of multiple game experiments demonstrates the existence and magnitude of ensemble effects. Our theory of behavioral spillovers can explain some of the findings in these experiments (Section 4.3).

The interest in multiple game experiments can be attributed partly to their ability to generate deeper insights into both individual and collective behavior. For learning about individual behavior, these experiments provide a laboratory in which subjects find themselves in a more cognitively taxing environment, one that resembles real world situations in which multiple stimuli simultaneously demand a person's attention. At the collective level, the findings from multiple game experiments may contribute to an institutional explanation for behavioral variations in the play of common games. Distinct sets of experiences or cases lead distinct communities to draw different analogies when constructing strategies (Gilboa and Schmeidler, 1995). Institutional interventions that take into account the behavioral repertoire of the relevant individuals may be more likely to succeed. A better understanding of behavioral spillovers can contribute to the analyses and design of institutions.

We have organized the paper as follows. In Section 4.3, we summarize the relevant theoretical and experimental literature. Section 5.4 describes the specific games included in this study and our experimental design. Somewhat unusually, in Section 2.4 we first present the results from the control sessions, where participants play a single game, and then develop our multiple-game hypotheses in Section 5.5, which are based on theory as well as results from the control sessions. Section 2.6 reports our findings on the ensemble effects. In Section 2.7, we discuss what these findings might mean and comment on potential future directions.

2.2 Literature Review

In this section we review the theoretical and experimental literature on multiple games. Samuelson (2001) formally models behavioral spillovers and cognitive load when people play multiple games. He assumes that people pay a cognitive cost to analyze a strategic interaction. More sophisticated analyses require more cognitive load. In his model, individuals maintain a stock of analogies to organize their reasoning. In the analysis of three different bargaining games, the ultimatum game, the

Rubinstein (1982) alternating offer bargaining game, and a tournament, he characterizes two equilibria, one in which the two bargaining games are played separately, and one in which they are played jointly. In the latter, players apply common analogies to disparate bargaining situations.

Samuelson's analysis is restricted to bargaining games. Bednar and Page (2007) examine behavioral spillovers and cognitive load effects in a broader class of six 2×2 games. They prove conditions for the existence and efficiency of behavioral externalities, using computational agent based models (Miller and Page, 2007). Their agent based models show that simple learning rules could locate the proposed equilibria when played in isolation. When agents needed to solve multiple games simultaneously, the agents often created routines that they applied across strategic domains. The agent based model generates behavioral spillovers; agents employed identical strategies in distinct games. The model also shows evidence of cognitive load: some ensembles of games outstrip the capacity of the agents to play each game optimally. In those cases, they find especially strong ensemble effects. Given their focus on ensemble effects, Bednar and Page (2007) provide the main theoretical foundation for the current paper.

In comparison to the action-bundling results of Samuelson (2001) and Bednar and Page (2007), Jehiel (2005) uses a belief-bundling approach, where a player forms expectations about the behavior of the other players by pooling together several contingencies (analogy class) in which these other players must move, and forms an expectation about the average behavior in each analogy class. In his analogy-based expectation equilibrium, a player with coarser beliefs could still adopt different actions in different normal form games.

We now review the emerging multiple games experiments, and use our entropy measure to explain some of the findings in these experiments. Falk, Fischbacher and Gaechter (2010) investigate social interaction effects when two identical coordination or public goods games are played simultaneously with different opponents, and find no behavioral spillovers between the two games, which is consistent with our prediction that two identical games with the same entropy should not influence each other. In comparison, Savikhin and Sheremeta (2012) study simultaneous play of a public goods game (low entropy) and a competitive lottery contest (high entropy). They find that cooperation in the public goods game reduces overbidding in the lottery, while contributions to the public good are not affected by the simultaneous participation in the lottery. This result is consistent with our prediction that game influence flows from low entropy games to high entropy games. In a third study, using both a sequential and a simultaneous treatment, Cason, Savikhin and Sheremeta (2012) report cooperation spillovers from the median-effort game (low entropy) to the subsequent minimum-effort game (high entropy) when the games are played sequentially, but not simultaneously. Again, this result is consistent with our predictions based on entropy. Finally, Cason and Gangadharan (2010) investigate behavioral spillovers between a threshold public goods game and a competitive double auction market. They find that cooperation in public goods provision is less common when players simultaneously compete in the market.

There also exist sequential multiple games experiments. These studies identify significant framing (Tversky and Kahneman, 1986) and learning transfer effects. Both of these phenomena are related to behavioral spillovers (Cooper and Kagel, 2008; Haruvy and Stahl, 2010). Consider the experiments that first auction off the right to play in a game. These experiments produce different outcomes in subsequent games. If we think of the auction as an initial game, we can interpret the resulting improved outcomes as resulting from some sort of behavioral spillover. For example, Van Huyck, Battalio and Beil (1993) demonstrate that without a pre-play auction, the median-effort coordination game played in isolation leads to inefficient equilibrium but that auctioning off the right to play before the coordination game leads to the payoff-dominant equilibrium. Crawford and Broseta (1998) explore the efficiency-enhancing effect of auctions theoretically using a model of stochastic, history-dependent learning dynamics, giving an analytic explanation for these results.[2] Lastly, Huck, Jehiel and Rutter (2010) study feedback spillovers in sequential multiple games and find empirical support for an analogy-based expectation equilibrium (Jehiel, 2005). They use a different protocol from the multiple game experiments discussed above. In their experiment, a player plays one of two games in each round, and sometimes receives the aggregate distribution of the play of the opponents over the two games, a design feature aimed to compare the long-run behavior in the presence and absence of feedback spillovers. Grimm and Mengel (2010) use a similar protocol where a player plays one of several games each round, and find that their data can be rationalized by either action- or belief-bundling.

The experiments in this paper differ from these aforementioned studies, in that we consider pairs of games chosen from an ensemble of similar games. Therefore,

[2]A additional strand of research that complements our findings looks at the emergence of cooperation. Weber (2006) reports the results of a minimum-effort game experiment where successful coordination is achieved in large groups by starting with small groups and adding entrants who are aware of the group's history. Successful coordination in large groups can be interpreted as learning transfer from small groups that find it easier to coordinate.

we are not only able to show the presence of spillovers and cognitive load across a specific pair, but also we are able to demonstrate the comparative size of those effects based on the characteristics of the games being played. Thus, these experimental results provide a foundation for the behaviors exhibited in the follow up studies that we mention above.

2.3 Experimental Design

In this section, we describe the specific games included in our study and then provide a detailed description of our experimental procedures.

2.3.1 The Games

We choose four games which are variants of the six studied by Bednar and Page (2007). Here we test whether the phenomena derived within models and generated by artificial agents can be produced in a laboratory with real people. We focus on four 2×2 games: the Prisoner's Dilemma (PD), Strong Alternation (SA), Weak Alternation (WA), and a Self Interest game (SI).

The individual games belong to a class of two-person two-action games that contain a self-regarding action S and an alternative, C; in three of the games (PD, SA and WA), this alternative is cooperative. In these three games, cooperation lowers a player's own payoff and raises the payoff of the other, and being selfish does the opposite, so in the one shot game, the unique dominant strategy equilibrium involves both players choosing selfish. In the fourth game, Self-Interest (SI), S is both the stage game dominant strategy and Pareto dominant.

The first game is a standard Prisoner's Dilemma, where the stage game has a dominant strategy equilibrium, (S, S), which is Pareto dominated by (C, C). Note that (C, C) also maximizes the joint payoff of the two players.

Prisoner's Dilemma: (PD)

	C	S
C	7, 7	2,10
S	10,2	4,4

In the second and third games, Strong Alternation (SA) and Weak Alternation (WA), while (S, S) remains the dominant strategy equilibrium for the stage game, agents do best (i.e., maximize joint payoff) in repeated play by alternating between the off-diagonals, (C, S) and (S, C). In Strong Alternation, the incentives to alternate are much stronger than in Weak Alternation. The alternation games are a distant cousin to the conventional Battle of Sexes and Game of Chicken, where agents are rewarded for coordinating their behavior. In our alternation games, four outcomes are rewarded with positive payoffs: CC, SS, and the alternating strategies of CS then SC and SC then CS. Coordinating on CC or SS is much less taxing than working out an alternating behavior, and the positive payoffs for each reduce the focality of an alternating equilibrium.

Strong Alternation: (SA)

	C	S
C	7, 7	4,14
S	14,4	5,5

Weak Alternation: (WA)

	C	S
C	7, 7	4,11
S	11,4	5,5

In the final game, Self Interest (SI), the dominant strategy equilibrium, (S, S), also Pareto dominates all other outcomes. Furthermore, in the stage game, S uniformly dominates C.

Self Interest: (SI)

	C	S
C	7, 7	2,9
S	9,2	10,10

2.3.2 Experimental Procedure

Our experiments consist of four control sessions, each of which consists of a single game, and 14 treatment sessions, each of which consists of a pair of games. This experimental design enables us to determine the effects of ensemble on behavior by comparing the ensemble with the corresponding control sessions and to compare behavior across ensembles.

The control sessions follow a common protocol for infinitely repeated games in the laboratory. We have one 12-player session for each of the single games. Participants are randomly matched into pairs at the beginning of each session, and play the same match for the entire experiment. In each session, participants first play the game for 200 rounds. After round 200, whether the game will continue to the next round depends on the "throw of the die" that is determined by the computer's random number generator. At the end of each round after round 200, with 90% chance, the game will continue to the next round. With 10% chance, the game stops. In other words, we implement an infinitely repeated game, with a discount factor of 1 for the first 200 rounds, and 0.9 thereafter. With the chosen discount factors, (C, C) can be sustained as a repeated game equilibrium in PD, SA and WA. With 12 players in each control session, we have 6 independent observations for each single game.

In the ensemble treatment, we again use twelve players in each session. Within each session, at the beginning, each player is randomly matched with two other participants, both of whom will be her matches for the entire experiment. She plays two distinct games, one with each of her matches. For example, in the (PD_l, WA_r) ensemble, a player plays PD with her left match, and WA with her right match, as displayed in the following table using neutral action labels. This design allows us to analyze whether or not behavior in one game is influenced by the nature of the other game.

		Column Player			Column Player	
		Left (A)	Right (B)		Left (A)	Right (B)
Row	Top (A)	7, 7	2,10	Top (A)	7, 7	4,11
Player	Bottom (B)	10,2	4,4	Bottom (B)	11,4	5, 5

As in the control sessions, we implement an infinitely repeated game, with a discount factor of 1 for the first 200 rounds, and 0.9 thereafter. Within each session, the twelve players are partitioned into independent groups of 4 each, yielding 3 independent observations. We number our players from one to twelve. The matching protocol is the following: $\underbrace{4-2-1-3}$, $\underbrace{6-5-7-8}$, $\underbrace{10-9-11-12}$ form three independent groups, each with four participants positioned on a circle, and each participant plays her left and right match.

As the two games are displayed side by side, we conduct two independent sessions for each game ensemble, changing the order of the display to avoid the order effect within each round. For example, for the game ensemble of SA and WA, we display SA as the left game in one session, and WA as the left game in another session. This way, if a player always makes decisions from left to right, we have a balanced number of observations for each order.

We used z-Tree (Fischbacher, 2007) to program our experiments. As z-Tree does not record the mouse movements within each stage, we ran two additional sessions with ensembles, (SI, WA) and (WA, SI), where we use the software Morae to record the mouse movement. These two sessions enable us to determine the order of decisions within each round. The (SI, WA) session has 12 subjects, while the (WA, SI) has only eight subjects.[3]

Table 2.1 Features of Experimental Sessions

Control		Ensemble Treatment		
Game n	Pairs	(Left, Right)	n	Groups
		(PD, WA)	12	3
		(WA, PD)	12	3
		(PD, SI)	12	3
PD 12	6	(SI, PD)	12	3
		(SA, WA)	12	3
		(WA, SA)	12	3
		(SA, PD)	12	3
SA 12	6	(PD, SA)	12	3
		(SI, WA)	$12 + 12$	6
		(WA, SI)	$12 + 8$	5
		(SI, SA)	12	3
SI 12	6	(SA, SI)	12	3
WA 12	6			
Total 48	24		164	41

Table 4.2 reports features of experimental sessions, including the name of the game, the number of players in each session, the number of independent pairs for each control session, the ensemble of games, the number of players in each session, as well as the number of independent groups in each ensemble session.

Overall, 18 independent computerized sessions were conducted in the RCGD lab at the University of Michigan from March to October 2007, yielding a total of 212 subjects. Our subjects were students from the University of Michigan, recruited by email from a subject pool for economic experiments.[4] Participants were allowed to participate in only one session. Each ensemble treatment session lasted approximately 90 minutes, whereas each control session lasted about 45 minutes. The exchange rate was set to 100 tokens for $1. In addition, each participant was paid a $5 show-up fee. Average earnings per participant were $37.49 for those in the treatment sessions and $22.77 for those in the control sessions. Data are available from the authors upon request.

[3]We recruited for twelve subjects, however, only eight showed up.
[4]Graduate students from the Economics Department are excluded from the list.

2.4 Results: Control Sessions

In this section, we report the results from the control sessions at the outcome level. This analysis provides a benchmark from which we can identify the presence of cognitive load and behavioral spillovers results in Section 2.6. In Subsection 2.6.2, we infer the repeated game strategies emerged in each game in the control and compare them with those in the ensembles. In this section, we treat each pair as an independent observation.

We first introduce an empirical measure of cognitive load. To measure the behavioral variation in a game, we apply a standard entropy measure to the outcome distributions.[5] The *entropy* of a random variable X with a probability density function, $p(x) = Pr\{X = x\}$, is defined by

$$H(X) = -\sum_{x} p(x) \log_2 p(x).$$

Entropy measures the amount of stochastic variation in a random variable that can assume a finite set of values. Therefore it is also a measure of the amount of information required to describe that distribution. When using logarithms to base two, that measure captures the number of binary variables (bits) needed to describe the data.

For the analysis of two-person games, we model individual stage game strategies as a discrete random variable, X, with realizations in one of the four cells. Throughout the analysis, we use the convention that $0 \log 0 = 0$.[6] The entropy in a generic 2×2 game is in the interval $[0, 2]$, with the lower bound indicating certainty, i.e., all outcomes are in one cell, and the upper bound indicating a uniform distribution among the four cells. The cause of behavioral variation could be strategic uncertainty over what the other player will do.

In Figures 1-4, we present time series data for each pair in each of the control sessions, with the entropy for each pair presented at the bottom of each graph.

[Figure 2.1 about here.]

Figure 2.1 presents outcomes in the Self Interest game. In this game, all six pairs converge to the Pareto dominant stage game equilibrium quickly and stay there. The entropy for each pair ranges from 0 to 0.04, indicating very little behavioral variation.

[5] Shannon (1948) is credited with the development of the concept of entropy and the birth of information theory. Many basic concepts and findings in this field are summarized in Cover and Thomas (2006).

[6] This convention is easily justified by continuity, since $x \log x \to 0$ as $x \to 0$.

This behavioral consistency is likely attributable to the uniform dominance property of the dominant strategy equilibrium in the stage game. Additionally, participants take an average of 0.62 seconds per round to make a decision in SI, significantly shorter than in any other game ($p \leq 0.01$, one-sided permutation tests). Based upon the uniform dominance property of the unique Pareto efficient stage game equilibrium, its low entropy, and response time, we posit that SI imposes the least cognitive load.

[Figure 2.2 about here.]

Figure 2.2 presents behavior in the Prisoner's Dilemma game. In this game, over half of the pairs play CC, the efficient outcome, which is consistent with findings from previous experiments (Andreoni and Miller, 2002). Curiously, one pair also alternate for a fair number of rounds. The entropy for each pair ranges from 0.08 to 1.79, indicating changing behavioral variation. In addition, participants take an average of 1.00 second per round to make a decision in PD, significantly longer than SI, but shorter than SA ($p \leq 0.01$, one-sided permutation tests). As a "context" this game does not establish as strong a behavioral norm as the Self Interest game. Based upon this finding, we anticipate that PD will have a weaker behavioral pull than SI. The difficulty of learning to cooperate in the PD game may limit its spillover effects on play in other games.

[Figure 2.3 about here.]

Figure 2.3 presents behavior in the Strong Alternation game, where 5/6 of the pairs successfully establish the alternation outcomes. Pair 2 also attempts alternation on and off during the experiment. The entropy for each pair ranges from 1.29 to 1.90, indicating substantial behavioral variation.[7] In addition, participants take an average of 2.72 seconds per round to make a decision in SA, significantly longer than in any other game ($p \leq 0.01$, one-sided permutation tests). We interpret the longer response time in SA as evidence that coordinated alternation requires greater cognitive effort. Since successful alternation is established in five out of six pairs, this game also provides a strong context from which spillovers might occur.

[Figure 2.4 about here.]

[7]Perfect coordinated alternation results in an entropy of 1.

Last, Figure 2.4 presents the dynamics from the Weak Alternation game. In this game, only two out of six pairs develop an alternating behavior, two pairs cooperate, one (pair 4) converges to SS, and the last pair (pair 6) does not seem to have converged to a stable outcome. The entropy for each pair ranges from 0.44 to 1.91, with the highest aggregate entropy among all four games. In addition, participants take an average of 1.24 seconds per round to make a decision in WA, significantly longer than SI, shorter than SA ($p \leq 0.01$, one-sided permutation tests), not significantly different from PD ($p = 0.138$, one-sided permutation test). As WA results in higher behavioral variation, we speculate that subject behavior in WA is more likely to be influenced by the other game in an ensemble.

Table 2.2 Distribution of Outcomes and Entropy in Control Sessions

	SI		PD		SA		WA	
	C	S	C	S	C	S	C	S
C	0.00	0.14	55.68	11.67	5.02	39.81	33.18	21.57
S	0.00	99.86	14.82	17.82	40.37	14.81	22.74	22.51
Entropy	0.02		1.68		1.68		1.98	

To summarize our findings, Table 2.2 reports the aggregate distribution of outcomes in each of the four games in the control sessions, and the respective entropy for each game in the last line. The behavioral variation measured by entropy is the lowest in the Self Interest game (0.02), followed by Prisoner's Dilemma and Strong Alternation (1.68), and Weak Alternation (1.98). Using the entropy for each pair of players as an independent observation, we find that the ranking between SI and any other game is significant ($p = 0.001$, one-sided permutation tests) while other pairwise comparisons are not significant at the 10% level. Based on the entropy comparisons between games, we develop a partial ordering of the four games in terms of cognitive load: SI $<$ PD $\sim$ SA $\sim$ WA. This partial ordering is consistent with that obtained from the response time comparisons, which has been used to measure cognitive activities in psychology (Luce, 1986) and more recently, in behavioral economics (Rubinstein, 2007). We postulate that games with lower entropy might have stronger behavioral spillovers than those with higher entropy.

In our outcome level analysis, we focus on the Pareto efficient outcomes: (1) for both players to always play selfish each round (SS) in the Self Interest game; (2) for both players to always cooperate (CC) in Prisoner's Dilemma; and (3) coordinated alternation between S and C (ALT) in the Strong and Weak Alternation games, respectively. These outcomes also coincide with the simplest equilibria among the set of Pareto efficient ones in our set of games.[8]

Table 2.3 Average Proportion of Pareto Efficient Outcomes in Control Sessions

	% Outcomes			P-values of Permutation Tests		
Games	SS	CC	ALT	CC v. SS	CC v. ALT	SS v. ALT
SI	**99.86**	0.00	0.00	0.000	0.500	0.000
PD	17.82	**55.68**	15.44	0.039	0.031	0.389
SA	14.81	5.02	**71.12**	0.040	0.000	0.001
WA	22.51	33.18	**36.14**	0.308	0.430	0.317

Table 2.3 reports the proportion of three outcomes in each game over the entire series. Boldfaced numbers are the mode of the distribution. Within each game (row), we compare the proportion of pairs of outcomes, using one-sided permutation tests. The general null hypothesis is equal proportions, while the alternative hypothesis is that the proportion of the Pareto efficient outcome (boldfaced) is higher than any other outcome. The last three columns report the corresponding p-values for each pairwise test.

In the control sessions, the proportion of Pareto efficient outcomes (SS in SI, CC in PD, ALT in SA and WA) is significantly higher than any other outcomes in SI, PD and SA ($p < 0.05$). However, in WA, there is no significant difference in the proportions of any of the three outcomes ($p > 0.10$).

In sum, three distinct outcomes emerge in the control sessions: in the SI game, selfishness; in the PD game, cooperation; and in SA and WA an alternating form of cooperation, where subjects alternate between the cooperative and selfish actions. Weak Alternation has weaker incentives, so the coordinated alternation is not as prominent as with Strong Alternation.

2.5 Hypotheses of Ensemble Effects

In this section, we present a set of hypotheses testing the null of game independence against our two posited ensemble effects, behavioral spillovers and cognitive load. Our alternative hypotheses, Hypothesis 1 - 6, are based on the theoretical results from Bednar and Page (2007), as well as our empirical results from the control sessions presented in Section 2.4. These hypotheses are also broadly consistent with the

[8]When we represent a repeated game strategy as an automaton, the simplest strategy is defined as one with the least number of states (Baron and Kalai, 1993; Kalai and W.Stanford, 1988). We present our repeated game strategy analysis in Subsection 2.6.2.

analogy based model of Jehiel (2005) and the case based reasoning of Gilboa and Schmeidler (1995).

Our general null hypothesis is of game independence: play in one game is not be affected by the existence of another game to play. If the independence hypothesis is correct then we should see no difference between behaviors in the control sessions (games played in isolation) and when games are presented to subjects as part of ensembles, nor should we see any difference in behavior in one game when it is paired with different games.

Based on results from Bednar and Page (2007), we anticipate that the ensemble play will depend upon *which* other game is in the ensemble. Since SI has significantly lower entropy than any of the three other games, we expect the dominant outcome in SI more likely to appear in the game it is paired with, but not vice versa. Specifically, we expect:

Hypothesis 1 (Effects of SI)**.** *Compared to the corresponding control or other ensembles, games paired with Self-Interest will exhibit more selfishness.*

Since pairwise entropy comparisons among PD, SA and WA are not significant, behavior in each game could influence or be influenced by the game it is paired with. We base the following alternative hypotheses on other results from Bednar and Page (2007).

Hypothesis 2 (Effects of PD)**.** *Compared to the corresponding control or other ensembles, games (excluding SI) paired with the Prisoner's Dilemma will exhibit more cooperation.*

Hypothesis 3 (Effects of SA)**.** *Compared to the corresponding control or other ensembles, games (excluding SI) paired with Strong Alternation will exhibit more alternation.*

Hypothesis 4 (Effects of WA)**.** *Compared to the corresponding control or other ensembles, games (excluding SI) paired with Weak Alternation will exhibit more alternation.*

Payoff parameters for the two alternation games produce stronger incentives to alternate in Strong Alternation compared to Weak Alternation. This is confirmed by outcomes in the control sessions, with 71% (resp. 36%) of alternation in SA (resp. WA). Therefore, in ensembles, we expect to see more alternation in games paired with SA.

Hypothesis 5 (Efficient Outcomes: Ensemble vs. Control)**.** *Compared to the corresponding control sessions, subjects in a ensemble treatment will less often achieve efficient outcomes in any game with non-trivial entropy.*

Specifically, Hypothesis 5 implies that subjects in PD, SA or WA (each with non-trivial entropy) will less often produce efficient outcomes when each game is part of an ensemble compared to the corresponding control, while those in SI (with trivial entropy) will not behave differently in ensembles compared to the SI control.

In Section 2.4, we develop a partial ordering of the four games based upon the entropy of each game in the control sessions, i.e., the behavioral variation follows the order of SI $<$ PD $\sim$ SA $\sim$ WA. Based on the entropy, response time and the payoff structure of each game, we posit that Self Interest is the only easy game to play so it will be the only game for which we do not expect to see a significant falloff in efficient outcomes in the game SI is paired with. Thus, we formulate a hypothesis based on the cognitive load of the context game.

Hypothesis 6 (Efficient Outcomes: Ensemble vs. Ensemble)**.** *Participants are more likely to achieve Pareto efficient outcomes in a game when it is paired with SI than when it is paired with other games.*

2.6 Results: Ensemble Effects

In this section, we present ensemble effects at the outcome level (subsection 2.6.1) as well as those at the strategy level (subsection 2.6.2). In all our analysis in this section, a pair in a control session or a group of four in an ensemble session is treated as an independent observation.

2.6.1 Ensemble Effects at the Outcome Level

Our anticipation was that subjects would play particular games differently between the control sessions, where they played a single game, and when that game appeared as part of an ensemble. This prediction emerges from the two core hypotheses: both behavioral spillovers and cognitive load will affect play in ensembles. Consequently, we expect different outcomes between the control sessions and the corresponding ensembles. To establish the existence of behavioral spillovers in the presences of cognitive load, we compare outcomes between ensembles.

Table 2.4 Distribution of Outcomes in the Control and Ensembles

	SI			PD			SA			WA		
	(SS	CC	ALT)	(SS	CC	ALT)	(SS	CC	ALT)	(SS	CC	ALT)
SI	(100	0	0)	(99	0	0)	(99	0	0)	(99	0	0)
PD	(46	42	5)	(18	56	15)	(23	41	21)	(39	40	9)
SA	(32	7	48)	(24	15	48)	(15	5	71)	(39	10	38)
WA	(28	44	21)	(40	31	18)	(40	11	37)	(23	33	36)

Note: The diagonal is the outcome distribution for the control sessions.

Table 2.4 presents the outcome distribution for the control and ensemble treatments. Each cell contains the proportion of SS, CC and ALT for the row game, when it is paired with the column game. The dominant diagonal contains the outcome distribution for the control sessions. We can then test Hypotheses 1 to 4 using information in this table. For example, Hypothesis 1 predicts that, row-by-row (excluding the SI row), the proportion of SS is largest when a game is paired with SI. In what follows, we present the results testing each hypothesis and tabulate the statistical support in a separate table for each hypothesis.

Result 1 (Effects of SI)**.** *The proportion of SS in PD is weakly higher when it is paired with SI (46%) than when it is played in isolation (18%) or when it is paired with SA (23%). Similarly, the proportion of SS in SA is weakly higher when it is paired with SI (32%) than when it is played in isolation (15%). The latter becomes significant after Round 100 (32% versus 6%).*

Support. *In Table 2.4, pairwise comparisons between the SS columns represent the effects of SI on the likelihood of selfishness in other games. Permutation tests of the general null hypothesis of game independence against Hypothesis 1 are reported in Table 2.5 below. After Round 100, the difference between (**SA**, SI) and **SA** control becomes significant ($p = 0.031$, one-sided permutation test).*

By Result 1, we reject the null in favor of Hypothesis 1 at the 10% level for two sets of comparisons. When paired with SI, the proportion of SS is larger for PD (resp. SA) than when it is played in isolation or when it is paired with SA (resp. PD). After 100 rounds, one set of comparison is significant at the 5% level. We also note that SI does not have such an effect on WA.

We now examine the effects of PD on other games. Hypothesis 2 predicts that, in Table 2.4, row-by-row (excluding the PD row), the proportion of CC is largest when a game is paired with PD. While the outcome distribution of SI is not affected by

Table 2.5 Effects of SI on Other Games: Permutation Tests

SS in			
	PD control	(**PD**,SA)	(**PD**,WA)
(**PD**,SI)	0.070	0.109	0.371
	(**SA**,PD)	**SA** control	(**SA**,WA)
(**SA**,SI)	0.290	0.065	0.680
	(**WA**,PD)	(**WA**,SA)	**WA** control
(**WA**,SI)	0.846	0.813	0.364

any game it is paired with, the effects of PD on cooperation is present in SA, and to a lesser extent, in WA.

Result 2 (Effects of PD)**.** *Comparing (**SA**, PD) and **SA** control sessions, the proportion of CC is significantly higher when SA is paired with PD (15% versus 5%). Comparing (**SA**, PD) and (**SA**, SI), we observe CC more often in SA when paired with PD (15% versus 7%). The difference is more pronounced and significant after Round 100 (14% versus 1%).*
*Similarly, comparing (**WA**, PD) and (**WA**, SA), the proportion of CC is significantly higher when WA is paired with PD (31% versus 11%).*

Support. *In Table 2.4, pairwise comparisons between the CC columns when a game is paired with PD versus when it is paired with another game represent the effects of PD on the likelihood of CC in other games. One-sided permutation tests of the general null hypothesis of game independence against Hypothesis 2 are reported in Table 2.6. One-sided permutation test comparing the proportion of CC in in SA in ensembles (**SA**, PD) and (**SA**, SI) yields* $p = 0.023$ *after 100 rounds.*

Table 2.6 Effects of PD on Other Games: Permutation Tests

CC in			
	SI control	(**SI**,SA)	(**SI**,WA)
(**SI**,PD)	1.000	1.000	1.000
	(**SA**,SI)	**SA** control	(**SA**,WA)
(**SA**,PD)	0.119	0.049	0.299
	(**WA**,SI)	(**WA**,SA)	**WA** control
(**WA**,PD)	0.897	0.024	0.566

By Result 2, we reject the general null hypothesis in favor of Hypothesis 2 for the effects of PD on SA and WA for some ensemble comparisons. However, PD has no effect on SI.

Result 3 (Effects of SA)**.** *When PD (resp. WA) is paired with SA, the proportion of ALT is larger than when it is played in isolation or when it is paired with any other game. However, none of the pairwise comparisons is significant.*

Support. *In Table 2.4, pairwise comparisons between the ALT columns when a game is paired with SA versus when it is paired with another game represent the effects of SA on the likelihood of ALT in other games. One-sided permutation tests of the general null hypothesis of game independence against Hypothesis 3 are reported in Table 2.7.*

Table 2.7 Effects of SA on Other Games: Permutation Tests

ALT in	**SI** control	(**SI**,PD)	(**SI**,WA)
(**SI**,SA)	1.000	1.000	0.647
	(**PD**,SI)	**PD** control	(**PD**,WA)
(**PD**,SA)	0.110	0.355	0.183
	(**WA**,SI)	(**WA**,PD)	**WA** control
(**WA**,SA)	0.142	0.119	0.485

By Result 3, we fail to reject the general null hypothesis in favor of Hypothesis 3. Furthermore, as the proportion of ALT is the largest when a game is paired with SA than when it is played in isolation or when it is paired with any other game, Hypothesis 4 is not supported in our data.[9]

Although WA does not have any effects on the likelihood of alternation on other games as hypothesized by the behavioral spillovers hypothesis, it does affect selfish play in PD and SA, consistent with the effects of cognitive load. When PD is paired with WA, subjects play SS nearly as often as when PD was paired with SI. For SA, the effect is even stronger: more subject pairs play SS when SA is paired with WA than when it is paired with SI.

Result 4 (Effects of WA)**.** *The proportion of SS is significantly higher in (**PD**, WA) (39%) than in the PD control sessions (18%) or in (**PD**, SA) (23%). Likewise, the proportion of SS is significantly higher in (**SA**, WA) (39%) than in the SA control sessions (15%).*

Support. *Pairwise comparisons between the SS columns when a game is paired with WA versus when it is paired with another game in Table 2.4 indicate the effects of*

[9]We omit the table reporting p-values of the corresponding pairwise one-sided permutation tests, none of which is significant at the 10% level.

WA on the likelihood of selfishness in other games. Permutation tests of the general null hypothesis of game independence against the hypothesis of higher SS in a game when it is paired with WA are reported in Table 2.8.

Table 2.8 Effects of WA on Other Games: Permutation Tests

SS in	**SI** control	(**SI**,PD)	(**SI**,SA)
(**SI**,WA)	0.997	0.695	0.630
	(**PD**,SI)	**PD** control	(**PD**,SA)
(**PD**,WA)	0.630	0.030	0.036
	(**SA**,SI)	(**SA**,PD)	**SA** control
(**SA**,WA)	0.321	0.174	0.043

We conjecture that the increase of SS in PD (resp. SA) when it is paired with WA might be due to increased cognitive load, as WA played in isolation has the highest entropy among the four games and subjects are significantly less likely to reach Pareto efficient outcomes in WA than in the other games.[10] Thus, when either PD or SA is played with WA, subjects might resort to the dominant strategy in the stage game when they play PD and SA.

We now turn to Hypothesis 5 where we aggregate the previous analysis by comparing Pareto efficient outcomes in the control and ensembles. In the control sessions, the Pareto-efficient outcomes (SS in SI, CC in PD, ALT in SA and WA) emerge as the mode among all the three outcomes in SI, PD and SA played in isolation (diagonal in Table 2.4.[11]

Table 2.9 reports the proportion of Pareto efficient outcomes in the control and ensembles, as well as p-values from one-sided permutation tests. The general null hypothesis is game independence, i.e., the proportion of Pareto efficient outcomes of a particular game is the same between the control and ensemble treatments, while the alternative hypothesis is that the the proportion of Pareto efficient outcomes is higher when a game is played alone. We summarize the results below.

Result 5 (Pareto Efficient Outcomes in Ensembles)**.** *In every game, the proportion of the Pareto efficient outcome decreases when a game is part of an ensemble than when the same game is played in isolation. This decrease is significant for ALT in SA.*

[10]There is no dominant outcome in WA control sessions, but the difference in entropy is not statistically significant.

[11]ALT is also the mode in WA control sessions although the proportion of ALT in WA is not statistically higher than other two outcomes.

Table 2.9 Pareto Efficient Outcomes in Ensemble vs. Control

	SS in SI	CC in PD	ALT in SA	ALT in WA
% in Control	0.999	0.557	0.711	0.361
% in Ensemble	0.991	0.410	0.449	0.244
H_1: control > ensemble one-sided p-values	0.002	0.171	0.031	0.213
# of observations (control, ensemble)	(6, 23)	(6, 18)	(6, 18)	(6, 23)

Note: In the permutation tests, we treat each pair (group of four) in the control (ensemble) sessions as one observation.

Support. *In Table 2.9, we reject the null of game independence in favor of H_1 for SS in SI ($p = 0.002$), and ALT in SA ($p = 0.031$).*

Result 5 indicates a general decrease of Pareto efficient outcomes when a game is played in an ensemble compared to the same game played in isolation. In particular, in Strong Alternation, subjects alternate significantly less ($p = 0.031$) when it is in an ensemble than when it is played in isolation. The Self Interest game, however, is not affected by the presence of other games. It is the easiest game to play. Whether in control or ensemble treatments, subjects quickly converge to SS, with over 99% of selfishness across the rounds. Thus, by Result 5, we reject the null of game independence in favor of Hypothesis 5 in SA.

Lastly, we examine Hypothesis 6 which postulates that the proportion of Pareto efficient outcomes in a game is higher when it is paired with SI than when it is paired WA. While this hypothesis correctly forecasts the direction of change, none of the pairwise comparisons is significant at the 10% level.

We next compare the efficiency generated in each game. Following convention, we capture efficiency by the percentage of potential payoff above the minimum payoff that the players receive. Our normalized efficiency measure is defined as follows.

$$\text{Efficiency} = \frac{\text{Actual joint payoffs} - \text{Minimum joint payoffs}}{\text{Maximum joint payoffs} - \text{Minimum joint payoffs}} \tag{2.1}$$

Table 2.10 presents the average efficiency in the control and the ensemble sessions. In each row, we present the efficiency of the row game when it is played alone (diagonal, boldfaced) and when it is paired with another game. Consistent with outcome level results, the efficiency in SI, PD and SA control sessions is higher than that of the

Table 2.10 Efficiency in the Control and Ensembles

	SI	PD	SA	WA
SI paired with	**99.86**	99.32	99.23	98.99
PD paired with	51.5	**73.35**	63.59	53.97
SA paired with	64.24	68.1	**82.69**	55.66
WA paired with	63.06	53.97	57.89	**70.86**

corresponding games in ensemble sessions. In particular, the following comparisons, **SI** > (**SI**, SA), **SI** > (**SI**, WA), and **SA** > (**SA**, WA), are significant at the 5% level, while **SA** > (**SA**, SI), **SA** > (**SA**, PD), and **PD** > (**PD**, WA) are significant at the 10% level (see Table 2.11 for p-values of one-sided permutation tests).[12]

Table 2.11 Efficiency Comparison between Control and Ensemble: Permutation Tests

Control	(**SI**,PD)	(**SI**,SA)	(**SI**,WA)
SI	0.130	0.002	0.004
	(**PD**,SI)	(**PD**,SA)	(**PD**,WA)
PD	0.120	0.225	0.066
	(**SA**,SI)	(**SA**,PD)	(**SA**,WA)
SA	0.052	0.099	0.034
	(**WA**,SI)	(**WA**,PD)	(**WA**,SA)
WA	0.290	0.134	0.217

Table 2.12 reports the p-values of one-sided permutation tests of pairwise efficiency comparisons of a common game when it is paired with different games. None of the pairwise comparisons is significant at the 10% level.

Table 2.12 Efficiency Comparison between Ensembles: Permutation Test Results

Game in	Game Ensemble			P-Values of Permutation Tests		
	(1)	(2)	(3)	(1) vs. (2)	(1) vs. (3)	(2) vs. (3)
PD in	(**PD**,SA)	(**PD**,SI)	(**PD**,WA)	0.240	0.170	0.428
SA in	(**SA**,PD)	(**SA**,SI)	(**SA**,WA)	0.387	0.185	0.273
WA in	(**WA**,PD)	(**WA**,SI)	(**WA**,SA)	0.191	0.371	0.335
SI in	(**SI**,PD)	(**SI**,SA)	(**SI**,WA)	0.411	0.341	0.417

In general, the experimental results agree with our alternative hypotheses: game

[12]Subjects also played WA more efficiently in the control than in ensembles, but the difference is not statistically significant.

independence is not supported, but instead subjects are influenced by contextual effects of behavioral spillovers and cognitive load. While analyses reported in this subsection are based on non-parametric tests, we recognize that income effects might influence behavior. Thus, we use probit regressions with standard errors clustered at the pair (respectively group) level for Results 1 to 5, controlling for income and learning effects in each specification. Following Ham, Kagel and Lehrer (2005), we use a participant's cash balance prior to period t to control for income effects. We find that, while cash balance is statistically significant in most specifications, its marginal effect on behavior is never greater than 1%.[13] Further, Results 1 to 5 continue to hold in our regression analysis.

2.6.2 Ensemble Effects at the Strategy Level

In this subsection, we analyze repeated game strategies in each game, and the ensemble effects at the strategy level. Following Rubinstein (1986) and Abreu and Rubinstein (1988), we use automaton (or the Moore machine) to represent repeated game strategies. In Figure 2.5, we present 29 repeated game strategies, including the exhaustive set of 26 1- and 2-state automata, and three 3-, 4- and 5-state automata which have been widely discussed in previous literature (Engle-Warnick and Slonim, 2006). For each subject i, we calculate the *fitting proportion* for each automaton, M_j, for the entire sequence of observed actions $\{a_i^t\}_{t=1}^T$, defined as $F_i(M_j) = \sum_{t=1}^T I(a_i^t, M_j^t)/T$, where the indicator function $I(a_i^t, M_j^t) = 1$ if $a_i^t = M_j^t$. At the treatment level, we define the average fitting proportion as $\bar{F}(M_j) = \sum_{i=1}^n F_i(M_j)/n$, where n is the number of players.

To highlight the main results, in subsequent discussions, we only compare the performance of a subset of 1- and 2-state automata.[14] Table 2.13 presents 9 out of 26 1- and 2-state automata, including all six strategies with at least 50% fitting proportion across all games and ensembles, and three which do not survive the 50% fitting threshold but have been extensively discussed in the literature (Bednar and Page, 2007; Engle-Warnick and Slonim, 2006; Hanaki, Sethi, Erev and Peterhansl, 2005). These 9 strategies can be divided into three categories: (1) cooperative strategies, including Always Cooperate (AC), Forgive Once (F1), Suspicious Forgive Once (sF1); (2) reciprocal strategies, including Tit-for-Tat (TFT), and Suspicious Tit-for-Tat (sTFT); and (3) selfish strategies, including Always Selfish (AS), Grim Trigger

[13] Regression tables are available from the authors upon requests.

[14] The complete analysis of all 29 automata are available from the authors upon request.

Table 2.13 Description of Nine Strategies

Strategy Category	Name of Strategy	Strategy Number	Initial Action	Continued Play
Cooperative	Always Cooperate (AC)	M1	C	Always play cooperate
	Forgive once (F1)	M8	C	Go to S if other plays S and go to C when the last period is S
	Susipicious Forgive once(sF1)	M9	S	Go to S if other plays S and go to C when the last period is S
Reciprocal	Tit for Tat (TFT)	M4	C	Copy other's previous action
	Susicious Tift For Tat (sTFT)	M5	S	Copy other's previous action
Selfish	Always Selfish (AS)	M2	S	Always play selfish
	Grim Trigger (GT)	M3	C	C until other plays S, then S forever
	Switch After C(SAC)	M6	C	After C, play S until other plays C
	Susipicious Switch After C(sSAC)	M7	S	After C, play S until other plays C

(GT),[15] Switch after Cooperate (SAC), suspicious Switch after Cooperate (sSAC). Of these nine strategies, AC, AS and GT do not survive the 50% fitting threshold.

The strategy level analysis in the control sessions reveals few surprises. In the PD game, reciprocal strategies, TFT and sTFT, have significantly higher fitting proportion, $\bar{F}(TFT) = \bar{F}(sTFT) = 0.84$, than every other strategy.[16] This result is consistent with the literature where TFT is widely and successfully used in the repeated PD simulations (Axelrod and Hamilton, 1981; Bednar and Page, 2007; Hanaki et al., 2005). These two strategies have even better fitting proportion, $\bar{F}(TFT) = \bar{F}(sTFT) = 0.90$, in the Weak Alternation game.[17] In the Strong Alternation game, TFT, sTFT as well as SAC and sSAC are all best fitting strategies, each with a fitting proportion of $\bar{F} = 0.85$, and each significantly outperforming every other strategy outside this set.[18] This result is consistent with Hanaki et al (2005) where in a simulation of coordination games similar to Strong Alternation, SAC is the best performing strategy. Finally, in the Self Interest game, due to the lack of variation in actions, player behavior can be explained by any different strategy that produces all selfish behavior. Among these strategies are TFT, sTFT, SAC, sSAC, AS, and GT, each has a fitting proportion of $F = 1.00$.

The strategy level analysis for the ensemble sessions also aligns with expectations. Comparing strategies used in PD when it is paired with SI and those used in the PD control sessions, we find that the fitting proportion of selfish strategies increases from control to ensembles. For example, the fitting proportion of strategies that choose to be selfish after cooperating (SAC and sSAC) increases from 0.59 in the control to 0.71 in the ensemble ($p = 0.08$, one-sided permutation test). It is still the case though that TFT and sTFT retain the greatest fitting proportion in PD ($\bar{F} = 0.91$) even when it is paired with SI. Likewise, when PD is paired with WA, the fitting proportion of selfish strategies again increases, with $\bar{F}(AS) = 0.31$ in the control sessions and 0.49 in the (PD, WA) ensemble ($p = 0.09$), $\bar{F}(SAC) = 0.59$ in the control and 0.66 in the

[15] We classify GT as a selfish strategy because, when faced with any match that plays S with some positive probability, GT is identical to AS from then on. However, we recognize that it is subgame perfect punishment. Further, GT-GT is a subgame perfect equilibrium which achieves perfect cooperation. Thus, this classification is not perfect.

[16] $\bar{F}(TFT) > \bar{F}(M_j)$: $p < 0.05$ for any $M_j \neq \{AC, sTFT\}$, and $p < 0.10$ for $M_j = AC$. Similarly, $\bar{F}(sTFT) > \bar{F}(M_j)$: $p < 0.05$ for any $M_j \neq \{sF1, AC, TFT\}$, and $p < 0.10$ for $M_j \in \{sF1, AC\}$, one-sided Wilcoxon signed rank tests.

[17] $\bar{F}(TFT) > \bar{F}(M_j)$: $p < 0.05$ for any $M_j \neq \{SAC, sSAC, sTFT\}$, and $p < 0.10$ for $M_j \in \{SAC, sSAC\}$. Similarly, $\bar{F}(sTFT) > \bar{F}(M_j)$: $p < 0.05$ for any $M_j \neq \{F1, SAC, sSAC, TFT\}$, and $p < 0.10$ for $M_j \in \{F1, SAC, sSAC\}$, one-sided Wilcoxon signed rank tests.

[18] SAC or sSAC has significantly greater fitting proportion than any other strategy outside the set ($p < 0.05$), while TFT or sTFT has significantly greater fitting proportion than any other strategy outside the set ($p < 0.05$) except for F1 ($p < 0.10$), one-sided Wilcoxon signed rank tests.

ensemble ($p = 0.06$), and $\bar{F}(sSAC) = 0.59$ in the control and 0.66 in the ensemble ($p = 0.07$).

Comparing strategies used in SA when it is paired with SI and those used in the SA control sessions, we find that the fitting proportion of selfish strategies significantly increases from the control to the ensemble, while that of cooperative strategies weakly decreases. Specifically, $\bar{F}(GT) = 0.55$ in the control and 0.65 in the ensemble ($p = 0.02$), while $\bar{F}(AC) = 0.45$ in the control and 0.37 in the ensemble ($p = 0.10$). We obtain similar results comparing the (SA, WA) ensemble and SA control sessions.

Last, when we compare repeated game strategies between ensembles, several of the results from the outcome level analysis survive, albeit in a slightly different form. Specifically, we find behavioral spillover effects of PD and SA, as well as cognitive load effects.

Consistent with Result 2 where the proportion of CC is higher when a game is paired with PD, we find that, in the (**WA**, PD) vs.(**WA**, SA) ensembles, when WA is paired with PD, the fitting proportion of cooperative strategies, such as always cooperate (AC), is weakly higher than when it is paired with SA (45 vs. 36%, $p = 0.10$), while that of selfish strategies, such as AS (55 vs. 64%), GT (55 vs. 64%), SAC (72 vs. 81%) and sSAC (72 vs. 81%) is weakly higher when WA is paired with SA ($p < 0.1$ for each comparison, one-sided permutation tests). Likewise, when SA is paired with PD, the fitting proportion of cooperative strategies including AC, F1 and sF1 is higher than when it is paired with SI or WA, however, none of these pairwise comparison is significant at 10% level.

Consistent with Result 3 where the proportion of ALT is higher when a game is paired with SA, we find that, in the (**WA**, SA) vs.(**WA**, SI) ensembles, when WA is paired with SA, the fitting proportion of SAC (sSAC) in WA is significantly higher than when it is paired with SI (SAC: 81 vs. 71%, $p = 0.033$; sSAC: 81 vs. 71%, $p = 0.035$, one-sided permutation tests).

While Hypothesis 6 on cognitive load is only supported directionally at the outcome level, we find statistically significant effects at the strategy level. Comparing PD strategies in the (**PD**, SI) vs. (**PD**, WA) ensembles, when PD is paired with SI, the fitting proportion of TFT and sTFT is significantly higher than when it is paired with WA (TFT: 91 vs. 83%, $p = 0.039$; sTFT: 91 vs. 83%, $p = 0.042$, one-sided permutation tests).

In sum, analysis of behavioral spillovers at both the outcome and strategy levels yields largely consistent results, i.e., when games are paired in ensembles, play differs from isolated controls, and in predictable ways. In some cases, the two levels of analysis provide different lenses on the same phenomenon. For example, both analyses demonstrate that a game paired with the PD game exhibits more cooperative outcomes and strategies than observed in the control or when it is paired with another game. In other cases, the strategy level analysis highlights a different feature of the results. For example, the strategy analysis allows us to see how often pairs are able to coordinate on alternating strategies.

2.7 Discussion

In this paper, we present an experimental study to test for ensemble effects in game playing behavior. We test for these effects looking both at outcomes and strategies. Our study reveals evidence of behavioral spillovers that depend in predictable ways on features of the games in the ensemble. In particular, if subjects play one game in an ensemble that encourages selfishness or cooperation, then they are more likely to exhibit that behavior in the other game in their ensemble, even though they play the other game with a different player. We also see evidence of cognitive load.

To derive our hypotheses about cognitive load and behavioral spillovers, we introduce a new measure of behavioral variance, *entropy*, which is consistent with response time, an alternative measure of cognitive load used in psychology. We compute the entropy of a game played in isolation and make out-of-sample predictions of its effect in game ensembles. We posit that in ensembles that include games that produce high entropy outcomes, cognitive load will be most pronounced. Consistent with our expectations, cognitive load has the greatest effect when ensembles include Weak Alternation, our highest-entropy game. In contrast, we hypothesize that low entropy games would produce stronger behavioral spillovers and are less influenced by other games. Both predictions are supported by our data.

Our findings provide an initial demonstration of how a person's behavior in a given game depends on the ensemble of strategic situations that the person faces. In doing so, they call into question the focus on isolated games in most theoretical and empirical analyses. This critique extends to mechanism design, which assumes that incentives can be considered independent of the broader behavioral context.

To summarize, our multiple games experiments demonstrate that significant ensemble effects emerge in the laboratory setting. Outcome level and strategy level analysis show consistent ensemble effects. Subjects with incentives to behave cooperatively (resp. selfishly) in one game, tend to behave similarly in another game even if that behavior is not efficient. Our results provide a behavioral explanation for observations in other ensemble experiments using more complex games, such as the public goods game and lottery contest (Savikhin and Sheremeta, 2012), the median- and minimum-effort games (Cason et al., 2012), and the threshold public goods and double auctions (Cason and Gangadharan, 2010). This emerging literature indicates that contexts affect behavior.

2.8 Appendix: Experimental Instructions

We present the instructions for the (PD, WA) ensemble. Instructions for other ensemble treatments are identical except for the specific game forms. Instructions for the control sessions are identical to the ensemble instructions except that two games and two other participants are replaced with one game and one other participant everywhere. Hence we omit them here.

Name: PCLAB: Total Payoff:

Introduction

- You are about to participate in a decision process in which you will play two games with two other participants. Each game will be played with a different participant and will be played for many rounds. This is part of a study intended to provide insight into certain features of decision processes. If you follow the instructions carefully and make good decisions, you may earn a considerable amount of money. You will be paid in cash at the end of the experiment.
- During the experiment, we ask that you please do not talk to each other. If you have a question, please raise your hand and an experimenter will assist you.

Procedure

- Matching: At the beginning of the experiment, you will be matched randomly with two other participants, both of whom will be your matches for the entire experiment. You will be matched with these same two people in all rounds. You will play a different game with each of these people.
- Roles: Throughout the game, you will be designated as the “row” player and your matches will be the “column” players. You will be a “row” player in all rounds, and your matches will be “column” players in all rounds
- Actions: In each round, you and your two matches will simultaneously and independently make decisions in two different games. One is the left game and the other is the right game. You will play the left game with one of your matches (Left Game Match) and play the right game with the other match (Right Game Match). In each game, the row player (you) will click either the Top (A) or the Bottom (B) button. The column player (your Left or Right Game Match) will choose either the Left (A) or Right (B) button. These choices determine which part of the matrix is relevant (Top Left, Top Right, Bottom Left, Bottom Right).
- Interdependence: A player’s earnings depend on the decision made by the player and on the decision made by his or her two matches as shown in the matrix below. In each cell, the row player’s payoff is shown in red and the column player’s payoff is shown in blue.

		Column Player			Column Player	
		Left	Right		Left	Right
Row	Top	7, 7	2,10	Top	7, 7	4,11
Player	Bottom	10,2	4,4	Bottom	11,4	5, 5

For example, if the row player (you) chooses Top (A) and the column player (your left game match) chooses Right (B) in the left game, then the row player (you) will get 2 points, while the column player (your left game match) will get 10 points in this game. Meanwhile, if the row player (you) chooses Bottom (B) and the column player (your right game match) chooses right (B) in the right game, then the row player (you) will get 5 points, and the column player (your right game match) will also get 5 points in this game. So as the row player in both games, you will get 7 points in this round totally.

- Rounds: You will first play the two games for 200 rounds. After round 200, whether the games will continue to the next round depends on the "throw of a die" that is determined by the computer's random number generator. At the end of each round after round 200, with 90% chance, the games will continue to the next round. With 10% chance, the games stop.
- Earnings: Your earnings are determined by the choices that you and your two matches make in every round. Your total earning is the sum of your earnings in all rounds.
  The exchange rate is $1 for 100 points.
  You can round up your total earning to the next dollar. For example, if you earn $15.23, you can round it up to $16.
- History: In each round, your and your two matches' decisions in all previous rounds will be displayed in a history window.

We encourage you to earn as much money as you can. Do you have any questions?

Figure 2.1 Control: Behavior in Self Interest

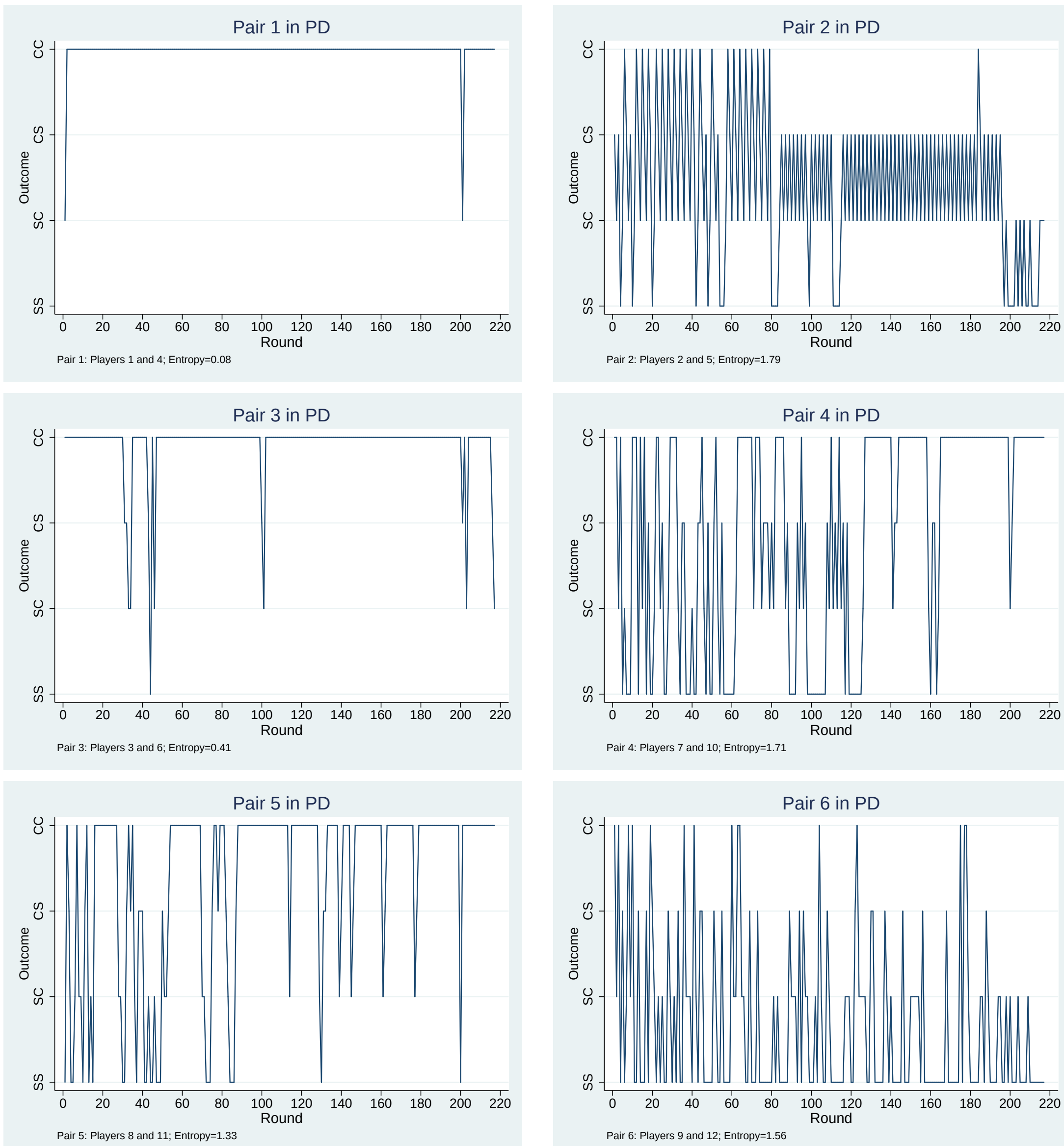


Figure 2.2 Control: Behavior in Prisoner's Dilemma

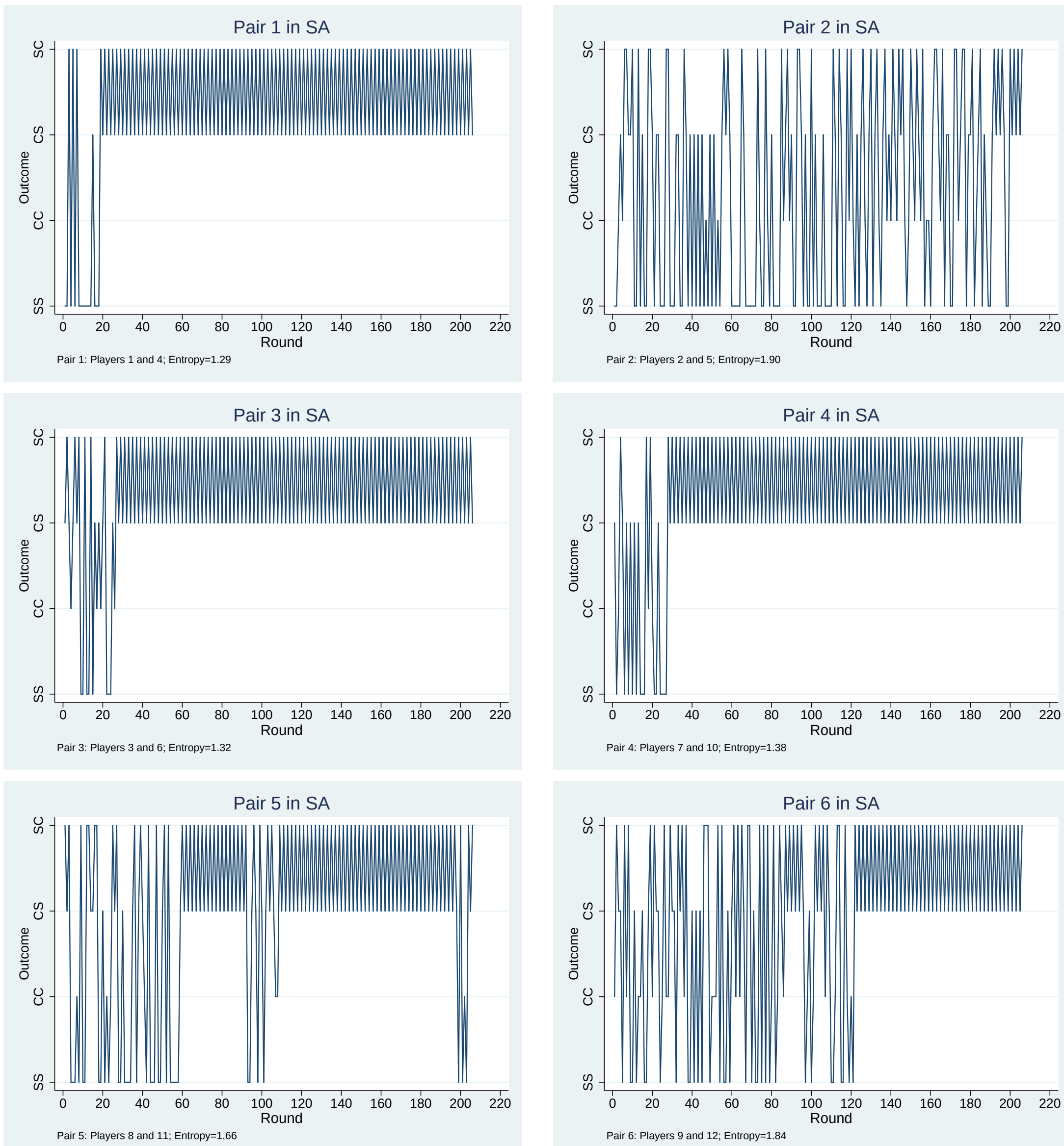
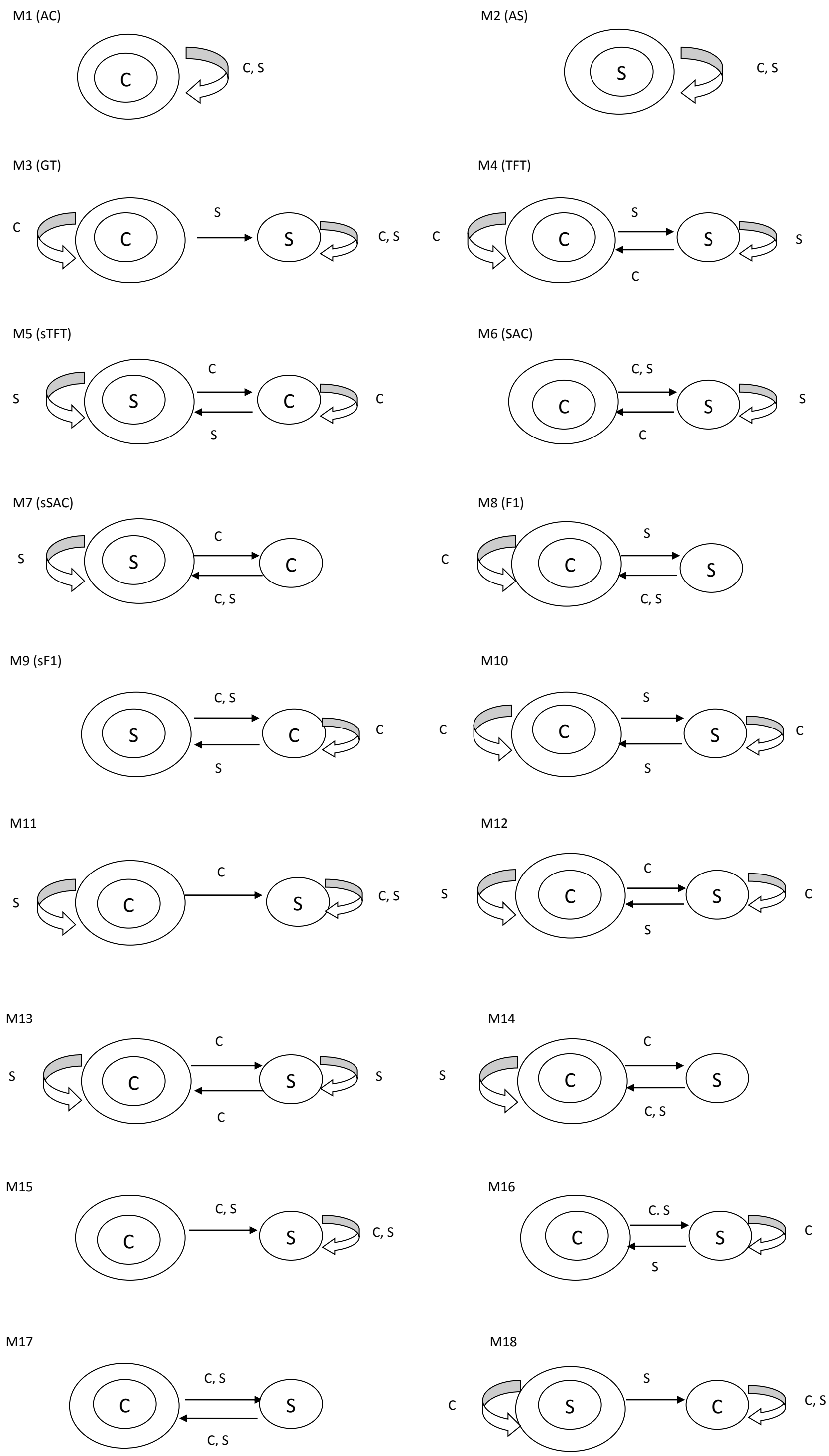


Figure 2.3 Control: Behavior in Strong Alternation

Figure 2.4 Control: Behavior in Weak Alternation



M1 (AC)
C
C, S
M2 (AS)
S
C, S
M3 (GT)
C
C
S
S
C, S
M4 (TFT)
C
C
S
C
S
S
M5 (sTFT)
S
S
C
S
C
C
M6 (SAC)
C
C, S
C
S
S
M7 (sSAC)
S
S
C
C, S
C
M8 (F1)
C
C
S
C, S
S
M9 (sF1)
S
C, S
S
C
C
M10
C
C
S
S
S
C
M11
S
C
C
S
C, S
M12
S
C
C
S
S
C
M13
S
C
C
C
S
S
M14
S
C
C
C, S
S
M15
C
C, S
S
C, S
M16
C
C. S
S
S
C
M17
C
C, S
C, S
S
M18
C
S
S
C
C, S

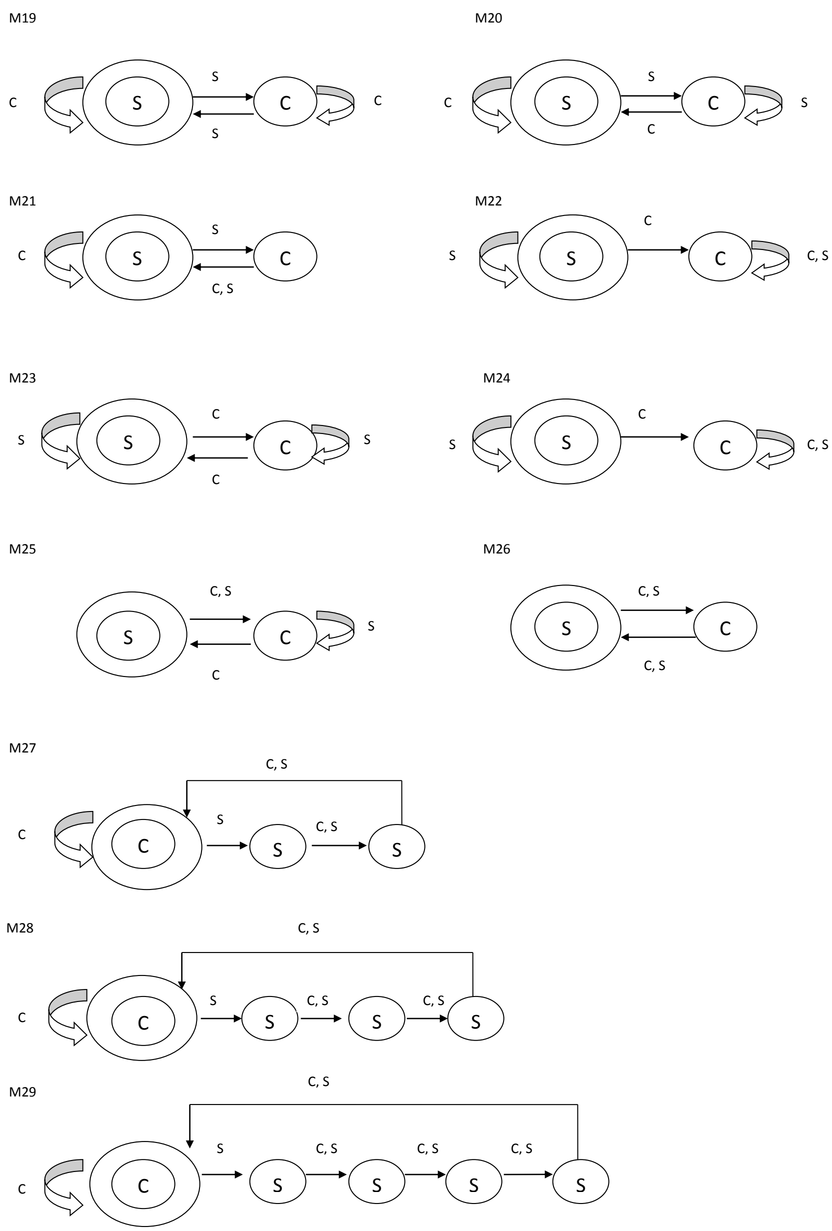


Figure 2.5 Automata Representation of Repeated Game Strategies

Chapter 3

Which Hat to Wear? Impact of Natural Identities on Coordination and Cooperation

3.1 Introduction

As the world becomes increasingly integrated and the workforce becomes more diverse, motivating individuals from diverse backgrounds to work together effectively is a major challenge facing organizations today. While increasing diversity in groups has been found to elicit positive outcomes such as enhancing thoughtful decision processes (Nemeth, 1986), expanding access to social networks and resources (Tushman, 1977), promoting innovation (Van Der Zee and Paulus, 2008), and facilitating problem solving (Hong and Page, 2001), increasing diversity also introduces group biases that may contribute to conflict among group members (Hargreaves Heap and Zizzo, 2009; Pelled, Eisenhardt and Xin, 1999). As a result, organizations wishing to obtain the benefits associated with diversity must also learn how to manage diversity in order to facilitate coordination, cooperation and positive interpersonal relationships among their members.

Research findings underscore the importance of effectively promoting coordination, cooperation and positive interpersonal relationships among members of an organization. Positive relationships have been associated with a host of important outcomes such as more effective sharing of resources and information, greater trust and better performance (Blatt and Camden, 2006; Gruenfeld, Mannix, Williams and Neale, 1996). Thus, integrating a diverse workforce, and motivating members who come from different backgrounds to work effectively towards a common goal is an important task facing many organizations.

However, despite this importance, organizations trying to promote better coordination and cooperation in diverse groups face several challenges in accomplishing this goal. First, work on minimal groups in psychology and near-minimal groups in economics finds that individuals are predisposed to favor the ingroup over the outgroup to enhance and maintain positive self-esteem (Tajfel and Turner, 1979). As a consequence, individuals perceive their ingroup members to be more similar to them than members of the outgroup (Allen and Wilder, 1975) and ascribe more positive traits to ingroup members (Brewer, 1979). Individuals are also more likely to help members of the ingroup over the outgroup (Crosby, Bromley and Saxe, 1980), to allocate more rewards to ingroup members (Wilder, 1986), and to show more charity, less envy, more positive reciprocity, less negative reciprocity, and more social welfare maximizing actions towards ingroup members (Chen and Li, 2009). In sum, research on minimal and near-minimal groups has collected a great deal of evidence showing that highlighting different social identities may fragment a group by introducing group biases that lead to counterproductive outcomes.

However, in the real world, people can be simultaneously identified along many dimensions of identity (Hewstone, 1996). Consider an African American male accountant who is a partner in his firm. He may be identified by his gender (male), his race (black), his role (partner), his occupation (accountant) or his organization (firm). Some of these identities may be shared by other members of the group, while other identities may not. Thus, highlighting these different identities may call forth different group orientations and their consequent behaviors within an organization. Furthermore, research finds that feelings of similarities to others within a group can be situationally altered by manipulating the salience of different social identities (Chatman, Polzer, Barsade and Neale, 1998). While highlighting uncommon identities may fragment a group, highlighting common identities might unify a group.

In practice, common identities have been used to create common goals and values. For example, Nike founder Phil Knight and many of his employees have tattoos of the Nike "swoosh" logo on their left calves as a sign of group membership (Camerer and Malmendier, 2007). To create a common identity, organizations have attempted various team-building exercises, such as simulated space missions where the crew works together to overcome malfunctions while navigating through space (Ball, 1999). While standard economic theory does not have an explanation for such phenomena, research on social identity shed lights on the effects of common identity on organization outcomes.

Social psychology work on intergroup relations finds that highlighting a common ingroup identity can reduce intergroup bias (Dovidio, Gaertner and Saguy, 1980;

Gaertner and Dovidio, 2000). For instance, college roommates from differing ethnic backgrounds who perceived more common identities were less likely to show decline in their friendship than roommates who did not (West, Pearson, Dovidio, Shelton and Trail, 2009). In another study, emphasizing a common ingroup identity increased satisfaction with coworkers in ethnically diverse workgroups (Cunningham, 2005).

Moreover, evidence in experimental economics finds that a common group identity increases cooperation in public goods games (Eckel and Grossman, 2005) and prisoner's dilemma games (Goette, Huffman and Meier, 2006), where the dominant strategy is to completely free ride or defect. Furthermore, it improves coordination in the battle of sexes game (Charness, Rigotti and Rustichini, 2007), the provision point mechanism (Croson, Marks and Snyder, 2008), and the minimum effort game (Chen and Chen, 2011). The latter two games have multiple Pareto ranked equilibria; a salient common identity leads to the selection of a more efficient equilibrium.

This study extends previous research on the effects of a common identity on economic behavior. In particular, we investigate the effects of highlighting a common vs. fragmenting identity on coordination and cooperation in a series of prisoner's dilemma games with varying incentives for cooperation. Using subjects from two large public universities with comparable academic standing (the University of Michigan and the University of California at Los Angeles), we prime participant school identity as their common identity, and ethnic identity as the fragmenting identity.

Our results show that Asian and Caucasian participants respond differently to priming. While priming ethnic identity significantly increase the ingroup favoritism and outgroup discrimination for both UM and UCLA Asians compared to the control, it has no significant effect among Caucasians. Secondly, priming a common (school) identity reduces group bias for UM Asians in the coordination game, resulting in a significant increase of both ingroup and outgroup cooperation. However, in games with a unique inefficient Nash equilibrium, the effects of priming a common identity are more complex. While priming alleviates the negative effects of the competitiveness stereotype on cooperation among UCLA Asians, it enhances such negative effects among UM Asians. The differential response to priming from first-generation (UM Asian students) and second- or third-generation (UCLA Asian students) ethnic minorities has policy implications for socializing new immigrants which we will elaborate in Section 5.7.

This paper contributes to the literature on the effects of group identities on cooperation and coordination in several ways. First, rather than inducing group identity in the laboratory, we study two naturally existing social identities - ethnic identity and organization identity. Thus, compared to studies using induced group identity, our results can be more easily applied to relevant real-life work environments. Second, this study goes beyond documenting the intergroup bias in individual choices. We use the identity priming technique from social psychology to manipulate the salience of the respective identities to investigate the extent to which evoking different dimensions of these identities impacts intergroup bias. Third, this study is among the first in economics to empirically evaluate the effectiveness of using a common identity as a design tool to increase cooperation among an ethnically diverse group of participants. Lastly, compared to social psychology studies of natural identities, we demonstrate that identity priming interacts with the strategic properties of games. The same priming technique can have different effects in different games.

The rest of the paper is organized as follows. Section 5.4 presents the experimental design. In Section 4.7, we present our analysis and results. Section 5.7 discusses the results and concludes.

3.2 Experimental Design

Our experimental design simulates a work environment in an organization in which employees have multi-dimensional social identities and engage in strategic interactions with one another involving potential tradeoffs between self interest and group interest. Although our participants share a common organization identity, they come from diverse ethnic backgrounds. The incentivized tasks in the experiment involve choices to cooperate or coordinate with another employee in the organization. Thus, the experiment design captures three important factors that may influence individual choices at a workplace: self interest, group interest, and intergroup relations. We use the priming method from social psychology to make one of the participants' natural identities salient before they participate in a sequence of one-shot prisoner's dilemma games.

In this study, we are interested in several questions. First, do people exhibit ingroup favoritism and outgroup discrimination, even in the absence of priming, when the other player's ethnic identity is known? Second, does group behavior intensify when we prime a fragmenting (ethnic) identity? Lastly, can we alleviate ingroup favoritism and outgroup discrimination by priming a common organization identity? In what follows, we describe the priming method, introduce the games and present the experimental procedure.

3.2.1 Identity Priming

Priming is an experimental technique in psychology that introduces certain stimuli ("primes") to activate individuals' social knowledge structures (Bargh, 2006). The types of primes include text (e.g., a questionnaire, an article, or a word scrambling game), image, or audio.

Priming social identities can impact people's behavior and attitudes outside of their awareness and control (see Bargh and Chartrand, 1999), as demonstrated by social psychologists in a large body of work on identity priming. In these laboratory studies, psychologists have found that making social identities salient often induces study participants to adopt behaviors that are consistent with the stereotypes associated with the identity. These effects occur even when participants are not aware that they are being primed. In one study, college students primed with stereotypes of the elderly walk more slowly as they exit the study than those who are not primed with stereotypes of the elderly (Bargh, Chen and Burrows, 1996). In another study, Steele and Aronson (1995) find that African American students who are stereotyped to be poor students underperform on academic tests when asked to indicate their race prior to taking the test. These effects have also been documented in other groups such as Hispanic Americans (Aronson, Quinn and Spencer, 1998), individuals from lower socio-economic status (Croizet and Claire, 1998) and women in math (Spencer, Steele and Quinn, 1999).

On the other hand, while activating negative stereotypes can hurt performance, activating positive stereotypes can boost performance. In one experiment, Shih, Pittinsky and Ambady (1999) examined the performance of Asian women on a mathematics test. Women are stereotyped to have inferior quantitative skills (Benbow, 1995; Hedges and Nowell, 1995) while Asians are stereotyped to have superior quantitative skills (Steen, 1987). Shih et al. (1999) find that Asian American women perform better on a mathematics test when their ethnic identity is primed, but worse when their gender identity is primed, compared to a control group with neither identity primed. In contrast, Asian Americans taking a verbal test showed the reverse pattern of performance. In this case, women are stereotyped to be verbally talented while Asians are not. Asian American women perform higher on the verbal test when their gender is salient, and worse when their ethnicity is made salient (Shih, Pittinsky and Trahan, 2006). These priming techniques have also been applied to study risk and time preferences in economics (Benjamin, Choi and Strickland, 2010).

Identity priming can also activate intergroup bias. Simply exposing individuals to words indicating ingroup or outgroup identity can elicit differential judgements from people. Perdue, Dovidio, Gurtman and Tyler (1990) find that subliminally exposing individuals to words associated with the ingroup and the outgroup (i.e. "us", "them") affects how quickly study participants judge positive and negative words. Participants are more quick to judge positive to be positive if exposed to ingroup words such as "us," and more quick to judge negative words to be negative if exposed to outgroup words such as "them". In the present study, we use identity priming methods to examine if individuals automatically exhibit intergroup bias in prisoner's dilemma games.

We choose two ethnic groups, Caucasians and Asians, which can be differentiated by their last names. For Asian participants, we focus on those with Chinese last names in order to avoid potential complex intergroup preferences among different Asian groups, e.g., Chinese and Japanese.

We adopt the priming technique from Shih et al. (1999), and subtly activate a social category outside of participants' awareness in the identity treatments. The stimuli are introduced through a pre-experiment questionnaire. In the ethnic identity treatment, the questions pertain to an individual's ethnic background, family history ("How many generations has your family lived in America?" and "From which countries did you family originate?"), and cultural heritage ("What languages do you speak?"). In the school identity treatment, subjects are asked about which school they attend. They are then asked to reflect on their choices of schools when applying for college ("Did you consider any other school? If yes, what other schools?", "Why did you decide to choose your specific school?"). Since the subjects in each experimental session study at the same university (UM or UCLA), these questions pertain to an individual's common identity of being part of her university. Because the two universities share comparable academic standings, we minimize the possibility that the impact of the common identity priming may be influenced by participants' perception on the standing of their universities.[1] In the control sessions, the questions are designed to be identity neutral, i.e., related to neither the ethnic nor the school identities. Subjects are asked about their activities in leisure time, for example, "How often do you watch television?" "How often do you eat out?" and "How often do you attend movies?" The identity neutral questionnaire is designed to preserve the direct comparability with the two identity treatments. These procedures are adopted from

[1]Li, de Oliveira and Eckel (2010) design a controlled field experiment in two neighborhoods in Dallas, TX, to study the impact of having a common identity on individual contributions to local public goods. They find that the same common identity priming leads to opposite outcomes. While it *decreases* the likelihood of giving in the poor neighborhood, it *increases* the likelihood of giving in the mid-income neighborhood.

those used in past psychology experiments and the questionnaires are modified versions of those used in Shih et al. (1999). The primes are designed to make salient the appropriate social identity and activate the constructs associated with the identity. A social identity is attached to a whole host of associated traits, stereotypes, social expectations, and schemas (Deaux, 1996). The questionnaires are included in Appendix A.

3.2.2 The Games

To investigate intergroup and intragroup coordination and cooperation under conditions when a fragmenting or a common identity is made salient, we choose variants of the prisoner's dilemma games. This class of games is among the simplest of those which capture the tension between individual and group interests. It has also been used in the social identity literature in psychology to investigate the causes of group bias (Simpson, 2006; Yamagishi and Kiyonari, 2000).

Figure 3.1 presents the extensive forms of the five sequential prisoner's dilemma games in our experiment. In each game, player 1 has two strategies, cooperate (C) or defect (D), whereas player 2 has four strategies:

- Always cooperate (CC): cooperate if player 1 cooperates, and cooperate if player 1 defects.
- Always defect (DD): defect if player 1 cooperates, and defect if player 1 defects.
- Reciprocal (CD): cooperate if player 1 cooperates, and defect if player 1 defects.
- Opposite (DC): defect if player 1 cooperates, and cooperate if player 1 defects.

Note that, while we use C and D throughout the paper for the ease of exposition, the subjects are given neutral terminologies. Player 1 (2), called player A (B) in the instructions, has actions A1 (B1) and A2 (B2), corresponding to C and D, respectively.

In one-shot scenarios, a sizeable literature on social preferences uncovers a non-negligible number of conditional cooperators in social dilemma types of games (Fehr and Gaechter, 2000; Healy, 2007). Healy (2007) models the sequential prisoner's dilemma game as a game of incomplete information about player 2's types. Specifically, let p be player 1's belief that 2 is a conditional cooperator. Assuming risk neutrality, player 1 will choose to cooperate if the expected value from cooperation is at least as great as the expected value from defection, i.e.,

$$p\pi_1(C,C) + (1-p)\pi_1(C,D) \geq \pi_1(D,D).$$

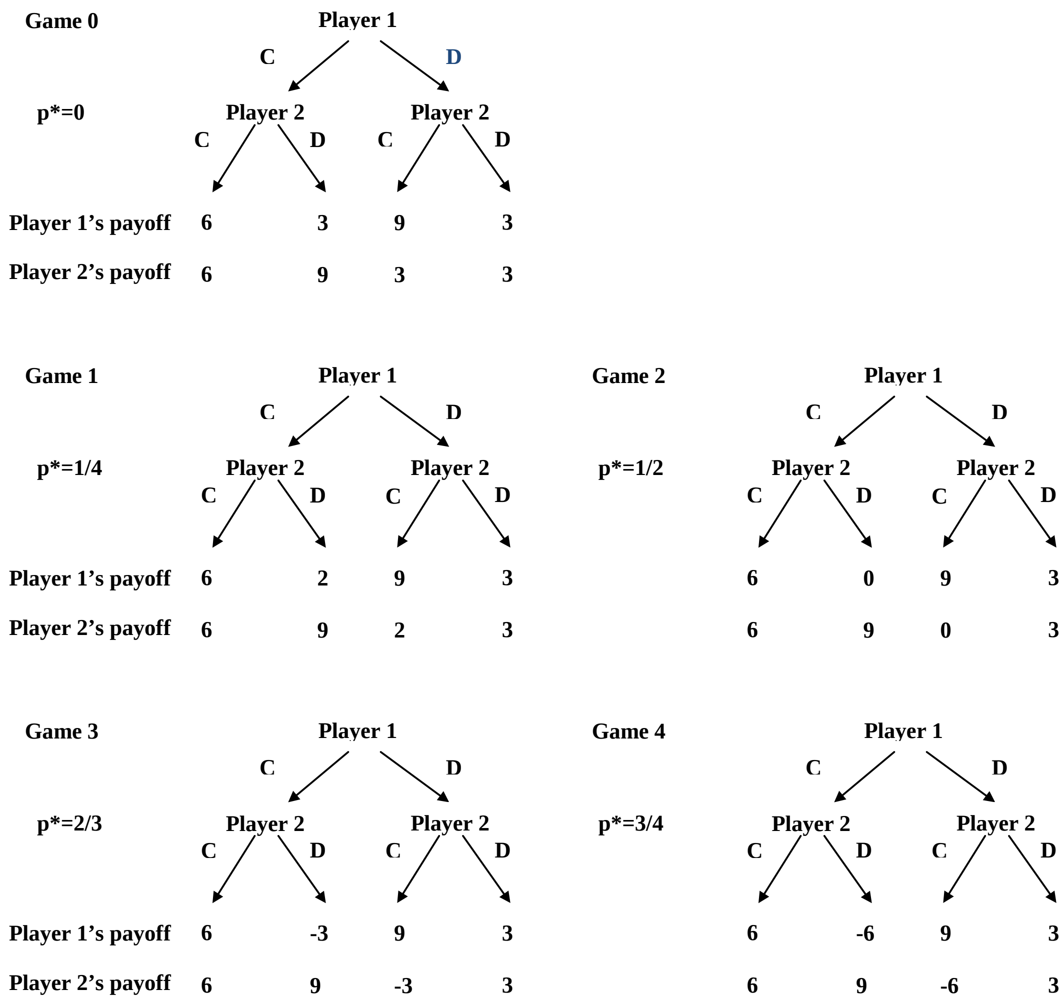


Figure 3.1 Extensive Form Representation of Games Used in the Experiment

Therefore, player 1 prefers to choose the lottery rather than choosing Defect if and only if the likelihood that player 2 is a conditional cooperator is sufficiently high, or $p \geq p^*$, where

$$p^* = \frac{\pi_1(D, D) - \pi_1(C, D)}{\pi_1(C, C) - \pi_1(C, D)}.$$

In our experiment, payoffs in each game are chosen such that $p^* \in \{0,\ 1/4,\ 1/2,\ 2/3,\ 3/4\}$, which corresponds to games 0 to 4. In game 1, player 1 should cooperate if she believes that at least 1/4 of player 2s are conditional cooperators. In contrast, in game 4, player 1 will cooperate when she believes that the proportion of conditional cooperators exceeds 3/4. Other things being equal, we expect to see the likelihood of player 1's cooperation decrease from game 0 to game 4.

In this design, the range of thresholds for cooperation enables us to measure the sensitivity and robustness of group behavior under varying incentives. This design feature is an improvement over previous studies, where only one threshold is implemented, such as in Yamagishi and Kiyonari (2000) who implement a sequential prisoner's dilemma game with $p^* = 1/2$.

To accurately elicit player 2's type, we use the strategy method. Specifically, player 2 is asked to submit a complete strategy without knowing player 1's choice, in the form of "if A chooses A1, I choose __(B1 or B2); if A chooses A2, I choose __ (B1 or B2)." The use of the strategy method effectively transforms the extensive form games in Figure 3.1 into the normal form games in Figure 3.2.

In normal form representation, game 0 has four Nash equilibria, {(D, DD), (C, DD), (D, CC), (D, DC)}, while each game in games 1-4 has a unique pure strategy Nash equilibrium, (D, DD). Thus, behavior in game 0 measures group effects on *coordination*, while behavior in games 1-4 measures group effects on *cooperation*.

Of player 2's four strategies, DC (i.e., doing the opposite to what player 1 does) warrants more discussion. In games 1-4, DC is weakly dominated by DD, and as expected, empirically adopted least often (Section 4.7). In game 0, however, DC is a *weakly dominant* strategy for player 2. Comparing player 2's two weakly dominant strategies, DD and DC, we note that DC maximizes joint payoffs and Pareto dominates DD. Specifically, if player 1 chooses to defect, DC leads to a higher joint payoff without sacrificing own payoff (3 regardless); however, if player 1 chooses to cooperate (which leads to a joint payoff of 12 regardless what player 2 does), player 2 chooses to defect to maximize self interest. Therefore, we name DC as the rational joint-payoff-maximizing strategy (hereafter rJPM) in game 0. Note that player

Game 0	CC	DD	CD	DC
C	6,6	3,9	6,6	3,9
D	9,3	3,3	3,3	9,3

Game 1	CC	DD	CD	DC
C	6,6	2,9	6,6	2,9
D	9,2	3,3	3,3	9,2

Game 2	CC	DD	CD	DC
C	6,6	0,9	6,6	0,9
D	9,0	3,3	3,3	9,0

Game 3	CC	DD	CD	DC
C	6,6	-3,9	6,6	-3,9
D	9,-3	3,3	3,3	9,-3

Game 4	CC	DD	CD	DC
C	6,6	-6,9	6,6	-6,9
D	9,-6	3,3	3,3	9,-6

Figure 3.2 Normal Form Representation of Games Used in the Experiment

2's other joint-payoff-maximizing strategy, CC, is weakly dominated, and thus not rational.

3.2.3 Experimental Procedure

At both UM and UCLA, we implement one control condition and two identity treatments, each of which has five independent sessions at each university. The two treatments include an ethnic identity treatment where we prime participants' (fragmenting) ethnic identities and a school identity treatment where we prime participants' common school identity. We explain our experimental procedure in detail below.

Common to all three experimental conditions, each session consists of eight subjects and three stages: a pre-experiment questionnaire to prime a participant's natural identity in the treatments and an identity-neutral questionnaire for the control condition, four rounds of two-person prisoner's dilemma (PD) games, each with a different match, and a post-experiment questionnaire to elicit demographics information and to check the effects of priming.

In the first stage, participants in each experimental session fill out a pre-experiment questionnaire designed to prime ethnic or school identity in the two respective treatments, or an identity-neutral questionnaire in the control condition.

In the second stage, eight subjects in each session are randomly assigned as player 1 or 2 in the two-person PD games for four rounds. Although their player roles are fixed during the experiment, their match in each round is different in order to minimize repeated game effects. In each round, each participant plays the five PD games with her match. To control for any game order effect within a treatment, we use a Latin Square design, whereby each of the five sessions in a treatment has a different game order.[2]

Unlike most laboratory experiments that use anonymous matching,[3] we provide the co-player's ethnic background information in all three treatments. Specifically, the co-player's last name appears on the screen in the UM sessions. For example, a participant is told that she is matched with "Chen" or "Smith" while making the decision. The displayed name is the co-player's real last name. At UCLA, we display an acronym that combines three pieces of information including the the co-player's grade standing (Freshman, Sophomore, etc.), ethnicity, and player ID.[4] For example, a participant is told that she is matched with "FreshAsianCA1" or "SophCaucasianCA3." The grading standing and player ID are added to alleviate any potential experimenter demand effect.

Furthermore, since the participants go through several rounds, we expose them to photos as an unobtrusive means to reinforce the primes.[5] We select four pictures for each treatment, and display one picture at a time on the computer screen for five seconds before subjects proceed to the next round. In the ethnic identity treatment, pictures of architecture from China and Europe are shown, while in the school identity treatment, subjects see pictures of their university landmarks. In the control sessions, identity-neutral landscape pictures are shown. These photos were pretested to establish that they primed the appropriate identities and that they were equally positive

[2]The game orders include 0-1-2-3-4, 1-2-3-4-0, 2-3-4-0-1, 3-4-0-1-2, and 4-0-1-2-3, so that each game has appeared once in each position.

[3](Andreoni and Petrie, 2004) is a notable exception where subjects' digital photos are presented to their partners in a laboratory fundraising experiment.

[4]We were not able to obtain UCLA IRB approval to display subject last names.

[5]The use of posters and pictures to prime stereotypes is common procedure in psychological priming studies. For instance, Cheryan, Plaut, Davies and Steele (2009) used posters to make salient stereotypes in the computer sciences. Chen and Bargh (1997) exposed participants to picture of Black and White faces to prime stereotypes associated with race.

in valence.[6] Additionally, we elicit individual beliefs about her match's decision in each game, and reward each correct guess with 2 points. Feedbacks on their matches' actual decisions are not provided until the end of the experiment. The experimental instructions and the pictures (Figures 3.3 and 3.4) are included in Appendix B.

Note in all the treatments, including the control condition, co-player's surname (or ethnicity) is provided to subjects before they make decisions. We choose this design to make the setting more comparable to real-life social interactions at workplaces. When people interact with one another at work, they have the information on their co-workers' ethnicity. Therefore, compared to an alternative design in which no information is provided on the co-player, the current control condition serves as a better benchmark and carries more natural generalization to organization design.

In the third stage, we conduct a post-experiment survey, which collects information on demographics, self-statements, strategies used during the experiments, and evaluation of ethnic stereotypes. The post-experiment questionnaire and summary responses are included in Appendix C.

Table 3.1 Features of Experimental Sessions

	Participants					
Site	UM		UM	UCLA		UCLA
Treatments	Caucasian	Asian	Total	Caucasian	Asian	Total
Fragmenting ID	19	21	8×5	17	23	8×5
Common ID	19	21	8×5	19	21	8×5
Control	20	20	8×5	21	19	8×5

Table 4.2 summarizes the features of the experimental sessions, including treatments, number of participants, and ethnic compositions by treatment. Overall, 30 independent computerized sessions were conducted. Fifteen sessions were conducted at the School of Information Lab at the University of Michigan from May to July 2008, with 62 Asian and 58 Caucasian participants. Another 15 sessions were conducted in the California Social Science Experimental Laboratory (CASSEL) at UCLA in May 2009, with 63 Asian and 57 Caucasian participants. All 240 of our subjects were students from UM and UCLA.

For each session at UM, we pre-screened the last names of potential participants, with a threshold of at least three participants with European last names, and three

[6]For the pretest, we had coders rate the photos on how ethnic, UM/UCLA related, and positive they were. We found that the ethnic architecture were rated as more ethnic than the other photos. The UM/UCLA photos were more UM/UCLA related than the other photos. Furthermore, there were no differences in how positive the photos were.

with Chinese last names. For each session at UCLA, as CASSEL does not allow any ethnic screening, we over-recruited subjects for each session to ensure the same minimal number of Asian and Caucasian students in each experimental session as in UM. Extra subjects were directed into a separate room for a survey session unrelated to this experiment. At each site, each subject participated in only one session. We use z-Tree (Fischbacher, 2007) to program our experiments. Each treatment session lasts approximately one hour, with the first 15 minutes used for instructions. The exchange rate is set to 8 points for $1. In addition, each participant is paid a $5 show-up fee. Average earnings per participant are $20 at UM ($18 at UCLA), including the show-up fee. Data are available from the authors upon request.

3.3 Results

Before we present the results several data issues warrant some discussions. Recall that information on subject's ethnicity is revealed to co-players through last names in the UM experiment. In 5% of UM observations, subjects are matched with their acquaintances. Among the acquaintances, 88% of them come from the same ethnic group, which makes it impossible to disentangle the acquaintance effect from inter-group preference.[7] We thus exclude them from the main analysis. The second issue is that some subjects miscategorize their matches' ethnicities. The post-experiment survey shows this affects 8% of UM observations. For these observations, the match ethnicity is recoded to reflect subjects' perception.[8] We also report results using the actual ethnic identities (i.e., without recoding) in footnotes whenever the recoding affects statistical significance. In addition, seven subjects at UM, self-identified as economics graduate students or post-doc, are significantly more likely to choose to defect compared to other subjects.[9] We include them in the analysis, but also report the results if these observations are excluded. In the UCLA experiment, 15 out of 120 subjects self-identify as being from ethnicities other than Asian or Caucasian. We include them in the analysis, although excluding them does not change the results.

Two common features apply throughout our analysis. First, all the analysis

[7]Among these acquaintance pairings, the proportion of player 2s choosing rJPM in game 0 (or DD in games 1-4) is 69% (62%), compared to 59% (80%) for non-acquaintance pairings.

[8]The matching type is coded as "outgroup" in the analysis if one categorizes the match's ethnicity as "other" or "don't know".

[9]Among economics graduate students and post-doc, the cooperation rate as player 1 is 12.5% in game 0, and 0 in games 1-4. As player 2, the proportion of them choosing rJPM (DD) in game 0 (games 1-4) is 35% (75%).

controls for potential interdependency of individual decisions across games. In the analysis based on aggregate data, test of proportions is used with standard errors clustered at the individual level, and one-sided p values are reported. Standard errors are clustered at the individual level in regressions.[10] Second, we use a 5-percent statistical significance level as our threshold to establish the significance of an effect.

We are interested in the extent to which the ethnic and school priming influences coordination (game 0) and cooperation (games 1-4), respectively. Since earlier studies in social psychology (Brewer, 1999) suggest that favoritism towards ingroup and discrimination against outgroup may occur separately, we examine the treatment effects on ingroup favoritism, outgroup discrimination, and intergroup differentials, respectively. The analysis focuses on individual strategies although the results are largely consistent with actions-based analysis.

Since game 0 has multiple Pareto ranked equilibria, while games 1-4 each have a unique but inefficient Nash equilibrium, we report results separately for game 0 in Table 3.2 and for games 1-4 in Table 3.3. Each table contains summary statistics on the left and treatment effects on the right. The top panel pertains to player 1s' choice of cooperative strategy, and the bottom panel to player 2s' choice of rJPM (DD) strategy in game 0 (1-4). Note player 2 has four strategies. We focus on rJPM in game 0 and DD in games 1-4, since the two strategies are weakly dominant in theory and also the mode of empirical distribution in the respective games.[11] In each panel of Tables 3.2 and 3.3, results are presented separately for each university, and then for the pooled data. The proportions are italicized and bolded if the ingroup-outgroup comparison within treatment is statistically significant ($p \leq 0.05$), whereas the p-value is highlighted in boldface only if the treatment effect is statistically significant ($p \leq 0.05$). Results will be discussed in more detail in the following subsections.

3.3.1 Control

Recall subjects in the control sessions are given information (i.e., last name or acronym) that reveals the match's ethnicity, although the pre-survey is intended to be identity neutral. This design enables us to identify possible group effects associated with the revelation of ethnicity *information*. It makes the setting comparable

[10]Participants make their decisions independently, without any feedback on their decisions until the end of the experiment.

[11]For game 0, we also complete a set of analysis for player 2 which includes both rJPM and CC as the two cooperative strategies. As very few subjects play CC, the results do not differ significantly from those reported in the current version.

to real-life workplaces where co-workers have information about others' ethnicity. The results, presented in columns 1-2 in Tables 3.2 and 3.3, establish a baseline for comparison with the two identity treatments.

Table 3.2 Summary Statistics and Treatment Effects in Game 0

	Proportion of Cooperation						Treatment Effects			
Player 1	Control		Ethnic		School		Ethnic v. Control		School v. Control	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UM	In	Out	In	Out	In	Out	In	Out	In	Out
Asian	64	41	37	20	40	25	0.101	0.082	0.147	0.177
White	24	23	41	36	40	25	0.196	0.196	0.175	0.430
UCLA										
Asian	44	29	31	22	27	27	0.194	0.367	0.140	0.473
White	41	35	36	18	28	33	0.392	0.149	0.239	0.472
UM+UCLA										
Asian	53	36	33	21	33	26	0.060	0.112	0.069	0.228
White	33	29	39	27	32	29	0.349	0.433	0.463	0.478
	Proportion of rJPM						Treatment Effects			
Player 2	Control		Ethnic		School		Ethnic v. Control		School v. Control	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UM	In	Out	In	Out	In	Out	In	Out	In	Out
Asian	41	26	45	43	77	76	0.414	0.214	0.016	0.012
White	69	62	54	73	70	71	0.220	0.269	0.483	0.296
UCLA										
Asian	39	38	38	18	50	50	0.491	0.091	0.289	0.296
White	50	36	71	67	56	55	0.158	0.109	0.396	0.208
UM+UCLA										
Asian	40	33	42	30	64	64	0.450	0.406	0.038	0.021
White	57	53	63	70	61	63	0.351	0.120	0.406	0.239

Pairwise comparisons between columns 1-2 in the top panels of Tables 3.2 and 3.3 show higher rate of ingroup cooperation than outgroup cooperation by almost all player 1s in both sets of games, except UCLA Asians in games 1-4. This intergroup difference is statistically significant for Asian player 1s in game 0, and Caucasian player 1s in games 1-4.[12] Similar comparisons in the bottom panel of Table 3.2 indicate that player 2s in game 0 are more likely to choose a joint payoff maximizing strategy rJPM with an ingroup than with an outgroup match. These observations in

[12] UM Caucasian player 1s in games 1-4 are significantly more cooperative with an ingroup match (28% with ingroup vs. 18% with outgroup, $p = 0.026$). If we do not correct for misperceptions of the match's ethnicity in the UM data, this effect becomes weakly significant (25% vs. 20%, $p = 0.057$).

Table 3.3 Summary Statistics and Treatment Effects in Games 1-4

	Proportion of Cooperation						Treatment Effects			
Player 1	Control		Ethnic		School		Ethnic v. Control		School v. Control	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UM	In	Out	In	Out	In	Out	In	Out	In	Out
Asian	30	23	38	24	20	19	0.325	0.470	0.260	0.389
White	28	18	41	39	33	24	0.215	0.070	0.389	0.346
UCLA										
Asian	36	38	24	36	30	18	0.183	0.462	0.305	0.044
White	33	26	23	20	35	38	0.266	0.320	0.453	0.202
UM+UCLA										
Asian	34	28	30	30	25	18	0.369	0.452	0.201	0.119
White	31	22	33	30	34	30	0.419	0.225	0.388	0.241
	Proportion of DD						Treatment Effects			
Player 2	Control		Ethnic		School		Ethnic v. Control		School v. Control	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UM	In	Out	In	Out	In	Out	In	Out	In	Out
Asian	68	58	67	76	83	90	0.485	0.122	0.144	0.028
White	88	76	92	93	90	85	0.355	0.018	0.441	0.218
UCLA										
Asian	85	80	63	52	50	44	0.043	0.066	0.015	0.027
White	55	50	71	69	61	70	0.192	0.124	0.363	0.113
UM+UCLA										
Asian	76	71	65	64	66	69	0.126	0.292	0.190	0.457
White	67	67	81	83	71	77	0.131	0.061	0.376	0.156

the control sessions thus suggest, at least qualitatively, ingroup favoritism and outgroup discrimination as a result of the match's ethnicity information being revealed. These results are partially consistent with Fershtman and Gneezy (2001), who find that Israeli Jewish participants exhibit mistrust towards men of Eastern origin in trust games, where ethnic origins are inferred from the names of their matches.

In contrast to player 2s' choices in game 0, their choices in games 1-4 show outgroup favoritism. The rate of the always-defect strategy DD is at least as high with an ingroup than with an outgroup match (columns 1-2 of bottom panel in Table 3.3).[13] This lower rate of DD with an outgroup match is due to player 2s' increased

[13]The intergroup difference is significant for UM Caucasian player 2s (88% vs. 76%, $p < 0.01$). However, this comparison is not significant if we do not correct the misperceptions of ethnic identities (83% vs. 78%, $p = 0.190$).

positive reciprocity towards an outgroup match in games 1-4.[14] This outgroup favoritism, while opposite to the findings in earlier studies with near-minimal groups in the lab (Chen and Li, 2009), has been reported in studies with natural identities.[15] Nevertheless, since we are primarily interested in how different identity priming influences individual choices and their intergroup preferences, these findings in the control sessions serve as a benchmark for the analysis of the treatment effects of ethnic and school priming.

3.3.2 Ethnic Priming: Fragmenting Identities

This subsection investigates behavioral changes in the ethnic priming treatment relative to the control sessions. Recall subjects in both treatments are given information on the match's ethnicity. The only difference is that the pre-survey in the ethnic priming treatment is used to activate ethnic identities, whereas that in the control is designed to be identity neutral. We focus on the treatment effect, i.e., how ethnic priming, in addition to information on match's ethnicity (surnames/acronyms), influences the way subjects treat ingroup and outgroup as well as intergroup differentials, relative to the control.

Compared to the control sessions, we expect that ingroup favoritism (and outgroup discrimination) will be stronger in the ethnic priming treatment when the ethnicity of participants is made more salient.

Hypothesis 7 (Ethnic Priming)**.** *Compared to the control, players are more cooperative with those from the same ethnic group, and less cooperative with those from the other ethnic group in the ethnic priming treatment.*

Hypothesis 7 implies that, compared to the control, player 1s will cooperate more with an ingroup match in the ethnic priming treatment, whereas player 2s will be more (less) likely to choose rJPM (DD) with an ingroup match in game 0 (games 1-4) under the ethnic priming treatment.

Summary statistics for the ethnic priming treatment are reported in columns 3-4, and the treatment effects in columns 7-8 of Tables 3.2 (for game 0) and 3.3 (for

[14]For UM Caucasian player 2s in the control (games 1-4), when player 1 cooperates, the proportion of cooperation is 10% for an ingroup match and 20% for an outgroup match ($p = 0.001$, one-sided). When player 1 defects, the proportion of cooperation is 16% for an ingroup match and 17% for an outgroup match ($p = 0.156$, one-sided).

[15]For example, Friesen, Arifovic, Ludwig, Wright, Giamo and Baray (2011) report outgroup favoritism exhibited by East Asian children in a dictator game.

games 1-4). Table 3.2 indicates that Asian player 2s choose rJPM in game 0 more often with ingroup than with outgroup. This observation indicates that, for Asians, the ethnic treatment preserves ingroup favoritism and outgroup discrimination for the coordination game found in the control. The ethnic priming also influences the ingroup-outgroup gap relative to the control - the gap significantly increases for UCLA Asian player 2s as a result of the ethnic priming. This leads to Result 6.

Result 6 (Ethnic Priming on Coordination: Asians)**.** *In the ethnic priming treatment, Asian player 2s are significantly more likely to choose rJPM in game 0 with an ingroup than with an outgroup match. Furthermore, compared to the control sessions, this intergroup difference is significantly stronger for UCLA Asian player 2s.*

Support. *Table 3.2 indicates that Asian player 2s (UM+UCLA) choose rJPM in game 0 more often with ingroup than with outgroup (42% vs. 30%,* $p = 0.035$*).*[16] *Ethnic priming makes UCLA Asian player 2s differentially more likely to choose rJPM in game 0 with an ingroup than with an outgroup match (38% vs. 18%,* $p = 0.006$*), compared to the control (39% vs. 38%). For the treatment effect, the one-sided p-value of the difference-in-difference analysis is* 0.031.

Result 6 indicates that, in the coordination game (game 0), ethnic priming exacerbates the intergroup gap compared to the control for Asian player 2s. We now turn to games 1-4, each of which has a unique Nash equilibrium. Table 3.3 indicates that in games 1-4 ethnic priming makes UM player 2s more likely to choose DD strategy with an outgroup than with an ingroup match, increasing the degree of discrimination against the outgroup and hence reversing their outgroup favoritism in the control. Specifically, in the UM experiment, the likelihood to choose DD with outgroup increases from 58% in the control to 76% in the ethnic priming treatment by Asian player 2s (from 53% to 83% if economics graduate students/post-doc are excluded, $p = 0.017$), and increases from 76% to 93% by Caucasian player 2s ($p = 0.018$). In the UCLA experiment, the ethnic priming tends to enhance ingroup favoritism - the likelihood to choose DD with ingroup decreases from 85% in the control to 63% in the ethnic priming treatment by Asian player 2s ($p = 0.043$). In addition, the ethnic priming significantly increases UM Asian player 2s' intergroup differentials - in favor of ingroup - in cooperation, which leads to Result 7.

Result 7 (Ethnic Priming on Cooperation: UM Asians)**.** *Ethnic priming makes UM Asian player 2s differentially more likely to choose DD with an outgroup than with an*

[16] This effect is stronger if UM economic graduate students and postdocs are excluded (43% vs. 28%, $p = 0.012$).

ingroup match (76% vs. 67%), compared to the control (58% vs. 68%). This effect is stronger if economics graduate students and postdocs are excluded.

Support. *One-sided p-value in the difference-in-difference analysis is* 0.075*. When we exclude the economic graduate students and post-doc who choose to defect most time regardless the matching type, the rate of choosing DD with ingroup (outgroup) becomes 63% (53%) in control, and 66% (83%) in the ethnic treatment ($p = 0.021$, difference-in-difference analysis). In this case, the increase in intergroup differentials is primarily driven by the sharp increase in the choice of DD with outgroup ($p = 0.017$).*

In sum, compared to the control, while ethnic priming increases ingroup favoritism and outgroup discrimination for Asians and Caucasians alike, statistically significant impact is found only for Asians. Thus, we reject the null in favor of Hypothesis 7 for Asians. Analysis in this subsection suggests that Asians' social identities may be more malleable compared to Caucasians.

3.3.3 School Priming: Common Identity

We next evaluate how school identity priming influences individual behavior compared to the control sessions. Recall the common identity prime, implemented in the pre-survey, is designed to subtly activate individual's common identity of being part of her university. We again focus on the treatment effect, i.e., how school identity priming, in addition to the information on a match's ethnicity, influences ingroup favoritism, outgroup discrimination, and intergroup differentials, relative to the control.

In the school priming treatment, we expect less intergroup bias compared to the control. More specifically, we expect to observe an overall increase of cooperation for player 1s and an increase (decrease) of the rational joint-payoff-maximizing (always-defect) strategy for player 2s in game 0 (games 1-4).

Hypothesis 8 (School Priming)**.** *Compared to the control, in the school priming treatment, a player will be more cooperative with an ingroup and an outgroup match.*

Hypothesis 8 implies that, compared to the control, in the school priming treatment, player 1 will be more likely to cooperate with an ingroup and an outgroup match. For player 2, the likelihood of adopting the rational joint-payoff-maximizing (DD) strategy increases (decreases) from the control to the school priming treatment in game 0 (games 1-4).

Summary statistics for the school priming treatment are reported in columns 5-6, and the treatment effects in last two columns of Tables 3.2 (for game 0) and 3.3 (for games 1-4). A comparison between columns 5-6 with 1-2 in the upper panel of Tables 3.2 and 3.3 indicates that there is no treatment effect for player 1s, except that UCLA Asians cooperate less with an outgroup match in the treatment than in the control ($p = 0.044$, column 10 in Table 3.3). However, a comparison between columns 5-6 with 1-2 in the lower panel of Table 3.2 reveals positive treatment effects. We find that the likelihood that player 2s choose rJPM in game 0 increases for both the ingroup and outgroup matching, from the control to the school priming treatment. We also find that the ingroup-outgroup gap in the proportion of rJPM reduces from 7% in the control to zero in the school priming treatment for Asians. This suggests that, for Asians, common identity priming is effective in enhancing ingroup favoritism while alleviating outgroup discrimination, which consequently reduces the degree of intergroup discrimination. The impact of school identity priming is statistically significant for Asian player 2s, which leads to Result 8.[17]

Result 8 (School Priming on Coordination: Asians). *The common school identity priming makes Asian player 2s significantly more likely to choose rJPM in game 0 for ingroup matching (64% compared to 40% in control;* $p = 0.038$*), and for outgroup matching (64% compared to 33% in control;* $p = 0.021$*).*

By Result 8, we reject the null in favor of Hypothesis 8 for Asians in the coordination game. Compared to the impact of school priming on coordination in game 0, its impact on cooperation in games 1-4 is more complex. Pooling data from both schools yields no statistically significant results, as school priming impacts UM and UCLA Asians differently. Specifically, Asian player 2s react to school priming in opposite ways at UM and UCLA (Table 3.3 lower panel). We will first summarize the results and then try to identify possible reasons behind the difference in results.

Result 9 (School Priming on Cooperation: UCLA Asians). *In games 1-4, the common school identity priming makes UCLA Asian player 2s significantly less likely to choose the always-defect strategy (DD) for both the ingroup (50% compared to 85% in control,* $p = 0.015$*) and the outgroup match (44% compared to 80% in control,* $p = 0.027$*). The proportion of DD reduces significantly from 82% in the control to 48% in this treatment (*$p = 0.022$*).*

[17]Results continue to hold if UM economic graduate students/postdoc are excluded. The likelihood for them to choose rJPM significantly increase for ingroup (47% in control vs. 75% in school priming; $p = 0.086$) and outgroup matches (29% vs. 74%; $p = 0.036$). The intergroup difference reduces from 18% to 1%, and the average increases from 38% to 74% ($p = 0.021$).

Result 10 (School priming on Cooperation: UM Asians)**.** *The common school identity priming makes UM Asian player 2s significantly more likely to choose the always-defect strategy (DD) for both the ingroup (83% compared to 68% in control,* $p = 0.144$*) and the outgroup match (90% compared to 58% in control,* $p = 0.028$*). The proportion of DD increases from 63% in the control to 87% in the treatment (*$p = 0.049$*).*[18]

Results 9 and 10 show that Asian player 2s react to the school priming differently at the two universities. While it makes UCLA Asian player 2s more cooperative, it makes UM Asian player 2s more competitive, compared to their counterparts in the control sessions. Social psychology research shows that when social identities are primed, individual behavior tends to conform to stereotypes (i.e., some innate statistical models of characteristics or behaviors) of the social categories associated with the primed identities (Shih et al., 1999). We thus conjecture that, while subtly activating a common identity, the school identity priming may also introduce *school specific* cues for behavior (e.g., being competitive) that subsequently influence individual decisions.

This conjecture is formulated based on subject responses to post-experiment survey question of why they chose to attend their university. Among Asian player 2s, 64% at UM report that they chose UM for academic reasons, such as good programs, reputation, and high ranking, significantly higher than the corresponding 20% at UCLA ($p = 0.044$, one-sided χ^2 test), where a substantially higher proportion refer to nonacademic reasons, such as location, food and in-state tuition. Statistics on individual perceptions are also consistent with this conjecture. In the post-experiment survey, subjects are asked to report, on a 1 to 7 Likert scale, their perceptions of the *competitiveness* of each ethnicity. UM Asians report 6.27 for Asians in the school priming treatment, significantly higher than the 5.4 in the control ($p = 0.05$, one-sided Wilcoxon rank-sum test). By contrast, UCLA Asians report 5.40 for Asians in the school priming treatment, comparable to the 5.45 in the control. In other words, the school identity priming may have influenced the ethnic stereotype of being competitive, particularly among UM Asians. In the analysis below, we use probit regressions to study determinants of player 2 choices of DD in games 1-4. We are interested in how perceptions of one's own ethnicity being competitive (coded as *SelfCompetitiveness*) interacts with school priming, and whether the interaction effects can explain

[18] Result 10 still holds even when economics graduate students/post-doc are excluded. In games 1-4, school priming significantly increases UM Asian player 2s' choice of DD for ingroup (85% vs. 63% in the control, $p = 0.062$) and outgroup (86% vs. 53% in the control, $p = 0.056$). The average likelihood of DD increases from 58% in the control to 85% ($p = 0.041$).

the discrepancy between Results 9 and 10.[19]

Table 3.4 Effects of School Priming on Stereotypes: Games 1-4

	Likelihood of Always Defect (DD)			
	UM		UCLA	
	(1) Asian	(2) Caucasian	(3) Asian	(4) Caucasian
School Priming	-4.917**	-0.770	2.273	2.225
	(2.485)	(2.045)	(1.675)	(1.360)
Ingroup	-0.034	0.462**	0.251	0.008
	(0.307)	(0.189)	(0.186)	(0.230)
School×Ingroup	-0.209	-0.126	-0.015	-0.157
	(0.506)	(0.174)	(0.239)	(0.246)
SelfCompetitive	-0.090	0.091	0.422*	0.227
	(0.239)	(0.320)	(0.228)	(0.199)
School×SelfCompetitive	0.993**	0.200	-0.620**	-0.377
	(0.447)	(0.388)	(0.302)	(0.287)
Women	0.440	-1.115**	0.454	0.684
	(0.495)	(0.478)	(0.525)	(0.454)
Age	0.071	-0.014	0.037	-0.097
	(0.062)	(0.069)	(0.192)	(0.184)
Game2	0.314	0.148	0.144	0.091
	(0.192)	(0.304)	(0.133)	(0.239)
Game3	-0.175	-0.030	0.138	-0.062
	(0.114)	(0.297)	(0.172)	(0.193)
Game4	0.177	0.233	0.103	-0.063
	(0.161)	(0.313)	(0.143)	(0.211)
Observations	316	280	336	304
Log Pseudo L.	-137.551	-118.132	-173.845	-177.169
Pseudo R^2	0.216	0.079	0.197	0.135

Notes:
a. Robust standard errors in parentheses are clustered at the individual level.
b. Significant at: * 10-percent level; ** 5-percent level; *** 1-percent level.

Results are reported in Table 3.4 separately for UM (column 1 for Asian player 2s, and column 2 for Caucasian player 2s) and UCLA (column 3 for Asian and column 4 for Caucasian). We pool data from the control and the school priming treatment for each university. The dependent variable is the likelihood of player 2s choosing the always-defect strategy DD. The independent variables include the school prim-

[19]The SelfCompetitiveness stereotype has the strongest predictive power among all stereotype variables. It takes on a value between 1 and 7, depending on subject's belief about the competitiveness of her own ethnic group.

ing treatment dummy (the control in the omitted category), the ingroup matching dummy, their interaction, the SelfCompetitiveness variable and its interaction with the school treatment dummy. We also control for gender, age and game fixed effects. Standard errors are clustered at the individual level. Results hold if the economics graduate students/post-doc are excluded from the UM data.

Result 11 (Stereotypes)**.** *In games 1-4, school priming significantly alleviates the negative effect of the competitiveness stereotype on cooperation by UCLA Asian player 2s, but enhances this negative effect on cooperation by UM Asian player 2s.*

Support. *The effect of SelfCompetitive in column 3 of Table 3.4 (0.422, $p < 0.10$) suggests that in the control, UCLA Asian player 2s who perceive Asians to be competitive are more likely to act competitively by choosing DD in games 1-4. The impact of the self perception of competitiveness is largely offset by school priming, as shown by the interaction effect of School×SelfCompetitive -0.620 ($p < 0.05$). In contrast, school priming at UM (unexpectedly) triggers Asian subjects' self perception of competitiveness, as suggested by the interaction effect of School×SelfCompetitive (0.993, $p < 0.05$), leading to higher rate of DD by Asian player 2s in the school treatment relative to control.*[20]

While Result 8 reports the positive impact of UM school priming on coordination for Asians player 2s in game 0, Result 10 reports its negative impact on cooperation in games 1-4. The two results, however, do not contradict each other. As discussed earlier, the net influence of school priming depends on two competing factors: the common identity that may improve the level of cooperation and coordination, and individual perception of self-competitiveness triggered by the school identity priming at UM. Which of the two forces dominates depends on the game structure. For example, it is costless for player 2s to choose rJPM over DD in game 0, but it is costly, on absolute terms, for them to do so in games 1-4. So the payoff structure in games 1-4 leaves room for Asian player 2s, who believe Asians to be more competitive than their counterparts in control do, to act competitively by choosing DD to maximize own payoffs. By contrast, the payoff structure in game 0 makes player 2s less likely to be affected by such competitiveness stereotype, since the cost is zero for them to act pro-socially. Thus, the effect of common identity dominates, which leads to an increase in the rate of coordination. It is worth noting that this positive impact of

[20] For UCLA Asians, the marginal effects are 0.114 for SelfCompetitive and -0.166 for School×*SelfCompetitive*. For UM Asians, the marginal effects are -0.028 and 0.193, respectively.

school identity priming on coordination survives despite a relative payoff disadvantage that player 2s have to face when choosing rJPM.

Overall, our results suggest that the impact of common identity priming depends crucially on the incentive structure, as well as organization-specific stereotypes triggered for each social group. Thus, while the social psychology literature focuses on the positive effects of a common ingroup identity (Cunningham, 2005), our study suggests that its effects might be incentive, institution and stereotype specific. Further, in both the UM and UCLA experiments, Asians are generally responsive to both the ethnic identity and common identity priming, but Caucasians are not. It remains an open question whether minorities tend to be more responsive to identity priming.

3.4 Discussions

As the workforce becomes increasingly diverse, organizations more frequently encounter the issue of motivating individuals from different backgrounds to work together towards a common goal. Our paper investigates the effects of priming a fragmenting (ethnic) versus a common organization identity on coordination and cooperation among Asian and Caucasian students in a controlled laboratory experiment.

We have several new findings. First, priming a fragmenting (ethnic) identity has a significant impact on the intergroup preference for Asians compared to the control, increasing both their ingroup favoritism and their outgroup discrimination, while it has no effect on Caucasians. Moreover, priming a common (school) identity reduces group bias for Asians in the coordination game, resulting in a significant increase in both ingroup and outgroup cooperation. However, in games with a unique inefficient Nash equilibrium, the effects of priming a common identity are more complex. While priming alleviates the negative effects of the competitiveness stereotype on cooperation among UCLA Asians, it enhances such negative effects among UM Asians. This result suggests that identity priming might work through its interaction with the activated stereotypes. Lastly, in both our treatments, Asians are responsive to priming, while Caucasians are not. It remains an open question whether minorities might be more responsive to identity priming than majorities. To answer this question, more studies are needed to investigate behavior among other ethnic groups.

This paper suggests that Asians, especially first-generation immigrants (UM), are more likely to be influenced by intergroup preferences than are Caucasians. In addition, the findings suggest that their identities are malleable, which consequently influences their behavior. Since first-generation Asians are more responsive to both fragmenting and common identity priming, our results offer new insights into socializing new immigrants.

Immigrants have become a substantial and increasingly important segment of the labor force in the United States and many other parts of the world. In 2004, one in seven workers in the United States, i.e., more than 21 million workers, were foreign born. These foreign-born workers accounted for more than half of the growth of the U.S. labor force during the past decade. Among these foreign-born workers, 40 percent come from Mexico and Central America, 25 percent from Asia, and the rest from the Western Hemisphere and Europe. More than 30 percent held bachelor's or more advanced degrees. Due to the native-born baby-boomers' exit from the labor force and the injection of these immigrant workers into the labor force, workplaces will continue to become more diverse. The U.S. Congressional Budget Office predicts that "[u]nless native fertility rates increase, it is likely that most of the growth in the U.S. labor force will come from immigration by the middle of the century."

Although economic assimilation of immigrants, i.e., the change in the wage gap between immigrant and native-born workers (Borjas, 1994, 1999), has been extensively studied in labor economics, immigrant social assimilation, especially at workplaces, has been significantly understudied.[21] This study underscores the importance to understand the factors that influence immigrant workers' social assimilation and the impact on their social interactions with others at workplaces. It also has important policy implications for organizational management. For example, building employees' common identity in an organization may serve as an identity-based mechanism to raise the cooperation and coordination level among employees in strategic environments and, consequently, increase the overall productivity of the organization. Organizations may also benefit from helping their immigrant workers' social assimilation process and promoting social networking across ethnic lines, or between native-born and foreign-born workers within the organizations.

It would be interesting for future research to study the impact of these policies on behaviors by workers from other ethnic groups (e.g., workers from Mexico and Central America), and to study whether the results can be generalized beyond ethnic lines to other "group" contexts at diverse workplaces, such as gender groups or different professional groups. Finally, we hope to extend this study to the field, and investigate the extent to which organizational policy design that focuses on common

[21] An exception is Cox and Orman (2010) who study immigrants' trust and trustworthiness in a lab experiment.

identity building may influence cooperation and coordination among workers.

3.5 Appendix A: Pre-experiment Questionnaire

A.1 Control sessions

We are interested in your opinions and experiences about certain aspects of young adult life.

1. Name: ______________ (UM only)
2. Age: ______________ (UM: *Mean 23.3, Std Dev 4.3, Median 22, Min 19, Max 42*) (UCLA: *Mean 19.8, Std Dev 1.6, Median 19, Min 17, Max 24*)
3. Grade/Year:
   - (a) Freshmen (UM: *0%*) (UCLA: *42.5%*)
   - (b) Sophomore (UM: *0%*) (UCLA: *17.5%*)
   - (c) Junior (UM: *17.5%*) (UCLA: *17.5%*)
   - (d) Senior (UM: *30%*) (UCLA: *17.5%*)
   - (e) $>$ 4 years (UM: *5%*) (UCLA: *2.5%*)
   - (f) Graduate student (UM: *47.5%*) (UCLA: *2.5%*)
4. How often do you watch television?
   - (a) every day (UM: *17.5%*) (UCLA: *20%*)
   - (b) $4-5$ times a week (UM: *22.5%*) (UCLA: *15%*)
   - (c) $2-3$ times a week (UM: *22.5%*) (UCLA: *32.5%*)
   - (d) a few times a month (UM: *25%*) (UCLA: *17.5%*)
   - (e) a few times a year (UM: *5%*) (UCLA: *5%*)
   - (f) rarely if ever (UM: *5%*) (UCLA: *10%*)
   - (g) Never (UM: *2.5%*) (UCLA: *0%*)
5. Do you have cable television?
   - (a) yes (UM: *70%*) (UCLA: *67.5%*)
   - (b) no (UM: *30%*) (UCLA: *32.5%*)
6. How often do you eat out?
   - (a) every day (UM: *7.5%*) (UCLA: *2.5%*)
   - (b) $4-5$ times a week (UM: *12.5%*) (UCLA: *2.5%*)
   - (c) $2-3$ times a week (UM: *27.5%*) (UCLA: *42.5%*)
   - (d) a few times a month (UM: *42.5%*) (UCLA: *45%*)
   - (e) a few times a year (UM: *7.5%*) (UCLA: *5%*)
   - (f) rarely if ever (UM: *0%*) (UCLA: *2.5%*)
   - (g) Never (UM: *2.5%*) (UCLA: *0%*)
7. How often do you attend movies?
   - (a) every day (UM: *0%*) (UCLA: *0%*)
   - (b) $4-5$ times a week (UM: *0%*) (UCLA: *0%*)
   - (c) $2-3$ times a week (UM: *2.5%*) (UCLA: *2.5%*)
   - (d) a few times a month (UM: *32.5%*) (UCLA: *15%*)
   - (e) a few times a year (UM: *52.5%*) (UCLA: *70%*)
   - (f) rarely if ever (UM: *7.5%*) (UCLA: *12.5%*)
   - (g) Never (UM: *5%*) (UCLA: *0%*)

A.2 Ethnic Priming Treatment

We are interested in your opinions and experiences about certain aspects of young adult life.

1. Name: ______________ (UM only)
2. Age: ______________ (UM: *Mean 23.8, Std Dev 4.6, Median 22, Min 18, Max 40*) (UCLA: *Mean 20, Std Dev 1.4, Median 20, Min 18, Max 23*)
3. Grade/Year:
   (a) Freshmen (UM: *2.6%*) (UCLA: *27.5%*)
   (b) Sophomore (UM: *12.8%*) (UCLA: *30%*)
   (c) Junior (UM: *5.1%*) (UCLA: *25%*)
   (d) Senior (UM: *18%*) (UCLA: *12.5%*)
   (e) > 4 years (UM: *10.3%*) (UCLA: *5%*)
   (f) Graduate student (UM: *51.3%*) (UCLA: *0%*)
4. Ethnicity:
   (a) African
   (b) Asian (UM: *48.7%*) (UCLA: *55%*)
   (c) European (UM: *51.3%*) (UCLA: *35%*)
   (d) Hispanic
   (e) Native
   (f) other (UCLA: *10%*)
   if it is other, please specify:______________

5. How many generations has your family lived in America?
   (a) First Generation (UM: *48.7%*) (UCLA: *30%*)
   (b) Second Generation (UM: *35.9%*) (UCLA: *30%*)
   (c) More than Two Generations (UM: *15.4%*)(UCLA: *40%*)
6. From which countries did you family originate? ______________
7. What languages do you speak? ______________
8. Are you involved in any student organizations?
   (a) yes (UM: *46.2%*) (UCLA: *82.5%*)
   (b) no (UM: *53.9%*) (UCLA: *17.5%*)
   If yes, which ones? ______________

A.3 School Priming Treatment

We are interested in your opinions and experiences about certain aspects of young adult life.

1. Name: ______________ (UM only)
2. Age: ______________ (UM: *Mean 22.2, Std Dev 3.0, Median 21, Min 18, Max 30*) (UCLA: *Mean 20.1, Std Dev 1.4, Median 20, Min 18, Max 24*)
3. Grade/Year:
   (a) Freshmen (UM: *0%*) (UCLA: *30%*)
   (b) Sophomore (UM: *18.9%*) (UCLA: *17.5%*)

(c) Junior (UM: *10.8%*) (UCLA: *25%*)
(d) Senior (UM: *35.1%*) (UCLA: *20%*)
(e) > 4 years (UM: *0%*) (UCLA: *2.5%*)
(f) Graduate student (UM: *35.1%*) (UCLA: *5%*)

4. School: ______________
5. Did you consider any other schools?
   (a) yes (UM: *62.2%*) (UCLA: *77.5%*)
   (b) no (UM: *37.8%*) (UCLA: *22.5%*)
   If yes, what other schools? ______________

6. Why did you decide to choose your specific school? ______________

3.6 Appendix B: Experimental Instruction

This is an experiment in decision making. You will be asked to fill out a survey at the beginning of the experiment. You will then make a series of decisions, and fill out another survey at the end of the experiment.

The amount of money you earn will depend upon the decisions you make and on the decisions other people make. In addition, you will be paid $5 for participation. Everyone will be paid in private and you are under no obligation to tell others how much you earned.

Please do not communicate with each other during the experiment. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

Roles: This experiment has 8 participants, four of whom are player As and the other four are player Bs. Your assigned role will be the same for all the games. Therefore, if you are a player A, you will always be a player A. Similarly, if you are a player B, you will always be a player B.

Matching: In each of the four rounds, a player A will be matched with a player B. You will never be matched with the same player twice.

Procedure: In each of the four rounds, both players A and B will make decisions on each of five games. The outcome of each game depends on the decisions of both players.

For instance, in the Example for Review Questions on the next page, player A moves first, by choosing A1 or A2. After A makes a decision, A will be asked to guess what B will choose.

Without knowing A's decision, player B will be asked to first guess what player A has chosen. Then player B decides whether to choose B1 or B2 under each of two scenarios: (1) Player A chooses A1; (2) Player A chooses A2.

Payoff for each game is determined by both players' decisions. For example, if player A chooses A1, and player B's decision is B2 if A chooses A1, and B1 if A chooses A2, the outcome of the game is (A1, B2), with payoffs 40 for A and 30 for B. Note that all of A's decisions and payoffs are in red, while B's are in blue.

In addition, a player earns 2 points for each correct guess. For example, if player A's guess is that B will choose B2. If it turns out to be correct, A will get 2 points. Otherwise, A will get zero point.

Feedback: You will not get any feedback after each game. At the very end of the experiment, you will be shown a history screen, with your decisions, your match's decisions, the accuracy of your guesses, and your payoff for each of the twenty games.

Total Payoffs: In each of the four rounds, your payoff will be the sum of your payoffs in all five games. Your total payoff will be the sum of your payoffs in all four rounds, i.e., in all 20 games. Your earnings are given in points. At the end of the experiment you will be paid based on the following exchange rate:

$1 = 8 points.

In addition, you will be paid $5 for participation, and 25 cents for answering each of the review questions correctly.

Review Questions: To help you understand the game, we will go over a number of review questions about the following made-up example. Each correct question is worth 2 points.

1. If Player A chooses A1, and player B chooses B1 when A chooses A1, A's payoff is ______________, and B's payoff is ______________.
2. If Player A chooses A1, and player B chooses B2 when A chooses A1, A's payoff is ______________, and B's payoff is ______________.
3. If Player A chooses A2, and player B chooses B1 when A chooses A2, A's payoff is ______________, and B's payoff is ______________.
4. If Player A chooses A2, and player B chooses B2 when A chooses A2, A's payoff is ______________, and B's payoff is ______________.
5. Player B guessed that Player A had chosen A1.
   If Player A actually chooses A1, Player B's payoff from her guess is ______________ points.
   If Player A actually chooses A2, Player B's payoff from her guess is ______________ points.
6. True or False: you are always matched with the same player throughout the Experiment.
   (a) True
   (b) False

Please raise your hand if you are finished with the review questions. An experimenter will come over and grade it. Please check that you have written down your name and ID number on the first page.

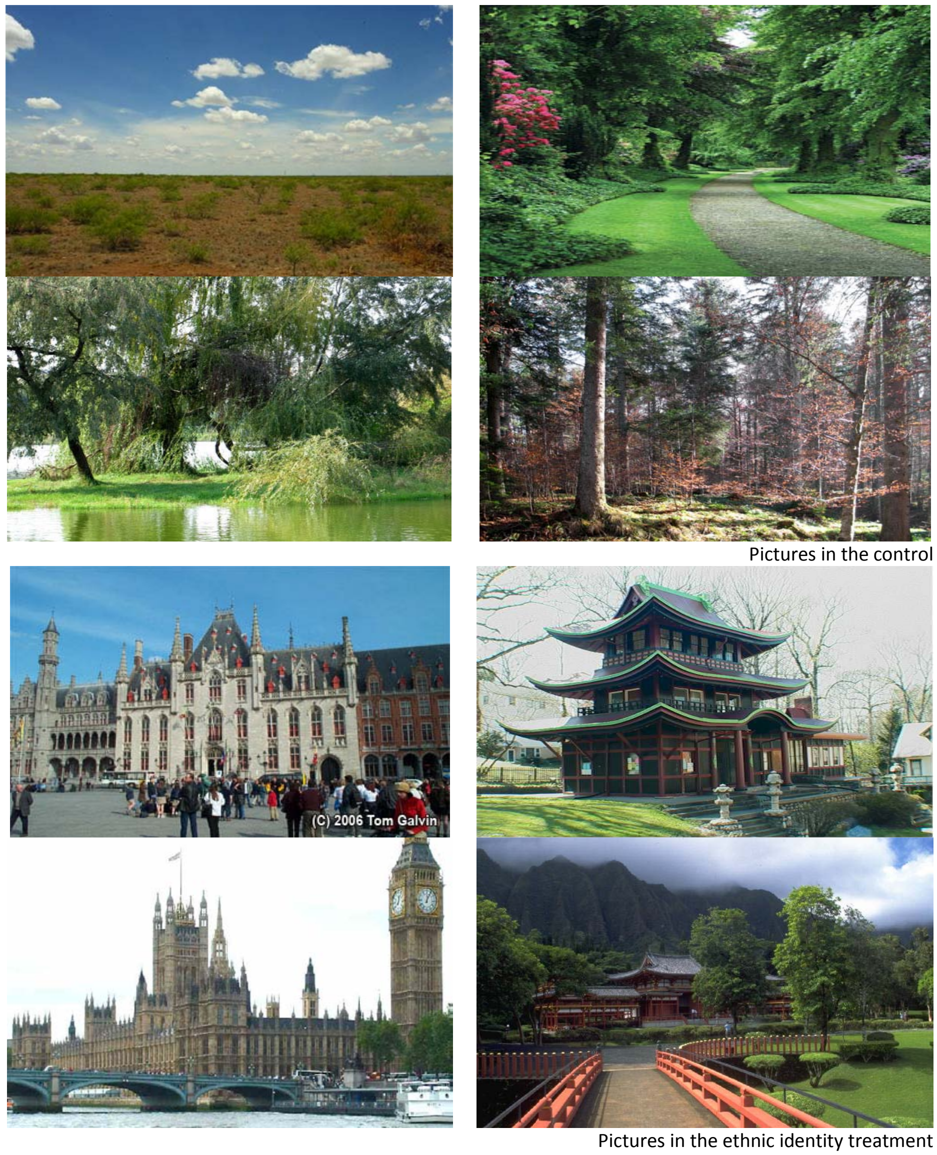


Figure 3.3 Priming Pictures: Control/Ethnic Priming

Pictures in the UM identity treatment

Pictures in the ULCA identity treatment

Figure 3.4 Priming Pictures: School Priming

3.7 Appendix C: Post-experiment Questionnaire

1. Please write five statements in answer to the question:“ Who am I?”
2. Gender
   (a) Male (UM: *43.1%*) (UCLA: *44.2%*)
   (b) Female (UM: *56.9%*) (UCLA: *55.8%*)
3. Ethnicity:
   (a) African (UM: *0.9%*) (UCLA: *0%*)
   (b) Asian (UM: *48.3%*) (UCLA: *50.8%*)
   (c) European (UM: *48.3%*) (UCLA: *35.8%*)
   (d) Hispanic (UM: *0%*) (UCLA: *0%*)
   (e) Native (UM: *1.7%*) (UCLA: *0%*)
   (f) other (UM: *0.9%*) (UCLA: *13.3%*)
   if it is other, please specify:______________

4. From which countries did you family originate?
5. What do you think is the experiment about?
6. How common do you think these stereotypes are in society?
   (a) Asian Americans are strategic (UM: *Mean 5.1, Std Dev 1.4, Median 5, Min 1, Max 7*) (UCLA: *Mean 4.7, Std Dev 1.6, Median 5, Min 1, Max 7*)
   (b) Asian Americans are trustworthy (UM: *Mean 4.0, Std Dev 1.4, Median 4, Min 1, Max 7*) (UCLA: *Mean 4.1, Std Dev 1.4, Median 4, Min 1, Max 7*)
   (c) Asian Americans are cooperative (UM: *Mean 4.3, Std Dev 1.7, Median 4, Min 1, Max 7*) (UCLA: *Mean 4.3, Std Dev 1.5, Median 4, Min 1, Max 7*)
   (d) Asian Americans are naive (UM: *Mean 3.5, Std Dev 1.6, Median 3, Min 1, Max 7*) (UCLA: *Mean 3.8, Std Dev 1.6, Median 4, Min 1, Max 7*)
   (e) Asian Americans are sneaky (UM: *Mean 3.8, Std Dev 1.5, Median 4, Min 1, Max 7*) (UCLA: *Mean 3.9, Std Dev 1.6, Median 4, Min 1, Max 7*)
   (f) Asian Americans are competitive (UM: *Mean 5.9, Std Dev 1.4, Median 6, Min 1, Max 7*) (UCLA: *Mean 5.6, Std Dev 1.4, Median 6, Min 1, Max 7*)
   (g) European Americans are strategic (UM: *Mean 4.0, Std Dev 1.8, Median 4, Min 1, Max 7*) (UCLA: *Mean 5.0, Std Dev 1.6, Median 5, Min 1, Max 7*)
   (h) European Americans are trustworthy (UM: *Mean 4.2, Std Dev 1.6, Median 4, Min 1, Max 7*) (UCLA: *Mean 4.2, Std Dev 1.6, Median 4, Min 1, Max 7*)
   (i) European Americans are cooperative (UM: *Mean 4.3, Std Dev 1.5, Median 4, Min 1, Max 7*) (UCLA: *Mean 4.4, Std Dev 1.5, Median 4, Min 1, Max 7*)
   (j) European Americans are naive (UM: *Mean 3.4, Std Dev 1.6, Median 3.5, Min 1, Max 7*) (UCLA: *Mean 3.7, Std Dev 1.6, Median 4, Min 1, Max 7*)
   (k) European Americans are sneaky (UM: *Mean 3.4, Std Dev 1.5, Median 4, Min 1, Max 7*) (UCLA: *Mean 4.0, Std Dev 1.6, Median 4, Min 1, Max 7*)
   (l) European Americans are competitive (UM: *Mean 4.8, Std Dev 1.3, Median 5, Min 1, Max 7*) (UCLA: *Mean 5.2, Std Dev 1.6, Median 5, Min 1, Max

7)

7. Generally speaking, would you say that people can be trusted or that you can't be too careful in dealing with people?
   (a) Always trusted (UM: *3.5%*) (UCLA: *1.7%*)
   (b) Usually trusted (UM: *69.8%*) (UCLA: *68.3%*)
   (c) Usually not trusted (UM: *24.1%*) (UCLA: *25.8%*)
   (d) Always not trusted (UM: *2.6%*) (UCLA: *4.2%*)
8. How many siblings do you have: ______________
   (UM: *Mean 1.3, Std Dev 1.3, Median 1, Min 0, Max 7*) (UCLA: *Mean 1.3, Std Dev 1.0, Median 1, Min 0, Max 6*)
9. How trusting are you?
   (a) Always trusting (UM: *16.4%*) (UCLA: *10%*)
   (b) Usually trusting (UM: *66.4%*) (UCLA: *65%*)
   (c) Usually not trusting (UM: *16.4%*) (UCLA: *25%*)
   (d) Always not trusting (UM: *0.9%*) (UCLA: *0%*)
10. There should be diversity programs to level the playing field for people from minority groups
    (a) Agree (UM: *73.3%*) (UCLA: *45%*)
    (b) Disagree (UM: *26.7%*) (UCLA: *55%*)
11. We should not allow special treatment based on race or gender. Merit should be the sole criteria
    (a) Agree (UM: *67.2%*) (UCLA: *79.2%*)
    (b) Disagree (UM: *32.8%*) (UCLA: *20.8%*)
12. Please write down the *Last *Name of your ten friends:
13. How strong is your University of Michigan (UCLA) school spirit?(UM: *Mean 5.3, Std Dev 1.8, Median 6, Min 1, Max 7*) (UCLA: *Mean 4.7, Std Dev 1.9, Median 5, Min 1, Max 7*)
14. During the experiment, how much did you pay attention to who your partner was? (UM: *Mean 3.0, Std Dev 2.0, Median 2, Min 1, Max 7*) (UCLA: *Mean 2.9, Std Dev 1.7, Median 3, Min 1, Max 7*)
15. During the experiment, I tried to maximize my own payoffs. (UM: *Mean 5.7, Std Dev 1.7, Median 6, Min 1, Max 7*) (UCLA: *Mean 5.5, Std Dev 1.6, Median 6, Min 1, Max 7*)
16. During the experiment, I tried to maximize joint payoffs. (UM: *Mean 3.7, Std Dev 1.9, Median 4, Min 1, Max 7*) (UCLA: *Mean 4.0, Std Dev 1.9, Median 4, Min 1, Max 7*)
17. For player As, during the experiment, if I chose A1 (the more generous option), I hoped player B would see it as a sign of trust and reciprocate.
    (a) Agree (UM: *56.7%*) (UCLA: *68.3%*)
    (b) Disagree (UM: *20.0%*) (UCLA: *13.3%*)
    (c) Not applicable as I never chose A1(UM: *23.3%*) (UCLA: *18.3%*)
18. For player Bs, during the experiment, if player A chose A1 (the more generous option), I felt I needed to reciprocate
    (a) Agree (UM: *38.3%*) (UCLA: *40.0%*)
    (b) Disagree (UM: *53.3%*) (UCLA: *53.3%*)

(c) Not applicable as A never chose A1(UM: *8.3%*) (UCLA: *6.7%*)

19. Do you know any participants in today's experiment
    (a) Yes (UM: *69.0%*) (UCLA: *19.2%*)
    (b) No (UM: *31.0%*) (UCLA: *80.8%*)
20. If so, please write down their last name: ______________
    (UM only)
21. What do you think is the ethnicity of the person with this name? (UM only)
    (a) Chen
        i. Asian
        ii. European
        iii. Other
            if it is other, please specify: ______________
        iv. I don't know

    (UM: *overall accuracy 91%; ingroup 85%; outgroup 97%.*)

Chapter 4

Crowdsourcing with All-Pay Auctions: a Field Experiment on Taskcn

4.1 Introduction

The Internet has transformed how work is done, from allowing geographically dispersed workers to collaborate to enabling tasks to be globally crowdsourced through public solicitation (Howe, 2006, 2008; Kleeman, Voss and Rieder, 2008). The term crowdsourcing typically refers to the open solicitation of effort on a well-defined task to a community (crowd) to obtain a submitted solution before a deadline. Crowdsoucing has become an increasingly popular choice for tasks, such as translation, programming, and website design. Due to the open nature of effort solicitation in crowdsourcing, it is important to understand how the incentives accompanying a task affect both participation and the quality of the output produced.

Historically, intrinsically interesting small tasks have been crowdsourced through voluntary contributions. A well-known example is the Audubon Society's Christmas Bird Count, the longest-running wildlife census. For 111 years, tens of thousands of birders and scientists have been recording the birds they encounter during the holiday season and contributing their data to the Audubon Society for scientific research. A more recent example is the ESP game, which uses an online coordination game to label large quantities of images on the Internet, generating high quality metadata which improve image search results while helping the visually impaired to navigate images on the Internet.[1]

In comparison, problems requiring substantial expertise and effort tend to be crowdsourced with monetary incentives. In such cases, a requester posts a problem

[1] Source: `http://www.gwap.com/gwap/gamesPreview/espgame/`, retrieved on October 10, 2011.

and indicates a respective reward. In some crowdsourcing designs, individuals can claim the task and complete it without direct competition. In other designs, the task is open to an unlimited number of participants, and one or several of the submissions are chosen as winners. A recent example of the latter is the 2011 NASA Planetary Data System Idea Challenge on TopCoder.com, which calls for ideas for potential applications to enable exploration and analysis of the scientific data from NASA planetary missions. This challenge awarded five top submission prizes within two weeks of the start of the challenge.[2]

To obtain the best quality submissions, crowdsourcing sites experiment with different incentive mechanisms. For instance, some crowdsourcing designs allow an individual to claim a task as her own, thus becoming a monopolist provider of the solution. An example is the now defunct question-answering site, Google Answers, which employed 500 pre-selected researchers to answer consumer questions. Once a researcher chose a question, she locked it and became the sole provider of an official answer. If the answer was satisfactory, she received the reward. Similarly, Amazon's Mechanical Turk allows workers to choose small human intelligence tasks (HITs). Once a HIT is chosen, the worker becomes the sole provider of a solution. Field experiments conducted on Google Answers find that, while a higher reward increases the likelihood that a question is answered, it has no effect on answer quality (Chen, Ho and Kim, 2010; Harper, Raban, Rafaeli and Konstan, 2008; Jeon, Kim and Chen, 2010). Similarly, on Amazon's Mechanical Turk, a higher reward is found to increase participation but to have no impact on quality (Mason and Watts, 2009).

In contrast to Google Answers and Mechanical Turk, some crowdsourcing sites, such as Taskcn in China and TopCoder in the United States, introduce a competitive element in the form of contests. In the simplest form of the contest, after a requester posts a task and reward, any user can submit a solution to the task. Each task may receive many submissions. Since every user who submits a solution expends effort regardless of whether or not she wins, this simplest form of contest mechanism is equivalent to a first-price all-pay auction, where everyone expends effort, but only the winner receives the reward. To our knowledge, no field experiment has been performed to understand the effect of the reward levels and reserve quality on either participation or submission quality in such a competitive setting.

In addition to allowing for competition, crowdsourcing sites experiment with other features of all-pay auctions. On Taskcn, for example, sequential all-pay auctions,

[2] Source: `http://community.topcoder.com/tc?module=ProjectDetail&pj=30016974`, retrieved on September 16, 2011.

where late entrants can observe the content of earlier submissions, used to be the only exchange mechanism. Recently, users were given the ability to password-protect their solutions.[3] Theoretically, if all users password-protect their solutions, a sequential all-pay auction is transformed into a simultaneous all-pay auction. On the other hand, if only a fraction of users password-protect their solutions, the contest becomes a hybrid sequential/simultaneous all-pay auction. In comparison, on TopCoder, every submission is sealed. Given the options available, an evaluation of the various design features in all-pay auction mechanisms can potentially inform and thus improve the design and outcome of crowdsourcing mechanisms.

To evaluate the effects of both reward size and early high-quality submission (i.e., a soft reserve) on overall participation and submission quality, we conduct a field experiment on Taskcn. In our experiment, we post translation and programming tasks on Taskcn. The tasks are of similar difficulty, but the reward is varied. In addition, for a subset of tasks, we submit a high quality solution early in the bidding period. Unlike earlier field experiments on Google Answers and Mechanical Turk, in the competitive setting of Taskcn, we find significant reward effects on participation and submission quality, which is consistent with our theoretical predictions. However, unpredicted by theory, we find that experienced users respond to our experimental treatments differently from inexperienced ones. Specifically, experienced users are more likely to enter auctions with a high reward than inexperienced users. Furthermore, they are less likely to enter an auction where a high quality solution is already posted. As a result, our reserve treatments result in significantly lower average submission quality than those without a reserve.

The rest of the paper is organized as follows. We first give an overview of Taskcn in Section 4.2. In Section 4.3, we review the crowdsourcing and all-pay auction literature. Section 5.3 presents a theoretical framework for sequential and simultaneous all-pay auctions. Section 5.4 presents the experimental design. Based on our theoretical framework and experimental design, we present our hypotheses in Section 4.6. Experimental results are presented in Section 4.7. Finally, in Section 4.8, we discuss our results and their design implications.

[3]There are two methods of protecting solution content utilized on Taskcn. One is to use a prepaid service provided by the site; the other is to submit the solution with password protection and send the password to the requester by email.

4.2 Taskcn: An Overview

Since the crowdsourcing site Taskcn (`http://www.taskcn.com/`) was founded in 2006, it has become one of the most popular online labor markets in China. On Taskcn, a requester first fills out an online request form with the task title, the reward amount(s), the closing date for submissions, and the number of submissions that will be selected as winners. When the closing date is reached, the site sends a notice to the requester who posts the task, asking her to select the best solution(s) among all the submissions. The requester can also choose the best solution(s) before the closing date. Once the task is closed, the winner receives 80% of the reward and the site retains 20% of the reward as a transaction fee. As of August 24, 2010 Taskcn had accumulated 39,371 tasks, with rewards totaling 27,924,800 CNY (about 4.1 million USD).[4] Of the 2,871,391 registered users on Taskcn, 243,418 have won rewards.

To inform our field experiment, we crawled and analyzed the full set of tasks posted on Taskcn from its inception in 2006 to March 2009. As of the time of our crawl, tasks were divided into 15 categories, including requests for graphic, logo and web designs; translations; business names and slogan suggestions; and computer coding. Challenging tasks, such as those involving graphic design and website building, have the highest average rewards (graphic design: 385 CNY; web building: 460 CNY) as they require higher levels of expertise, whereas tasks asking for translations, or names and slogan suggestions offer lower average rewards (translation: 136 CNY; name/slogan: 170 CNY). In addition, most tasks (76.5%) select a single submission to win the reward.

Within the site, each ongoing task displays continually updated information on the number of users who have registered for the task and the number of submissions. Unless protected, each solution can be viewed by all users. Taskcn started a solution protection program, which hides the content of one's submission from other users. To protect a submission, a user pays a fee for the program.[5] Password-protected submissions are displayed to the requester ahead of other submissions. Instead of paying for solution protection, many users on Taskcn protect their solution content by submitting an encrypted solution and sending the password to the requester.

Once on the site, after reading task specifications and submitted solutions, a user can decide whether to register for a task and submit a solution before the closing date. A user can also view the number of credits accrued by previous submitters. The num-

[4]The exchange rate between the US dollar and the Chinese yuan was 1 USD = 6.8 CNY in 2009 and 2010.

[5]The fee ranges from 90 CNY for three months to 300 CNY for a year.

ber of credits corresponds to the hundreds of CNY a user has won by competing in previous tasks, and may signal either expertise or likelihood of winning. Even after a user registers for a task, she may decide not to submit a solution. Furthermore, there is no filter to prevent low quality solutions.

Given Taskcn's design, it is of interest to understand how users respond to incentives induced by different design features. For example, does a higher reward induce more submissions and higher quality? Does the early entry of a high quality submission deter the submission of low quality solutions? What tasks are more likely to elicit password-protected solutions? Do experienced users respond to incentives differently from inexperienced users? We address each of these questions in our subsequent experimental design and data analysis.

4.3 Literature Review

Our research is closely related to two streams of literature: the empirical literature on crowdsourcing, and the theoretical and experimental literature on all-pay auctions. Thus, we discuss each area of research separately below.

4.3.1 The Crowdsourcing Literature

Spurred by the increasing number of both simple and complex tasks solicited and fulfilled online, research on crowdsourcing has emerged in several academic disciplines, including economics, sociology, and information and computer sciences. In this section, we focus on studies that specifically investigate incentive structures. In what follows, we organize the findings by the incentive structures provided by various crowdsourcing sites, in the order of monopoly, contest, and non-pecuniary incentive. A key question explored in this literature is what factors affect participation levels and overall solution quality.

We first examine the provider-as-a-monopolist incentive structure, where the solution provider becomes the sole provider for a request once she claims it. One example of such a site is Google Answers, which employed 500 researchers to provide answers to customer questions, each priced between \$2 and \$500. While Harper, Raban, Rafaeli and Konstan (2008) find that a higher reward leads to significantly better answer quality on Google Answers, Chen et al. (2010) find no such reward effects. To reconcile these findings, Jeon et al. (2010) conduct a meta-analysis using data from both studies. Using the Heckman correction for selection bias (Heckman, 1979), they find that, while a higher reward increases the likelihood that a question is answered, conditional on getting an answer, a higher reward has no effect on quality. In addition, they find that researcher reputation is the only significant predictor of answer quality. We refer the reader to Chen et al. (2010) for a comprehensive review of the literature related to incentive effects on answer quality in online question-answering communities.

With a similar monopolist incentive structure as Google Answers, the Amazon Mechanical Turk (MTurk) is designed to employ human labor to perform small human intelligence tasks (HITs) that computers are not yet good at, such as tagging images and describing products. Unlike Google Answers, these tasks are typically priced at less than a dollar per task. A field experiment conducted by Mason and Watts (2009) on MTurk finds that, while a higher reward leads to higher participation, it does not lead to better solution quality, consistent with the findings in Jeon et al. (2010). We conjecture that the absence of a reward effect on quality is due to the monopolist incentive structure, i.e., a lack of competition once a provider takes up a question or a task.

Compared to the monopolist incentive structure, a contest ensures competition after a provider takes up a task, where the competitiveness depends on the number of entrants as well as their expertise and efforts. The best-known crowdsourcing sites using contests include Taskcn and TopCoder. However, to our knowledge, existing studies of these two sites use naturally occurring field data.

As described in Section 4.2, on Taskcn, a requester posts a task to the site, along with a designated monetary reward amount and a deadline for submission. Users then submit their solutions. At the deadline, the requester chooses the winner, who receives 80% of the reward while the site receives 20%. In a study using data crawled from Taskcn, Yang, Adamic and Ackerman (2008*a*) find that, while a higher reward is correlated with more views ($\rho = 0.64$), the correlation between reward and the number of submissions is lower ($\rho = 0.43$). Importantly, using human coders for a random sample of 157 tasks, the authors find a positive and significant correlation between reward size and the skill requirements for the corresponding task, indicating that reward size is *endogenously* related to task difficulty. Therefore, to investigate the causality between reward and contestant behavior, it is important to *exogenously* vary the reward levels while controlling for task difficulty, which we do in this paper by conducting a randomized field experiment. Building on the empirical findings of Yang et al. (2008*a*), DiPalantino and Vojnovic (2009) construct a theoretical all-pay auction model for crowdsourcing (subsection 4.3.2). Using a subsample of Taskcn data, they find that participation rates are increasing with reward at a decreasing rate, consistent with their theoretical prediction. However, neither study addresses the issue of quality. Thus, our study contributes to the research on crowdsourcing by investigating both participation and solution quality using a randomized field experiment.

Another well-known contest-based crowdsourcing site, TopCoder.com, is the largest competitive software development community in the world. Unlike Taskcn where the sequential all-pay auction is the prevalent mechanism, TopCoder uses simultaneous all-pay auctions, where a contestant cannot observe the content of others' submissions before submitting her own solution, although both the identities and ratings of other entrants are known. Using historical data from the TopCoder website, Archak (2010) finds that reward is a significant determinant of solution quality. Furthermore, Archak finds that highly-rated contestants tend to sign up early in the registration phase to deter the entry of other contestants. In an empirical analysis of the effects of competition, Boudreau, Lacetera and Lakhani (forthcoming) find that, while the average solution quality decreases with a larger number of competitors for easier problems, greater competition increases solution quality for more challenging tasks.

In addition to monetary incentives, crowdsourcing sites may also use non-pecuniary incentives. The most prevalent non-pecuniary incentives are reputation systems. While a contestant on Taskcn earns one credit point for every 100 CNY she wins on the site, a contestant on TopCoder receives an average grade from three expert reviewers for each submission. On Google Answers, each requester rates the official answer to her question on a 1-5 star scale. Each answerer's overall rating is public information.

On other sites, including Yahoo! Answers (Adamic, Zhang, Bakshy and Ackerman, 2008), Naver Knowledge-In (Nam, Adamic and Ackerman, 2009), and Baidu Knows (Yang and Wei, 2009), each asker selects the best answer for her question and each site provides a system of status levels, awarding higher status to more active and valued contributors. Another form of non-pecuniary incentive is virtual currency. On Yahoo! Answers, a fixed number of points are awarded to a given activity, such as answering a question. On Baidu Knows and Knowledge-In, a requester may transfer points to the provider of the best answer to her question. In a field experiment comparing answer quality across different Q&A sites, Harper et al. (2008) find that Google Answers generates answers of higher quality than sites relying exclusively on non-pecuniary incentives.

Compared to the studies reviewed above, our study represents the first randomized field experiment on an all-pay auction crowdsourcing site. By exogenously varying reward levels and the presence of a soft reserve, our method enables us to more precisely evaluate the reward and reserve effects on participation and solution quality, while preserving the realism of a natural field setting (Harrison and List, 2004).

4.3.2 All-pay Auction Literature

As Taskcn uses all-pay auctions as its contest mechanism, we review the theoretical and experimental literature on all-pay auctions in this subsection. In the economics literature, the first-price all-pay auction is often used to model rent-seeking, R&D races, lobbying, and tournaments. While most of this literature focuses on simultaneous all-pay auctions where players submit their bids without knowing others' bids, there is also a body of literature investigating properties of sequential all-pay auctions. Table 4.1 summarizes the theoretical and experimental studies, organized by the timing of bids and the relevant information structures.

Table 4.1 All-Pay Auction Literature: Theoretical Studies and Laboratory Experiments

	Simultaneous All-Pay Auctions	
	Theory	Laboratory Experiments
Complete Information	Baye et al. (1996) Bertoletti (2010) Anderson et al. (1998)	Potters et al. (1998) Davis and Reilly (1998) Gneezy and Smorodinsky (2006) Lugovskyy et al. (2010) Liu (2011)
Incomplete Information	Amann and Leininger (1996) Krishna and Morgan (1997) Fibich et al. (2006) DiPalantino and Vojnovic (2009)	Noussair and Silver (2006)
	Sequential All-Pay Auctions	
	Theory	Laboratory Experiments
Complete Information	Konrad and Leininger (2007)	Liu (2011)
Incomplete Information	Segev and Sela (2011)	

Within this area of research, Baye, Kovenock and de Vries (1996) provide a theoretical characterization of the mixed strategy Nash equilibrium for a simultaneous all-pay auction under complete information. Bertoletti (2010) extends this model to investigate the role of a reserve price and finds that a strict reserve price increases allocation efficiency. In an incomplete information setting, both Krishna and Morgan (1997) and Amann and Leininger (1996) characterize the symmetric Bayesian Nash equilibrium.[6] While the previous studies focus on a single auction, DiPalantino and Vojnovic (2009) investigate a multiple all-pay auction model, where contestants choose between tasks with different rewards. In particular, DiPalantino and Vojnovic (2009) show that a higher reward increases participation levels. However, they do not examine the effect of reward on submission quality.

A number of laboratory experiments test the predictions of simultaneous all-pay auction models (Table 4.1, right column). Under complete information, most studies find that players overbid relative to the risk neutral Nash equilibrium predictions in early rounds but learn to reduce their bids with experience (Davis and Reilly, 1998; Gneezy and Smorodinsky, 2006; Liu, 2011). An exception to this finding is Potters, de Vries and van Winden (1998), who find bidding behavior consistent with Nash equilibrium predictions.[7] Rent overdissipation as a result of overbidding can be (partially) explained by a logit equilibrium (Anderson, Goeree and Holt, 1998). In comparison, in an incomplete information and independent private value environment, Noussair and Silver (2006) find that revenue exceeds the risk neutral Bayesian Nash equilibrium prediction, due to aggressive bidding by players with high valuations and passive bidding by those with low valuations. Both overbidding and behavioral heterogeneity among different types of players are consistent with risk aversion (Fibich, Gavious and Sela, 2006).

Compared to research on simultaneous all-pay auctions, fewer studies investigate sequential all-pay auctions. Relevant to our study, in a complete information sequential all-pay auction model with endogenous entry, Konrad and Leininger (2007) characterize the subgame perfect Nash equilibrium, where players with the lowest bidding cost enter late, while others randomize between entering early and late. Extending this work to an incomplete information sequential all-pay auction setting, Segev and Sela (2011) demonstrate that giving a head start to preceding players can improve contestant effort. In a laboratory test of the Konrad and Leininger

[6] Krishna and Morgan's model assumes that everyone's value for the object is randomly drawn from the same distribution, whereas Amann and Leininger (1996) prove the existence and uniqueness of a Bayesian Nash equilibrium in a two-player incomplete information all-pay auction with an asymmetric value distribution.

[7] The combination of several design features might explain the results in Potters et al. (1998), including small group size ($n = 2$), stranger matching and more periods (30).

(2007) model, Liu (2011) finds that players learn to enter late with experience in all treatments.

In addition to the above, there is also a growing literature comparing all-pay auctions with other mechanisms in the fundraising context, which has a public good component, differentiating it from our study. We refer the reader to Carpenter, Matthews and Schirm (2010) for a summary of this literature and the references therein.

Compared to the existing literature on all-pay auctions, we conduct a field experiment on Taskcn, where features of sequential and simultaneous all-pay auctions coexist. As such, our results have the potential to inform the design of all-pay auctions for crowdsourcing sites.

4.4 Theoretical Framework

In this section, we outline our theoretical framework to derive comparative statics results which will serve as the basis for our experimental design and hypotheses. In doing so, we follow the model in Segev and Sela (2011), and extend their results to incorporate the effects of a reward and a reserve price on bidding strategies in sequential and simultaneous all-pay auctions.

In our model, there is a single task to be crowdsourced through an all-pay auction. The reward for the task is $v \geq 1$. There are n users, each differing in ability. Let $a_i \geq 0$ be user i's ability, which is her private information. User abilities are i.i.d. draws from the interval [0,1] according to the cumulative distribution function, $F(x)$, which is common knowledge. Additional assumptions on $F(x)$ will be introduced in subsections 4.4.1 and 4.4.2 respectively. The user with the best quality solution wins the reward, while other users incur time and effort in preparing their solutions.

To examine the effects of a reserve quality, i.e., a threshold solution representing the minimum acceptable quality, on participation levels and submission quality, we include a reserve quality, $q_0 \geq 0$. In this case, user i wins a reward equal to v if and only if the quality of her submission is the highest among the submissions and is at least as high as the reserve, i.e., $q_i \geq \max\{q_j, q_0\}$, $\forall j \neq i$. For technical reasons, we assume that ties are broken in favor of the late entrant.[8] For user i, a submission of quality q_i costs q_i/a_i, indicating that it is less costly for a high ability user to submit the same quality solution than a low ability user. In what follows, we separately

[8] This is a technical assumption to derive strict subgame perfect equilibria instead of ϵ-equilibria.

characterize the comparative statics results in the sequential and simultaneous all-pay auctions under incomplete information. All proofs and examples are relegated to Appendix A.

4.4.1 Sequential All-pay Auctions under Incomplete Information

Without allowing for password protection of solutions, the competitive process on Taskcn approximates a sequential all-pay auction, where solutions are submitted sequentially and the best solution is selected as the winner. Following Segev and Sela (2011), we first characterize the subgame perfect equilibria of a sequential all-pay auction under incomplete information. For theoretical results presented in this subsection, we need the additional assumption that $F(x) = x^c$, $0 < c < 1$, as in Segev and Sela (2011).[9]

In a sequential auction, each of n users enters the auction sequentially. In period i where $1 \leq i \leq n$, user i submits a solution with quality, $q_i \geq 0$, after observing previous submissions. Using backward induction, we characterize the equilibrium bidding functions of users n through 1, to derive the following comparative statics.

Proposition 1 (Reward Effect on Participation Level)**.** *In a sequential all-pay auction under incomplete information, a higher reward weakly increases the likelihood that user i submits a solution of positive quality.*

Proposition 1 indicates that we expect reward size to have a non-negative effect on user participation. Intuitively, a user's likelihood of participation *ex ante* depends on the reward size and the highest quality submissions before hers. When the reward size increases, the highest quality among earlier submissions also increases. With a zero reserve and risk neutrality, these two effects cancel each other. In comparison, with a positive reserve, the reward effect on participation dominates that from the increase of the highest quality among earlier submissions, resulting in a strict increase in a user's likelihood of participation.

A requester's satisfaction with the auction outcome typically depends not only on the quantity of submissions, but more importantly, on the quality.

[9] Due to the complexity of player i's winning function, closed-form bidding functions for the general ability distribution function have not been obtained (Segev and Sela, 2011).

Proposition 2 (Reward Effect on Expected Submission Quality)**.** *In a sequential all-pay auction under incomplete information, a higher reward increases user i's expected submission quality.*

Proposition 2 indicates that we expect reward size to have a positive effect on the expected submission quality. In Appendix A, we present a two-player example (Example 1) with closed-form solutions for the quality and likelihood of submissions, as well as the average and highest quality.

We now examine the effect of a positive reserve on participation. The following proposition parallels the equivalent reserve price effect on participation in winner-pay auctions where a positive reserve price excludes bidders with low values (Krishna, 2009).

Proposition 3 (Reserve Effect on Participation Level)**.** *In a sequential all-pay auction under incomplete information, a higher reserve quality decreases the likelihood that a user submits a solution with positive quality.*

Intuitively, the higher the reserve quality, the less likely it is that a user with a low ability will participate in the auction. In Appendix A, we present Example 2, a continuation of Example 1, to demonstrate the relevant comparative statics with respect to reserve quality.

As we do not have a general solution for the optimal reserve quality, we present a numerical example to illustrate the effects of reserve quality on the expected highest and average quality. Figure 4.1 presents the expected highest quality (left panel) and average quality (right panel) as a function of the reserve when $F(x) = x^c$, $c = 0.5$, $v = 1$ and $n = 2$. In this example, the reserve quality which maximizes the expected highest quality is 0.47, whereas the one which maximizes the expected average quality is 0.43.

The optimal reserve quality in this example is in the middle of the quality range. Intuitively, an appropriate reserve should exclude users with low abilities and thus increase the expected highest and average quality.

4.4.2 Simultaneous All-pay Auctions under Incomplete Information

In this subsection, we investigate the extreme case when all solutions to a task are submitted under password protection. In this scenario, the competitive process is approximated by a simultaneous all-pay auction, where users do not see others' solutions

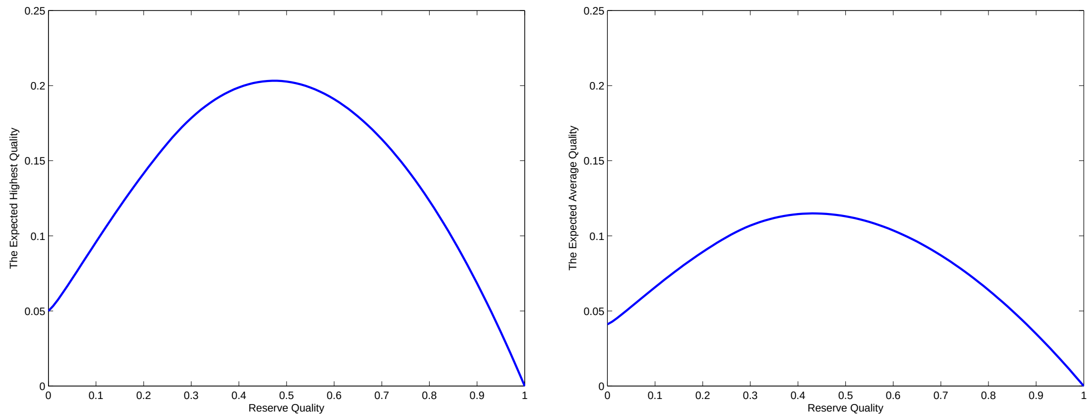


Figure 4.1 Effects of Reserve Quality on the Expected Highest and Average Quality: $c = 0.5$; $v = 1$; $n = 2$

before submitting their own. The crowdsourcing process on TopCoder is an example of a simultaneous all-pay auction. We derive comparative statics for simultaneous all-pay auctions under incomplete information to examine the effects of reward size and reserve quality.

In a simultaneous all-pay auction, each user submits her solution without observing those of others. Each user's ability is again an i.i.d. draw from the cumulative distribution function $F(x)$ with support $[0, 1]$. To prove Propositions 4 through 6, we need the additional assumption that $H_i(x) = \prod_{j \neq i}^{n} F(x)$ is strictly concave and $H_i(0) = 0$. However, the assumption that $F(x) = x^c$ is not necessary for results in this subsection. We now state three propositions. The first two propositions examine the reward effects on participation level and submission quality, respectively.

Proposition 4 (Reward Effect on Participation Level)**.** *In a simultaneous all-pay auction under incomplete information, a higher reward weakly increases the likelihood that user i submits a solution of positive quality.*

Proposition 5 (Reward Effect on Expected Submission Quality)**.** *In a simultaneous all-pay auction under incomplete information, a higher reward increases the expected submission quality.*

We now state the reserve effect on participation level.

Proposition 6 (Reserve Effect on Participation Level)**.** *In a simultaneous all-pay auction under incomplete information, a higher reserve decreases participation.*

Unlike the sequential case, since every user in a simultaneous all-pay auction is symmetric *ex ante*, the reserve which maximizes the expected highest quality is the same as that which maximizes the expected average quality. We now present two numerical examples to illustrate the effects of reserve quality on the expected quality for each player in a simultaneous all-pay auction. The left panel in Figure 4.2 presents the expected quality for each player when $c = 0.2$, $v = 1$ and $n = 2$ and the optimal reserve quality is $q_0 = 0.4$. The right panel in Figure 4.2 presents the expected quality for each player when $c = 0.8$, $v = 1$ and $n = 2$ and the optimal reserve quality is $q_0 = 0$. The examples suggest that, in a simultaneous all-pay auction, the effect of a positive reserve on expected quality depends on the distribution of abilities. Thus, requesters might not always be better off by setting a positive reserve.

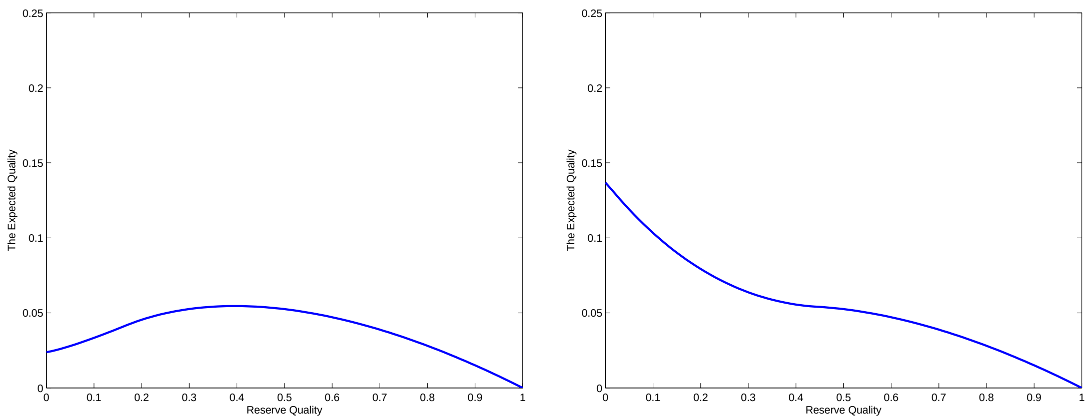


Figure 4.2 Effects of Reserve Quality on the Expected Quality: $c = 0.2$ (left), $c = 0.8$ (right); $v = 1$; $n = 2$

In this section, we separately characterize the reward and reserve effects on participation and submission quality under sequential and simultaneous all-pay auctions, respectively. We find that reward and reserve quality have similar effects on both participation and quality under each auction format. While these characterizations provide benchmarks for our experimental design and hypotheses, in reality, most all-pay auctions on Taskcn are hybrid sequential/simultaneous auctions, where participants endogenously determine whether to password protect their solutions. Two other features of the field not captured by our theoretical models are endogenous entry timing and the choice among multiple auctions, each of which is modeled by Konrad and Leininger (2007) and DiPalantino and Vojnovic (2009), respectively. A more realistic model which incorporates endogenous selection of the auction format, endogenous entry and choice among multiple auctions is left for future work. Nonetheless, our experiment provides a useful framework with which to study the effect of reward level and reserve presence on participation and submission quality.

4.5 Experimental Design

In this section, we outline our experimental design. We use a 2×3 factorial design to investigate the reward and reserve quality effects on user behavior on Taskcn. Specifically, we investigate whether tasks with a higher reward will attract more submissions and generate solutions of a higher quality. We are also interested in determining whether a high-quality solution posted early, playing the role of a soft reserve, will deter the entry of low quality solutions, especially if it is posted by a user with a history of winning on the site.

4.5.1 Task Selection: Translation and Programming

Of the task categories on Taskcn, we choose to use translation and programming tasks for our field experiment, as the nature of the respective solutions is fairly standard and objective.

Our translation tasks fall into two categories: personal statements collected from Chinese graduate students at the University of Michigan and company introductions downloaded from Chinese websites. We choose these two categories as they are challenging, each requiring a high level of language skills and effort compared to translating other types of documents, such as resumes. In Appendix B, we provide an example of a personal statement and an example of a company introduction, as well as a complete list of Taskcn IDs and URLs for all translation tasks used in our experiment.

For the programming tasks, we construct 28 programming problems, including 14 Javascript and 14 Perl tasks. None of our programming tasks is searchable and each has practical use. A complete list of the programming tasks is provided in Appendix B. One example of such a task reads: "Website needs a password security checking function. Show input characters as encoded dots when user types password. Generate an information bar to indicate the security level of the password, considering these factors: (1) length of the password; (2) mixture of numbers and characters; (3) mixture of upper and lower case letters; (4) mixture of other symbols. Please provide source code and html for testing." The functionality of such programming tasks can be assessed by qualified programmers.

To prepare for our field experiment, we crawled all the tasks on Taskcn posted from its inception in 2006 to March 27, 2009. Table 4.2 presents summary statistics for these two types of tasks. While translation and programming tasks have the

Table 4.2 Summary Statistics about Tasks on Taskcn from 2006 to March 27, 2009

	Reward (in CNY)			# of Submissions		
	Mean	Median	SD	Mean	Median	SD
Translation	137	100	164	109	42	163
Programming	176	100	378	10	6	17

same median reward on the site, the former generate a higher median number of submissions, possibly due to the ability to submit a machine-generated solution.

4.5.2 Treatments

Our parameter choices are based on the summary statistics in Table 4.2. To investigate the reward effects, we choose two reward levels for our tasks, 100 CNY and 300 CNY, based on the following considerations. First, using the median reward for our low reward treatments guarantees a certain amount of participation. Second, the two reward levels have a monetarily salient difference.

As translation tasks posted by Taskcn users have a relatively large number of submissions on Taskcn (Table 4.2), we investigate whether the early entry of a high quality submission can influence participation, similar to the effect of a reserve price in an auction. Thus, for each reward level, we vary the reserve conditions, including No-Reserve, Reserve-without-Credit, and Reserve-with-Credit.[10] The two reserve conditions differ in whether the user posting the high quality solution has credits from previous wins. In the Reserve-without-Credit treatments, each early submission is posted by a user without a winning history on the site, whereas in the Reserve-with-Credit treatments, our submissions are posted by a user with four credits. To ensure the quality of the translations used in the reserve treatments, we ask a bilingual student (the owner of the personal statement when applicable) to provide the first round of English translations, and then a native English speaker to polish them.

Table 4.3 summarizes our six treatments. The number in brackets indicates the number of tasks posted in a treatment. A total of 120 translation (28 programming) tasks are randomly assigned to six (two) treatments. Thus the full 2×3 factorial design is applied to translation tasks, while programming tasks are used to check for the robustness of any reward effects. We use a greater number of translation tasks

[10] Recall that users earn 1 credit whenever they earn 100 CNY on the site. We created our own user account and obtained winning credits by winning tasks before the launch of our experiment.

Table 4.3 Number of Tasks by Experimental Treatment

	No-Reserve	Reserve-without-Credit	Reserve-with-Credit
Low-Reward (100 CNY)	Programming (14) Translation (20)	Translation (20)	Translation (20)
High-Reward (300 CNY)	Programming (14) Translation (20)	Translation (20)	Translation (20)

in the field experiment in part because of the relative difficulty in generating unique, plausible, and comparable programming tasks.

4.5.3 Experimental Procedure

We posted 148 tasks on Taskcn between June 3 and 22, 2009, eight tasks per day (one translation and one programming task from each treatment) so as not to drastically increase the total number of tasks posted daily on the site.[11]

Each task was posted for seven days, with one reward per task. To avoid reputation effects from the requester side, we created a new user account for each task. After a task was posted, any user could participate and submit a solution within seven days. At the end of the seven-day period, we selected a winner for each task, excluding our reserve submissions.[12]

Table 4.4 Summary Statistics for User Credits

	Mean	Median	Min	Max	Standard Deviation
Translation	0.43	0	0	96	4
Programming	4	0	0	62	11

During our experiment, 948 users participated in the translation tasks, submitting a total of 3751 solutions, and 82 users participated in the programming tasks, submitting a total of 134 solutions. Table 4.4 presents the summary statistics of user credits among our participants. In addition to the number of submissions, participants also vary in their password protection behavior between these two types of tasks. We find that 8% of the translation and 53% of the programming solutions are submitted

[11]From January to March 2009, the average number of new tasks posted on the site per day is 12. Since each task is open between one week to a month, and all open tasks are listed together, users select from dozens to hundreds of tasks at a time.

[12]We find that the average quality of the winning solutions (4.33) is not significantly different from that of our reserve submissions (4.36) based on the evaluation of raters blind to the research design and hypotheses ($p = 0.423$, one-sided paired t-test).

with password protection. The difference in the proportion of password-protected submissions is significant ($p < 0.01$, test of proportion, two-sided).

4.5.4 Rating Procedure

To determine submission quality, we recruited raters blind to the research hypotheses to evaluate each submission. All raters were graduate students from the University of Michigan. Our rating procedures follow the standard practice in content analysis (Krippendorff, 2003). To evaluate the translation submissions, we proceeded in two stages. First, we recruited three bilingual Chinese students to independently judge whether a submission was machine-translated. If two of them agreed that a submission was machine-translated, we categorized it as a machine translation. Second, we recruited nine different bilingual Chinese students, whom we randomly assigned into three rating groups. For this stage, all valid translations plus one randomly-selected machine translation for each task were independently evaluated by three raters.[13] Raters for translation tasks each had scored above 600 on the TOEFL. To evaluate the programming submissions, we recruited three Chinese students, each with an undergraduate degree in computer science and several years of web programming experience. We conducted training and rating sessions for all raters. Raters within each rating group independently evaluated the same set of task-submission pairs. Details of the rating procedures and instructions can be found in Appendix C.

Table 4.5 Rating Task Quantities and Inter-rater Reliabilities (ICC[3,3])

	Group	# Tasks	# Submissions	Task Difficulty	Submission Quality
Translation	1	43	265	0.62	0.90
	2	35	215	0.88	0.88
	3	42	284	0.72	0.68
Programming	1	28	108	0.55	0.77

From October 2009 to February 2010, we conducted 45 rating sessions at the University of Michigan School of Information Laboratory. Each session lasted no more than two hours. Students were paid a flat fee of $15 per hour to compensate them for their time. We used intra-class correlation coefficients, ICC[3,3], to measure inter-rater reliability.

[13]Note that the machine translations were not marked in the second stage. Thus, this procedure provides an additional consistency check among raters.

Table 4.5 presents the number of rating tasks and the inter-rater reliability for each rating group. The last two columns present the inter-rater reliability for each rating group. Good to excellent reliability is observed for all rating groups.[14] Additionally, machine translations are rated as having significantly lower quality than other valid translations in the second stage,[15] providing further evidence of rating consistency between the first- and second-stage raters. In our subsequent analysis, we use the median evaluation for the task difficulty and overall submission quality.[16]

4.6 Hypotheses

In this section, we describe our hypotheses comparing user behavior in different treatments based on the theoretical predictions outlined in Section 5.3. We are interested in two outcome measures: participation and submission quality.

Based on Propositions 1 and 4, we expect that a task with a higher reward will receive more participation.

Hypothesis 9 (Reward Effect on Participation)**.** *A task with a high reward attracts more submissions than a task with a low reward.*

We now discuss the reserve effects on participation. Based on Propositions 3 and 6, we predict that the early entry of a high quality solution will decrease overall participation. Even though our reserve is not binding, we predict that users who cannot produce a translation with a higher quality will not participate. Thus, we expect to observe less participation in the reserve treatments compared to the no-reserve treatments. This effect should be stronger for the reserve-with-credit treatments.

Hypothesis 10 (Reserve Effect on Participation)**.** *The number of submissions in the reserve treatments will be less than that in the no-reserve treatments.*

For submission quality, based on Propositions 2 and 5, we expect that a task with a higher reward will attract higher quality submissions.

[14] In general, values above 0.75 represent excellent reliability, values between 0.40 and 0.75 represent fair to good reliability, and values below 0.40 represent poor reliability.

[15] On a 1-7 Likert scale, the average median quality of machine and valid translations is 2 and 5, respectively, significantly different from each other ($p < 0.01$, using ordered probit regressions with standard errors clustered at the task level).

[16] Task difficulty is measured by the median evaluation for questions 1(d) in translation and 1(b) in programming, whereas overall submission quality is measured by the median evaluation for questions 3 in translation and 2(d) in programming. See Appendix C for rating instructions.

Hypothesis 11 (Reward Effect on Submission Quality). *A task with a high reward attracts submissions of higher quality than a task with a low reward.*

Lastly, despite a lack of theoretical characterizations of the reserve effect on *expected* submission quality, we formulate a hypothesis based on *ex post* submission quality. We expect that the average submission quality in the reserve treatments will be higher than that in the no-reserve treatments, since only users who can generate a solution better than the reserve are expected to participate.

Hypothesis 12 (Reserve Effect on Submission Quality). *The average submission quality will be higher in the reserve treatments than in the no-reserve treatments.*

4.7 Results

Of the 120 translation and 28 programming tasks posted, we received submissions for every task. On average, each translation (programming) task received 1830 (1211) views, 46 (9) registrations and 31 (5) submissions. Although it might at first appear that participation is several times greater for translation tasks relative to programming tasks, most of the submissions for translation tasks are machine-generated. The average number of valid translations per task (5) is equal to that of solutions to programming tasks. Of the submissions received, 8% (53%) of the translation (resp. programming) solutions are password protected, making them hybrid sequential/simultaneous all-pay auctions.

A total of 948 (82) unique users participate in the translation (programming) tasks.[17] We categorize the participants based on their prior winning experience. We define *experienced users* as those who have won at least 100 CNY (with at least one reputation credit) prior to our experiment, whereas we define *inexperienced users* as those who have not. Table 4.6 reports the summary statistics of participant credits.[18] Specifically, we find that 4% (27%) of the participants in the translation (programming) tasks are experienced users.

Before analyzing our results, we first check that our randomization of tasks across treatments works. Pairwise Kolmogorov-Smirnov tests comparing task difficulty across treatments yield $p > 0.10$ for both translation and programming tasks, indi-

[17]We treat each unique ID as a unique user, as the reputation system on the site encourages users to keep a single identity across tasks.

[18]These summary statistics are computed based on field data from Taskcn from 2006 through June 2, 2009, the day before our experiment.

Table 4.6 The Percentage of Each User Type in the Experiment

Task		Number of Users	Percentage	Median Credit	Mean Credit
Translation	Experienced Users	42	4	3	10
	Inexperienced Users	906	96	0	0
Programming	Experienced Users	22	27	5	10
	Inexperienced Users	60	73	0	0

cating that task difficulty is comparable across different treatments. In what follows, we evaluate the specific treatment effects on participation and submission quality.

We first examine whether different reward levels affect participation. Due to the existence of machine translations and solutions copied from others, we separately examine the effect of reward level on both the total number of translation submissions and that of valid translations. To qualify for a valid translation, a submission must be neither machine-translated nor copied from previous submissions. Similarly, we separate programming submissions into valid and invalid solutions. Of the 134 programming submissions, we find that 26 are invalid due to either incompleteness or copying from previous submissions. In both types of tasks, valid solutions involve a certain amount of effort in the preparation process, while invalid ones involve minimum effort. In our separate analyses, we find no significant difference between the reserve-with-credit and reserve-without-credit treatments in their effect on either participation or submission quality (participation: $p > 0.10$, two-sample t-tests; quality: $p > 0.10$, ordered probit regressions with standard errors clustered at the task level). Therefore, in subsequent analyses, we pool the two treatments together as a single reserve treatment.

Figure 4.3 presents the reward effect on participation in both the translation (top panel) and programming tasks (bottom panel). For each type of task, we present the participation data for all submissions and valid submissions separately. The average number of submissions and standard errors for the high- and low-reward treatments are presented in each graph. We summarize the results below.

Result 12 (Reward Effect on Participation)**.** *Translation (programming) tasks in the high-reward treatments receive significantly more submissions compared to those in*

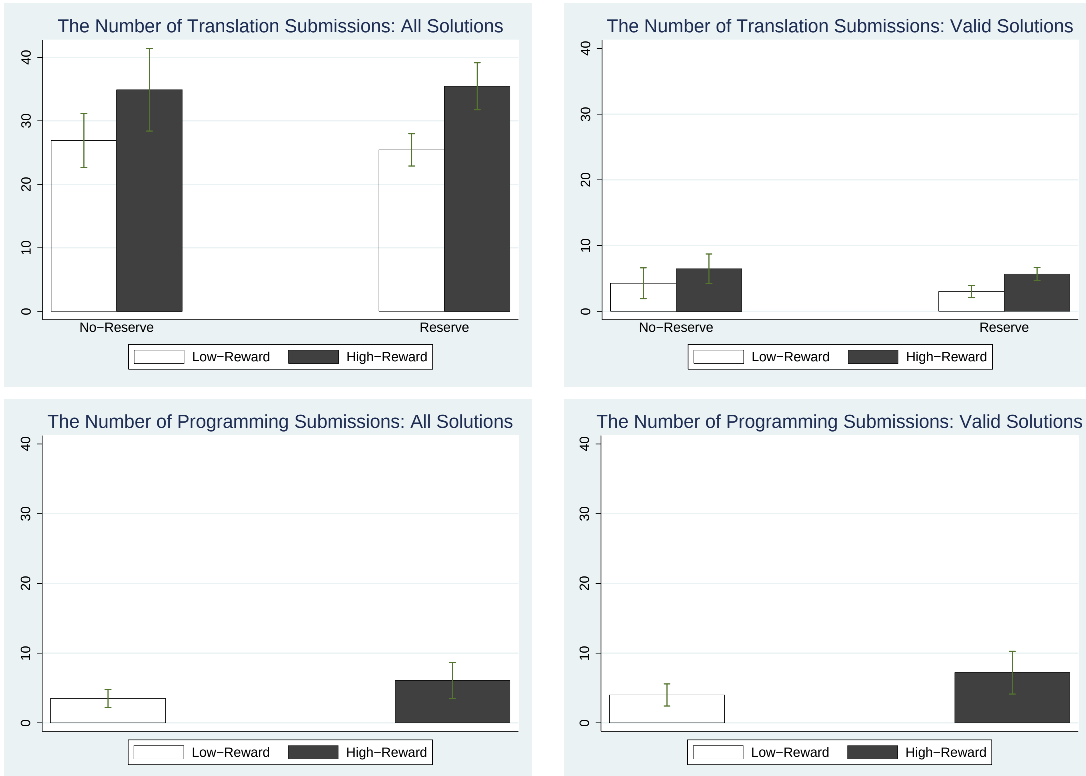


Figure 4.3 Reward Effect on Participation Level

the low-reward treatments.

Support. *Table 4.7 presents the summary statistics and treatment effects for both the translation and programming tasks. Specifically, we find that the average number of translation submissions per task is significantly higher in the high-reward than in the low-reward treatments (no-reserve: $p = 0.019$; reserve: $p < 0.01$, one-sided two-sample t-tests). Furthermore, this difference is (weakly) significant for the subset of valid translations (no-reserve: $p = 0.081$; reserve: $p < 0.01$, one-sided two-sample t-tests). For programming tasks, one-sided permutation tests yield $p = 0.037$ for all submissions and $p = 0.031$ for valid submissions. Note that non-parametric tests are used for programming tasks due to the small number of tasks in each treatment.*

By Result 12, we reject the null hypothesis in favor of Hypothesis 9, that a higher reward induces more submissions. This result is consistent with our theoretical predictions in Propositions 1 and 4, as well as empirical findings on Taskcn (DiPalantino and Vojnovic, 2009).

We now analyze the reserve effects on participation. Recall that Proposition 3 (6) predicts fewer submissions in the reserve treatments than in the no-reserve treatments

Table 4.7 Treatment Effects on the Average Number of Submissions Per Task

All Solutions	Translation			Programming	
	No-Reserve	Reserve	Reserve Effect		All
High-Reward	35	35	$p = 0.436$	High-Reward	6
Low-Reward	27	25	$p = 0.260$	Low-Reward	4
Reward Effect	$p = 0.019$	$p = 0.000$		Reward Effect	$p = 0.037$
Valid Solutions	**Translation**			**Programming**	
	No-Reserve	Reserve	Reserve Effect		All
High-Reward	6	6	$p = 0.251$	High-Reward	7
Low-Reward	4	3	$p = 0.153$	Low-Reward	4
Reward Effect	$p = 0.081$	$p = 0.000$		Reward Effect	$p = 0.031$

in sequential (simultaneous) all-pay auctions. Interestingly, we find no difference in the number of submissions between the reserve and no-reserve treatments (Table 4.7, column 4). Thus, we fail to reject the null hypothesis in favor of Hypothesis 10.

Summarizing all treatments, Table 4.8 reports OLS regression analysis to enable a comparison of the relative effectiveness of the different treatments on participation in translation tasks. The dependent variables are the number of submissions for all solutions (1) and valid solutions (2). Independent variables include the following variables (with omitted variables in parentheses): high-reward (low-reward), reserve (no-reserve), and task difficulty. From Table 4.8, we see that the coefficient of the high-reward dummy is positive and significant at the 1% level in both (1) and (2), indicating a robust reward effect on participation when we control for other factors. Specifically, from low-reward to high-reward tasks, the average number of submissions increases by 10 (3) for all (valid) solutions. The coefficient for task difficulty is negative and significant, indicating that more difficult tasks receive fewer submissions.

In addition to participation, we are also interested in factors affecting the quality of valid submissions. Two outcome measures are used to evaluate quality: the quality of all valid submissions and the quality of the best solution for each task. For some tasks, e.g. programming, only the quality of the best solution may matter. However, in modularizeable tasks, the requester might care about the average quality of the submitted solutions. One example is translation, where different translations may be combined at the sentence or paragraph level. Thus, we examine the reward effect on

Table 4.8 OLS: Determinants of the Number of Submissions in Translation Tasks

Dependent Variable	# of Submissions (All)	# of Submissions (Valid)
	(1)	(2)
High-Reward	9.749***	2.626***
	(1.859)	(0.671)
Reserve	-1.511	-1.328*
	(1.996)	(0.717)
Task Difficulty	-2.995***	-0.840**
	(0.970)	(0.349)
Constant	38.90***	7.681***
	(4.505)	(1.626)
Observations	120	112
R^2	0.232	0.167

Notes:
1. Standard errors are in parentheses.
2. Significant at: * 10%; ** 5%; *** 1%.

both the average and the highest submission quality.

Table 4.9 Ordered Probit: Determinants of Submission Quality for Valid and Best Translations

Dependent Variable	Quality of Valid Translations		Quality of Best Translations	
	(1)	(2)	(3)	(4)
High Reward	0.312**	-0.004	0.277	-0.152
	(0.146)	(0.187)	(0.255)	(0.418)
Reserve	-0.596***	-1.127***	-0.962***	-1.665***
	(0.153)	(0.200)	(0.263)	(0.393)
Task Difficulty	0.138**	0.169**	0.200	0.295*
	(0.0645)	(0.074)	(0.142)	(0.172)
User Fixed Effects	No	Yes	No	Yes
Observations	533	533	178	178
R^2	0.038	0.387	0.087	0.465

Notes:
1. Robust standard errors in parentheses are clustered at the task level in (1) and (3).
2. Significant at: * 10%; ** 5%; *** 1%.

Table 4.9 presents four ordered probit specifications, which investigate factors affecting submission quality for all valid translations and best translations. The dependent variables are the quality of valid translations (specifications 1 and 2) and that of best translations (specifications 3 and 4), while the independent variables include the following variables (with omitted variables in parentheses): high reward (low reward), reserve (no-reserve), and task difficulty. Specifications (1) and (3) report pooled models with standard errors clustered at the task level. We find that the coefficient of the high-reward dummy is positive and significant in (1), indicating a reward effect on the average submission quality, whereas the same coefficient is positive but insignificant in (3), indicating the absence of a reward effect on the quality of the best translations. In comparison, the coefficient of the reserve dummy is negative and significant in both specifications, indicating a negative reserve effect on quality when we control for other factors. The coefficient of task difficulty is positive and significant at the 5% level in (1), indicating that solutions to more difficult tasks are more likely to get higher quality ratings. As 43% (38%) of the users who submit a valid (best) solution participate in more than one task, we then report fixed effects models in specifications (2) and (4) to investigate whether the estimation in the pooled model is driven by the within-user variation in the submission quality over tasks. In the fixed effects model, we fail to find significant reward effect on submission quality within each user. However, the reserve dummy remains negative and significant, indicating that each user produces lower submission quality for tasks with a reserve compared to those without a reserve. We summarize the results below.

Result 13 (Reward Effect on Submission Quality)**.** *The average quality of valid translations is significantly higher in the high-reward treatments than in the low-reward treatments. Furthermore, this effect is not driven by within-user variations over different reward levels.*

Support. *Table 4.9 presents four ordered probit specifications with and without user fixed effects. The high-reward dummy is positive in (1) and (2), significant in (1) but not in (2).*

By Result 13, we reject the null hypothesis in favor of Hypothesis 11 that a task with a high reward attracts submissions of higher quality than a task with a low reward. Furthermore, our fixed effects model indicates that, among users who participate in multiple tasks, there is no evidence that a user produces submissions of higher quality for high-reward tasks than those for low-reward tasks. In comparison, we find that, while programming tasks in the high-reward treatment attract higher average quality submissions than those in the low-reward treatment, this difference is not statistically significant (average quality of valid solutions: 3.88 vs. 3.75, $p = 0.220$, using ordered probit with standard errors clustered at the task level; average quality of best solutions: 5.00 vs. 4.78, $p = 0.380$, using one-sided permutation tests). Using similar analysis, we find that the reserve effects on quality is significant whether user fixed effects are controlled for, indicating that within-user variation accounts for part of the reserve effects on quality.

Result 14 (Reserve Effect on Submission Quality)**.** *The quality of valid and best translations is significantly lower in the reserve treatments than in the no-reserve treatments. This effect is partially driven by within-user variation over the presence of a reserve.*

Support. *Table 4.9 presents four ordered probit specifications with and without user fixed effects. The reserve dummy is negative and significant in all four specifications.*

Note that Result 14 contradicts Hypothesis 12. While a fully rational user should submit a solution only when its quality exceeds that of any previous submissions including the reserve, our participants do not always follow this simple rule. Our subsequent analysis suggests that the quality of valid translations submitted after the reserve are lower because the reserve deters the entry of more experienced users. By extension, experienced users differ from inexperienced ones in their ability to recognize high quality submissions.

The crucial factor which drives both Results 13 and 14 is user entry decisions. We now analyze user entry decisions by type, computed from two perspectives: user quality exhibited within our experiment, and their winning history prior to the start of our experiment.

We first investigate entry decisions using user quality computed within our experiment. We hypothesize that one possibility which might have led to the significant results in the pooled model is that tasks with a high reward (reserve) are more likely to attract (deter) high-quality users. To test this hypothesis, we construct a two-stage model.[19] In the first stage, we regress submission quality of each solution on the user dummies. Consequently, the estimated coefficient for user i, $\hat{\mu}_i$, approximates her user quality compared to that of the omitted user. This measure of user quality might be determined by various factors, including user ability, effort and reputation. We then construct a new statistic, $\bar{\hat{\mu}}_t = \frac{1}{n_t}\sum_{i=1}^{n_t} \hat{\mu}_t$, representing the average user quality per task, and regress $\bar{\hat{\mu}}_t$ on the reward size of each task, the reserve dummy and the task difficulty.

Table 4.10 reports two OLS specifications investigating determinants of average user quality among valid (specification 1) and best (specification 2) translation sub-

[19] We thank Jeff Smith for suggesting this approach.

Table 4.10 OLS: Determinants of the User Quality in Translation Tasks

Dependent Variable	Average User Quality Among Valid Solutions	Average User Quality Among Best Solutions
	(1)	(2)
High-Reward	0.757***	1.490**
	(0.219)	(0.623)
Reserve	-0.519**	-0.879
	(0.224)	(0.585)
Task Difficulty	-0.005	-0.125
	(0.118)	(0.433)
Constant	-1.742***	-0.106
	(0.428)	(1.303)
Observations	112	103
R^2	0.145	0.085

Notes:
1.Robust standard errors are in parentheses
2.Significant at: * 10%; ** 5%; *** 1%.

missions. In specification (1), we find that the coefficient of the high-reward dummy is positive and significant, indicating that a high-reward task attracts higher-quality users. In comparison, the coefficient of the reserve dummy is negative and significant, indicating that average user quality in a task with a reserve is lower. For best solutions (2), the coefficient of the high-reward dummy is positive and significant, indicating that, among users who provide best solutions, average user quality is significantly higher for a high-reward task compared to that for a low-reward task. In comparison, the coefficient of the reserve dummy is negative but insignificant ($p = 0.136$, two-sided).

Having analyzed entry decisions based on user quality exhibited within our experiment, we proceed to investigate entry decisions using user winning history prior to the start of our experiment. To do so, we first compute the median user credit per task. Considering all valid solutions for a task, we find that the average median user credit is higher in the high-reward treatment than that in the low-reward treatment. This difference is significant in the no-reserve treatments.

Result 15 (Reward Effect on Entry)**.** *Average user quality among valid and best translations is significantly higher in the high-reward than in the low-reward treatments. Furthermore, the average median user credit is significantly higher in the high-reward-no-reserve than in the low-reward-no-reserve treatment.*

Support. *Table 4.10 reports two OLS specifications investigating determinants of average user quality in translation tasks. The coefficient for the high-reward dummy is positive and significant in both specifications. Using user credit prior to our experiment, we find that, in the no-reserve treatments, the average median user credit is 0.45 in the high-reward treatment, and 0.05 in the low-reward treatment. A one-sided two-sample t-test rejects the null hypothesis in favor of the alternative hypothesis that the average median credit is higher in the high-reward treatment ($p = 0.048$). For the reserve treatments, the relationship holds but is not significant (0.14 vs. 0.09, $p = 0.290$, one-sided two-sample t-tests). In comparison, among all valid programming solutions for each task, again, the relationship holds but is not significant (2.09 vs. 1.34, $p = 0.196$, one-sided permutation tests).*

Using similar analysis, we now summarize the reserve effects on entry decisions, using user quality observed during our experiment (Table 4.10) and user credits accumulated prior to our experiment. Using user credit history as an indication of their type, we find that, among all valid solutions for a high-reward task, the average median user credit is weakly lower in the reserve treatment.

Result 16 (Reserve Effect on Entry)**.** *Average user quality among valid translations is significantly lower in the reserve than in the no-reserve treatments. Furthermore, the average median user credit is weakly lower in the reserve-high-reward than in the no-reserve-high-reward treatments.*

Support. *Table 4.10 reports two OLS specifications investigating determinants of user quality in translation tasks. The coefficient for the reserve dummy is negative and significant in specification (1). Using user credit prior to our experiment, we find that, in the high-reward treatment, the average median user credit is 0.14 in the reserve treatment and 0.45 in the no-reserve treatment. A one-sided two-sample t-test rejects the null hypothesis in favor of the alternative hypothesis that the average median credit is lower in the reserve treatment at the 10% level ($p = 0.097$). For the low-reward treatments, the comparison between the reserve and no-reserve treatments is not significant (0.05 vs. 0.09, p=0.323, one-sided two-sample t-tests).*

Overall, Result 16 indicates that the early entry of a high quality translation is more likely to deter high-quality (experienced) users compared to low-quality (inexperienced) users. The differential entry response to the presence of a high quality reserve partially explains our finding that the reserve has a negative effect on subsequent submission quality (Result 14).

Lastly, to test the theoretical predictions on entry timing in sequential all-pay auctions from Konrad and Leininger (2007), we investigate factors influencing submission time. Using naturally occurring field data on Taskcn, Yang, Adamic and Ackerman (2008*b*) find a positive correlation between reward size and later submission. A possible explanation is that users, especially experienced ones, strategically wait to submit solutions for high reward tasks. An alternative explanation is that higher rewards are offered for more difficult tasks, which require more time to complete. As reward level is endogenously determined in naturally occurring field data but exogenously determined in our experiment, we are able to separate the effects of reward size and task difficulty on submission timing.

Table 4.11 Determinants of Submission Time

Dependent Variable	Submission Time (All)		Submission Time (Valid)	
	(1)	(2)	(3)	(4)
High-Reward	0.200***	0.133***	0.466**	0.305
	(0.045)	-0.045	(0.232)	(0.208)
Valid Solution		1.236***		
		(0.109)		
Reserve		-0.032		-0.0261
		(0.049)		(0.195)
Task Difficulty		0.013		0.216**
		(0.024)		(0.092)
Experienced Users		0.074		0.754***
		(0.102)		(0.280)
Constant	0.571***	0.402***	1.457***	0.632
	(0.032)	(0.115)	(0.200)	(0.415)
Observations	3,515	3,515	485	485
R^2	0.004	0.088	0.011	0.040

Notes:
1. Standard errors in parentheses are clustered at the task level.
2. Significant at: * 10%; ** 5%; *** 1%.

In Table 4.11, we report four OLS specifications to investigate factors affecting submission time for all (specifications 1 and 2) and valid translations (specifications 3 and 4). To replicate results from Yang, Adamic and Ackerman (2008*b*), specifications (1) and (3) include the high-reward dummy as the only independent variable. In comparison, specifications (2) and (4) include the following additional independent variables (with omitted variables in parentheses): reserve (no reserve), task difficulty, and experienced users (inexperienced users). When other variables are not controlled for, we replicate the finding in Yang et al. (2008*b*) that a high reward has a positive and significant effect on submission time. However, this significance disappears for valid solutions after controlling for task difficulty and user experience. We summarize these results below.

Result 17 (Submission Time)**.** *For valid translations, experienced users submit their translations significantly later than do inexperienced ones, controlling for task difficulty.*

Support. *In specification (4) of Table 4.11, the coefficient of the experienced user dummy is positive and significant at the 1% level, indicating that experienced users submit their solutions later than others. On average, experienced users submit their solutions 0.754 days later than inexperienced ones do.*

Among all solutions, we find that those for a high-reward task are submitted 0.13 days later. Furthermore, a valid translation is submitted 1.236 days later than a machine-translation. Restricting our analysis to valid submissions, translations for a high-reward task are still submitted significantly later than those for a low-reward task. After controlling for task difficulty, however, we find that experienced users submit their solutions 0.754 days later than inexperienced users do, while the reward effect on submission time is no longer significant. Furthermore, the task difficulty coefficient is positive and significant, indicating that users take 0.216 days longer to submit a valid solution for each additional level of difficulty (on a 1-7 Likert scale).

In summary, we find significant reward effects on participation levels and submission quality. Furthermore, a higher reward also attracts higher quality (more experienced) users, indicating that a monetary incentive is effective in inducing more submissions and better solutions. Interestingly, while the early entry of a high quality solution does not significantly affect the number of submissions, we find that solution quality dramatically decreases with the presence of a reserve, as it deters the entry of high quality (experienced) users. In addition to their entry decisions, experienced users also submit their solutions later than inexperienced users do, controlling for task difficulty.

4.8 Discussion

As crowdsourcing has become an increasingly important problem-solving method, utilized by individuals, non-profit and for-profit organizations alike, evaluating the behavioral response of various design features will help improve the performance of crowdsourcing institutions and increase user satisfaction. In this study, we examine the effect of different design features of a crowdsourcing site on participation levels, submission quality and user entry decisions. Conducting a field experiment on Taskcn, a nascent online labor market, we find that a higher reward induces more participation and higher submission quality. By controlling for the existence of a reserve in the form of a high quality early submission, we find that a reserve lowers subsequent submission quality, as it preferentially deters the entry of experienced users. Experienced users also distinguish themselves from inexperienced ones by being more likely to select higher reward tasks over lower reward ones, and by submitting their solutions later.

Through our field experiment, we are able to observe interesting patterns that likely would not have emerged had the experiment been conducted in a lab setting. The most surprising finding of our experiment is that the entry decisions of high quality (experienced) users drive the reward and reserve effects on submission quality. A higher reward attracts more experienced users, while a high quality reserve deters them. This finding not only informs the design of crowdsourcing institutions, but also provides useful feedback to theory (Samuelson, 2005). While most existing theoretical models of all-pay auctions ignore entry decisions, the model with endogenous entry (DiPalantino and Vojnovic, 2009) treats every user as fully rational, which cannot explain our reserve effects on quality.[20] Our results suggest that a more accurate theory for predicting behavior in the field should incorporate behavior of both naive and sophisticated types, such as an extension of the cognitive hierarch model (Camerer, Ho and Chong, 2004) to the all-pay auction domain.

The second is the way the site actually provides users with the power to transform a sequential all-pay auction into a simultaneous all-pay auction, by allowing users to hide solutions from other participants. We find that valid solution providers are more likely to protect their solutions compared to those who provide machine-generated translations (10% vs. 2%, $p < 0.01$, one-sided test of proportions), suggesting that the result of true effort is more likely to be protected from being copied by others. Again, the endogenous choice of auction format has not been evaluated theoretically. Our study provides the first empirical evaluations of such mechanisms, which might inform future theoretical research.

Lastly, we find that the majority of translations submitted are machine translations, which require very little effort on the part of the participants but increase the

[20]Morgan, Orzen and Sefton (2010) presents a theoretical model with endogenous participation in the Tullock contest, which differs from an all-pay auction.

screening effort of requesters. This finding reveals the need for an entry barrier or censoring mechanism if a site wants to provide a better user experience for requesters. In addition, while a reserve in the form of an early high quality solution deters the entry of high quality (experienced) users in the experiment, it does not deter low-quality submissions, which indicates the need for additional incentives to attract high quality and deter low-quality users. One possible way to encourage earlier entry by experts is a tie-breaking rule favoring an earlier entry, as has been tested and confirmed in a lab setting by Liu (2011).

Finally, a feature of the Taskcn site we did not explore in this study is the option of designating multiple winners as opposed to a single winner for a task. Using multiple rewards to induce greater effort than a single reward is well-modeled (Moldovanu and Sela, 2001) and examined in laboratory experiments (Muller and Schotter, 2010). However, to our knowledge, there has not yet been a field experiment investigating the effect of allowing multiple winners on submission behavior. This can be a natural extension of our present work.

4.9 Appendix A: Proofs and Examples

Recall that Propositions 1 through 3 require the assumption that the ability distribution function is from the family, $F(x) = x^c$, where $0 < c < 1$.

Proof of Proposition 1: In what follows, we will consider two cases, the case with a zero reserve, and one with a positive reserve.

Case 1: Zero Reserve. We first derive the equilibrium bidding function for each user, when the reserve is zero, i.e., $q_0 = 0$.

Using backward induction, we expect that user n will win the auction if the quality of her solution is higher than or equal to the best quality among all previous submissions, which is $\max\{q_j(a_j)\}_{j<n}$, and if her ability is sufficiently high, $a_n \geq \frac{1}{v}\max\{q_j(a_j)\}_{j<n}$. If her ability is not high enough, i.e., $a_n < \frac{1}{v}\max\{q_j(a_j)\}_{j<n}$, her benefit from winning (v) is less than her bidding cost, thus she should bid zero. Therefore, the equilibrium bidding function of the last user, n, is given by:

$$q_n(a_n) = \begin{cases} 0 & \text{if } 0 \leq a_n < \frac{1}{v}\max\{q_j(a_j)\}_{j<n}, \\ \max\{q_j(a_j)\}_{j<n} & \text{if } \frac{1}{v}\max\{q_j(a_j)\}_{j<n} \leq a_n \leq 1. \end{cases} \tag{4.1}$$

Next, we derive the equilibrium bidding function for user i, where $i = 2, \ldots, n-1$. We do so by solving the following constrained optimization problem. Applying the Revelation Principle, user i with ability a_i will choose to behave as a user with ability s who maximizes her expected payoff:

$$\begin{array}{ll} \max\limits_{s} & \{v\prod_{j=i+1}^{n} F_j(q_j = 0) - \frac{q_i(s)}{a_i}\} \\ s.t. & q_i(s) \geq \max\{q_j(a_j)\}_{j<i}. \end{array} \tag{4.2}$$

As $F(x) = x^c$, the probability that user i wins the auction conditional on her submitting a solution with quality at least as high as the best previous submission becomes:

$$\begin{aligned} \prod_{j=i+1}^{n} F_j(q_j = 0) &= \left[\frac{q_i(s)}{v}\right]^{c(1-c)^{(n-(i+1))}} \left[\frac{q_i(s)}{v}\right]^{c(1-c)^{(n-(i+2))}} \cdots \left[\frac{q_i(s)}{v}\right]^{c} \\ &= \left[\frac{q_i(s)}{v}\right]^{1-(1-c)^{n-i}}. \end{aligned}$$

When the constraint is not binding, the first-order condition is then:

$$v[1-(1-c)^{n-i}]\left[\frac{q_i(s)}{v}\right]^{-(1-c)^{n-i}}\frac{q_i'(s)}{v}-\frac{q_i'(s)}{a_i}=0.$$

Assuming the interior part of the equilibrium bidding function is strictly monotone, i.e., $q_i'(s)>0$, the first-order condition becomes:

$$[1-(1-c)^{n-i}]\left[\frac{q_i(s)}{v}\right]^{-(1-c)^{n-i}}-\frac{1}{a_i}=0. \tag{4.3}$$

Therefore, the interior solution is $q_i(a_i)=v[a_i(1-(1-c)^{n-i})]^{\frac{1}{(1-c)^{n-i}}}$. Let $d_i\equiv(1-c)^{n-i}$. Thus, we can rewrite the interior solution as:

$$q_i(a_i)=v[a_i(1-d_i)]^{\frac{1}{d_i}}. \tag{4.4}$$

The second-order condition is then:

$$\begin{aligned}
&q_i''(s)\left[(1-d_i)\left(\tfrac{q_i(s)}{v}\right)^{-d_i}-\tfrac{1}{a_i}\right]+\tfrac{(q_i'(s))^2}{v}\left\{-d_i\,(1-d_i)\left[\tfrac{q_i(s)}{v}\right]^{-d_i-1}\right\}\\
&=-\tfrac{(q_i'(s))^2}{v}\left\{d_i\,(1-d_i)\left[\tfrac{q_i(s)}{v}\right]^{-d_i-1}\right\}, \text{as the first term is zero by Equation (4.3),}\\
&<0.
\end{aligned}$$

To characterize the equilibrium bidding function, we define two boundaries as:

$$\overleftarrow{a}_i=\left[\frac{\max\{q_j(a_j)\}_{j<i}}{v}\right]^{d_i},\ \text{and}\ \overrightarrow{a}_i=\frac{1}{1-d_i}\left[\frac{\max\{q_j(a_j)\}_{j<i}}{v}\right]^{d_i}.$$

These boundaries partition the support of abilities into three ranges:

1. When $0\le a_i<\overleftarrow{a}_i$, the expected payoff from submitting a positive bid is negative. Thus, the user should submit a zero bid.
2. When $\overleftarrow{a}_i\le a_i<\overrightarrow{a}_i$, as $\max\{q_j(a_j)\}_{j<i}>v[(1-d_i)a_i]^{\frac{1}{d_i}}$, bidding $\max\{q_j(a_j)\}_{j<i}$ dominates $v[(1-d_i)a_i]^{\frac{1}{d_i}}$. Therefore, the constraint is binding, and we obtain a corner solution.
3. When $\overrightarrow{a}_i\le a_i\le 1$, Equation (4.4) is the interior solution of the constrained optimization problem (4.2) while the constraint is not binding.

Summarizing the above analysis, we characterize the equilibrium bidding function for user $i \in \{2, \cdots, n-1\}$ as follows:

$$q_i(a_i) = \begin{cases} 0 & \text{if } 0 \leq a_i < \overleftarrow{a}_i, \\ \max\{q_j(a_j)\}_{j<i} & \text{if } \overleftarrow{a}_i \leq a_i < \overrightarrow{a}_i, \\ v[a_i(1-d_i)]^{\frac{1}{d_i}} & \text{if } \overrightarrow{a}_i \leq a_i \leq 1. \end{cases} \tag{4.5}$$

Note that when $\max\{q_j(a_j)\}_{j<i} \geq v(1-d_i)^{\frac{1}{d_i}}$, the third range of Equation (4.5) does not exist.

Lastly, user 1's bidding function is $q_1(a_1) = v\left[a_1(1-d_1)\right]^{\frac{1}{d_1}}$, where $0 \leq a_1 \leq 1$.

Now we derive the comparative statics of the reward effect on participation. Let $P_i(q_i = 0)$ be the probability that user i bids zero. For user $i > 1$, the probability of bidding zero depends on $\overleftarrow{a}_i$. Since $\max\{q_j(a_j)\}_{j<i} = \max\{v\left[a_j(1-d_j)\right]^{\frac{1}{d_j}}\}_{j<i} = v\max\{[a_j(1-d_j)]^{\frac{1}{d_j}}\}_{j<i}$, we obtain $\overleftarrow{a}_i = \max\{[(1-d_j)a_j]^{\frac{d_i}{d_j}}\}_{j<i}$, which is independent of the reward level, v. In addition, for user 1, $q_1(a_1) > 0$, $\forall a_1 > 0$, and $a_1 = 0$ is a measure zero event. Therefore, $P_1(q_1 = 0) = 0$. In summary, with a zero reserve, the probability of participation for any user i, $1 - P_i(q_i = 0)$, is independent of v.

Case 2: Positive Reserve. We now consider the positive reserve case, i.e., $q_0 > 0$.

As in Case 1, we first characterize the equilibrium bidding function of the last user, n, in the following two scenarios:

1. If the maximum bid from previous users does not exceed the reserve, i.e., $\max\{q_j(a_j)\}_{j<n} \leq q_0$, the only constraint for user n is the reserve, q_0. Thus, user n's bidding function becomes:

$$q_n(a_n) = \begin{cases} 0 & \text{if } 0 \leq a_n < \frac{q_0}{v}, \\ q_0 & \text{if } \frac{q_0}{v} \leq a_n \leq 1. \end{cases} \tag{4.6}$$

   When $a_n < \frac{q_0}{v}$, the expected payoff from submitting a positive bid is negative. Thus, she should bid zero. When $a_n \geq \frac{q_0}{v}$, as $\max\{q_j(a_j)\}_{j<n} \leq q_0$, the optimal bidding strategy for user n is to bid q_0. Consequently, she wins the auction.
2. If the maximum bid from previous users exceeds the reserve, i.e., $\max\{q_j(a_j)\}_{j<n} > q_0$, user n's bidding function is characterized by Equation (4.1).

For user i, $i = 2, ..n-1$, her equilibrium bidding function, $q_i(a_i)$, is the solution to the optimization problem (4.2), with the additional constraint, $q_i(s) \geq q_0$. It is separately characterized in the following two scenarios:

1. If the maximum bid from previous users does not exceed the reserve, i.e.,

$\max\{q_j(a_j)\}_{j<i} \leq q_0$, the equilibrium bidding function is characterized by

$$q_i(a_i) = \begin{cases} 0 & \text{if } 0 \leq a_i < \overleftarrow{a}_i, \\ q_0 & \text{if } \overleftarrow{a}_i \leq a_i < \overrightarrow{a}_i, \\ v[a_i(1-d_i)]^{\frac{1}{d_i}} & \text{if } \overrightarrow{a}_i \leq a_i \leq 1, \end{cases} \tag{4.7}$$

where the boundaries are defined as $\overleftarrow{a}_i(v, q_0) = (\frac{q_0}{v})^{d_i}$ and $\overrightarrow{a}_i(v, q_0) = \frac{1}{1-d_i}(\frac{q_0}{v})^{d_i}$.

2. If the maximum bid from previous users exceeds the reserve, i.e., $\max\{q_j(a_j)\}_{j<i} > q_0$, the equilibrium bidding function is characterized by Equation (4.5).

Lastly, user 1's equilibrium bidding function is characterized by Equation (4.7) with $i = 1$.

To characterize user i's *ex ante* likelihood of submitting a solution with positive quality, e.g., $P_i(q_i > 0)$, we first compute her probability of bidding zero, $P_i(q_i = 0)$.

Define $q_i^* \equiv [a_i(1-d_i)]^{\frac{1}{d_i}}$ as user i's bid in the third range of her equilibrium bidding function when $v = 1$.

First, when $i = 1$, the probability of bidding 0 is $F(\overleftarrow{a}_1(v, q_0)) = F\left((\frac{q_0}{v})^{d_1}\right)$ and $\{\partial F\left((\frac{q_0}{v})^{d_1}\right)\}/\{\partial v\} = (-cd_1)v^{-cd_1-1}q_0^{cd_1} < 0$.

Next, for user $i > 1$, we define a sequence of conditional probabilities, $N_i(v, q_0)^{(j)}$, where $1 \leq j < i$, as follows:

$$N_i(v, q_0)^{(1)} = P_i(q_i = 0 | q_1 \leq q_0),$$

$$\ldots$$

$$N_i(v, q_0)^{(j)} = P_i(q_i = 0 | q_1, \ldots, q_j \leq q_0),$$

$$\ldots$$

$$N_i(v, q_0)^{(i-1)} = P_i(q_i = 0 | q_{j<i} \leq q_0) = F\left(\overleftarrow{a}_i(v, q_0)\right) = F\left(\left(\frac{q_0}{v}\right)^{d_i}\right).$$

Therefore, $N_i(v, q_0)^{(j)} = P_i(q_i = 0 | q_1, \ldots, q_j \leq q_0)$ is the conditional probability for user i to bid 0 when none of the first j users' bids exceeds the reserve. In particular, $N_i(v, q_0)^{(i-1)}$ is the conditional probability for user i to bid 0 when none of the previous bids exceeds the reserve, which is equivalent to user i being the first active user in the new sequence with $n - i + 1$ users.

Define another sequence of conditional probabilities for each user $i > 1$, $O_i(a_j)$, where $1 \leq j < i$, as follows:

$$O_i(a_1) = P_i(q_i = 0 | q_1 = vq_1^*),$$

$$O_i(a_2) = P_i(q_i = 0 | q_1 \leq q_0, q_2 = vq_2^*),$$

$$\dots$$

$$O_i(a_j) = P_i(q_i = 0 | q_1, \dots, q_{j-1} \leq q_0, q_j = vq_j^*),$$

$$\dots$$

$$O_i(a_{i-1}) = P_i(q_i = 0 | q_1, \dots, q_{i-2} \leq q_0, q_{i-1} = vq_{i-1}^*) = F\left(\overleftarrow{a}_i(q_{i-1}^*)\right) = F\left((q_{i-1}^*)^{d_i}\right).$$

Therefore, $O_i(a_j) = P_i(q_i = 0 | q_1, \dots, q_{j-1} \leq q_0, q_j = vq_j^*)$ is the conditional probability for user i to bid 0 when none of the bids before user j exceeds the reserve, $q_1, \dots, q_{j-1} \leq q_0$, and user j's bid is the equilibrium bid in the third range of her equilibrium bidding function, vq_j^*.

Moreover, $\forall 1 < j \leq i$, we characterize the conditional probability for user i to bid 0 when none of the first $j-1$ bids exceeds the reserve as:

$$N_i(v, q_0)^{(j-1)} = \int_0^{\overrightarrow{a}_j(v,q_0)} N_i(v, q_0)^{(j)} f(a_j) da_j + \int_{\overrightarrow{a}_j(v,q_0)}^1 O_i(a_j) f(a_j) da_j. \tag{4.8}$$

The first term is the conditional probability for user i to bid 0 with a random variable $q_j \leq q_0$, $P_i(q_i = 0, q_j \leq q_0 | q_1, \dots, q_{j-1} \leq q_0)$. The second term is the conditional probability for user i to bid 0 with a random variable $q_j = vq_j^* \geq q_0$, $P_i(q_i = 0, q_j = vq_j^* | q_1, \dots, q_{j-1} \leq q_0)$. Differentiating Equation (4.8) with respect to v and using the Leibniz integral rule, we obtain:

$$
\begin{aligned}
\frac{\partial N_i(v,q_0)^{(j-1)}}{\partial v} &= \frac{\partial \overrightarrow{a}_j(v,q_0)}{\partial v} N_i(v,q_0)^{(j)} f\left(\overrightarrow{a}_j(v,q_0)\right) + \int_0^{\overrightarrow{a}_j(v,q_0)} \frac{\partial N_i(v,q_0)^{(j)}}{\partial v} f(a_j)da_j \\
& -\frac{\partial \overrightarrow{a}_j(v,q_0)}{\partial v} O_i\left(\overrightarrow{a}_j(v,q_0)\right) f\left(\overrightarrow{a}_j(v,q_0)\right).
\end{aligned}
$$

By continuity of the equilibrium bidding function at $\overrightarrow{a}_j(v,q_0)$, we have $N_i(v,q_0)^{(j)} = O_i\left(\overrightarrow{a}_j(v,q_0)\right)$. Therefore, the first and third terms on the RHS cancel each other, which simplifies the RHS:

$$
\begin{aligned}
\frac{\partial N_i(v,q_0)^{(j-1)}}{\partial v} &= \int_0^{\overrightarrow{a}_j(v,q_0)} \frac{\partial N_i(v,q_0)^{(j)}}{\partial v} f(a_j)da_j \\
&= \frac{\partial N_i(v,q_0)^{(j)}}{\partial v} F\left(\overrightarrow{a}_j(v,q_0)\right),
\end{aligned}
\tag{4.9}
$$

as $\{\partial N_i(v,q_0)^{(j)}\}/\{\partial v\}$ is independent of a_j.

Therefore, the probability of bidding 0 for user i, $P_i(q_i = 0)$, can be rewritten as:

$$
P_i(q_i = 0) = \int_0^{\overrightarrow{a}_1(v,q_0)} N_i(v,q_0)^{(1)} f(a_1)da_1 + \int_{\overrightarrow{a}_1(v,q_0)}^1 O_i(a_1)f(a_1)da_1.
$$

Expanding $N_i(v,q_0)^{(1)}$ and $O_i(a_1)$, we have:

$$
\begin{aligned}
N_i(v,q_0)^{(1)} &= P_i(q_i = 0 | q_1 \leq q_0) \\
&= \int_0^{\overrightarrow{a}_2(v,q_0)} \int_0^{\overrightarrow{a}_3(v,q_0)} \cdots \int_0^{\overrightarrow{a}_{i-1}(v,q_0)} \left[\int_0^{\overleftarrow{a}_i(v,q_0)} f(a_i)da_i\right] \ldots f(a_2)da_2 \\
& + \int_0^{\overrightarrow{a}_2(v,q_0)} \int_0^{\overrightarrow{a}_3(v,q_0)} \cdots \int_{\overrightarrow{a}_{i-1}(v,q_0)}^1 \left[\int_0^{\overleftarrow{a}_i(q^*_{i-1})} f(a_i)da_i\right] \ldots f(a_2)da_2 \\
& + \ldots \\
& + \int_{\overrightarrow{a}_2(v,q_0)}^1 \int_{\overrightarrow{a}_3(q^*_2)}^1 \cdots \int_{\overrightarrow{a}_{i-1}(q^*_{i-2})}^1 \left[\int_0^{\overleftarrow{a}_i(q^*_{i-1})} f(a_i)da_i\right] \ldots f(a_2)da_2,
\end{aligned}
$$

and

$$\begin{aligned}
O_i(a_1) &= P_i(q_i = 0 | q_1 = vq_1^*) \\
&= \int_0^{\overrightarrow{a}_2(q_1^*)} \int_0^{\overrightarrow{a}_3(q_1^*)} \dots \int_0^{\overrightarrow{a}_{i-1}(q_1^*)} \left[\int_0^{\overleftarrow{a}_i(q_1^*)} f(a_i)da_i \right] \dots f(a_2)da_2 \\
&\quad + \int_0^{\overrightarrow{a}_2(q_1^*)} \int_0^{\overrightarrow{a}_3(q_1^*)} \dots \int_{\overrightarrow{a}_{i-1}(q_1^*)}^{1} \left[\int_0^{\overleftarrow{a}_i(q_{i-1}^*)} f(a_i)da_i \right] \dots f(a_2)da_2 \\
&\quad + \dots \\
&\quad + \int_{\overrightarrow{a}_2(q_1^*)}^{1} \int_{\overrightarrow{a}_3(q_2^*)}^{1} \dots \int_{\overrightarrow{a}_{i-1}(q_{i-2}^*)}^{1} \left[\int_0^{\overleftarrow{a}_i(q_{i-1}^*)} f(a_i)da_i \right] \dots f(a_2)da_2,
\end{aligned}$$

where the boundaries are defined as $\overleftarrow{a}_i(q_j^*) = (q_j^*)^{d_i}$ and $\overrightarrow{a}_i(q_j^*) = \frac{1}{1-d_i}(q_j^*)^{d_i}$.

Using the Leibniz integral rule, we have:

$$\begin{aligned}
\frac{\partial P_i(q_i = 0)}{\partial v} &= \frac{\partial \overrightarrow{a}_1(v, q_0)}{\partial v} N_i(v, q_0)^{(1)} f\left(\overrightarrow{a}_1(v, q_0)\right) + \int_0^{\overrightarrow{a}_1(v,q_0)} \frac{\partial N_i(v, q_0)^{(1)}}{\partial v} f(a_1)da_1 \\
&\quad - \frac{\partial \overrightarrow{a}_1(v, q_0)}{\partial v} O_i\left(\overrightarrow{a}_1(v, q_0)\right) f\left(\overrightarrow{a}_1(v, q_0)\right) \\
&= \int_0^{\overrightarrow{a}_1(v,q_0)} \frac{\partial N_i(v, q_0)^{(1)}}{\partial v} f(a_1)da_1 \\
&= \frac{\partial N_i(v, q_0)^{(1)}}{\partial v} F\left(\overrightarrow{a}_1(v, q_0)\right).
\end{aligned}$$

The second equality obtains as the first and third terms cancel each other. The third equality obtains as $\{\partial N_i(v, q_0)^{(1)}\}/\{\partial v\}$ is independent of a_1. Moreover, by iteratively applying Equation (4.9), we obtain:

$$\begin{aligned}
\frac{\partial P_i(q_i = 0)}{\partial v} &= \frac{\partial N_i(v, q_0)^{(1)}}{\partial v} F\left(\overrightarrow{a}_1(v, q_0)\right) \\
&= \left(\frac{\partial N_i(v, q_0)^{(2)}}{\partial v} F\left(\overrightarrow{a}_2(v, q_0)\right) \right) F\left(\overrightarrow{a}_1(v, q_0)\right) \\
&= \left(\frac{\partial N_i(v, q_0)^{(i-1)}}{\partial v} F\left(\overrightarrow{a}_{i-1}(v, q_0)\right) \right) F\left(\overrightarrow{a}_{i-2}(v, q_0)\right) \cdots F\left(\overrightarrow{a}_1(v, q_0)\right) \\
&= \frac{\partial F\left((\frac{q_0}{v})^{d_i} \right)}{\partial v} F\left(\overrightarrow{a}_{i-1}(v, q_0)\right) \cdots F\left(\overrightarrow{a}_1(v, q_0)\right). \\
&= (-cd_i) v^{-cd_i - 1} q_0^{cd_i} F\left(\overrightarrow{a}_{i-1}(v, q_0)\right) \cdots F\left(\overrightarrow{a}_1(v, q_0)\right)
\end{aligned}$$

As $F\left(\overrightarrow{a}_{i-1}(v, q_0)\right) \cdots F\left(\overrightarrow{a}_1(v, q_0)\right) > 0$, we obtain $\{\partial P_i(q_i = 0)\}/\{\partial v\} < 0$. In summary, with a positive reserve, $q_0 > 0$, the probability of participation for user i

strictly increases in v.

$\blacksquare$

We now use a two-user example, adapted from Segev and Sela (2011), to illustrate our theoretical results.

Example 1. *Consider a sequential all-pay auction with two users whose abilities are i.i.d. draws from a concave distribution function $F(x) = x^{0.5}$ with support on $[0, 1]$. In addition, the reward is $v \geq 1$.*

In this example, the equilibrium bidding function for user 1 is $q_1(a_1) = \frac{a_1^2}{4}v$. After observing 1's submission, user 2 bids according to the following equilibrium bidding function,

$$q_2(a_2) = \begin{cases} 0 & \text{if } 0 \leq a_2 < \frac{a_1^2}{4}, \\ \frac{a_1^2}{4}v & \text{if } \frac{a_1^2}{4} \leq a_2 \leq 1. \end{cases}$$

The likelihood that user 1 submits a positive bid is 1, while the conditional likelihood that user 2 submits a positive bid is

$$\text{Prob}(q_2 > 0 \mid a_1) = 1 - F(\frac{a_1^2}{4}) = 1 - \frac{a_1}{2}.$$

In addition, the likelihood that user 2 submits a positive bid is:

$$\text{Prob}(q_2 > 0) = \int_0^1 (1 - \frac{a_1}{2})\frac{1}{2\sqrt{a_1}}da_1 \approx 0.83.$$

Lastly, the expected quality for each user, Q_1, Q_2, the average and the highest quality, AQ and HQ, can be characterized as follows:

$$\begin{aligned} Q_1 &= v\int_0^1 \frac{a_1^2}{4}\frac{1}{2\sqrt{a_1}}da_1 = 0.05v, \\ Q_2 &= \int_0^1 \int_{\frac{a_1^2}{4}}^1 \frac{a_1^2}{4}vf(a_2)da_2 f(a_1)da_1 \approx 0.03v, \\ AQ &= \frac{v}{2}\int_0^1 (2 - \frac{a}{2})\frac{a_1^2}{4}\frac{1}{2\sqrt{a_1}}da_1 \approx 0.04v, \\ HQ &= Q_1 = 0.05v. \end{aligned}$$

Note, with zero reserve, user i's expected quality, Q_i, is less than Q_{i-1}. Therefore, the expected highest quality is $HQ = Q_1$.

Proof of Proposition 2:

We now prove that user i's expected submission quality, $Q_i(q_i)$, strictly increases in the reward level, v.

First, when $i = 1$, we show that $\{\partial Q_1(v, q_0)\}/\{\partial v\} > 0$, i.e., $\forall\ v_2 > v_1$, $Q_1(v_2, q_0) > Q_1(v_1, q_0)$.

$$
\begin{aligned}
Q_1(v_1, q_0) &= \int_{\overleftarrow{a}_1(q_0,v_1)}^{\overrightarrow{a}_1(q_0,v_1)} q_0 f(a_1)da_1 + \int_{\overrightarrow{a}_1(q_0,v_1)}^{1} v_1 q_1^* f(a_1)da_1 \\
&= \int_{(\frac{q_0}{v_1})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}} q_0 f(a_1)da_1 + \int_{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}}^{1} v_1 q_1^* f(a_1)da_1 \\
&< \left[\int_{(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}} q_0 f(a_1)da_1 + \int_{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}} q_0 f(a_1)da_1\right] + \\
&\quad \int_{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}}^{1} v_1 q_1^* f(a_1)da_1 \\
&< \left[\int_{(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}} q_0 f(a_1)da_1 + \int_{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}} v_2 q_1^* f(a_1)da_1\right] + \\
&\quad \int_{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}}^{1} v_2 q_1^* f(a_1)da_1 \\
&= \int_{(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}} q_0 f(a_1)da_1 + \\
&\quad \left[\int_{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}} v_2 q_1^* f(a_1)da_1 + \int_{\frac{1}{1-d_1}(\frac{q_0}{v_1})^{d_1}}^{1} v_2 q_1^* f(a_1)da_1\right] \\
&= \int_{(\frac{q_0}{v_2})^{d_1}}^{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}} q_0 f(a_1)da_1 + \int_{\frac{1}{1-d_1}(\frac{q_0}{v_2})^{d_1}}^{1} v_2 q_1^* f(a_1)da_1 \\
&= Q_1(v_2, q_0).
\end{aligned}
$$

Next, for user $i > 1$, we define a sequence of conditional expected quality, $T_i(v, q_0)^{(j)}$, where $1 \le j < i$, as follows:

$$T_i(v, q_0)^{(1)} = Q_i(q_i | q_1 \le q_0),$$

$$\ldots$$

$$T_i(v, q_0)^{(j)} = Q_i(q_i | q_1, \ldots, q_j \leq q_0),$$

$$\ldots$$

$$\begin{aligned} T_i(v, q_0)^{(i-1)} &= Q_i(q_i | q_{j<i} \leq q_0) \\ &= \int_{\overleftarrow{a}_i(q_0, v)}^{\overrightarrow{a}_i(q_0, v)} q_0 f(a_i) da_i + \int_{\overrightarrow{a}_i(q_0, v)}^{1} v q_i^* f(a_i) da_i \\ &= \int_{(\frac{q_0}{v})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v})^{d_i}} q_0 f(a_i) da_i + \int_{\frac{1}{1-d_i}(\frac{q_0}{v})^{d_i}}^{1} v \left((1-d_i)a_i\right)^{\frac{1}{d_i}} f(a_i) da_i. \end{aligned}$$

Therefore, $T_i(v, q_0)^{(j)} = Q_i(q_i | q_1, , , q_j \leq q_0)$ is the conditional expected quality for user i when none of the first j bids exceeds the reserve. In particular, $T_i(v, q_0)^{(i-1)}$ is the conditional expected quality for user i when none of the previous bids exceeds the reserve, which is equivalent to user i being the first active user in the new sequence with $n - i + 1$ users.

Define another sequence of conditional expected quality for user i, $S_i(v, a_j)$, where $1 \leq j < i$, as follows:

$$S_i(v, a_1) = Q_i(q_i | q_1 = v q_1^*),$$

$$S_i(v, a_2) = Q_i(q_i | q_1 \leq q_0, q_2 = v q_2^*),$$

$$\ldots$$

$$S_i(v, a_j) = Q_i(q_i | q_1, \ldots, q_{j-1} \leq q_0, q_j = v q_j^*),$$

$$\ldots$$

$$S_i(v, a_{i-1}) = Q_i(q_i | q_1, \ldots, q_{i-2} \leq q_0, q_{i-1} = v q_{i-1}^*).$$

Therefore, $S_i(v, a_j) = Q_i(q_i | q_1 \ldots q_{j-1} \leq q_0, q_j = v q_j^*)$ is the conditional expected quality for user i when none of the bids before user j exceeds the reserve,

$q_1, \ldots, q_{j-1} \leq q_0$, and user j's bid is the equilibrium bid in the third range of her equilibrium bidding function, vq_j^*.

Moreover, $\forall 1 < j \leq i$, we characterize the conditional expected quality for user i when none of the first $j-1$ bids exceeds the reserve as:

$$T_i(v, q_0)^{(j-1)} = \int_0^{\overrightarrow{a}_j(v,q_0)} T_i(v, q_0)^{(j)} f(a_j) da_j + \int_{\overrightarrow{a}_j(v,q_0)}^1 S_i(v, a_j) f(a_j) da_j. \quad (4.10)$$

The first term is the conditional expected quality for user i with a random variable $q_j \leq q_0$, $Q_i(q_i, q_j \leq q_0 | q_1, \ldots, q_{j-1} \leq q_0)$. The second term is the conditional expected quality for user i with a random variable $q_j = vq_j^* \geq q_0$, $Q_i(q_i, q_j = vq_j^* | q_1, \ldots, q_{j-1} \leq q_0)$. Differentiating Equation (4.10) with respect to v and using the Leibniz integral rule, we obtain:

$$\begin{aligned}
\frac{\partial T_i(v, q_0)^{(j-1)}}{\partial v} = & \frac{\partial \overrightarrow{a}_j(v, q_0)}{\partial v} T_i(v, q_0)^{(j)} f\left(\overrightarrow{a}_j(v, q_0)\right) + \int_0^{\overrightarrow{a}_j(v,q_0)} \frac{\partial T_i(v, q_0)^{(j)}}{\partial v} f(a_j) da_j \\
& - \frac{\partial \overrightarrow{a}_j(v, q_0)}{\partial v} S_i\left(v, \overrightarrow{a}_j(v, q_0)\right) f\left(\overrightarrow{a}_j(v, q_0)\right) + \\
& \int_{\overrightarrow{a}_j(v,q_0)}^1 \frac{\partial S_i(v, a_j)}{\partial v} f(a_j) da_j.
\end{aligned}$$

By continuity of the equilibrium bidding function at $\overrightarrow{a}_j(v, q_0)$, we have $T_i(v, q_0)^{(j)} = S_i\left(v, \overrightarrow{a}_j(v, q_0)\right)$. Therefore, the first and third terms on the RHS cancel each other, which simplifies the RHS:

$$\begin{aligned}
\frac{\partial T_i(v, q_0)^{(j-1)}}{\partial v} &= \int_0^{\overrightarrow{a}_j(v,q_0)} \frac{\partial T_i(v, q_0)^{(j)}}{\partial v} f(a_j) da_j + \int_{\overrightarrow{a}_j(v,q_0)}^1 \frac{\partial S_i(v, a_j)}{\partial v} f(a_j) da_j \\
&= \frac{\partial T_i(v, q_0)^{(j)}}{\partial v} F\left(\overrightarrow{a}_j(v, q_0)\right) + \int_{\overrightarrow{a}_j(v,q_0)}^1 \frac{\partial S_i(v, a_j)}{\partial v} f(a_j) da_j, \quad (4.11)
\end{aligned}$$

as $\partial T_i(v, q_0)^{(j)} / \partial v$ is independent of a_j. Furthermore, $S_i(v, a_j) = Q_i(q_i | q_1 \ldots q_{j-1} \leq q_0, q_j = vq_j^*)$ is equivalent to user i's conditional expected quality with $q_j = vq_j^*$ in the zero reserve case, as q_0 is no longer binding. Thus, user i's submitted quality can take on the value of (i) user j's interior solution, or (ii) user k's interior solution, where $k \in \{j+1, \cdots, i-1\}$, or (iii) user i's own interior solution, each of which linearly increases in v by Equation (4.5). Therefore, $S_i(v, a_j)$ linearly increases in v

and $\partial S_i(v,a_j)/\partial v > 0$.

Therefore, the expected quality for user i, $Q_i(v,q_0)$, can be rewritten as:

$$Q_i(v,q_0) = \int_0^{\overrightarrow{a}_1(v,q_0)} T_i(v,q_0)^{(1)} f(a_1)da_1 + \int_{\overrightarrow{a}_1(v,q_0)}^1 S_i(v,a_1)f(a_1)da_1.$$

Expanding $T_i(v,q_0)^{(1)}$ and $S_i(v,a_1)$, we have:

$$\begin{aligned}
T_i(v,q_0)^{(1)} = {} & Q_i(q_i = 0 | q_1 \le q_0) \\
= {} & \int_0^{\overrightarrow{a}_2(v,q_0)} \int_0^{\overrightarrow{a}_3(v,q_0)} \cdots \int_0^{\overrightarrow{a}_{i-1}(v,q_0)} [\int_{\overleftarrow{a}_i(v,q_0)}^{\overrightarrow{a}_i(v,q_0)} q_0 f(a_i)da_i \\
& + \int_{\overrightarrow{a}_i(v,q_0)}^1 v q_i^* f(a_i)da_i] \dots f(a_2)da_2 \\
& + v \int_0^{\overrightarrow{a}_2(v,q_0)} \int_0^{\overrightarrow{a}_3(v,q_0)} \cdots \int_{\overrightarrow{a}_{i-1}(v,q_0)}^1 [\int_{\overleftarrow{a}_i(q_{i-1}^*)}^{\overrightarrow{a}_i(q_{i-1}^*)} q_{i-1}^* f(a_i)da_i \\
& + \int_{\overrightarrow{a}_i(q_{i-1})}^1 q_i^* f(a_i)da_i] \dots f(a_2)da_2 \\
& + \dots \\
& + v \int_{\overrightarrow{a}_2(v,q_0)}^1 \int_{\overrightarrow{a}_3(q_2^*)}^1 \cdots \int_{\overrightarrow{a}_{i-1}(q_{i-2}^*)}^1 [\int_{\overleftarrow{a}_i(q_{i-1}^*)}^{\overrightarrow{a}_i(q_{i-1}^*)} q_{i-1}^* f(a_i)da_i \\
& + \int_{\overrightarrow{a}_i(q_{i-1})}^1 q_i^* f(a_i)da_i] \dots f(a_2)da_2,
\end{aligned}$$

and

$$\begin{aligned}
S_i(v,a_1) = {} & Q_i(q_i | q_1 = v q_1^*) \\
= {} & v \int_0^{\overrightarrow{a}_2(q_1^*)} \int_0^{\overrightarrow{a}_3(q_1^*)} \cdots \int_0^{\overrightarrow{a}_{i-1}(q_1^*)} [\int_{\overleftarrow{a}_i(q_1^*)}^{\overrightarrow{a}_i(q_1^*)} q_1^* f(a_i)da_i \\
& + \int_{\overrightarrow{a}_i(q_1^*)}^1 q_i^* f(a_i)da_i] \dots f(a_2)da_2 \\
& + v \int_0^{\overrightarrow{a}_2(q_1^*)} \int_0^{\overrightarrow{a}_3(q_1^*)} \cdots \int_{\overrightarrow{a}_{i-1}(q_1^*)}^1 [\int_{\overleftarrow{a}_i(q_{i-1}^*)}^{\overrightarrow{a}_i(q_{i-1}^*)} q_{i-1}^* f(a_i)da_i \\
& + \int_{\overrightarrow{a}_i(q_{i-1}^*)}^1 q_i^* f(a_i)da_i] \dots f(a_2)da_2 \\
& + \dots \\
& + v \int_{\overrightarrow{a}_2(q_1^*)}^1 \int_{\overrightarrow{a}_3(q_2^*)}^1 \cdots \int_{\overrightarrow{a}_{i-1}(q_{i-2}^*)}^1 [\int_{\overleftarrow{a}_i(q_{i-1}^*)}^{\overrightarrow{a}_i(q_{i-1}^*)} q_{i-1}^* f(a_i)da_i
\end{aligned}$$

$$+\int_{\overrightarrow{a}_i(q_{i-1}^*)}^{1} q_i^* f(a_i)da_i]\dots f(a_2)da_2.$$

Using the Leibniz integral rule, we have:

$$\begin{aligned}
\frac{\partial Q_i(v,q_0)}{\partial v} &= \frac{\partial \overrightarrow{a}_1(v,q_0)}{\partial v} T_i(v,q_0)^{(1)} f\left(\overrightarrow{a}_1(v,q_0)\right) + \int_0^{\overrightarrow{a}_1(v,q_0)} \frac{\partial T_i(v,q_0)^{(1)}}{\partial v} f(a_1)da_1 \\
&\quad - \frac{\partial \overrightarrow{a}_1(v,q_0)}{\partial v} S_i\left(v, \overrightarrow{a}_1(v,q_0)\right) f\left(\overrightarrow{a}_1(v,q_0)\right) + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1. \\
&= \int_0^{\overrightarrow{a}_1(v,q_0)} \frac{\partial T_i(v,q_0)^{(1)}}{\partial v} f(a_1)da_1 + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1 \\
&= \frac{\partial T_i(v,q_0)^{(1)}}{\partial v} F\left(\overrightarrow{a}_1(v,q_0)\right) + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1.
\end{aligned}$$

The second equality obtains as the first and third terms cancel each other. The third equality obtains as $\{\partial T_i(v,q_0)^{(1)}\}/\{\partial v\}$ is independent of a_1. Moreover, by iteratively applying Equation (4.11), we obtain:

$$\begin{aligned}
\frac{\partial Q_i(v,q_0)}{\partial v} &= \frac{\partial T_i(v,q_0)^{(1)}}{\partial v} F\left(\overrightarrow{a}_1(v,q_0)\right) + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1 \\
&= \left[\frac{\partial T_i(v,q_0)^{(2)}}{\partial v} F\left(\overrightarrow{a_2}(v,q_0)\right) + \int_{\overrightarrow{a_2}(v,q_0)}^{1} \frac{\partial S_i(v,a_2)}{\partial v} f(a_2)da_2\right] F\left(\overrightarrow{a}_1(v,q_0)\right) \\
&\quad + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1 \\
&= \frac{\partial T_i(v,q_0)^{(2)}}{\partial v} F\left(\overrightarrow{a_2}(v,q_0)\right) F\left(\overrightarrow{a}_1(v,q_0)\right) \\
&\quad + F\left(\overrightarrow{a}_1(v,q_0)\right) \int_{\overrightarrow{a_2}(v,q_0)}^{1} \frac{\partial S_i(v,a_2)}{\partial v} f(a_2)da_2 + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1 \\
&= \frac{\partial T_i(v,q_0)^{(i-1)}}{\partial v} F\left(\overrightarrow{a}_{i-1}(v,q_0)\right) \cdots F\left(\overrightarrow{a}_1(v,q_0)\right) \\
&\quad + F\left(\overrightarrow{a}_{i-1}(v,q_0)\right) \cdots F\left(\overrightarrow{a}_1(v,q_0)\right) \int_{\overrightarrow{a}_{i-1}(v,q_0)}^{1} \frac{\partial S_i(v,a_{i-1})}{\partial v} f(a_{i-1})da_{i-1} \\
&\quad + \dots \\
&\quad + \int_{\overrightarrow{a}_1(v,q_0)}^{1} \frac{\partial S_i(v,a_1)}{\partial v} f(a_1)da_1.
\end{aligned}$$

As $\forall x > 0$, $F(x) > 0$ and $\{\partial S_i(v,a_j)\}/\{\partial v\} > 0$, the sign of $\{\partial Q_i(v,q_0)\}/\{\partial v\}$ depends on $\{\partial T_i(v,q_0)^{(i-1)}\}/\{\partial v\}$.

Applying the same technique used for user 1, we show below that $\{\partial T_i(v, q_0)^{(i-1)}\}/\{\partial v\} > 0$, i.e., $\forall\ v_2 > v_1$, $T_i(v_2, q_0)^{(i-1)} > T_i(v_1, q_0)^{(i-1)}$.

$$\begin{aligned}
T_i(v_1, q_0)^{(i-1)} &= \int_{\overleftarrow{a}_i(q_0,v_1)}^{\overrightarrow{a}_i(q_0,v_1)} q_0 f(a_i)da_i + \int_{\overrightarrow{a}_i(q_0,v_1)}^{1} v_1 q_i^* f(a_i)da_i \\
&= \int_{(\frac{q_0}{v_1})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}} q_0 f(a_i)da_i + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}}^{1} v_1 q_i^* f(a_i)da_i \\
&< \left[\int_{(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}} q_0 f(a_i)da_i + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}} q_0 f(a_i)da_i\right] \\
&\quad + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}}^{1} v_1 q_i^* f(a_i)da_i \\
&< \left[\int_{(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}} q_0 f(a_i)da_i + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}} v_2 q_i^* f(a_i)da_i\right] \\
&\quad + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}}^{1} v_2 q_i^* f(a_i)da_i \\
&= \int_{(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}} q_0 f(a_i)da_i \\
&\quad + \left[\int_{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}} v_2 q_i^* f(a_i)da_i + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_1})^{d_i}}^{1} v_2 q_i^* f(a_i)da_i\right] \\
&= \int_{(\frac{q_0}{v_2})^{d_i}}^{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}} q_0 f(a_i)da_i + \int_{\frac{1}{1-d_i}(\frac{q_0}{v_2})^{d_i}}^{1} v_2 q_i^* f(a_i)da_i \\
&= T_i(v_2, q_0)^{(i-1)}.
\end{aligned}$$

In summary, the expected quality, Q_i, strictly increases in the reward level, v. In particular, when $q_0 = 0$, the expected quality for each user linearly increases in v.

∎

Now we use a two-user sequential all-pay auction example to show the comparative statics of both the reward and reserve effect.

Example 2. *Using the parameters of Example 1, we add a reserve,* $0 < q_0 < v$.

The equilibrium bidding functions thus become:

$$q_1(a_1) = \begin{cases} 0 \ \text{ if } 0 \le a_1 < \sqrt{\frac{q_0}{v}}, \\ q_0 \ \text{ if } \sqrt{\frac{q_0}{v}} \le a_1 < 2\sqrt{\frac{q_0}{v}}, \\ \frac{a_1^2}{4} v \text{ if } 2\sqrt{\frac{q_0}{v}} \le a_1 \le 1. \end{cases}$$

Note, when $q_0 \geq \frac{v}{4}$, the third range of $q_1(a_1)$ does not exist.

If $0 \leq a_1 \leq 2\sqrt{\frac{q_0}{v}}$,

$$q_2(a_2) = \begin{cases} 0 \text{ if } 0 \leq a_2 < \frac{q_0}{v}, \\ q_0 \text{ if } \frac{q_0}{v} \leq a_2 \leq 1. \end{cases}$$

If $2\sqrt{\frac{q_0}{v}} \leq a_1 \leq 1$,

$$q_2(a_2) = \begin{cases} 0 \text{ if } 0 \leq a_2 < \frac{a_1^2}{4}, \\ \frac{a_1^2}{4}v \text{ if } \frac{a_1^2}{4} \leq a_2 \leq 1. \end{cases}$$

The probability that user 1 submits a positive bid becomes:

$$P_1(q_1 > 0) = 1 - F(\sqrt{\frac{q_0}{v}}) < 1.$$

When q_0 increases, $P_1(q_1 > 0)$ decreases. When v increases, $P_1(q_1 > 0)$ increases.

Next, the probability that user 2 participates, denoted as $P_2(q_2 > 0)$, becomes:

$$1 - [\int_0^{2\sqrt{\frac{q_0}{v}}} \int_0^{\frac{q_0}{v}} f(a_2)da_2 f(a_1)da_1 + \int_{2\sqrt{\frac{q_0}{v}}}^1 \int_0^{\frac{a_1^2}{4}} f(a_2)da_2 f(a_1)da_1]$$
$$\approx 1 - \left(0.94 * (\frac{q_0}{v})^{0.75} + \frac{1}{6}\right).$$

When q_0 increases, $P_2(q_2 > 0)$ decreases. When v increases, $P_2(q_2 > 0)$ increases.

Consequently, the expected quality for each user, the average and highest quality are characterized as follows:

$$\begin{aligned}
Q_1 &= \int_{\sqrt{\frac{q_0}{v}}}^{2\sqrt{\frac{q_0}{v}}} q_0 * 0.5a_1^{-0.5}da_1 + v\int_{2\sqrt{\frac{q_0}{v}}}^1 \frac{a_1^2}{4} * 0.5a_1^{-0.5}da_1 \approx 0.05v + 0.13\frac{q_0^{1.25}}{v^{0.25}}, \\
Q_2 &= \int_0^{2\sqrt{\frac{q_0}{v}}} \int_{\frac{q_0}{v}}^1 q_0 * 0.5a_2^{-0.5}da_2 0.5a_1^{-0.5}da_1 + v\int_{2\sqrt{\frac{q_0}{v}}}^1 \int_{\frac{a_1^2}{4}}^1 \frac{a_1^2}{4} * 0.5a_2^{-0.5}da_2 0.5a_1^{-0.5}da_1 \\
&\approx 0.03v + 1.13\frac{q_0^{1.25}}{v^{0.75}} - 1.2\frac{q_0^{1.75}}{v^{0.75}}, \\
AQ &= \frac{Q_1 + Q_2}{2} \approx \frac{1}{2}\left(0.08v + 0.13(\frac{q_0^{1.25}}{v^{0.25}}) + 1.13\frac{q_0^{1.25}}{v^{0.75}} - 1.2\frac{q_0^{1.75}}{v^{0.75}}\right), \\
HQ &= \int_0^{\sqrt{(\frac{q_0}{v})}} \int_{\frac{q_0}{v}}^1 q_0 f(a_2)da_2 f(a_1)da_1 + \int_{\sqrt{(\frac{q_0}{v})}}^{2\sqrt{(\frac{q_0}{v})}} \int_0^1 q_0 f(a_2)da_2 f(a_1)da_1
\end{aligned}$$

$$+\int_{2\sqrt{(\frac{q_0}{v})}}^{1}\int_0^1 v\frac{a_1^2}{4}f(a_2)da_2f(a_1)da_1$$
$$\approx 0.05v + 1.13\frac{q_0^{1.25}}{v^{0.25}} - \frac{q_0^{1.75}}{v^{0.75}}.$$

∎

Proof of Proposition 3:

The proof is similar to that for Proposition 1, so we omit it.

∎

Recall that, for Propositions 4 through 6, we assume that $H_i(x) = \prod_{j \neq i} F(x) = F^{n-1}(x)$ is strictly concave and that $H_i(0) = 0$. In addition, we do not assume that $F(x) = x^c$.

Proof of Proposition 4 and 6:

Case 1: Zero Reserve. We first derive the equilibrium bidding function for each user, when the reserve is zero, i.e., $q_0 = 0$. We do so by solving the following maximization problem for each user i:

$$\max_{q_i} \{v \prod_{j \neq i} F_j(q_j < q_i) - \frac{q_i}{a_i}\}. \tag{4.12}$$

As we characterize a symmetric equilibrium, we omit i in subsequent proofs. We define the inverse of $q(a)$ as $a(q)$.

$$\max_{q} \{vH(a(q)) - \frac{q}{a}\}. \tag{4.13}$$

As $H(a(q)) = F^{n-1}(a(q))$, the first-order condition is:

$$vf\left(a(q)\right)(n-1)F^{n-2}\left(a(q)\right)a'(q) - \frac{1}{a} = 0.$$

As $a(q)$ is the inverse of $q(a)$, then $a'(q) = \frac{1}{q'(a)}$ and we have:

$$q'(a) = a(n-1)vF^{n-2}(a)f(a).$$

We next integrate $q'(a)$ from 0 to a:

$$q(a) = (n-1)v \int_0^a sF^{n-2}(s)f(s)ds + C.$$

As $q(0) = 0$, we have $C = 0$. Therefore, the equilibrium bidding function becomes:

$$\begin{aligned} q(a) &= (n-1)v \int_0^a sF^{n-2}(s)dF(s) \\ &= v\left[aF^{n-1}(a) - \int_0^a F^{n-1}(s)ds\right]. \end{aligned} \tag{4.14}$$

The second-order condition is satisfied by following the same proof in Moldovanu and Sela (2001).

By Equation (4.14), $q(a) \geq 0$, $\forall\ a > 0$. Additionally, $a = 0$ is a measure zero event. Therefore, with a zero-reserve, the probability of participation for any user i, $1 - P_i(q_i = 0)$, is 1 in simultaneous all-pay auctions.

Case 2: Positive Reserve. We now consider the positive reserve case, i.e., $q_0 > 0$.

When $q_0 > 0$, we solve the same maximization problem as (4.12) with an additional constraint, $q_i \geq q_0$.

$$\begin{aligned} &\max_{q_i} \{v \textstyle\prod_{j \neq i} F_j(q_j < q_i) - \frac{q_i}{a_i}\} \\ &s.t. \qquad\qquad q_i \geq q_0. \end{aligned} \tag{4.15}$$

To characterize the equilibrium bidding function, we define two boundaries as:

$$\overleftarrow{a} = \frac{\frac{q_0}{v}}{H(\frac{q_0}{v})}, \text{ and } q_0 = v[\overrightarrow{a} F^{n-1}(\overrightarrow{a}) - \int_0^{\overrightarrow{a}} F^{n-1}(s)ds].$$

These boundaries partition the support of abilities into three ranges:

1. When $0 \leq a < \overleftarrow{a}$, the expected payoff from submitting a positive bid is negative. Thus, the user should submit a zero bid.
2. When $\overleftarrow{a} \leq a < \overrightarrow{a}$, as $q_0 > v\left[aF^{n-1}(a) - \int_0^a F^{n-1}(s)ds\right]$, bidding q_0 dominates $v[aF^{n-1}(a)$ $- \int_0^a F^{n-1}(s)ds]$. Therefore, the constraint is binding, and we obtain a corner solution.
3. When $\overrightarrow{a} \leq a \leq 1$, Equation (4.14) is the interior solution of the constrained optimization problem (4.15) while the constraint is not binding.

Summarizing the above analysis, we characterize the equilibrium bidding function for user i as follows:

$$q(a) = \begin{cases} 0 & \text{if } 0 \leq a \leq \overleftarrow{a}, \\ q_0 & \text{if } \overleftarrow{a} \leq a \leq \overrightarrow{a}, \\ v[aF^{n-1}(a) - \int_0^a F^{n-1}(s)ds] & \text{if } \overrightarrow{a} \leq a \leq 1, \end{cases} \tag{4.16}$$

Note that when $q_0 > v[1 - \int_0^1 F^{n-1}(s)ds]$, the third range of Equation (4.16) does not exist.

Now we examine the reward and the reserve effect on participation in simultaneous all-pay auctions, i.e., $P(q = 0) = F(\overleftarrow{a}) = F(\frac{q_0/v}{H(q_0/v)})$ strictly decreases in v and strictly increases in q_0.

Defining $Z(a) \equiv \frac{a}{H(a)}$, we first show $Z(a)$ strictly increases in a.

Differentiating Z(a) w.r.t. a, we obtain:

$$\frac{dZ(a)}{da} = \frac{H(a) - aH'(a)}{H^2(a)}.$$

As $H(a)$ is strictly concave and $H(0) = 0$, $\forall\ a > 0$, we have $H(a) > aH'(a)$. Therefore,

$$\frac{dZ(a)}{da} = \frac{H(a) - aH'(a)}{H^2(a)} > 0.$$

Consequently, $Z(a)$ strictly increases in a. Moreover, as $F(x)$ also strictly increases in x, when v increases, $F(\overleftarrow{a})$ strictly decreases. Therefore, the probability of participation, $P(q > 0)$, strictly increases with v.

Similarly, when q_0 increases, $F(\overleftarrow{a})$ strictly increases and the probability of participation, $P(q > 0)$, strictly decreases.

∎

Proof of Proposition 5:

Using Equation (4.16), the expected quality for each user i in simultaneous all-pay auctions is

$$Q = \int_{\overleftarrow{a}}^{\overrightarrow{a}} q_0 f(s)ds + \int_{\overrightarrow{a}}^{1} v \left[aH(a) - \int_0^a H(s)ds\right] f(s)ds,$$

where $H(s) = F^{n-1}(s)$.

When the reward size, v, increases, $\overleftarrow{a}$ strictly deceases by following the proof of Proposition 4. Now we show that $\overrightarrow{a}$ also strictly decreases in v.

As $q_0 = v[\overrightarrow{a} F^{n-1}(\overrightarrow{a}) - \int_0^{\overrightarrow{a}} F^{n-1}(s)ds]$, we obtain:

$$\frac{q_0}{v} = \overrightarrow{a} F^{n-1}(\overrightarrow{a}) - \int_0^{\overrightarrow{a}} F^{n-1}(s)ds.$$

Define $M(\overrightarrow{a}) \equiv \frac{q_0}{v}$, which has a corresponding inverse function $\overrightarrow{a} \equiv M^{-1}(\frac{q_0}{v})$. We first show that $M(\overrightarrow{a})$ is strictly increases in $\overrightarrow{a}$.

Applying the Leibniz integral rule and differentiating $M(\overrightarrow{a})$ w.r.t. $\overrightarrow{a}$, we obtain:

$$\frac{dM(\overrightarrow{a})}{d\overrightarrow{a}} = (n-1)\overrightarrow{a} F^{n-2}(\overrightarrow{a}) f(\overrightarrow{a}) > 0.$$

As $M(\overrightarrow{a})$ strictly increases in $\overrightarrow{a}$, the inverse function $\overrightarrow{a} \equiv M^{-1}(\frac{q_0}{v})$ also strictly increases in $t \equiv \frac{q_0}{v}$. Therefore, when v increases, t strictly decreases and $\overrightarrow{a}$ strictly decreases.

Similar to the proof for user 1's expected quality in sequential all-pay auctions

(Proposition 2), the expected quality for each user in simultaneous all-pay auctions strictly increases in v.

$\blacksquare$

4.10 Appendix B: Sample Tasks and List of Taskcn IDs and URLs for All Tasks

We provide sample translation tasks for both the personal statements and company introductions, with excerpts of a reserve submission and a machine translation for each task. For the programming tasks, we also provide a sample task and the corresponding solutions. For each task used in our experiment, we provide the complete list of Taskcn IDs and URLs.[21] *Interested readers can browse each task and the corresponding solutions from the Taskcn online archive by directly clicking on the URLs or by entering the TaskID from the search window on* `http://www.taskcn.com/`.

1. **Sample Personal Statement**
   (a) **TaskID 40883** (excerpt)

   个人申请文书

   带着与生俱来的好奇心，从社会现象，科学迷题，到现代技术，

   我对很多事物都充满了好奇。也因为这样，我选择了信息管理

   系统作为本科专业，并在数学，计算机科学以及其他社会科学

   方面得到了严格的训练。。。。。。

   (b) **Reserve Submission**
   Born with strong curiosity about the world, I am interested in diverse topics including social phenomena, scientific puzzles, and modern technologies. Therefore, I chose Management Information Systems as my undergraduate major, in which I received rigorous training in mathematics, computer science and other social sciences.

   (c) **Machine Translation**
   With innate curiosity, from a social phenomenon, the scientific puzzle to modern technology, I have a lot of things are full of curiosity. Also for this reason that I chose the information management system as a degree, get a rigorous mathematics, computer science and other social science training.

[21] The URLs were effective as of September 25, 2011.

2. **Sample Company Introduction**
   (a) **TaskID 40614** (excerpt)

   公司简介

   tp 汽车保险股份有限公司是 2004 年 12 月经中国保险监督管理委员会批准设立的全国性金融机构，是中国第一家专业汽车保险公司。公司总部设在上海浦东陆家嘴金融区，注册资本 5.5 亿元人民币，主要经营机动车交通事故责任强制保险和机动车商业保险，同时还经营企业财产险、家财险、货运险、责任险、短期意外险和健康险等业务。。。。。。

   (b) **Reserve Submission**
   TP Auto Insurance Co., Ltd. is a national financial institute approved by China Insurance Regulatory Commission on December 2004. It is the first professional Chinese Auto insurance company. The headquarters are in the Pu Dong Lu Jiazui financial district in Shanghai, with a registered capital of 550 million RMB. The company mainly operates Compulsory Traffic Accident Liability Insurance for Motor Vehicles and Commercial Insurance for Motor Vehicles. It also operates Enterprise Property Insurance, Family Property Insurance, Shipping insurance, Liability Insurance, Short-time Accident Insurance and Health Insurance, etc.

   (c) **Machine Translation**
   TP Automobile Insurance Company is a 12 period in 2004 the China Insurance Regulatory Commission approved the establishment of a national financial institutions, is China's first professional automobile insurance. The company is headquartered in Shanghai Pudong Lujiazui financial district, the registered capital of 550 million yuan, mainly engaged in the compulsory motor vehicle traffic accident liability insurance and commercial insurance of motor vehicles, as well as property insurance enterprises, home Insurance, cargo insurance, liability insurance short-term accident insurance and health insurance services.

3. **Sample Programming Task**
   (a) **TaskID 40707**
   Website needs a password security checking function. Show input characters as encoded dots when user types password. Generate an information bar to indicate the security level of the password, considering these factors:
      i. length of the password;
      ii. mixture of numbers and characters;
      iii. mixture of upper and lower case letters;
      iv. mixture of other symbols.

   Please provide source code and html for testing.

(b) The sample solution can be found on the first author's website: http://sitemaker.umich.edu/liuxiao/files/spt.pdf.

4. **Taskcn IDs and URLs for All Translation Tasks**
   (a) The High-Reward-No-Reserve Treatment
   40570 http://www.taskcn.com/w-40570.html;
   41106 http://www.taskcn.com/w-41106.html;
   40627 http://www.taskcn.com/w-40627.html;
   41211 http://www.taskcn.com/w-41211.html;
   40678 http://www.taskcn.com/w-40678.html;
   41232 http://www.taskcn.com/w-41232.html;
   40766 http://www.taskcn.com/w-40766.html;
   41289 http://www.taskcn.com/w-41289.html;
   40820 http://www.taskcn.com/w-40820.html;
   41356 http://www.taskcn.com/w-41356.html;
   40855 http://www.taskcn.com/w-40855.html;
   41388 http://www.taskcn.com/w-41388.html;
   40896 http://www.taskcn.com/w-40896.html;
   41460 http://www.taskcn.com/w-41460.html;
   40993 http://www.taskcn.com/w-40993.html;
   41513 http://www.taskcn.com/w-41513.html;
   41034 http://www.taskcn.com/w-41034.html;
   41567 http://www.taskcn.com/w-41567.html;
   41068 http://www.taskcn.com/w-41068.html;
   41623 http://www.taskcn.com/w-41623.html.
   (b) The High-Reward-Reserve-Without-Credit Treatment
   40614 http://www.taskcn.com/w-40614.html;
   41115 http://www.taskcn.com/w-41115.html;
   40650 http://www.taskcn.com/w-40650.html;
   41156 http://www.taskcn.com/w-41156.html;
   40694 http://www.taskcn.com/w-40694.html;
   41243 http://www.taskcn.com/w-41243.html;
   40761 http://www.taskcn.com/w-40761.html;
   41282 http://www.taskcn.com/w-41282.html;
   40812 http://www.taskcn.com/w-40812.html;
   41353 http://www.taskcn.com/w-41353.html;
   40883 http://www.taskcn.com/w-40883.html;
   41393 http://www.taskcn.com/w-41393.html;
   40940 http://www.taskcn.com/w-40940.html;
   41427 http://www.taskcn.com/w-41427.html;
   40991 http://www.taskcn.com/w-40991.html;
   41491 http://www.taskcn.com/w-41491.html;
   41015 http://www.taskcn.com/w-41015.html;
   41548 http://www.taskcn.com/w-41548.html;
   41055 http://www.taskcn.com/w-41055.html;
   41596 http://www.taskcn.com/w-41596.html.

(c) The High-Reward-Reserve-With-Credit Treatment
40612 `http://www.taskcn.com/w-40612.html`;
41103 `http://www.taskcn.com/w-41103.html`;
40646 `http://www.taskcn.com/w-40646.html`;
41175 `http://www.taskcn.com/w-41175.html`;
40695 `http://www.taskcn.com/w-40695.html`;
41235 `http://www.taskcn.com/w-41235.html`;
40764 `http://www.taskcn.com/w-40764.html`;
41294 `http://www.taskcn.com/w-41294.html`;
40816 `http://www.taskcn.com/w-40816.html`;
41360 `http://www.taskcn.com/w-41360.html`;
40863 `http://www.taskcn.com/w-40863.html`;
41384 `http://www.taskcn.com/w-41384.html`;
40919 `http://www.taskcn.com/w-40919.html`;
41430 `http://www.taskcn.com/w-41430.html`;
40985 `http://www.taskcn.com/w-40985.html`;
41470 `http://www.taskcn.com/w-41470.html`;
41008 `http://www.taskcn.com/w-41008.html`;
41534 `http://www.taskcn.com/w-41534.html`;
41046 `http://www.taskcn.com/w-41046.html`;
41606 `http://www.taskcn.com/w-41606.html`.

(d) The Low-Reward-No-Reserve Treatment
40585 `http://www.taskcn.com/w-40585.html`;
41176 `http://www.taskcn.com/w-41176.html`;
40673 `http://www.taskcn.com/w-40673.html`;
41199 `http://www.taskcn.com/w-41199.html`;
40699 `http://www.taskcn.com/w-40699.html`;
41261 `http://www.taskcn.com/w-41261.html`;
40765 `http://www.taskcn.com/w-40765.html`;
41291 `http://www.taskcn.com/w-41291.html`;
40826 `http://www.taskcn.com/w-40826.html`;
41364 `http://www.taskcn.com/w-41364.html`;
40897 `http://www.taskcn.com/w-40897.html`;
41385 `http://www.taskcn.com/w-41385.html`;
40945 `http://www.taskcn.com/w-40945.html`;
41459 `http://www.taskcn.com/w-41459.html`;
40995 `http://www.taskcn.com/w-40995.html`;
41509 `http://www.taskcn.com/w-41509.html`;
41035 `http://www.taskcn.com/w-41035.html`;
41566 `http://www.taskcn.com/w-41566.html`;
41078 `http://www.taskcn.com/w-41078.html`;
41636 `http://www.taskcn.com/w-41636.html`.

(e) The Low-Reward-Reserve-Without-Credit Treatment
40591 `http://www.taskcn.com/w-40591.html`;
41123 `http://www.taskcn.com/w-41123.html`;

40663 http://www.taskcn.com/w-40663.html;
41190 http://www.taskcn.com/w-41190.html;
40704 http://www.taskcn.com/w-40704.html;
41234 http://www.taskcn.com/w-41234.html;
40759 http://www.taskcn.com/w-40759.html;
41284 http://www.taskcn.com/w-41284.html;
40814 http://www.taskcn.com/w-40814.html;
41336 http://www.taskcn.com/w-41336.html;
40882 http://www.taskcn.com/w-40882.html;
41410 http://www.taskcn.com/w-41410.html;
40939 http://www.taskcn.com/w-40939.html;
41439 http://www.taskcn.com/w-41439.html;
40988 http://www.taskcn.com/w-40988.html;
41492 http://www.taskcn.com/w-41492.html;
41023 http://www.taskcn.com/w-41023.html;
41533 http://www.taskcn.com/w-41533.html;
41065 http://www.taskcn.com/w-41065.html;
41610 http://www.taskcn.com/w-41610.html;

(f) The Low-Reward-Reserve-With-Credit Treatment

40625 http://www.taskcn.com/w-40625.html;
41111 http://www.taskcn.com/w-41111.html;
40643 http://www.taskcn.com/w-40643.html;
41171 http://www.taskcn.com/w-41171.html;
40691 http://www.taskcn.com/w-40691.html;
41242 http://www.taskcn.com/w-41242.html;
40754 http://www.taskcn.com/w-40754.html;
41288 http://www.taskcn.com/w-41288.html;
40822 http://www.taskcn.com/w-40822.html;
41343 http://www.taskcn.com/w-41343.html;
40793 http://www.taskcn.com/w-40793.html;
41383 http://www.taskcn.com/w-41383.html;
40941 http://www.taskcn.com/w-40941.html;
41431 http://www.taskcn.com/w-41431.html;
40981 http://www.taskcn.com/w-40981.html;
41473 http://www.taskcn.com/w-41473.html;
41029 http://www.taskcn.com/w-41029.html;
41526 http://www.taskcn.com/w-41526.html;
41056 http://www.taskcn.com/w-41056.html;
41576 http://www.taskcn.com/w-41576.html.

5. **Taskcn IDs and URLs for All Programming Tasks**
   (a) The High-Reward Treatment
   40599 `http://www.taskcn.com/w-40599.html`;
   41053 `http://www.taskcn.com/w-41053.html`;
   40652 `http://www.taskcn.com/w-40652.html`;
   41142 `http://www.taskcn.com/w-41142.html`;
   40707 `http://www.taskcn.com/w-40707.html`;
   41423 `http://www.taskcn.com/w-41423.html`;
   40778 `http://www.taskcn.com/w-40778.html`;
   41454 `http://www.taskcn.com/w-41454.html`;
   40846 `http://www.taskcn.com/w-40846.html`;
   41519 `http://www.taskcn.com/w-41519.html`;
   40904 `http://www.taskcn.com/w-40904.html`;
   41664 `http://www.taskcn.com/w-41664.html`;
   40999 `http://www.taskcn.com/w-40999.html`;
   42091 `http://www.taskcn.com/w-42091.html`.
   (b) The Low-Reward Treatment
   40654 `http://www.taskcn.com/w-40654.html`;
   41144 `http://www.taskcn.com/w-41144.html`;
   40726 `http://www.taskcn.com/w-40726.html`;
   41424 `http://www.taskcn.com/w-41424.html`;
   40780 `http://www.taskcn.com/w-40780.html`;
   41456 `http://www.taskcn.com/w-41456.html`;
   40848 `http://www.taskcn.com/w-40848.html`;
   41574 `http://www.taskcn.com/w-41574.html`;
   40959 `http://www.taskcn.com/w-40959.html`;
   41665 `http://www.taskcn.com/w-41665.html`;
   41000 `http://www.taskcn.com/w-41000.html`;
   41983 `http://www.taskcn.com/w-41983.html`;
   41054 `http://www.taskcn.com/w-41054.html`;
   42092 `http://www.taskcn.com/w-42092.html`.

4.11 Appendix C: Rating Instructions

To improve the reliability of students' ratings, we conducted training sessions before the rating sessions began. For the translation tasks, we gave raters one sample personal statement and company introduction, then asked them to rate the difficulty of both questions.[22] *We also gave them two submissions for each task and asked them to rate the quality of each submission. One of the submissions was written by either the personal statement provider or our two undergraduate research assistants, while the other was randomly drawn from the submissions that we received from the pilot session. For the programming task, we followed the same procedure with two sample tasks. In addition, to help raters develop and refine their own personal rating scales, we asked them to individually give reasons for their rating scores for each task-submission pair.*

C.1. Translations

All translation raters were asked to provide ratings for the following items for each task-submission pair:

1. Please rate the question for the following factors:
   (a) Please rate the effort level in terms of time needed for a proficient translator.
   (0: 0-0.5 hour; ...; 10: 5-7 days)
   (b) It requires deep understanding of a specific field.
   (1 = strongly disagree; ...; 7 = strongly agree)
   (c) It requires highly advanced English writing skills.
   (1 = strongly disagree; ...; 7 = strongly agree)
   (d) Please rate the overall translation difficulty of the original text.
   (1 = very easy; ...; 7 = very difficult)
2. Please rate the answer for the following factors: (1 = strongly disagree; ...; 7 = strongly agree)
   (a) Overall, the translation is accurate.
   (b) The translation is complete.
   (c) The translator has a complete and sufficient understanding of the original document.
   (d) The translation is coherent and cohesive (it can be smoothly read).
   (e) The translation properly conforms to the correct usage of English expression.
3. Please rate the overall quality of this translation work.
   (1 = very low quality; ...; 7 = very high quality.)

[22] These two tasks were used in the pilot session before the experiment. The purpose of the pilot session was to check the reward and task duration parameters.

C.2. Programming

For the programming tasks, raters were asked to rate the following items for each task-submission pair:

1. Please rate the task for the following factors:
   (a) Please rate the task by the level of expertise it requires to fulfill the task description:
      1: The task requires minimal knowledge and expertise in programming in the language. A person with normal college education can accomplish it without training.
      2: ...
      3: ...
      4: The task requires substantial knowledge and expertise comparable to that of a trained programmer with 2-3 years of relevant programming experience in the language.
      5: ...
      6: ...
      7: The task requires very high level of knowledge and expertise that professional expert would have. The expert should have deep and comprehensive understanding on the philosophy of the language, as well as more than 5 years of professional experience.
   (b) Please rate the task on the required effort level in terms of time needed for a trained programmer to accomplish the task as described. A trained programmer is defined as someone with 2 - 3 years of programming experience with Javascript or other language as required. The work can be done within (including everything such as coding, testing, packing etc.):
      0: 0-0.5 hour;
      1: 0.5 - 1 hour;
      2: 1 - 2 hours;
      3: 2 - 3 hours;
      4: 3 - 5 hours;
      5: 5 - 8 hours;
      6: 8 - 12 hours;
      7: 12 - 24 hours;
      8: 2 - 3 days;
      9: 4 - 5 days;
      10: 5 - 7 days.
2. Please rate the solution for the following factors:
   (a) **Functionality:** Please rate the solution by the degree to which it realized the function requirement as the task description. (1-7)
      1: The solution does not realize any of the required functions.
      2: ...
      3: ...
      4: The solution realizes most of the required functions.
      5: ...

6: ...
7: The solution not only realizes all required functions, but also enhances some important functions beyond the requirement, and presents thoughtful considerations.

(b) **Programming professionalism and skill:** Please rate the solution in terms of its methods, structure, and terminology involved in design, which can be directly reflected as its readability, extendability, and testability:
1: The solution shows total novice.
2: ...
3: ...
4: The solution presents basic considerations above all three perspectives. Professional skills are employed in the major areas of the coding process.
5: ...
6: ...
7: The solution is a master piece in terms of professionalism.

(c) **Time:** Please rate the solution on the effort level in terms of how much time a trained programmer needs to accomplish the present solution. A trained programmer is defined as someone with 2-3 years of programming experience with Javascript or other language as required. The work can be done within (including everything such as coding, testing, packing etc.)
0: 0-0.5 hour;
1: 0.5 - 1 hour;
2: 1 - 2 hours;
3: 2 - 3 hours;
4: 3 - 5 hours;
5: 5 - 8 hours;
6: 8 - 12 hours;
7: 12 - 24 hours;
8: 2 - 3 days;
9: 4 - 5 days;
10: 5 - 7 days.

(d) **Overall Quality:** Please rate the overall quality of this programming work.
(1 = very low quality; ...; 7 = very high quality)

Chapter 5

Endogenously Sequential v. Simultaneous All-Pay Auctions: an Experimental Study

5.1 Introduction

All-pay auctions have been used to model such diverse activities as political lobbying, research and development (R&D) races, and tournament contests. Most theoretical and experimental work on all-pay auctions concerns simultaneous all-pay auctions where bidders bid simultaneously and independently. However, in many real life applications of all-pay auctions, individuals or groups move sequentially instead of simultaneously, and a player who enters late can decide how much effort to expend after observing others' behavior. For example, the patent competition between an incumbent and an entrant is sequential (Leininger, 1991). Similarly, in the United States presidential election process, the sitting president and his affiliated party hold their convention later than does the opposing party (Morgan, 2003). Therefore, it is interesting to explore individual behavior in sequential all-pay auctions.

Furthermore, whether sequential or simultaneous all-pay auctions can induce greater efforts from participants is an open question for mechanism designers or decision makers. For instance, both these two all-pay auctions have been used as mechanisms for crowdsourcing contest sites, which we call all-pay auction crowdsourcing markets. Specifically, on these websites, a task is posted by a requester along with a reward for the best solution; any user on the site may submit a solution. The requester then selects the best solution and rewards the corresponding user, however, other users' efforts in the task are not compensated. If we equate users' efforts with bids in an auction, such a system is equivalent to an all-pay auction. While these sites have been successful in encouraging use, which mechanism is the most effective one for encouraging effort is under experimentation. Many crowdsourcing sites, e.g., Taskcn.com, implement sequential all-pay auctions where users submit their solutions sequentially and thus late entrants can observe the content of prior solutions. In contrast, other sites such as Topcoder.com use simultaneous all-pay auctions which prevent users from reading each other's submissions. Recently, Taskcn uses a fee-based solution protection program. When all submissions are password-protected, it converts the sequential all-pay auction to simultaneous all-pay auction. Although it is difficult to compare the performance of these different mechanisms in the field, laboratory experiments can be used to test their performance in a controlled setting. Thus, the results from this laboratory study can shed light on the design of crowdsourcing markets.

Compared to most all-pay auction models, where bidders' entry timing are exogenous, users in crowdsourcing markets endogenously decide to enter early or late. For instance, in a field experiment conducted on Taskcn, the results show that experienced users submit their solutions significantly later than do others. Additionally, the early entry of a high quality solution can deter the entry of experienced users (Liu, Yang, Adamic and Chen, 2011). Therefore, this paper studies individual entry behavior based on a endogenously sequential all-pay auction model. In addition, given the observation that many requesters on Taskcn explicitly announce that the early submission will be selected if there are multiple best solutions,[1] this study also investigates whether this simple favor-early tie-breaking rule is an effective tool to increase the ratio of early entries.

Altogether, whereas most experimental work in the all-pay auction literature investigates simultaneous all-pay auctions, this study is the first laboratory experiment which reports individual behavior in sequential all-pay auctions and compares their performance with that of simultaneous all-pay auctions. In this study, we observe lower or equivalent average bids in sequential all-pay auctions compared to theoretical predictions, while we observe that the average bids in simultaneous all-pay auctions are either greater than or equal to theoretical predictions. Consequently, we find that the revenue in sequential all-pay auctions is lower than that in simultaneous all-pay auctions. In comparison, we find that the ratio of efficient allocation and the average earnings are higher in sequential all-pay auctions than in simultaneous all-pay auctions. In summary, these results suggest that simultaneous all-pay auctions are more effective than sequential all-pay auctions in aggregating effort and knowledge from users in corwdsourcing markets. Finally, individuals are more likely to enter the

[1] Source: http://www.taskcn.com/w-60017.html, retrieved on October 15, 2011.

early bidding stage when a favor-early tie-breaking rule is present, but this effect is attenuated as individuals gain experience.

The rest of this paper is organized as follows. In Section 5.2, we review the relevant literature. Section 5.3 presents a theoretical model of endogenously sequential all-pay auctions. Section 5.4 presents the experimental design. Section 5.5 presents our hypotheses. In Section 5.6, we present our analysis and results. Section 5.7 concludes.

5.2 Literature Review

In the auction literature, most studies of all-pay auctions explore simultaneous all-pay auctions where bidders submit their bids independently and simultaneously.[2] For complete information all-pay auctions where each bidder's value (or bidding cost) is common knowledge, Baye, Kovenock and de Vries (1996) provide a complete characterization of the mixed strategy Nash equilibria, where total bids are expected to be less than or equal to the value of the object. In contrast to this theoretical prediction, a series of experiments (Davis and Reilly, 1998; Gneezy and Smorodinsky, 2006; Lugovskyy, Puzzello and Tucker, 2010) report overdissipation where the total bids exceed the value of the object at the aggregate level. This phenomenon is attributed to various factors, including (1) the size of the group, (2) individual experiences, and (3) the matching protocol.[3] In an experimental study with a minimum group size (2 players), a sufficient learning opportunity (30 rounds), and a random rematching protocol, Potters, de Vries and van Winden (1998) find an average bid consistent with the Nash equilibrium prediction. Specifically, at the individual level, they find that bidders do not use a mixed strategy, as predicted by the risk-neutral model. Furthermore, they find that bidders can be categorized into different types. For example, a substantial proportion of bidders choose to overweight strategies with higher realized payoffs in previous rounds; consequently, they are more likely to use the same strategy across rounds instead of randomizing their bids. In complete information all-pay auctions where bidders have different bidding costs, Anderson and Stafford (2003) find that the bids from low-cost bidders are significantly higher than those

[2]There is extensive literature on lottery contests where the winning probability is not deterministic but proportional to the number of bids (Tullock, 1980). We refer the reader to Sheremeta (2011) for a summary of this literature and the references therein.

[3]Assuming individuals are bounded rational, Anderson, Goeree and Holt (1998) characterize the existence of overdissipation in logit equilibria for complete-information all-pay auctions.

from high-cost bidders, indicating that bidders with a higher winning probability are likely to bid more aggressively.

A series of studies have also examined simultaneous all-pay auctions with incomplete information. Assuming individuals are risk neutral, Krishna and Morgan (1997) characterize the symmetric Bayesian Nash equilibrium where bidders have a common value distribution. By relaxing the assumption that everyone has the same bidding distribution, Amann and Leininger (1996) prove the existence and uniqueness of the Bayesian Nash equilibrium for two bidders. Consistent with results in the complete information all-pay auction experiments, overdissipation is also observed to decrease with experience in the incomplete information all-pay auction context(Noussair and Silver, 2006). Finally, studies of simultaneous all-pay auctions with incomplete information find that bidders with a low value often bid lower than the equilibrium prediction whereas high-value bidders tend to bid higher than the equilibrium prediction (Noussair and Silver, 2006).[4]

Relative to the extensive literature on simultaneous all-pay auctions, there are few current studies on sequential all-pay auctions. Konrad and Leininger (2007) characterize the Subgame Perfect Nash Equilibrium (SPNE) in a endogenously sequential all-pay auction with complete-information. In the model, bidders first endogenously choose to enter between the early and late bidding stages before the auction starts, and when all bidders enter the same stage, it is a simultaneous all-pay auction. Otherwise, it is a sequential all-pay auction where late bidders bid after observing early bidders' bids. Segev and Sela (2011) derive the perfect equilibrium of exogenously sequential all-pay auctions with incomplete information where bidders' entry timing is exogenously determined. Extending the theoretical framework of Segev and Sela (2011), Liu et al. (2011) characterize the reward and reserve-price effects on participation and contribution quality in both incomplete information sequential and simultaneous all-pay auctions and test the predictions in a field experiment on Taskcn. In particular, they find that experienced users are more likely to enter late compared to inexperienced users and conjecture that experienced users may either need more time to work on submissions or strategically wait to observe others' submissions.

Building upon the current literature on all-pay auctions, this study provides the first experimental study to examine individual behavior in sequential all-pay auctions and compare their performances with simultaneous all-pay auctions under complete information. Moreover, in contrast to previous experimental literature where the en-

[4]The low value of an object is equivalent to a high bidding cost and the high value of an object is equivalent to a low bidding cost.

try decision of an individual is exogenous, we allow individuals to choose entry timing in our endogenous-entry treatments and examine the entry timing dynamic by controlling for individual heterogeneity in abilities and others. Therefore, this design is more closely aligned with many real-life situations.

5.3 A Theoretical Framework of Endogenously Sequential All-Pay Auctions

In this section, we first outline a theoretical framework to characterize the equilibria for endogenously sequential all-pay auctions with complete information. In doing so, we follow both the model and notation in Konrad and Leininger (2007), which characterizes the SPNE with heterogeneous bidders. We also consider homogeneous bidders to show that different tie-breaking rules affect individual entry decisions in the equilibrium.[5] This complete model captures competitive ability differences among contestants in crowdsourcing and tournaments. For instance, all-pay auctions with homogeneous bidders characterize a highly competitive environment whereas those with heterogeneous bidders represent a less competitive environment. Additionally, we discuss the relevant theoretical predictions in Baye et al. (1996) that apply to simultaneous all-pay auctions.

In Konrad and Leininger's (2007) model, prior to the auction, each bidder first decides between entering the early and late bidding stages independently and simultaneously and then announces the entry decision to other bidders. In this scenario, let E be the set of bidders who enter early with cardinality $\#E$ and L be the set of bidders who enter late with cardinality $\#L$. The probability of entering the early bidding stage is q_i for bidder i. Individual entry decisions jointly determine the structure of the subsequent all-pay auctions. In the extreme case where everyone enters at the same stage, the subsequent game becomes a simultaneous all-pay auction.

After all bidders have entered the auction, an object with value v is auctioned among n bidders. Each bidder i has a linear cost function, $c_i(x_i) = c_i x_i$, where x_i is bidder i's bid and $\bar{x}_i$ is determined by $c_i \bar{x}_i = v$. Without loss of generality, we assume that $c_1 \leq c_2 \leq ... \leq c_n$. In a simultaneous all-pay auction, everyone enters at the same bidding stage, whereas in a sequential all-pay auction, bidders in the late stage submit their bids after observing early bids. Assuming everyone is risk-neutral,

[5] In contrast with homogenous bidders, the SPNE for heterogeneous bidders does not vary with tie-breaking rules (Konrad and Leininger, 2007).

Table 5.1 The Entry Game with a Favor-Early Tie-Breaking Rule: Homogeneous Bidders

		Bidder 2	
		Early	Late
Bidder 1	Early	0, 0	0, 0
	Late	0, 0	0, 0

conditional on other bidders' bids x_{-i}, bidder i's payoff, $\pi_i(x_i, x_{-i})$ is:

$$\pi_i(x_i, x_{-i}) = \begin{cases} v - c_i x_i \text{ if } x_i > \max x_{-i} \\ \quad -c_i x_i \quad \text{if } x_i < \max x_{-i} \end{cases}$$

In the case of a tie, we now discuss both the favor-early and favor-late tie-breaking rules. The favor-early tie-breaking rule states that, when there are multiple highest bids, the winner is randomly selected from the group of early bidders with the highest bids. By contrast, the favor-late tie-breaking rule randomly selects the winner from the group of late bidders with the highest bids. Note that when all of the highest bids are from bidders in the same stage, the winner is randomly selected with equal probability.[6] We next derive the SPNE in sequential all-pay auctions for homogeneous bidders ($c_1 = c_2 = ... = c_n \equiv c$) to show that different tie-breaking rules yield different equilibrium predictions. Specifically, a tie-breaking rule that favors early entries leads to the following proposition.

Proposition 7 (Sequential All-Pay Auctions with a Favor-Early Tie-Breaking Rule: Homogeneous Bidders)**.**

1. *In a sequential all-pay auction, the unique SPNE is that one of the early bidders will bid $\frac{v}{c}$ and all others will bid 0. In equilibrium, the expected revenue is equal to $\frac{v}{c}$, while the expected payoff for player i is 0.*
2. *With endogenous entry, player i randomly chooses to enter the auction between early and late entry stages.*

Proof. See Appendix A. □

Table 5.1 presents the corresponding entry game with two bidders. The number in each cell is the expected payoff for each bidder with different game structures. In this game, there are four pure-strategy Nash equilibria: (E,E), (E,L), (L,E) and (L,L).

By contrast, when the tie-breaking rule favors late entries, the expected payoff for entering late is at least as high as that for entering early; therefore, it is a weakly

[6]In simultaneous all-pay auctions, as everyone is in the same bidding stage, the winner is also randomly selected with equal probability.

Table 5.2 The Entry Game with a Favor-Late Tie-Breaking Rule: Homogeneous Bidders

		Bidder 2	
		Early	Late
Bidder 1	Early	0, 0	0, v
	Late	v, 0	0, 0

dominant strategy for player i to enter late. We summarize this prediction in the following proposition.

Proposition 8 (Sequential All-Pay Auctions with a Favor-Late Tie-Breaking Rule: Homogeneous Bidders)**.**

1. *In a sequential all-pay auction where there is one and only one player in the late stage, the unique SPNE is that all early bidders will bid 0 and the single late bidder will bid 0 as well. In this equilibrium, the expected revenue is equal to 0, while the expected payoff is v for the late bidder and 0 for others.*[7]
2. *With endogenous entry,* $q_i = 0$ *is a weakly dominant strategy,* $\forall i \in N$.

Proof. As the proof is similar to that relating to heterogeneous bidders in Konrad and Leininger (2007), we omit it here. □

Table 5.2 presents the corresponding entry game with two bidders. This game has three pure-strategy Nash equilibria: (E,L), (L,E) and (L,L). Furthermore, entering late is both bidders' weakly dominant strategy.

When bidders have heterogeneous costs, where $c_1 < c_2 < ... < c_n$, we now summarize the theoretical predictions in Konrad and Leininger (2007). Note, for the bidder with the second-lowest cost among all bidders in the late stage, $c_{l(2)}\overline{x}_{l(2)} = v$. For the bidder with the second-lowest cost among N bidders, $c_2\overline{x}_2 = v$.

Proposition 9 (Sequential All-Pay Auctions with Heterogeneous Bidders: Proposition 2 in Konrad & Leininger 2007)**.**

1. *Given a sequential all-pay auction, the equilibrium payoff for each bidder under the unique SPNE is:*

$$\pi_{i>1} = 0$$

$$\pi_1 = \begin{cases} 1 & \textit{if } L = \{1\} \\ v - c_1\overline{x}_{l(2)} & \textit{if } 1 \in L \textit{ and } \#L > 1 \\ v - c_1\overline{x}_2 & \textit{if } 1 \in E \end{cases} \tag{5.1}$$

[7]In a sequential all-pay auction where there are $1 < i \leq N$ bidders in the late stage, this scenario is equivalent to a simultaneous all-pay auction with i players.

Table 5.3 The Entry Game for Heterogeneous Bidders

		Bidder 2	
		Early	Late
Bidder 1	Early	$v - c_1\bar{x}_2$, 0	$v - c_1\bar{x}_2$, 0
	Late	v, 0	$v - c_1\bar{x}_2$, 0

2. *$q_1 = 0$ is a weakly dominant strategy for bidder 1.*
3. *Bidders 2 through n randomly choose between early and late entry, as their expected payoff is always 0.*
4. *A tie-breaking rule does not change the equilibrium.*[8]

Table 5.3 presents the entry game with two bidders. In particular, there are three pure-strategy Nash equilibria: (E,L), (L,E) and (L,L), and entering late is bidder 1's weakly dominant strategy.

After characterizing the SPNE in sequential all-pay auctions, we summarize the theoretical predictions in Baye et al.'s (1996) model of simultaneous all-pay auctions. To do so, we first specify the following predictions for homogeneous bidders.

Proposition 10 (Simultaneous All-Pay Auctions with Homogeneous Bidders: Theorem 1 in Baye et al. 1996)**.**

1. *The unique symmetric Nash equilibrium is that all bidders randomize continuously on [0, $\frac{v}{c}$].*[9]
2. *In any equilibrium, the expected payoff for each bidder is 0.*
3. *All equilibria are revenue equivalent; the expected revenue is $\frac{v}{c}$.*

In simultaneous all-pay auctions with two bidders, the bidding CDF for risk-neutral bidders is $\frac{cx}{v}$ and the corresponding PDF is $\frac{c}{v}$.

Although there are multiple Nash equilibria with homogeneous bidders in this game when $n > 2$, the Nash equilibrium with heterogeneous bidders is always unique.

Proposition 11 (Simultaneous All-Pay Auctions with Heterogeneous Bidders: Theorem 3 in Baye et al. 1996)**.**

1. *In the Nash equilibrium, bidder 1 randomizes uniformly on [0, $\bar{x}_2$]. Bidder 2 randomizes uniformly on (0, $\bar{x}_2$] and her probability of bidding 0 is $1 - c_1$. Each bidder $i > 2$ bids 0 without exception.*
2. *The expected payoff is $v(\frac{1}{c_1} - 1)$ for bidder 1 and 0 for all other bidders.*
3. *The expected revenue is $\frac{v}{2c_2} + \frac{vc_1}{2c_2^2}$.*

[8]The favor-late tie-breaking rule generates a unique SPNE, whereas both the favor-early and random tie-breaking rules generate the ϵ equilibria which converge to the unique SPNE.

[9]When $n = 2$, the unique Nash equilibrium exists; when $n > 2$, a continuum of asymmetric Nash equilibria exist.

In summary, assuming individuals are risk-neutral, we expect that there will be more early entries by homogeneous bidders with a favor-early tie-breaking rule as late entry is the weakly dominant strategy for them with a favor-late tie-breaking rule. In addition, for heterogeneous bidders, late entry is the weakly dominant strategy for the bidder with the lowest bidding cost regardless of tie-breaking rules. In terms of the revenue comparison, for homogeneous bidders, sequential all-pay auctions (with a favor-early tie-breaking rule) will generate revenue **equal** to that generated by simultaneous all-pay auctions.[10] Finally, for heterogeneous bidders, the revenue in sequential all-pay auctions (with only the lowest cost player in the late stage) will be **lower** than that in simultaneous all-pay auctions. These theoretical predictions guide our subsequent experimental design and data analysis.

5.4 Experimental Design

As discussed in Section 5.3, we expect individual behavior to vary across bidding cost environments and with different tie-breaking rules. Specifically, we are interested in investigating whether a favor-early tie-breaking rule induces more early entries, particularly for homogeneous bidders. More importantly, we want to compare individuals' bids and revenue between sequential and simultaneous all-pay auctions. Therefore, we use a 2×3 factorial design. We identify two bidding cost environments where two bidders bid against each other. First, we identify the homogenous environment where both bidders have the same bidding cost. Specifically, the marginal bidding cost is 1 token/bid for everyone in the experiment. Secondly, we identify the heterogeneous environment where at the beginning of each round, we randomly choose one bidder as the high-cost bidder. This person's bidding cost is 1 token/bid while the bidding cost of the other bidder is 0.8 token/bid. In our endogenous-entry treatments where both the sequential and simultaneous all-pay auctions exist, we include both favor-early and favor-late tie-breaking rules. To establish a benchmark and to replicate results from the relevant literature, we also include exogenous-entry treatments where only simultaneous all-pay auctions exist.

This 2×3 factorial design yields six treatments, as shown in Table 5.4. Each treatment has three independent sessions, with 12 subjects in each session. At the beginning of each round, subjects are randomly matched into groups of two. Since

[10]As entering late is the weakly dominant strategy for homogenous bidders with a favor-late tie-breaking rule, we do not expect sequential all-pay auctions there.

Table 5.4 Experimental Design

Auction Format	Entry Timing	Tie-Breaking Rule	Homogeneous Bidders	Heterogeneous Bidders	Total Subjects
Sim	Exogenous	Random	12×3	12×3	72
Sim or Seq	Endogenous	Favor-Early	12×3	12×3	72
Sim or Seq	Endogenous	Favor-Late	12×3	12×3	72

all-pay auction models assume one-shot interactions, the random re-matching protocol minimizes repeated-game effects. The value of the object is always set at 100 tokens and each person's bidding cost is common knowledge to represent a complete information environment. To prevent any potential bankruptcy problem, we give every subject 125 tokens as an endowment at the beginning of each round. Finally, each session lasts 30 rounds to capture any learning effect.

In our exogenous-entry treatments, there is no entry decision stage and every subject submits a bid independently and simultaneously. At the end of each round, the high bidder wins the object and each bidder pays her own bid. If there is a tie for the high bid, the winner is randomly chosen.

By contrast, in our endogenous-entry treatments, each participant first chooses whether to enter the auction early or late. After the entry decisions are shared between contestants, each bidder chooses a bid in her respective bidding stage. If a subject enters early, her bid will be observed by the late bidder. If both subjects are in the same stage, they bid independently and simultaneously, as in the exogenous-entry treatments. For treatments with a favor-early (late) tie-breaking rule, when bidders enter in different bidding stages and submit the same bid, the early (late) bidder is the winner. When bidders are tied and enter at the same bidding stage, the winner is randomly selected.

After players participate in 30 rounds, we implement the lottery choices outlined in Tanaka, Camerer and Nguyen (2010) to measure individual risk preference. A sample of the instructions is included in Appendix C.

At the end of the experiment, we give each participant a post-experiment questionnaire which includes demographic and personality trait questions, as well as menstrual cycle questions for female participants.[11] The post-experiment questionnaire

[11]We collect the menstrual cycle data for another study which investigates gender effects on individual behaviors across different games.

is included in Appendix D. Altogether, we conducted 18 independent computerized sessions at the School of Information Lab at the University of Michigan in May 2010, utilizing a total of 216 subjects. Our subjects are students from the University of Michigan, recruited by email from a subject pool for economic experiments. We allow subjects to participate in only one session. We use z-Tree (Fischbacher 2007) to program our experiments. Each session lasts approximately one hour, with the first 15 minutes used for instructions. The exchange rate is set to 8 tokens per \$1. In addition, each participant is paid a \$5 show-up fee. The average amount participants earn is \$20, including the show-up fee. Data are available from the author upon request.

5.5 Hypotheses

Based on the theoretical predictions outlined in Section 5.3 and given our experimental design, we now state our alternative hypotheses, while our general null hypothesis is that there is no difference between games or treatments.[12]

First, Proposition 7 implies that in the endogenous-entry-favor-early treatment, the expected payoff for homogeneous bidders is always 0, regardless of entry stage. Furthermore, Proposition 8 predicts that, in the endogenous-entry-favor-late treatment, the expected payoff for homogeneous bidders who enter late is always higher than or equal to that of early entry bidders. Therefore, we expect more early entries by homogeneous bidders in the endogenous-entry-favor-early treatment than in the endogenous-entry-favor-late treatment. In contrast, heterogeneous bidders' expected payoffs do not change with tie-breaking rules. The weakly dominant strategy for low-cost bidders is to enter late, while it does not matter if high-cost bidders randomly choose between early and late entry. This discussion leads to our first hypothesis:

Hypothesis 13 (Effects of Tie-Breaking Rules on Entry Decisions)**.** *In the entry decision stage, (a) homogeneous bidders are more likely to enter early when a favor-early tie-breaking rule is implemented than when a favor-late tie-breaking rule is implemented. (b) Different tie-breaking rules do not affect entry decisions for bidders with heterogeneous costs.*

Next, we discuss the expected bidder behavior in both sequential and simultaneous all-pay auctions. In particular, Proposition 7 predicts that homogeneous bidders in

[12] The only exception is Hypothesis 2. In the alternative hypothesis, we expect equal revenue between endogenous- and exogenous-entry treatments for homogeneous bidders.

sequential all-pay auctions with a favor-early tie-breaking rule should have total bids equal to those in simultaneous all-pay auctions. Specifically, the early bidder bids 100 and the late bidder bids 0. Although the expected revenue in sequential all-pay auctions is 0 when a favor-late tie-breaking rule is implemented, entering late is the weakly dominant strategy for bidders in the endogenous-entry-favor-late treatments (Proposition 8). Therefore, the revenue in the endogenous-entry-favor-late treatments should be also the same as that in the exogenous-entry treatments.

Hypothesis 14 (Sequential v. Simultaneous All-Pay Auctions: Homogeneous Bidders)**.** *For homogenous bidders, the revenue in the endogenous-entry treatments should be the same as that in the exogenous-entry treatments.*

Hypothesis 14 implies that the revenue for homogeneous bidders should be the same across treatments. In addition, Proposition 10 implies that in simultaneous all-pay auctions, both bidders randomize uniformly on [0, 100] and the bidding PDF is 0.01.[13]

For heterogeneous bidders, the equilibrium predictions do not change with different tie-breaking rules. Propositions 9 and 11 imply the following hypothesis.

Hypothesis 15 (Sequential v. Simultaneous All-Pay Auctions: Heterogeneous Bidders)**.** *For heterogeneous bidders, the revenue in the endogenous-entry treatments should be less than that in the exogenous-entry treatments.*

Hypothesis 15 implies that in sequential all-pay auctions with a high-cost early bidder, both bidders will bid 0 and the probability for low-cost bidders to win the auction is 1. In contrast, in simultaneous all-pay auctions, the low-cost bidder randomizes uniformly on [0, 100] and the bidding PDF is 0.01. Furthermore, the high-cost bidder randomizes uniformly on (0, 100] and the bidding PDF is 0.008. Additionally, the probability of bidding 0 for the high-cost bidder is 0.2. Therefore, the probability for low-cost bidders to win the auction is 0.6 in simultaneous all-pay auctions, which is lower than that in sequential all-pay auctions.

[13] As subjects can only bid integers between 0 and 100, there is a continuum of asymmetric Nash equilibria and the symmetric Nash equilibrium is that bidders uniformly bid any integers between 0 and 99(Baye, Kovenock and de Vries, 1994). With the large strategy space here, this symmetric Nash equilibrium converges to the unique Nash equilibrium under the continuous strategy space.

5.6 Results

In this section, we first report participant entry decisions in our endogenous-entry treatments. We then report individual bids in sequential and simultaneous all-pay auctions and compare endogenous- and exogenous-entry treatments, including revenue, earnings, and efficiency. We also investigate whether individuals randomize continuously in simultaneous all-pay auctions, as predicted by a risk-neutral model.[14] Finally, we propose a simultaneous all-pay auction model with risk-averse and loss-averse bidders, which fits the pattern of individual bids better than the risk-neutral model.

5.6.1 Entry Decisions

In this subsection, we compare the effects of tie-breaking rules on individual entry decisions. Figure 5.1 presents the proportion of early entries for homogeneous bidders under different tie-breaking rules in each round. The x-axis denotes the round and the y-axis indicates the proportion of early entries. The dashed line represents the proportion of early entries in the favor-early treatment while the solid line represents that of the entries in the favor-late treatment. Although the proportion of early entries decreases with time, for homogeneous bidders, it is higher in the favor-early treatment than in the favor-late treatment.[15] However, no significant difference exists for either low-cost or high-cost bidders (Figure 5.2). Using a test of proportions with standard errors clustered at the session level, we find a significant difference only for homogeneous bidders.

Result 18 (Entry Decisions)**.** *Homogeneous bidders are significantly more likely to enter early when the tie-breaking rule favors early over late entry (9% vs. 2%,* $p < 0.01$*, one-sided). Among heterogeneous bidders, there is no significant difference for either low-cost (9% vs. 5%,* $p = 0.308$*, one-sided) or high-cost bidders (8% vs. 5%,* $p = 0.308$*, one-sided).*

By Result 18, we reject the null hypothesis in favor of Hypothesis 13. Table 5.5 presents the proportion of different entry decision combinations. For each type of bidder, both bidders opt to enter late most often.

Although the favor-early tie-breaking rule induces homogeneous bidders to enter early, these bidders do not choose between early and late entry with equal probability.

[14] We include both simultaneous all-pay auctions in endogenous- and exogenous-entry treatments.
[15] The only two exceptions are rounds 12 and 29.

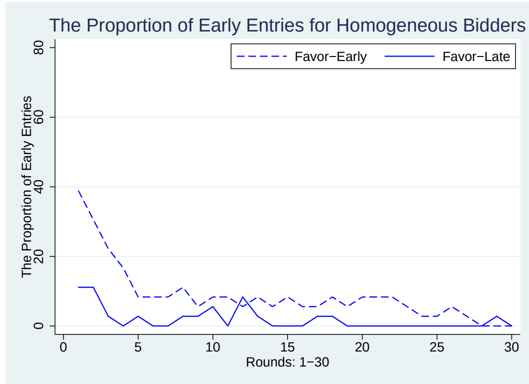


Figure 5.1 The Early Entry Rate for Homogeneous Bidders: Favor-Early v. Favor-Late

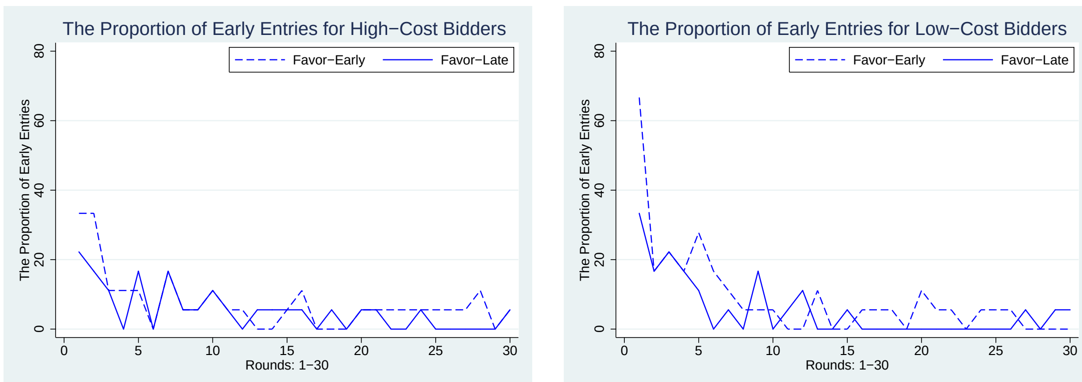


Figure 5.2 The Early Entry Rate for Heterogeneous Bidders: Favor-Early v. Favor-Late

Interestingly, the majority of high-cost bidders also enter late. We conjecture that these bidders would rather take a risk to win with a positive payoff. For example, in the equilibrium, high-cost bidders have a 40% chance of winning in simultaneous all-pay auctions, which may lead them to choose a late entry.

5.6.2 Endogenously Sequential v. Simultaneous All-Pay Auctions

In this section, we report individual bidder behavior across the sequential and simultaneous all-pay auctions in endogenous-entry treatments.[16] We also compare revenue, earnings and efficiency between endogenous- and exogenous-entry treatments.

[16]As only 1% of the total number of pairs of bidders participating in the simultaneous all-pay auctions are pairs where both bid in the early stage, we focus on simultaneous all-pay auctions where both bidders enter the late bidding stage.

Table 5.5 Percentage of Sequential and Simultaneous All-Pay Auctions in Endogenous-Entry Treatments

Bidder Type	Game Type	Favor-Early		Favor-Late	
		All Rounds	Rounds 16-30	All Rounds	Rounds 16-30
Homogeneous	Seq	15	9	4	1
	Sim with Both Early	1	0	0	0
	Sim with Both Late	84	91	96	99
Heterogeneous	Seq with High-Cost Early	6	5	5	2
	Seq with Low-Cost Early	7	4	5	1
	Sim with Both Early	1	0	0	0
	Sim with Both Late	85	91	90	97

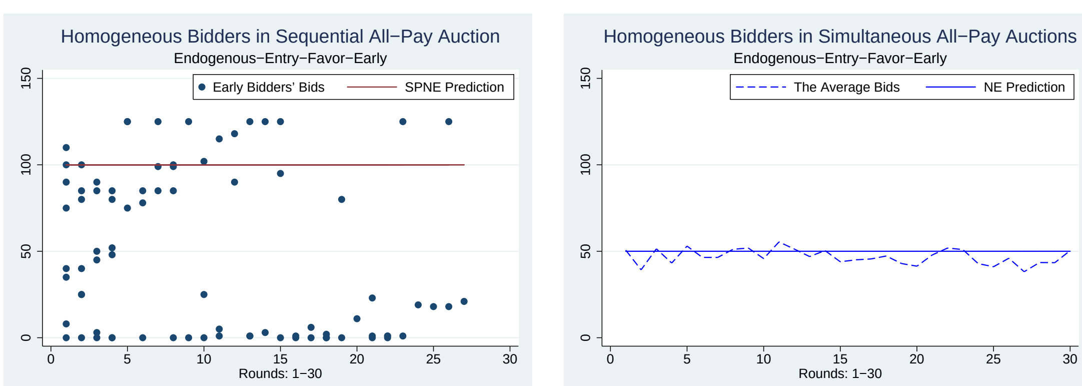


Figure 5.3 Sequential v. Simultaneous All-Pay Auctions: Homogeneous Bidders in the Endogenous-Entry-Favor-Early Treatment

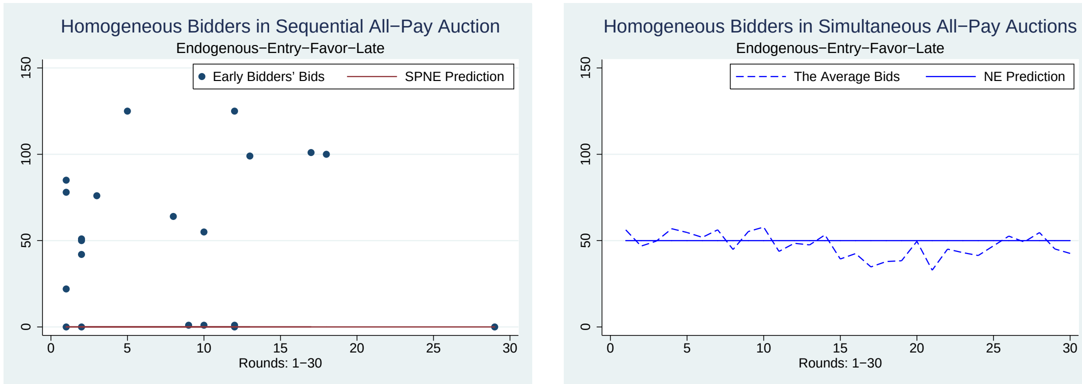


Figure 5.4 Sequential v. Simultaneous All-Pay Auctions: Homogeneous Bidders in the Endogenous-Entry-Favor-Late Treatment

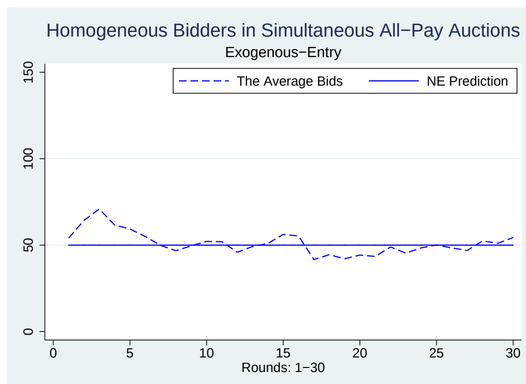


Figure 5.5 Simultaneous All-Pay Auctions: Homogeneous Bidders in the Exogenous-Entry Treatment

For homogeneous bidders in the endogenous-entry-favor-early treatment, Figure 5.3 (left panel) presents the early bids in sequential all-pay auctions. The dots represent individual bids and the line represents the SPNE prediction. We observe that these bids are bifurcated: 20% are equal to or above the SPNE of 100, while 36% are equal to or near 0.[17] In the post-experiment questionnaire, we ask subjects about their reasons for entering early and find the influence of social preference. For example, one player responded, "If I entered early and bid 0, I was guaranteed 125 tokens, and the other person could make max profit by winning." In addition, we

[17]Following Gneezy and Smorodinsky (2006), we use bids ≤ 5 as the cutoff for "near-zero" bids. We find that one subject bids irrationally even after round 15, making bids of 19, 18, 18, and 21 during rounds 24 to 27. We suspect that this subject did not understand the experiment well. For example, in the quiz to test subjects' understanding of the instructions for the experiment, this subject answered only two out of ten questions correctly, while the median number of correct answers is six in the endogenous-entry treatments.

find that 80% of our bidder 2s best respond to their respective bidder 1s. Specifically, when a bidder 1 bids less than 100, the bidder 2 adds one more bid to his match's bid. When bidder 1 bids 100 or higher, we find that bidder 2 bids 0. Although our (early) bidders behave differently from SPNE predictions in the sequential all-pay auctions, the average bids for homogeneous bidders in simultaneous all-pay auctions are consistent with Nash equilibrium predictions. Figures 5.3 (right panel) presents the average bids in each round for simultaneous all-pay auctions with homogeneous bidders. There is overbidding at the beginning which decreases with time. Using each session as one independent observation, we find no significant difference between the average bids and the NE prediction (47 v. 50, $p = 0.143$, one-sided one-sample Wilcoxon signed-rank tests).

For homogeneous bidders in the endogenous-entry-favor-late treatment, Figure 5.4 (left panel) presents the early bids in the sequential all-pay auctions. Consistent with Result 18, very few homogeneous bidders enter early. Their bids are also bifurcated. In contrast, Figure 5.4 (right panel) presents the average bids in simultaneous all-pay auctions and it is not statistically significantly different from the NE prediction (47 v. 50, $p = 0.5$, one-sided one-sample Wilcoxon signed-rank tests).

In addition, Figure 5.5 presents the average bids in simultaneous all-pay auctions in the exogenous-entry treatments. First, we find no significant difference between the average bids and the NE prediction (51 v. 50, $p = 0.5$, one-sided one-sample Wilcoxon signed-rank tests). Furthermore, the pairwise comparison of the average bids in simultaneous all-pay auctions between treatments is not significantly different (Favor-Early v. Favor-Late: 47 v. 47, $p = 0.414$; Favor-Early v. Exogenous: 47 v. 51, $p = 0.414$; Favor-Late v. Exogenous: 47 v. 51, $p = 0.414$, one-sided two-sample Wilcoxon rank-sum tests), indicating whether individuals endogenously choose or are exogenously assigned to simultaneous all-pay auctions does not affect their aggregate bidding behavior.

Table 5.6 The Average Revenue for Homogeneous Bidders

Treatment		Rounds	Revenue	P-values (1-sided)		
				(1) v. (2)	(1) v. (3)	(2) v. (3)
Exogenous-Entry	(1)	1-30	102	0.267		
		26-30	101	0.038		
Endogenous-Entry-Favor-Early	(2)	1-30	92			0.408
		26-30	88			0.138
Endogenous-Entry-Favor-Late	(3)	1-30	94		0.292	
		26-30	97		0.256	

Next, we report the revenue for homogeneous bidders between treatments. Note that there is a significant proportion of sequential all-pay auctions with low bids in the endogenous-entry-favor-early treatment. From Table 5.6, we see that the revenue in the endogenous-entry-favor-early treatments is lower than that in the endogenous-entry-favor-late and exogenous-entry treatments. In particular, the comparison in the last five rounds between the endogenous-entry-favor-early and exogenous-entry treatments is significant.

Result 19 (Revenue: Homogeneous Bidders)**.** *For homogeneous bidders, the revenue in the endogenous-entry-favor-early is lower than in exogenous-entry treatments (All Rounds: 92 v. 102,* $p = 0.267$*; Rounds 26-30: 88 v. 101,* $p = 0.038$*, one-sided two-sample Wilcoxon rank-sum tests).*

By Result 19, we fail to reject the null hypothesis in favor of Hypothesis 14. In addition, we find that the average earnings in each round is (significantly) higher in the endogenous-entry-favor-early than in the exogenous-entry treatments (All Rounds: 4 v. -1, $p = 0.256$; Rounds 26-30: 6 v. -1, $p = 0.038$, one-sided two-sample Wilcoxon rank-sum tests).

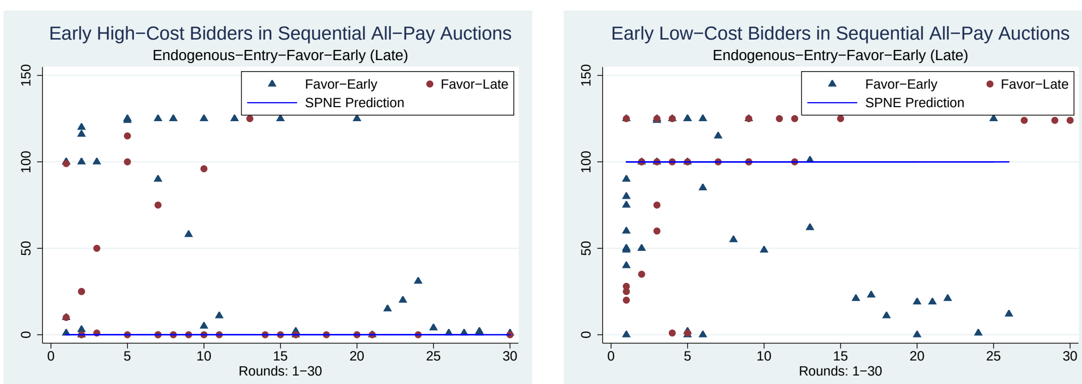


Figure 5.6 The Early High (Low)-Cost Bidders' Bids in Sequential All-Pay Auctions

Next, we examine the behavior of heterogeneous bidders across sequential and simultaneous all-pay auctions. Figure 5.6 plots early bids in both the endogenous-entry-favor-early and endogenous-entry-favor-late treatments. The triangles represent individual bids with a favor-early tie-breaking rule, while the dots represent individual bids with a favor-late tie-breaking rule, with the line representing the SPNE prediction. For high-cost bidders, regardless of the tie-breaking rule, 50% make "near-zero" bids while 28% make bids equal to or greater than 100. Bid amounts decrease significantly with experience. After round 15, the proportion of "near-zero" bids is 79%, and moves to 100% in the last five rounds. For low-cost bidders, the bids are bifurcated even in the last five rounds.

Figure 5.7 Simultaneous All-Pay Auctions: Heterogeneous Bidders in the Endogenous-Entry Treatment

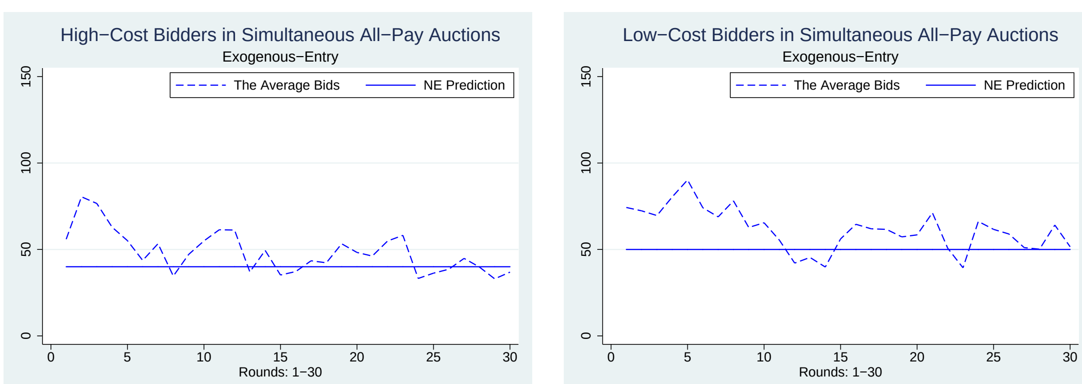


Figure 5.8 Simultaneous All-Pay Auctions: Heterogeneous Bidders in the Exogenous-Entry Treatment

In the simultaneous all-pay auctions when both bidders enter late, Figure 5.7 presents the average bid for both high- and low-cost bidders. In contrast to the sequential all-pay auctions, high-cost bidders in the simultaneous all-pay auctions overbid at the beginning and decrease their bids with experience. In particular, in the last five rounds, the average bids for the high-cost bidders are (weakly) significantly lower than the NE prediction (Favor-Early: 27 v. 40, $p = 0.054$; Favor-Late: 31 v. 40, $p = 0.143$; Pooled: 29 v. 40, $p = 0.023$, one-sided one-sample Wilcoxon signed-rank tests). Low-cost bidders also overbid at first and subsequently decrease their bids; however, their average bids are still higher than the NE prediction even in the last five rounds (Favor-Early: 60 v. 50, $p = 0.054$; Favor-Late: 56 v. 50, $p = 0.143$; Pooled: 58 v. 50, $p = 0.058$, one-sided one-sample Wilcoxon signed-rank tests).[18]

In addition, we observe similar bidding behaviors in simultaneous all-pay auctions in the exogenous-entry treatments. Figure 5.8 presents the average bid for both high- and low-cost bidders. Specifically, high-cost bidders decreases their bids with experience (All Rounds: 49 v. 40, $p = 0.054$; Rounds 26-30: 39 v. 40, $p = 0.297$, one-sided one-sample Wilcoxon signed-rank tests) and the average bids for low-cost bidders is higher than the NE prediction (All Rounds: 61 v. 50, $p = 0.054$; Rounds 26-30: 55 v. 50, $p = 0.297$, one-sided one-sample Wilcoxon signed-rank tests). Furthermore, the average bid for high-cost bidders in the endogenous-entry-favor-early treatment is significantly lower than in the other two treatments (Favor-Early v. Favor-Late: 36 v. 48, $p = 0.025$; Favor-Early v. Exogenous: 36 v. 49, $p = 0.025$; Favor-Late v. Exogenous: 48 v. 49, $p = 0.256$, one-sided two-sample Wilcoxon rank-sum tests) while the pairwise comparison is not significant in the last five rounds (Favor-Early v. Favor-Late: 27 v. 31, $p = 0.138$; Favor-Early v. Exogenous: 27 v. 39, $p = 0.138$; Favor-Late v. Exogenous: 31 v. 39, $p = 0.414$, one-sided two-sample Wilcoxon rank-sum tests). The average bid for the low-cost bidders in the endogenous-entry-favor-late treatment is marginally significantly higher than in the other two treatments (Favor-Early v. Favor-Late: 60 v. 69, $p = 0.063$; Favor-Early v. Exogenous: 60 v. 61, $p = 0.256$; Favor-Late v. Exogenous: 69 v. 61, $p = 0.063$, one-sided two-sample Wilcoxon rank-sum tests) and there is no significant difference in the last five rounds (Favor-Early v. Favor-Late: 60 v. 56, $p = 0.414$; Favor-Early v. Exogenous: 60 v. 55, $p = 0.256$; Favor-Late v. Exogenous: 56 v. 55, $p = 0.414$, one-sided two-sample Wilcoxon rank-sum tests).

Finally, we compare the revenue for heterogeneous bidders between treatments.

[18] This overbidding phenomenon is driven by bidders who continue to bid above 100. In the endogenous-entry treatments, for low-cost bidders, 74% of their bids are lower than 100 and the average is 47.

Table 5.7 The Average Revenue for Heterogeneous Bidders

Treatment		Rounds	Revenue	P-values (1-sided)		
				(1) v. (2)	(1) v. (3)	(2) v. (3)
Exogenous-Entry	(1)	1-30	110	0.025		
		26-30	94	0.256		
Endogenous-Entry-Favor-Early	(2)	1-30	96			0.025
		26-30	82			0.414
Endogenous-Entry-Favor-Late	(3)	1-30	116		0.256	
		26-30	88		0.256	

Table 5.7 reports summary statistics for the revenue across all treatments. From Table 5.7, we find that the revenue in the endogenous-entry-favor-early treatments is lower than that in the other two treatments. However, there is no significant difference in the last five rounds.

Result 20 (Revenue: Heterogeneous Bidders)**.** *For heterogeneous bidders, revenue in the endogenous-entry-favor-early treatments is significantly lower than that in the exogenous-entry and endogenous-entry-favor-late treatments (Favor-Early v. Exogenous: 96 v. 110, $p = 0.025$; Favor-Early v. Favor-Late: 96 v. 116, $p = 0.025$, one-sided two-sample Wilcoxon rank-sum tests), while the difference between treatments is not significant in the last five rounds ($p > 0.1$, one-sided two-sample Wilcoxon rank-sum tests).*

Result 20 indicates that for heterogeneous bidders, sequential all-pay auctions also generate lower revenue than simultaneous all-pay auctions. However, in the last five rounds, as there is few sequential all-pay auctions in the endogenous-entry treatments and individual bidding behavior in simultaneous all-pay auctions does not vary across treatments, the treatment difference is not significant anymore. Meanwhile, the average earnings in each round is higher in the endogenous-entry-favor-early treatments than in both the exogenous-entry (All Rounds: 8 v. 1, $p = 0.025$; Rounds 26-30: 15 v. 9, $p = 0.256$, one-sided two-sample Wilcoxon rank-sum tests) and endogenous-entry-favor-late treatments (All Rounds: 8 v. -1, $p = 0.025$; Rounds 26-30: 15 v. 12, $p = 0.414$, one-sided two-sample Wilcoxon rank-sum tests).

We also find that the ratio of efficient allocation where the winner is the low-cost bidder in the endogenous-entry treatments is higher than the ratio in the exogenous-entry treatments in both all rounds (Favor-Early v. Exogenous: 64% v. 59%, $p = 0.256$; Favor-Late v. Exogenous: 66% v. 59%, $p = 0.063$, one-sided two-sample Wilcoxon rank-sum tests) and in the last five rounds (Favor-Early v. Exogenous:

76% v. 67%, $p = 0.184$; Favor-Late v. Exogenous: 70% v. 67%, $p = 0.411$, one-sided two-sample Wilcoxon rank-sum tests).

5.6.3 Individual Strategy Analysis in Simultaneous All-Pay Auctions

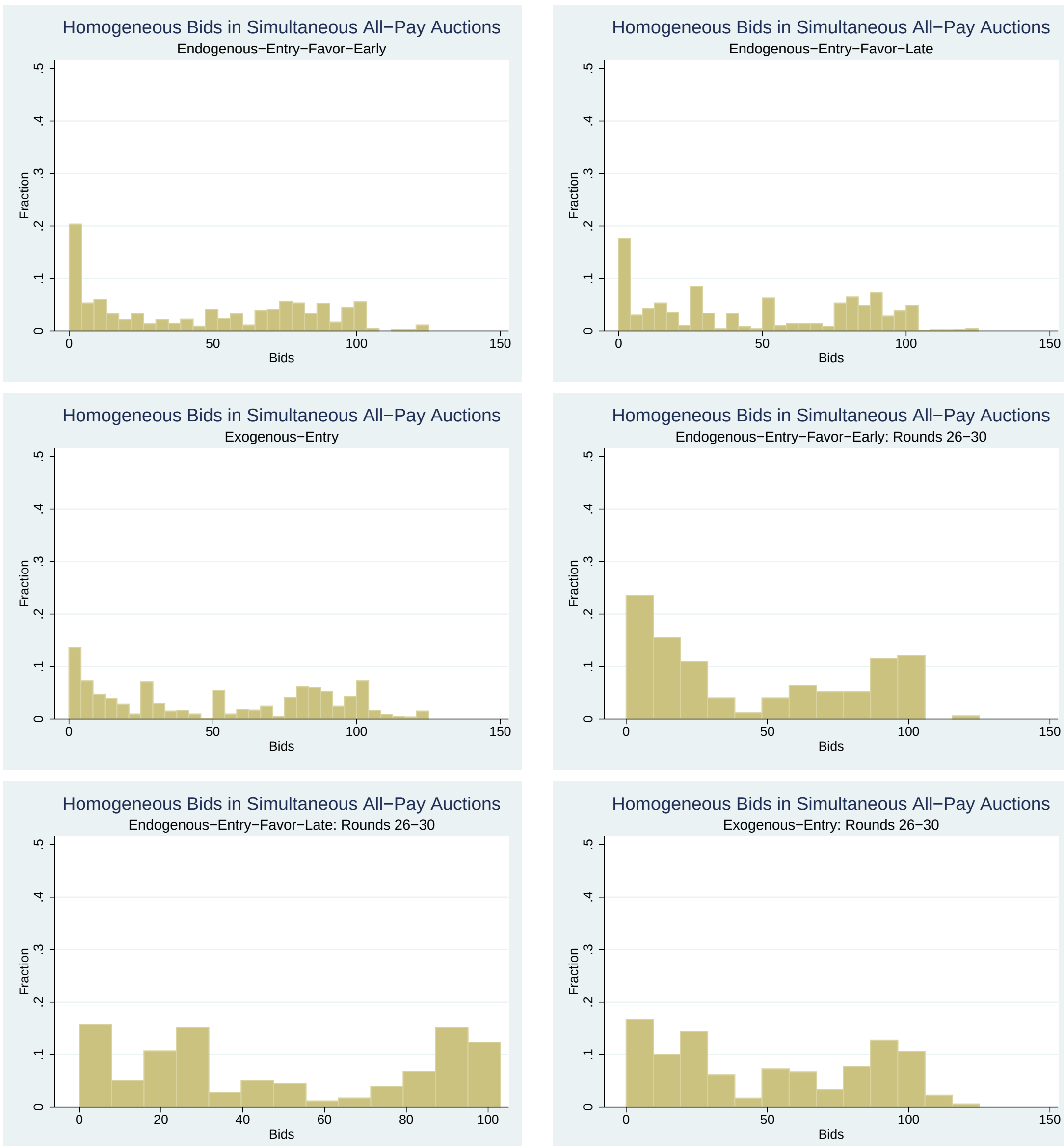


Figure 5.9 Bid Distribution of Homogeneous Bidders in Simultaneous All-Pay Auctions

In this section, we investigate bidding strategies in simultaneous all-pay auctions. We include simultaneous all-pay auctions where both bidders enter the late

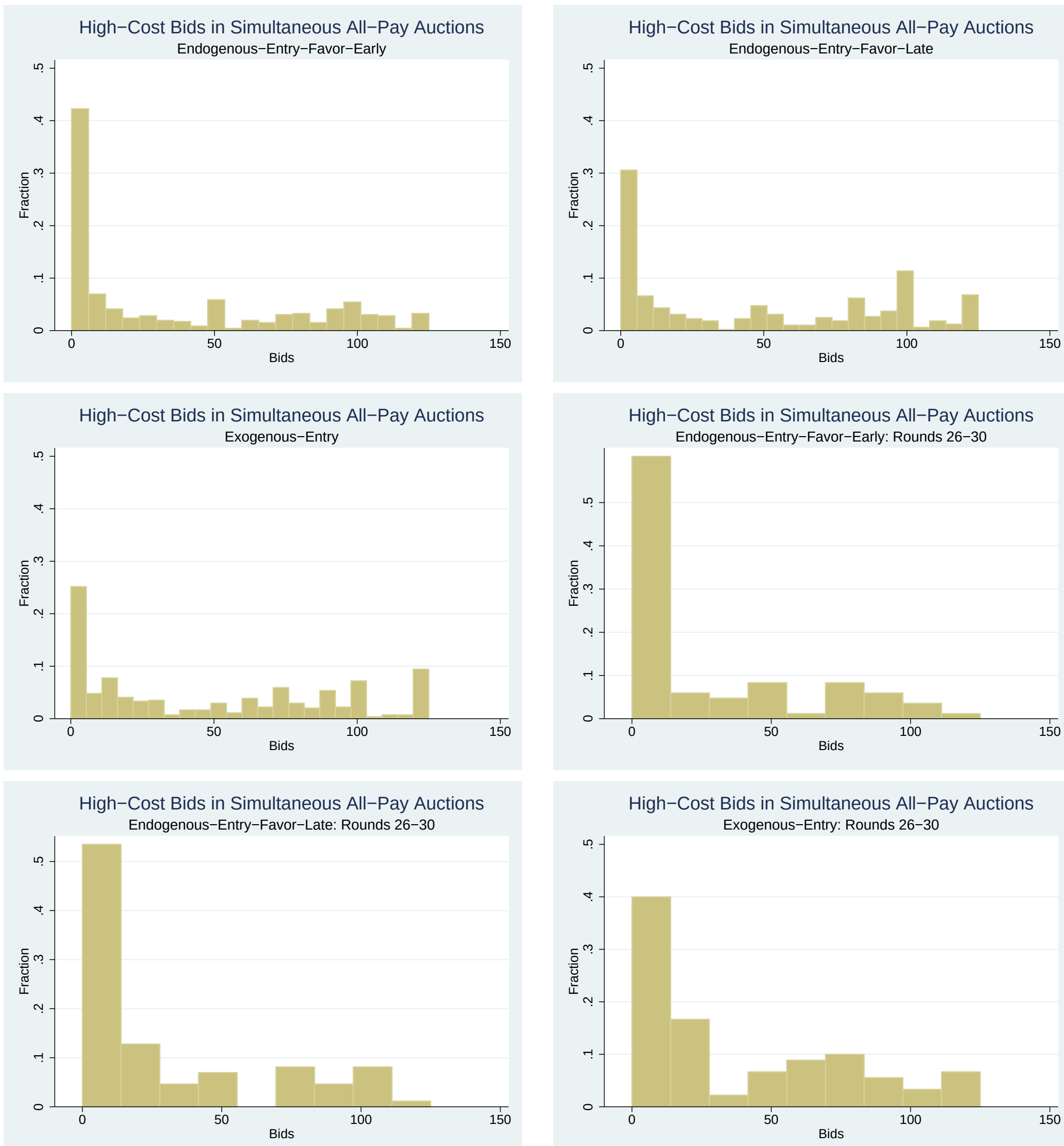


Figure 5.10 Bid Distribution of High-Cost Bidders in Simultaneous All-Pay Auctions

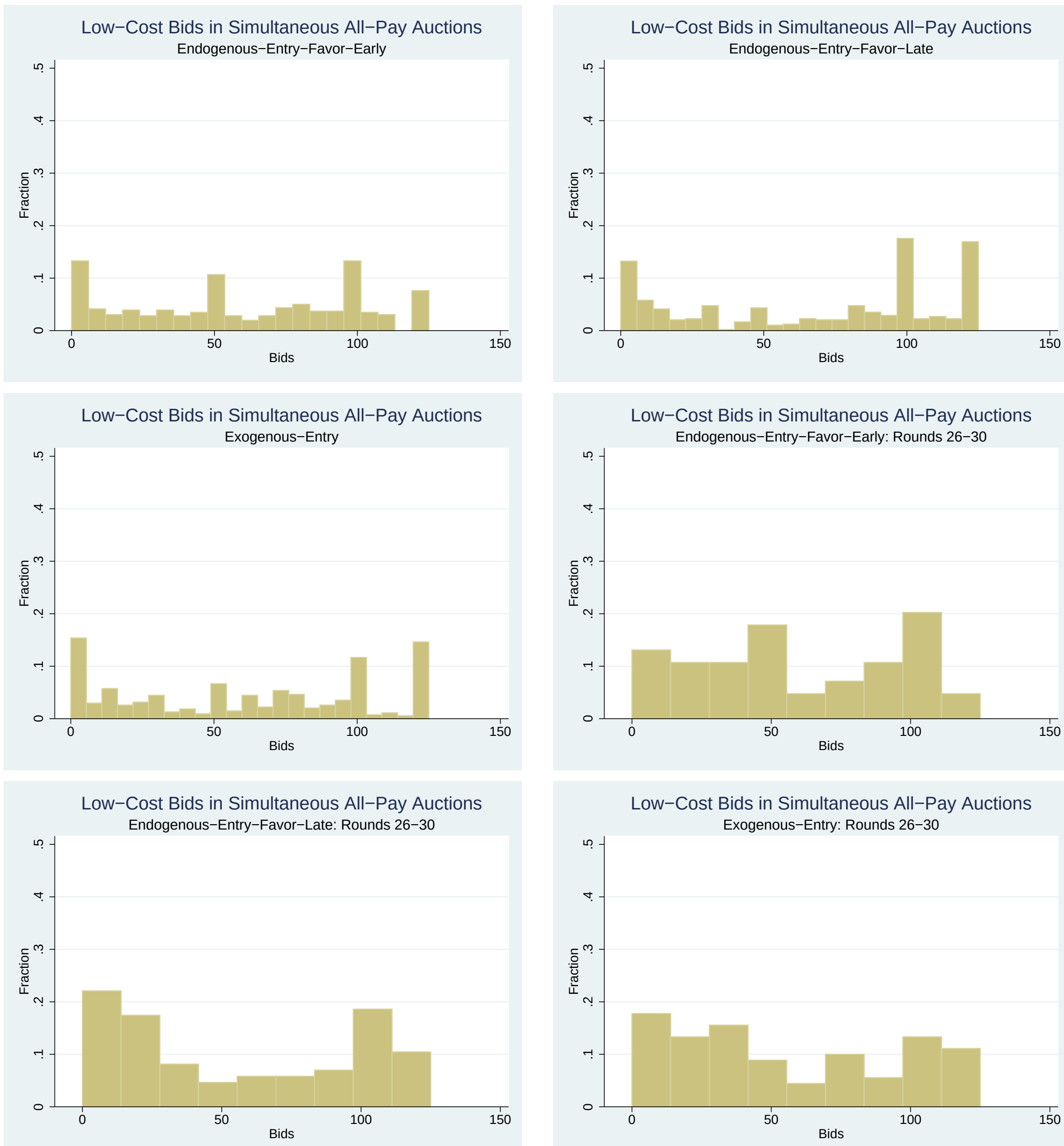


Figure 5.11 Bid Distribution of Low-Cost Bidders in Simultaneous All-Pay Auctions

bidding stage in endogenous-entry treatments, and simultaneous all-pay auctions in exogenous-entry treatments. Specifically, we examine whether homogeneous bidders, whose average bids are consistent with NE predictions, have the same bidding distribution as that predicted by a simultaneous all-pay auction model with risk-neutral bidders. Second, studying individual bidding strategies enables us to understand why the average bids for high- and low-cost bidders differ from NE predictions. Figures 5.9, 5.10 and 5.11 present the respective bidding histograms of each type across each treatment, including data from all rounds as well as data from rounds 26 to 30. In contrast with the NE prediction with risk-neutral bidders, where individuals are equally likely to bid any number in [0,100], we find that 0, 25, 50, 75 and 100 are empirical focal points. Furthermore, the bidding distribution is skewed toward lower bids for both the homogeneous and high-cost bidders and the proportion of near-zero bids is higher than the Nash Equilibrium prediction for risk-neutral high-cost bidders. For example, the proportion of near-zero bids for high-cost bidders is 42% in the endogenous-entry-favor-early treatment. In addition, low-cost bidders often choose to bid 125. Applying the econometric model in Wooders (2010) and Walker and Wooders (2001), we find that, in our simultaneous all-pay auctions, bids are not uniformly distributed in [0, 100]. The details of the analysis are described in Appendix B. Using the expected average bids under Nash equilibrium predictions with risk-neutral players as the cutoff, which is defined as c_0, where c_0 is 50 for homogeneous and low-cost bidders, and 40 for high-cost bidders, we categorize players into four types: (1) under-bidders: those who consistently bid less than c_0; (2) over-bidders: those who consistently bid more than c_0; (3) random-bidders: those who are equally like to bid between $[0, c_0]$ and $[c_0, 100]$; and (4) others: those who neither randomize nor keep the same strategy. Table 5.8 lists the percentage of each bidder type. We summarize this result below.

Result 21 (Individual Bids in Simultaneous All-pay Auctions)**.** *In simultaneous all-pay auctions, 40% of the bidders randomize their bids on their strategy space while 8% consistently overbid and 11% consistently underbid.*

Altogether, Result 21 indicates that bidding strategies are heterogeneous in simultaneous all-pay auctions. As risk-averse bids are bifurcated in incomplete information all-pay auctions (Fibich, Gavious and Sela 2006; Noussair and Silver 2006), we extend the complete information all-pay auction model to include risk- and loss-averse bidders. Assuming α is the risk-aversion parameter, where $0 < \alpha \leq 1$, λ is the loss-aversion parameter with $\lambda \geq 1$. Normalizing $c = 1$ for homogeneous bidders and $c_2 = 1$ for heterogeneous bidders, we obtain the following proposition.

Table 5.8 Percentage of Each Bidder Type in Simultaneous All-Pay Auctions

Bidder Type	Homogeneous	Low-Cost	High-Cost	All
Under-Bidders	14	11	8	11
Over-Bidders	13	4	7	8
Random-Bidders	20	47	53	40
Others	53	38	31	41

Proposition 12 (Simultaneous All-Pay Auctions with Risk- and Loss-Averse Bidders)**.**

1. *The unique symmetric Nash equilibrium is the mixed strategy Nash equilibrium where individuals randomize on [0, v].*
2. *In equilibrium, the bidding CDF is $G(x)$, where $G''(x) < 0$ with $0 \leq x \ll v$ and $G''(x) > 0$ with $0 \leq v - x \ll v$.*
3. *For bidder 1 among heterogeneous bidders, $G_1(0) \geq G_1^{rn}(0)$.*[19]

Proof. See Appendix A. □

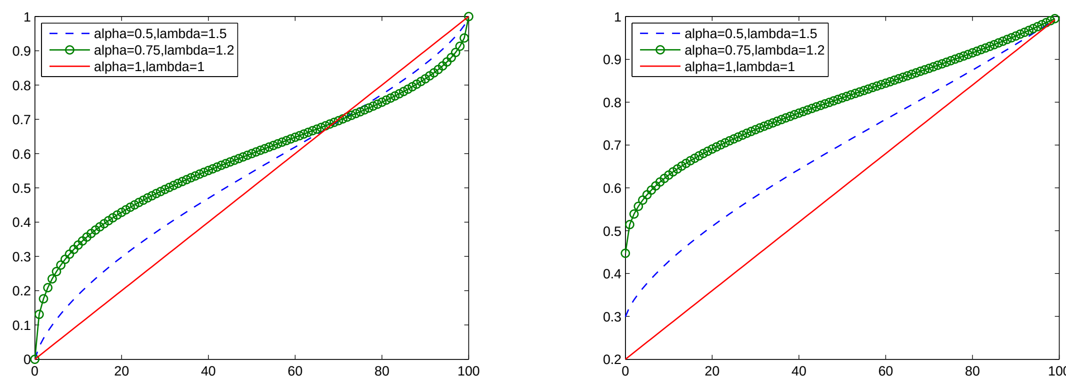


Figure 5.12 Bidding CDF for Risk-Averse and Loss-Averse Bidders in Simultaneous All-Pay Auctions: Homogeneous (Low-Cost) Bidders (left) and High-Cost Bidders (right)

Figure 5.12 presents two numerical examples. In each graph, the solid line represents the risk-neutral model, the dots represent relatively higher risk- and loss-aversion, and the dashed line represents relatively lower risk- and loss-aversion. Note that the bidding CDF exhibits an inverse S-shape; in particular, high-cost bidders have a higher probability of bidding 0 than those in the risk-neutral model. In addition, higher risk-aversion implies more bifurcated individual bids. Furthermore, higher loss-aversion implies a greater number of lower bids.

[19] $G_1(0)$ is bidder 1's probability of bidding 0. $G_1^{rn}(0)$ is her probability of bidding 0 with risk and loss neutrality.

Next, we estimate the risk- and loss-aversion parameters using data from all rounds as well as data from rounds 26-30. Table 5.9 reports the estimation results. In the last five rounds, where no learning effect is captured, the estimated risk-aversion parameter, α for high-cost bidders is 0.87 while the estimated loss-aversion parameter, λ is 2.92.[20] Thus the expected bids is 31.72.[21] These results explain why the average bids for high-cost bidders in later rounds are significantly lower than the NE prediction with risk-neutral bidders. Overall, we find that our data fit the model with risk- and loss-averse bidders significantly better than the model with risk- and loss-neutral bidders or the model with risk-averse and loss-neutral bidders ($p < 0.01$, one-sided Likelihood Ratio Tests).

5.7 Discussion

In this study, we present the results of an experiment to investigate individual behavior in endogenously sequential all-pay auctions. Our results show that a favor-early tie-breaking rule induces homogeneous bidders to enter the auction earlier. On a practical level, these results suggest that a favor-early tie-breaking rule may attract earlier inexperienced entrants on crowdsourcing sites such as Taskcn (Liu et al., 2011). However, it does not imply the same effect for experienced users. For example, Liu et al. (2011) find that experienced users on Taskcn submit their solutions significantly later than inexperienced users.

Furthermore, as the first experimental investigation comparing the two all-pay auction mechanisms, this study shows that the revenue generated in sequential all-pay auctions is less than that in simultaneous all-pay auctions. In addition, the average earnings for bidders in sequential all-pay auctions is greater than that in simultaneous all-pay auctions. We also find that the proportion of efficient allocations for sequential all-pay auctions is significantly higher than that for simultaneous all-pay auctions. These results have implications for the design of all-pay auction crowdsourcing sites as well as other all-pay auction contests. For example, simultaneous all-pay auctions may be more effective in encouraging greater efforts from contestants.

Finally, we study individual strategies in simultaneous all-pay auctions. In doing so, we find that a significant proportion of bidders do not randomize as predicted by the risk-neutral model. Instead, bidders are more likely to have either extremely high

[20] Tversky and Kahneman (1992) estimate $\alpha = 0.88$ and $\lambda = 2.25$.

[21] The average amount of bids in the last five rounds is 32.23 in the experiment.

Table 5.9 Risk and Loss Aversion Estimation: MLE

	Homogeneous			High-Cost			Low-Cost		
All Rounds	Model			Model			Model		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
α	1	0.69	0.70	1	0.73	0.75	1	0.67	0.67
(σ^2)		(0.03)	(0.03)		(0.03)	(0.03)		(0.05)	(0.05)
λ	1	1	1.33	1	1	1.29	1	1	1.23
(σ^2)			(0.13)			(0.22)			(0.24)
-Log(L)	13304	13017	12982	5369	5280	5274	5126	4997	4989
LR Test (P-Values)									
1 v. 2			0.000			0.000			0.000
1 v. 3			0.000			0.000			0.000
2 v. 3			0.000			0.001			0.000
Obs.	2889	2889	2889	1306	1306	1306	1113	1113	1113
	Homogeneous			High-Cost			Low-Cost		
Rounds 26-30	Model			Model			Model		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
α	1	0.67	0.68	1	0.74	0.87	1	0.78	0.80
(σ^2)		(0.05)	(0.06)		(0.03)	(0.05)		(0.05)	(0.07)
λ	1	1	1.23	1	1	2.92	1	1	1.52
(σ^2)			(0.09)			(0.59)			(0.37)
-Log(L)	2362	2300	2297	1024	1006	986	939	930	924
LR Test (P-Values)									
1 v. 2			0.000			0.000			0.000
1 v. 3			0.000			0.000			0.000
2 v. 3			0.009			0.000			0.000
Obs.	513	513	513	246	246	246	204	204	204

or low bids. This finding is consistent with the predictions of an enriched model with risk and loss aversion.

In the future, we are interested in studying the performance of all-pay auctions with endogenous entry timing in more realistic environments, such as those featuring incomplete information. The majority of subjects in our experiment learn to enter late, as do experienced users on real all-pay auction labor markets (Liu et al., 2011). Therefore, understanding how to design mechanisms to effectively induce a greater number of early entries is an important topic for both theoretical and experimental studies.

5.8 Appendix A: Proofs

Proof of Proposition 7:

We first characterize the SPNE in sequential all-pay auctions with homogeneous bidders when the tie-breaking rule favors early entries.

Following the process in Konrad and Leininger (2007), we first consider the subgame at the late stage (L) assuming that maximum bids from the early stage (E) are characterized by $\bar{x}_E \equiv max_{i \in E}\{x_i\}$.

Next, we show that the following candidate is an equilibrium in the subgame at stage L as bidder strategies are mutually best response. x_L^i represents the bids in stage L for bidder i.

1. If $c\bar{x}_E \geq v$, $x_L^i = 0$ and the payoff for bidder i in stage L is $\pi_L^i = 0$. Otherwise, the payoff for bidder i in stage L is $\pi_L^i < 0$.
2. If $c\bar{x}_E < v$, each bidder i in stage L will randomize on $[\bar{x}_E, \bar{x}]$ where $c\bar{x} = v$. In addition, $\pi_L^i = 0$.

Now we consider stage E. As $\bar{x}_L \equiv max_{i \in L}\{\bar{x}_i\}$ and $\bar{x}_i = \bar{x}$ for homogenous bidders, we show that the following candidate is an equilibrium in the subgame at stage E as bidders' strategies are mutually best response. x_E^i represents the bids in stage E for bidder i.

1. If there is one and only one bidder i in stage E, $x_E^i = \bar{x}$ and the payoff for this bidder is $\pi_E^i = 0$. Otherwise, the payoff $\pi_E^i \leq 0$.
2. If there is more than one bidder in stage E, one will bid $\bar{x}$ and all others will bid 0. For bidder i with $x_E^i = \bar{x}$, the payoff in stage E is $\pi_E^i = 0$; otherwise, the payoff is $\pi_E^i \leq 0$. For bidder k with $x_E^k = 0$, the payoff in stage E is $\pi_E^k = 0$; otherwise $\pi_E^k \leq 0$. Consequently, no one has an incentive to deviate from the strategy.

The uniqueness of this equilibrium follows the same proof in Baye et al. (1996). In addition, when all bidders enter either stage E or stage L together, the game become a simultaneous all-pay auction and the unique symmetric Nash equilibrium is that bidders randomize continuously on $[0, \bar{x}]$. The expected payoff for everyone is always 0.

Furthermore, when the entry timing is endogenous, since the expected payoff is always 0, bidders should randomly choose between early and late entry. ∎

Proof of Proposition 12:

For heterogeneous bidders, the bidding cost is c_1 for bidder 1 and c_2 for bidder 2. We normalize $c_2 = 1$. Following the line of argument in Baye et al. (1996), we show that the unique Nash equilibrium is the mixed strategy Nash equilibrium where bidders 1 and 2 randomly play between 0 and v and bidder 2 has a mass point at 0. Furthermore, we specify that each bidder $n \geq 3$ bids 0 with probability 1. The equilibrium bidding CDF is defined as $G_1(x)$ for bidder 1 and $G_2(x)$ for bidder 2. We first characterize $G_1(x)$.

As the expected payoff U_2 for bidder 2 is always 0 on the equilibrium, we obtain:

$$U_2(x) = G_1(x)u(v - x) + (1 - G_1(x))u(-x) = 0$$

Consequently,

$$G_1(x) = \frac{-u(-x)}{u(v - x) - u(-x)}.$$

The utility function is defined below:

$$u(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda(-x)^{\alpha} & \text{otherwise,} \end{cases}$$

where $0 < \alpha \leq 1$ measures individual risk aversion and $\lambda \geq 1$ represents individuals loss aversion. In particular, when $\alpha = 1$ and $\lambda = 1$, the bidding CDF is:

$$G_1^{rn}(x) = \frac{x}{v}.$$

We first show that $G_1(\frac{v}{2}) \geq \frac{1}{2}$.

When $x = \frac{v}{2}$, we obtain:

$$G_1(\frac{v}{2})u(\frac{v}{2}) + (1 - G_1(\frac{v}{2}))u(-\frac{v}{2}) = 0.$$

Then,

$$G_1(\frac{v}{2}) = \frac{\lambda}{1 + \lambda} \geq \frac{1}{2};$$

when $\lambda = 1$, we obtain $G_1(\frac{v}{2}) = \frac{1}{2}$.

As $\forall x$, $U_2(x) = 0$ in the equilibrium, $U_2'(x) = 0$ and $U_2''(x) = 0$. Defining $g_1(x) = G_1'(x)$, we obtain:

$$g_1(x) = \frac{(1 - G_1(x))u'(-x) + G_1(x)u'(v - x)}{u(v - x) - u(-x)} > 0.$$

Furthermore, $U''(x) = 0$ yields the following equation:

$$g_1'(x)(u(v-x)-u(-x)) = 2g_1(x)(u'(v-x)-u'(-x))-G_1(x)u''(v-x)-(1-G_1(x))u''(-x)$$

We next define $Z(x) = 2g_1(x)(u'(v-x) - u'(-x)) - G_1(x)u''(v-x) - (1 - G_1(x))u''(-x)$ to show that $Z(x) \leq 0$ when $0 \leq x \ll v$ and that $Z(x) \geq 0$ when $0 \leq v - x \ll v$.

As $u(x)$ is concave with $x \geq 0$, then $\exists \frac{v}{2} \leq x_1 = \frac{v}{1+\lambda^{\frac{1}{\alpha-1}}} < v$, $u'(v-x_1) = u'(-x_1)$. Therefore, $\forall\ x \geq x_1$, $u'(v-x) \geq u'(-x) > 0$, and $u''(v-x) \leq -u''(-x) < 0$. In addition, as $G(x) \geq \frac{1}{2}$, $Z(x) \geq 0$.

Next, we show that $Z(x) \leq 0$ when $0 \leq x \ll v$. As $G(\frac{v}{2}) \geq \frac{1}{2}$, then $\exists x_0 \leq \frac{v}{2}$, $G(x_0) = \frac{1}{2}$. Therefore, $\forall\ x \leq x_0$, $0 < u'(v-x) \leq u'(-x)$, and $-u''(-x) \leq u''(v-x) < 0$. In addition, as $G(x) \leq \frac{1}{2}$, $Z(x) \leq 0$.

Because of the continuity of $Z(x)$, $\exists y$ where $x_0 \leq y \leq x_1$, $Z(y) = 0$ and $g_1'(y) = 0$.[22] Now we show that there is one and only one y.

If $\exists\ Z(y_1) = Z(y_2) = 0$ and $y_1 < y_2$, then $Z'(y_1) < 0$. However, we know that

$$\begin{aligned} Z'(x) &= 2g_1'(x)(u'(v-x) - u'(-x)) + 2g_1(x)(-u''(v-x) + u''(-x)) \\ &-g_1(x)u''(v-x) + G_1(x)u'''(v-x) + g_1(x)u''(-x) + (1 - G_1(x))u'''(-x) \end{aligned}$$

When $x = y_1$, $Z'(y_1) = 3g_1(x)(u''(-x) - u''(v-x)) + G_1(x)u'''(v-x) + (1 - G_1)u'''(-x) \geq 0$, which is a contradiction.

Therefore, $\forall\ x \leq y$, $Z(x) \leq 0$ and $\forall\ x \geq y$, $Z(x) \geq 0$.

In addition, when $\lambda = 1$, $x_0 = x_1 = y = \frac{v}{2}$. Consequently, $\forall\ x \leq \frac{v}{2}$, $Z(x) \leq 0$ and $\forall\ x \geq \frac{v}{2}$, $Z(x) \geq 0$.

Now, we characterize bidder 2's CDF, $G_2(x)$. As the expected utility for bidder 1 is always $u((1-c_1)v)$ in the equilibrium, thus:

$$u((1-c_1)v) = G_2(x)u(v - c_1x) + (1 - G_2(x))u(-c_1x).$$

By extension,

$$G_2(x) = \frac{u((1-c_1)v) - u(-c_1x)}{u(v-c_1x) - u(-c_1x)}.$$

[22] As $Z(\frac{v}{2}) \leq 0$, $y \geq \frac{v}{2}$.

In particular, when $\alpha = 1$, $\lambda = 1$, then:

$$G_2^{rn}(x) = \frac{(1-c_1)v + c_1 x}{v}.$$

We now show that $G_2(0) \geq G_2^{rn}(0)$.

As $G_2(0) = \frac{u((1-c_1)v)}{u(v)}$ and $G_2^{rn}(0) = \frac{(1-c_1)v}{v}$, it is equivalent to show $\frac{u((1-c_1)v)}{(1-c_1)v} \geq \frac{u(v)}{v}$.

Because $u(x)$ is concave with $x \geq 0$, $u(t*x + (1-t)*y) \geq t*u(x) + (1-t)u(y)$ for $0 \leq t \leq 1$; therefore, if $y = 0$, we obtain $u(t*x) \geq t*u(x)$.

If we denote $z = t*x$, then $t = \frac{z}{x}$ for $\forall x \neq 0$.

Since $0 \leq t \leq 1$, $0 \leq z \leq x$, then $u(z) \geq \frac{u(x)}{x}*z$; therefore, $\frac{u(z)}{z} \geq \frac{u(x)}{x}$ $\forall 0 \leq z \leq x$. Defining $g(x) = \frac{u(x)}{x}$, $\forall x \geq 0$, we show that $g(x)$ decreases with x. Similarly, $\forall x \leq 0$, $g(x)$ decreases with x.

As $g((1-c_1)v) = \frac{u((1-c_1)v)}{(1-c_1)v}$, $g(v) = \frac{u(v)}{v}$, and $0 < (1-c_1)v < v$, we have $\frac{u((1-c_1)v)}{(1-c_1)v} \geq \frac{u(v)}{v}$.

Along the same line of proof for $G_1(x)$, $G_2''(x) < 0$ with $0 \leq x \ll v$ and $G_2''(x) > 0$ with $0 \leq v - x \ll v$.

In the end, for homogeneous bidders, $c_1 = c_2 = ..c_n = 1$. Along the same line of proof in Baye et al. (1996), the unique symmetric Nash equilibrium is that bidder i randomizes continuously on [0, v]. $H(x)$ is the CDF of bidder i in this symmetric Nash equilibrium. As the expected payoff for bidder i is 0, then:

$$U(x) = H(x)^{n-1}u(v-x) + (1-H(x))^{n-1}u(-x) = U(0) = 0$$

Therefore,we obtain:

$$\frac{H(x)}{1-H(x)} = \frac{(-u(-x))^{\frac{1}{n-1}}}{(u(v-x))^{\frac{1}{n-1}}}$$

By extension, we obtain:

$$\frac{H(x)}{1-H(x)+H(x)} = \frac{(-u(-x))^{\frac{1}{n-1}}}{(u(v-x))^{\frac{1}{n-1}} + (-u(-x))^{\frac{1}{n-1}}}$$

$$H(x) = \frac{\lambda^{\frac{1}{n-1}} x^{\frac{\alpha}{n-1}}}{(v-x)^{\frac{\alpha}{n-1}} + \lambda^{\frac{1}{n-1}} x^{\frac{\alpha}{n-1}}}$$

As we know from the above proof, for $G_1(x) = \frac{\lambda x^\alpha}{(v-x)^\alpha + \lambda x^\alpha}$, $\forall\, 0 < \alpha \leq 1$ and $\lambda \geq 1$, $G_1''(x) < 0$ with $0 \leq x \ll v$ and $G_1''(x) > 0$ with $0 \leq v - x \ll v$. Because $\lambda^{\frac{1}{n-1}} \geq 1$ and $0 < \frac{\alpha}{n-1} \leq 1$, $H(x)$ is also first concave and then convex. ∎

5.9 Appendix B: Testing Mixed Strategies in Simultaneous All-Pay Auctions

Following Wooders (2010), we categorize homogeneous bidders' bids in simultaneous all-pay auctions on [0, 100] into two categories: (1) [0,50]: representing low bids and (2) (50,100]: representing high bids. 50 is the expected average bid under the Nash equilibrium predictions. Table 5.10 shows the bidding counts for each category and the corresponding randomized binomial test results for homogeneous bidders.[23] With a null hypothesis that the probability for a bidder to choose a bid less than 50 is 0.5 and we find more rejection numbers than expected for each treatment.[24] Specifically, the null hypothesis that bidders choose low bids with probability 0.5 is rejected at the 5% level for 21 out of 36 bidders in our exogenous-entry treatments, 22 rejections in our endogenous-entry-favor-early treatments, and 24 in our endogenous-entry-favor-late treatments.[25]

To test the joint null hypothesis that all bidders in each treatment are equally likely to bid between these two different bidding categories, we examine the empirical distribution of the 36 p-values for the low bids from random binomial tests in each treatment. Under the null hypothesis that bidders choose low bids with probability 0.5, the p-value should be uniformly distributed in [0,1] for each treatment. Figure 5.13, left column presents the empirical CDF of p-values in each treatment. Kolmogorov-Smirnov (KS) test shows that none of these distributions is uniform ($p < 0.01$, two-sided).

Furthermore, we check the serial independence of bids by applying the method outlined in Walker and Wooders (2001). The null hypothesis is that each bid between low and high is serially independent. We reject the null hypothesis if there are too many or too few runs.[26] Tables 5.11, 5.12, and 5.13 report the data and results for the serial independence test. $F(r)$ denotes the probability of obtaining r, or few, runs. The null hypothesis is rejected at the 5% level if $F(r) < 0.025$ or $1 - F(r-1) < 0.025$. In summary, there are 17 out of 36 rejections in the exogenous-entry treatments, 14 rejections in the endogenous-entry-favor-early treatments, and

[23]As bids above 100 are not predicted by NE and an expected probability of each type of bid is required in a binomial test, we exclude this category in our binomial tests. The proportion of bids above 100 is 4% for homogeneous bidders. 12% for high-cost bidders and 25% for low-cost bidders.

[24]The expected rejection number is 1.8 bidders ($36 * 5\% = 1.8$).

[25]As the probability that bidders choose high bids is also 0.5, we obtain identical numbers of rejections for high bids.

[26]A run is a maximal string of consecutive identical symbols, either all low bids or high bids. For example, the bidding sequence s={L,L,H,L} has three runs.

17 rejections in the endogenous-entry-favor-late treatments.[27] In particular, these rejections occur because there are too few runs, $F(r) < 0.025$, indicating that bidders keep over or underbidding. In addition, to test the joint null hypothesis that bidders are serially independent in each treatment, we construct a statistic t^i by randomly drawing a number from the uniform distribution U[F(r-1),F(r)]. A particular realization of this statistic is given in the right column of Table 5.11, 5.12, and 5.13. Under the null hypothesis of serial independence, t^i is uniformly distributed in [0,1]. Figure 5.13, right column presents the empirical CDF of the realized values in each treatment. Kolmogorov-Smirnov (KS) tests show that these distributions are not uniform ($p < 0.01$, two-sided).

For high-cost and low-cost bidders, we apply the same technique to examine whether individuals play a mixed strategy as predicted by a risk-neutral model. We categorize individual bids on [0, 100] into two categories. As the average bids predicted by the Nash equilibrium for high-cost and low-cost bidders are 40 and 50 respectively, 40 is the cutoff for high-cost bidders and 50 is the cutoff for low-cost bidders. Tables 5.14 and 5.18 show bidding counts for each category and the corresponding randomized binomial test result for each bidder. For high-cost bidders, the null hypothesis that bidders choose low bids with probability 0.52 is rejected at the 5% level for 16 out of 36 bidders in the exogenous-entry treatments,[28] 12 bidders in the endogenous-entry-favor-early treatments, and 12 bidders in the endogenous-entry-favor-late treatments.[29] For low-cost bidders, the null hypothesis that bidders choose low bids with probability 0.5 is rejected at the 5% level for 18 out of 36 bidders in the exogenous-entry treatments, 9 bidders in the endogenous-entry-favor-early treatments, and 10 bidders in the endogenous-entry-favor-late treatments.[30] Kolmogorov-Smirnov (KS) test also rejects the joint null hypothesis that high(low)-cost bidders in each treatment bid low bids with probability 0.52 (0.5). Figures 5.14 and 5.15, left column present the empirical CDF of p-values for high- and low-cost bidders in each treatment. In addition, regarding serial independence, for high-cost bidders, there are 8 rejections in the exogenous-entry treatments, 12 rejections in the endogenous-entry-favor-early treatments, and 12 rejections in the endogenous-entry-favor-late treatments. For low-cost bidders, there are 15 rejections in the exogenous-entry treatments, 6 rejections in the endogenous-entry-favor-early treatments, and 6 rejections in the endogenous-entry-favor-late treatments. Tables 5.15,

[27] The expected rejection number is 1.8 bidders ($36 * 5\% = 1.8$).

[28] There is a mass point at bid 0 with probability 0.2 for high-cost bidders.

[29] The expected rejection number is 1.8 bidders ($36 * 5\% = 1.8$).

[30] The expected rejection number is 1.8 bidders ($36 * 5\% = 1.8$).

5.16, 5.17, 5.19, 5.20, and 5.21 report the data and results for the test of serial independence. Kolmogorov-Smirnov (KS) tests also reject the joint null hypothesis that bidders are serially independent in each treatment. Figures 5.14 and 5.15, right column present the empirical CDF of the realized values of t^i in each treatment for both high- and low-cost bidders and are significantly different with uniform distributions ($p < 0.01$, two-sided).

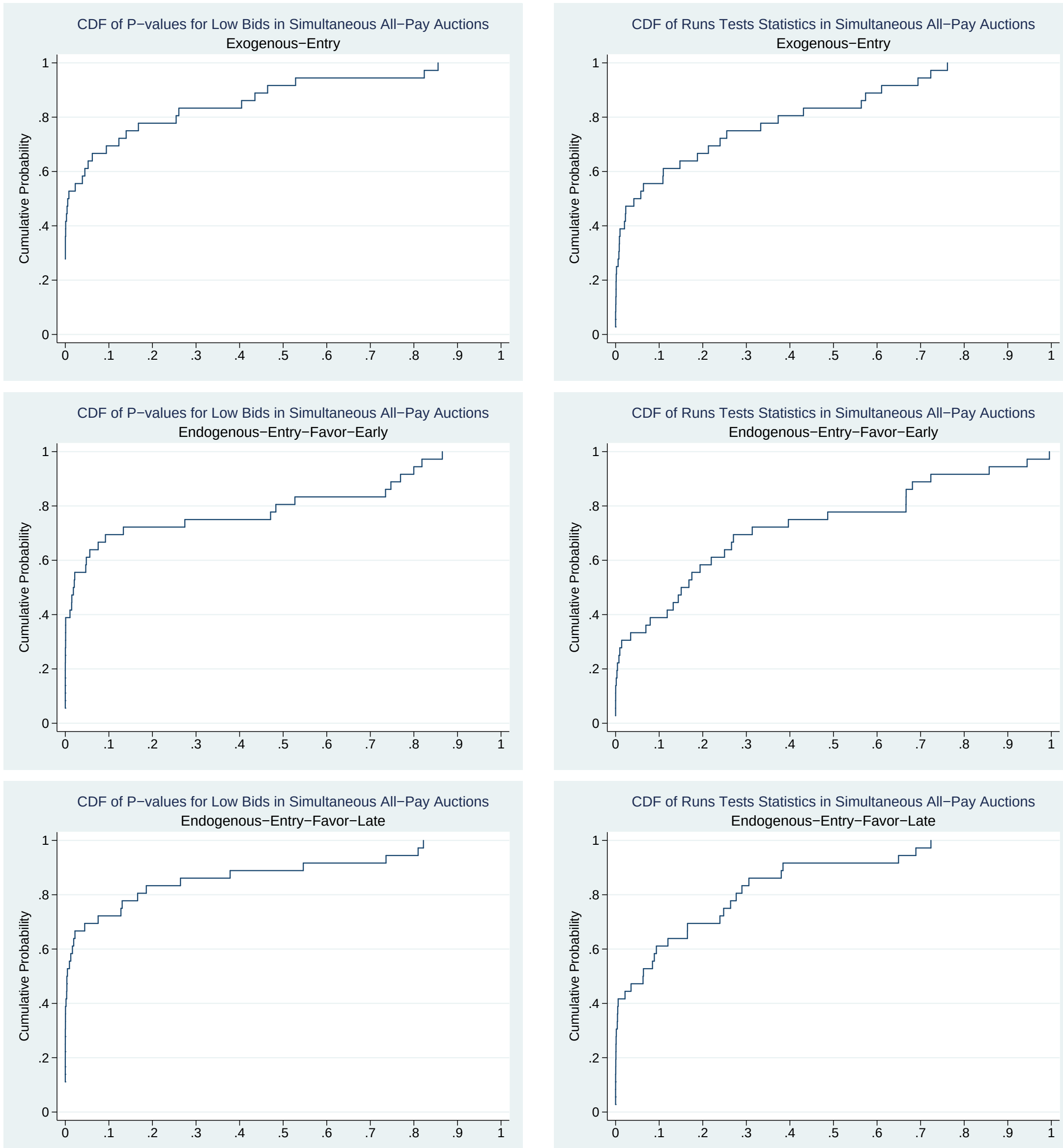


Figure 5.13 Randomized Binomial Tests and Runs Tests in Simultaneous All-Pay Auctions: Homogeneous Bidders

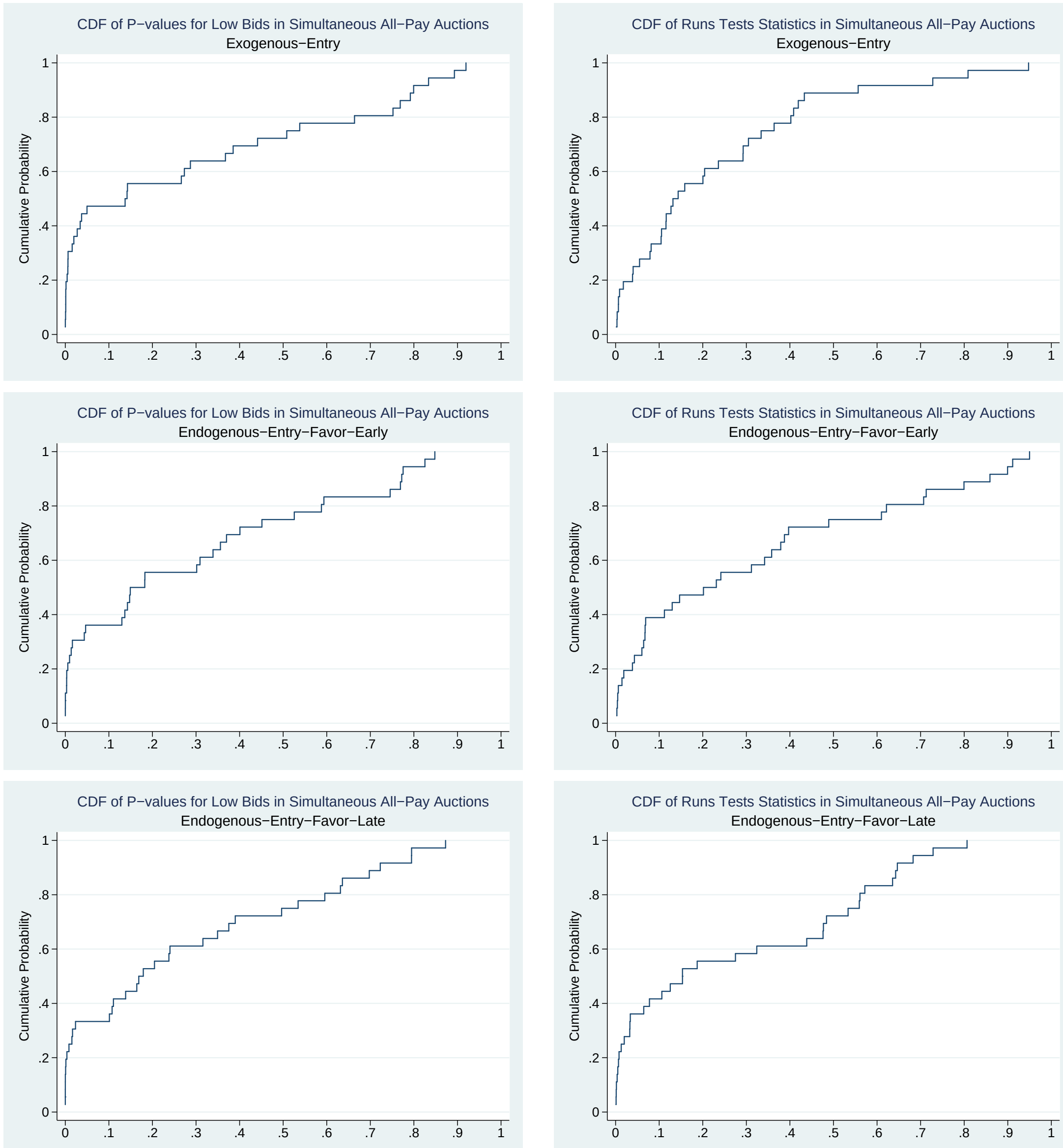


Figure 5.14 Randomized Binomial Tests and Runs Tests in Simultaneous All-Pay Auctions: High-Cost Bidders

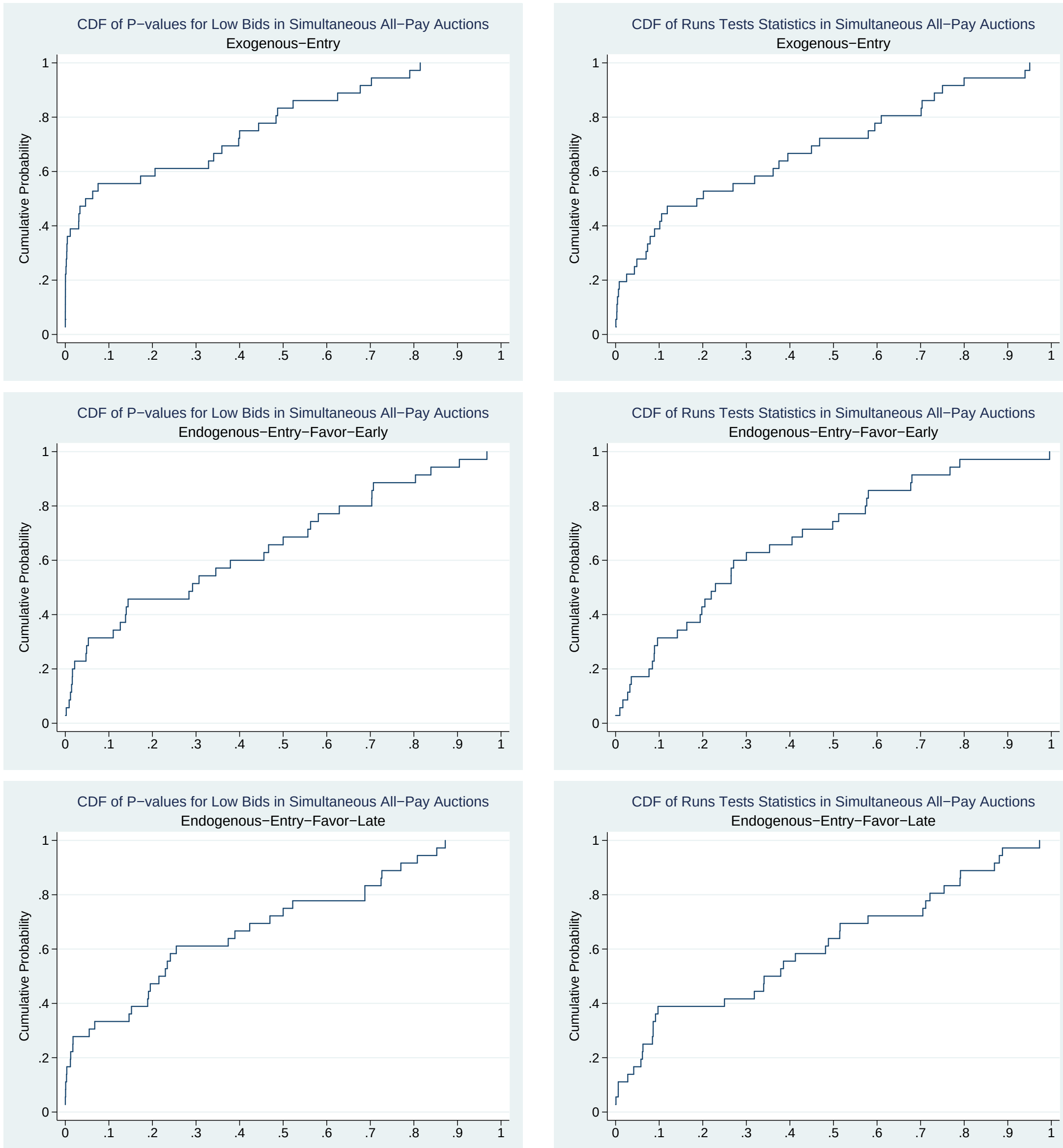


Figure 5.15 Randomized Binomial Tests and Runs Tests in Simultaneous All-Pay Auctions: Low-Cost Bidders

Table 5.10 Bids for Homogeneous Bidders in Simultaneous All-Pay Auctions

Treatment		Exogenous-Entry		Endogenous-Entry-Favor-Early		Endogenous-Entry-Favor-Late	
Session	Subject	[0,50]	(50,100]	[0,50]	(50,100]	[0,50]	(50,100]
1	1	7**	23**	13	15	24**	4**
	2	20	9**	13	14	18	12
	3	14	7	23 **	6 **	30 **	0 **
	4	0**	30**	10	19	29 **	1 **
	5	20	10	26 **	1 **	24 **	6 **
	6	23**	7**	23 **	5 **	16	12
	7	21**	7**	28 **	1 **	7 **	23 **
	8	2**	28**	12	15	23 **	6 **
	9	18	10	13	14	9	17
	10	26**	4**	24 **	4 **	20 **	5 **
	11	24**	0**	16	9	30 **	0 **
	12	26**	4**	3 **	25 **	29 **	1 **
2	1	7**	18**	8 **	19 **	14	12
	2	13	17	21 **	8 **	21 **	8 **
	3	18	10	8	12	12	15
	4	3**	16**	21 **	4 **	19	10
	5	29**	1**	23 **	5 **	22 **	7 **
	6	16	11	12	16	3 **	24 **
	7	22**	7**	19 **	8	6 **	21 **
	8	3**	27**	5 **	17 **	21 **	9 **
	9	10	10	12	15	1 **	15 **
	10	13	17	8 **	18 **	4 **	25 **
	11	16	12	1 **	20 **	13	15
	12	5**	25**	12	10	19	11
3	1	29**	1**	4 **	23 **	24 **	5**
	2	9	15	17	7	3**	26**
	3	15	15	11	10	4 **	25**
	4	8	18**	3 **	24 **	21 **	8 **
	5	21**	9**	0 **	24 **	19	9
	6	29**	1**	3 **	0	19	10
	7	3**	17**	5**	16 **	24**	1**
	8	18	12	24**	0**	7**	20**
	9	11	19	9**	1**	7**	21**
	10	0**	21**	23**	0**	5**	23**
	11	23**	5**	10	12	15	14
	12	25**	5**	7	15	11	18

Table 5.11 Runs Tests for Homogeneous Bidders in Exogenous-Entry Treatments

Session	Subject	[0,50]	(50,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	7	23	6	0.001	0.004**	0.002
	2	20	9	8	0.005	0.016**	0.010
	3	14	7	4	0.000	0.002**	0.001
	4	0	30	1	0.000	1**	0.993
	5	20	10	10	0.022	0.055	0.038
	6	23	7	8	0.018	0.048	0.034
	7	21	7	12	0.482	0.639	0.591
	8	2	28	2	0.000	0.005**	0.004
	9	18	10	15	0.536	0.661	0.569
	10	26	4	7	0.108	0.371	0.310
	11	24	0	1	0.000	1**	0.421
	12	26	4	2	0.000	0.000**	0.000
2	1	7	18	1	0.363	0.393	0.375
	2	13	17	19	0.632	0.608	0.621
	3	18	10	12	0.139	0.258	0.156
	4	3	16	3	0.002	0.020**	0.007
	5	29	1	2	0.000	0.067	0.035
	6	16	11	11	0.076	0.124	0.114
	7	22	7	11	0.259	0.459	0.456
	8	3	27	5	0.033	0.2	0.196
	9	10	10	13	0.673	0.731	0.676
	10	13	17	7	0.000	0.001**	0.001
	11	16	12	6	0.000	0.000**	0.000
	12	5	25	5	0.002	0.010**	0.002
3	1	29	1	3	0.067	1	0.536
	2	9	15	6	0.001	0.005**	0.004
	3	15	15	14	0.123	0.24	0.172
	4	8	18	8	0.017	0.047	0.037
	5	21	9	4	0.000	0.000**	0.000
	6	29	1	3	0.067	1	0.109
	7	3	17	6	0.298	0.509	0.427
	8	18	12	10	0.011	0.029	0.019
	9	11	19	3	0.000	0.000**	0.000
	10	0	21	1	0.000	1**	0.763
	11	23	5	5	0.002	0.013**	0.002
	12	25	5	5	0.002	0.010**	0.002

Table 5.12 Runs Tests for Homogeneous Bidders in Endogenous-Entry-Favor-Early Treatments

Session	Subject	[0,50]	(50,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	13	15	12	0.064	0.149	0.065
	2	13	14	8	0.002	0.009**	0.006
	3	23	6	5	0.001	0.003**	0.003
	4	10	19	9	0.010	0.028	0.015
	5	26	1	3	0.000	0.926	0.588
	6	23	5	3	0.000	0.000**	0.000
	7	28	1	3	0.069	1	0.548
	8	12	15	5	0.000	0.000**	0.000
	9	13	14	13	0.189	0.248	0.238
	10	24	4	6	0.049	0.123	0.090
	11	16	9	5	0.000	0.214	0.024
	12	3	25	5	0.038	0.214	0.135
2	1	8	19	13	0.388	0.415	0.411
	2	21	8	9	0.028	0.077	0.075
	3	8	12	8	0.067	0.159	0.082
	4	21	4	9	0.617	1	1.000
	5	23	5	2	0.000	0.000**	0.000
	6	12	16	14	0.206	0.358	0.349
	7	19	8	11	0.197	0.334	0.300
	8	5	17	9	0.398	0.696	0.410
	9	12	15	6	0.000	0.001**	0.000
	10	8	18	7	0.005	0.017**	0.008
	11	1	20	3	0.095	1	0.887
	12	12	10	10	0.142	0.271	0.168
3	1	4	23	2	0.000	0.000**	0.000
	2	17	7	11	0.397	0.591	0.531
	3	11	10	13	0.606	0.681	0.660
	4	3	24	7	0.395	1	0.828
	5	0	24	1	0.000	1**	0.607
	6	3	0	1	0.000	1**	0.000
	7	5	16	4	0.001	0.007**	0.003
	8	24	0	1	0.000	1**	0.521
	9	9	1	2	0.000	0.2	0.008
	10	23	0	1	0.000	1**	0.546
	11	10	12	10	0.142	0.271	0.231
	12	7	15	6	0.006	0.022**	0.015

Table 5.13 Runs Tests for Homogeneous Bidders in Endogenous-Entry-Favor-Late Treatments

Session	Subject	[0,50]	(50,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	24	4	9	0.568	1	0.871
	2	18	12	4	0.000	0.000**	0.000
	3	30	0	1	0.000	1**	0.914
	4	29	1	2	0.000	0.067	0.041
	5	24	6	2	0.000	0.000**	0.000
	6	16	12	8	0.002	0.007**	0.005
	7	7	23	5	0.000	0.001**	0.001
	8	23	6	7	0.013	0.05	0.031
	9	9	17	5	0.000	0.001**	0.000
	10	20	5	8	0.179	0.325	0.257
	11	30	0	1	0.000	1**	0.921
	12	29	1	3	0.067	1	0.873
2	1	14	12	11	0.085	0.129	0.104
	2	21	8	7	0.002	0.009**	0.002
	3	12	15	9	0.010	0.026	0.024
	4	19	10	8	0.003	0.010**	0.003
	5	22	7	9	0.055	0.144	0.132
	6	3	24	6	0.222	0.395	0.343
	7	6	21	10	0.322	0.486	0.406
	8	21	9	11	0.085	0.161	0.092
	9	1	15	3	0.125	1	0.446
	10	4	25	4	0.001	0.007**	0.002
	11	13	15	6	0.000	0.000**	0.000
	12	19	11	3	0.000	0.000**	0.000
3	1	24	5	8	0.135	0.254	0.178
	2	3	26	7	0.371	1	0.503
	3	4	25	2	0.000	0.000**	0.000
	4	21	8	6	0.000	0.002**	0.001
	5	19	9	6	0.000	0.002**	0.001
	6	19	10	13	0.228	0.345	0.315
	7	24	1	3	0.080	1	0.728
	8	7	20	5	0.000	0.002**	0.000
	9	7	21	11	0.281	0.306	0.294
	10	5	23	10	0.583	0.732	0.590
	11	15	14	12	0.048	0.114	0.061
	12	11	18	6	0.000	0.000**	0.000

Table 5.14 Bids for High-Cost Bidders in Simultaneous All-Pay Auctions

Treatment		Exogenous-Entry		Endogenous-Entry-Favor-Early		Endogenous-Entry-Favor-Late	
Session	Subject	[0,40]	(40,100]	[0,40]	(40,100]	[0,40]	(40,100]
1	1	4	0**	4	4	3	8
	2	10**	2**	12**	0**	12	5
	3	6	6	7	6	3	5
	4	7	13	4	5	8	4
	5	8	7	9	3	4	3
	6	3	6	12**	2**	5	3
	7	11**	0**	14**	4**	11**	0**
	8	4	5	9**	0**	8	2
	9	14**	3**	12	4	7	3
	10	14**	1**	9	5	0**	13**
	11	8	5	10	7	6	5
	12	4	4	14**	5	10	4
2	1	2**	10**	12	5	12**	1**
	2	8	8	12**	0**	10	3
	3	11**	1**	6	5	9	4
	4	12**	3**	2	3	5	5
	5	11	5	2	3	3	5
	6	0**	11**	10**	1**	8	9
	7	5	9	1	0	0**	13**
	8	9	8	4	1	1**	12**
	9	10	6	6	3	14**	3**
	10	6	8	9**	0**	6	3
	11	13**	1**	9**	0**	5	9
	12	4	8	3**	12**	1**	8**
3	1	1**	9**	1	1	5	5
	2	4	11	9	3	7	3
	3	3**	12**	1	5	3	3
	4	14**	1**	13**	2**	10	4
	5	1**	14**	11	4	15**	2**
	6	1	3	7	3	11**	2**
	7	10	8	2	6	0**	14**
	8	6	6	10	4	19**	0**
	9	11**	2**	8	8	4	5
	10	7	5	6	6	1	5**
	11	12**	1**	8	4	6	3
	12	18**	0**	17**	0**	12**	3**

Table 5.15 Runs Tests for High-Cost Bidders in Exogenous-Entry Treatments

Session	Subject	[0,40]	(40,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	4	0	1	0.000	1**	0.614
	2	10	2	4	0.182	0.455	0.401
	3	6	6	3	0.002	0.013**	0.006
	4	7	13	5	0.002	0.010**	0.008
	5	8	7	6	0.051	0.149	0.137
	6	3	6	5	0.345	0.643	0.469
	7	11	0	1	0.000	1**	0.290
	8	4	5	4	0.071	0.262	0.117
	9	14	3	2	0.000	0.003**	0.001
	10	14	1	2	0.000	0.133	0.130
	11	8	5	8	0.576	0.793	0.602
	12	4	4	2	0.000	0.029	0.005
2	1	2	10	4	0.182	0.455	0.185
	2	8	8	6	0.032	0.100	0.060
	3	11	1	2	0.000	0.167	0.044
	4	12	3	4	0.033	0.130	0.082
	5	11	5	5	0.022	0.077	0.068
	6	0	11	1	0.000	1**	0.421
	7	5	9	3	0.001	0.007**	0.005
	8	9	8	9	0.319	0.500	0.404
	9	10	6	5	0.013	0.047	0.026
	10	6	8	9	0.646	0.821	0.764
	11	13	1	2	0.000	0.143	0.001
	12	4	8	5	0.109	0.279	0.246
3	1	1	9	3	0.200	1.000	0.961
	2	4	11	6	0.176	0.374	0.319
	3	3	12	4	0.033	0.130	0.082
	4	14	1	3	0.133	1.000	0.214
	5	1	14	2	0.000	0.133	0.073
	6	1	3	2	0.000	0.500	0.461
	7	10	8	8	0.117	0.251	0.156
	8	6	6	6	0.175	0.392	0.348
	9	11	2	4	0.167	0.423	0.308
	10	7	5	6	0.197	0.424	0.421
	11	12	1	2	0.000	0.154	0.073
	12	18	0	1	0.000	1**	0.109

Table 5.16 Runs Tests for High-Cost Bidders in Endogenous-Entry-Favor-Early Treatments

Session	Subject	[0,40]	(40,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	4	4	4	0.114	0.371	0.192
	2	12	0	1	0.000	1**	0.142
	3	7	6	8	0.500	0.733	0.584
	4	4	5	3	0.016	0.071	0.037
	5	9	3	2	0.000	0.009**	0.002
	6	12	2	4	0.154	0.396	0.184
	7	14	4	5	0.031	0.121	0.058
	8	9	0	1	0.000	1**	0.193
	9	12	4	9	0.819	1	0.933
	10	9	5	10	0.902	0.972	0.967
	11	10	7	9	0.355	0.549	0.468
	12	14	5	8	0.299	0.496	0.488
2	1	12	5	3	0.000	0.003**	0.002
	2	12	0	1	0.000	1**	0.579
	3	6	5	6	0.262	0.522	0.304
	4	2	3	2	0.000	0.2	0.028
	5	2	3	4	0.500	0.9	0.520
	6	10	1	2	0.000	0.182	0.153
	7	1	0	1	0.000	1**	0.866
	8	4	1	2	0.000	0.4	0.193
	9	6	3	4	0.107	0.345	0.115
	10	9	0	1	0.000	1**	0.686
	11	9	0	1	0.000	1**	0.960
	12	3	12	3	0.004	0.033	0.017
3	1	1	1	2	0.000	1	0.381
	2	9	3	2	0.000	0.009**	0.009
	3	1	5	2	0.000	0.333	0.217
	4	13	2	3	0.019	0.143	0.042
	5	11	4	6	0.176	0.374	0.365
	6	7	3	3	0.017	0.083	0.056
	7	2	6	5	0.643	1	0.949
	8	10	4	4	0.014	0.068	0.066
	9	8	8	6	0.032	0.1	0.065
	10	6	6	3	0.002	0.013**	0.008
	11	8	4	3	0.004	0.024**	0.016
	12	17	0	1	0.000	1**	0.954

Table 5.17 Runs Tests for High-Cost Bidders in Endogenous-Entry-Favor-Late Treatments

Session	Subject	[0,40]	(40,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	3	8	6	0.533	0.788	0.637
	2	12	5	4	0.003	0.017**	0.015
	3	3	5	5	0.429	0.714	0.696
	4	8	4	6	0.279	0.533	0.519
	5	4	3	2	0.000	0.057	0.046
	6	5	3	4	0.143	0.429	0.195
	7	11	0	1	0.000	1**	0.901
	8	8	2	5	0.533	1	0.688
	9	7	3	4	0.083	0.283	0.190
	10	0	13	1	0.000	1**	0.127
	11	6	5	6	0.262	0.522	0.423
	12	10	4	2	0.000	0.002**	0.000
2	1	12	1	3	0.154	1	0.437
	2	10	3	2	0.000	0.007**	0.004
	3	9	4	2	0.000	0.003**	0.002
	4	5	5	4	0.040	0.167	0.054
	5	3	5	5	0.429	0.714	0.585
	6	8	9	4	0.001	0.005**	0.005
	7	0	13	1	0.000	1**	0.588
	8	1	12	2	0.000	0.154	0.001
	9	14	3	6	0.350	0.579	0.535
	10	6	3	3	0.024	0.107	0.071
	11	5	9	4	0.007	0.039	0.028
	12	1	8	2	0.000	0.222	0.191
3	1	5	5	3	0.008	0.04	0.039
	2	7	3	4	0.083	0.283	0.270
	3	3	3	4	0.300	0.7	0.554
	4	10	4	2	0.000	0.002**	0.002
	5	15	2	2	0.000	0.015**	0.000
	6	11	2	2	0.000	0.026	0.001
	7	0	14	1	0.000	1**	0.367
	8	19	0	1	0.000	1**	0.694
	9	4	5	6	0.500	0.786	0.706
	10	1	5	3	0.333	1	0.998
	11	6	3	5	0.345	0.643	0.375
	12	12	3	4	0.033	0.13	0.050

Table 5.18 Bids for Low-Cost Bidders in Simultaneous All-Pay Auctions

Treatment		Exogenous-Entry		Endogenous-Entry-Favor-Early		Endogenous-Entry-Favor-Late	
Session	Subject	[0,50]	(50,100]	[0,50]	(50,100]	[0,50]	(50,100]
1	1	7**	0**	*bids* > 100		2**	9**
	2	1	3	4	12	4	3
	3	7	9	8	7	1	7
	4	2	5	8	8	3	6
	5	4	10	7	9	3	3
	6	15**	1**	5	8	14**	3**
	7	14**	0**	3	7	5	10
	8	8	2**	2	6	3	4
	9	13**	0**	4	5	1	5
	10	12**	0**	5	7	0	3**
	11	10	6	8	5	3	4
	12	0	1	4	5	7	6
2	1	12**	4**	1	1	11**	1**
	2	0**	6**	1	0	4	4
	3	4	9	2**	7	3**	12**
	4	13**	2**	2	3	3	3
	5	7	7	1	2	3	4
	6	6	8	5	3	5	2
	7	6	9	3	1	2	4
	8	1**	12**	9	2	13**	0**
	9	12**	2**	11	3	7	2
	10	1**	12**	6	3	7	13
	11	3**	10	8**	1**	10	3**
	12	4	5	9**	0**	1	2
3	1	7	4	4	7	8	5
	2	0**	13**	11**	3**	1	2
	3	2	2	0**	7**	1**	14**
	4	2**	13**	11**	2**	14**	1**
	5	1**	13**	4	9	4	2
	6	0**	15**	7	3	0**	6**
	7	7	5	8	7	0	2**
	8	9	9	11**	2**	1	1
	9	6	3	5	7	5	1
	10	6	4	12**	3**	1	3
	11	13**	4**	3	8	19**	0**
	12	10**	0**	11**	0**	11**	2**

Table 5.19 Runs Tests for Low-Cost Bidders in Exogenous-Entry Treatments

Session	Subject	[0,50]	(50,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	7	0	1	0.000	1**	0.431
	2	1	3	3	0.500	1.000	0.696
	3	7	9	3	0.000	0.001**	0.000
	4	2	5	3	0.095	0.333	0.102
	5	4	10	6	0.203	0.419	0.386
	6	15	1	3	0.125	1.000	0.527
	7	14	0	1	0.000	1**	0.203
	8	8	2	3	0.044	0.222	0.155
	9	13	0	1	0.000	1**	0.306
	10	12	0	1	0.000	1**	0.288
	11	10	6	4	0.002	0.013**	0.004
	12	0	1	1	0.000	1**	0.240
2	1	12	4	6	0.154	0.335	0.303
	2	0	6	1	0.000	1**	0.623
	3	4	9	4	0.018	0.085	0.076
	4	13	2	4	0.143	0.371	0.199
	5	7	7	5	0.025	0.078	0.035
	6	6	8	8	0.413	0.646	0.499
	7	6	9	6	0.063	0.175	0.095
	8	1	12	2	0.000	0.154	0.112
	9	12	2	5	0.396	1.000	0.708
	10	1	12	3	0.154	1.000	0.850
	11	3	10	2	0.000	0.007**	0.005
	12	4	5	5	0.262	0.500	0.319
3	1	7	4	7	0.606	0.833	0.711
	2	0	13	1	0.000	1**	0.038
	3	2	2	2	0.000	0.333	0.197
	4	2	13	2	0.000	0.019**	0.002
	5	1	13	2	0.000	0.143	0.074
	6	0	15	1	0.000	1**	0.230
	7	7	5	2	0.000	0.003**	0.000
	8	9	9	6	0.012	0.044	0.018
	9	6	3	3	0.024	0.107	0.060
	10	6	4	2	0.000	0.010**	0.001
	11	13	4	4	0.007	0.037	0.008
	12	10	0	1	0.000	1**	0.208

Table 5.20 Runs Tests for Low-Cost Bidders in Endogenous-Entry-Favor-Early Treatments

Session	Subject	[0,50]	(50,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1						
	2	4	12	4	0.009	0.045	0.025
	3	8	7	7	0.149	0.296	0.295
	4	8	8	6	0.032	0.1	0.096
	5	7	9	2	0.000	0.000**	0.000
	6	5	8	3	0.002	0.010**	0.003
	7	3	7	4	0.083	0.283	0.212
	8	2	6	3	0.071	0.286	0.204
	9	4	5	5	0.262	0.5	0.496
	10	5	7	7	0.424	0.652	0.576
	11	8	5	6	0.152	0.347	0.156
	12	4	5	4	0.071	0.262	0.204
2	1	1	1	2	0.000	1	0.604
	2	1	0	1	0.000	1**	0.670
	3	2	7	3	0.056	0.25	0.113
	4	2	3	2	0.000	0.2	0.051
	5	1	2	3	0.667	1	0.758
	6	5	3	5	0.429	0.714	0.711
	7	3	1	2	0.000	0.5	0.325
	8	9	2	2	0.000	0.036	0.023
	9	11	3	5	0.148	0.423	0.345
	10	6	3	4	0.107	0.345	0.200
	11	8	1	2	0.000	0.222	0.085
	12	9	0	1	0.000	1**	0.062
3	1	4	7	4	0.033	0.142	0.119
	2	11	3	5	0.148	0.423	0.287
	3	0	7	1	0.000	1**	0.640
	4	11	2	5	0.423	1	0.430
	5	4	9	6	0.236	0.471	0.284
	6	7	3	5	0.283	0.583	0.370
	7	8	7	6	0.051	0.149	0.101
	8	11	2	2	0.000	0.026	0.020
	9	5	7	5	0.076	0.197	0.077
	10	12	3	6	0.396	0.637	0.407
	11	3	8	5	0.236	0.533	0.421
	12	11	0	1	0.000	1**	0.086

Table 5.21 Runs Tests for Low-Cost Bidders in Endogenous-Entry-Favor-Late Treatments

Session	Subject	[0,50]	(50,100]	Runs	F(r-1)	F(r)	U[F(r-1),F(r)]
1	1	2	9	2	0.000	0.036	0.008
	2	4	3	2	0.000	0.057	0.016
	3	1	7	3	0.250	1.000	0.456
	4	3	6	4	0.107	0.345	0.133
	5	3	3	4	0.300	0.700	0.527
	6	14	3	5	0.101	0.350	0.345
	7	5	10	4	0.005	0.029	0.026
	8	3	4	5	0.543	0.800	0.791
	9	1	5	2	0.000	0.333	0.237
	10	0	3	1	0.000	1**	0.856
	11	3	4	4	0.200	0.543	0.308
	12	7	6	7	0.296	0.500	0.500
2	1	11	1	3	0.167	1.000	0.419
	2	4	4	2	0.000	0.029	0.010
	3	3	12	4	0.033	0.130	0.124
	4	3	3	2	0.000	0.100	0.049
	5	3	4	4	0.200	0.543	0.529
	6	5	2	3	0.095	0.333	0.129
	7	2	4	4	0.400	0.800	0.713
	8	13	0	1	0.000	1**	0.872
	9	7	2	3	0.056	0.250	0.233
	10	7	13	2	0.000	0.000**	0.000
	11	10	3	4	0.045	0.171	0.090
	12	1	2	3	0.667	1.000	0.869
3	1	8	5	5	0.054	0.152	0.109
	2	1	2	2	0.000	0.667	0.479
	3	1	14	3	0.133	1.000	0.416
	4	14	1	2	0.000	0.133	0.111
	5	4	2	4	0.400	0.800	0.425
	6	0	6	1	0.000	1**	0.809
	7	0	2	1	0.000	1**	0.094
	8	1	1	2	0.000	1.000	0.161
	9	5	1	2	0.000	0.333	0.215
	10	1	3	3	0.500	1.000	0.918
	11	19	0	1	0.000	1**	0.809
	12	11	2	3	0.026	0.167	0.133

5.10 Appendix C: Experimental Instruction

This is an experiment in decision making. The experiment will proceed in two parts and you will make a series of decisions in each part. At the end, you will fill out a post-experiment questionnaire.

This experiment has 12 participants. Each of you has been randomly assigned an experiment ID at the beginning of the experiment. The experimenter will use this ID to pay you at the end of the experiment.

Part 1 of the Experiment Rounds: The first part of the experiment consists of 30 rounds of two-person auctions.

Endowment: Each of you has 125 tokens as the endowment at the beginning of each round.

Prize Values: At the beginning of each round, an object with a value of 100 tokens will be auctioned within each two-person group.

Matching: At the beginning of each round, you will be randomly matched with another person. You are equally likely to be matched with any other person in the room.

Decisions: In each round, you must make two decisions. First, both you and your match choose independently and simultaneously which bidding stage you want to enter. The entry decisions are then announced to both of you. Second, you and your match each choose a bid in your respective bidding stage.

Bids: There are two bidding stages: the early stage and the late stage.

1. If you enter early and your match enters late, you will choose a bid first. After observing your bid, your match will choose his or her bid.
2. If you enter late and your match enters early, your match will choose a bid first. After observing his or her bid, you will choose your own bid.
3. If both of you choose the same stage, you will each bid simultaneously.

Cost of the Bid: The cost of the bid captures the idea that it is sometimes more or less costly to submit a bid. In the experiment, it is determined by a random number generator at the beginning of each round. For each round, with 50% chance, the cost of the bid is 1 token for you and 0.8 tokens for your match. With 50% chance, the cost of the bid is 0.8 tokens for you and 1 token for your match. Here is a numerical example:

1. If the cost of your bid is 1 token and you bid 50, then you will pay 50 tokens.
2. If the cost of your bid is 0.8 tokens and you bid 50, then you will pay 40 tokens.

Bid Range: Your bid can be any integer between 0 and 125, inclusive.

Profits: In each round, your profits will be determined by (1) your bid; (2) your match's bid; (3) the cost of your bid; and (4) the entry decisions in the event of a tie.

Profits= Your Endowment- the cost of your bid*your bid + the value of the object if you win = 125 - the cost of your bid*your bid+ 100 if you win

For example, in a given round, if you bid 40 and the cost of your bid is 0.8 in this round, then

1. If you win the auction, then your profit is $125 - 0.8 * 40 + 100 = 193$ tokens
2. If you lose the auction, then your profit is $125 - 0.8 * 40 = 93$ tokens

Note: You will always pay for your bid, which is equal to the cost of your bid* your bid, no matter whether you win or lose.

The tie-breaking rule: If you and your match bid exactly the same amount, and

1. Both of you enter in the same stage, we will randomly choose one as the winner.
2. If one and only one of you enter in the early stage, then the early bidder will be the winner.

Cumulative Profits: Your cumulative profits will be the sum of your profits in all rounds.

Feedback: At the end of each round, you will get the following feedback on your screen:

1. Your entry decision
2. Your match's entry decision
3. Your bid
4. Your match's bid
5. Your profits
6. Your match's profits
7. Your cumulative profits

History: In each round, your and your matches' bids and entry decisions in each previous round, your and your matches' profits in each previous round, as well as your cumulative profits up to the last round will be displayed in a history box.

Review Questions: To help you understand the experiment, we will go over nine review questions before we start the auction. You can also find these review questions in the appendix for your reference. You will get 25 tokens for answering each of the review questions correctly.

Exchange Rate: $1 = 250 tokens.

Please do not communicate with each other during the experiment. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

Part 2 of the Experiment

After 30 rounds of auction, you will be making choices for three series of paired lotteries, such as those represented as "Option A" and "Option B" below. In both series 1 and 2, there are 14 lottery questions. In series 3, there are 7 lottery questions.

For each series, you are asked to choose a "switch" question from Option A to Option B. For example, you can choose to switch from Option A to Option B at

Question 6, which means you will choose Option A from Question 1 to Question 5 while you will choose Option B from Question 6 to the end of this series. You can also choose to never switch to Option B, which means you always choose Option A for all questions in this series. You can also choose to switch to Option B from Question 1, which means you always choose Option B for all questions in this series.

Even though there are 35 lottery questions, only one of them will end up being used. The selection of the one to be used depends on a random number generator, which is the equivalent of throwing a 35-sided die. Each lottery (Series 1: 1-14; Series 2: 1-14; Series 3: 1-7) is equally likely to be chosen.

After the lottery question is chosen, the money prize that you receive is determined by another random number generator, which is equivalent of throwing a ten-sided die. Each outcome, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10, is equally likely to be chosen. For example, if the number drawn from the first random number generator is 29, then the 29th lottery, which is the first lottery in series 3, is chosen. Furthermore, consider if the number drawn from the second random number generator is 4. If you chose Question 2 as the switch question for series 3, which means you choose Option A for question 1, then you will get 25 tokens. If you chose Question 1 as the switch question for series 3, which means you chose Option B for Question 1, you will get 30 tokens.

Review Questions: To help you understand the lottery, we will go over one review question before we start it. You will also get 25 tokens for answering the review question correctly.

Recall the exchange rate is still $1 = 250 tokens in the lottery.

Final Payment: Your final payment in this experiment will be

Your Earnings in the Review Question Part (for both auction and lottery) + Your Cumulative Profits in the Auction (Part1) + Your Profits in the lottery (Part2) + Participation Fee ($ 5)

Lottery Choice

Series1-Lottery Number	Option A	Option B
1	40 if the die is 1-3 10 if the die is 4-10	68 if the die is 1 5 if the die is 2-10
2	40 if the die is 1-3 10 if the die is 4-10	75 if the die is 1 5 if the die is 2-10
3	40 if the die is 1-3 10 if the die is 4-10	83 if the die is 1 5 if the die is 2-10
4	40 if the die is 1-3 10 if the die is 4-10	93 if the die is 1 5 if the die is 2-10
5	40 if the die is 1-3 10 if the die is 4-10	106.5 if the die is 1 5 if the die is 2-10
6	40 if the die is 1-3 10 if the die is 4-10	125 if the die is 1 5 if the die is 2-10
7	40 if the die is 1-3 10 if the die is 4-10	150 if the die is 1 5 if the die is 2-10
8	40 if the die is 1-3 10 if the die is 4-10	185 if the die is 1 5 if the die is 2-10
9	40 if the die is 1-3 10 if the die is 4-10	220 if the die is 1 5 if the die is 2-10
10	40 if the die is 1-3 10 if the die is 4-10	300 if the die is 1 5 if the die is 2-10
11	40 if the die is 1-3 10 if the die is 4-10	400 if the die is 1 5 if the die is 2-10
12	40 if the die is 1-3 10 if the die is 4-10	600 if the die is 1 5 if the die is 2-10
13	40 if the die is 1-3 10 if the die is 4-10	1000 if the die is 1 5 if the die is 2-10
14	40 if the die is 1-3 10 if the die is 4-10	1700 if the die is 1 5 if the die is 2-10

Series2-Lottery Number	Option A	Option B
1 (15)	40 if the die is 1-9 30 if the die is 10	54 if the die is 1-7 5 if the die is 8-10
2 (16)	40 if the die is 1-9 30 if the die is 10	56 if the die is 1-7 5 if the die is 8-10
3 (17)	40 if the die is 1-9 30 if the die is 10	58 if the die is 1-7 5 if the die is 8-10
4 (18)	40 if the die is 1-9 30 if the die is 10	60 if the die is 1-7 5 if the die is 8-10
5 (19)	40 if the die is 1-9 30 if the die is 10	62 if the die is 1-7 5 if the die is 8-10
6 (20)	40 if the die is 1-9 30 if the die is 10	65 if the die is 1-7 5 if the die is 8-10
7 (21)	40 if the die is 1-9 30 if the die is 10	68 if the die is 1-7 5 if the die is 8-10
8 (22)	40 if the die is 1-9 30 if the die is 10	72 if the die is 1-7 5 if the die is 8-10
9 (23)	40 if the die is 1-9 30 if the die is 10	77 if the die is 1-7 5 if the die is 8-10
10 (24)	40 if the die is 1-9 30 if the die is 10	83 if the die is 1-7 5 if the die is 8-10
11 (25)	40 if the die is 1-9 30 if the die is 10	90 if the die is 1-7 5 if the die is 8-10
12 (26)	40 if the die is 1-9 30 if the die is 10	100 if the die is 1-7 5 if the die is 8-10
13 (27)	40 if the die is 1-9 30 if the die is 10	110 if the die is 1-7 5 if the die is 8-10
14 (28)	40 if the die is 1-9 30 if the die is 10	130 if the die is 1-7 5 if the die is 8-10

Series3-Lottery Number	Option A	Option B
1 (29)	25 if the die is 1-5 -4 if the die is 6-10	30 if the die is 1-5 -21 if the die is 6-10
2 (30)	4 if the die is 1-5 -4 if the die is 6-10	30 if the die is 1-5 -21 if the die is 6-10
3 (31)	1 if the die is 1-5 -4 if the die is 6-10	30 if the die is 1-5 -21 if the die is 6-10
4 (32)	1 if the die is 1-5 -4 if the die is 6-10	30 if the die is 1-5 -16 if the die is 6-10
5 (33)	1 if the die is 1-5 -8 if the die is 6-10	30 if the die is 1-5 -16 if the die is 6-10
6 (34)	1 if the die is 1-5 -8 if the die is 6-10	30 if the die is 1-5 -14 if the die is 6-10
7 (35)	1 if the die is 1-5 -8 if the die is 6-10	30 if the die is 1-5 -11 if the die is 6-10

5.11 Appendix D: Post-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
   if it is other, please specify:______________
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. From which countries did your family originate?
8. Would you describe yourself as (Please choose one)
   (a) Optimistic
   (b) Pessimistic
   (c) Neither
9. 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
10. In general, do you see yourself as someone who is willing, even eager, to take risks? (1-7 likert scale)

11. Concerning just personal finance decisions, do you see yourself as someone who is willing, even eager, to take risks?
12. In general, do you see yourself as someone who, when faced with an uncertain situation, worries a lot about possible losses?
13. Concerning just personal finance decisions, are you someone who, when faced with an uncertain situation, worries a lot about possible losses?
14. In general, how competitive do you think you are?
15. Concerning just sports and leisure activities, how competitive do you think you are?

    How much do you agree with the following statements? (1-7 likert scale)

    I see myself as someone who
16. is helpful and unselfish with others
17. can be cold and aloof
18. is considerate and kind to almost everyone
19. likes to cooperate with others
20. is often on bad terms with others
21. feels little concern for others
22. is on good terms with nearly everyone
23. I can make my own decisions, uninfluenced by public opinion.
24. It is achievement, rather than popularity with others, that gets you ahead nowadays.
25. I will stick to my opinion if I think I am right, even if others disagree.
26. I will change the opinion I express as a result of an onslaught of criticism, even though I really do not change the way I feel.
27. The important thing in being successful nowadays is not how hard you work, but how well you fit in with the crowd.
28. I am more likely to express my opinion in a group when I see others agree with me.
29. In a given round, you have 125 tokens in your deposit account and you lose 25 of them after participating in the auction. How much money do you think you win or lose? I.e. given that you have 125 tokens in your deposit account already, do you consider it a loss if you end the auction with fewer than your original 125 tokens, or only if you lose the entirety of your endowment?
    (a) I lost 25; I consider ending the auction with any amount less than the original 125 tokens in my personal account to be a loss.
    (b) I won 100 tokens; I only consider the auction's outcome a loss if I lose the entirety of my 125-token endowment.
30. If you chose to enter the early bidding stage in any round, please write down the reason (sequential all-pay treatments only)
31. If you chose to enter the late bidding stage in any round, please write down the reason (sequential all-pay treatments only)

    For female participants only:
32. Are you currently menstruating?
    (a) Yes

(b) No

33. If yes, for how many days have you been menstruating?
34. If no, how many days away are you from the first day of your next menstrual period?
35. How many times do you menstruate a year?
36. On average, how many days are there between your menstrual cycles?
37. How many days does your menstruation last on average?
38. Are you on the pill?
    (a) Yes
    (b) No
39. If yes, what is the name of the birth control pill you are taking?
    (a) Name of the pill
    (b) I don't remember
40. What date was the first day of your last menstrual period?
    (a) Month:
    (b) Day:
41. Do you currently experience any symptoms of PMS (Premenstrual Syndrome)? (please choose one)
    (a) None
    (b) Mild
    (c) Severe

Chapter 6

Conclusion

My dissertation studies two streams of research in experimental and behavioral economics. The first one studies the effect of social context and social identity on individuals' behavior and the second one evaluates the performance of all-pay auctions on crowdsourcing labor markets.

In the first study, we present evidence from laboratory experiments of behavioral spillovers and cognitive load that spread across strategic contexts. In the experiments, subjects play two distinct games simultaneously with different opponents. We find that the strategies chosen and the efficiency of outcomes in one game depend on the other game that the subject plays. Furthermore, the play of the subjects in the experiment is altered in predictable directions. Specifically, using entropy as a measure of behavioral variation in a normal form game, we find that prevalent strategies in games with low outcome entropy are more likely to be used in the games with high outcome entropy, but not vice versa. Our findings suggest that behavior within a particular institution may depend upon the other incentive structures in play and, as a result, institutional outcomes may be context-dependent. In a follow-up experiment, my coauthors, Jenna Bednar, Yan Chen, Scott Page and I study the vertical multiple game effects instead of the horizontal multiple effects. Specifically, we ask subjects to play the same game with two different matches in the first 100 rounds and then replace one of the games with another. Surprisingly, the behavioral spillover effect from the game with the low entropy to others with high entropy does not hold in sequential multiple games. For example, the significant behavioral spillover from SI to other games in the simultaneous multiple game setting disappears when it becomes the historical game.

In the second study, we conduct experiments at two large public universities in the United States and manipulate the salience of participants' multidimensional natural identities to investigate the effects of identity on coordination and cooperation. By priming a fragmenting (ethnic) identity, we find that, compared to the control, Asians exhibit significantly more in-group cooperation and out-group discrimination, while Caucasians are not responsive to ethnic priming. In comparison, priming a common organizational (school) identity effectively reduces intergroup bias for Asians in the coordination game, resulting in a significant increase of both in-group and out-group cooperation. However, in games with a unique inefficient Nash equilibrium, the effects of priming a common organizational identity are more complex. While priming alleviates the negative effects of the competitiveness stereotype on cooperation among Asian students at UCLA, it enhances such effects among Asian student as the University of Michigan. This study suggests that common identity can be an effective non-pecuniary incentive to encourage individuals to choose the pareto efficient equilibrium in coordination games, whereas its effect in games with dominant strategies is weaker and institutionally dependent.

The third and fourth studies investigate the performances of different all-pay auction mechanisms on crowdsourcing labor markets. In a field experiment conducted on Taskcn.com, which is one of the largest Chinese crowdsourcing sites, we evaluate the effect of reward size and a reserve in the form of the early entry of a high-quality submission. Consistent with our theory, a higher reward attracts significantly more submissions and marginally higher submission quality. Unpredicted by the standard theory, a task with a reserve deters entries from high-quality users but not from low-quality users, consequently, we observe significantly lower quality for tasks with a reserve compared to those without a reserve. Moreover, the failure of using a reserve submission to deter the entry of low-quality submissions necessitates further discussions of efficient mechanisms to solve this problem. Taken together, compared to most laboratory experimental studies in the all-pay auction literature, crowdsourcing sites are excellent online labor markets in which to assess all-pay auction mechanisms and our results challenge standard all-pay theories. For example, users on crowdsourcing sites endogenously decide whether to password-protect their submissions or not, therefore, the mechanism is a hybrid between sequential and simultaneous all-pay auctions. Future work should examine the questions of when users decide to password-protect their submissions and which types of users are more likely to password-protect their submissions.

As we observe that both sequential and simultaneous all-pay auctions are implemented on different crowdsourcing websites and as it is unlikely to be able to compare the performance of these mechanisms in the field, we conduct a laboratory experiment to study this question. Our results show that the amount of bids, which approximates submission quality, is significantly higher in simultaneous all-pay auctions than in sequential all-pay auctions. Moreover, we endogenize individuals' entry decisions, which approximates users' entry decisions on crowdsourcing sites, and find that a favor-early tie-breaking rule, which is similar to the "first come, first serve" rule among multiple best submissions, attracts more early entries. However, this effect is attenuated as subjects acquire more experience. A natural extension of this work is to compare the performance of these two mechanisms in incomplete information settings.

In addition, we observe that many research and development contests are carried out at the group level and the individual contributions within a group are often unobservable. It will be interesting to explore individuals' behaviors in group contests and how they are different from those in individual contests.

Bibliography

Abreu, D. and Rubinstein, A. (1988), "The structure of nash equilibrium in repeated games with finite automata", *Econometrica* , Vol. 56, The Econometric Society, pp. pp. 1259–1281.

Adamic, L. A., Zhang, J., Bakshy, E. and Ackerman, M. S. (2008), Knowledge sharing and yahoo answers: Everyone knows something, *in* 'WWW 2008'.

Allen, V. L. and Wilder, D. A. (1975), "Categorization, belief similarity, and intergroup discrimination", *Journal of Personality and Social Psychology* , Vol. 32, pp. 971–977.

Amann, E. and Leininger, W. (1996), "Asymmetric all-pay auctions with incomplete information: The two-player case", *Games and Economic Behavior* , Vol. 14, pp. 1–18.

Anderson, L. R. and Stafford, S. L. (2003), "An experimental analysis of rent seeking under varying competitive conditions", *Public Choice* , Vol. 115, pp. 199–216.

Anderson, S. P., Goeree, J. K. and Holt, C. A. (1998), "Rent seeking with bounded rationality: An analysis of the all-pay auction", *Journal of Political Economy* , Vol. 106, pp. 828–853.

Andreoni, J. and Miller, J. (2002), "Giving according to garp: An experimental test of the consistency of preferences for altruism", *Econometrica* , Vol. 70, pp. 737–753.

Andreoni, J. and Petrie, R. (2004), "Public goods experiments without confidentiality: a glimpse into fund-raising", *Journal of Public Economics* , Vol. 88, pp. 1605 – 1623.

Archak, N. (2010), Money, glory and cheap talk: Analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on topcoder.com, *in* 'Proceedings of the 19th international conference on World Wide Web', Raleigh, North Carolina.

Aronson, J., Quinn, D. M. and Spencer, S. (1998), Stereotype threat and the academic underperformance of minorities and women, *in* J. K. Swim and C. Stangor, eds, 'Prejudice: The targets perspective', Academic Press, New York.

Axelrod, R. and Hamilton, W. D. (1981), "The evolution of cooperation", *Science* , Vol. 211 of *New Series*, pp. 1390–1396.

Ball, J. R. (1999), "'space missions' focus on team building", *The Greater Baton Rouge Business Report* , Vol. 17, p. 31.

Bargh, J. A. (2006), "What have we been priming all these years? on the development, mechanisms, and ecology of nonconscious social behavior", *European Journal of Social Psychology* , Vol. 36, p. 147168.

Bargh, J. A. and Chartrand, T. L. (1999), "The unbearable automaticity of being", *American Psychologist* , Vol. 54, pp. 462–479.

Bargh, J., Chen, M. and Burrows, L. (1996), "Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action", *Journal of Personality and Social Psychology* , Vol. 71, pp. 230–244.

Baron, D. and Kalai, E. (1993), "The simplest equilibrium of a majority-rule division game", *Journal of Economic Theory* , Vol. 61, pp. 290–301.

Baye, M. R., Kovenock, D. and de Vries, C. G. (1994), "The solution to the tullock rent-seeking game when r¿2: Mixed-strategy equilibria and mean dissipation rates", *Public Choice* , Vol. 81, pp. 363–380.

Baye, M. R., Kovenock, D. and de Vries, C. G. (1996), "The all-pay auction with complete information", *Economic Theory* , Vol. 8, pp. 291–305.

Bednar, J. and Page, S. E. (2007), "Can game(s) theory explain culture? the emergence of cultural behavior in multiple games", *Rationality and Society* , Vol. 19, pp. 65–97.

Benbow, C. P. (1995), "Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes", *Behavioral and Brain Sciences* , Vol. 11, pp. 169–232.

Benjamin, D. J., Choi, J. J. and Strickland, A. J. (2010), "Social identity and preferences", *American Economic Review* , Vol. 100, pp. 1913–1928.

Bertoletti, P. (2010), On the reserve price in all-pay auctions with complete information and lobbying games. Manuscript.

Blatt, R. and Camden, C. T. (2006), Positive relationships and cultivating community, *in* J. E. Dutton and B. R. Ragins, eds, 'Exploring positive relationships at work: Building a theoretical and research foundation', Lawrence Erlbaum Associates, Mahwah, NJ, pp. 243–264.

Borjas, G. (1994), "The economics of immigration", *Journal of Economic Literature* , Vol. 32, pp. 1667–1717.

Borjas, G. (1999), The economic analysis of immigration, *in* O. Ashenfelter and D. E. Card, eds, 'Handbook of Labor Economics', North Holland, Amsterdam.

Boudreau, K. J., Lacetera, N. and Lakhani, K. R. (forthcoming), "Incentives and problem uncertainty in innovation contests: An empirical analysis", *Management Science* .

Brewer, M. B. (1979), "Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis", *Psychological Bulletin* , Vol. 86, pp. 307–324.

Brewer, M. B. (1999), "The psychology of prejudice: Ingroup love and outgroup hate?", *Journal of Social Issues* , Vol. 55, pp. 429–444.

Camerer, C. F., Ho, T.-H. and Chong, J.-K. (2004), "A cognitive hierarchy model of games", *The Quarterly Journal of Economics* , Vol. 119, Oxford University Press, pp. 861–898.

Camerer, C. F. and Malmendier, U. (2007), Behavioral organizational economics, *in* P. Diamond and H. Vartiainen, eds, 'Behavioral Economics and Its Applications', Princeton University Press, Princeton and Oxford.

Carpenter, J., Matthews, P. H. and Schirm, J. (2010), "Tournaments and office politics: Evidence from a real effort experiment", *American Economic Review* , Vol. 100, pp. 504–17.

Cason, T. N. and Gangadharan, L. (2010), Cooperation spillovers and price competition in experimental markets. Working Paper.

Cason, T. N., Savikhin, A. and Sheremeta, R. M. (2012), "Behavioral spillovers in coordination games", *European Economic Review* , Vol. 56, pp. 233–245.

Charness, G., Rigotti, L. and Rustichini, A. (2007), "Individual behavior and group membership", *American Economic Review* , Vol. 97, pp. 1340 – 1352.

Chatman, J. A., Polzer, J. T., Barsade, S. G. and Neale, M. A. (1998), "Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes", *Administrative Science Quarterly* , Vol. 43, pp. 741–780.

Chen, M. and Bargh, J. A. (1997), "Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation", *Journal of Experimental Social Psychology* , Vol. 33, pp. 541–560.

Chen, R. and Chen, Y. (2011), "The potential of social identity for equilibrium selection", *American Economic Review* , Vol. 101, pp. 2562–2589.

Chen, Y., Ho, T.-H. and Kim, Y.-M. (2010), "Knowledge market design: A field experiment at google answers", *Journal of Public Economic Theory* , Vol. 12, pp. 641–664.

Chen, Y. and Li, S. X. (2009), "Group identity and social preferences", *American Economic Review* , Vol. 99, pp. 431–457.

Cheryan, S., Plaut, V. C., Davies, P. G. and Steele, C. M. (2009), "Ambient belonging: How stereotypical cues impact gender participation in computer science", *Journal of Personality and Social Psychology* , Vol. 97, pp. 1045–1060.

Cooper, D. J. and Kagel, J. H. (2008), "Learning and transfer in signaling games", *Economic Theory* , Vol. 34, pp. 415 – 439.

Cover, T. M. and Thomas, J. A. (2006), *Elements of Information Theory*, Wiley-Interscience.

Cox, J. and Orman, W. (2010), Immigrant assimilation, trust and social capital. IZA Discussion Paper No. 5063.

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

Croizet, J.-C. and Claire, T. (1998), "Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds", *Personality and Social Psychology Bulletin* , Vol. 24, pp. 588–594.

Crosby, F., Bromley, S. and Saxe, L. (1980), "Recent unobtrusive studies of black and white discrimination and prejudice: a literature review", *Psychological Bulletin* , Vol. 87, pp. 546–563.

Croson, R. T. A., Marks, M. B. and Snyder, J. (2008), "Groups work for women: Gender and group identity in the provision of public goods", *Negotiation Journal* , Vol. 24, pp. 411–427.

Cunningham, G. B. (2005), "The importance of a common in-group identity in ethnically diverse groups", *Group Dynamics: Theory, Research and Practice* , Vol. 9, pp. 251–260.

Davis, D. D. and Reilly, R. J. (1998), "Do too many cooks spoil the stew? an experimental analysis of rent-seeking and the role of a strategic buyer", *Public Choice* , Vol. 95, pp. 89–115.

Deaux, K. (1996), Social identification, *in* E. T. Higgins and A. W. Kruglanski, eds, 'Social Psychology: Handbook of Basic Principles', The Guilford Press, New York.

DiPalantino, D. and Vojnovic, M. (2009), Crowdsourcing and all-pay auctions, *in* 'Proceedings of the 10th ACM Conference on Electronic Commerce'.

Dovidio, J. F., Gaertner, S. L. and Saguy, T. (1980), "Commonality an the complexity of we: Social attitudes and social change", *Personality and Social Psychology Review* , Vol. 13, pp. 3–20.

Eckel, C. C. and Grossman, P. J. (2005), “Managing diversity by creating team identity”, *Journal of Economic Behavior & Organization* , Vol. 58, pp. 371–392.

Engle-Warnick, J. and Slonim, R. L. (2006), “Inferring repeated-game strategies from actions: evidence from trust game experiments”, *Economic Theory* , Vol. 28, pp. 603–632.

Falk, A., Fischbacher, U. and Gaechter, S. (2010), “Living in two neighborhoods - social interaction effects in the lab”, *Economic Inquiry* . Forthcoming.

Fehr, E. and Gaechter, S. (2000), “Cooperation and punishment in public goods experiments”, *American Economic Review* , Vol. 90, pp. 980–994.

Fershtman, C. and Gneezy, U. (2001), “Discrimination in a segmented society: An experimental approach”, *Quarterly Journal of Economics* , Vol. 116, pp. 351–377.

Fibich, G., Gavious, A. and Sela, A. (2006), “All-pay auctions with risk-averse players”, *International Journal of Game Theory* , Vol. 34, pp. 583–599.

Fischbacher, U. (2007), “z-tree: Zurich toolbox for ready-made economic experiment”, *Experimental Economics* , Vol. 10, pp. 171–178.

Friesen, J., Arifovic, J., Ludwig, A., Wright, S. C., Giamo, L. and Baray, G. (2011), Ethnic identity and discrimination among children. Simon Fraser University Working Paper.

Gaertner, S. L. and Dovidio, J. F. (2000), *Reducing intergroup bias: The common ingroup identity model*, The Psychology Press, Philadelphia, P.A.

Gilboa, I. and Schmeidler, D. (1995), “Case-based decision theory”, *The Quarterly Journal of Economics* , Vol. 110, The MIT Press, pp. 605–639.

Gneezy, U. and Smorodinsky, R. (2006), “All-pay auction: An experimental study”, *Journal of Economic Behavior and Organization* , Vol. 61, pp. 255–275.

Goette, L., Huffman, D. and Meier, S. (2006), “The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups”, *American Economic Review* , Vol. 96, pp. 212–216.

Grimm, V. and Mengel, F. (2010), An experiment on learning in a multiple games environment. Working Paper.

Gruenfeld, D. H., Mannix, E. A., Williams, K. Y. and Neale, M. A. (1996), “Group composition and decision making: How member familiarity and information distribution affect process and performance”, *Organizational Behavior and Human Decision Processes* , Vol. 67, pp. 1–15.

Ham, J. C., Kagel, J. H. and Lehrer, S. F. (2005), “Randomization, endogeneity and laboratory experiments: the role of cash balances in private value auctions”, *Journal of Econometrics* , Vol. 125, pp. 175 – 205. Experimental and non-experimental evaluation of economic policy and models.

Hanaki, N., Sethi, R., Erev, I. and Peterhansl, A. (2005), “Learning strategies”, *Journal of Economic Behavior and Organization* , Vol. 56, pp. 523–542.

Hargreaves Heap, S. P. and Zizzo, D. J. (2009), “The value of groups”, *American Economic Review* , Vol. 99, pp. 295–323.

Harper, F. M., Raban, D., Rafaeli, S. and Konstan, J. A. (2008), Predictors of answer quality in online q&a sites, *in* ‘CHI ’08: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems’, ACM, New York, NY, pp. 865–874.

Harrison, G. W. and List, J. A. (2004), “Field experiments”, *Journal of Economic Literature* , Vol. 42, pp. 1009–1055.

Haruvy, E. and Stahl, D. O. (2010), Learning transference between dissimilar symmetric normal-form games. Working Paper.

Healy, P. J. (2007), “Group reputations, stereotypes, and cooperation in a repeated labor market”, *American Economic Review* .

Heckman, J. J. (1979), “Sample selection bias as a specification error”, *Econometrica* , Vol. 47, pp. 153–161.

Hedges, L. V. and Nowell, A. (1995), “Sex differences in mental test scores, variability, and numbers of high-scoring individuals”, *Science* , Vol. 269, pp. 41–45.

Hewstone, M. (1996), Contact and categorization: Social psychological interventions to change intergroup relations, *in* N. Macrae, C. Stangor and M. Hewstone, eds, ‘Stereotypes and stereotyping’, Guilford Press, New York, p. 323368.

Hong, L. and Page, S. E. (2001), “Problem solving by heterogeneous agents”, *Journal of Economic Theory* , Vol. 97, pp. 123–163.

Howe, J. (2006), “The rise of crowdsourcing”, *Wired* , p. 4.

Howe, J. (2008), “Crowdsourcing: Why the power of the crowd is driving the future of business”, *New York: Crown Business* .

Huck, S., Jehiel, P. and Rutter, T. (2010), “Feedback spillover and analogy-based expectations: A multi-game experiment”, *Games and Economic Behavior* .

Jehiel, P. (2005), “Analogy-based expectation equilibrium”, *Journal of Economic Theory* , Vol. 123, pp. 81–104.

Jeon, G. Y., Kim, Y.-M. and Chen, Y. (2010), Re-examining price as a predictor of answer quality in an online q&a site, *in* 'Proceedings of the 28th International Conference on Human Factors in Computing Systems 2010 (CHI 2010)', Atlanta, Georgia, pp. 325–328.

Kalai, E. and W.Stanford (1988), "Finite rationality and strategic complexity in repeated games", *Econometrica* , Vol. 56, pp. 397–410.

Kleeman, F., Voss, G. and Rieder, K. (2008), "Un(der)paid innovators: The commercial utilization of consumer work through crowdsourcing", *Science, Technology & Innovation Studies* , Vol. 4, pp. 5–26.

Konrad, K. A. and Leininger, W. (2007), "The generalized stackelberg equilibrium of the all-pay auction with complete information", *Review of Economic Design* , Vol. 11, pp. 165–174.

Krippendorff, K. (2003), *Content analysis: An introduction to its methodology*, 2nd edn, Sage Publications, Thousand Oaks, CA.

Krishna, V. (2009), *Auction Theory*, 2nd edition edn, Academic Press, Waltham, Massachusetts.

Krishna, V. and Morgan, J. (1997), "An analysis of the war of attrition and the all-pay auction", *Journal of Economic Theory* , Vol. 72, pp. 343–362.

Leininger, W. (1991), "Patent competition, rent dissipation, and the persistence of monopoly: The role of research budgets", *Journal of Economic Theory* , Vol. 53, pp. 146–172.

Li, S. X., de Oliveira, A. and Eckel, C. (2010), Common identity and provision of public goods: An experimental investigation. University of Texas at Dallas CBEES Working Paper.

Liu, T. X. (2011), Sequential vs. simultaneous all-pay auction: An experimental study. University of Michigan, Manuscript.

Liu, T. X., Yang, J., Adamic, L. A. and Chen, Y. (2011), Crowdsourcing with all-pay auctions: a field experiment on taskcn. Manuscript, University of Michigan.

Luce, R. D. (1986), *Response Times: Their Role in Inferring Elementary Mental Organization*, Oxford University Press, New York.

Lugovskyy, V., Puzzello, D. and Tucker, S. (2010), "An experimental investigation of overdissipation in the all pay auction", *European Economy Review* , Vol. 54, pp. 974–997.

Mason, W. and Watts, D. J. (2009), Financial incentives and the 'performance of crowds', *in* 'KDD-HCOMP. 2009', Paris, France.

Miller, J. H. and Page, S. E. (2007), *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, Princeton University Press, Princeton, NJ.

Moldovanu, B. and Sela, A. (2001), "The optimal allocation of prizes in contests", *American Economic Review* , Vol. 91, pp. 542–558.

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

Morgan, J., Orzen, H. and Sefton, M. (2010), "Endogenous entry in contests", *Economic Theory* , pp. 1–29.

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

Nam, K., Adamic, L. A. and Ackerman, M. S. (2009), Questions in, knowledge in? a study of naver's question answering community, *in* 'The 27th Annual SIGCHI Conference on Human Factors in Computing Systems', Boston, MA.

Nemeth, C. J. (1986), "The differential contributions of majority and minority influence", *Psychological Review* , Vol. 93, pp. 23–32.

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

Pelled, L. H., Eisenhardt, K. M. and Xin, K. R. (1999), "Exploring the black box: An analysis of work group diversity, conflict and performance", *Administrative Science Quarterly* , Vol. 44, pp. 1–28.

Perdue, C. W., Dovidio, J. F., Gurtman, M. B. and Tyler, R. B. (1990), "Us and them: Social categorization and the process of ingroup bias", *Journal of Personality and Social Psychology* , Vol. 59, pp. 475–486.

Potters, J., de Vries, C. and van Winden, F. (1998), "An experimental examination of rational rent-seeking", *European Journal of Political Economy* , Vol. 14, pp. 783–800.

Rubinstein, A. (1982), "Perfect equilibrium in a bargaining model", *Econometrica* , Vol. 50, pp. 97–109.

Rubinstein, A. (1986), "Finite automata play the repeated prisoner's dilemma", *Journal of Economic Theory* , Vol. 39, pp. 83 – 96.

Rubinstein, A. (2007), "Instinctive and cognitive reasoning: A study of response times", *The Economic Journal* , Vol. 117, pp. 1243–1259.

Samuelson, L. (2001), "Analogies, adaptations, and anomalies", *Journal of Economic Theory* , Vol. 97, pp. 320–366.

Samuelson, L. (2005), "Economic theory and experimental economics", *Journal of Economic Literature* , Vol. 43, pp. 65–107.

Savikhin, A. and Sheremeta, R. M. (2012), "Simultaneous decision-making in competitive and cooperative environments", *Economic Inquiry* . Forthcoming.

Segev, E. and Sela, A. (2011), Sequential all-pay auctions with head starts. Manuscript.

Shannon, C. E. (1948), "A mathematical theory of communication", *Bell System Technical Journal* , Vol. 27, pp. 379–423.

Sheremeta, R. (2011), "Contest design: An experimental investigation", *Economic Inquiry* , Vol. 49, p. 573590.

Shih, M., Pittinsky, T. L. and Ambady, N. (1999), "Stereotype susceptibility: Identity salience and shifts in quantitative performance", *Psychological Science* , Vol. 10, pp. 81–84.

Shih, M., Pittinsky, T. L. and Trahan, A. (2006), "Domain-specific effects of stereotypes on performance", *Self and Identity* , Vol. 5, pp. 1–14.

Simpson, B. (2006), "Social identity and cooperation in social dilemmas", *Rationality and Society* , Vol. 18, pp. 443–470.

Spencer, S., Steele, C. M. and Quinn, D. M. (1999), "Stereotype threat and women's math performance", *Journal of Experimental Social Psychology* , Vol. 35, pp. 4–28.

Steele, C. M. and Aronson, J. (1995), "Stereotype threat and the intellectual test performance of african-americans", *Journal of Personality and Social Psychology* , Vol. 69, pp. 797–811.

Steen, L. A. (1987), "Mathematics education: A predictor of scientific competitiveness", *Science* , Vol. 237, pp. 251–253.

Tajfel, H. and Turner, J. (1979), An integrative theory of intergroup conflict, *in* S. Worchel and W. Austin, eds, 'The Social Psychology of Intergroup Relations', Brooks/Cole, Monterey, CA.

Tanaka, T., Camerer, C. F. and Nguyen, Q. (2010), "Risk and time preferences: Linking experimental and household survey data from vietnam", *American Economic Review* , Vol. 100, pp. 557–571.

Tullock, G. (1980), Efficient rent seeking, *in* 'Toward a Theory of the Rent-Seeking Society', Texas A&M University Press Press, pp. 97–112.

Tushman, M. L. (1977), "Special boundary roles in the innovation process", *Administrative Science Quarterly* , Vol. 22, pp. 587–605.

Tversky, A. and Kahneman, D. (1986), "Rational choice and the framing of decisions", *The Journal of Business* , Vol. 59, p. S251.

Tversky, A. and Kahneman, D. (1992), “Advances in prospect theory: Cumulative representation of uncertainty”, *Journal of Risk and Uncertainty* , pp. 297–323.

Van Der Zee, K. and Paulus, P. (2008), Social psychology and modern organizations: Balancing between innovativeness and comfort, *in* L. Steg, A. P. Buunk and T. Rothengatter, eds, ‘Applied social psychology: Understanding and managing social problems’, Cambridge University Press, New York, p. 271290.

Van Huyck, J. B., Battalio, R. C. and Beil, R. O. (1993), “Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication”, *Games and Economic Behavior* , Vol. 5, pp. 485 – 504.

Walker, M. and Wooders, J. (2001), “Minimax play at wimbledon”, *American Economic Review* , Vol. 91, pp. 1521–1538.

Weber, R. A. (2006), “Managing growth to achieve efficient coordination in large groups”, *American Economic Review* , Vol. 96, pp. 114 – 126.

West, T. V., Pearson, A. R., Dovidio, J. F., Shelton, J. N. and Trail, T. E. (2009), “Superordinate identity and intergroup roommate friendship development”, *Journal of Experimental Social Psychology* , Vol. 45, pp. 1266–1272.

Wilder, D. A. (1986), Social categorization: Implications for creation and reduction of intergroup bias, *in* L. Berkowitz, ed., ‘Advances in experimental social psychology’, Vol. 19, Academic Press, New York, p. 291355.

Wooders, J. (2010), “Does experience teach? professionals and minimax play in the lab”, *Econometrica* , Vol. 78, pp. 1143–1154.

Yamagishi, T. and Kiyonari, T. (2000), “The group as the container of generalized reciprocity”, *Social Psychology Quarterly* , Vol. 63, pp. 116–132.

Yang, J., Adamic, L. A. and Ackerman, M. S. (2008*a*), Competing to share expertise: the taskcn knowledge sharing community, *in* ‘Proceedings of International Conference on Weblogs and Social Media’, Seattle.

Yang, J., Adamic, L. A. and Ackerman, M. S. (2008*b*), Crowdsourcing and knowledge sharing: Strategic user behavior on taskcn, *in* ‘Proceedings of the 8th ACM Conference on Electronic Commerce’, Chicago.

Yang, J. and Wei, X. (2009), Seeking and offering expertise across categories: A sustainable mechanism works for baidu knows, *in* ‘International Conference on Weblogs and Social Media.’, San Jose.